A Treatise on Econometric Forecasting

Thesis by

Alfredo Martinez Estrada

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2007

(Defended February 13, 2007)

To my parents

Acknowledgements

My greatest gratitude and appreciation goes to my adviser, John Doyle. The support and intellectual freedom John provided through the years allowed me to develop academically and encouraged me to strive for the integrity of my work. The opportunities John has provided have open a world of possibilities for which I'm forever thankful.

This work would not have been possible without the encouragement and guidance of Ivana Komunjer. I am eternally grateful to Ivana for all the time, patience, and the knowledge she has imparted to me for the past four years. I would also like to thank my thesis committee: Yaser Abu-Mostafa, Mani Chandy, and Richard Murray.

Through all my years at Caltech, many people have been an integral part of my life. I would like to thank my small Pasadena family: Pablo Parrilo, for all his encouragement and advice, I would not have made it without him, Federico Spedalieri, for his kindness and sense of humor, Monica Giannelli, for her sincerity and pure, unconditional friendship, Sergey Pekarsky and Evgenia Pekarskaya, for the good parties and travels, Chris Mouton for all our lunches, dinners, and long conversations, and Luz Vela for her companionship during our many adventures. I also want to thank Andrew Doherty, Ted Wider, and Martin Basch for all the fun nights. I also want to thank my officemates Melvin Flores and Abhishek Tiwari and Xin Liu for all our travels to San Diego. Many other people have been much help through the years: Gabriela Surpi, Mario Sznaier, Octavia Camps and Mauricio Barahona.

Finally, I would like to express my greatest and deepest gratitude to my family, Manuel Martinez, Silvia Martinez, and Alejandro Martinez, for their support and unconditional love.

Abstract

We investigate the effects of model misspecification and stochastic dynamics in the problem of forecasting. In economics and many fields of engineering, many researchers are guilty of the dangerous practice of treating their mathematical models as the true data generating mechanisms responsible for the observed phenomena and downplaying or omitting all together the important step of model verification. In recent years, econometricians have acknowledged the need to account for model misspecification in the problems of estimation and forecasting. In particular, a large body of work has emerged to address properties of estimators under model misspecification, along with a plethora of misspecification testing methodologies. In this work, we investigate the combined effects of model misspecification and various types of stochastic dynamics on forecasts based on linear regression models. The data generating process (DGP) is assumed unknown to the forecaster except for the nature of process dependencies, i.e., independent identically distributed, covariance stationary, or nonstationary. Estimation is carried out by means of ordinary least squares, and forecasts are evaluated with the mean squared forecast error (MSFE) or mean square error of prediction. We investigate the sample size dependence of the MSFE. For this purpose, we develop an algorithm to approximate the MSFE by an expression depending only on the sample size n and moments of the processes. The approximation is constructed by Taylor series expansions of the squared forecast error which do not require knowledge of the functional form of the DGP. The approximation can be used to determine the existence of optimal observation windows which result in the minimum MSFE. We assess the accuracy of the approximating algorithm with Monte Carlo experiments.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The two main objectives in the fields of engineering, the social sciences, and the natural sciences are description of phenomena and prediction of phenomena. In engineering and most of the natural sciences, the ability to perform controlled experiments is of fundamental importance for testing theories and building models that explain underlying relationships. For most of the social sciences, and in particular for economics, researchers lack the important tool of repetitive experimentation. This missing link between empirical reality and theoretical modeling has been regarded as a considerable handicap in the development of economics as a science. Two influential developments in the early twentieth century addressed this quandary: the introduction of formal probability theory in economic modeling, and the development of the field of econometric forecasting. Probability based models allow for statistical hypothesis testing to evaluate results from estimation. Econometric forecasting, and in particular the use of out of sample forecasts, have become indispensable in the use of empirical studies to validate theoretical models.

The aim of this thesis is to answer the question: How much past data is optimal to use in the construction of a forecast? Our approach to the subject is to use tools from econometrics to determine the dependence of a common evaluation scheme, the mean square forecast error, on the sample size.

Many forecasting methodologies have been developed, with the most commonly used being time series and econometric models [9, 32]. A strategy for building forecasts must include three major steps: specification, estimation, and verification. In the work that follows, we keep with the convention of simplicity and specify linear models. Standard practice in estimation makes use of three possible mechanisms to determine the temporal significance of data: one is to use an expanding window, which includes all available data to form estimators; a second is to apply a rolling window of fixed size; the third applies a predetermined monotonic decreasing weighting function. These procedures are ad hoc with no basis for their application other than the researcher's intuition. Verification can consist of evaluating a forecast constructed with the estimated model by comparing the forecast to realizations outside the estimation sample.

Determining the temporal significance of data for the problems of estimation and forecasting is of great consequence for optimal accuracy. The most intuitive reason for this is that data may simply "get too old" to be informative, and in many cases may, in fact, hinder the discovery of the underlying relationships. This phenomena manifests itself, for example, in certain types of bias of estimators. The characteristic which encompasses the evolving nature of data is the dynamics of the data generating process. For the mathematical description of stochastic processes, the dynamics are summarized by the probability joint distribution. Mathematical convention categorizes process dynamics based on the joint distribution as either stationary or nonstationary. Proper selection of data is clearly an important matter for estimation and forecasting when considering nonstationary processes which are characterized by the dynamic nature of the joint distribution. For example, structural breaks in economic data due to institutional, political, financial, and technological changes are well documented [3, 13, 32, 33, 53, 139] and can lead to serious bias in estimation and unacceptable prediction errors. Less intuitive is the need to be concern about the temporal significance of data in the case of stationary processes which are generated by constant probability structures. Data temporal significance has ramifications for the treatment of stationary processes when model misspecification is inevitable. The concept of misspecification arises from the acknowledgment that researchers in general work with models of the data generating processes which suffer from discrepancies. For the treatment of economics, this idea is best described by White in [152],

> Because of the exceeding complexity of economic behavior, because of the extreme difficulty of measuring or even properly defining relevant aspects of economic phenomena, and because the economist typically has little or no control over the economic phenomenon under study, economic theory is

> fundamentally and inherently limited in the degree to which it can describe economic reality or make legitimate falsifiable statements about economic reality. Because the empirical economist must deal with nature in all her complexity, it is optimistic in the extreme to hope or believe that standard parametric economic models or probability models are sufficiently adequate to capture this complexity.
>
> A realistic attitude in such circumstances is that an economic model or a probability model is in fact only a more or less crude approximation to whatever might be the "true" relationships among the observed data, rather than necessarily providing an accurate description of either the actual economic or probabilistic relationships. Consequently, it is necessary to view economic and/or probability models as misspecified to some greater or lesser degree.

The ramifications of model misspecification for estimation have been studied mainly for linear regression models of non-stochastic variables under a very restricted class model misspecifications [20, 72, 73, 80, 98, 108, 120, 121, 124, 145, 146, 157]. More generally, the work in [43, 149, 150, 152] addresses stochastic process and provides large sample properties of estimators, such as the quasi-maximum likelihood estimator, ordinary least squares and weighted least squares, in the presence of general model misspecifications.

In this work, the goal is to construct a data based procedure for determining the temporal significance of data for the problem of forecasting. In particular, we are interested in the behavior of forecast evaluating schemes for finite size samples. To do this, we develop a forecasting strategy which integrates the estimation and verification steps into one step. We now describe this strategy. As mentioned, the model is specified as a linear regression of observed stochastic processes, $\{X_\tau\}$, that act as explanatory variables for the dependent stochastic process, $\{Y_\tau\}$. The regression parameters are estimated with the ordinary least squares (OLS) estimator. At this point in most forecasting strategies, the OLS estimator is completely defined as a function of the most recent or available n observations of the processes, and the estimation procedure is finished. In our strategy, the value of n is a variable to be determined in the verification step. As such, the OLS estimator is implicitly a function of the variable n. The one step ahead forecast $\hat{Y}_{t+1,n}$ at the origin t of the variable Y_{t+1} is given by the linear form of the regression and by using the OLS. Again, the forecast, through the OLS, is implicitly a function of the variable n. As in much of the literature, [32, 33], the forecast evaluating scheme of choice is the mean square forecast error (MSFE), defined as $MSFE_{t,n} = E[(Y_{t-1} - \hat{Y}_{t+1,n})^2]$, and, through the forecast, the MSFE is implicitly a function of the variable n. The verification and estimation steps are linked together by the determination of n. The optimal value of n is determined in the verification stage by evaluating the forecast performance by means of the MSFE. For this, we define an optimal observation window of size $n^\star$, as the solution to the optimization problem:

$$\min_{n \in \mathbb{N}^+} MSFE_{t,n}.$$

$n^\star$ can be either finite or infinite[1]. The case when $n^\star$ is infinite implies all data available should be used for forecasting. The case when $n^\star$ is finite describes the optimal continuous compact observation window to be used for forecasting. The key question is therefore to study the behavior of the MSFE as a function of the sample size variable n. Analyzing the sample size dependence (SSD) of the MSFE is a difficult task, especially under the assumption of misspecification. The significance of misspecification in forecasting has been studied in [16, 17, 91, 126]. This work gives expressions for the MSFE that only apply to the case where the data generating process is known to be an autoregressive process of order m and the forecast is constructed with a model which is an autoregressive process of order $p \neq m$. Clearly this violates our assumption of not knowing the functional form of the data generating process in the course of the analysis.

Up to now, no method has been developed to study the SSD of the unconditional MSFE[2]. In the chapters to follow, we construct an approximation of the MSFE for forecasting problems involving processes with different types of stochastic dependencies which can be used to study the SSD under the assumption arbitrary misspecifications.

[1] If the minimum MSFE occurs at more than one value of n, $n^\star$ refers to the smallest value of n.

[2] In [113] the authors obtain a first order approximation for the MSFE under the assumption of i.i.d. normally distributed processes.

1.1 Outline and contributions

The thesis is organized into nine chapters. Chapter 2 presents basic concepts of forecasting, e.g. methodologies, principles, and definitions, and introduces the main problem of interest. Section 2.2 provides a historic exposition of important developments in economic forecasting. Section 2.3 describes different forecasting methodologies. Section 2.4 describes the forecast problem of predicting an unknown data generating process using a linear forecasting model. Section 2.5 provides a short exposition on the subject of misspecification in terms of density functions. Section 2.6 presents some motivating examples and section 2.7 provides theoretical intuition for the problem of determining optimal observation widows.

Chapter 3 presents notation of probability, random variables, and expectations. The concept of truncated expectation, and properties based on the standard notation of expectations are developed. Truncated expectations are crucial to the development of the forecasting algorithms based on Taylor approximations which are presented in chapters to follow.

Chapter 4 presents an algorithm to approximate the expectation of functions of random variables based on Taylor series expansions. The technique is used in Chapters 5, 6, and 7 to approximate the mean square forecast error (MSFE).

Chapter 5 presents the algorithm which yields an approximation of the MSFE for a forecasting problem involving independent and identically distributed processes. This Taylor algorithm approximation is meant to be used as a tool to describe the sample size dependence (SSD) of the MSFE. Section 5.2 reviews some properties of the OLS and MSFE under the assumption of a correctly specified forecast model. Section 5.3 describes properties of the ordinary least squares (OLS) under the assumption of a functionally misspecified model. Section 5.4 presents the derivation of the Taylor algorithm for the scalar case, and Section 5.5 presents the derivation for the multi-variate case. Section 5.6 evaluates the performance of the Taylor algorithm for the MSFE of a scalar forecasting problem with Monte Carlo experiments.

Chapter 6 presents the algorithm which yields an approximation of the mean square forecast error for a forecasting problem involving stationary processes. Section 6.2

presents results in the literature concerning estimation under misspecification with dependent observations. Section 6.3 presents the algorithm and Section 6.4 presents Monte Carlo experiments to evaluate the MSFE approximation.

Chapter 7 presents the algorithm which yields an approximation of the mean square forecast error for a forecasting problem involving independent and identically distributed processes which undergo structural breaks. Section 7.2 presents the algorithm and Section 7.3 presents Monte Carlo experiments to evaluate the MSFE approximation.

Chapter 8 presents a literature review of the Delta method as well as new results for a wider class of functions. Chapter 9 discusses the conditions needed for application of the Delta method results presented in Chapter 8.

The main contributions of this thesis are as follow:

- We develop an algorithm to approximate the MSFE in forecasting problems formulated with models which may be misspecified. Unlike anything in the literature, our algorithm makes no assumptions on the specific form of the data generating process and can be applied to real empirical problems.

- We employ the MSFE approximation to investigate the sample size dependence of the MSFE and determine the existence of optimal observation windows for three classes of processes: i.i.d. processes, covariance stationary processes, and structural break processes.

- We prove some Delta method theorems for unbounded functions which provide bounds on the error of approximation.

- We provide an extensive treatment on the use of Taylor series to approximate statistics.

Chapter 2

Forecasting

2.1 Introduction

In this chapter, we present basic concepts of forecasting, e.g., methodologies, principles, and definitions, and introduce the main problem of interest. Section 2.2 provides a historic exposition of the most important developments and contributions in economic modeling and forecasting. Section 2.3 describes different forecasting methodologies. The two main methodologies of interest are time series models and econometric models. Section 2.4 describes the forecast problem of predicting an unknown data generating process using a linear forecasting model, with Section 2.4.4 focusing on the analysis of the mean square forecast error (MSFE). Section 2.5 provides a short exposition on the subject of misspecification in terms of density functions. Section 2.6 presents some motivating examples and Section 2.7 provides theoretical intuition for the problem of determining optimal observation widows.

2.2 History and background

Many would agree that the two main goals in the study of econometrics are optimal estimation and forecasting. To provide a clear prospective of the contribution of this thesis to the field of forecasting, we present a short overview of the history and methods of forecasting in economics. In the broadest sense, forecasting is any set of rules or procedure which is carried out with the intent of predicting the outcome of a future event, or some particular characteristic of a future event. We refer the reader to the references [32, 41, 62, 99, 103] for further details on the survey that follows.

To realize a comprehensive understanding of the development and present state of economic forecasting, it is paramount to assess the progression of macroeconomic theory and modeling. The reason for this is that the first attempts at forecasting came about as methods for evaluating macroeconomic models. The origins of economic forecasting can be traced to the work of economists of the nineteenth and early twentieth centuries in the two main branches of macroeconomics, business cycle and demand analysis. Morgan [103] gives an account of attempts by early econometricians to model these economic phenomena. William Stanley Jevons and Henry Ludwell Moore were two of the first economists to apply the econometric approach of combining economic theory with statistical tools to give evidence for hypotheses concerning the business cycle.

Jevons was one of the first economists to combine theory with statistical data on many events to explain the business cycle. Jevons' initial hypothesis on trade cycles was that the sunspot cycle of 11.1 years was responsible for a weather cycle which in turn caused a harvest cycle and ultimately led to a price cycle ([79] in paper VI). Jevons' analysis consisted of laying out data for a number of price series for different crops over a 140 year period on an 11 year grid. The analysis, based on agricultural data from the thirteenth and fourteenth centuries, showed similar patterns of variation in the prices of each of the crops. The results of this work were inconclusive since the analysis revealed similar patterns for grids of 3,5,7,9 and 13 years. Jevons also investigated cycles in commercial credit. His analysis of nineteenth century financial crises exhibited an average cycle of 10.8 years, short of the sunspot cycle of 11.1 years. Jevons suggested that his sunspot theory combined with the theory of credit cycle would produce the observed averaged cycle for financial crises. Most of his contemporaries dismissed the work of Jevons. Some of the strongest criticism concerned the lack evidence and weak explanation of the casual mechanisms of his theory. Nonetheless, the idea behind Jevons' work of combining endogenous and exogenous causes became an important element in econometric models of the business cycle in the 1930's.

Much like Jevons, Henry Ludwell Moore developed theories on the exogenous causes of the business cycle. Moore [101] found evidence to attribute the business cycle to weather cycles, and later [102] extended the casual reasoning back to movements of the planet Venus. For Moore, the casual chain of explanation between the weather cycle and the business cycle was the primary subject of study. He abandoned the standard methodologies of the time on the grounds that the real dynamic factors of the economy could not be captured by comparative statistics. Morgan [103] gives as example of the contemporary mainstream methods, the work by Robertson [123] on the business cycle, which made use of comparative static arguments with statistical data but without any statistical analysis or explanation of the dynamic path of the economy. Moore's efforts focused on discovering and verifying statistically the casual connections in the chain of evidence in order to explain the business cycle. His treatment of evidence, according to Morgan [103], was highly technological compared to his predecessors and contemporaries. Moore's statistical methods included harmonic analysis, correlation, multiple regression, and time series decomposition. His analysis of business cycles was far superior to any other statistical treatment of the period.

In a 1933 paper, Ragnar Frisch [52] made important progress in the application of the econometric method by developing a dynamic mathematical model of the business cycle, which not only enabled theorists to explore for insights into how the economy might work but also was amenable to econometric analysis. The work of Moore and others explained and estimated the business cycle by fitting the dynamic patterns of a particular time. Frisch's model was not built to fit any particular data set, instead, the purpose of the model design was to generate economic cycles through the interactions of the equations in the system by estimating parameters based on the particular data set at hand. The second important econometric design of Frisch's model was the interaction of random shocks with a deterministic system. The role of random shocks transformed the model from a solely theoretical model producing the cyclical components to one which could produce the jagged appearance of economic data. The shocks changed the dynamic economic model into a formal econometric stochastic model of how real economic data might be produced [103].

The first crucial event in the annals of forecasting was the formulation of the first practical macroeconomic model of the business cycle by Jan Tinbergen in 1936. Morgan [103] examines in detail the work and contribution of Tinbergen's macrodynamic models. The following summarizes Morgan's account. Tinbergen's contribution consists of three major reports in which he estimates and tests models of economies and the business cycle. The first of these reports was published in 1936 in response to a request by the Dutch Economic Association to study policies to help relieve the depression [140]. In it, Tinbergen builds and estimates the first macrodynamic model of the business cycle. The model is also used to simulate the likely impact of policies. As his starting point, Tinbergen takes the basic idea of Frisch [52] that a business cycle should consist of two parts, an economic mechanism and the outside influences or shocks. The Dutch model contained 31 variables and 22 relationships which were divided into technical equations, definitional equations, and direct casual relationships which provided explanations of price movements, sales, competition, and the formation and disposal of incomes. Each of the equations was estimated separately. The formation of each individual equation and the choice of variables were found by iterating between theoretical ideas and empirical evidence. Graphical methods were used by plotting dependent and explanatory variables to reveal specific causes of a crisis or revival. To understand the behavior of the model, Tinbergen reduced the system of 22 equations to one difference equation of one variable, non-labor income, by a process of substitution and elimination. The final equation gave a representation of the structure of the Dutch economy which Tinbergen in turn used to find the time path of the system. The Dutch model showed that the economy had a damped cyclic path which would tend to an equilibrium provided there were no disturbances. In practice, determining the dynamic character the model was complicated by the presence of disturbances. Extrapolation of the model was used as a test of the power of the model to provide a theory of the business cycle and led to the investigation of optimal policy based on the model predicted time paths. Policy changes affected the relations in the model through additive disturbance terms or by changing coefficients and causing structural change. In this way, Tinbergen's model originated the practice of determining policy based on econometric forecasts. The second report made by Tinbergen [141] was a commission made by the League of Nations to undertake statistical tests of business cycle theories presented by G. Haberler [63]. Tinbergen developed and estimated mathematical models for verbally expressed theories of the business cycle, but the emphasis of this report was on testing using procedures involving economic and statistical criteria. The first of these procedures involved testing the models on different countries and time periods. Second, Tinbergen tried the models on different subperiods to test for for structural changes. Third, prediction tests were carried out by extrapolating the fitted equations. The third important report made by Tinbergen was also part of the League of Nations report [141]. In it, Tinbergen developed a three-stage procedure for evaluating theories of the business cycle. The stages were first to test whether the verbal model could be expressed as an econometric model; second to statistically verify the relations of the model; and third to test and verify if the final equation had a cyclic solution. To evaluate his procedure, Tinbergen built the first large scale macroeconometric model of the USA. Tinbergen's general conclusion was that a depression can originated from inherent disproportionalities in the economy, and that policy changes might intervene to prevent the rise or fall of a depression. By the 1940s, the war had vanquished the depression and theories of the business cycle had gone out of fashion. Nonetheless, the econometric methods for estimation and testing set forth by Tinbergen had great influence on the work of economists in the second part of the twentieth century.

As we have noted, since the beginning of the twentieth century, econometricians had been using statistical methods to measure and verify economic theory. And yet, it was the prevalent belief at the time that probability theory was inapplicable to economic data. The paradox lay in the theoretical basis for statistical methods being probability theory, and economists using statistical methods at the same time they rejected probability theory. Applied economists at this time believed in the existence of real laws of economics waiting to be discovered. Thus, the primary goal in early econometric work was that of measurement. No importance was paid to inference, so that if measured values were put in question, blame was attributed to the quality of the data and no doubt was cast on the theory.

In areas where a generally agreed theory existed, i.e., demand theory, statistical methods were simply tools to measure the parameters of the laws. The theory was not in doubt, the measured laws were taken to be true, and questions of inference did not arise. In other areas where theoretical laws were in doubt, such as business cycle research, statistical methods were used to uncover the true laws from the data. Inference again found a limited role. Therefore, inference methods based on probability theory as tools to compare theoretical laws to empirical relationships were neglected and deemed unnecessary by the economists of the time. This neglect was due in part to econometricians' belief that economic data did not meet the criteria necessary for the application of probability reasoning. In work on demand, for example, the least squares method was used as an estimation device without any reference to probability distributions. This is due to the fact that the relationship between two variables can be measured by a least squares line, and the distribution of the variables does not come into question unless one is interested in inference about whether it is a good measure. The application of probability theory was rejected in such work, based on the argument that observations were rarely the result of sampling procedures. One of the earliest rejections of the application of probability theory to economics was that offered by Warren Persons, [112, 111], in his 1923 presidential address to the American Statistical Association. Persons rejected mathematical probability theory in business cycle analysis and forecasting, and cited as a reason the fact that economic data are time-related and "cannot be considered a random sample except in an unreal, hypothetical sense."

The first comprehensive discourse on the rejection of the application of probability theory in economics and the validity of economic forecasting is the work of Morgenstern (1928) [104]. Morgenstern delineated the problems with probability theory as the lack of homogeneity of the underlying conditions, the non-independence of observed time series and the limited availability of data. Besides his objections towards probability theory, Morgenstern also argued against economic and business forecasting on the basis that forecasts would be invalidated by reactions to them. This is reminiscent of the "Lucas critique" [97]. Because of the impossibility of economic forecasts and the impact of adverse effects of decisions made based on them, Morgenstern censured the use of forecasting for stabilization and social control.

The work of Morgenstern was critically reviewed by Marget (1929) [99]. In his work, Marget outlines the following three main propositions offered by Morgenstern:

I. Forecast in economics by the methods of economic theory and statistics is "in principle" impossible.

II. Even if it were possible to develop a technique of economic forecasting, such a technique would be incomplete, by virtue of its necessary limitation to methods based on a knowledge of economics alone; it would therefore be incapable of application in actual situations.

III. Moreover, such forecasts can serve no useful purpose. All attempts to develop a formal technique for forecast are therefore to be discouraged.

Morgenstern provides support for each these propositions with further subsidiary sub-propositions. We review the arguments given by Morgenstern for these propositions and the counter arguments of Marget.

The sub-propositions given by Morgenstern for the first proposition, **I**, that "forecasting in economics, by methods of economic theory and statistics, is in principle impossible" are as follows:

A. The data with which the economic forecaster must deal are of such a nature as to make it certain that the prerequisites for adequate induction must always be lacking.

B. Economic processes, and therefore the data in which their action is registered, are not characterized by a degree of regularity sufficient to make their future course amenable to forecast, such "laws" as are discoverable being by nature "inexact" and loose, and therefore unreliable.

C. Forecasting in economics differs from forecasting in all other sciences in the characteristic that, in economics, the very fact of forecast leads to "anticipations" which are bound to make the original forecast false.

For sub-proposition **A**, Morgenstern first argues on the incompatibility of economic data and probability analysis, as a method of scientific induction, as a major obstruction to the problem of economic forecasting. The criteria required by Morgenstern on economic data for the application of formal probability theory include homogeneity and independence. Marget argues that the level of homogeneity and independence required by Morgenstern is so extreme as to make use of probability theory in other scientific areas – where its usefulness is well established – inconceivable. Marget, like most economists of the time, agrees with Morgenstern on the partial failure of probability theory as a tool for induction in economic forecasting. Nonetheless, Marget does not see this failure as a coup de grace for the principle of forecasting, and argues probability analysis is by no means the only tool available for scientific forecasting. For instance, prediction of day to day weather is cited by Marget as an example of forecasting which is primarily based on a theory of causation rather than techniques of probability. As a second point in support of his argument, Morgenstern points to the inadequacy of economic statistics in providing a complete description of economic processes and ultimately being used for forecast. Marget argues that even if economic statistics alone can not provide a basis for induction, which in turn serves as basis for forecast, there is no reason why new methods can not be developed which can further the paths of progress in forecasting.

For sub-proposition **B**, Morgenstern addresses the concept of an "economic law" by distinguishing between two types of "law." The first interpretation given is in the sense of a "rule of adequate causation", and the second as a tendency to "continuous repetition." The latter description of "law" is used by Morgenstern to refer to a tendency of data to conform to measurable patterns that can be predicted by mathematical formulas. According to Morgenstern, by the nature of economic processes, one can not expect to discover regularities of the kind described by the second type of economic "law" and furthermore

> The discovery of such regularities by purely empirical means would carry with it no assurance of the indefinite continuance of these regularities, and so would represent no reliable basis for forecast.

Marget views the second type of "law", which concerns itself with regularities, to be in some sense naive, and argues that the concept of law which best exemplifies the basis for most scientific endeavor is a law as a "rule of adequate causation." Marget views as reasonable the possibility of explaining movements in statistical data based on the concept of causation. Indeed, if this were not the case, Marget explains, all validity of scientific explanation in economics would be futile. Marget presents the explanation of processes based on causation as the path to follow in order to make progress in the lines forecasting, and ties such rules of causation to the study of economic theory.

The third sub-proposition, **C**, of Morgenstern is seen by Marget as the most important. If the third sub-proposition of Morgenstern is found to be sound, all other arguments in favor of the possibility of forecasting in economics become irrelevant. Marget's position on the third sub-proposition — regarding the invalidity of a forecast due to the causal influence of the forecast itself — is that forecasting should be feasible by including the possible reactions to the forecast as one of the potential factors affecting the final result. Marget also questions whether the anticipatory actions need necessarily to be of the disruptive sort which invalidates the original forecast. In some instances, Marget argues, all that might result from these anticipations is an "intensification, instead of a contradiction, of the actions that would have been inaugurated in any case." Furthermore, Marget insists there is no reason to assume that the new datum from the forecast must outweigh all other data available, and necessarily cause agents to abandon the course of action that would be taken in the absence of the original forecast.

In his second principal proposition, **II**, Morgenstern argues that even if a "positive theory of forecasting in economics" were possible, it would not be adequate in practice, since the data in use are the result of forces other than just economic forces. Marget begins his counterpoint by suggesting that the objection is as valid against explanation of economic theory as it is for attempts at forecasting. Morgenstern sees as a major obstruction to further progress in economic forecasting the ramifications that can be attributed to different branches of knowledge. Sociology, for example, is cited by Morgenstern as a field not yet sufficiently advanced to be of practical use to a business forecaster. Marget responds that the incompleteness of knowledge cannot be used to deny the possibility of the attainment of further knowledge. Morgenstern argues that, for an economic forecast, only economic theory and the data refined by economic statistics may be used, while at the same time stating that economic data is not sufficient for the problem in practice. Marget states there is no reason why an economist interested in forecasts of cotton prices, for example, should not combine her own knowledge on how to economize on the basis of a particular situation with the first hand knowledge of meteorologists and agronomists as to what the situation might be.

In his third and final proposition, **III**, Morgenstern asserts that the attempt to forecast economic events is "without purpose." Morgenstern concludes that the possible use of forecasting as an instrument for social control of industry, in particular the possibility of stabilization, may well endanger those efforts by threatening the "rationality" of the economic processes. For this proposition, Marget does not refer back to the earlier analysis of the disruptive feedback effects between the forecast and anticipations. Instead, Marget challenges the fundamental argument that presents stabilization as a test for the usefulness of attempts at forecast, and the view that forecast itself can have significance only for economic policy and not for the development of economic theory. Marget believes Morgenstern fails to recognize the value which persistent attempts to forecast have for the development of economic theory. Marget sustains that failures in forecasting, like failures in attempts at verification of economic theory, should be greeted with enthusiasm, since it is likely such failures are due to inadequate attention to important factors. Marget believes the test of successful forecasting has the inestimable advantage of pointing out new variables and new possibilities of mechanisms which might never have otherwise been discovered or estimated.

The views expressed by Morgenstern and Marget regarding the validity of economic forecasting set the stage for further development at a time where forecasting techniques were at their infancy. Economic forecasting was not doomed as Morgenstern might have one believe, but at the same time, the arguments of Marget needed to be substantiated by formal protocols. In 1944, the publication of Trygve Haavelmo's *The probability approach in econometrics* [62] provided the first basis for such protocols in the form of probability techniques. According to Haavelmo, econometric research aims at a conjunction of economic theory and actual measurements through the use of the theory and techniques of statistical inference. Haavelmo summarizes the state of the art in econometrics.

> So far, the common procedure has been, first to construct an economic theory involving exact functional relationships, then to compare this theory with some actual measurements, and, finally, "to judge" whether the correspondence is "good" or "bad." Tools of statistical inference have been introduced, in some degree, to support such judgment, e.g., the calculation of a few standard errors and multiple-correlation coefficients. The application of such simple "statistics" has been considered legitimate, while, at the same time, the adoption of definite probability models has been deemed a crime in economic research, a violation of the very nature of economic data. That is to say, it has been considered legitimate to use some of the tools developed

> in statistical theory without accepting the very foundation upon which statistical theory is built. For no tool developed in the theory of statistics has any meaning — except, perhaps, for descriptive purposes — without being referred to some stochastic scheme.

Haavelmo attributes the reluctance of economists to accept probability theory as a basis for economic theory to a very narrow concept of probability theory. Most economists of the time believed probability schemes applied only to phenomena consisting of series of observations where each observation originated as an independent drawing from a single population. Economic time series do not conform to such a narrow model of probability "because the successive observations are not independent." Haavelmo's premise is that it is not necessary for observations to be independent or to follow the same one-dimensional probability law, that in fact, it is sufficient to consider the whole set of n observations as one observation of n variables following an n-dimensional joint probability law. One can test the hypothesis regarding the joint probability law and draw inference as to its form based on one n-dimensional sample point.

The general principles of statistical inference introduced by Haavelmo are based on the Neyman-Pearson theory of testing statistical hypotheses. Haavelmo addresses many issues including: a general discussion on the connection between abstract models and economic reality; the question of establishing "constant relationships" in economics, and the degree of invariance of economic relations with respect to changes in structure; the nature of stochastic models and their applicability to economic data; demonstration that a hypothetical system of economic relations can be expressed as statements of the joint probability law of the economic variables involved, and that such a system can be regarded as a statistical hypothesis in the Neyman-Pearson sense; the well posed problem of estimation; and an outline of the problem of predictions.

We describe the general probability formulation of Haavelmo's prediction problem. By a statistical prediction or forecast, one means a probability statement about the location of a sample point to be observed in the future. If one considers n random variables, $X_1, X_2, \ldots, X_n$, with a known joint probability law, one may calculate the probability of a sample point falling into a given region of the sample space. If the actual joint probability law of the variables to be predicted is known, the problem of deriving a prediction formula is one of probability calculus, while the question of choosing a "best" prediction formula is subjective matter. More often, the probability law is not known and the prediction problem becomes closely connected with the problems of testing hypotheses and estimation.

Consider n time series of random variables $X_{i,t}$, $i = 1, 2, \ldots, n$ observable from $t = 1$ on. Suppose we can observe values up to some time, $t = s_i$, for each of the n series, and the problem is to predict later observations. The total of random variables to be considered are

$$X_{i,t} = (X_{i,1}, X_{i,2}, \ldots, X_{i,s_i}, X_{i,s_{i+1}}, X_{i,s_{i+2}}, \ldots), \quad i = 1, 2 \ldots, n.$$

One might want to predict any joint system of M variables among the variables $X_{i,s_{i+\tau}}$ for $i = 1, 2, \ldots, n$; $\tau = 1, 2, \ldots$. The M to be predicted variables, relabeled as $X_{N+1}, \ldots,$ X_{N+M}, together with the $s_1 + s_2 + \cdots + s_n = N$ observed variables, relabeled as $X_1, \ldots, X_N$, form a system of $N + M$ variables. We assume, regardless of the values $s_1, \ldots, s_n$, and regardless of the set of M future variables, the joint probability law of the $N + M$ variables exists even if it might not be known to the forecaster. Let this joint probability be denoted as $p = p(X_1, \ldots, X_N, X_{N+1}, \ldots, X_{N+M})$, which usually can be described implicitly by a system of stochastic relations between the variables. Let $p_1 = p_1(X_1, \ldots, X_N)$ denote the joint probability law of the N variables $X_1, \ldots, X_N$, and denote the conditional probability law of the M variables $X_{N+1}, \ldots, X_{N+M}$, conditional on the N variables $X_1, \ldots, X_N$ by $p_2 = p_2(X_{N+1}, \ldots, X_{N+M}|X_1, \ldots, X_N)$. If p is known, one can calculate p_2, given the N variables $X_1, \ldots, X_N$ and $p = p_1 \cdot p_2$.

Let E_1 denote any sample values of the observable variables $X_1, \ldots, X_N$, and E_2 denote any sample values of the future variables $X_{N+1}, \ldots, X_{N+M}$. Any E_1 can be represented by a point in the N dimensional sample space R_1 of the variables $X_1, \ldots, X_N$, and any E_2 can be represented by a point in the M dimensional sample space R_2 of the variables $X_{N+1}, \ldots, X_{N+M}$. Similarly, we let E denote a point in the sample space R of all $N + M$ variables. Now, given any particular E_1, one can calculate from p_2 the probability that E_2 will fall in a given point set of the sample space R_2. The resulting probability would be a function of E_1. Furthermore, for any given E_1 and any given probability level P, one can derive a system of point set regimes in R_2 with probability of E_2 falling in one of such sets equal to P. Any such point set in R_2 is referred to as a region of prediction and denoted by W_2. One is usually interested in a region W_2 of probability P, which is in some sense the "narrowest" possible. The choice of probability level and region W_2 will depend on the particular intended use, and such choice is therefore is not a problem of statistics.

If p_2 is known, the problem of prediction is one of probability calculus and not one of statistical inference from a sample. In practice, p_2 is unknown and information about p_2 must be obtained from samples E_1 of previous observation. This procedure is made possible by the following important basic assumption:

> The probability law, p, of the $N+M$ variables $X_1, \ldots, X_N, X_{N+1}, \ldots, X_{N+M}$ is of such a type that the specification of p_1 implies the complete specification of p and, therefore, of p_2.

That is, if p is characterized by a number of unknown parameters, then all these parameters must also characterize p_1 so that p_2 contains no other parameters. This assumption therefore implies that for prediction to be possible, a certain persistence in the mechanism which produces the data must be present.

Haavelmo also describes a method by which to derive prediction formulae. Given that E_2 denotes a point in the sample space R_2 of $X_{N+1}, \ldots, X_{N+M}$, we denote by $\hat{E}_2$ a point in R_2 to be used as a prediction of E_2. The problem is one of defining $\hat{E}_2$ as a function of $X_1, \ldots, X_N$, such that the probability of $\hat{E}_2$ being close, in some sense, to E_2 is high. $\hat{E}_2$ is called a prediction function. Furthermore, one can assign a system of weights to the possible errors in prediction by defining a weight or loss function $\mathcal{L}(E_2, \hat{E}_2)$, such that $\mathcal{L}(E_2, E_2) = 0$ and $\mathcal{L}(E_2, \hat{E}_2) > 0$ for $E_2 \neq \hat{E}_2$. The expected value of the loss function in repeated samples is given by:

$$r = \int_R \mathcal{L}(E_2, \hat{E}_2) p dE.$$

The choice of $\hat{E}_2$ as a function of $X_1, \ldots, X_N$ should be so that r is as small as possible. The problem of deriving the best prediction function is closely related to the problem of deriving best estimates. Although there is always some level of subjectivity when it comes to choosing a prediction function and loss function, the procedure given by Haavelmo describes precisely where and how the subjective elements enter the prediction problem.

Haavelmo's interpretation of economic processes as realizations of stochastic processes rather than realizations of independent processes gave way to the acceptance of probability theory for modeling in economics. Furthermore, his methodology for prediction based on the concepts of probability laws, prediction formulae, and loss functions set forth the development of mathematically precise protocols to study the validity of forecasting.

By the end of the 1940s, Haavelmo's probability approach had been generally accepted in the USA, and became the basis for the macroeconomic model built by Lawrence R. Klein for the Cowles Commission in 1950 [85]. Klein recognized the importance of the contributions made by Tinbergen in his two League of Nations reports, and considered his own work an extension of Tinbergen's work. The structural form of Klein's models also reflects the influence of Keynes' "General Theory." Klein sought to emphasize the discovery of economic theories through his models as well as performing forecasts.

> If we know the quantitative characteristics of the economic system, we shall be able to forecast with a specified level of probability the course of certain economic magnitudes such as employment, output, or income; and we shall also be able to forecast with a specified level of probability the effect upon the system of various economic policies. ([85], p.1)

Klein considers as his main contribution the ability to accept or reject admissible hypotheses of economic theory based on their suitability for the purpose of forecasting. Klein classifies the variables to be used in the model as endogenous or exogenous. Endogenous variables are those determined by the economic system and include output, employment, prices, profits, rents, investment. Exogenous variables are those representing forces outside the economic system such as those originating from natural, technological, sociological, political, or institutional events. Klein argues, economists have developed theories of economic behavior which can be used to determine the endogenous variables and their relations expressed as structural equations. Klein defines $y_{i,t-k}$ as the ith endogenous variable in the $t-k$ period, z_i as the ith exogenous variable, u_{it} as the ith random disturbance of the tth period and the model of the economic system is given as

follows:

$$f_i(y_{1,t},\ldots,y_{n,t},\ldots,y_{1,t-p},\ldots,y_{n,t-p},z_1,\ldots,z_m) = u_{it}, \quad i=1,\ldots,n. \tag{2.2.1}$$

The f_i functions define the structural equations, which equal in number to the endogenous variables, and the econometric problem of interest is the estimation of the structural parameters of the f_i functions. Klein also offers an alternative problem when the main aim at hand is forecasting rather than explanation and description. The argument for the alternative procedure is that not all structural parameters in (2.2.1) might be needed to construct a forecast. Klein solves (2.2.1) for the endogenous variables to be forecasted, such that the new set of equations, referred to as the reduced form, are as follows:

$$y_{it} = g_i(y_{1,t-1},\ldots,y_{n,t-1},\ldots,y_{1,t-p},\ldots,y_{n,t-p},z_1,\ldots,z_m,u_{1t},\ldots,u_{nt}), \quad i=1,\ldots,n. \tag{2.2.2}$$

The parameters of (2.2.2) will be different from the parameters of (2.2.1). Klein studies three statistical models. The first of these models is a simple three equation system by which he "sacrificed details of economic behavior patterns in order to illustrate different methods of structural estimation in dynamical economic systems" (p. 84). In his second model, Klein estimated parameters which were deemed necessary for purposes of forecasting. Finally, in his last model, Klein developed the same procedures, but for a large structural model of the economy.

The importance of econometric modeling and forecasting was further strengthened by the work of H. Theil in the 1960s [137, 138]. In [137], Theil outlines the three main problems of forecast analysis: verification and accuracy analysis; the analysis of the generation of predictions; and the use of forecasts for policy purposes. Furthermore, Theil provides new measures to evaluate forecast accuracy with empirical application for the Dutch and Scandinavian economies. Theil also addresses two problems of methodology: the particular type of data analyzed, and statistical inference. For the problem of statistical inference, Theil discusses the desirable properties of econometric and statistical approaches, and generalizes the method of least-squares for the complications of autocorrelated disturbances and simultaneous equations. Finally, Theil turns to the problem of determining the relationship between forecasting and policy by addressing the uncertainty characterizing decision processes. In [138], Theil deals with general problems of methodology and the consequences of prediction errors at the decision-making level. In several chapters, Theil introduces information theory as a tool for evaluation of forecasts and to deal with data obtained from surveys.

In the post-war period, apart from the work of Klein and Theil, the development of theoretical methods for forecasting focused on time series analysis [32]. Among others, the work on time series analysis can be exemplified by that of Wiener [156], Kalman [82], Whittle [154], Box and Jenkins [25] and Harvey [67, 68]. Also, by the end of the 1970s, Keynesian macroeconomic models such as those of Tinbergen, Kelin, and Theil were in decline, as was structural Keynesian macroeconomic forecasting [41]. In response to the failures of Keynesian structural models, econometricians began to explore nonstructural forecasting methods. Work on nonstructural methods predates the Keynesian period, but this work was overlooked mainly for the popularity of Keynesian methods. Beginning in the 1920s, the work of Slutsky [132] and Yule [161] focused on the use of simple linear difference equations driven by random stochastic shocks, autoregressions, for modeling and forecasting a variety of economic and financial time series [41]. The key insight in the use of autoregressions is that system dynamics convert random inputs into serially correlated outputs, a phenomenon called the Slutsky-Yule effect. In the 1930s, H. Wold [158] made a ground breaking contribution by showing that given sufficient stability of the underlying probabilistic mechanism generating the series, the stochastic part can be represented as a model of the Slutsky-Yule type. N. Wiener [156] and A. Kolmogorov [88, 89] worked out the mathematical formulae for optimal forecasts from models of the type studied by Slutsky, Yule, and Wold. In the late 1950s and late 1960s, R. Kalman extended the theory by relaxing conditions imposed by Wiener and Kolmogorov. His forecasting formula is known as the Kalman filter, which is designed to work with a state-space representation of the system. The Wold-Wiener-Kolmogorov-Kalman theory is exposited in Whittle [155]. A major push in the direction of nonstructural methods came in 1970 with the publication of Box and Jenkins' book [25] on nonstructural time series analysis and forecasting.

Box and Jenkins' model allowed for stochastic trends to be driven by cumulative effects of the random shocks, rather than just modeling trends via a linear deterministic function of time. The concept of stochastic trends had wide-range implications, since shocks to series have permanent effects. The most important contribution of the Box and Jenkins methodology consists of a framework for nonstructural forecasting formulated as iterative cycles of model formation, estimation, diagnostic testing, and forecasting. The main tool at the core of the Box-Jenkins framework are autoregressive moving average (ARMA) models. The need for modeling cross-variable relationships in macroeconomics led to the expansion of the Box-Jenkins program by the creation of vector autoregressions (VAR) to handle multivariate modeling and forecasting. VAR models are less restrictive than the system-of-equations used in structural models, because variables do not need to be label as endogenous or exogenous. Instead, with VAR models, all variables are considered to be endogenous. Early contributions to multivariate work of time series include the work of Granger [57] and Sims [130, 131]. Dynamic factor models originated from a need to make VAR models more flexible. In dynamic factor models, some economic shocks are common across sectors while others are particular to only a few sectors. Contributions to dynamic factor models include the work of Sargent and Sims [127], Geweke [54], Stock and Watson [134, 135], Quah and Sargent [118], and Forni and Reichlin [50]. The concept of cointegration, where two or more series contain a stochastic trend but their linear combination does not, was developed by Granger [58], and Engle and Granger [48].

As for nonlinear models, one of the most important applications of the time series Box-Jenkins methods is the modeling of volatility dynamics, which allows forecasting of the unobservable volatility of observable processes. The literature of volatility forecasting began with the seminal papers of Engle in 1982 [47] and Bollerslev in 1986 [22]. Their models allow the conditional variance of the shocks to vary with time, as a function of past errors in the case of the former, and as a function of past errors and past conditional variances in the case of the latter. These nonlinear models have become of great importance in finance, and extensive surveys of volatility forecasting include Bollerslev, Chou and Kroner [23], Bollerslev, Engle and Nelson [24], and Poon and Granger [117]. A second important category of nonlinear time series models is regime-switching models. In regime-switching, or threshold models, an indicator variable determines the occurrence of a switch. Important contributions to threshold models include the work of Tong [142], Granger and Teräsvirta [60], and Hamilton [65].

We thus conclude this survey of some of the most important developments and contributions in economics and econometrics to the problem of forecasting. This survey, although not exhaustive, attempts to give a taste of the progression in modeling and forecasting that has led to the methodology applied in the work to follow. We note many important areas of research have not been covered, such as neural networks and machine learning. We refer the interested reader to other more extensive surveys of economic forecasting [9], [147].

2.3 Forecasting methodologies

We next provide an overview of forecasting methodologies, and extensively describe the one particular methodology which is put into practice in the core of this thesis. A most extensive catalog of forecasting methodologies can be found in Armstrong's book [9], and the following sketch of methodologies is based on his work.

Forecasting methodologies can be categorized into two classes: judgmental and statistical. The first class of methodologies described by Armstrong, judgmental methodologies, include role playing, intentions, and expert opinions. Role playing is a forecasting methodology which attempts to predict decisions and actions of people and groups by requiring participants to act and respond to fictitious situations that replicate possible conflicts. Role playing is most effective in prediction when the conflicting parties must respond to large changes. Examples of situations where role playing might be applicable include companies designing product and predicting consumer reactions, labor issues, military strategies, forming strategies in court cases, and negotiating contracts. Intentions, as a methodology, outlines procedures to use individuals' plans, goals, or expectations about the future to forecast individuals' actions. Basic principles of intentions measurement require that intentions should be quantified using probability scales, that intentions should be adjusted to remove biases, that respondents should be segmented, and that intentions can be used to form best and worst case forecasts. Intentions can be applied to problems such as marketers measuring consumers' purchase intentions, and the design of political polls. Expert opinion, as a forecast methodology, consists of principles using a collection of experts' forecasts to construct one unifying forecast.

The second class of forecasting methodologies outlined by Armstrong, statistical methodologies, can be divided into two subcategories: extrapolation models and econometric models. These two subcategories have been addressed in the history survey of the previous section. Extrapolation models are also known as nonstructural models, and econometric models are also referred to as structural models.

Armstrong [11] presents an extensive account on principles and strategies for forecasting with extrapolation models, and the following is a summary of time series models. The main principle behind extrapolation of time series is that all necessary information is contained in the historical values of the time series being forecasted, while the principle behind cross-sectional extrapolation is that characteristics of one set of data can be generalized to another set. The strengths of using extrapolation of time series are that past behavior tends to be a good indicator of future behavior, it is objective, it is replicable, and its is inexpensive. Time series extrapolation is also known as univariate time series forecasting. Armstrong's first principle for extrapolation of time series is that, when selecting data, one should use all relevant data and adjust the data for important past events. Second, one should make seasonal adjustments when seasonal effects are expected. A third principle, when extrapolating, is the use of simple functional forms. By far, the most influential models of time series are the univariate models proposed by Box and Jenkins [25]. Most time series models can be expressed as Box-Jenkins models. The dominant class of scalar time series models are integrated autoregressive moving average models (ARIMAs). There are several reasons for the success of the Box-Jenkins framework. Generally, the order of the AR and MA polynomials required for adequate fit of time series is relatively low. Many economic time series are non-stationary but in many cases can be made stationary by differencing; in such cases, ARIMA models are amenable for analysis. Excellent surveys of the Box-Jenkins framework include [25, 27, 64, 67].

The following summary of principles and strategies for econometric forecasting is based on the work of P. Allen and R. Fildes [2]. At the core of econometric methods lie statistical procedures which are employed to estimate models specified primarily by economic theory. Early econometric models focused on collecting as many casual variables as possible, if deemed by theory as relevant. This strategy led to much failure because little attention was given to the dynamic structure. The application of vector autoregression (VAR) methods in the 1980s resolved much of the problem. Contemporary econometricians use economic theory as a guide to describe long-term cause and effect relationships, and use data to determine the structure of the model, in terms of lags on variables and differencing, which best describes the short-term dynamics. The principal tool available to the econometrician is regression analysis. Allen and Fildes suggest the fundamental principle for econometric forecasting is to aim for a relatively simple model specification. We now describe an eight-step strategy for forecasting, as proposed by Allen and Fildes, based on time series econometrics. The eight steps comprise: defining the objectives, determining the set of variables, collecting the data, forming an initial specification, estimating the model, misspecification testing, model simplification, and comparing the out-of-sample performance.

By defining the objective, Allen and Fildes refer to deciding whether the purpose of the study is to explain or to forecast. For the purpose of explanation, such as analyzing policy, model structure is the important factor, and conditional forecasts should be used to test the model. For the purpose of forecasting, one must be able to forecast the explanatory variables used in the model with certain level of accuracy. When it comes to determining the set of variables to be included in the model, it is suggested that one considers casual variables based on guidelines from theory and previous empirical research. Armstrong [10] gives four criteria for including a variable in a model:

1. a strong casual relationship is expected,
2. the casual relationship can be estimated accurately,
3. the casual variable changes substantially over time,
4. the change in the casual variable can be forecasted accurately.

For collecting data, Allen and Fildes suggest gathering all data available. This does not imply that all data is ultimately used for estimation and forecasting, but rather, the claim is that knowledge of factors such as structural breaks can result in improved models and superior forecast accuracy.

Once the list of variables to be used in the model has been determined, in the step of initial specification, the econometrician designates the variables that occur in a specific equation and the functional form of the equation. Part of determining the functional form consists of deciding on the number of lags on each variable. Determining what variables to include in an equation is usually based on theory. The use of a vector autoregression model avoids the task of assigning variables as dependent or explanatory, since each left hand side variable depends on lags of itself and the other variables on the right hand side. Allen and Fildes suggest one must take into account all previous work when specifying a preliminary model. This concept of encompassing can be described as follows: a theory encompasses a rival theory if the former explains at least as much as the latter explained. [106] and [49] are examples of work on forecast encompassing. The common approach used by time series econometricians (e.g., [70],[69]) to model building relies upon a general-to-specific principle. In this approach, a model with certain degree of generality is tested for misspecification, and failure leads to a new simpler model for testing.

For the step of estimation, Allen and Fildes suggest there seems to be no advantage in using any other procedure other than ordinary least squares (OLS). Some support for this conclusion is that OLS seems to be robust to violations of underlying assumptions. OLS has stood up well against theoretically superior estimation methods. In the case of estimating systems of equations, OLS is biased, but according to Kennedy [83], this bias is not much worse that that of other methods. OLS is robust to misspecification, and OLS has the smallest variance among estimators. Monte Carlo studies have shown OLS to be less sensitive than other estimators to problems of multicollinearity, errors in variables, and misspecification in small samples. Dielman and Rose [42] compare out-of-sample forecasts from OLS, least absolute value (LAV), and Prais-Winsten methods on a bivariate model with first order autocorrelated errors and find that OLS was frequently better.

Once a model has been estimated, misspecification tests can be applied. The failure of a specification test is an indication that the model as estimated is an inadequate summary of the data. Unfortunately, Allen and Fildes point out, there is not much evidence to tie misspecification tests to forecasting performance. Some econometricians view the failure of a misspecification test as a reason to explore new specifications rather than focus on new estimation methods. Such econometricians view theory as a guide, although incomplete, for selecting casual variables, and consider testing essential in the construction of models. When a model fails a number of misspecification tests, the econometrician must consider additional casual variables, restructure the dynamic interdependencies, or re-evaluate the functional form. Some important misspecification tests include parameter stability, specification error, omitted variables, nonlinearities, autoregressive residuals, and linear versus log-linear specification. Once a model satisfies a number of misspecification tests, one can consider simplifying the model. As mentioned, for the purpose of forecasting, one should aim towards simplicity rather than correct specification. In time series, reducing the lag length is the primary method of simplification and should be done one equation at a time in VAR models. Beginning with a general equation, reducing the lag successively guarantees the residual sum of squares of the new restricted model will not be statistically worse than the residual sum of squares for the previous more general model. Finally, it is important to test model performance with data not used for estimation. This out-of-sample forecasting method gives clues to the generality of the model since, it might do well in explaining the past but it may perform poorly in predicting the future. Much of the work presented in this thesis is mainly concerned with univariate time series in the Box-Jenkins framework, although some treatment of multivariate processes is presented in Chapter 5. The reason for restricting mainly to univariate processes is to maintain simplicity in computation and exposition. There are no theoretical obstructions to expand the computational work to VAR models.

Econometricians making use of ARIMA or VAR models face four main sources of error. Specification error can be present due to inappropriate choice of explanatory variables, use of an incorrect functional form, or the presence of structural breaks. Conditioning error results from inaccuracies in the information used to form the conditional forecast. When constructing a forecast, parameters are estimated based on a sample of observations; the inaccuracies involved in estimating these parameters result in sampling error in the forecast. Finally, random error is present in a forecast, even under correct specification, due to the residuals used in the modeling and estimation.

The evaluation of forecasts is a critical step that must be carried out before implementing a forecast. Clements and Hendry [32] provide a complete and systematic treatment of forecast evaluation for time series models, and we summarize their primary principles. Granger and Newbold [59] presented a critique of evaluation methods available at the time, and Clements and Hendry summarize the main contention:

> Methods for gauging forecast accuracy cannot usefully be based on comparison of the time series, or the distributional properties, of the actual and predicted series. It makes more sense to analyze the difference between the two.

A general criterion to measure ex post forecast accuracy, based on the actual values (A_t) of a series and the predicted values (P_t), can be given as follows:

$$I(P_t, A_t). \tag{2.3.1}$$

An optimal prediction is one for which (2.3.1) obtains an extremum. Based on the main contention of Granger and Newbold, the criterion can be made more specific by writing it as follows:

$$I(P_t, A_t) = I(A_t - P_t, A_t) = I(\epsilon_t, A_t) = C(\epsilon_t), \tag{2.3.2}$$

with $\epsilon_t = A_t - P_t$, and the costs are only a function of the forecast error, ϵ_t. If $C(\cdot)$ is a quadratic function, the criterion is squared in the error and averaging over errors leads to the mean square forecast error (MSFE) criterion. Reasons for choosing a quadratic form for C include mathematical tractability, large errors are proportionately more serious than small errors, and in many situations over and under prediction have similar costs. We list other measures of forecast accuracy:

1. Mean absolute error (MAE): This is the average of the absolute values of the forecast error, and is best applicable when the cost of forecast errors is proportional to the absolute size of the forecast error.

2. Root mean square error (RMSE): This is the square root of the average of the squared values of the forecast error. This measure implicitly weights large errors more than small errors. This is simply the square root of the MSFE.

3. Mean absolute percentage error (MAPE): This is the average of the absolute values of the percentage errors. It is dimensionless, and its use is appropriate when the cost of the error is closely related to the percentage error, rather than to the numerical size of the error.

4. Median absolute percentage error (MdAPE)

5. Relative absolute error (RAE): This measure compares the error for a proposed forecasting model to that for the naive forecast.

6. Correlation of forecasts with actual values: In this measure, changes, rather than levels of the variable being forecasted are regressed on the forecasts of these changes and the resulting R^2 is used as a measure of forecast accuracy. Armstrong [9] warns against using R^2 to compare forecasting models.

7. Conditional efficiency: A forecast A is conditionally efficient relative to forecast B if B contributes no useful information beyond that contained in A, and can be evaluated by regressing the variable being forecasted on A and B and testing the null that the coefficient of B is zero.

The work in this thesis evaluates forecasts using the MSFE. One reason for using the MSFE is its computational tractability. Another reason for using the MSFE is due to the generality of our methods. Since no specific economic phenomena is considered in developing our algorithms, we select the MSFE for its generality over other context-specific loss functions.

The first assumption we adhere to in the work to follow is that the observed process to be forecasted originates from a data generating process (DGP) which might depend on a parameter vector $\theta \in \Theta \subset \mathbb{R}^k$. Clements and Hendry ([32], p.11) present a framework for the forecasting problem with six facets: (A) the nature of the DGP; (B) the knowledge level about the DGP; (C) the dimensionality of the system to be studied; (D) the form of the analysis; (E) the forecast horizon; and (F) the linearity of the system. The principal aim of this thesis is to develop algorithms, under the assumption of unknown DGP and unknown θ, for different dynamic structures of the DGP.

2.4 Forecast problem

As described in section 2.3, the two principal forecasting methodologies used in economics and finance are univariate time series models in the Box-Jenkins tradition, and vector autoregressive (VAR) econometric models. In what follows, we describe the forecast problem of interest. The scope of our approach in constructing the problem is general enough to allow for application of both the univariate time series and VAR methodologies.

2.4.1 Notation and setup

Consider a stochastic process $Z \equiv \{Z_\tau : \Omega \longrightarrow \mathbb{R}^{m+1}, m \in \mathbb{N}, \tau = 1, \dots, T+1\}$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\mathcal{F} = \{\mathcal{F}_\tau, \tau = 1, \dots, T+1\}$ and $\mathcal{F}_\tau$ is the σ-field $\mathcal{F}_\tau \equiv \sigma\{Z_s, s \leq \tau\}$. In what follows, we denote by Y_τ the component of interest of the observed vector Z_τ, $Y_\tau \in \mathbb{R}$, and interpret the remaining components, denoted W_τ, as being an $m \times 1$ vector of other variables. In other words, we let $Z_\tau \equiv (Y_\tau, W_\tau^\top)^\top$. The random variable Y_τ is further assumed to be continuously distributed.

The forecasting problem considered involves forecasting the variable Y_{t+s}, where s is the prediction horizon of interest, $s \geq 1$, and t is the forecast origin with $t < T$. In what follows, we set $s = 1$ and examine the one-step-ahead predictions of Y_{t+1}, knowing that all results developed in this case can readily be generalized to any $s > 1$. In standard notation, the subscript τ on the expectation, $E_\tau[\cdot]$, denotes conditioning on the entire information set $\mathcal{F}_\tau$. In particular, we shall assume the forecaster employs the expected value of Y_{t+1} conditional on the entire information set $\mathcal{F}_t$, $E_t[Y_{t+1}]$ to specify the forecast model. We denote by X_t an $m \times 1$ column vector of $\mathcal{F}_t$-measurable variables that are used to forecast Y_{t+1}, $X_t = (X_t^1, \dots, X_t^m)^\top$. For the case $m = 1$, $X_t = X_t^1$. In applications, X_t can contain (1) various lags of the variable of interest Y_τ, (2) realizations of the other variables W_τ, as well as (3) any function of the previous two. As such, our setup will allow for applications involving both time series and cross-section data. In what follows, we use the following notation for the time series $\{Y_\tau\}_{\tau=t-n+1}^{t}$ and $\{X_\tau\}_{\tau=t-n}^{t-1}$:

$$\begin{aligned} X_{t,n} &\equiv (X_{t-n}, ..., X_{t-1})^\top \in \mathbb{R}^{n\times m}, \\ Y_{t,n} &\equiv (Y_{t-n+1}, \ldots, Y_t)^\top \in \mathbb{R}^{n\times 1}, \\ Q_{t,n} &\equiv X_{t,n}^\top X_{t,n} \in \mathbb{R}^{m\times m}. \end{aligned}$$

We assume the forecaster does not know the data generating process (DGP) responsible for the observed time series $\{Y_\tau\}$. Instead, she uses some, possibly misspecified, forecasting model to produce her forecasts, which are then evaluated using a loss function $\mathcal{L}$. In practice, the most commonly encountered situation is the one in which the forecasting model employed is linear and the loss function is quadratic. In what follows, we derive the mean square forecast error (MSFE) for linear forecasting models under the possible presence of model misspecification.

2.4.2 Forecast construction

As mentioned, we assume the forecaster specifies the forecast model based on the conditional expectation $E_t[Y_{t+1}]$ of the observed process $\{Y_\tau\}$. This is as a consequence of the well known fact that the prediction with the smallest MSFE is, in fact, the conditional expectation $E_t[Y_{t+1}]$, ([64], p. 72). For example, for a DGP with additive innovations of the form

$$DGP : Y_{t+1} = \psi(X_t) + U_{t+1}, \tag{2.4.1}$$

where $\{U_\tau\}$ is the innovation process with $E_\tau[U_\tau] = 0$ and $Var(U_\tau) = \sigma_U^2 < \infty$ for all τ, $E_t[Y_{t+1}] = \psi(X_t)$. It is common practice in econometrics to assume a specific form for $E_t[Y_{t+1}]$ in estimation and forecasting. In what follows, we use the notation $\Psi_{t,n} \equiv (E_t[Y_{t-n+1}], \ldots, E_t[Y_t])^\top \in \mathbb{R}^{n\times 1}$ and $U_{t,n} = (U_{t-n+1}, \ldots, U_t)^\top \in \mathbb{R}^{n\times 1}$, so that

$$Y_{t,n} = g(\Psi_{t,n}, U_{t,n}), \tag{2.4.2}$$

for some functional form g. Not knowing the exact form of the DGP, we assume the forecaster's prediction of Y_{t+1} is based on a model for $E_t[Y_{t+1}]$ which is linear in X_t, and an innovation process $\{V_\tau\}$ such that

$$Y_{t+1} = \beta^\top X_t + V_{t+1}, \tag{2.4.3}$$

where β is an $m \times 1$ parameter vector, $\beta \in B$, B compact in $\mathbb{R}^m$, and V_τ is such that $E_\tau[V_\tau] = 0$.

It is important to note that — while being linear — the forecasting model is not assumed to be correctly specified. In other words, we do not make the assumption that $E_t(Y_{t+1})$ is a linear function of X_t. In fact, a major aim of the work in this thesis is to investigate the ramifications of the phenomena of misspecification in the context of forecasting. Misspecification of the forecasting model can result from a variety of causes. For example, in her choice of X_t, the forecaster might omit some of the $\mathcal{F}_t$–measurable variables that enter $E_t(Y_{t+1})$; in this case, the forecasting model is dynamically misspecified. Moreover, even if $E_t(Y_{t+1})$ is a function of X_t alone, its functional form might be highly nonlinear; in this case the forecasting model is functionally misspecified.

The parameter β is assumed to be estimated by an ordinary least squares (OLS) estimator. The OLS estimator of β can be computed by using sample sets of various sizes. When the sample set is a continuous interval in time, we refer to it as an observation window. In the construction of a forecast, one important aspect to determine is the nature of the sample used for the estimation of the forecast model. In the case of an observation window, this corresponds to determining its length. In chapters to follow, we develop quantitative methods for determining the length of an observation window used in the forecasting problem. Two prevalent methods found in the literature are: (1) a rolling window forecasting scheme, or (2) a recursive (also known as expanding window) forecasting scheme. Under the rolling window forecasting scheme, the forecaster re-estimates the parameter β of the linear forecasting model in (2.4.3) at each point t, $T - R \leq t < T$. The estimation sample contains the n most recent observations—X_{t-n} to X_{t-1} and Y_{t-n+1} to Y_t— so the OLS estimator of β has the form

$$\hat{\beta}_{t,n} \equiv Q_{t,n}^{-1} X_{t,n}^\top Y_{t,n}. \tag{2.4.4}$$

For example, in the single regressor case, the above expression for $\hat{\beta}_{t,n}$ reduces to $\hat{\beta}_{t,n} =$

$(\sum_{s=t-n}^{t-1} X_s^2)^{-1} \ (\sum_{s=t-n}^{t-1} Y_{s+1}X_s)$. The above OLS estimator $\hat{\beta}_{t,n}$ is then used to construct the forecast $\hat{Y}_{t+1,n}$ of Y_{t+1} as follows

$$\hat{Y}_{t+1,n} = \hat{\beta}_{t,n}^{\top} X_t. \tag{2.4.5}$$

This procedure is repeated R times over the out-of-sample period $[T - R, T]$, and the forecaster re-estimates β each time there are new observations available. The value of n — which enters the forecast $\hat{Y}_{t+1,n}$ through the OLS estimator $\hat{\beta}_{t,n}$ — is most often chosen in an ad hoc manner, since there are no systematic methods in the literature to obtain an optimal value.

The recursive (or expanding) window scheme involves using all past observations available, i.e., the observations from date 1 to t. Hence, if the forecaster uses a recursive window forecasting scheme, at any time t, $T-R \leq t < T$, she computes $\hat{\beta}_{t,t} \equiv Q_{t,t}^{-1} X_{t,t}^{\top} Y_{t,t}$, and constructs $\hat{Y}_{t+1,t} = \hat{\beta}_{t,t}^{\top} X_t$. In other words, the recursive scheme corresponds to the case where $n = t$ in the OLS expression (2.4.4) above. As previously, the OLS estimator $\hat{\beta}_{t,t}$ is computed T times, only now the estimate of β relies on all the data prior to time t. Both the rolling window and recursive forecasting schemes have great shortcomings. For instance, neither of these schemes is likely to be optimal if the DGP for the time series $\{Y_s\}$ undergoes a structural break. A rolling window of a short fixed size might work well immediately after the break but valuable information will be lost as the distance from the break increases. The recursive scheme will produce significantly biased forecasts after the break, until the post break information out weighs the pre-break information. It is our ultimate goal to develop and evaluate a new optimal forecasting scheme which relies on the nature of the processes $\{Y_s\}$ and $\{X_s\}$ for the choice of the forecasting window. Before tackling this in the chapters that follow, we examine forecast evaluation based on the decomposition of the MSFE.

2.4.3 Forecast evaluation

In our evaluation of the accuracy of the forecasts $\hat{Y}_{t+1,n}$, we abide by common practice, and represent the accuracy criterion by means of a cost or loss function. Assuming the forecast evaluator uses a quadratic loss function, an optimal forecasting scheme consists of minimizing mean square forecast error (MSFE). Hence, we are interested in examining the dependence of the expected squared forecast error on the window size n. Following the standard approach [56, 32], the expected squared forecast error can be defined in one of two ways, depending on its intended use. For calculating specific errors given past realizations of the explanatory variables, $\mathcal{X}_t = \sigma\{X_{t-n}, \ldots, X_t\}$, we define the criterion

$$CMSFE_{t,n} \equiv E[(Y_{t+1} - \hat{Y}_{t+1,n})^2 | \mathcal{X}_t], \tag{2.4.6}$$

where $\hat{Y}_{t+1,n}$ is as defined in (2.4.5). We refer to this criterion as the conditional MSFE. On the other hand, if we wish to analyze general properties of the MSFE, independent of specific realizations of the explanatory variables, the unconditional MSFE or simply the MSFE, is given by

$$MSFE_n \equiv E[(Y_{t+1} - \hat{Y}_{t+1,n})^2] = E[\epsilon_{t+1,n}^2], \tag{2.4.7}$$

where $\epsilon_{t+1,n}$ is the time-$t+1$ forecast error, $\epsilon_{t+1,n} \equiv Y_{t+1} - \hat{Y}_{t+1,n}$. In the work to follow, as in [113, 114], we use the latter form of the MSFE for forecast accuracy evaluation.

2.4.4 Decomposition of the MSFE

It is common for analysis to decompose the MSFE into component parts. The squared bias and variance decomposition consists of the sum of two terms, as traditionally done in the forecasting literature (see, e.g., [56, 32, 113]), and has the following form:

$$E[\epsilon_{t+1,n}^2] = b_n^2 + v_n, \tag{2.4.8}$$

where $b_n^2 \equiv (E[\epsilon_{t+1,n}])^2$ is the squared bias of the forecast error, and $v_n \equiv Var(\epsilon_{t+1,n})$ is the variance of the forecast error. (2.4.8) is easily derived from the definition of the variance.

Writing the MSFE as the sum of the squared bias and variance of the error allows for a revealing analysis of the first two moments of the error in the forecast. The bias term refers to the level of model misspecification in the forecast, while the variance captures the level of homogeneity in the processes. Both the bias and variance terms are affected by the accuracy of the estimator employed. In what follows, we present some properties of the CMSFE and the MSFE which concern their sample size dependence (SSD), i.e., properties regarding the observation window size n.

We assume the DGP has the general form $Y_{t,n} = \Psi_{t,n} + U_{t,n}$, and rewrite the OLS estimator in (2.4.4) as $\hat{\beta}_{t,n} = \Theta_{t,n} + \Lambda_{t,n}$, where

$$\Theta_{t,n} \equiv Q_{t,n}^{-1}\, X_{t,n}^{\top}\, \Psi_{t,n}, \quad \Lambda_{t,n} \equiv Q_{t,n}^{-1}\, X_{t,n}^{\top}\, U_{t,n}. \tag{2.4.9}$$

The forecast error evaluated at $t+1$ is given by

$$\epsilon_{t+1,n} = \psi(X_t) + U_{t+1} - (\Theta_{t,n} + \Lambda_{t,n})^{\top} X_t. \tag{2.4.10}$$

The CMSFE can be written as the sum of a conditional squared bias term and a conditional variance term as follows:

$$CMSFE_{t,n} = b^2_{\mathcal{X}_{t,n}} + v_{\mathcal{X}_{t,n}}, \tag{2.4.11}$$

$$b^2_{\mathcal{X}_{t,n}} = E^2[\epsilon_{t+1,n}|\mathcal{X}_t] = \Big(\psi(X_t) - \Theta_{t,n}^{\top} X_t\Big)^2, \tag{2.4.12}$$

$$v_{\mathcal{X}_{t,n}} = Var(\epsilon_{t+1,n}|\mathcal{X}_t) = \sigma_U^2 + \mathrm{var}(\Lambda_{t,n}^{\top} X_t) = \sigma_U^2 + \sigma_U^2 X_t^{\top} Q_{t,n}^{-1} X_t. \tag{2.4.13}$$

It is clear that both components depend on the particular realization $\mathcal{X}_t$. The following proposition describes the n dependence of the conditional variance component.

Proposition 2.1

(i) For a given realization $\mathcal{X}_t$, $v_{\mathcal{X}_{t,n}} \downarrow n$.

(ii) For a correctly specified linear model and a given realization $\mathcal{X}_t$, the optimal forecasting scheme is recursive.

The proposition implies the variance decreases as the amount of data used to form the forecast increases. To gain some intuition on the variance decay with n, consider the scalar case $m = 1$. In this case, $Q = \sum_{i=t-n}^{t-1} X_i^2$ and the conditional variance is

$$v_{\mathcal{X}_{t,n}} = \sigma_U^2 + \sigma_U^2 (\sum_{i=t-n}^{t-1} X_i^2)^{-1} X_t^2. \tag{2.4.14}$$

The variance decay is clear from the fact the denominator increases as n increases, while the numerator is constant. The monotonic behavior of the conditional variance suggests that any interesting behavior of the CMSFE as the sample size increases is due entirely to the conditional squared bias term.

The conditional squared bias component for a misspecified model (2.4.12), on the other hand, does not exhibit a clear monotonic dependence on n. In fact, the conditional squared bias for a misspecified model inherits the erratic nature of the particular realization $\mathcal{X}_t$, making the CMSFE unfit for any analysis of an optimal observation window. We can see this clearly in the scalar case $m = 1$ where the term $\Theta_{t,n}^{\top} X_t$ in the conditional squared bias is given by

$$\Theta_{t,n}^{\top} X_t = X_t (\sum_{i=t-n}^{t-1} X_i^2)^{-1} \sum_{i=t-n}^{t-1} \psi(X_i) X_i. \tag{2.4.15}$$

The absence of an n dependent decay in the squared bias can be seen by comparing (2.4.14) and (2.4.15). The following example illustrates these ideas.

Example 2.1 *Consider the nonlinear univariate DGP given by* $Y_{t+1} = X_t^2 + U_{t+1}$*, where* $\{U_\tau\} \sim IIN(0,1)$*. Furthermore, assume the process* $\{X_\tau\}$ *follows an AR(1):* $X_{t+1} = (1-a) + aX_t + V_{t+1}$*, where* $a = 0.9$ *and* $\{V_\tau\} \sim IIN(0, 0.4)$*. We investigate the SSD of the conditional variance and the conditional bias through a Monte Carlo experiment for three realizations of the process* $\{X_\tau\}$*. The results of the experiment, given in figure* 2.1*, show the erratic nature of the conditional squared bias component.* □

Due to the failure of the CMSFE in revealing optimal forecasting schemes, we turn to the unconditional MSFE as defined in (2.4.7), written in terms of the squared bias and variance components. The forecast error evaluated at $t+1$ given by (2.4.10) leads to the unconditional squared bias component

$$b_n^2 = E^2[\epsilon_{t+1,n}] = \Big(E[\psi(X_t)] - E[\Theta_{t,n}^{\top} X_t] \Big)^2, \tag{2.4.16}$$

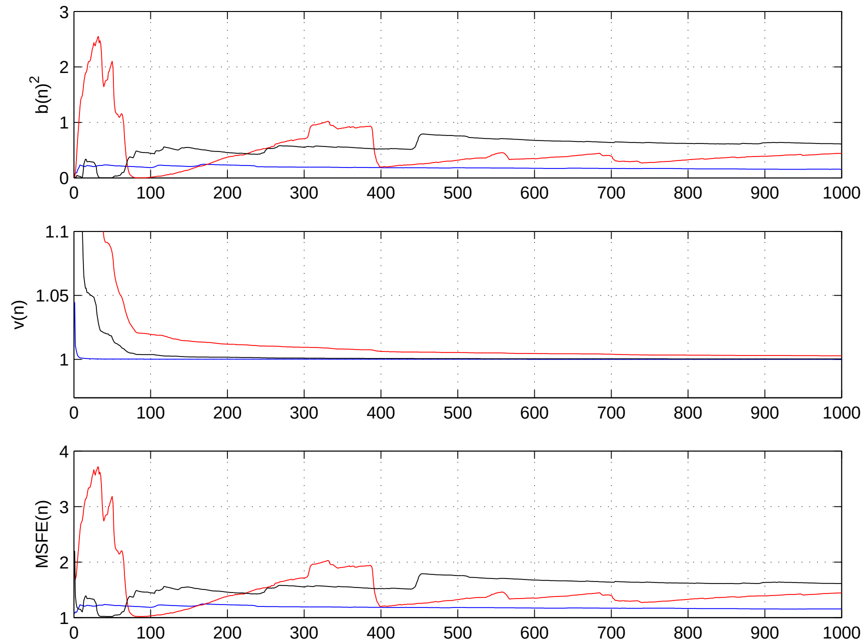


Figure 2.1: Conditional squared bias, conditional variance and CMSFE for three realizations of the process $\{X_t\}$

and the unconditional variance component

$$v_n = Var(\epsilon_{t+1,n}) = \sigma_U^2 + \mathrm{Var}(\psi(X_t)) + \mathrm{Var}(\Theta_{t,n}^\top X_t) + \mathrm{var}(\Lambda_{t,n}^\top X_t) \\ - 2\mathrm{Cov}(\psi(X_t), \Theta_{t,n}^\top X_t), \quad (2.4.17)$$

where $\mathrm{Var}(\Lambda_{t,n}^\top X_t) = \sigma_U^2 E[(X_t^\top Q_{t,n}^{-1} X_t)]$. As expected, neither component depends on a particular realization of the process $\{X_\tau\}$.

The unconditional variance component, (2.4.17), of the MSFE under misspecification contains noise from parameter estimation, $\sigma_U^2 + \sigma_U^2 E[(X_t^\top Q_{t,n}^{-1} X_t)]$, as well as variance terms which are associated with the misspecification of the model, $\mathrm{Var}(\psi(X_t)) + \mathrm{Var}(\Theta_{t,n}^\top X_t) - 2\mathrm{Cov}(\psi(X_t), \Theta_{t,n}^\top X_t)$. The presence of these latter terms makes the SSD of the unconditional bias and variance ambiguous. We note that the SSD of both the squared bias and variance components is manifested in the terms $\Theta_{t,n}^\top X_t$ and $X_t^\top Q_{t,n}^{-1} X_t$. To understand some aspects of the SSD, the following proposition characterizes the SSD of the term $E[X_t^\top Q_{t,n}^{-1} X_t]$ and the SSD of the unconditional variance in the case of a linear DGP.

Proposition 2.2

(i) $E[X_t^\top Q_{t,n}^{-1} X_t] > E[X_t^\top Q_{t,n+1}^{-1} X_t]$ *for any* n *and* t.

(ii) When the DGP is linear, $v_n \downarrow n$.

Proof. (i) First we show $X_t^\top Q_{t,n}^{-1} X_t > X_t^\top Q_{t,n+1}^{-1} X_t$ a.s-P. We write

$$\bar{X}_{t,n+1} = \left[X_{t-n-1} \bar{X}_{t,n}^\top \right] \in \mathbb{R}^{n+1 \times k},$$

and substitute in the expression for $Q_{t,n+1}$ so that

$$X_t^\top Q_{t,n+1}^{-1} X_t \quad = \quad = X_t^\top \left(\bar{X}_{t,n}^\top \bar{X}_{t,n} + X_{t-n-1} X_{t-n-1}^\top \right)^{-1} X_t, \text{ a.s-P} \qquad (2.4.18)$$

Using the following inverse formula

$$(A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1},$$

we can rewrite (2.4.18) as

$$X_t^\top Q_{t,n+1}^{-1} X_t = X_t^\top Q_{t,n}^{-1} X_t - X_t^\top Q_{t,n}^{-1} X_{t-n-1} (1 + X_{t-n-1}^\top Q_{t,n}^{-1} X_{t-n-1})^{-1} X_{t-n-1}^\top Q_{t,n}^{-1} X_t, \text{ a.s-P},$$

where we used $A_{11} = Q_{t,n}$, $A_{12} = X_{t-n-1}$, $A_{21} = X_{t-n-1}^\top$ and $A_{22} = -1$. It follows, since $Q_{t,n}$ is positive definite, $Q_{t,n}^{-1}$ is positive semidefinite so $X_{t-n-1}^\top Q_{t,n}^{-1} X_{t-n-1}$ is positive semidefinite and $1 + X_{t-n-1}^\top Q_{t,n}^{-1} X_{t-n-1}$ and its inverse are positive scalars. Finally we have

$$X_t^\top Q_{t,n}^{-1} X_{t-n-1} X_{t-n-1}^\top Q_{t,n}^{-1} X_t = (X_{t-n-1}^\top Q_{t,n}^{-1} X_t)^\top (X_{t-n-1}^\top Q_{t,n}^{-1} X_t) \geq 0, \text{ a.s-P}$$

and the result follows.

(ii) Substituting $\bar{F}_{t,n} = \bar{X}_{t,n} \beta$ in the expression for $\Theta_{t,n}$ one obtains $\Theta_{t,n} = \beta$. Substituting $\Theta_{t,n} = \beta$ and $f(X_t) = \beta^\top X_t$ in (2.4.17) one obtains $v_n = \sigma_U^2 + 2\text{Var}(\beta^\top X_t)) + \sigma_U^2 E[(X_t^\top Q_{t,n}^{-1} X_t)] - 2\text{Cov}(\beta^\top X_t, \beta^\top X_t) = \sigma_U^2 + \sigma_U^2 E[(X_t^\top Q_{t,n}^{-1} X_t)]$. The result follows from part (i). ■

The proposition implies the term $E[X_t^\top Q_{t,n}^{-1} X_t]$ of the variance decreases as more data is used to construct the forecast. Next, we look at the term $\Theta_{t,n}^\top X$. The term $\Theta_{t,n}^\top X_t$ can

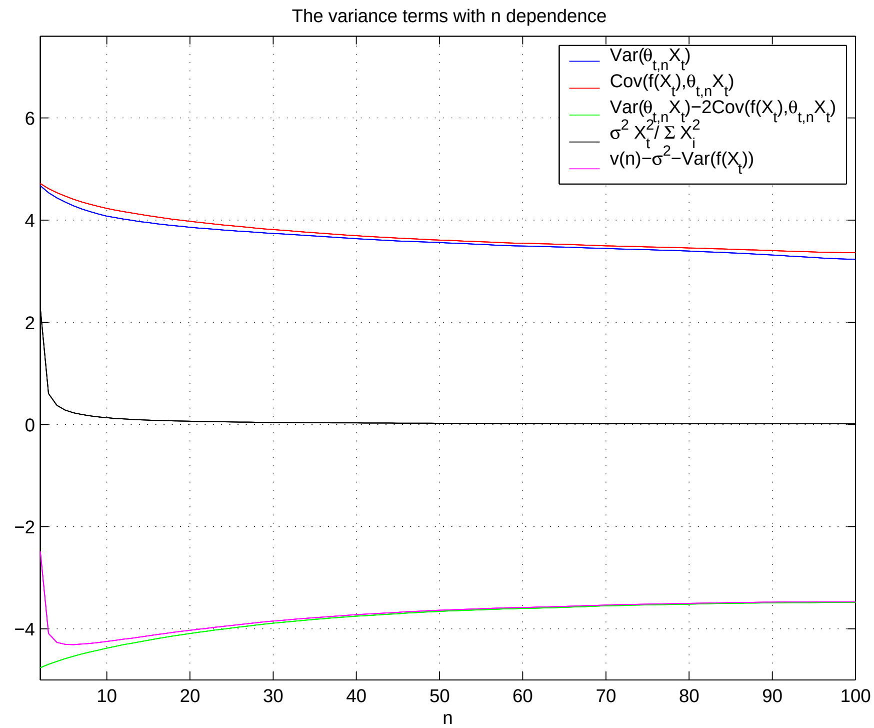


Figure 2.2: Terms of bias components of the MSFE for a misspecified linear model for a quadratic DGP

be found in both the bias and variance components in the expectation, $E[\Theta_{t,n}^{\top}X_t]$, the variance, $\mathrm{Var}(\Theta_{t,n}^{\top}X_t)$, and the covariance with $\psi(X_t)$, $\mathrm{Cov}(\psi(X_t), \Theta_{t,n}^{\top}X_t)$. The presence of the term $\Theta_{t,n}^{\top}X_t$ in both the squared bias component and the variance component makes it difficult to establish a trade-off with respect to the window size n. In fact, as the following example demonstrates, a trade-off between the unconditional bias and the unconditional variance is not warrantied to exist.

Example 2.2 *Consider the nonlinear DGP and process $\{X_\tau\}$ given in example* 2.1*. We investigate the SSD of the variance covariance terms $Var(\Theta_{t,n}^{\top}X_t)$, $Cov(\psi(X_t), \Theta_{t,n}^{\top}X_t)$, and $\sigma_U^2 E[(X_t^{\top}Q_{t,n}^{-1}X_t)]$ through a Monte Carlo experiment. For each value of window size n, $n = 1, \ldots, 100$, we compute the probability limits of different components of $Var(\Theta^{\top}X_t)$, as sample averages across* 10,000 *replications of the series $\{X_t\}$, $\{Y_t\}$, and $\{U_t\}$. The results of the experiment are shown in figure* (2.2)*.* $\square$

Furthermore, the following proposition shows that the only conclusions about optimal forecasting schemes which one can arrive at are for the simple correctly specified linear

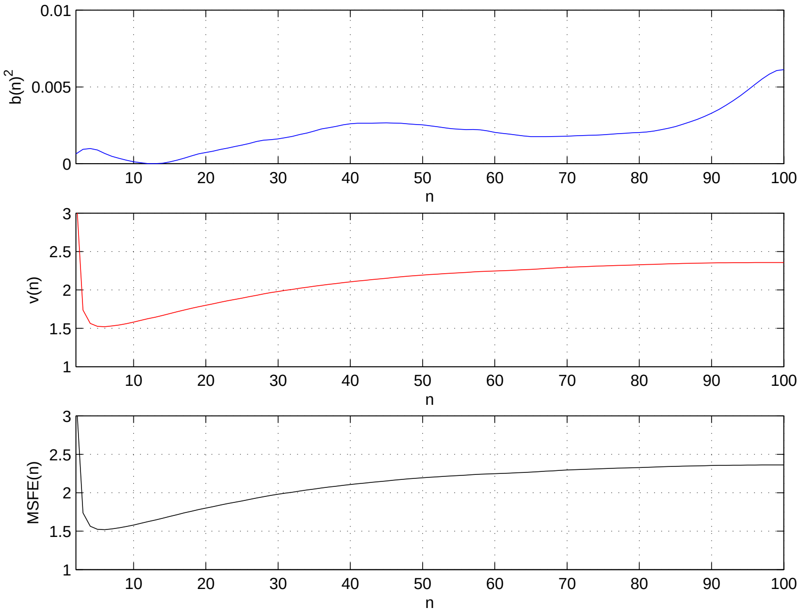


Figure 2.3: The three plots correspond to the unconditional bias, variance, and MSFE, respectively, for a misspecified linear model for a quadratic TDGP.

case.

Proposition 2.3 *If the DGP* (2.4.1) *is linear, the squared bias* b_n^2 *is zero and it is mean square optimal to use a recursive forecast scheme.*

Proof. Since $\psi(X_t) = \Theta_{t,n}^\top X_t$, the unconditional bias in (2.4.16) is zero and the unconditional variance in (2.4.17) reduces to $\sigma_U^2 + \sigma_U^2 E[(X_t^\top Q_{t,n}^{-1} X_t)]$. The result follows from proposition 2.2. ■

Concluding, we have seen there are two possible ways to define the MSFE, a conditional form and an unconditional form. We examined the squared bias and variance decomposition for both conditional and unconditional forms. For the conditional form, the dependence of the conditional squared bias on the particular realization $\mathcal{X}_t$ made the conditional MSFE unfit for analyzing optimal forecasting window schemes. Furthermore, for the unconditional bias and variance decomposition, the presence of variance covariance terms made the SSD ambiguous and its analysis infeasible. The work in the chapters to follow provide tools to assist in the analysis of the SSD of the MSFE.

2.5 Misspecification

Economic and econometric models are parsimonious mathematical devices used to approximate complex generating processes. As such, models fail to capture the complete dynamic relationships responsible for the observed behavior and misspecification becomes ubiquitous. One of the main goals of this thesis is to understand the ramifications of misspecification for the problem of forecasting. In particular, we are interested in the nature of the sample size dependence of the mean square forecast error under misspecified conditions. This subject is addressed in Chapters 5, 6, and 7.

Some common types of misspecification include the omission of relevant variables, inclusion of irrelevant variables, incorrect functional form, errors-in-variables, autocorrelation, heteroscedasticity, incompleteness of systems, and incorrect distributional assumptions. A formal and concrete treatise of misspecification can be conducted by the use of maximum likelihood techniques in the tradition of Cox [34, 35], Berk [14, 15], Huber [74], and White [152]. The following is based on [152].

Empirical phenomena is viewed as the realization of a stochastic processes as given in the following assumption.

Assumption 2.1 *The observed data are a realization of a stochastic process* $Z \equiv \{Z_\tau : \Omega \to \mathbb{R}^v, v \in \mathbb{N}, \tau = 1, 2, ...\}$ *on a complete probability space* $(\Omega, \mathcal{F}, P_0)$*, where* $\Omega = \mathbb{R}^{v\infty} \equiv \times_{\tau=1}^{\infty}\mathbb{R}^v$ *and* $\mathcal{F} = \mathcal{B}^{v\infty} \equiv \mathcal{B}(\mathbb{R}^{v\infty})$*.*

As an element ω of Ω ranges over Ω, the realization $Z_\tau(\omega)$ ranges over $\mathbb{R}^v$. For concreteness and convenience the choice $\Omega = \mathbb{R}^{v\infty}$ is made so that Z_τ is the projection operator that selects z_τ as the τth coordinate of ω, $Z_\tau(\omega) = z_\tau$. The $v \times 1$ observation vector Z_τ is often partitioned as $Z_\tau = (Y_\tau^\top, X_\tau^\top)^\top$, where Y_τ is $l \times 1$ and X_τ is $v - l \times 1$, where Y_τ is a set of dependent variables to be determined, explained or forecasted partly on the basis of other variables X_τ.

The probability measure P_0 provides a complete description of the stochastic behavior of the sequence Z and is viewed as the true data generating mechanism or data generating process. The problems of estimation and inference arise because P_0 is unknown. Given a realization of the of the sequence Z, knowledge of P_0 can be inferred from Z. Usually, one has available a realization z^n of a finite history, $Z^n \equiv (X_1^\top, \ldots, X_n^\top)^\top$, referred to as a sample of size n. The stochastic generating process of any sample of size n is completely described by its distribution $P_0^n(B) \equiv P_0[X^n \in B]$ for $B \in \mathcal{B}^{vn}$. The goal of estimation and inference is to learn about P_0^n from information contained in the sample generated by Z^n. A description of the stochastic nature of any sample equivalent to that provided by P_0^n is given by the Radon-Nikodýn density.

Theorem 2.4 (Theorem 2.1 in [152]) *Given assumption* 2.1 *and if* P_0^n *is absolutely continuous with respect to given* σ*-finite measures* v^n *on* $(\mathbb{R}^{vn}, \mathcal{B}^{vn})$*, there exists a measurable non-negative Radon-Nikodýn density* $g^n \equiv dP_0^n/dv^n$*, unique up to a set of* v^n*-measure zero, such that*

$$P_0^n(B) = \int_B g^n dv^n,$$

for all $B \in \mathcal{B}^{vn}$.

As long as v^n is properly chosen, the theorem warranties the existence of the relevant density function. Given v^n, knowledge of g^n is tantamount to knowledge of P_0^n. One can recover P_0^n by using the sample to learn about g^n. This can be done by constructing an approximation to g^n based on Z^n. A criterion to evaluate such an approximation was introduced by Kullback and Leibler [90].

Definition 2.5 (KLIC) *Let* $(\Omega, \mathcal{F}, v)$ *be a measure space, let* $g : \Omega \to \mathbb{R}^+$ *be a measurable function satisfying* $\int g dv < \infty$ *and* $\int_S g \log g dv < \infty$*, where* $S \equiv \{\omega \in \Omega : g(\omega) > 0\}$*, and let* $f : \Omega \to \mathbb{R}^+$ *be a measurable function satisfying* $\int_S g \log f dv < \infty$*. The Kullback-Leibler Information Criterion (KLIC) is defined as*

$$\mathbb{I}(g : f) \equiv \int_S g \log(g/f) dv.$$

The KLIC measures the discrepancy between g and f as described by the information inequality.

Theorem 2.6 (Information inequality, theorem 2.3 in [152]) *Let* f, g, v, S *and* $\mathbb{I}$ *be as in definition* 2.5*. If* $\int_S (g - f) dv \geq 0$*, then* $\mathbb{I}(g : f) \geq 0$ *and* $\mathbb{I}(g : f) = 0$ *if and only if* $g = f$ *almost everywhere* -v *on* S.

$\mathbb{I}(g : f)$ can serve as a measure of the closeness of f to g as discussed by Akaike [1]. Comparison of the adequacy of two approximations f_1 and f_2 by means of the KLIC is based on

$$\mathbb{I}(g : f_1) - \mathbb{I}(g : f_2) = \int_S \log(f_2/f_1) g dv,$$

where the latter quantity can be estimated without knowledge of g.

Approximations of g^n can be based on a probability model as defined below.

Definition 2.7 (Probability model) *Let $(\Omega, \mathcal{F})$ be a measurable space. A probability model is a collection $\mathcal{P}$ of distinct probability measures on $(\Omega, \mathcal{F})$.*

An element P of $\mathcal{P}$ is a model element.

Definition 2.8 (Correctly specified probability model) *The probability model $\mathcal{P}$ is correctly specified for Z if $\mathcal{P}$ contains P_0, the data generating process of assumption* 2.1*. Otherwise, $\mathcal{P}$ is misspecified for Z.*

In many cases, P_0 is assumed to belong to some probability model with elements indexed by a finite parameter vector, $\mathcal{P} = \{P_\theta : \theta \in \Theta \subseteq \mathbb{R}^p, p \in \mathbb{N}\}$. Such a model is referred to as a parametric probability model and written $\mathcal{P} = \{P_\theta\}$. A parametric probability model tends to be a small subset of $\mathcal{P}^*$, the collection of all probability measures on $(\Omega, \mathcal{F})$.

Theorem 2.9 (Theorem 2.6 in [152]) *Let $\mathcal{P} = \{P_\theta\}$ be a parametric probability model. Define P_θ^n as $P_\theta^n(B) \equiv P_\theta[Z^n \in B]$, $B \in \mathcal{B}^{vn}$, $n = 1, 2, \ldots$, $\theta \in \Theta$. Suppose there exists a σ-finite measure η^n on $(\mathbb{R}^{vn}, \mathcal{B}^{vn})$ such that for each θ in Θ, P_θ^n is absolutely continuous with respect to η^n, $n = 1, 2, \ldots$. Then there exists a nonnegative Radon-Nikodýn density $f^n(\cdot, \theta) = dP_\theta^n / d\eta^n$ measurable-$\mathcal{B}^{vn}$ for each θ in Θ, $n = 1, 2, \ldots$.*

The density $f^n(\cdot, \theta)$ is said to be constructed "from the top down" by first positing a parametric probability model $\mathcal{P}$ and then applying theorem 2.9. In economics, approximations to g^n are rarely constructed from the top down. The mapping $f^n(x^n, \cdot) : \Theta \to \mathbb{R}^+$ is referred to as the likelihood function generated by the probability model $\mathcal{P}$ with respect to η^n for the realization z^n, or simply the likelihood function generated by $\mathcal{P}$. An important representation of g^n for the construction of approximations is given in the next theorem.

Theorem 2.10 (Theorem 2.7 in [152]) *Given assumption* 2.1 *and given* P_0^n *is absolutely continuous with respect to given* σ*-finite measures* v^n *on* $(\mathbb{R}^{vn}, \mathcal{B}^{vn})$*, the densities* g^n*,* $n = 1, 2, \ldots$ *can be chosen such that* $z^n \in S^n \equiv \{z^n : g^n(z^n) > 0\}$ *implies* $z^{n-1} \in S^{n-1}$ *for all* z^n *in* S^n*. We refer to densities* g^n *with this property as standard. Then for all* z^n *in* S^n

$$\log\ g^n(z^n) = \sum_{\tau=1}^{n} \log\ g_\tau(x^\tau), \quad n = 1, 2, \ldots,$$

where $g_\tau(z^\tau) \equiv g^\tau(z^\tau)/g^{\tau-1}(z^{\tau-1})$*,* $\tau = 1, 2, \ldots,$ *and* $g_1(z^1) \equiv g^1(z^1) = g^1(x_1)$*.*

Often, g_τ can be interpreted as a conditional density of Z_τ given $Z^{\tau-1}$ with respect to a measure v_τ. An approximation to g^n can be constructed "from the bottom up" with functions $f_\tau : \mathbb{R}^{v\tau} \times \Theta \to \mathbb{R}^+$ as approximations to g_τ, $\tau = 1, 2, \ldots$ as follows:

$$f^n(z^n, \theta) \equiv \prod_{\tau=1}^{n} f_\tau(z^\tau, \theta).$$

This approximation is referred to as a quasi-likelihood function. A probability model $\mathcal{P}$ is constructed "from the bottom up" if the model is generated by a sequence of function $\{f^n = \prod_{\tau=1}^n f_\tau\}$ as defined below.

Definition 2.11 *Let* η^n *be a measure on* $(\mathbb{R}^{vn}, \mathcal{B}^{vn})$ *and let* $f^n : \mathbb{R}^{vn} \times \Theta \to \mathbb{R}^+$ *be measurable-*$\mathcal{B}^{vn}$ *for each* θ *in* Θ*, an arbitrary set,* $n = 1, 2, \ldots$*. For each* θ *in* Θ*, define the measure*

$$P_\theta^n(B) = \int_B f^n(z^n, \theta) d\eta^n(z^n), \quad B \in \mathcal{B}^{vn},$$

We say that $\{f^n\}$ *generates the probability model* $\mathcal{P} = \{P_\theta\}$ *with respect to* $\{\eta^n\}$ *if for each* θ *in* Θ *there exists a probability measure* P_θ *on* $(\mathbb{R}^{v\infty}, \mathcal{B}^{v\infty})$ *such that for each* n *the restriction of* P_θ *to* $(\mathbb{R}^{vn}, \mathcal{B}^{vn})$ *is given by* P_θ^n*,* $n = 1, 2, \ldots$*.*

To generate a probability model, it is necessary that f_τ be a conditional density for Z_τ, given $Z^{\tau-1}$ for all θ in Θ and all τ. This requirement is often violated in econometric practice, rendering probability models too narrow a class of approximations to P_0. For this reason, attention is focused to a wider class of approximations referred to as parametric stochastic specifications.

Definition 2.12 (Parametric stochastic specifications) *A parametric stochastic specifications on* $(\Omega, \mathcal{F})$ *is a collection* $\mathcal{S}$ *of sequences of functions* $f(\theta) \equiv \{f_\tau(\cdot, \theta) : \mathbb{R}^{v\tau} \to \mathbb{R}^+, \tau = 1, 2, \dots\}$ *obtained by letting* θ *range over* $\Theta \subseteq \mathbb{R}^p$, $p \in \mathbb{N}$ *where for each* $\tau = 1, 2, \dots$ *and each* $\theta \in \Theta$, $f_\tau(\cdot, \theta) : \mathbb{R}^{v\tau} \to \mathbb{R}^+$ *is measurable-*$\mathcal{B}^{v\tau}$, *i.e.* $\mathcal{S} \equiv \{f(\theta) : \theta \in \Theta\}$.

$\mathcal{S} = \{f_\tau\}$ is a specification for Z when the conditions of the definition are met and $f^n = \prod_{\tau=1}^n f_\tau$ is referred to as the quasi-likelihood specified by $\mathcal{S}$. Stochastic specifications may be correctly or incorrectly specified to varying degrees. For some applications, f_τ is allowed to depend on n, $\{f_{n\tau} : \mathbb{R}^{v\tau} \times \Theta \to \mathbb{R}^+, n, t = 1, 2, \dots\}$. The following assumption is useful in construction specifications.

Assumption 2.2 *The functions* $f_\tau : \mathbb{R}^{v\tau} \times \Theta \to \mathbb{R}^+$ *are such that* $f_\tau(\cdot, \theta)$ *is measurable-*$\mathcal{B}^{v\tau}$ *for each* θ *in* Θ, *a compact subset of* $\mathbb{R}^p$, $p \in \mathbb{N}$, *and* $f_\tau(Z^\tau, \cdot)$ *is continuous on* Θ *a.s.-*P_0, *i.e.,* $f_\tau(z^\tau, \cdot)$ *is continuous on* Θ *for all* z^τ *in some* $F_\tau \in \mathcal{B}^{v\tau}$, $P_0^\tau[F_\tau] = 1$, $\tau = 1, 2, \dots$.

Under assumption 2.2, the quasi-likelihood $f^n = \prod_{\tau=1}^n f_\tau$ can be viewed as an approximation to g^n as measured by the KLIC

$$\mathbb{I}(g^n : f^n; \theta) \equiv \int_{S^n} [\log g^n(z^n)/f^n(z^n, \theta)]\, g^n(z^n) dv^n(z^n).$$

Choosing θ to minimize $\mathbb{I}(g^n : f^n; \theta)$ is equivalent to choosing θ to maximize the following

$$\begin{aligned}
\tilde{L}_n(\theta) &= \int_{S^n} \log f^n(z^n, \theta) g^n(z^n) dv^n(z^n) \\
&= \int_{S^n} \log f^n(z^n, \theta) dP_0^n(z^n) \\
&= E[\log f^n(Z^n, \theta)].
\end{aligned}$$

When $f^n(z^n, \theta)$ is correctly specified, $f^n(z^n, \theta_0) = g^n(z^n)$ for a unique vector θ_0 in Θ so that choosing θ to maximize $\tilde{L}_n(\theta)$ yields θ_0 by the information inequality. In practice, θ cannot be chosen in this way since $\tilde{L}_n(\theta)$ is an expected value determined by the unknown

g^n. This can often be solved approximately using sample information. For this purpose, note that maximizing $\tilde{L}_n(\theta)$ is equivalent to maximizing

$$\bar{L}_n(\theta) \equiv n^{-1}\tilde{L}_n(\theta) = E[n^{-1}\log f^n(Z^n,\theta)].$$

Furthermore, it follows that

$$n^{-1}\log f^n(Z^n,\theta) = n^{-1}\sum_{\tau=1}^{n}\log f_\tau(Z^\tau,\theta).$$

If a law of large numbers applies to the sum, for n sufficiently large, $E[n^{-1}\log f^n(Z^n,\theta)]$ can be approximated by $L_n(Z^n,\theta) \equiv n^{-1}\log f^n(Z^n,\theta)$. Therefore, the value of θ which provides the best approximation to g^n can be approximated by the solution $\hat{\theta}_n$ to the problem

$$\max_{\theta\in\Theta} L_n(Z^n,\theta) \equiv n^{-1}\sum_{\tau=1}^{n}\log f_\tau(Z^\tau,\theta).$$

L_n is the quasi-log-likelihood function and $\hat{\theta}_n$ is the quasi-maximum likelihood estimator (QMLE). We give an existence theorem.

Theorem 2.13 (Theorem 2.12 in [152]) *Given assumptions* 2.1 *and* 2.2 *and a sequence* $\{\Theta_n\}$ *of compact subsets of* Θ, *for each* $n = 1, 2, \ldots$ *there exists a function* $\hat{\theta}_n : \mathbb{R}^{vn} \to \Theta_n$ *measurable-*$\mathcal{B}^{vn}$ *and a set* $F_n \in \mathcal{B}^{vn}$ *with* $P_0^n(F_n) = 1$ *such that for all* z^n *in* F_n

$$L_n(z^n,\hat{\theta}(z^n)) = \max_{\theta\in\Theta_n} L_n(z^n,\theta).$$

$\hat{\theta}_n$ is a random variable with stochastic properties such as consistency and an asymptotic distribution. White [149, 152] studies the consistency of the QMLE. The idea is that because $\hat{\theta}_n$ maximizes $L_n(Z^n,\theta)$ and $L_n(Z^n,\theta)$ tends to $\bar{L}_n(\theta) \equiv E[L_n(Z^n,\theta)]$, then $\hat{\theta}_n$ should tend to the value of θ, θ_n^*, which maximizes $\bar{L}_n$. Under assumptions 2.1 and 2.2 and assumptions on the continuity of $E[\log f_\tau(Z^\tau,\cdot)]$, $\{\log f_\tau(Z^\tau,\theta)\}$ obeying a law of large numbers and a uniqueness of the maximizers of $\{\bar{L}_n\}$, White proves $\hat{\theta}_n - \theta_n^* \to 0$ as $n \to \infty$ a.s.-P_0. We note $\bar{L}_n$ depends on the chosen parametric stochastic specification $\mathcal{S} = \{f_\tau\}$ and as such, θ_n^* does not necessarily coincide with the parameter θ_0 of the correctly specified parametric stochastic specification. White [152] and Domowits and White [151] give conditions for the asymptotic normality of the QMLE.

In chapters 5 and 6, we review some large sample results for the OLS under assumptions of misspecification and develop approximations to understand finite sample properties of the OLS and the MSFE under misspecification. For sake of brevity, we omit a description of the vast field of misspecification tests but direct the interested reader to the comprehensive monograph by Godfrey [55].

2.6 Motivating examples

The following examples serve as motivation for the work in chapters to follow by illustrating the sample size dependence (SSD) of the MSFE under different circumstances. The principal phenomena that we try to capture with these examples is the effect of model misspecification on the SSD of the MSFE.

In this first example, we investigate the SSD of the MSFE for the forecast of a DGP consisting of a linear regression with a correctly specified model.

Example 2.14 *We consider the forecast problem where the DGP is generated by a regression process of the form:*

$$Y_t = \phi X_{t-1} + U_t,$$

with $\{U_\tau\} \sim IIN(0, \sigma_U)$ *and* $\{X_\tau\} \sim IIN(\mu, \sigma_x)$*. The forecaster applies a correctly specified model of the form* $Y_t = \beta X_{t-1} + V_t$*, resulting in the forecast* $\hat{Y}_{t+1} = \hat{\beta} X_t$*. The OLS formed from the* n *most recent observations is given by the following:*

$$\hat{\beta}_{t,n} = \Big[\sum_{\tau=t-n}^{t-1} X_\tau^2\Big]^{-1} \sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau = \phi + \Big[\sum_{\tau=t-n}^{t-1} X_\tau^2\Big]^{-1} \sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau,$$

and its expectation is $E[\hat{\beta}_{t,n}] = \phi$*. The square of the OLS is as follows :*

$$\hat{\beta}_{t,n}^2 = \phi^2 + 2\phi\Big[\sum_{\tau=t-n}^{t-1} X_\tau^2\Big]^{-1} \sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau + \Big[\sum_{\tau=t-n}^{t-1} X_\tau^2\Big]^{-2} \Big[\sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau\Big]^2$$

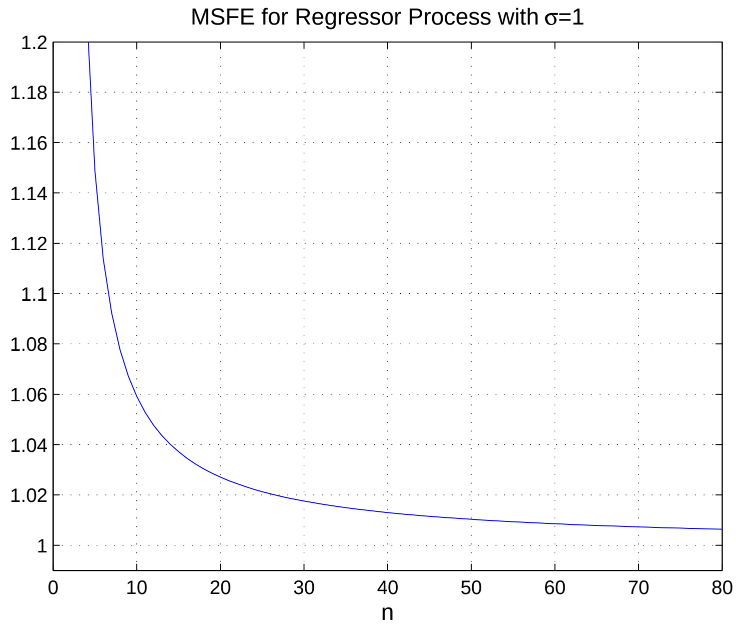


Figure 2.4: MSFE with $\sigma_U = 1$ and $\mu = 1$ and $\sigma_x = 1$

and its expectation is $E[\hat{\beta}^2_{t,n}] = \phi^2 + \sigma^2_U E[1/\sum_{\tau=t-n}^{t-1} X^2_\tau]$. *The MSFE is as follows:*

$$\begin{aligned} MSFE =& E[Y^2_{t+1}] - 2E[Y_{t+1}X_t]E[\hat{\beta}_{t,n}] + E[X^2_t]E[\hat{\beta}^2_{t,n}] \\ =& \sigma^2_U \left(1 + E[X^2_t]E\left[\Big(\sum_{\tau=t-n}^{t-1} X^2_\tau\Big)^{-1}\right]\right). \end{aligned}$$

To investigate the sample size dependence of the MSFE expression above, we conduct a Monte Carlo experiment. The conditional MSFE is given by

$$CMSFE = \sigma^2_U \left(1 + X^2_t \Big(\sum_{\tau=t-n}^{t-1} X^2_\tau\Big)^{-1}\right).$$

We produce one hundred thousand i.i.d realizations of the sequence $\{X_1, \ldots, X_{80}\}$ *with* $\{X_\tau\}_{\tau=1}^{80} \sim IIN(1,1)$ *and* $\sigma_U = 1$. *The MSFE from the Monte Carlo experiment is shown in Figure* 2.4. *As expected, the MSFE decreases monotonically with increasing sample size and it is optimal to use as much data as available.*

□

In the second example, we investigate the sample size dependence of the MSFE for the forecast of a DGP consisting of the same regression process as in example 2.14. In contrast to the previous case, we assume the forecaster uses a misspecified model.

Example 2.15 *We consider the forecast problem where the DGP is generated by a regression process of the form:*

$$Y_t = \phi X_{t-1} + U_t,$$

with $\{U_\tau\} \sim IIN(0, \sigma_U)$ *and* $\{X_\tau\} \sim IIN(\mu, \sigma_x)$. *The forecaster applies a misspecified white noise model of the form* $Y_t = \beta + V_t$, *resulting in the forecast* $\hat{Y}_{t+1} = \hat{\beta}$. *The OLS formed from the* n *most recent observations is given by the following:*

$$\hat{\beta}_{t,n} = \frac{1}{n} \sum_{\tau=t-n}^{t-1} Y_{\tau+1},$$

and its expectation is $E[\hat{\beta}_{t,n}] = E[Y_t]$. *The square of the OLS is as follows:*

$$\hat{\beta}^2_{t,n} = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}\Big)^2 = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y^2_{\tau+1} + \sum_{i\neq j,t-n}^{t-1} Y_{i+1}Y_{j+1}\Big),$$

and its expectation is

$$E[\hat{\beta}^2_{t,n}] = \frac{1}{n}E[Y_t^2] + \left(1 - \frac{1}{n}\right) E^2[Y_t].$$

The MSFE is given by

$$\begin{aligned} MSFE =& E[Y^2_{t+1}] - 2E[Y_{t+1}]E[\hat{\beta}_{t,n}] + E[\hat{\beta}^2_{t,n}] \\ =& Var(Y_t)\left(1 - \frac{1}{n}\right). \end{aligned}$$

The MSFE decreases monotonically with increasing sample size and it is optimal to use as much data as available.

□

Example 2.16 *We consider the forecast problem where the DGP is generated by an*

$AR(1)$ process of the form:

$$Y_t = \mu + \phi Y_{t-1} + U_t.$$

The forecaster applies a white noise model of the form $Y_t = \beta + V_t$, resulting in the forecast $\hat{Y}_{t+1} = \hat{\beta}_{t,n}$. The MSFE takes the following form:

$$MSFE = E[Y_{t+1}^2] - 2E[Y_{t+1}\hat{\beta}_{t,n}] + E[\hat{\beta}_{t,n}^2].$$

We are interested in the sample size dependence of the MSFE, which translates in part to the sample size dependence of the OLS. The OLS formed from the n most recent observations is given by the following:

$$\hat{\beta}_{t,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} Y_{\tau+1}.$$

The second term of the MSFE is as follows:

$$E[Y_{t+1}\hat{\beta}_{t,n}] = \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[Y_{t+1}Y_{\tau+1}] = E^2[Y_t] + \frac{1}{n}\sum_{i=1}^{n}\gamma_i,$$

where $\gamma_i = Cov(Y_t, Y_{t-i})$. Substituting the expression for the autocovariance of the process $\{Y_\tau\}$, $\gamma_i = \phi^i\sigma_U^2/(1-\phi^2)$, the expression for the variance of Y_t, $Var(Y_t) = \sigma_U^2/(1-\phi^2)$, and the summation

$$\sum_{i=1}^{n}\phi^i = \phi\frac{(1-\phi^n)}{(1-\phi)},$$

we obtain the following expression

$$E[Y_{t+1}\hat{\beta}_{t,n}] = E^2[Y_t] + \frac{1}{n}Var(Y_t)\,\phi\left(\frac{1-\phi^n}{1-\phi}\right). \tag{2.6.1}$$

The square of the OLS is as follows:

$$\hat{\beta}_{t,n}^2 = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}\Big)^2 = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 + \sum_{i\neq j,t-n}^{t-1} Y_{i+1}Y_{j+1}\Big),$$

and its expectation is

$$E[\hat{\beta}_{t,n}^2] = \frac{1}{n}E[Y_t^2] + \frac{1}{n^2}\sum_{i\neq j,t-n}^{t-1} E[Y_{i+1}Y_{j+1}].$$

For the second term we have

$$\begin{aligned}\sum_{i\neq j,t-n}^{t-1} E[Y_{i+1}Y_{j+1}] =& \sum_{i\neq j,t-n}^{t-1}\left[Cov(Y_{i+1},Y_{j+1}) + \frac{\mu^2}{(1-\phi)^2}\right]\\ =&2\sum_{i=1}^{n-1}(n-i)\gamma_i + (n^2-n)\frac{\mu^2}{(1-\phi)^2}.\end{aligned}$$

Substituting the expression for the autocovariance $\gamma_i = \phi^i\sigma_U^2/(1-\phi^2)$, *the expression for the variance* $Var(Y_t) = \sigma_U^2/(1-\phi^2)$, *and the summations*

$$\sum_{i=1}^{n-1}\phi^i = \phi\frac{(1-\phi^{n-1})}{(1-\phi)}, \quad \sum_{i=1}^{n-1} i\phi^i = \frac{\phi - n\phi^n - \phi^{n+1} + n\phi^{n+1}}{(\phi-1)^2},$$

we obtain

$$E[\hat{\beta}_{t,n}^2] = E^2[Y_t] + \frac{1}{n}Var(Y_t)\left(1+\frac{2\phi}{1-\phi}\right) - \frac{2}{n^2}Var(Y_t)\,\phi\left(\frac{1-\phi^n}{(1-\phi)^2}\right). \tag{2.6.2}$$

Substituting expressions (2.6.1) *and* (2.6.2) *in the MSFE, we obtain the following expression for the MSFE:*

$$MSFE = Var(Y_t)\left[1 + \left(1+\frac{2\phi^{n+1}}{1-\phi}\right)\frac{1}{n} - 2\phi\left(\frac{1-\phi^n}{(1-\phi)^2}\right)\frac{1}{n^2}\right].$$

Figures 2.5 *through* 2.9 *show the MSFE for the case with* $\sigma_U = 1$ *and different values of* ϕ.

□

Example 2.17 *We consider the forecast problem where the DGP is generated by an*

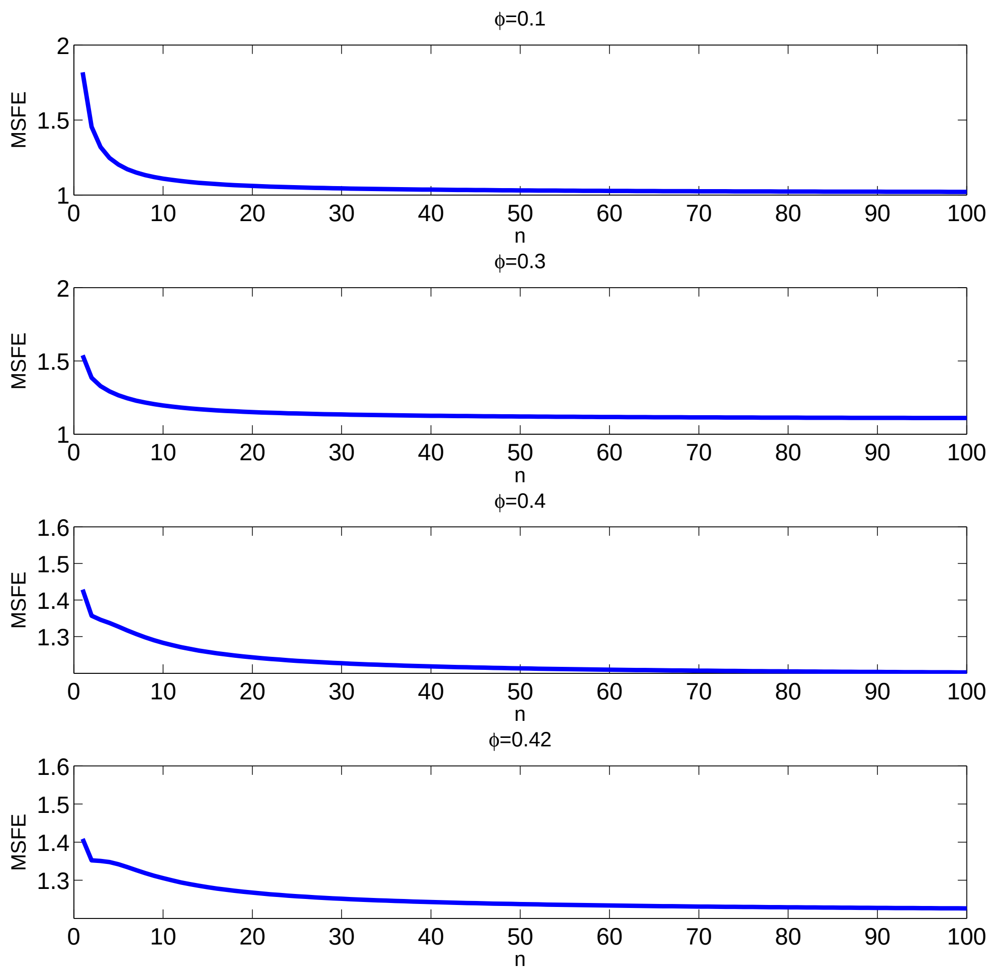


Figure 2.5: MSFE for a constant forecast model which misspecifies an AR(1) DGP with $\sigma_U = 1$

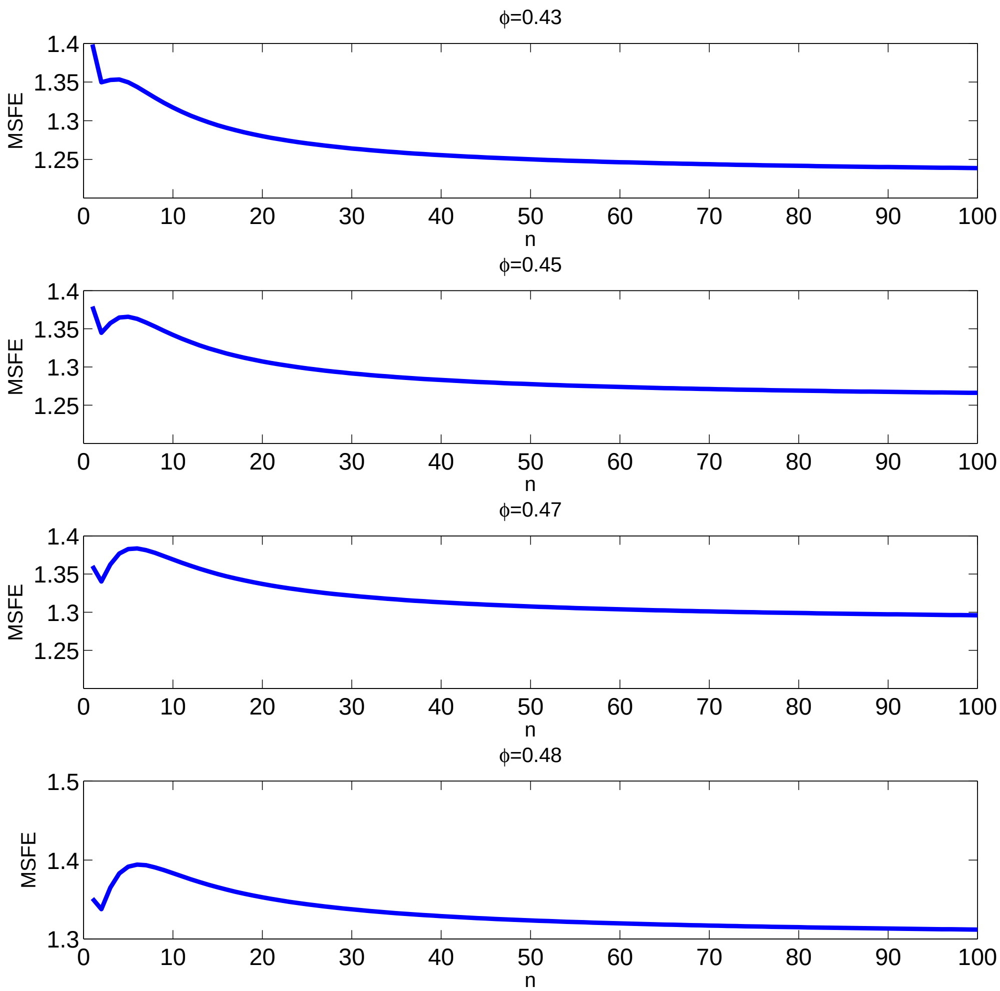


Figure 2.6: MSFE for a constant forecast model which misspecifies an AR(1) DGP with $\sigma_U = 1$

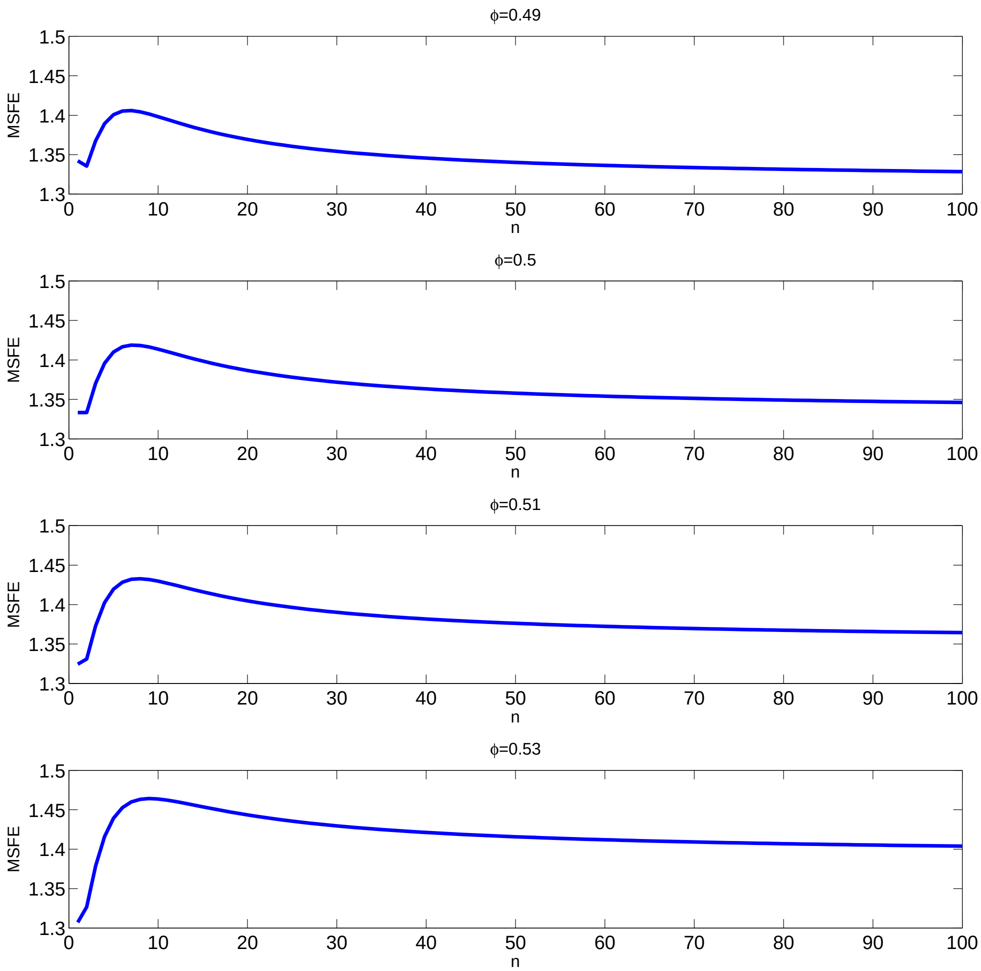


Figure 2.7: MSFE for a constant forecast model which misspecifies an AR(1) DGP with $\sigma_U = 1$

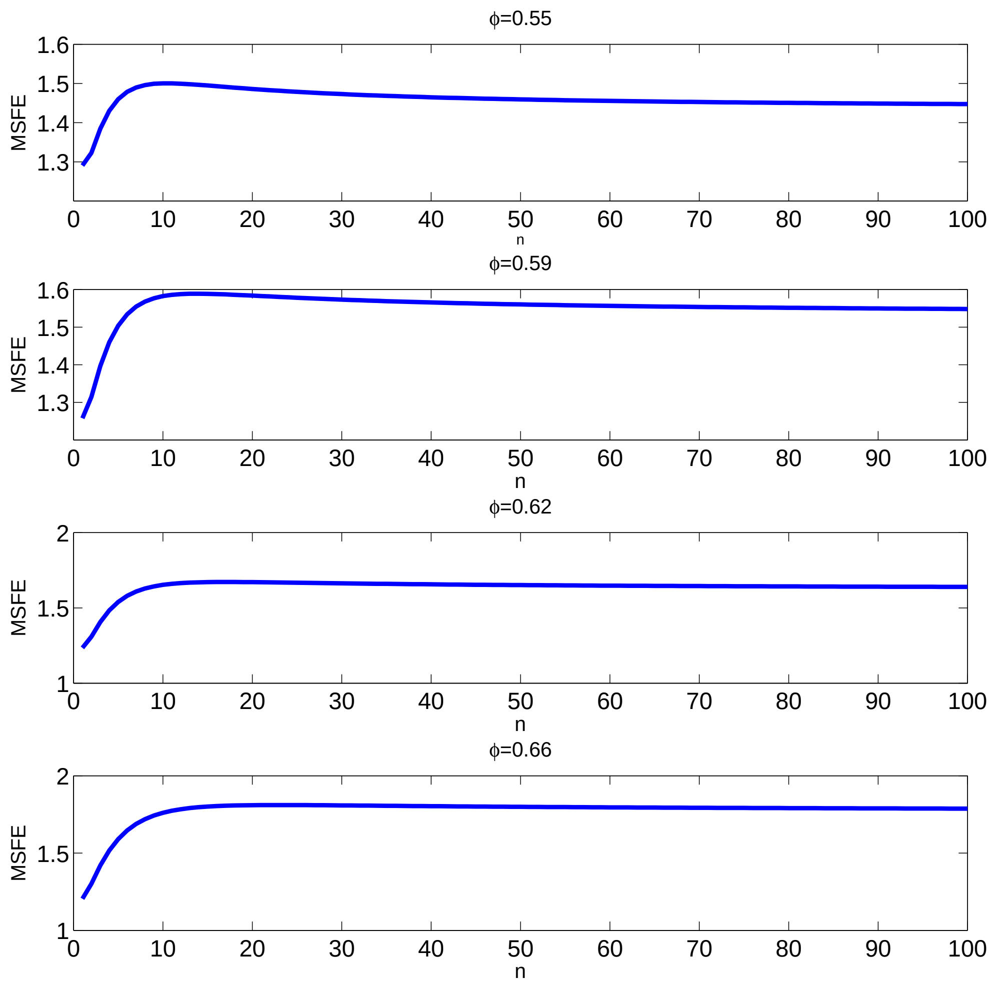


Figure 2.8: MSFE for a constant forecast model which misspecifies an AR(1) DGP with $\sigma_U = 1$

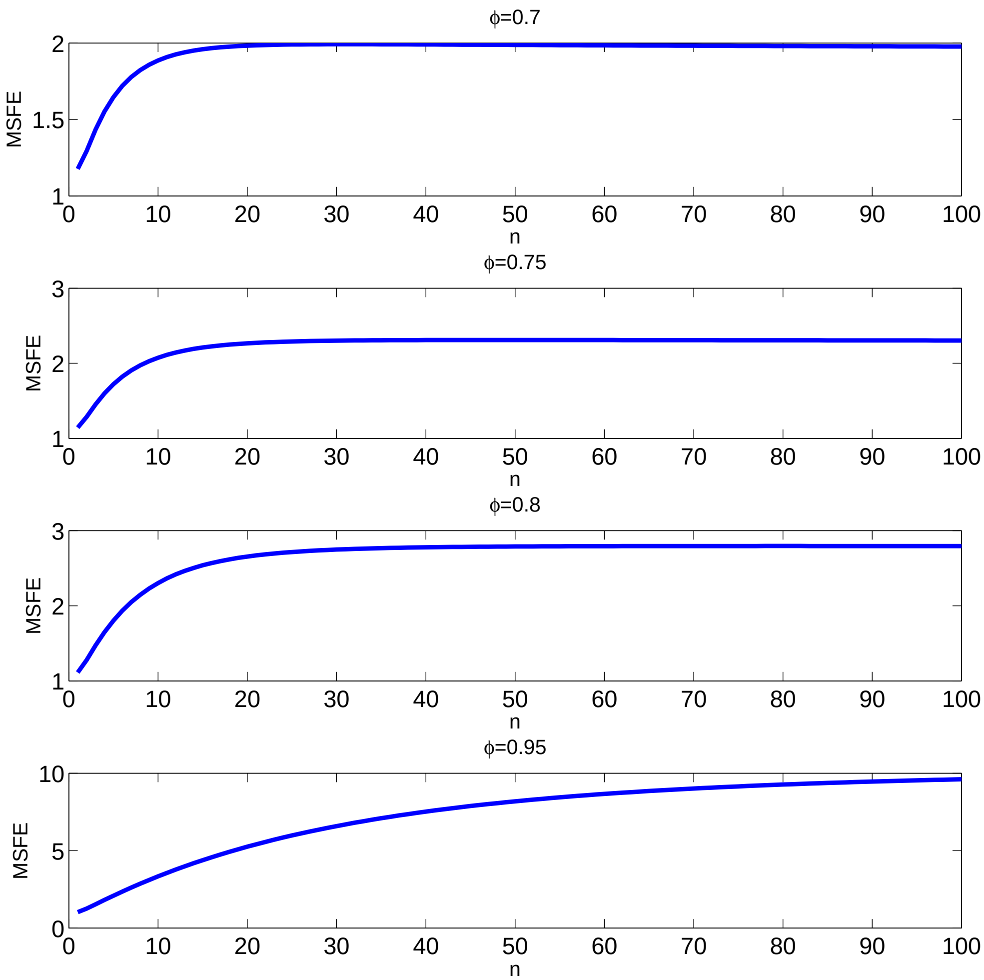


Figure 2.9: MSFE for a constant forecast model which misspecifies an AR(1) DGP with $\sigma_U = 1$

ϕ	Figure			
	A	B	C	D
ϕ_1	0.08	0.3	0.4	0.6
ϕ_2	0.08	0.1	0.3	0.2
ϕ_3	0.08	0.05	0.2	0.01
ϕ_4	0.08	0.05	0.01	0.01
ϕ_5	0.08	0.05	0.01	0.01
ϕ_6	0.08	0.05	0.01	0.01
ϕ_7	0.08	0.05	0.01	0.01
ϕ_8	0.08	0.05	0.01	0.01
ϕ_9	0.08	0.05	0.01	0.01
ϕ_{10}	0.08	0.05	0.01	0.01
ϕ_{11}	0.08	0.05	0.01	0.01
ϕ_{12}	0.08	0.05	0.01	0.01

Table 2.1: Autoregressive parameters

$AR(12)$ process of the form:

$$\begin{aligned} Y_t =& \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \phi_5 Y_{t-5} + \phi_6 Y_{t-6} \\ &+ \phi_7 Y_{t-7} + \phi_8 Y_{t-8} + \phi_9 Y_{t-9} + \phi_{10} Y_{t-10} + \phi_{11} Y_{t-11} + \phi_{12} Y_{t-12} + U_t, \end{aligned}$$

where U_t is zero mean white noise. The forecaster applies an $AR(1)$ model of the form $Y_t = \beta Y_{t-1} + V_t$, resulting in the forecast $\hat{Y}_{t+1} = \hat{\beta}_{t,n} Y_t$. The table below provides the parameter values for the plots shown in Figure 2.10. The MSFEs are generated by means of Monte Carlo simulations.

□

Example 2.18 *We consider the forecast problem where the DGP is generated by a deterministic trend process of the form:*

$$Y_t = \mu + \delta t + U_t.$$

The forecaster applies a white noise model of the form $Y_t = \beta + V_t$, resulting in the forecast $\hat{Y}_{t+1} = \hat{\beta}_{t,n}$. The MSFE takes the following form

$$MSFE = E[Y_{t+1}^2] - 2E[Y_{t+1}\hat{\beta}_{t,n}] + E[\hat{\beta}_{t,n}^2].$$

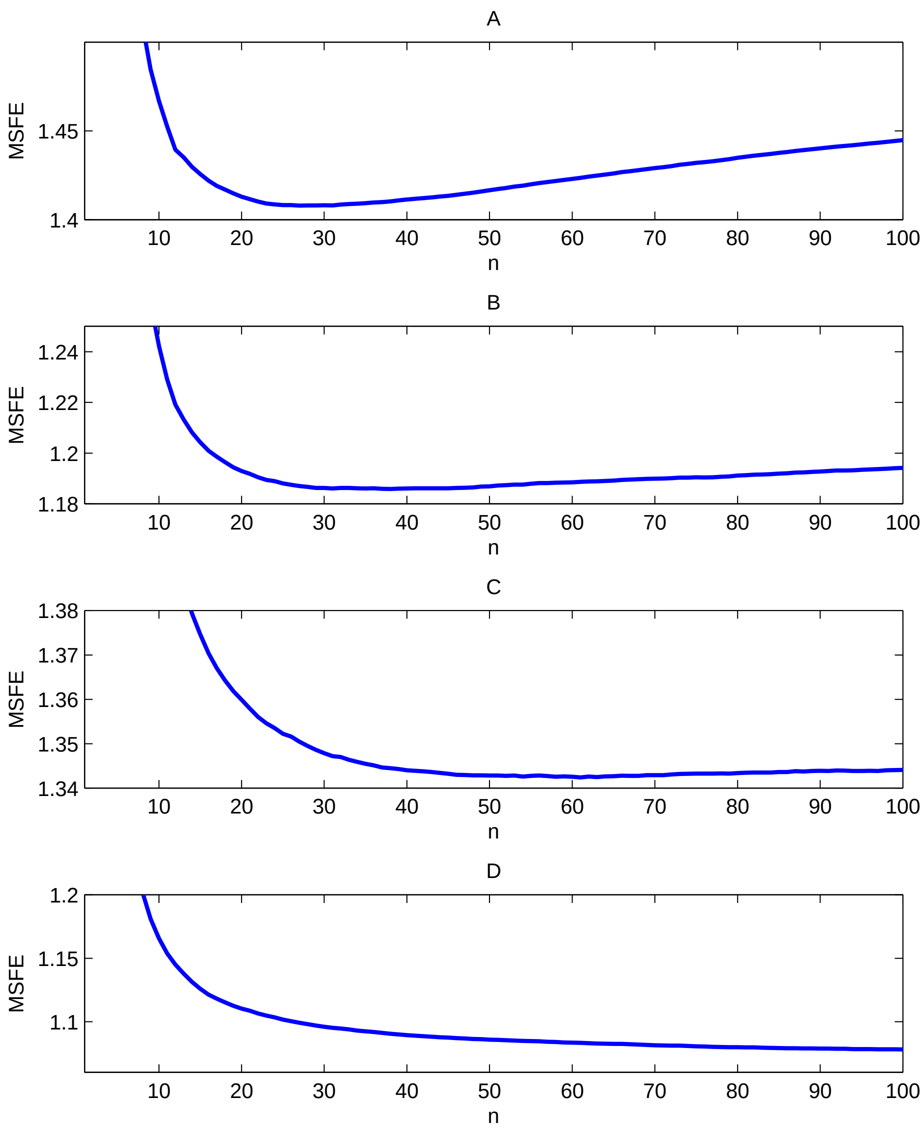


Figure 2.10: MSFE for an AR(1) forecast model which misspecifies an AR(12) DGP with $\sigma_U = 1$

The OLS formed from the n most recent observations is given by the following

$$\hat{\beta}_{t,n} = \frac{1}{n} \sum_{\tau=t-n}^{t-1} Y_{\tau+1}.$$

The first term of the MSFE is as follows:

$$E[Y_{t+1}^2] = (\mu + \delta(t+1))^2 + \sigma_U^2.$$

The second term of the MSFE is as follows:

$$E[Y_{t+1}\hat{\beta}_{t,n}] = (\mu + \delta(t+1))(\mu + \frac{\delta}{2} + \delta t - \frac{\delta n}{2}).$$

The square of the OLS is as follows:

$$\hat{\beta}_{t,n}^2 = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}\Big)^2 = \frac{1}{n^2}\Big(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 + \sum_{i\neq j,t-n}^{t-1} Y_{i+1}Y_{j+1}\Big),$$

and its expectation is

$$E[\hat{\beta}_{t,n}^2] = \frac{1}{4}(2\mu + \delta + 2\delta t)^2 + \frac{\sigma_U^2}{n} - \frac{\delta}{2}(2\mu + \delta + 2\delta t)n + \frac{\delta^2}{4}n^2.$$

Combining terms, the MSFE has the following form

$$MSFE = \frac{\delta^2}{4} + \sigma_U^2 + \frac{\sigma_U^2}{n} + \frac{\delta^2}{2}n + \frac{\delta^2}{4}n^2.$$

□

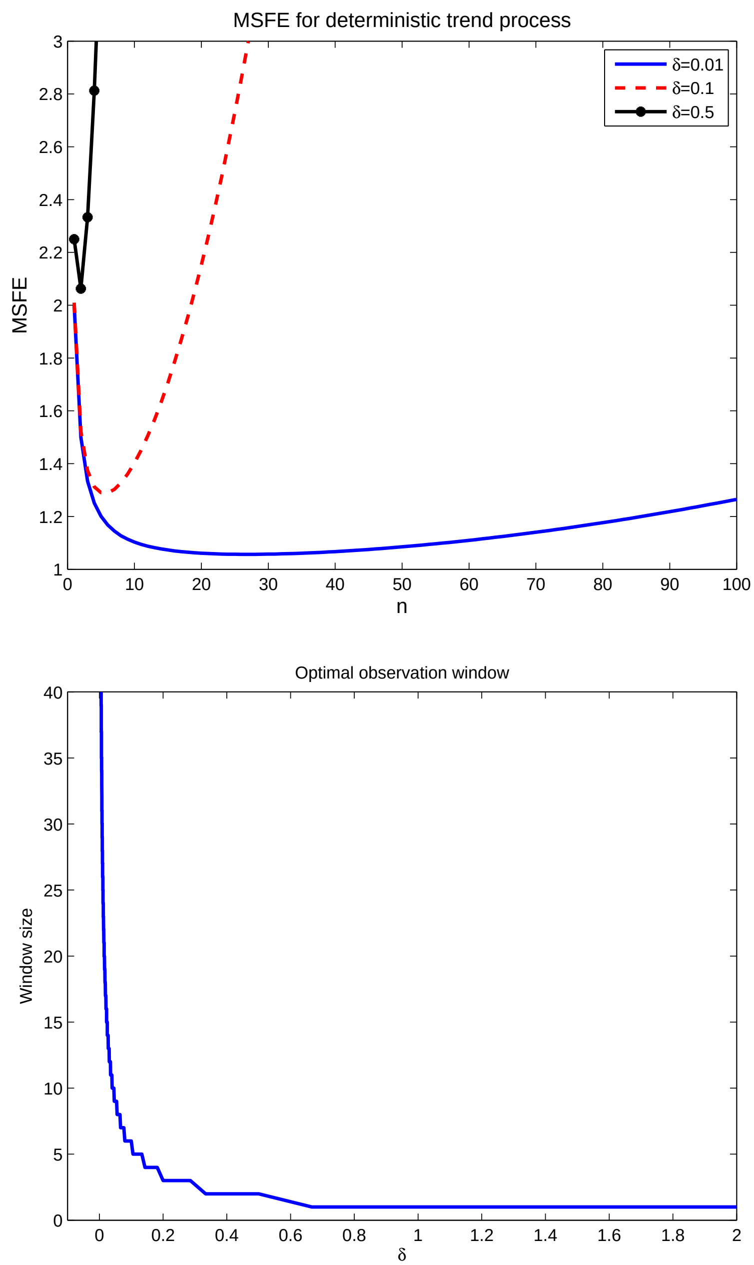


Figure 2.11: MSFE with $\sigma_U = 1$ for the deterministic trend example

2.7 Intuition behind our approach

In Section 2.4, we construct the forecasting problem based on a linear regression model. The explanatory variables can consist of casual variables, as well as time lags of the dependent variable. In the latter case, the resulting formulation is a time series model. Restricting the problem by selecting a linear model is in line with common practice in the forecasting literature, which favors in most situations simple models over correct specification ([2], p. 306). For estimation, ordinary least squares (OLS) is the estimator of choice. The evaluation of forecasts is to be carried out by means of the MSFE. The methods chosen for estimation and evaluation allow for the most general framework possible in the sense of the processes being analyzed.

The primary aim of this thesis is to understand how the accuracy of a forecast might depend on the amount of data used in the estimation of the model. By amount of data, we refer to the temporal element of the series. Should we use the last month, quarter, year, or decade of a particular time series in formulating a forecast? In some of the literature, this is referred to as selecting an observation window. We do not try to address the question of determining casual dependencies of different cross-sectional data.

To better understand how one might go about determining such an observation window, we recall the eight step strategy outlined by Allen and Fildes [2] to construct econometric forecasts. For the first step, the objective is forecasting. For determining the set of variables, we assume the relevant casual relationships have been established and the list of variables to be used in the problem are given. We further assume the forecaster has access to the longest available series for each of the variables and has some relative knowledge of events such as past structural breaks. The specification of the model as a linear regression has been established as well as the use of OLS for estimation. It is at the stage of estimation that the issue of an observation window can first be raised. This is made particularly simple by the use of the OLS, which depends on an estimation sample consisting of the last n observations, as can be seen in (2.4.4). The question to ask is: Can we determine an optimal observation window at the stage of estimation alone? The answer is no. To understand why this is so, consider what the estimation problem entails. For the case of least squares, the estimation problem is given by the optimization problem:

$$\hat{\beta}_{t,n} = \text{argmin}_{\beta \in B} \sum_{\tau=t-n}^{t-1} (Y_{\tau+1}(\theta) - \beta X_{\tau})^2. \tag{2.7.1}$$

The aim of the estimation problem (2.7.1) is to choose a $\hat{\beta}_{t,n}$ which, on average, replicates the process as close as possible with the linear model, i.e., the aim is optimal fit and explanation. Since the objective of the forecast problem is not explanation, but rather prediction, one must question the appropriateness of choosing an observation window at this stage. Consider for example, a process with unstable parameters which has undergone a structural break in the past and is modeled with a correct functional form. If the observation window is determined at the estimation stage, the answer would be, in most situations, to use all post-break data and to ignore all pre-break data. This would assure that the model fits the post-break process as close as possible. Nonetheless, it is well known that, in many situations, such as having a short post-break data history, optimal forecasts make use of pre-break data. In the case of the MSFE, this is due to the bias-variance trade-off. Consequently, the task of evaluating the temporal significance of data for the purpose of forecasting must be carried beyond the estimation stage.

The existing methods used to discriminate data based on a temporal criteria include: using an expanding window; using a fixed-size window; and using exponential declining weights. These methods are ad hoc and are always applied at the estimation stage, making them sub-optimal for the purpose of forecasting. A major contribution of this thesis is the reformulation of the standard forecasting strategies to allow for evaluation of the temporal significance of data in a setting more appropriate than the estimation problem. This reformulation of forecasting strategies is essential to make the temporal evaluation of data a systematic procedure which relies on the dynamic nature of the observed processes.

The standard way of solving the estimation problem (2.7.1) assumes the use of series $\{X_{t-n}, \ldots, X_{t-1}\}$ and $\{Y_{t-n+1}, \ldots, Y_t\}$ of a predetermined length n and possibly with predetermined weights. The selection of this length n and weights, as noted earlier, is done in an ad hoc manner by qualitative means which are very loosely based on some theoretic aspects of the processes being observed. For example, a forecaster might be aware the economic phenomenon of interest undergoes small but frequent structural shifts, and she might choose to use a fixed size observation window of length equal to the average length of the periods between shifts.

Once the estimation problem (2.7.1) is solved, $\hat{\beta}_{t,n}$ is a fixed quantity leading to the forecast $\hat{Y}_{t+1} = \hat{\beta}_{t,n} X_t$. This estimation problem does not evaluate data directly in terms of its temporal significance, and at no point during the eight steps of the forecasting strategy is the accuracy of the forecast tested for sensitivity to the length of the data set.

We intend to make the selection of an observation window systematic and quantitative. Instead of blindly predetermining the length of the series in the problem, we propose a reformulation of the estimation problem which treats the length n as a variable to be determined simultaneously with the estimator $\hat{\beta}_{t,n}$. The criteria for determining the length of the series is maximizing accuracy of the forecast as a function of n. In the case of forecast evaluation with the MSFE, this criteria translates to minimizing the MSFE as a function of n. For evaluation by means of the MSFE, the reformulated estimation problem is as follows:

$$n^* = \operatorname{argmin}_{n \in \mathbb{N}} E[(Y_{t+1} - \hat{\beta}_{t,n} X_t)^2], \tag{2.7.2}$$

$$\hat{\beta}_{t,n} = \operatorname{argmin}_{\beta \in B} \sum_{\tau = t-n}^{t-1} (Y_{\tau+1}(\theta) - \beta X_\tau)^2. \tag{2.7.3}$$

These ideas can be developed in a more general setting. Consider the forecast problem of predicting the variable Y_{t+1} where the DGP and forecast model are as follows:

$$\text{DGP}: \ Y_t = g(W_t, \theta), \quad \text{Model}: \ Y_t = f(X_t, \beta). \tag{2.7.4}$$

We assume $W_t \in \mathbb{R}^m$ and $X_t \in \mathbb{R}^k$ are t-measurable vectors of random variables and $\theta \in \mathbb{R}^p$ and $\beta \in \mathbb{R}^q$ are parameter vectors. The vectors W_t and X_t can contain any set of causal cross-sectional variables, as well as time lags of the dependent variable Y_t. This setup allows for modeling of any degree of misspecification. Up to this point, the problem does not differ from standard strategies. We solve the forecast problem by means of the following forecast equation and error:

$$\text{Forecast equation}: \ \hat{Y}_{t+1,n}(K_n) = f(X_t, \hat{\beta}_{t,n}(K_n)),$$
$$\text{Error}: \ \epsilon_{t+1,n}(K_n) = Y_{t+1} - \hat{Y}_{t+1,n}(K_n).$$

$\hat{\beta}_{t,n}$ is the estimator of β, and K_n is a real valued function $K_n : \mathbb{R}^n \to [0,1]$, which plays the role of a kernel assigning weights with values in the interval $[0,1]$ to each of the datum used in forming $\hat{\beta}_{t,n}$. K_n has as domain $\mathbb{R}^n$ because each particular weight in $[0,1]$ assigned to a variable must be determined based on the information contained by all of the explanatory variables $\{X_{t-n}, \ldots, X_{t-1}\}$. In particular, we will demonstrate autocovariances among the data play an important role in determining the kernel K_n. It is important to note the kernel K_n is a time or temporal kernel, as opposed to the typical spatial kernels used in nonparametric econometrics. Spatial kernels weight data according to the distance of the value of a particular datum to a mean. We make explicit the dependence of the forecast $\hat{Y}$ and the error ϵ on the kernel K_n to emphasize how our strategy differs from contemporary forecasting strategies which do not analyze the temporal dependence of a forecast.

Under the unrealistic assumption of correct specification, the DGP and model coincide and are given by $Y_t = g(W_t, \theta)$. The forecast equation becomes $\hat{Y}_{t+1,n}(K_n) = g(W_t, \hat{\theta}_{t,n}(K_n))$. Under reasonable assumptions, an unbiased, $E[\hat{\theta}_{t,n}] = \theta$, and consistent, $\hat{\theta}_{t,n} \xrightarrow{P} \theta$, estimator $\hat{\theta}_{t,n}$ can be obtained. The forecast evaluation under these conditions should lead us to the choice of the trivial kernel $K_n = 1$. The reason being that using the trivial kernel, one obtains the following highly desirable relations:

$$\hat{Y}_{t+1,n} \xrightarrow{P} Y_{t+1}, \quad \epsilon_{t+1,n} \xrightarrow{P} 0.$$

Under misspecification, the DGP and model would be given by (2.7.4). The forecast equation becomes $\hat{Y}_{t+1,n}(K_n) = f(X_t, \hat{\beta}_{t,n}(K_n))$. Estimators for this problem will be biased and the kernel is determined, for a chosen estimator $\hat{\beta}_{t,n}$ and a chosen cost function

$\mathcal{L}$, by the following optimization problem:

$$\min_{n} \mathcal{L}(\epsilon_{t+1,n}(K_n)).$$

For the work presented in this thesis, we focus attention to step kernels of the form

$$K_n = 1_{I(n)} = \begin{cases} 1 & \text{if} \quad I(n) \quad \text{is true} \\ 0 & \text{otherwise} \end{cases}.$$

The kernel is the indicator function which is one if condition $I(n)$ is satisfied, and zero otherwise. Returning to the case of estimation with OLS and evaluation with the MSFE, the estimation problem described in (2.7.2) and (2.7.3) can be written in terms of a temporal step kernel as follows:

$$n^* = \text{argmin}_{n \in \mathbb{N}} E[(Y_{t+1} - \hat{\beta}_{t,n}(K_n)X_t)^2],$$
$$\hat{\beta}_{t,n}(K_n) = \text{argmin}_{\beta \in B} \sum_{\tau=t-n}^{t-1} (K_{t-\tau} Y_{\tau+1}(\theta) - \beta X_\tau)^2,$$
$$K_n = \begin{cases} 1, & n < n^* \\ 0 & n > n^* \end{cases}.$$

This system of relations cannot be solved explicitly for n^*. Instead, one can apply a search method for the optimal window size n^* by calculating the MSFE, $E[(Y_{t+1} - \hat{\beta}_{t,n}(K_n)X_t)^2]$, for different values of n starting with $n = 1$. This procedure would reveal the sample size dependence (SSD) of the MSFE. The difficulty in applying the search method as suggested lies in that the squared forecast error is a non-trivial function of the explanatory variables. This function can not be simplified with the usual properties of the expectation in order to obtain a functional form depending explicitly on the value of n. To tackle the problem of discerning the SSD of the MSFE, we propose an approximation method. This method has as a main goal to approximate functionally complex statistics such as the squared forecast error by simple statistics with tractable expectations.

In the work of this thesis, we consider a Taylor polynomial of order m, P_m, to approximate the squared forecast error,

$$\epsilon_{t+1,n}^2 = (Y_{t+1} - \hat{\beta}_{t,n}(K_n)X_t)^2 \approx P_m(X_{t-n}, \ldots, X_{t-1}, Y_{t-n+1}, \ldots, Y_t).$$

A main contribution of this thesis is to provide an extensive exposition on the use of Taylor polynomials to approximate statistics and apply those approximations in the context of forecasting. Of particular interest is that, in general, the resulting approximation can be written as a linear combination of moments and real autocovariances which can easily be approximated, making the method suitable for empirical applications. To carry out approximations of statistics with Taylor polynomials, attention must be given to the fact that there has to be some agreement between the radius of convergence and the range of the random variables involved in the statistic. Chapter 3 and Chapter 4 address this and related questions.

Chapter 3

Expectations and truncated expectations

3.1 Introduction

This chapter presents basic standard notation of probability, random variables, and expectations. We develop the concept of truncated expectation and describe properties based on the standard notation of expectations. Truncated expectations are crucial to the development of the forecasting algorithms based on Taylor approximations which are presented in chapters to follow.

3.2 Expectations

Let $(\Omega, \mathcal{F}, P)$ be a probability measure space and $(\mathbb{R}, \mathcal{B})$ a measurable space. A random variable X is an $\mathcal{F}/\mathcal{B}$ measurable function $X : \Omega \to \mathbb{R}$. That is, $X(\omega)$ induces an inverse mapping from $\mathcal{B}$ to $\mathcal{F}$ such that $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{B}$, where $\mathcal{B}$ is the linear Borel field. The symbol μ will denote a probability measure on the real line, while P is used for the probability measure on the underlying space Ω. The following theorem relates P and μ.

Theorem 3.1 (Theorem 3.1.3 in [31]) *Each random variable on the probability space $(\Omega, \mathcal{F}, P)$ induces a probability space $(\mathbb{R}, \mathcal{B}, \mu)$ by means of the following correspondence:*

$$\mu(B) = PX^{-1}(B) = P(X^{-1}(B)) = P(\omega : X(\omega) \in B), \quad \forall B \in \mathcal{B}.$$

□

The measure μ, induced by X, is called the probability distribution or law, and has an associated distribution function F_X given by

$$F_X(x) = \mu((-\infty, x]) = P(\omega : X(\omega) \leq x).$$

If X is a r.v. on $(\Omega, \mathcal{F}, P)$ which induces the space $(\mathbb{R}, \mathcal{B}, \mu)$ and $g : \mathbb{R} \to \mathbb{R}$ is a Borel function, then $g \circ X(\omega) = g(X(\omega))$ is a random variable on the probability space $(\mathbb{R}, \mathcal{B}, \mu g^{-1})$. The distribution of $g(X)$ is μg^{-1} with

$$\mu g^{-1}(A) = \mu(g^{-1}A) = P(\omega : g(X(\omega)) \in A) = P(\omega : X(\omega) \in g^{-1}A).$$

We now define the integral of a measurable function and present some properties of integrals which are essential to define the expectation of functions of random variables. Let ϕ denote a real measurable function on the probability space $(\Omega, \mathcal{F}, P)$. If ϕ is nonnegative, the integral of ϕ with respect to the measure P is defined as follows:

$$\int_\Omega \phi(\omega) dP(\omega) = \sup \sum_i \left[\inf_{\omega \in \Lambda_i} \phi(\omega) \right] P(\Lambda_i),$$

where the supremum extends over all finite decompositions $\{\Lambda_i\}$ of Ω into $\mathcal{F}$-sets. For a general function ϕ, define its positive part, ϕ^+, and negative part, ϕ^- as follows

$$\phi^+(\omega) = \begin{cases} \phi(\omega), & 0 \leq \phi(\omega) \leq \infty \\ 0, & -\infty \leq \phi(\omega) \leq 0 \end{cases},$$

$$\phi^-(\omega) = \begin{cases} -\phi(\omega), & -\infty \leq \phi(\omega) \leq 0 \\ 0, & 0 \leq \phi(\omega) \leq \infty \end{cases},$$

so that $\phi = \phi^+ - \phi^-$. The general integral is defined by

$$\int_\Omega \phi(\omega) dP(\omega) = \int_\Omega \phi^+(\omega) dP(\omega) - \int_\Omega \phi^-(\omega) dP(\omega).$$

For a set $\Lambda \in \mathcal{F}$, the integral of ϕ over Λ is defined by

$$\int_{\Lambda} f(\omega)dP(\omega) = \int_{\Omega} 1_{\omega\in\Lambda} \cdot \phi(\omega)dP(\omega),$$

where $1_{\omega\in\Lambda}$ is the indicator function of the set Λ. Given δ is a nonnegative measurable function on the measure space $(\Omega, \mathcal{F}, P)$, a measure ν defined by

$$\nu(\Lambda) = \int_{\Lambda} \delta(\omega)dP(\omega), \quad \Lambda \in \mathcal{F}$$

is said to have density δ with respect to P. A random variable X on $(\Omega, \mathcal{F}, P)$ and its distribution μ have density f with respect to the Lebesgue measure λ if f is a nonnegative Borel function on $\mathbb{R}$ and

$$P(\omega : X(\omega) \in A) = \mu(A) = \int_{A} f(x)dx, \quad A \in \mathbb{R}.$$

For any random variable the density is assumed to be with respect to the Lebesgue measure λ if no other measure is specified. The density f and distribution function F_X of a random variable X are related by the following Lebesgue integral

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

The following theorem presents important relations involving integration and the density of a measure.

Theorem 3.2 (Theorem 16.11 in [19]) *If ν has density δ with respect to P, then*

$$\int_{\Omega} \phi(\omega)d\nu(\omega) = \int_{\Omega} \phi(\omega)\delta(\omega)dP(\omega), \tag{3.2.1}$$

holds for nonnegative ϕ. Moreover, ϕ, not necessarily nonnegative, is integrable with respect to ν if and only if $\phi\delta$ is integrable with respect to P, in which case (3.2.1) *and*

$$\int_{\Lambda} \phi(\omega)d\nu(\omega) = \int_{\Lambda} \phi(\omega)\delta(\omega)dP(\omega),$$

both hold. $\square$

We now address change of variables by a mapping and integration. Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be measurable spaces and $T : \Omega \to \Omega'$ a $\mathcal{F}/\mathcal{F}'$ measurable mapping. For a measure P on $\mathcal{F}$, PT^{-1} defines a measure on $\mathcal{F}'$ given by $PT^{-1}(\Lambda') = P(T^{-1}\Lambda')$, for $\Lambda' \in \mathcal{F}'$. The following theorem gives change of variable formulas for integration.

Theorem 3.3 (Theorem 16.13 in [19]) *If ϕ is nonnegative, then*

$$\int_{\Omega} \phi(T\omega)P(d\omega) = \int_{\Omega'} \phi(\omega')PT^{-1}(d\omega'). \tag{3.2.2}$$

A function ϕ, not necessarily nonnegative, is integrable with respect to PT^{-1} if and only if ϕT is integrable with respect to P, in which case (3.2.2) *and*

$$\int_{T^{-1}\Lambda'} \phi(T\omega)P(d\omega) = \int_{\Lambda'} \phi(\omega')PT^{-1}(d\omega'),$$

hold. $\square$

We can now use all the concepts of integration to define expectation. The expected value of a random variable X on $(\Omega, \mathcal{F}, P)$ is the integral of X with respect to the measure P:

$$E[X] = \int_{\Omega} X(\omega)dP(\omega).$$

For each Λ in $\mathcal{F}$, the truncated expectation is given by

$$E[X(\omega) \cdot 1_{\omega\in\Lambda}] = \int_{\Lambda} X(\omega)dP(\omega). \tag{3.2.3}$$

The following assumptions are made in the theorem that follows which shows different representations of the expectation.

Assumption 3.1 *The r.v. X on $(\Omega, \mathcal{F}, P)$ induces the probability space $(\mathbb{R}, \mathcal{B}, \mu)$.*

Assumption 3.2 *$g : \mathbb{R} \to \mathbb{R}$ is a Borel function so that $g(X)$ is a r.v. on $(\mathbb{R}, \mathcal{B}, \mu g^{-1})$.*

The following theorem shows the dual characterization of the expectation of a function.

Theorem 3.4 (Theorem 3.2.2 in [31]) *Under assumptions* 3.1 *and* 3.2

$$E[g(X)] = \int_{\Omega} g(X(\omega))dP(\omega) = \int_{\mathbb{R}} g(x)d\mu(x). \tag{3.2.4}$$

□

(3.2.4) follows directly from theorem 3.3, replacing $T : \omega \to \Omega'$ with $X : \Omega \to \mathbb{R}$, ϕ by g, setting $\omega' = x$, and noting $PX^{-1}(d\omega') = \mu(dx) = d\mu(x)$. Furthermore, under assumptions 3.1 and 3.2 and if X has density f with respect to the Lebesgue measure, we have

$$E[g(X)] = \int_{\mathbb{R}} g(x)d\mu(x) = \int_{\mathbb{R}} g(x)f(x)d\lambda = \int_{-\infty}^{\infty} g(x)f(x)dx. \tag{3.2.5}$$

(3.2.5) follows from theorem 3.2 by replacing ν with μ, P with λ, ω with x, ϕ with g, Ω with $\mathbb{R}$ and δ with f. If X has distribution function F_X with continuous derivatives we have $dF_X(x) = f(x)dx$ and

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx = \int_{-\infty}^{\infty} g(x)dF_X(x).$$

We now extend the results and definitions to multiple random variables. In $\mathbb{R}^k$, the k-dimensional Borel field $\mathcal{B}^k$ is $\sigma(\mathbb{R}^k)$, where $\mathbb{R}^k$ denotes the measurable rectangles, $B_1 \times B_2 \times \cdots \times B_k$ where $B_i \in \mathcal{B}$ for $i = 1, \ldots, k$, of $\mathbb{R}^k$. We call a measurable mapping X into $\mathbb{R}^k$, $X : \Omega \to \mathbb{R}^k$ a random vector on the space $(\Omega, \mathcal{F}, P)$ and write $X(\omega) = (X_1(\omega), \ldots, X_k(\omega))^\top$. X is measurable $\mathcal{F}$ if and only if each component mapping X_i is measurable $\mathcal{F}$. For a k-dimensional random vector $X = (X_1, \ldots, X_k)^\top$, the distribution μ, which is a probability measure on $\mathcal{B}^k$, and the distribution function are given by

$$\mu(A) = P(\omega : (X_1(\omega), \ldots, X_k(\omega)) \in A), \quad A \in \mathcal{B}^k,$$
$$F(x_1, \ldots, x_k) = P(\omega : X_1(\omega) \leq x_1, \ldots, X_k(\omega) \leq x_k) = \mu(S_x),$$

where $S_x = \{y : y_i \leq x_i, i = 1, \ldots, k\}$. A random vector X and its distribution μ have density f with respect to the k-dimensional Lebesgue measure λ if f is a nonnegative Borel function on $\mathbb{R}^k$ and

$$P(\omega : X(\omega) \in A) = \mu(A) = \int_A f(x_1, \ldots, x_k)dx_1 \cdots dx_k, \quad A \in \mathbb{R}^k.$$

If X is a k-dimensional random vector with distribution μ and $g : \mathbb{R}^k \to \mathbb{R}^i$ is measurable, then $g(X)$ is an i-dimensional random vector with distribution μg^{-1}. If $g_j : \mathbb{R}^k \to \mathbb{R}$ is defined by $g_j(x_1, \ldots, x_k) = x_j$, it follows $g_j(X) = X_j$ has distribution $\mu_j = \mu g_j^{-1}$ given by $\mu_j(A) = \mu[(x_1, \ldots, x_k) : x_j \in A] = P(\omega : X_j(\omega) \in A)$, for $A \in \mathbb{R}$. The μ_j are referred to as the marginal distributions of μ. If μ has density f with respect to the k-dimensional Lebesgue measure, μ_j has density f_j with respect to the one dimensional Lebesgue measure given by

$$f_j(x) = \int_{R^{k-1}} f(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k) dx_1 \cdots dx_{j-1} dx_{j+1} \cdots dx_k.$$

The random variables $X_1, \ldots, X_k$ are defined to be independent if the σ-fields they generate $\sigma(X_1)$,...,$\sigma(X_k)$ are independent. $X_1, \ldots, X_k$ are independent if and only if $P(X_1 \in H_1, \ldots, X_k \in H_k) = P(X_1 \in H_1) \cdots P(X_k \in H_k)$, and if and only if $P(X_1 \leq x_1, \ldots, X_k \leq x_k) = P(X_1 \leq x_1) \cdots P(X_k \leq x_k)$.

Given the random vector $(X_1, \ldots, X_k)$ with distribution μ having density f and distribution function F and each X_i with marginal distribution μ_i having density f_i and marginal distribution function F_i, $X_1, \ldots, X_k$ are independent if and only if μ is the product measure with $\mu = \mu_1 \times \cdots \times \mu_k$, if and only if $F(x_1, \ldots, x_k) = F_1(x_1) \cdots F_k(x_k)$, and if and only if $f(x) = f_1(x_1) \cdots f_k(x_k)$. For Borel measurable function $g : \mathbb{R}^k \to \mathbb{R}$ with $g^{-1}(B) \in \mathcal{B}^k$ for every $B \in \mathcal{B}$, $h(\omega) = g(X_1(\omega), \cdots, X_k(\omega))$ is a $\mathcal{F}/\mathcal{B}$ measurable r.v. and we have the expectation

$$E[g(X_1(\omega), \ldots, X_k(\omega))] = \int_\Omega h(\omega) dP(\omega).$$

Similarly, applying theorem 3.3,

$$E[g(X_1(\omega), \ldots, X_k(\omega))] = \int_{\mathbb{R}^k} g(x_1, \ldots, x_k) d\mu(x_1, \cdots, x_k),$$

and if μ has density f with respect to the k-dimensional Lebesgue measure λ by theorem 3.2 and Fubini's theorem

$$E[g(X_1(\omega), \ldots, X_k(\omega))] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_k) f(x_1, \ldots, x_k) dx_1 \cdots dx_k.$$

3.3 Truncated expectations

Under assumptions 3.1 and 3.2 for each $A \in \mathcal{B}$

$$E[g(X) \cdot 1_{\omega \in X^{-1}A}] = \int_{X^{-1}A} g(X(\omega))dP(\omega) = \int_A g(x)d\mu(x). \tag{3.3.1}$$

(3.3.1) follows from theorem 3.3 by replacing $T : \omega \to \Omega'$ with $X : \Omega \to \mathbb{R}$, ϕ by g, Λ' by A, setting $\omega' = x$ and noting $PX^{-1}(d\omega') = \mu(dx) = d\mu(x)$. Furthermore, under assumptions 3.1 and 3.2 and if X has density f with respect to the Lebesgue measure and $A = [a, b]$, we have

$$E[g(X) \cdot 1_{\omega \in X^{-1}A}] = \int_A g(x)d\mu(x) = \int_A g(x)f(x)d\lambda = \int_a^b g(x)f(x)dx. \tag{3.3.2}$$

(3.3.2) follows from theorem 3.2 by replacing ν with μ, P with λ, ω with x, ϕ with g, Λ with A and δ with f. If X has distribution function F_X with continuous derivatives we have $dF_X(x) = f(x)dx$ and

$$E[g(X) \cdot 1_{\omega \in X^{-1}A}] = \int_a^b g(x)f(x)dx = \int_a^b g(x)dF_X(x). \tag{3.3.3}$$

We refer to the expectation given by (3.3.1), (3.3.2), and (3.3.3) as the truncated expectation of $g(X)$ to A and write

$$\bar{E}[g(X), A] = E[g(X) \cdot 1_{\omega \in X^{-1}A}].$$

Truncated moments to A and truncated central moments to A about x_0 are

$$\bar{E}[X^k, A] = \int_{X^{-1}A} X^k(\omega)dP(\omega) = \int_A x^k d\mu(x),$$

$$\bar{E}[(X - x_0)^k, A] = \int_{X^{-1}A} (X(\omega) - x_0)^k dP(\omega) = \int_A (x - x_0)^k d\mu(x),$$

respectively. When the interval A is clear from context, we write $\bar{E}[X]$ for (3.2.3). For $A \in \mathbb{R}^k$ and $A = A_1 \times \cdots \times A_k$, the truncated expectation of $g(X_1(\omega), \ldots, X_k(\omega))$ to A

is given by

$$\begin{aligned}E[g(X_1(\omega),\dots,X_k(\omega))\cdot 1_{\omega\in X^{-1}A}] &= \int_{X^{-1}A} g(X_1(\omega),\cdots,X_k(\omega))dP(\omega)\\ &= \int_A g(x_1,\dots,x_k)d\mu(x_1,\dots,x_k)\\ &= \int_{A_1}\cdots\int_{A_k} g(x_1,\dots,x_k)f(x_1,\dots,x_k)dx_1\cdots dx_k.\end{aligned}$$

This expectation will be denoted as follows

$$\bar{E}[g(X_1(\omega),\dots,X_k(\omega)),A] = E[g(X_1(\omega),\dots,X_k(\omega))\cdot 1_{\omega\in X^{-1}A}].$$

We now present some properties of truncated expectations.

Assumption 3.3 *$X=(X_1,\dots,X_k)$ is a random vector on the space $(\Omega,\mathcal{F},P)$ into $\mathbb{R}^k$.*

Assumption 3.4 *$A\in\mathbb{R}^k$ and $A=A_1\times\cdots\times A_k$ where each A_i is an interval in $\mathbb{R}$.*

Proposition 3.5 (Martinez) *Given c is a real constant:*

1. *Under assumption 3.3, $\bar{E}[c,A]=cP(\omega: X(\omega)\in A)$,*
2. *Under assumptions 3.3 and 3.4, for $X_1,\dots,X_k$ independent, $\bar{E}[c,A]=cP(X_1\in A_1)\cdots P(X_k\in A_k)$.*

Proof. For 1, we write

$$\begin{aligned}\bar{E}[c,A] =& E[c\cdot 1_{\omega\in X^{-1}A}] = \int_{X^{-1}A} c dP(\omega) = \int_A c d\mu(x_1,\dots,x_k)\\ =& \int_{A_1}\cdots\int_{A_k} cf(x_1,\dots,x_k)dx_1\cdots dx_k = cP(\omega: X(\omega)\in A).\end{aligned}$$

For 2, with $X_1,\dots,X_k$ independent, it follows

$$\begin{aligned}&\int_{A_1}\cdots\int_{A_k} cf(x_1,\dots,x_k)dx_1\cdots dx_k = \int_{A_1}\cdots\int_{A_k} cf_1(x_1)\cdots f_k(x_k)dx_1\cdots dx_k\\ =& c\int_{A_1} f_1(x_1)dx_1\cdots\int_{A_k} f_k(x_k)dx_k = cP(X_1\in A_1)\cdots P(X_k\in A_k).\end{aligned}$$

■

Proposition 3.6 (Martinez) *Given assumptions* 3.3 *and* 3.4*, for* $X_1, \dots, X_k$ *independent*

$$\bar{E}[X_i, A] = \bar{E}[X_i, A_i]P(X_1 \in A_1) \cdots P(X_{i-1} \in A_{i-1})P(X_{i+1} \in A_{i+1}) \cdots P(X_k \in A_k),$$

for $i = 1, \dots, k$.

Proof. For $X_1, \dots, X_k$ independent

$$\begin{aligned}
\bar{E}[X_1, A] &= \int_{A_1} \cdots \int_{A_k} x_1 f_1(x_1) \cdots f_k(x_k) dx_1 \cdots dx_k \\
&= \int_{A_1} x_1 f_1(x_1) dx_1 \int_{A_2} f_2(x_2) dx_2 \cdots \int_{A_k} f_k(x_k) dx_k \\
&= \bar{E}[X_1, A_1] P(X_2 \in A_2) \cdots P(X_k \in A_k).
\end{aligned}$$

The general result follows if X_1 is replaced by any of the X_i's. ∎

Proposition 3.7 (Martinez) *Given assumptions* 3.3 *and* 3.4*, it follows:*

1. *Given assumption* 3.3*,* $\bar{E}\left[\sum_{i=1}^k c_i X_i, A\right] = \sum_{i=1}^k c_i \bar{E}[X_i, A]$.

2. *Given assumptions* 3.3 *and* 3.4*, for* $X_1, \dots, X_k$ *independent*

$$\begin{aligned}
\bar{E}\left[\sum_{i=1}^k c_i X_i, A\right] = &\sum_{i=1}^k c_i \bar{E}[X_i, A_i] P(X_1 \in A_1) \cdots P(X_{i-1} \in A_{i-1}) \\
&\cdot P(X_{i+1} \in A_{i+1}) \cdots P(X_k \in A_k).
\end{aligned}$$

Proof. For 1,

$$\begin{aligned}
\bar{E}\left[\sum_{i=1}^k c_i X_i, A\right] &= \int_{A_1} \cdots \int_{A_k} \sum_{i=1}^k c_i x_i f(x_1, \dots, x_k) dx_1 \cdots dx_k \\
&= \sum_{i=1}^k c_i \int_{A_1} \cdots \int_{A_k} x_i f(x_1, \dots, x_k) dx_1 \cdots dx_k = \sum_{i=1}^k c_i \bar{E}[X_i, A].
\end{aligned}$$

For $X_1, \dots, X_k$ independent or i.i.d the result follows from

$$\bar{E}\left[\sum_{i=1}^k c_i X_i, A\right] = \sum_{i=1}^k c_i \int_{A_1} f_1(x_1) dx_1 \cdots \int_{A_i} x_i f_i(x_i) dx_i \cdots \int_{A_k} f_k(x_k) dx_k.$$

■

Proposition 3.8 (Martinez) *Given assumptions* 3.3 *and* 3.4*, for* $X_1, \ldots, X_k$ *independent, it follows:*

1. $\bar{E}[X_1 X_2, A] = \bar{E}[X_1, A_1]\bar{E}[X_2, A_2]P(X_3 \in A_3) \cdots P(X_k \in A_k)$,
2. $\bar{E}[X_1 \cdots X_k, A] = \bar{E}[X_1, A_1] \cdots \bar{E}[X_k, A_k]$.

Proof. For $X_1, \ldots, X_k$ independent or i.i.d

$$\bar{E}[X_1 X_2, A] = \int_{A_1} x_1 f_1(x_1) dx_1 \int_{A_2} x_2 f_2(x_2) dx_2 \int_{A_3} f_3(x_3) dx_3 \cdots \int_{A_k} f_k(x_k) dx_k.$$

2 is a simple extension of 1. ■

Chapter 4

Taylor series approximations of expectations

4.1 Introduction

Evaluating the expectation of a function of random variables is an important problem with many applications. In econometrics, for estimators, which are functions of random variables, determining their moments is important to understand small and large sample properties. In economics and finance, approximating the expectation of utility functions is necessary to solve portfolio optimization problems, [95, 93, 71, 40]. This chapter presents an algorithm to approximate the expectation of functions of random variables based on Taylor series expansions. These techniques will be used in later chapters to approximate the expectation of functions with complicated dependencies on sums of random variables and other statistics.

4.2 Algorithm

We begin by considering univariate functions. Given a random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$ with continuous density function $f(x)$ and a Borel function $\varphi : \mathbb{R} \to \mathbb{R}$, the expected value of $Y \equiv \varphi(X)$ is given by

$$E[Y] = \int_{-\infty}^{\infty} s g(s) ds,$$

where g is the density function of Y. This expectation can be rewritten, as presented in (3.2.5), in the following form:

$$E[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(s) f(s) ds.$$

Obtaining an explicit analytic expression for this expectation by integration can be done in very few cases. Numerical integration is the most viable option. Most numerical procedures would involve knowing the functional form of the density. Such algorithms applied to real empirical problems would require estimating the distribution from data. In many situations, one would prefer to work with an expression of the expected value $E[\varphi(X)]$, which consists of a function of moments of the argument variable X. We study algorithms based on Taylor approximations which require estimation of only a few central moments. Such algorithms have been a standard device for computing expected utilities for portfolio optimization [125, 143]. In this literature, there has been much debate on the accuracy of approximating expectation of functions by means of a Taylor series expansion. But as we will discuss, much of the confusion can be settled with some basic theorems of integration and by putting aside issues concerning the appropriateness of utility functions.

The idea of approximating the expectation of a function by means of a Taylor series relies on the hope that taking the expectation of the function is equivalent to taking the expectation of its series representation, and in turn that the expectation of the series expansion is equivalent to summing the series of expected values of the series elements. There are two important mathematical issues which must be addressed to assess the viability of such an approximation. The first issue is the convergence of a Taylor series to the function it represents. The second issue has to do with term-by-term integrability of an infinite series. We begin by reviewing some concepts of convergent power series.

From the theory of infinite series of non-random variables, the Taylor series

$$\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(x - x_0)^k, \tag{4.2.1}$$

is a particular type of power series, and can represent the function $\varphi(x)$ in a neighborhood of x_0. Such neighborhood is referred to as the neighborhood of convergence of the series and is defined by the radius of convergence. The radius of convergence of (4.2.1) is given by

$$r = 1/\alpha \quad \text{with} \quad \alpha = \lim_{k\to\infty} (|\varphi^{(k)}(x_0)/k!|)^{1/k}. \tag{4.2.2}$$

For any $x \in B = \{x : |x - x_0| < r|\}$, the series (4.2.1) converges to $\varphi(x)$. For any $x \in B^c = \{x : |x - x_0| \geq r|\}$, the series (4.2.1) diverges.

We would like to understand a similar relation between a function of a random variable and a Taylor series with random elements. When considering random variables X and $\varphi(X)$ with density functions $f(x)$ and $g(x)$ respectively, the Taylor series

$$\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(X - x_0)^k, \tag{4.2.3}$$

has radius of convergence as defined by (4.2.2). The almost sure convergence of (4.2.3) with a finite radius of convergence r can be written as

$$\varphi(X)I(X \in B) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(X - x_0)^k I(X \in B) \quad \text{a.s.}, \tag{4.2.4}$$

where $I(\cdot)$ is the indicator function. For $r = \infty$ we have simply

$$\varphi(X) = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(X - x_0)^k \quad \text{a.s.}. \tag{4.2.5}$$

When considering the approximation of a function of non-random variables by a Taylor series, the approximation is only true within the radius of convergence. When considering the approximation of the expectation of a function of a random variable by a Taylor series, we must take into account not only the radius of convergence of the series but also the range of the random variable in question. The algorithm for computing the expected value of a function of a random variable based on a Taylor series approximation is therefore based on the following expression

$$E[\varphi(X)] = T_1 + T_2, \tag{4.2.6}$$

$$T_1 = E\left[\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(X - x_0)^k I(X \in A)\right] = \int_A \left[\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(s - x_0)^k\right] f(s)ds,$$

$$T_2 = E[\varphi(X)I(X \in A^c)] = \int_{A^c} \varphi(s)f(s)ds.$$

In (4.2.6), the interval of integration of the expectation is split. T_1 represents an integral whose interval A is a compact strict subset of the region of convergence B of the Taylor series of φ, and T_2 represents an integral over the complement of A denoted by A^c.

The objective is, given (4.2.6), to view T_1 as an approximation of $E[\varphi(X)]$ provided T_2 is small

$$E[\varphi(X)] \approx T_1.$$

T_1 gives an expression based on the central moments of the random variable X if the integral and summation can be interchanged. This is also known as integrating the series term-by-term. Therefore, the applicability of a Taylor's series expansion of $\varphi(X)$ to approximate the expectation $E[\varphi(X)]$ depends on the circumstances which allow for the following equality

$$\int_A \left[\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!}(x - x_0)^k f(x)\right] dx = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!} \int_A \left[(x - x_0)^k f(x)\right] dx. \tag{4.2.7}$$

Well known sufficient conditions concerning uniform convergence of series exist which allow the integral of a series to be computed term by term. Such conditions will be fundamental to the approximating algorithm we develop, and we state them in the following theorem.

Theorem 4.1 (Knopp, [87]) *The series $F(x) = \sum f_n(x)$ is assumed uniformly convergent in the interval J, and all the functions $f_n(x)$ are supposed integrable over the closed subinterval J': $a \leq x \leq b$, so that $F(x)$ is also continuous in that subinterval. Then $F(x)$ is also integrable over J' and the integral of $F(x)$ over the interval J' may be obtained by term-by-term integration*

$$\int_a^b \left[\sum_{k=0}^{\infty} f_n(x)\right] dx = \sum_{k=0}^{\infty} \left[\int_a^b f_n(x)dx\right]. \tag{4.2.8}$$

Similar sufficient conditions can be found in [7, 8, 51, 133]. The following theorem concerns uniform convergence of a Taylor series.

Theorem 4.2 (Apostol, [7]) *A power series converges uniformly on every compact subset interior to the neighborhood of convergence.*

The following theorem is necessary for the proposition to follow.

Theorem 4.3 (Knopp, [87]) *If $\sum f_n(x)$ is uniformly convergent in J, so is the series $\sum g(x) f_n(x)$, where $g(x)$ denotes any function defined and bounded in the interval J.*

We can now state a proposition.

Proposition 4.4 (Martinez) *Let $\varphi(x) : \mathbb{R} \to \mathbb{R}$ be a function whose Taylor series representation about the point x_0 has neighborhood of convergence $B = \{x : |x - x_0| < r\}$. Let X be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$ with bounded density function f, mean $E[X] = \mu$, with $E(X-\mu)^k < \infty$ for $k = 1, 2 \ldots$, and $E[\varphi(X)] < \infty$. Then*

$$E[\varphi(X)] = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!} \bar{E}[(X-\mu)^k, A] + E[\varphi(X) I\{X \in A^c\}], \tag{4.2.9}$$

where $A \subset B$ and

$$\bar{E}[(X-\mu)^k, A] = \int_A (s-\mu)^k f(s) ds, \quad k = 1, 2, \ldots$$

will be referred to as truncated central moments. Truncated expectations are defined in section 3.3.

Proof. $E[\varphi(X)]$ can be written as (4.2.6). It is only left to prove that (4.2.7) holds. By theorem 4.2, the Taylor series representation of $\varphi(x)$ converges uniformly on every compact subset A' of the neighborhood of convergence B. By theorem 4.3, the series

$$\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!} (x - x_0)^k f(x),$$

converges uniformly on the compact subset A'. By theorem 4.1, (4.2.7) holds with A a compact subset of A'. ■

It is important to note that the conditions of the proposition are sufficient and not necessary. Every time truncated central moments are used with $A \subset B$, the series in (4.2.9) will converge. The necessity of the conditions fail because there are series that can be integrated term-by-term which do not converge uniformly. Furthermore, the conditions of uniform convergence restrict the interval A to be compact.

4.3 Examples

Given a random variable X, the relevance of the radius of convergence of the Taylor series representation of a function φ when approximating the expected value of $\varphi(X)$ was first pointed out in [95]. Unfortunately, the author provides misleading explanations for the conclusion reached. The author concludes:

> The counterexamples confirm the analytic result that the interval of convergence prohibits the application of a Taylor's series expansion for a logarithmic and power utility function. Regardless of what sort of probability distribution is involved, the approximation does not work.... We can conclude that the hitherto common Taylor's series expansion yields an exact result for the normal distribution, exponential utility combination only.

The main problem with the author's conclusions is applying the uniform convergence conditions of theorem 4.1 as necessary rather than sufficient conditions.

> The integral of an infinite sum is equal to the sum of an infinite series of integrals only if the series converges uniformly.

Furthermore, the author fails to realize the need to use (4.2.6) and (4.2.7) in the approximation. No satisfactory alternative solution is given in [95] to the problem of erroneous approximations resulting from inappropriate use of the Taylor's series. Proposition 4.4 provides such alternative solution. In the following examples, we apply the results of proposition 4.4 to the numerical cases studied in [95].

We consider the utility functions examined in [95]. These include an exponential, a power, and a logarithmic utility function as given below:

$$U(x) = 1000(1 - \exp(-0.05x)), \tag{4.3.1}$$

$$U(x) = \begin{cases} \frac{x^{0.5}}{.5}, & x > 0 \\ 0, & x = 0 \\ -\frac{x^2}{2}, & x < 0 \end{cases}, \tag{4.3.2}$$

$$U(x) = \ln x, \quad x > 0. \tag{4.3.3}$$

The radius of convergence for the exponential utility is infinity. For the power utility as well as the logarithmic utility, the radius of convergence is equal the point x_0 around which the series expansion is made. In our examples, x_0 is equal to the mean of the random variable X.

As in [95], we investigate normally distributed returns with probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty,$$

with numerical parameters $\mu = 10$ and $\sigma^2 = 82$ and lognormally distributed returns with mean $m = 10$ and variance $s^2 = 82$ with probability distribution:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2\right], \quad 0 < x < \infty.$$

The parameters of the lognormal distribution are

$$\mu = \log\left[\frac{m^2}{(s+m^2)^{1/2}}\right] \approx 2.0032, \quad \sigma^2 = \log\left[\frac{s}{m^2}+1\right] \approx 0.5988.$$

The central moments of the normal distribution are given by

$$m_{2k-1} = 0, \quad m_{2k} = \frac{(2k)!\sigma^{2k}}{k!2^k}, \quad k = 1, 2, \ldots.$$

The central moments, m_j, of the lognormal distribution can be obtained from the raw moments M_j

$$M_j = \exp\left(j\mu + \frac{(j\sigma)^2}{2}\right), \quad m_j = \sum_{k=0}^{j} \left(\begin{array}{c} j \\ k \end{array}\right) (-1)^k M_{j-k} M_1^k.$$

We compare the approximations obtained using the inappropriate Taylor series algorithm in [95] to approximations obtained using the algorithm in proposition 4.4.

In table 4.1, we present results for the power utility function (4.3.2) and normally distributed returns. The second of five columns labeled central moments presents the approximation to the expected value of $E[\varphi(X)]$ obtained using the algorithm in [95]. Columns three, four, and five present the approximations to the expectation $E[\varphi(X)I\{X \in A\}]$ given by the expression

$$\sum_{k=0}^{n} \frac{\varphi^{(k)}(x_0)}{k!} \bar{E}[(X-\mu)^k, A],$$

which forms part of the expected value $E[\varphi(X)]$ according the algorithm in proposition 4.4 with $\mu = 10$, $A = \{x : a \leq x \leq b\}$, $a = 1 \times 10^{-10}$, and $b = 40$, $b = 19$, and $b = 10$ respectively.

The row labeled EU_{int} presents the expected value of utility computed by numerical integration. The entries below the label EU_{moment} in column two are the expected value of utility computed with the algorithm in [95], aggregating even order central moments from the second to the twentieth and the sixtieth. The entries below the label EU_{moment} in column three are the expected value of the truncated utility computed with the algorithm in proposition 4.4 using truncated central moments with $A = \{x : 1 \times 10^{-10} \leq x \leq 40\}$. Similarly for columns four and five with $b = 19$ and $b = 10$, respectively.

The results of Table 4.1 column two demonstrate as stated in [95] that the inappropriate algorithm provides diverging approximations to the expectation as the number of Taylor series terms increases. The results of table 4.1, column three, with truncated moments with $b = 40$, demonstrate approximations of $E[\varphi(X)I\{X \in A\}]$ will diverge when the condition $A \subset B$ is violated which is the case for column three, since for the power utility $B = \{x : 0 < x < 20\}$.

Table 4.1: Expected utility for the power function with normal distribution of returns

	Central moments	Truncated central moments		
		b=40	b=19	b=10
EU_{int}	3.29127	5.72865	4.17849	1.63261
EU_{moment}				
n				
2	5.67629	5.68382	4.23834	1.70708
4	5.17793	5.63197	4.20645	1.66358
6	4.10522	5.46344	4.19525	1.65037
8	0.17441	4.81374	4.18990	1.64443
10	-20.3739	1.88376	4.18687	1.64119
12	-160.436	-13.0124	4.18497	1.63920
14	-948.491	-95.9590	4.18369	1.63787
16	-8694.25	-590.983	4.18278	1.63693
18	-97456.3	-3707.77	4.18211	1.63625
20	-1.26×10^6	-24175.8	4.18159	1.63572
60	-1.75×10^{35}	-1.32×10^{22}	4.17910	1.63322

Table 4.2: Expected utility for the exponential function with lognormal distribution of returns

	Central moments	Truncated central moments		
		b=250	b=40	b=10
EU_{int}	348.45852	348.45579	334.95660	149.78575
EU_{moment}				
n				
2	331.30006	331.45016	331.10499	151.35090
4	332.96127	334.91574	334.64081	149.79519
6	298.42314	332.66927	334.94485	149.78578
8	-49.125352	330.33387	334.95632	149.78575
10	-4065.2351	331.54319	334.95659	149.78575
12	-46885.083	336.12332	334.95660	149.78575
14	-431291.72	341.38923	334.95660	149.78575
16	-3.27×10^6	345.21978	334.95660	149.78575
18	-2.06×10^7	347.24970	334.95660	149.78575
20	-1.31×10^8	348.08368	334.95660	149.78575
60	-2.86×10^{11}	348.45579	334.95660	149.78575

Table 4.3: Expected utility for the power function with lognormal distribution of returns

	Central moments	Truncated central moments		
		b=40	b=19	b=10
EU_{int}	5.86842	5.65357	4.67461	2.934209
EU_{moment}				
n				
2	5.67629	5.64162	4.75615	3.023612
4	1.71872	5.51842	4.69837	2.959522
6	-689.208	5.16860	4.68431	2.944414
8	-724895	3.57674	4.67933	2.939136
10	-2.02×10^9	-4.75683	4.67717	2.936872
12	-8.75×10^{12}	-52.4990	4.67611	2.935765
14	-4.70×10^{16}	-344.679	4.67555	2.935173
16	-2.86×10^{20}	-2224.52	4.67522	2.934834
18	-1.90×10^{24}	-14798.9	4.67502	2.934629
20	-1.32×10^{28}	-101559	4.67489	2.934500
60	-3.9×10^{106}	-7.53×10^{22}	4.67461	2.934212

Table 4.4: Expected utility for the logarithmic function with lognormal distribution of returns

	Central moments	Truncated central moments		
		b=40	b=19	b=10
EU_{int}	2.00317	1.94496	1.63163	1.01679
EU_{moment}				
n				
2	1.89259	1.96099	1.71776	1.10915
4	-2.28348	1.80780	1.66498	1.05186
6	-895.907	1.30455	1.64785	1.03370
8	-1.096×10^6	-1.25812	1.64059	1.02608
10	-3.450×10^9	-16.0976	1.63701	1.02234
12	-1.65×10^{13}	-108.832	1.63506	1.02032
14	-9.59×10^{16}	-720.681	1.63392	1.01914
16	-6.26×10^{20}	-4925.26	1.63321	1.01841
18	-4.40×10^{24}	-34743.5	1.63276	1.01794
20	-3.24×10^{28}	-251570	1.63245	1.01763
60	-1.7×10^{107}	-3.25×10^{23}	1.63165	1.01681

The results of columns four and five show approximations which converge towards the quantities obtained by numerical integration as the number of Taylor series terms increases. These results validate proposition 4.4.

Table 4.2 has the same format as table 4.1, and presents results for the exponential utility function with lognormally distributed returns. Again the results in column two demonstrate as stated in [95] that the inappropriate algorithm provides diverging approximations to the expectation of the exponential function as the number of Taylor series terms increases. For the exponential function we know $B = \mathbb{R}$. Therefore, proposition 4.4 implies any compact subset A will result in converging approximations to $E[\varphi(X)I\{X \in A\}]$. This is indeed confirmed by the results of columns three, four, and five.

Tables 4.3 and 4.4 provide similar results for the power function with lognormal returns, and for the logarithmic function with lognormal returns respectively. To conclude, all four numerical examples demonstrate that when the radius of convergence of $\varphi(x)$ is finite, the algorithm of proposition 4.4 with truncated moments and with A a compact subset of B provides a convergent Taylor series approximation of $E[\varphi(X)I\{X \in A\}]$. Otherwise, for the case of exponential utility and infinite radius of convergence, one can still choose a compact set A to obtain a convergent Taylor series approximation of $E[\varphi(X)I\{X \in A\}]$.

4.4 Approximation error

As stated above, using truncated expectations we can write

$$E[\varphi(X)] = \bar{E}[\varphi(X), A] + \bar{E}[\varphi(X), A^c],$$

where A is a compact subset of the neighborhood of convergence and A^c is its complement. Furthermore, for any finite n, and under the assumptions of proposition 4.4, we can write

$$\bar{E}[\varphi(X), A] = \sum_{k=0}^{n} \frac{\varphi^{(k)}(\mu)}{k!} \bar{E}[(X-\mu)^k, A] + \bar{E}[R(n), A],$$

where $R(n)$ is the Lagrange remainder of the Taylor series. It follows, the approximation error is given by

$$E[\varphi(X)] - \sum_{k=0}^{n} \frac{\varphi^{(k)}(\mu)}{k!} \bar{E}[(X-\mu)^k, A] = \bar{E}[R(n), A] + \bar{E}[\varphi(X), A^c]. \tag{4.4.1}$$

Proposition 4.4 and the examples in the previous section provide and demonstrate the methodology to approximate $E[\varphi(X)I\{X \in A\}]$ by means of converging Taylor series approximations. The accuracy of the approximation $E[\varphi(X)] \approx E[\varphi(X)I\{X \in A\}]$ depends on the size of $E[\varphi(X)I\{X \in A^c\}]$. In what follows, we attempt to find bounds for $E[\varphi(X)I\{X \in A^c\}]$ in order to make improvements on the approximation $E[\varphi(X)] \approx E[\varphi(X)I\{X \in A\}]$ by using an approximation which incorporates a measure on the size of $E[\varphi(X)I\{X \in A^c\}]$. To accomplish this, we must define a particular class of functions.

Consider a random variable X defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with $E[X] = \theta < \infty$. Let $\mathcal{A} \subset \mathbb{R}$ be an interval, possibly unbounded, with $P(X \in \mathcal{A}) = 1$. Define $\mathcal{G}_\alpha$ as the class of functions defined on $\mathcal{A}$ such that $\varphi \in \mathcal{G}_\alpha$ implies φ has a Taylor series expansion about θ with a possibly unbounded neighborhood of convergence $B \subseteq \mathcal{A}$ and with

$$|\varphi(x)| = O(|x|^\alpha) \quad \text{as} \quad |x| \to \infty.$$

Therefore, given $\varphi \in \mathcal{G}_\alpha$, there exists an $N > 0$ such that $|\varphi(x)| \le c|x|^\alpha$ for some constant c for all x with $|x - \theta| \ge N$. Otherwise, $|\varphi(x)| \le M$ for $|x - \theta| < N$ for some constant M. We now present a number of assumptions followed by a proposition which provides an approximation of $E[\varphi(X)]$ in terms of the infinite Taylor series and which takes into account a bound on the approximation error $\bar{E}[\varphi(X), A^c\}]$.

Assumption 4.1 $\varphi \in \mathcal{G}_\alpha$.

Assumption 4.2 *X is a random variable with $E[X] = \theta$ and $E(X-\theta)^k < \infty$ for $k = 1, 2, \ldots$.*

Assumption 4.3 $E[\varphi(X)] < \infty$.

Assumption 4.4 *$A = \{x : a \le x \le b\}$ is a compact subset of the neighborhood of convergence B of $\varphi(x)$, $\theta + N_1 \ge b$ and $\theta - N_2 \le a$.*

Proposition 4.5 (Martinez) *Under assumptions 4.1 through 4.4*

$$E[\varphi(X)] \le \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(\theta)}{k!} \bar{E}(X-\theta)^k + R_1 + R_2,$$

where $R_1 = MP(X \in \Lambda)$, $\quad R_2 = cE[|X|^\alpha I\{X \in \Lambda^c\}]$, *and*

$$\Lambda = \{x : \theta - N_2 < x < a\} \cup \{x : b < x < N_1 + \theta\},$$
$$\Lambda^c = \{x : x \ge N_2 + \theta\} \cup \{x : x \le \theta - N_1\},$$
$$\bar{E}(X-\mu)^k = \int_A (s-\mu)^k f(s)ds, \quad k = 1, 2, \ldots.$$

Proof. We write $E[\varphi(X)] = T_1 + T_2$ where,

$$T_1 = E[\varphi(X) I\{X \in A\}], \quad T_2 = E[\varphi(X) I\{X \in A^c\}].$$

We find expressions for T_1 and T_2 beginning with T_2. Since $\varphi \in \mathcal{G}_\alpha$,

$$E[\varphi(X) I\{X \in \Lambda\}] \le MP(X \in \Lambda),$$

where the constant M depends on b and N. Similarly $E[\varphi(X) I\{X \in \Lambda^c\}] \le cE[|X|^\alpha I\{X \in \Lambda^c\}]$. It follows $T_2 \le MP(X \in \Lambda) + cE[|X|^\alpha I\{X \in \Lambda^c\}]$. For T_1, by proposition 4.4 we can write

$$T_1 = \sum_{k=0}^{\infty} \frac{\varphi^{(k)}(x_0)}{k!} \bar{E}(X-\mu)^k \tag{4.4.2}$$

and the theorem is proven. ■

R_1 and R_2 are in terms of absolute moments and probabilities of the random variable X, both of which can be calculated or estimated easily with knowledge of X. The corollary that follows gives a bound similar to proposition 4.5 but with the infinite series replaced by a finite sum and a bounded remainder.

Assumption 4.5 *M_q is a number such that $|\varphi^{(q+1)}(x)| \leq M_q$ for every $x \in A$*

Corollary 4.6 *Under assumptions 4.1 through 4.5*

$$E[\varphi(X)] \leq \sum_{k=0}^{q} \frac{\varphi^{(k)}(\theta)}{k!} \bar{E}(X-\theta)^k + R_1 + R_2 + R_3,$$

where $R_1 = MP(X \in \Lambda), \quad R_2 = cE[|X|^\alpha I\{X \in \Lambda^c\}],$

$$R_3 = \frac{M_q}{(q+1)!} \bar{E}|X-\theta|^{q+1},$$

$$\Lambda = \{x : \theta - N_2 < x < a\} \cup \{x : b < x < N_1 + \theta\},$$

$$\Lambda^c = \{x : x \geq N_2 + \theta\} \cup \{x : x \leq \theta - N_1\},$$

$$\bar{E}(X-\mu)^k = \int_A (s-\mu)^k f(s)ds, \quad \bar{E}|X-\mu|^k = \int_A |s-\mu|^k f(s)ds \quad k = 1, 2, \ldots.$$

Proof. This follows from equating the Taylor series and the Taylor polynomial plus remainder as follows

$$\sum_{k=0}^{\infty} \frac{\varphi^{(k)}(\theta)}{k!} (x-\theta)^k = \sum_{k=0}^{q} \frac{\varphi^{(k)}(\theta)}{k!} (x-\theta)^k + R_q(x),$$

where $R_q(x) = \varphi^{(q+1)}(c)(x-\theta)^{q+1}/(q+1)!$ for some c in the interval (a, x). The result follows since $|R_q(x)| \leq M_q |x-\theta|^{q+1}/(q+1)!$ for every $x \in A$ ∎

Example 4.7 *We revisit two of the numerical examples studied in section* 4.3*. We apply the result of proposition* 4.5 *to the example of an exponential utility function with lognormal distribution of returns and to the example of a power utility function with lognormal distribution of returns. The results are presented in tables* 4.5 *and* 4.6*.*

There are two main ways to improve on the error from R_2. For functions in general, one is to do piecewise linear approximations of φ is Λ^c. Another, simpler, method can be applied to functions like the power and logarithmic utilities. These functions have radius of convergence equal to the point at which the Taylor series expansion is taken. Instead of evaluating the Taylor series around the mean θ, one can do the evaluation at some large value x_0. This effectively reduces the size of Λ^c.

Table 4.5: Expected utility for the exponential function with lognormal distribution of returns

	Truncated central moments		
	b=250	b=40	b=10
EU_{int}		348.45852	
EU_{moment}	R_1, R_2	R_1, R_2	R_1, R_2
n			
2	331.45289	345.79466	500.75764
4	334.91847	349.33048	499.20193
6	332.67200	349.63452	499.19252
8	330.33660	349.64600	499.19249
10	331.54591	349.64626	499.19249
12	336.12605	349.64626	499.19249
14	341.39195	349.64626	499.19249
16	345.22251	349.64626	499.19249
18	347.25243	349.64626	499.19249
20	348.08641	349.64626	499.19249
60	348.45851	349.64626	499.19249

Table 4.6: Expected utility for the power function with lognormal distribution of returns

	Truncated central moments	
	b=19	b=10
EU_{int}	5.86842	
EU_{moment}	R_1, R_2	R_1, R_2
n		
2	6.26983	7.18741
4	6.21205	7.12332
6	6.19798	7.10821
8	6.19300	7.10293
10	6.19085	7.10067
12	6.18979	7.09956
14	6.18922	7.09897
16	6.18889	7.09863
18	6.18869	7.09843
20	6.18857	7.09830
60	6.18829	7.09801

Table 4.7: Expected utility for the power function with lognormal distribution of returns

	Truncated central moments				
EU_{int}	5.86842				
EU_{moment}	$x_0 = 5$	$x_0 = 10$	$x_0 = 50$	$x_0 = 100$	$x_0 = 500$
	$b = 9$	$b = 19$	$b = 99$	$b = 199$	$b = 999$
n					
2	2.62402	4.75615	7.28777	8.95420	17.43726
4	2.61850	4.69837	6.50790	7.48044	13.18943
6	2.61721	4.68431	6.22911	6.89060	11.26424
8	2.61679	4.67933	6.09514	6.57976	10.13221
10	2.61661	4.67717	6.02043	6.39213	9.37765
12	2.61654	4.67611	5.97474	6.26909	8.83565
14	2.61649	4.67555	5.94496	6.18372	8.42643
16	2.61647	4.67522	5.92462	6.12200	8.10619
18	2.61646	4.67502	5.91021	6.07593	7.84872
20	2.61645	4.67489	5.89971	6.04066	7.63732
50	2.61644	4.67462	5.86317	5.89335	6.42874
100	2.61644	4.67461	5.85984	5.87196	6.05069
200	2.61644	4.67461	5.85949	5.86848	5.91294
5000	2.61644	4.67461	5.85947	5.86809	5.86842

Example 4.8 *We revisit one of the numerical examples studied in section* 4.3*. We apply the result of proposition* 4.5 *to the example of a power utility function with lognormal distribution of returns. The Taylor series is expanded at the point* x_0*. We evaluate different approximations as the point* x_0 *and the end point* b *of the interval* A *get larger. The results presented in table* 4.7 *demonstrate the approximation improves as* n *and the value of* x_0 *increase.*

In later chapters we present Delta method results where one is interested in how the expectation depends on some parameter. In these cases the results are equalities rather than the bound given in the above theorem. These equalities are obtained by including big-O expressions of certain order of the parameter.

Chapter 5

Taylor algorithm for independent identically distributed processes

5.1 Introduction

In this chapter, we construct an algorithm which yields an approximation, based on Taylor series, of the mean square forecast error (MSFE) for a forecasting problem involving independent and identically distributed processes. This Taylor algorithm approximation is meant to be used as a tool to describe the sample size dependence (SSD) of the MSFE.

Sample size dependence refers to the dependence of a statistic on a parameter or parameters which embody information concerning the amount of data involved in the formation of the statistic. For example, consider a stationary stochastic process $\{X_\tau\}_{\tau=1}^{N}$ with $E[X_i] = \mu_x$ and variance σ_x^2 $\forall i$. The sample mean of the process, $\bar{\mu}_{x,n} = 1/n \sum_{i=1}^{n} X_i$, is a random variable and a statistic with n describing the sample size. One might be interested in the behavior of this random variable for different values of n. Large sample theory would tell us $\bar{\mu}_{x,n}$ is consistent, $\bar{\mu}_{x,n} \xrightarrow{P} \mu_x$. Of more interest is the behavior of $\bar{\mu}_{x,n}$ for finite values of n. For this, we investigate the SSD of two moments of the sample mean: the expected value of $\bar{\mu}_{x,n}$, and the mean square error (MSE) of $\bar{\mu}_{x,n}$ and μ_x. The expected value of $\bar{\mu}_{x,n}$, $E[\bar{\mu}_{x,n}] = \mu_x$, is independent of n. This is the unbiased property of the sample mean and, again, not of much use for the purpose at hand. The MSE between $\bar{\mu}_{x,n}$ and μ_x, $MSE = E[(\bar{\mu}_{x,n} - \mu_x)^2]$, gives a measure of the average squared deviation of $\bar{\mu}_{x,n}$ from μ_x. This can be helpful to understand, on average, how much of

$\bar{\mu}_{x,n}$ differs from μ_x for finite values of n. This SSD can be derived explicitly. First

$$E[\bar{\mu}_{x,n}^2] = E\left[\frac{1}{n^2}\Big(\sum_{i=1}^{n} X_i\Big)^2\right] = \frac{1}{n^2}\Big(\sum_{i=1}^{n} E[X_i^2] + \sum_{i\neq j} E[X_i X_j]\Big)$$
$$= \frac{1}{n}(\sigma_x^2 + \mu_x^2) + \Big(1 - \frac{1}{n}\Big)(\gamma_{1,2} + \mu_x^2),$$

where $\gamma_{1,2} = Cov(X_1, X_2)$ and we make use of the stationarity assumption.

$$MSE_n = E[(\bar{\mu}_{x,n} - \mu_x)^2] = E[\bar{\mu}_{x,n}^2 - 2\bar{\mu}_{x,n}\mu_x + \mu_x^2] = E[\bar{\mu}_{x,n}^2] - \mu_x^2,$$

and finally we obtain

$$MSE_n = \frac{1}{n}(\sigma_x^2 - \gamma_{1,2}) + \gamma_{1,2}. \tag{5.1.1}$$

(5.1.1) gives the explicit dependence of the MSE on the sample size n. We can see that on average, the square difference between $\bar{\mu}_{x,n}$ and μ_x decays as $1/n$ to $\gamma_{1,2}$.

Just as the sample mean is a statistic, a forecast, in a forecasting problem as described in Chapter 2, is a statistic constructed from some predetermine functional form and an estimator, scalar, or vector. This estimator, another statistic, will depend on a variable n, describing the size of the sample used to form the estimator. Therefore, given a stochastic process $\{Y_\tau\}$, and a forecast $\hat{Y}_{t+1,n}$ of Y_{t+1}, we are interested in understanding how the average squared difference between $\hat{Y}_{t+1,n}$ and Y_{t+1} behaves for different values of the sample size n. We are interested in the SSD of the MSFE. In forecasting, understanding the SSD of the MSFE can be of great importance. This is especially true if we can find classes of processes for which analyzing the SSD results in an optimal observation window which provides the best forecast possible for a particular estimator.

In this chapter, we propose to understand the SSD of the MSFE for a forecasting problem involving independent processes. The forecasting model is assumed linear and the estimator of choice is the OLS. Unlike the motivating example of the sample mean given above, determining the SSD of the MSFE can not be done explicitly. This is due, in the scalar case, to the fractional functional form of the OLS and, in the multi-variate case, to the inversion of a matrix of sample data. This complication can not be simply solved by a different choice of estimator, since the OLS tends to be the simplest estimator available. One of the main contributions of this thesis is to overcome this difficulty by developing a methodology to extract the SSD from a statistic such as the MSFE with a complicated functional form.

The methodology proposed consists of writing the square forecast error (SFE) as a function of two statistics. This function is approximated by a Taylor expansion with respect to the two statistics about two points, the expectation of the two statistics. We obtain an approximation of the MSFE by taking the expectation of the Taylor approximation of the SFE. The expected value of the resulting Taylor approximation is a polynomial of central moments of the two statistics. These central moments are subsequently expanded and simplified to extract the explicit sample size dependence which is manifested in the sample size variable n. The final expression for the approximation of the MSFE is a polynomial in $1/n$ with coefficients consisting of functions of moments of the observed dependent and explanatory processes. The algorithm makes no assumptions on the form of the DGP for the dependent variable. This allows us to investigate the ramifications of misspecification in the forecasting problem and how these might manifest themselves in the SSD.

The rest of the chapter is organized as follows. In Section 5.2, we review some properties of the OLS and MSFE under the assumption of a correctly specified forecast model. Section 5.3 describes properties of the OLS under the assumption of a functionally misspecified model which have repercussions for the forecasting problem. Section 5.4 presents the derivation of the Taylor algorithm for the scalar case, and Section 5.5 presents the derivation for the multi-variate case. Finally, in Section 5.6, the performance of the Taylor algorithm for the MSFE of a scalar forecasting problem is evaluated with Monte Carlo experiments, and Section 5.7 concludes.

5.2 Properties for the OLS and MSFE under correct specification

Let $\{Y_\tau\}$ be an observable scalar process of interest to a forecaster. In general, the DGP is not known to the forecaster and therefore, in order to forecast, she must construct mathematical models which best capture empirical characteristics of the observed process. The forecaster might be interested in formulating her predictions based on linear models of the process. A linear regression model is a correspondence which relates the dependent variable Y_{t+1} to a $(m \times 1)$ vector of explanatory variables, X_t as follows:

$$Y_{t+1} = X_t^\top \phi + V_{t+1}, \tag{5.2.1}$$

where $\{V_\tau\}$ is a scalar innovation process and ϕ is a vector of parameters. If we identify t as the present time, we consider the sample of the n most resent observations $(y_{t-n+1}, \ldots, y_t, x_{t-n}, \ldots, x_{t-1})$. The OLS estimate based on such a sample is the value of ϕ which minimizes the residual sum of squares (RSS):

$$RSS \equiv \sum_{\tau=t-n}^{t-1} (y_{\tau+1} - x_\tau^\top \phi)^2. \tag{5.2.2}$$

In this section, as in most literature treatments of linear regression, we make the draconian assumption that the DGP can be described by a mathematical relation of the process $\{Y_\tau\}$ which coincides with the form of the regression model (5.2.1):

$$DGP : Y_{t+1} = X_t^\top \beta + U_{t+1}, \tag{5.2.3}$$

where U_t is a scalar innovation process and β is a vector of parameters. In other words, we assume the model (5.2.1) is correctly specified. β is often refereed to as the true parameter vector and the objective of it is to obtain the best possible estimate for this parameter vector, based on the observed sample. Under condition (5.2.3), the OLS estimate of β obtained from the minimization of the RRS is given by:

$$\hat{\beta}_{t,n} = \Big[\sum_{\tau=t-n}^{t-1} x_\tau x_\tau^\top \Big]^{-1} \sum_{\tau=t-n}^{t-1} x_\tau y_{\tau+1}. \tag{5.2.4}$$

The OLS sample residual for observation t is $\hat{v}_t \equiv y_t - x_t^\top \hat{\beta}_{t,n}$. We now return to deal with the random processes and define the following objects:

$$X_{t,n} \equiv (X_{t-n}, ..., X_{t-1})^\top \in \mathbb{R}^{n\times m},$$
$$Y_{t,n} \equiv (Y_{t-n+1}, \ldots, Y_t)^\top \in \mathbb{R}^{n\times 1},$$
$$Q_{t,n} \equiv X_{t,n}^\top X_{t,n} \in \mathbb{R}^{m\times m},$$
$$U_{t,n} \equiv (U_{t-n+1}, \ldots, U_t)^\top \in \mathbb{R}^{n\times 1}.$$

As a function of the random processes, the OLS is a statistic and can be written as follows:

$$\hat{\beta}_{t,n} = \beta + Q_{t,n}^{-1} X_{t,n}^\top U_{t,n}. \tag{5.2.5}$$

(5.2.5) gives the relation between the true parameter β and the OLS estimator $\hat{\beta}_{t,n}$. This relation is true because of the correctly specified assumption given by condition (5.2.3). Many results concerning the OLS exist based on different assumptions on the explanatory variables and the innovation process [61, 64]. We will focus on result for a specific set of assumptions.

Assumption 5.1 *X_t is stochastic and independent of U_s for all t, s.*

Assumption 5.2 *U_t is i.i.d with mean zero and variance σ_u^2.*

Taking expectations of (5.2.5) and exploiting assumption 5.2,

$$E[\hat{\beta}_{t,n}] = \beta + E[Q_{t,n}^{-1} X_{t,n}^\top] E[U_{t,n}] = \beta,$$

so that the OLS estimator is unbiased.

For asymptotic results, our interest is in the behavior of $\hat{\beta}_{t,n}$ as n becomes large. We begin by establishing consistency of the OLS for which we need the following assumption.

Assumption 5.3 $(1/n) \sum_{\tau=t-n}^{t-1} X_\tau X_\tau^\top \xrightarrow{P} Q$*, a positive definite matrix.*

From (5.2.5), we write

$$\hat{\beta}_{t,n} - \beta = \Big[(1/n) \sum_{\tau=t-n}^{t-1} X_\tau X_\tau^\top\Big]^{-1} \Big[(1/n) \sum_{\tau=t-n}^{t-1} X_\tau U_{\tau+1}\Big]. \tag{5.2.6}$$

For the first term of (5.2.6), assumption 5.3 and theorem $A.19$ imply

$$\Big[(1/n) \sum_{\tau=t-n}^{t-1} X_\tau X_\tau^\top\Big]^{-1} \xrightarrow{P} Q^{-1}. \tag{5.2.7}$$

For the second term of (5.2.6), note $X_\tau U_{\tau+1}$ is a martingale difference sequence with a finite variance-covariance matrix given by $E[X_\tau U_{\tau+1} X_\tau^\top U_{\tau+1}] = \sigma_u^2 E[X_\tau X_\tau^\top]$. By proposition $A.31$,

$$\Big[(1/n) \sum_{\tau=t-n}^{t-1} X_\tau U_{\tau+1}\Big] \xrightarrow{P} 0. \tag{5.2.8}$$

Applying proposition $A.20$ to (5.2.6), (5.2.7) and (5.2.8),

$$\hat{\beta}_{t,n} - \beta \xrightarrow{P} Q^{-1} \cdot 0 = 0,$$

confirming the consistency of the OLS estimator. For the asymptotic distribution of the OLS we require a further assumption.

Assumption 5.4 $E[X_\tau X_\tau^\top] = Q_\tau$, *a positive definite matrix with* $(1/T) \sum_{\tau=1}^{T} Q_\tau \to Q$.

Under the assumptions above, it can be shown (see [64], p . 210) that

$$\sqrt{T}(\hat{\beta}_{t,n} - \beta) \xrightarrow{L} N(0, \sigma_u^2 Q^{-1}). \tag{5.2.9}$$

Furthermore, the OLS estimate of the variance of the innovations, σ_u^2, is given by $s_n^2 = RSS/(n-m)$, which is unbiased, consistent, and satisfies

$$\sqrt{T}(s_n^2 - \sigma_u^2) \xrightarrow{L} N(0, \mu_4 - \sigma_u^4).$$

We now turn to properties of the MSFE under the assumption of a correctly specified model. For the scalar case, the OLS reduces to

$$\hat{\beta}_{t,n} = \beta + \Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-1} \sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau.$$

We calculate the large sample properties of the SFE. Substituting the OLS estimator in the expression for the SFE we obtain:

$$\begin{aligned} SFE_n =&(Y_{t+1} - \hat{Y}_{t+1})^2 = (\beta - \hat{\beta}_{t,n})^2 X_t^2 + 2(\beta - \hat{\beta}_{t,n}) X_t U_{t+1} + U_{t+1}^2 \\ =&\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-2}\Big(\sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau\Big)^2 X_t^2 \\ &- 2\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-1} \sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau X_t U_{t+1} + U_{t+1}^2. \end{aligned}$$

By theorem $A.13$ it follows

$$\frac{1}{n}\sum_{\tau=t-n}^{t-1} U_{\tau+1} X_\tau \xrightarrow{P} E[U_{t+1} X_t] = 0, \quad \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2 \xrightarrow{P} E[X_t^2]. \tag{5.2.10}$$

Multiplying and dividing the first term of the SFE by $1/n^2$, multiplying and dividing the second term of the SFE by $1/n$, applying (5.2.10) and theorem $A.18$ part 2, we obtain:

$$SFE_n \xrightarrow{P} U_{t+1}^2, \quad MSFE_n \xrightarrow{P} \sigma_u^2.$$

We can derive a simplified expression for the MSFE for the case where the elements of the explanatory process $\{X_\tau\}$ are mutually independent, i.e., $E[X_i X_j] = E[X_i]E[X_j]$ for $i \neq j$. First, we know the OLS is unbiased, $E[\hat{\beta}_{t,n}]$, and the expected value of the square of the OLS is given as follows:

$$E[\hat{\beta}_{t,n}^2] = \sigma_u^2 E\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-1}\Big].$$

Given these expressions for $E[\hat{\beta}_{t,n}]$ and $E[\hat{\beta}_{t,n}^2]$, the MSFE is as follows:

$$MSFE = \sigma_u^2 + \sigma_u^2 E\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-1}\Big].$$

This expression is simple, yet, the SSD is not transparent. Even under the assumption of correct specification, one can see the difficulty of determining the SSD of a statistic such as a MSFE which incorporates the OLS. In the next section, we present large sample results for the OLS under the assumption of misspecification.

5.3 Misspecification and the OLS

Much of what is known about estimation and inference relies on the assumption that the model in question coincides with the data generating process. For this reason, it is important to understand properties of commonly used estimators under the assumption of misspecification. The most important results in the literature regarding properties of the OLS when the regression model is misspecified were developed by White in [150]. These results are large sample properties of the OLS under functional misspecification. In this section, we present the assumptions and the main results of [150]. The first assumption describes the class of DGPs under consideration.

Assumption 5.5 *The true model is*

$$Y_\tau = g(Z_\tau) + \epsilon_\tau, \quad \tau = 1, \ldots, n,$$

where g is an unknown function and (Z_τ, ϵ_τ) are i.i.d. random $1 \times (p+1)$ vectors such that $E[Z_\tau] = 0$, $E[Z_\tau^\top Z_\tau] = M_{zz}$ is finite and nonsingular, $E[\epsilon_\tau] = 0$, $E[\epsilon_\tau^2] = \sigma_\epsilon^2 < \infty$, $E[Z_\tau^\top \epsilon_\tau] = 0$ and $E[g(Z_\tau)^2] = \sigma_g^2 < \infty$.

The linear model is of the form

$$Y_\tau = X_\tau \beta + u_\tau, \quad \tau = 1, \ldots, n,$$

where $u_\tau \equiv g(Z_\tau) - X_\tau\beta + \epsilon\tau$ is a random variable and the $1 \times k$ vector X_τ has elements which are functions of elements of Z_τ but some elements of Z_τ may be omitted.

Let $F_{z,\epsilon}$ denote the joint distribution of Z_τ, ϵ_τ. White writes the mean square error of approximation prediction as follows:

$$\sigma^2(\beta) = \int [g(z) - x\beta + \xi]^2 dF_{z,\epsilon}(x, \xi).$$

With the i.i.d assumption, this coincides with the definition of the MSFE at an arbitrary forecast origin. The OLS estimator is $\hat{\beta}_{OLS,n} = (X^\top X)^{-1} X^\top Y$, where X is the $n \times k$ matrix with rows X_τ. Further assumptions are as follows.

Assumption 5.6 *g and X are measurable functions of Z.*

Assumption 5.7 *$E[g(Z_\tau)\epsilon_\tau] = 0$, $E[X_\tau^\top] = 0$, $E[X_\tau^\top X_\tau] = M_{xx}$ is finite and nonsingular.*

β^* is defined as the parameter that uniquely solves the following optimization

$$\min_\beta \sigma^2(\beta). \tag{5.3.1}$$

The main result is given in the following theorem.

Theorem 5.1 (Theorem 2 in [150]) *Under assumptions* 5.5, 5.6, *and* 5.7, $\hat{\beta}_{OLS,n} \xrightarrow{a.s.} \beta^*$ *and* $s^2 \xrightarrow{a.s.} \sigma^2(\beta^*)$ *where* $s^2 = (n-k)^{-1} \sum_{\tau=1}^n (Y_\tau - X_\tau \hat{\beta}_{OLS,n})^2$.

If $g(z) = x\beta_0$, then $\beta^* = \beta_0$ for any distribution of the Z_τ otherwise, β^* depends crucially on the distribution of the Z_τ. As the sample size goes to infinity, $\hat{\beta}_{OLS,n}$ is approximately normally distributed, as shown in the next theorem.

Theorem 5.2 (Theorem 3 in [150]) *Under assumptions* 5.5, 5.6, *and* 5.7,

$$\sqrt{n}(\hat{\beta}_{OLS,n} - \beta^*) \overset{A}{\sim} N(0, M_{xx}^{-1} V(\beta^*) M_{xx}^{-1}),$$

provided $E[Y_i^2 X_i^\top X_i]$ *and* $E[X_{ij}^2 X_i^\top X_i]$, $j = 1, \ldots, k$ *are finite. Moreover,* $(X^\top X/n) \xrightarrow{a.s.} M_{xx}^{-1}$ *and*

$$\hat{V}_{OLS} = n^{-1} \sum_{\tau=1}^n (Y_\tau - X_\tau \hat{\beta}_{OLS,n})^2 X_\tau^\top X_\tau \overset{a.s.}{\to} V(\beta^*),$$

so that

$$(X^\top X/n)^{-1} \hat{V}_{OLS} (X^\top X/n)^{-1} \xrightarrow{a.s.} M_{xx}^{-1} V(\beta^*) M_{xx}^{-1}.$$

Theorem 5.1 is of great consequence for the problem of forecasting under misspecification. To see this, by its definition, β^* is the value attained by the linear parameter of the model which results in the smallest value of the MSFE, i.e., $\sigma^2(\beta^*)$. The objective of analyzing the SSD of the MSFE is to find the values of the sample size variable n for which the MSFE attains the value $\sigma^2(\beta^*)$. Theorem 5.1 describes the behavior of the MSFE as n goes to infinity by characterizing the behavior of $\hat{\beta}_{OLS,n}$ as n goes to infinity. Since $\hat{\beta}_{OLS,n}$ attains the value β^* at infinity, $\sigma^2(\beta)$ attains the value $\sigma^2(\beta^*)$ at infinity. Although theorem 5.1 is a good first start in understanding the SSD of the MSFE, the next issue one would like to address is the behavior of the MSFE for finite values of n. We would like to answer the question: Does there exist an $n^* < \infty$ so that $\sigma^2(\hat{\beta}_{OLS,n^*}) = \sigma^2(\beta^*)$. The main purpose of the work in this thesis is to understand the SSD for finite values of n.

In the next section, we develop an algorithm that can be used to construct an approximation of the MSFE in order to analyze the sample size dependence and determine the possible existence of optimal observation windows of finite length.

5.4 The algorithm: scalar case

As presented in chapter 2, the forecasting problem of interest consists of predicting the observed process $\{Y_\tau\}$ at $\tau = t+1$, $Y_{t+1} \in \mathbb{R}$, by means of a linear regression of the $k \times 1$ column vector X_t of $\mathcal{F}_t$-measurable variables. In this section we assume $k = 1$. The forecaster does not know the DGP which generates the series $\{Y_\tau\}$ and uses a linear model in X_t to approximate the conditional expectation $E_t[Y_{t+1}]$. The linear model used to forecast Y_{t+1} is of the form

$$Y_{t+1} = \beta X_t + V_{t+1}, \tag{5.4.1}$$

in which the parameter β, $\beta \in B$, B compact in $\mathbb{R}$, is estimated by OLS. The estimation sample contains the n most recent observations, $\{Y_{t-n+1}, \ldots, Y_t\}$ and $\{X_{t-n}, \ldots, X_{t-1}\}$, and the OLS estimator of β has the form

$$\hat{\beta}_{t,n} = \left(\sum_{\tau=t-n}^{t-1} X_\tau X_\tau^\top \right)^{-1} \left(\sum_{\tau=t-n}^{t-1} X_\tau Y_{\tau+1} \right). \tag{5.4.2}$$

The OLS estimator $\hat{\beta}_{t,n}$ is used to construct the forecast of Y_{t+1}, denoted $\hat{Y}_{t+1,n}$, given by

$$\hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t.$$

Using as cost function a squared loss function, the criterion which provides a measure of forecast accuracy is the MSFE given by

$$MSFE_n = E[(Y_{t+1} - \hat{Y}_{t+1,n})^2] = E[Y_{t+1}^2] - 2E[Y_{t+1}\hat{Y}_{t+1,n}] + E[\hat{Y}_{t+1,n}^2]. \tag{5.4.3}$$

The MSFE is the expected value of statistics which depend on the sample size parameter n. We construct a Taylor algorithm, as developed in Chapter 4, to approximate the MSFE in order to investigate the existence of an optimal observation window. The existence of such optimal observation window can be revealed by assessing the SSD of the MSFE. For this purpose, we begin the construction of the algorithm by focusing on the expectation of the following n-dependent terms

$$\Pi_{1,n} \equiv Y_{t+1}\hat{Y}_{t+1,n} = Y_{t+1} X_t \hat{\beta}_{t,n}, \tag{5.4.4}$$

$$\Pi_{2,n} \equiv \hat{Y}_{t+1,n}^2 = X_t^2 \hat{\beta}_{t,n}^2. \tag{5.4.5}$$

Substituting the scalar form of the OLS estimator $\hat{\beta}_{t,n}$, $\Pi_{1,n}$ and $\Pi_{2,n}$ become, respectively,

$$\Pi_{1,n} = Y_{t+1} X_t \Big(\sum_{s=t-n}^{t-1} X_s^2 \Big)^{-1} \sum_{s=t-n}^{t-1} Y_{s+1} X_s, \quad \Pi_{2,n} = \Big[\Big(\sum_{s=t-n}^{t-1} X_s^2 \Big)^{-1} X_t \sum_{s=t-n}^{t-1} Y_{s+1} X_s \Big]^2.$$

By defining the statistics $S_{1,n}$ and $S_{2,n}$ as follows:

$$S_{1,n} \equiv \frac{1}{n} \sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau, \quad S_{2,n} \equiv \frac{1}{n} \sum_{\tau=t-n}^{t-1} X_\tau^2,$$

the OLS estimator can be rewritten

$$\hat{\beta}_{t,n} = \frac{S_{1,n}}{S_{2,n}}, \tag{5.4.6}$$

and (5.4.4) and (5.4.5) become respectively

$$\Pi_{1,n} = Y_{t+1} X_t \frac{S_{1,n}}{S_{2,n}}, \quad \Pi_{2,n} = X_t^2 \left(\frac{S_{1,n}}{S_{2,n}} \right)^2 .$$

We assume the sequence of regressors $\{X_\tau\}$ to be independent and identically distributed. By independence, we can write

$$E[\Pi_{1,n}] = E[Y_{t+1} X_t] E[\hat{\beta}_{t,n}], \quad E[\Pi_{2,n}] = E[X_t^2] E[\hat{\beta}_{t,n}^2]. \tag{5.4.7}$$

We take a slight detour to explain a settle point involving (5.4.7). In an empirical situation, the independence assumption of the explanatory process $\{X_\tau\}$ can be tested. But (5.4.7) has the stronger implication that the random variable Y_{t+1} is independent of the random variables $\{X_{t-n}, \ldots, X_{t-1}\}$. In an empirical situation, this independence would have to be tested. The existence of such independence in the data would be the motivating force for constructing the forecasting model in the specification stage of the forecast methodology. In the case the independence between Y_{t+1} and $\{X_{t-n}, \ldots, X_{t-1}\}$ cannot be established, the algorithm would need to be modified. Chapter 6 develops a Taylor algorithm applicable for more general dependencies between the dependent and explanatory processes.

Continuing with our exposition, the next step in the construction of the algorithm is to apply the techniques of Chapter 4 to find approximations of $E[\hat{\beta}_{t,n}]$ and $E[\hat{\beta}_{t,n}^2]$. Such approximations are conducted by means of Taylor series expansions of $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$ about some points ω_1 and ω_2 respectively. From the theory developed in Chapter 4, we learned that approximating the expectation of a function of random variables by means of Taylor series requires one, in many instances, to approximate the expectation by a truncated expectation. Using truncated expectations is necessary because Taylor series approximations are valid only within the region of convergence and, at the same time, the random variables involved take values on a specific range. In the case of $\hat{\beta}_{t,n}$ and $\hat{\beta}^2_{t,n}$, the approximations will depend on truncated central moments of $S_{1,n}$ and $S_{2,n}$. Let $\mathcal{A}$ be a set inside the region of convergence $\mathcal{B}$ of the Taylor series of $\hat{\beta}_{t,n}$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$. Appendix C.1.1 provides details on the nature of the region of convergence of the Taylor series expansion of the OLS and on the nature of convergence sets such as $\mathcal{A}$. We write the expectation of the OLS estimator and its square as follows

$$E[\hat{\beta}_{t,n}] = \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}] + \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}^c], \quad E[\hat{\beta}^2_{t,n}] = \bar{E}[\hat{\beta}^2_{t,n}, \mathcal{A}] + \bar{E}[\hat{\beta}^2_{t,n}, \mathcal{A}^c], \tag{5.4.8}$$

where $\mathcal{A}^c$ is the complement of $\mathcal{A}$. Taylor series can be used within $\mathcal{A}$ to approximate $\hat{\beta}_{t,n}$ and $\hat{\beta}^2_{t,n}$. To obtain further analytic results, we assume $P(X \in \mathcal{A}) \approx 1$ so that $E[\hat{\beta}_{t,n}] \approx \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}]$ and $E[\hat{\beta}^2_{t,n}] \approx \bar{E}[\hat{\beta}^2_{t,n}, \mathcal{A}]$. We define the points about which to calculate the Taylor series as follows:

$$\omega_1 \equiv E[S_{1,n}] = E[Y_{t+1}X_t], \quad \omega_2 \equiv E[S_{2,n}] = E[X_t^2],$$

where the equalities follow from the i.i.d. assumption. The fourth order Taylor polynomial of $\hat{\beta}_{t,n}$ about the points ω_1 and ω_2 is as follows:

$$\begin{aligned}
Q(\hat{\beta}_{t,n}, 4) =& \frac{\omega_1}{\omega_2} + \frac{1}{\omega_2}(S_{1,n} - \omega_1) - \frac{\omega_1}{\omega_2^2}(S_{2,n} - \omega_2) - \frac{1}{\omega_2^2}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2) \\
&+ \frac{\omega_1}{\omega_2^3}(S_{2,n} - \omega_2)^2 + \frac{1}{\omega_2^3}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^2 - \frac{\omega_1}{\omega_2^4}(S_{2,n} - \omega_2)^3 \\
&+ \frac{\omega_1}{\omega_2^5}(S_{2,n} - \omega_2)^4 - \frac{1}{\omega_2^4}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3.
\end{aligned}$$

The fourth order Taylor polynomial of $\hat{\beta}^2_{t,n}$ about the points ω_1 and ω_2 is as follows:

$$\begin{aligned}
Q(\hat{\beta}^2_{t,n}, 4) =& \frac{\omega_1^2}{\omega_2^2} + 2\frac{\omega_1}{\omega_2^2}(S_{1,n} - \omega_1) - 2\frac{\omega_1^2}{\omega_2^3}(S_{2,n} - \omega_2) + \frac{1}{\omega_2^2}(S_{1,n} - \omega_1)^2 \\
&- 4\frac{\omega_1}{\omega_2^3}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2) + 3\frac{\omega_1^2}{\omega_2^4}(S_{2,n} - \omega_2)^2 \\
&- 2\frac{1}{\omega_2^3}(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2) + 6\frac{\omega_1}{\omega_2^4}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^2 \\
&- 4\frac{\omega_1^2}{\omega_2^5}(S_{2,n} - \omega_2)^3 + \frac{3}{\omega_2^4}(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^2
\end{aligned}$$

$$- 8\frac{\omega_1}{\omega_2^5}(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3 + 5\frac{\omega_1^2}{\omega_2^6}(S_{2,n} - \omega_2)^4.$$

We take expectations of the fourth order polynomials to obtain the approximations

$$E[\hat{\beta}_{t,n}] \approx \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}] \approx E[Q(\hat{\beta}_{t,n}, 4)],$$

$$E[\hat{\beta}^2_{t,n}] \approx \bar{E}[\hat{\beta}^2_{t,n}, \mathcal{A}] \approx E[Q(\hat{\beta}^2_{t,n}, 4)].$$

Using these approximations, the MSFE approximation becomes

$$MSFE_n \approx E[Y_{t+1}^2] - 2E[Y_{t+1}X_t]E[Q(\hat{\beta}_{t,n}, 4)] + E[X_t^2]E[Q(\hat{\beta}^2_{t,n}, 4)]. \tag{5.4.9}$$

The central moments involved in the expectation of the Taylor polynomials are expanded and simplified to derive the SSD in terms of the sample size variable n. Appendix C, Section $C.2$, presents the derivation of the central moments for the general case without assuming $P(X \in \mathcal{A}) \approx 1$. With $P(X \in \mathcal{A}) \approx 1$, the expectation of the term $(S_{1,n} - \omega_1)$ is as follows:

$$E[(S_{1,n} - \omega_1)] = \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[Y_{\tau+1}X_\tau] - \omega_1 = E[Y_t X_{t-1}] - \omega_1 = 0,$$

where the second equality follows from the i.i.d assumptions. We write the rest of the central moments involved in the expectation of $Q(\hat{\beta}_{t,n}, 4)$ and $Q(\hat{\beta}^2_{t,n}, 4)$ under the i.i.d and $P(X \in \mathcal{A}) \approx 1$ assumptions:

$$\begin{aligned}
&E[(S_{2,n} - \omega_2)] = E[X_{t-1}^2] - \omega_2 = 0,\\
&E[(S_{1,n} - \omega_1)^2] = \frac{1}{n}\Big[E[Y_t^2 X_{t-1}^2] - E^2[Y_t X_{t-1}]\Big] = \frac{1}{n}\mathrm{Var}(Y_t X_{t-1}),\\
&E[(S_{2,n} - \omega_2)^2] = \frac{1}{n}\Big[E[X_{t-1}^4] - E^2[X_{t-1}^2]\Big] = \frac{1}{n}\mathrm{Var}(X_t^2),\\
&E[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)] = \frac{1}{n}\Big[E[Y_t X_{t-1}^3] - E[Y_t X_{t-1}]E[X_{t-1}^2]\Big] = \frac{1}{n}\mathrm{Cov}(Y_t X_{t-1}, X_{t-1}^2),\\
&E[(S_{2,n} - \omega_2)^3] = \frac{1}{n^2}\Big[E[X_{t-1}^6] - 3E[X_{t-1}^4]E[X_{t-1}^2] + 2E^3[X_{t-1}^2]\Big],\\
&E[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^2] = \frac{1}{n^2}\Big[E[Y_t X_{t-1}^5] - E[Y_t X_{t-1}]E[X_{t-1}^4]\\
&\qquad - 2E[Y_t X_{t-1}^3]E[X_{t-1}^2] + 2E[Y_t X_{t-1}]E^2[X_{t-1}^2]\Big],
\end{aligned}$$

$$
\begin{aligned}
E[(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2)] = \frac{1}{n^2}\Big[& E[Y_t^2X_{t-1}^4] - E[Y_t^2X_{t-1}^2]E[X_{t-1}^2] \\
& - 2E[Y_tX_{t-1}]E[Y_tX_{t-1}^3] + 2E^2[Y_tX_{t-1}]E[X_{t-1}^2]\Big], \\
E[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)^3] = \frac{3}{n^2}\Big[& E[Y_tX_{t-1}^3]E[X_{t-1}^4] - E[Y_tX_{t-1}]E[X_{t-1}^2]E[X_{t-1}^4] \\
& - E^2[X_{t-1}^2]E[Y_tX_{t-1}^3] + E[Y_tX_{t-1}]E^3[X_{t-1}^2]\Big] \\
& + \frac{1}{n^3}\Big[E[Y_tX_{t-1}^7] - E[Y_tX_{t-1}]E[X_{t-1}^6] - 3E[Y_tX_{t-1}^5]E[X_{t-1}^2] \\
& - 3E[Y_tX_{t-1}^3]E[X_{t-1}^4] + 6E[Y_tX_{t-1}]E[X_{t-1}^2]E[X_{t-1}^4] \\
& + 6E[Y_tX_{t-1}^3]E^2[X_{t-1}^2] - 6E[Y_tX_{t-1}]E^3[X_{t-1}^2]\Big], \\
E[(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2)^2] = \frac{1}{n^2}\Big[& E[Y_t^2X_{t-1}^2]E[X_{t-1}^4] - E[Y_t^2X_{t-1}^2]E^2[X_{t-1}^2] \\
& - E^2[Y_tX_{t-1}]E[X_{t-1}^4] - 4E[Y_tX_{t-1}]E[X_{t-1}^2]E[Y_tX_{t-1}^3] \\
& + 2E^2[Y_tX_{t-1}^3] + 3E^2[Y_tX_{t-1}]E^2[X_{t-1}^2]\Big] \\
& + \frac{1}{n^3}\Big[E[Y_t^2X_{t-1}^6] - E[Y_t^2X_{t-1}^2]E[X_{t-1}^4] - 2E[Y_t^2X_{t-1}^4]E[X_{t-1}^2] \\
& + 2E[Y_t^2X_{t-1}^2]E^2[X_{t-1}^2] - 2E[Y_tX_{t-1}^5]E[Y_tX_{t-1}] \\
& + 2E^2[Y_tX_{t-1}]E[X_{t-1}^4] + 8E[Y_tX_{t-1}^3]E[Y_tX_{t-1}]E[X_{t-1}^2] \\
& - 2E^2[Y_tX_{t-1}^3] - 6E^2[Y_tX_{t-1}]E^2[X_{t-1}^2]\Big], \\
E[(S_{2,n}-\omega_2)^4] = \frac{3}{n^2}\Big[& E^2[X_{t-1}^4] - 2E[X_{t-1}^4]E^2[X_{t-1}^2] + E^4[X_{t-1}^2]\Big] \\
& + \frac{1}{n^3}\Big[E[X_{t-1}^8] - 4E[X_{t-1}^6]E[X_{t-1}^2] - 3E^2[X_{t-1}^4] + 12E[X_{t-1}^4]E^2[X_{t-1}^2] \\
& - 6E^4[X_{t-1}^2]\Big].
\end{aligned}
$$

These central moments can be derived from the general central moments given in $C.2$ by replacing truncated expectations with expectations and simplifying. Substituting the above central moments in the expression for $E[Q(\hat{\beta}_{t,n},4)]$, one obtains

$$
\begin{aligned}
E[Q(\hat{\beta}_{t,n},4)] = & \frac{\omega_1}{\omega_2} + \frac{1}{n}\left[\frac{\omega_1}{\omega_2^3}E[X_{t-1}^4] - \frac{1}{\omega_2^2}E[Y_tX_{t-1}^3]\right] \\
& + \frac{1}{n^2}\left[-\frac{\omega_1}{\omega_2^4}E[X_{t-1}^6] - \frac{\omega_1}{\omega_2^3}E[X_{t-1}^4] + \frac{1}{\omega_2^3}E[Y_tX_{t-1}^5] - 2\frac{1}{\omega_2^2}E[Y_tX_{t-1}^3]\right. \\
& \left. + 3\frac{\omega_1}{\omega_2^5}E^2[X_{t-1}^4] - \frac{3}{\omega_2^4}E[Y_tX_{t-1}^3]E[X_{t-1}^4] + \frac{3}{\omega_2^2}E[Y_tX_{t-1}^3]\right]
\end{aligned}
$$

$$+ \frac{1}{n^3}\left[\frac{\omega_1}{\omega_2^5}E[X_{t-1}^8] - 3\frac{\omega_1}{\omega_2^4}E[X_{t-1}^6] - 3\frac{\omega_1}{\omega_2^5}E^2[X_{t-1}^4] + 6\frac{\omega_1}{\omega_2^3}E[X_{t-1}^4]\right.$$
$$- \frac{1}{\omega_2^4}E[Y_tX_{t-1}^7] + \frac{3}{\omega_2^3}E[Y_tX_{t-1}^5] + \frac{3}{\omega_2^4}E[Y_tX_{t-1}^3]E[X_{t-1}^4]$$
$$\left.- \frac{6}{\omega_2^2}E[Y_tX_{t-1}^3]\right].$$

Similarly, substituting the above central moments in the expression for $E[Q(\hat{\beta}_{t,n}^2, 4)]$, one obtains

$$\begin{aligned}
E[Q(\hat{\beta}_{t,n}^2, 4)] =& \frac{\omega_1^2}{\omega_2^2} + \frac{1}{n}\left[3\frac{\omega_1^2}{\omega_2^4}E[X_{t-1}^4] - 4\frac{\omega_1}{\omega_2^3}E[Y_tX_{t-1}^3] + \frac{1}{\omega_2^2}E[Y_t^2X_{t-1}^2]\right] \\
&+ \frac{1}{n^2}\left[-4\frac{\omega_1^2}{\omega_2^5}E[X_{t-1}^6] - 3\frac{\omega_1^2}{\omega_2^4}E[X_{t-1}^4] + 6\frac{\omega_1}{\omega_2^4}E[Y_tX_{t-1}^5] + 4\frac{\omega_1}{\omega_2^3}E[Y_tX_{t-1}^3]\right. \\
&- \frac{2}{\omega_2^3}E[Y_t^2X_{t-1}^4] - \frac{1}{\omega_2^2}E[Y_t^2X_{t-1}^2] + 15\frac{\omega_1^2}{\omega_2^6}E^2[X_{t-1}^4] \\
&\left.- 24\frac{\omega_1}{\omega_2^5}E[Y_tX_{t-1}^3]E[X_{t-1}^4] + \frac{3}{\omega_2^4}E[Y_t^2X_{t-1}^2]E[X_{t-1}^4] + \frac{6}{\omega_2^4}E^2[Y_tX_{t-1}^3]\right] \\
&+ \frac{1}{n^3}\left[5\frac{\omega_1^2}{\omega_2^6}E[X_{t-1}^8] - 12\frac{\omega_1^2}{\omega_2^5}E[X_{t-1}^6] - 15\frac{\omega_1^2}{\omega_2^6}E^2[X_{t-1}^4] + 18\frac{\omega_1^2}{\omega_2^4}E[X_{t-1}^4]\right. \\
&- 8\frac{\omega_1}{\omega_2^5}E[Y_tX_{t-1}^7] + 18\frac{\omega_1}{\omega_2^4}E[Y_tX_{t-1}^5] + 24\frac{\omega_1}{\omega_2^5}E[Y_tX_{t-1}^3]E[X_{t-1}^4] \\
&- 24\frac{\omega_1}{\omega_2^3}E[Y_tX_{t-1}^3] + \frac{3}{\omega_2^4}E[Y_t^2X_{t-1}^6] - \frac{3}{\omega_2^4}E[Y_t^2X_{t-1}^2]E[X_{t-1}^4] \\
&\left.- \frac{6}{\omega_2^3}E[Y_t^2X_{t-1}^4] + \frac{6}{\omega_2^2}E[Y_t^2X_{t-1}^2] - \frac{6}{\omega_2^4}E^2[Y_tX_{t-1}^3]\right].
\end{aligned}$$

The construction of the algorithm is completed by substituting the expressions for $E[Q(\hat{\beta}_{t,n}, 4)]$ and $E[Q(\hat{\beta}_{t,n}^2, 4)]$ in the MSFE approximation (5.4.9). The approximation of the MSFE is as follows:

$$\begin{aligned}
MSFE_n &\approx E[Y_{t+1}^2] - 2\omega_1 E[Q(\hat{\beta}_{t,n}, 4)] + \omega_2 E[Q(\hat{\beta}_{t,n}^2, 4)] \\
&= \frac{1}{\omega_2^5}\left[C + \frac{A}{n} - \frac{\Delta}{n^2} + \frac{\Omega}{n^3}\right]
\end{aligned} \tag{5.4.10}$$

with $\Delta = A + 2B - D$, $\Omega = 6A - 6B - D + E$,

$$A = \omega_1^2\omega_2^2 E[X_{t-1}^4] - 2\omega_1\omega_2^3 E[Y_tX_{t-1}^3] + \omega_2^4 E[Y_t^2X_{t-1}^2],$$

$$B = \omega_1^2\omega_2 E[X_{t-1}^6] - 2\omega_1\omega_2^2 E[Y_t X_{t-1}^5] + \omega_2^3 E[Y_t^2 X_{t-1}^4],$$

$$C = E[Y_{t+1}^2]\omega_2^5 - \omega_1^2\omega_2^4,$$

$$D = 9\omega_1^2 E^2[X_{t-1}^4] - 18\omega_1\omega_2 E[Y_t X_{t-1}^3]E[X_{t-1}^4] + 3\omega_2^2 E[Y_t^2 X_{t-1}^2]E[X_{t-1}^4] + 6\omega_2^2 E^2[Y_t X_{t-1}^3],$$

$$E = 3\omega_1^2 E[X_{t-1}^8] - 6\omega_1\omega_2 E[Y_t X_{t-1}^7] + 3\omega_2^2 E[Y_t^2 X_{t-1}^6].$$

The fourth order MSFE approximation given in (5.4.10) depends on the sample size n up to a cubic term $1/n^3$. It can be shown the central moments $E[(S_{2,n} - \omega_2)^5]$, $E[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^4]$, and $E[(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^3]$ involved in the fifth order term of the Taylor series of $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$ do not alter the constant term, the $1/n$ term or the $1/n^2$ term of the fourth order MSFE approximation. In fact, the fifth order terms of the Taylor series approximation of $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$ only contribute a $1/n^3$ term and a $1/n^4$ term. Although the Taylor series approximation of the MSFE can be found up to any order required, further analytic results can be obtained by focusing on the MSFE approximation up to quadratic terms given by

$$MSFE_n \approx \frac{1}{\omega_2^5}\left[C + \frac{A}{n} - \frac{\Delta}{n^2}\right] \equiv \overline{MSFE}_n. \tag{5.4.11}$$

To determine the existence of an optimal observation window, we examine the solution to the following optimization problem

$$\min_n \left\{C + \frac{A}{n} - \frac{\Delta}{n^2}\right\}.$$

The extremum of the MSFE approximation (5.4.11) is given by

$$n_o = 2\frac{\Delta}{A}. \tag{5.4.12}$$

By analyzing this extremum, we can determine an approximation for the optimal observation window. Let $\bar{n}^*$ denote the size of the observation window which minimizes the MSFE approximation $\overline{MSFE}_n$. $\bar{n}^*$ is the approximation to the optimal observation window n^* which minimizes the true MSFE. Since in most practical applications the amount of data available is finite, we denote by $\bar{n}$ the size of the largest data window available for forecasting and estimation. To understand the SSD of the MSFE, we determine some properties of the MSFE approximation, $\overline{MSFE}_n$. First, the limit of $\overline{MSFE}_n$ as $n \to \infty$ is given by C/ω_2^4 and $C > 0$.

Proposition 5.3 $C \geq 0$.

Proof. See Appendix C.3. ■

Define $\underline{n} = \Delta\bar{n}/(A\bar{n} - \Delta)$. The main conclusion about the existence of an optimal observation window when the processes in question are i.i.d is summarized in the following proposition and its proof presents the analysis of the SSD of $\overline{MSFE}_n$.

Proposition 5.4 *If $\{X_s\}$ and $\{Y_s\}$ are i.i.d. processes and $\underline{n} < 1$, then $\bar{n}^* = \bar{n}$.*

Proof. First we rewrite A as follows

$$A = \omega_2 E\Big[(\omega_1 X_t^2 - \omega_2 Y_{t+1} X_t)^2\Big].$$

Since $\omega_2 > 0$, it follows $A > 0$. The partial derivative of the MSFE approximation (5.4.11) with respect to n is

$$\frac{\partial}{\partial n}\overline{MSFE}_n = \frac{1}{\omega_2^5}\left[\frac{-A}{n^2} + 2\frac{\Delta}{n^3}\right], \tag{5.4.13}$$

and the extremum is given by (5.4.12). We analyze the two cases $n_o \leq 0$ and $n_o > 0$.

Case $n_o \leq 0$:

Since the size of the forecasting window must be a positive integer, $n_o \leq 0$ is not a solution to the optimal forecasting window problem. Nonetheless, we examine the behavior of $\overline{MSFE}_n$ for positive values of n when $n_o \leq 0$. From the expression for n_o, $n_o \leq 0$ if and only if $\Delta \leq 0$ and, as a consequence, $\overline{MSFE}_n \to +\infty$ as $n_+ \to 0$, $\overline{MSFE}_n \to C/\omega_2^4$ as $n \to \infty$, and by (5.4.13) $\frac{\partial}{\partial n}\overline{MSFE}_n < 0$ whenever $n > 0$. Therefore, $\overline{MSFE}_n$ decreases monotonically as $n \to \infty$ suggesting it is optimal to use all available data to estimate $\hat{\beta}_{t,n}$ and obtain the smallest value of the MSFE.

Case $n_o > 0$:

First, note $n_o > 0$ if and only if $\Delta > 0$. To determine if n_o is a minimum or a maximum of $\overline{MSFE}_n$, we write the second partial derivative of $\overline{MSFE}_n$ with respect to

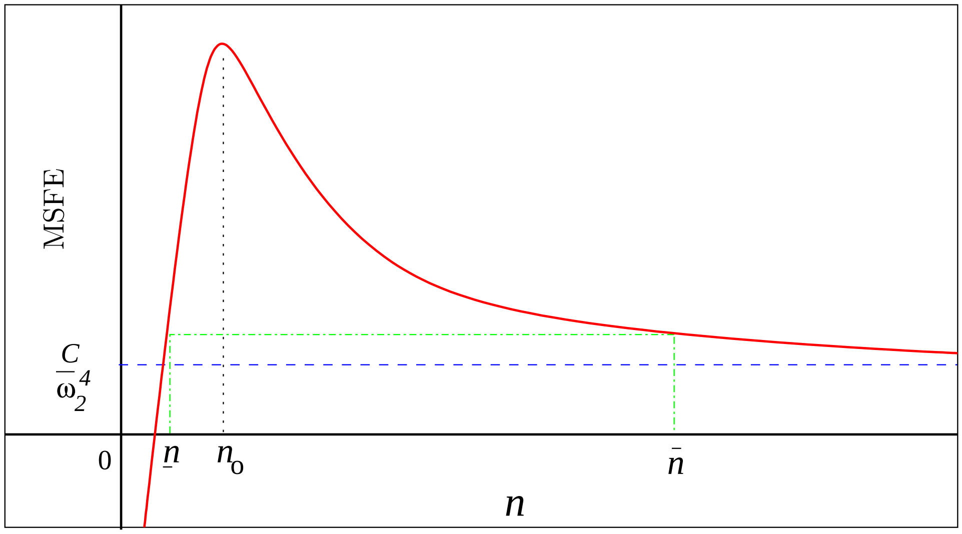


Figure 5.1: MSFE approximation for $n_o > 0$

n

$$\frac{\partial^2 \overline{MSFE}_n}{\partial n^2} = \frac{1}{\omega_2^5}\left[2\frac{A}{n^3} - 6\frac{\Delta}{n^4}\right]. \tag{5.4.14}$$

Substituting (5.4.12) in (5.4.14) leads to

$$\left.\frac{\partial^2 \overline{MSFE}_n}{\partial n^2}\right|_{n^*} = -\frac{1}{8\omega_2^5}\frac{A^4}{\Delta^3} < 0,$$

and it follows n_o is a maximum of $\overline{MSFE}_n$ whenever $n_o > 0$ and therefore $\bar{n}^* \neq n_o$. $\underline{n}$ is defined as the value of n, which is less $\bar{n}$, at which $\overline{MSFE}_n$ has the same value as at $\bar{n}$. The general shape of $\overline{MSFE}_n$ for $n_o > 0$ is illustrated in figure 5.1. Since $\bar{n}^*$ must be a positive integer, the result follows, $\bar{n}^* = \bar{n}$ when $\underline{n} < 1$. ■

As noted earlier, $\overline{MSFE}_n$ is an approximation of the MSFE truncated at the $1/n^2$ term. $\overline{MSFE}_n$ was implemented to provide further analytic results. Nonetheless, the approximation (5.4.10) can be used to graphically analyze the SSD of the MSFE, given the necessary moments.

5.5 The algorithm: multi-variate case

We now construct the approximation of the MSFE for the multi-variate case with $k = m$. As before, we denote by X_t a $m \times 1$ column vector, $X_t = (X_t^1, \ldots, X_t^m)^\top$, of $\mathcal{F}_t$-measurable variables that are used to forecast $Y_{t+1} \in \mathbb{R}$. As in the scalar case, using as cost function the squared loss function, the criterion which provides a measure of forecast accuracy is the MSFE given by

$$MSFE_n = E[(Y_{t+1} - \hat{Y}_{t+1,n})^2] = E[Y_{t+1}^2] - 2E[Y_{t+1}\hat{Y}_{t+1,n}] + E[\hat{Y}_{t+1,n}^2]. \tag{5.5.1}$$

The MSFE is the expected value of statistics, $\hat{Y}_{t+1,n}$ and $\hat{Y}_{t+1,n}^2$, which depend on the parameter n. To begin the construction of the algorithm, and since we are interested in the SSD of the MSFE, we restrict attention to the expectations of the following terms

$$\Pi_{1,n} \equiv Y_{t+1}\hat{Y}_{t+1,n} = Y_{t+1}X_t^\top \hat{\beta}_{t,n}, \tag{5.5.2}$$

$$\Pi_{2,n} \equiv \hat{Y}_{t+1,n}^2 = (\hat{\beta}_{t,n}^\top X_t)^2 = X_t^\top \hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top X_t. \tag{5.5.3}$$

Substituting the vector form of the OLS estimator $\hat{\beta}_{t,n}$, $\Pi_{1,n}$ and $\Pi_{2,n}$ become, respectively,

$$\Pi_{1,n} = Y_{t+1}X_t^\top \left(X_{t,n}^\top X_{t,n}\right)^{-1} X_{t,n}^\top Y_{t,n},$$
$$\Pi_{2,n} = X_t^\top (X_{t,n}^\top X_{t,n})^{-1} X_{t,n}^\top Y_{t,n} Y_{t,n}^\top X_{t,n} (X_{t,n}^\top X_{t,n})^{-1} X_t.$$

By defining the statistics $S_{1,n}$ and $S_{2,n}$ as follows

$$S_{1,n} \equiv \frac{1}{n} X_{t,n}^\top Y_{t,n} \in \mathbb{R}^{m\times 1}, \quad S_{2,n} \equiv \frac{1}{n} X_{t,n}^\top X_{t,n} \in \mathbb{R}^{m\times m},$$

we rewrite the OLS estimator

$$\hat{\beta}_{t,n} = S_{2,n}^{-1} S_{1,n} \in \mathbb{R}^{m\times 1}, \tag{5.5.4}$$

and (5.5.2) and (5.5.3) become respectively

$$\Pi_{1,n} = Y_{t+1}X_t^\top \, S_{2,n}^{-1}\, S_{1,n}\,, \quad \Pi_{2,n} = X_t^\top S_{2,n}^{-1}\, S_{1,n}^\top\, S_{1,n}\, S_{2,n}^{-1} X_t.$$

We assume X_t and X_s are independent for all $t \neq s$ and X_t has the same distribution for all t. Furthermore, for each t, X_t^i and X_t^j are independent for all $i \neq j$. By the independence assumptions, we can write

$$
\begin{aligned}
E[\Pi_{1,n}] &= E[Y_{t+1}X_t^\top]E[\hat{\beta}_{t,n}] = E[Y_{t+1}X_t^\top]E[S_{2,n}^{-1}\, S_{1,n}], \\
E[\Pi_{2,n}] &= E[(\text{cs }(X_tX_t^\top))^\top]E[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)] \\
&= E[(\text{cs }(X_tX_t^\top))^\top]E[\text{cs }(S_{2,n}^{-1}\, S_{1,n}^\top\, S_{1,n}\, S_{2,n}^{-1})],
\end{aligned}
$$

where cs stands for column string. The next step in the construction of the algorithm is to apply the techniques of Chapter 4 to find approximations of $E[\hat{\beta}_{t,n}]$ and $E[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)]$. Such approximations are conducted by means of Taylor series expansions of $\hat{\beta}_{t,n}$ and cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$ about points ω_1 and ω_2, respectively. From the theory developed in Chapter 4, and as in the scalar case, approximating the expectation of a function of random variables by means of Taylor series requires one, in many instances, to approximate the expectation by a truncated expectation. In the case of $\hat{\beta}_{t,n}$ and cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$, the approximations will depend on truncated central moments of $S_{1,n}$ and $S_{2,n}$. We write the expectation of the OLS estimator $\hat{\beta}_{t,n}$ and cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$ as follows

$$
\begin{aligned}
E[\hat{\beta}_{t,n}] &= \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}] + \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}^c], \\
E[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)] &= \bar{E}[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), \mathcal{A}] + \bar{E}[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), \mathcal{A}^c],
\end{aligned}
$$

where $\mathcal{A}$ is a region where Taylor series can be used to approximate $\hat{\beta}_{t,n}$ and cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$. To obtain further analytic results, we assume $P(X \in \mathcal{A}) \approx 1$ so that $E[\hat{\beta}_{t,n}] \approx \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}]$ and $E[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)] \approx \bar{E}[\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), \mathcal{A}]$. We define

$$
\omega_1 \equiv E[S_{1,n}] = E[X_tY_{t+1}] \in \mathbb{R}^{m\times 1}, \quad \omega_2 \equiv E[S_{2,n}] = E[X_tX_t^\top] \in \mathbb{R}^{m\times m},
$$

where the expectation of a matrix is equal to the matrix of the expectation of the elements and, similarly, for the expectation of a vector. We write matrix Taylor polynomials, applying notation, given in Appendix $C.4$, for derivatives of matrix valued functions of matrices with respect to matrices and vectors as defined in [144]. We define the vector b_n by stacking the $m \times 1$ vector $S_{1,n}$ and the column string of the $m \times m$ matrix $S_{2,n}$, and similarly define the vector $\bar{b}$ by stacking the $m \times 1$ vector ω_1 and the column string of the $m \times m$ matrix ω_2.

$$b_n \equiv \begin{pmatrix} S_{1,n} \\ cs\ S_{2,n} \end{pmatrix} \in \mathbb{R}^{m(1+m)\times 1}, \quad \bar{b} \equiv \begin{pmatrix} \omega_1 \\ cs\ \omega_2 \end{pmatrix} \in \mathbb{R}^{m(1+m)\times 1}.$$

The Mth order Taylor polynomial approximating $\hat{\beta}_{t,n}$, with respect to b_n and about the point $\bar{b}$, is as follows:

$$Q(\hat{\beta}_{t,n}, M) = \omega_2^{-1}\omega_1 + \sum_{i=1}^{M} \frac{1}{i!}\Big(\mathcal{D}^i_{b_n^{\top i}}\ \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}}\Big((b_n - \bar{b})^{\otimes i} \otimes I\Big).$$

The Mth order Taylor polynomial approximating cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})$ about the point $\bar{b}_2$ is as follows:

$$Q(\text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}), M) = \text{cs } (\omega_2^{-1}\omega_1^{\top}\omega_1\omega_2^{-1}) + \sum_{i=1}^{M} \frac{1}{i!}\Big(\mathcal{D}^i_{b_n^{\top i}}\ \text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})\Big)_{b_n=\bar{b}}\Big((b_n - \bar{b})^{\otimes i} \otimes I\Big),$$

where $\omega_2^{\top} = \omega_2$. We take expectations of the Mth order polynomials to obtain the approximations

$$\begin{aligned} &E[\hat{\beta}_{t,n}] \approx \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}] \approx E[Q(\hat{\beta}_{t,n}, M)], \\ &E[\text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})] \approx \bar{E}[\text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}), \mathcal{A}] \approx E[Q(\text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}), M)]. \end{aligned}$$

The expectations of the Mth order Taylor polynomials of $\hat{\beta}_{t,n}$ and cs $\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}$ are respectively

$$E[Q(\hat{\beta}_{t,n}, M)] = \omega_2^{-1}\omega_1 + \sum_{i=2}^{M} \frac{1}{i!}\Big(\mathcal{D}^i_{b_n^{\top i}}\ \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}} E\Big[(b_n - \bar{b})^{\otimes i}\Big] \otimes I, \tag{5.5.5}$$

$$\begin{aligned} E[Q(\text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}), M)] = {} & \text{cs } (\omega_2^{-1}\omega_1^{\top}\omega_1\omega_2^{-1}) \\ & + \sum_{i=1}^{M} \frac{1}{i!}\Big(\mathcal{D}^i_{b_n^{\top i}}\ \text{cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})\Big)_{b_n=\bar{b}} E\Big[(b_n - \bar{b})^{\otimes i}\Big] \otimes I. \end{aligned} \tag{5.5.6}$$

In this exposition, we use fourth order polynomials $M = 4$. We take expectations of the fourth order polynomials to obtain the approximations

$$E[\hat{\beta}_{t,n}] \approx E[Q(\hat{\beta}_{t,n}, 4)], \quad E[\mathrm{cs}(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})] \approx E[Q(\hat{\beta}_{t,n}^{2}, 4)].$$

Using these approximations, the MSFE approximation becomes

$$MSFE_n \approx E[Y_{t+1}^2] - 2E[Y_{t+1}X_t^{\top}]E[Q(\hat{\beta}_{t,n}, 4)] + E[(\mathrm{cs}(X_t X_t^{\top}))^{\top}]E[Q(\mathrm{cs}(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top}), 4)]. \tag{5.5.7}$$

In order to analyze the SSD of the MSFE approximation, we are interested in the n dependence of (5.5.5) and (5.5.6). First, we note, as before, ω_1 and ω_2 are n-independent. Next we examine the derivative terms

$$\left(\mathcal{D}^{i}_{b_n^{\top i}}\, \hat{\beta}_{t,n}\right)_{b,n=\bar{b}}, \quad \left(\mathcal{D}^{i}_{b_n^{\top i}}\, \mathrm{cs}\ (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})\right)_{b_n=\bar{b}}. \tag{5.5.8}$$

The n dependence of $\hat{\beta}_{t,n}$ and cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})$ occurs through the statistics $S_{1,n}$ and $S_{2,n}$. $S_{1,n}$ is a $m \times 1$ vector with terms of the form

$$S_{1i,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_{\tau}^{i} Y_{\tau+1}, \quad i = 1, \ldots, m.$$

$S_{2,n}$ is a $m \times m$ matrix with terms of the form

$$S_{2ij,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_{\tau}^{i} X_{\tau}^{j}, \quad i = 1, \ldots, m,\ j = 1, \ldots, m.$$

By the definition of the vectors b_n and $\bar{b}$, $\hat{\beta}_{t,n}$ evaluated at $b_n = \bar{b}$ is n-independent. Similarly, cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^{\top})$ evaluated at $b_n = \bar{b}$ is n-independent. This is clear from the zeroth order terms of the Taylor expansions (5.5.5) and (5.5.6). The first derivative of $\hat{\beta}_{t,n}$ with respect to $b_n^{\top}$ is given by

$$\mathcal{D}_{b_n^{\top}}\, \hat{\beta}_{t,n} = \left(\mathcal{D}_{b_{1,n}}\, \hat{\beta}_{t,n} \quad \mathcal{D}_{b_{2,n}}\, \hat{\beta}_{t,n} \cdots \mathcal{D}_{b_{m(1+m),n}}\, \hat{\beta}_{t,n}\right) \in \mathbb{R}^{m \times m(1+m)},$$

where $b_n^\top = (b_{1,n} \cdots b_{m(1+m),n})$. Each of the m elements of $\hat{\beta}_{t,n}$ is a rational function of the elements $S_{1i,n}$ $i = 1, \ldots, m$ and the elements $S_{2ij,n}$ $i = 1, \ldots, m,\ j = 1, \ldots, m$ and the only n dependence is through these $m(1+m)$ terms. Consequently, each of the m elements of $\mathcal{D}_{b_{h,n}}\ \hat{\beta}_{t,n}$, by the definition of b_n, is a rational function of the elements $S_{1i,n}$ $i = 1, \ldots, m$ and the elements $S_{2ij,n}$ $i = 1, \ldots, m,\ j = 1, \ldots, m$ for $h = 1, \ldots, m(m+1)$. The n-dependence of $\mathcal{D}_{b_{h,n}}\ \hat{\beta}_{t,n}$ is also only through the $m(1+m)$ terms $S_{1i,n}$ and $S_{2ij,n}$. When evaluated at $b_n = \bar{b}$, $\mathcal{D}_{b_n^\top}\ \hat{\beta}_{t,n}$ is n-independent. The same arguments are true for any derivative of $\hat{\beta}_{t,n}$ with respect to $b_n^\top$ and for any derivative of cs $(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$ with respect to $b_n^\top$.

Proposition 5.5 *Both expressions in* (5.5.8) *are n-independent for $i = 0, 1, \ldots$.*

By the proposition, the SSD dependence of (5.5.5) and (5.5.6) is restricted to the expectation term $E\Big[(b_n - \bar{b})^{\otimes\, i}\Big] \in \mathbb{R}^{m^i(m+1)^i \times 1}$, which corresponds to the central moments of the scalar case. The elements of the $m^i(m+1)^i \times 1$ vector $E\Big[(b_n - \bar{b})^{\otimes\, i}\Big]$ are central moments of the statistics $S_{1i,n}$ and $S_{2ij,n}$. We write these central moments, which are involved in the fourth order polynomials $Q(\hat{\beta}_{t,n}, 4)$ and $Q(\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), 4)$. For the second order Taylor terms, with indexes i, j, k, l running from 1 to m, the central moments are as follows:

$$
\begin{aligned}
E[(S_{1i,n} - \omega_{1i})(S_{1j,n} - \omega_{1j})] &= \frac{1}{n} V_{1,ij}^2 \ , \\
E[(S_{1i,n} - \omega_{1i})(S_{2jk,n} - \omega_{2jk})] &= \frac{1}{n} V_{2,ijk}^2 \ , \\
E[(S_{2ij,n} - \omega_{2ij})(S_{2kl,n} - \omega_{2kl})] &= \frac{1}{n} V_{3,ijkl}^2 \ .
\end{aligned}
$$

For the third order Taylor terms, with indexes i, j, k, l, o, p running from 1 to m, the central moments are as follows:

$$
\begin{aligned}
E[(S_{1i,n} - \omega_{1i})(S_{1j,n} - \omega_{1j})(S_{1k,n} - \omega_{1k})] &= \frac{1}{n^2} V_{1,ijk}^3 \ , \\
E[(S_{1i,n} - \omega_{1i})(S_{1j,n} - \omega_{1j})(S_{2kl,n} - \omega_{2kl})] &= \frac{1}{n^2} V_{2,ijkl}^3 \ , \\
E[(S_{1i,n} - \omega_{1i})(S_{2jk,n} - \omega_{2jk})(S_{2lo,n} - \omega_{2lo})] &= \frac{1}{n^2} V_{3,ijklo}^3 \ , \\
E[(S_{2ij,n} - \omega_{2ij})(S_{2kl,n} - \omega_{2kl})(S_{2op,n} - \omega_{2op})] &= \frac{1}{n^2} V_{4,ijklop}^3 \ .
\end{aligned}
$$

For the fourth order Taylor terms, with indexes i, j, k, l, o, p, q, r running from 1 to m, the central moments are as follows:

$$
\begin{aligned}
E[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{1k,n}-\omega_{1k})(S_{1l,n}-\omega_{1l})] &= \frac{1}{n^2}V^4_{1,ijkl}+\frac{1}{n^3}U^4_{1,ijkl}\,,\\
E[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{1k,n}-\omega_{1k})(S_{2lo,n}-\omega_{2lo})] &= \frac{1}{n^2}V^4_{2,ijklo}+\frac{1}{n^3}U^4_{2,ijklo}\,,\\
E[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{2kl,n}-\omega_{2kl})(S_{2op,n}-\omega_{2op})] &= \frac{1}{n^2}V^4_{3,ijklop}\\
&\quad+\frac{1}{n^3}U^4_{3,ijklop}\,,\\
E[(S_{1i,n}-\omega_{1i})(S_{2jk,n}-\omega_{2jk})(S_{2lo,n}-\omega_{2lo})(S_{2pq,n}-\omega_{2pq})] &= \frac{1}{n^2}V^4_{4,ijklopq}\\
&\quad+\frac{1}{n^3}U^4_{4,ijklopq}\,,\\
E[(S_{2ij,n}-\omega_{2ij})(S_{2kl,n}-\omega_{2kl})(S_{2op,n}-\omega_{2op})(S_{2qr,n}-\omega_{2qr})] &= \frac{1}{n^2}V^4_{5,ijklopqr}\\
&\quad+\frac{1}{n^3}U^4_{5,ijklopqr}\,.
\end{aligned}
$$

We present the expansion of the above central moments for orders one through four and the definition of the variables $V^2_{1,ij},\ldots,V^4_{5,ijklopqr}$ and $U^4_{1,ijkl},\ldots,U^4_{5,ijklopqr}$ in Appendix C.5.1. The vectors $E\left[(b_n-\bar{b})^{\otimes\, i}\right]$ $i=2,3,4$ are reformulated in Appendix C.5.1 in a form which emphasizes the SSD. The resulting expressions consist of n-independent terms multiplying $1/n$, $1/n^2$ and $1/n^3$. These expressions are as follows:

$$
\begin{aligned}
E[(b_n-\bar{b})^{\otimes\,2}] &= \frac{1}{n}\Big[E_{2,1}\,V^2_{1,[i[j]]}+E_{2,2}\,V^2_{2,[i\ \mathrm{cs}\ [jk]]}+E_{2,3}\,V^2_{2,\ \mathrm{cs}\ [[i]jk]}\\
&\quad+E_{2,4}\,V^2_{3,\ \mathrm{cs}\ [ij\ \mathrm{cs}\ [kl]]}\Big]\\
&\equiv \frac{1}{n}V^2,\\
E[(b_n-\bar{b})^{\otimes\,3}] &= \frac{1}{n^2}\Big[E_{4,1}\,V^3_{1,[i[j[k]]]}+E_{4,2}\,V^3_{2,[i[j\ \mathrm{cs}[kl]]]}+E_{4,3}\,V^3_{2,[i\ \mathrm{cs}[[j]kl]]}\\
&\quad+E_{4,4}\,V^3_{3,[i\ \mathrm{cs}[jk\ \mathrm{cs}[lo]]]}+E_{4,5}\,V^3_{2,[[i[j]]kl]}+E_{4,6}\,V^3_{3,\ \mathrm{cs}[[i\ \mathrm{cs}[jk]]lo]}\\
&\quad+E_{4,7}\,V^3_{3,\ \mathrm{cs}[\ \mathrm{cs}[[i]jk]lo]}+E_{4,8}\,V^3_{4,\ \mathrm{cs}[ij\ \mathrm{cs}[kl\ \mathrm{cs}[op]]]}\Big]\\
&\equiv \frac{1}{n^2}V^3,\\
E[(b_n-\bar{b})^{\otimes\,4}] &= \frac{1}{n^2}\Big[E_{6,1}\,V^4_{1,[i[j[k[l]]]]}+E_{6,2}\,V^4_{2,[i[j[k\ \mathrm{cs}[lo]]]]}+E_{6,3}\,V^4_{2,[i[j\ \mathrm{cs}[[k]lo]]]}\\
&\quad+E_{6,4}\,V^4_{3,[i[j\ \mathrm{cs}[kl\ \mathrm{cs}[op]]]]}+E_{6,5}\,V^4_{2,[i\ \mathrm{cs}[[j[k]]lo]]}+E_{6,6}\,V^4_{3,[i\ \mathrm{cs}[[j\ \mathrm{cs}[kl]]lo]]}
\end{aligned}
$$

$$
\begin{aligned}
&+ E_{6,7}\, V^4_{3,[i \text{ cs}[\text{ cs}[[j]kl]op]]} + E_{6,8}\, V^4_{4,[i \text{ cs}[jk \text{ cs}[lo \text{ cs}[pq]]]]} \\
&+ E_{6,9}\, V^4_{2, \text{ cs}[[i[j[k]]]lo]} + E_{6,10}\, V^4_{3, \text{ cs}[[i[j \text{ cs}[kl]]]op]} \\
&+ E_{6,11}\, V^4_{3, \text{ cs}[[i \text{ cs}[[j]kl]]op]} + E_{6,12}\, V^4_{4, \text{ cs}[[i \text{ cs}[jk \text{ cs}[lo]]]pq]} \\
&+ E_{6,13}\, V^4_{3, \text{ cs}[\text{ cs}[[i[j]]kl]op]} + E_{6,14}\, V^4_{4, \text{ cs}[\text{ cs}[[i \text{ cs}[jk]]lo]pq]} \\
&+ E_{6,15}\, V^4_{4, \text{ cs}[\text{ cs}[\text{ cs}[[i]jk]lo]pq]} + E_{6,16}\, V^4_{5, \text{ cs}[ij \text{ cs}[kl \text{ cs}[op \text{ cs}[pq]]]]}\Big] \\
&+ O\left(\frac{1}{n^3}\right) \\
&\equiv \frac{1}{n^2}V^4 + \frac{1}{n^3}U^4.
\end{aligned}
$$

The definition of all matrices $E_{2,1}, \ldots, E_{6,16}$, as well as a description of the subscript indexing notation used for the V variables, can be found in Appendix C.5.1. Substituting the above expressions for $E[(b_n - \bar{b})^{\otimes 2}]$, $E[(b_n - \bar{b})^{\otimes 3}]$, and $E[(b_n - \bar{b})^{\otimes 4}]$ in the expression for the expectation of the fourth order Taylor polynomial of $\hat{\beta}_{t,n}$, $E[Q(\hat{\beta}_{t,n}, 4)]$, we obtain

$$
\begin{aligned}
E[Q(\hat{\beta}_{t,n}, 4)] =& \omega_2^{-1}\omega_1 + \frac{1}{2n}\Big(\mathcal{D}^2_{b_n^\top{}^2}\, \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}} V^2 \otimes I \\
&+ \frac{1}{n^2}\left[\frac{1}{3!}\Big(\mathcal{D}^3_{b_n^\top{}^3}\, \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}} V^3 \otimes I + \frac{1}{4!}\Big(\mathcal{D}^4_{b_n^\top{}^4}\, \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}} V^4 \otimes I\right] \\
&+ \frac{1}{4!n^3}\Big(\mathcal{D}^4_{b_n^\top{}^4}\, \hat{\beta}_{t,n}\Big)_{b_n=\bar{b}} U^4 \otimes I.
\end{aligned}
$$

Similarly, substituting the above expressions for $E[(b_n - \bar{b})^{\otimes 2}]$, $E[(b_n - \bar{b})^{\otimes 3}]$ and $E[(b_n - \bar{b})^{\otimes 4}]$ in the expression for the expectation of the fourth order Taylor polynomial of $\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)$, $E[Q(\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), 4)]$, we obtain

$$
\begin{aligned}
E[Q(\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), 4)] =& \text{ cs }(\omega_2^{-1}\omega_1^\top\omega_1\omega_2^{-1}) + \frac{1}{2n}\Big(\mathcal{D}^2_{b_n^\top{}^2} \text{ cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\Big)_{b_n=\bar{b}} V^2 \otimes I \\
&+ \frac{1}{n^2}\left[\frac{1}{3!}\Big(\mathcal{D}^3_{b_n^\top{}^3} \text{ cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\Big)_{b_n=\bar{b}} V^3 \otimes I + \frac{1}{4!}\Big(\mathcal{D}^4_{b_n^\top{}^4} \text{ cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\Big)_{b_n=\bar{b}} V^4 \otimes I\right] \\
&+ \frac{1}{4!n^3}\Big(\mathcal{D}^4_{b_n^\top{}^4} \text{ cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\Big)_{b_n=\bar{b}} U^4 \otimes I.
\end{aligned}
$$

Substituting the above expressions of $E[Q(\hat{\beta}_{t,n}, 4)]$ and $E[Q(\text{cs }(\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top), 4)]$ in expression (5.5.7) for the MSFE fourth order approximation, we obtain the following expression for the MSFE approximation with explicit SSD

$$\begin{aligned}
\overline{MSFE}_n \equiv & \left[E[Y_{t+1}^2] - 2\omega_1^\top \omega_2^{-1}\omega_1 + (\text{cs } \omega_2)^\top \text{cs } (\omega_2^{-1}\omega_1^\top \omega_1 \omega_2^{-1})\right] \\
& + \frac{1}{2n}\left[-2\omega_1^\top \left(\mathcal{D}^2_{b_n^\top{}^2}\, \hat{\beta}_{t,n}\right)_{b_n=\bar{b}} + (\text{cs } \omega_2)^\top \left(\mathcal{D}^2_{b_n^\top{}^2} \text{ cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\right)_{b_n=\bar{b}} V^2 \otimes I\right] \\
& + \frac{1}{n^2}\left[\frac{1}{3!}\left(-2\omega_1^\top \left(\mathcal{D}^3_{b_n^\top{}^3}\, \hat{\beta}_{t,n}\right)_{b_n=\bar{b}} + (\text{cs } \omega_2)^\top \left(\mathcal{D}^3_{b_n^\top{}^3} \text{ cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\right)_{b_n=\bar{b}}\right) V^3 \otimes I\right. \\
& \left. + \frac{1}{4!}\left(-2\omega_1^\top \left(\mathcal{D}^4_{b_n^\top{}^4}\, \hat{\beta}_{t,n}\right)_{b_n=\bar{b}} + (\text{cs } \omega_2)^\top \left(\mathcal{D}^4_{b_n^\top{}^4} \text{ cs } (\hat{\beta}_{t,n}\hat{\beta}_{t,n}^\top)\right)_{b_n=\bar{b}}\right) V^4 \otimes I\right] \\
& \equiv C + \frac{A}{n} - \frac{\Delta}{n^2}.
\end{aligned}$$

The analysis of the SSD of the MSFE approximation above follows as in the scalar case.

5.6 Monte-Carlo evidence

In this section, we present two sets of Monte Carlo experiments designed to test the Taylor algorithm method developed above.

5.6.1 Robustness of the approximating algorithm

In the first set of Monte Carlo experiments, our goal is to assess qualitatively the robustness of the Taylor algorithm to changes in the region of convergence of the Taylor series employed in the approximation. The Taylor algorithm relies on specifying a set $\mathcal{A} \subseteq \mathcal{B}$ where $\mathcal{B}$ is the region of convergence of the Taylor series of $\hat{\beta}_{t,n}$ so that

$$E[\hat{\beta}_{t,n}] = \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}] + \bar{E}[\hat{\beta}_{t,n}, \mathcal{A}^c].$$

Within $\mathcal{B}$, and therefore within $\mathcal{A}$, the Taylor series of $\hat{\beta}_{t,n}$ converges. Letting $Q(\hat{\beta}_{t,n}, 4)$ be the 4th order Taylor polynomial of $\hat{\beta}_{t,n}$, the approximation of the OLS and the MSFE are as follows:

$$E[\hat{\beta}_{t,n}] \approx \bar{E}[Q(\hat{\beta}_{t,n}, 4), \mathcal{A}], \tag{5.6.1}$$

$$MSFE_n \approx E[Y_{t+1}^2] - 2E[Y_{t+1}X_t]\bar{E}[Q(\hat{\beta}_{t,n}, 4), \mathcal{A}] + E[X_t^2]\bar{E}[Q(\hat{\beta}_{t,n}^2, 4), \mathcal{A}]. \tag{5.6.2}$$

Clearly, if $P((X_{t-n},\dots,X_{t-1})\in\mathcal{A})\approx 1$, $\bar{E}[\hat{\beta}_{t,n},\mathcal{A}^c]$ will be small and (5.6.1) and (5.6.2) can be considered good approximations. In what follows, we evaluate the accuracy of (5.6.2) for varying values of the probability $P((X_{t-n},\dots,X_{t-1})\in\mathcal{A})$. This evaluation is carried out by constructing the approximation (5.6.2) with the truncated expectations of $Q(\hat{\beta}_{t,n},4)$ and $Q(\hat{\beta}^2_{t,n},4)$. The truncated expectations of $Q(\hat{\beta}_{t,n},4)$ and $Q(\hat{\beta}^2_{t,n},4)$ are constructed using the truncated central moments in Appendix C, Section $C.2$. The resulting approximation is compared to a benchmark MSFE.

We choose the following DGP for the experiment

$$Y_{t+1}=X_t^2+U_{t+1}, \tag{5.6.3}$$

with the process $\{U_\tau\}\sim IIN(0,\sigma_u)$ and $\{X_\tau\}\sim IIN(\mu_x,\sigma_x)$. We set $\mu_x=1$, $\sigma_x=0.1$, $\sigma_u=1$. The forecast model is given by $Y_{t+1}=\beta X_t+V_{t+1}$, the forecast is given by $\hat{Y}_{t+1,n}=\hat{\beta}_{t,n}X_t$, where $\hat{\beta}_{t,n}$ is the OLS estimator (5.4.2), and the forecast error is $\epsilon_{t+1,n}=Y_{t+1}-\hat{Y}_{t+1,n}$.

Since the MSFE can not be evaluated analytically, we calculate the benchmark MSFE by means of Monte Carlo simulations. The motivation behind using Monte Carlo simulations to determine a benchmark MSFE lies in that the MSFE is equal to the expected value of the conditional mean square forecast error (CMSFE)

$$MSFE=E[CMSFE],\quad CMSFE=E_t[\epsilon^2_{t+1,n}].$$

Given a realization of the processes $\{X_\tau\}_{\tau=t-n}^{t-1}$ and $\{Y_\tau\}_{\tau=t-n+1}^{t}$, it is simple to compute the CMSFE conditional on the given sample. Generating many such samples, M, by Monte Carlo simulations, we can construct M CMSFEs, $\{CMSFE_i\}_{i=1}^M$, and approximate the MSFE by the sample mean of the simulations

$$MSFE\approx\frac{1}{M}\sum_{i=1}^{M}CMSFE_i.$$

We now describe the details involved in the construction of the benchmark MSFE. For the given set of values of the parameters $\mathcal{P}=\{\mu_x,\sigma_x,\sigma_u\}$, twenty thousand Monte Carlo simulations are conducted ($M=20000$). We use the index m to denote a particular

Monte Carlo simulation. For the mth simulation, we generate the sample series $\{x_{\tau,m}\}_{\tau=1}^{T}$ of length $T = 101$ as a realization of the explanatory process $\{X_\tau\}_{\tau=t-n}^{t}$ such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $101 \leftrightarrow t$. Each x is a realization of a normally distributed random variable, $X \sim N(\mu_x, \sigma_x)$, and the population series is independent and identically distributed, $\{X_\tau\}_{\tau=t-n}^{t-1} \sim IID$. From this sample series, we calculate the sample series $\{f_{\tau,m}\}_{\tau=1}^{T}$ by means of the relation $f_{\tau,m} = x_{\tau,m}^2$, according to the DGP (5.6.3). Finally, with the sample series $\{x_{\tau,m}\}_{\tau=1}^{T}$, and $\{f_{\tau,m}\}_{\tau=1}^{T}$, at the forecast origin $\tau = T-1$, we construct the CMSFE as follows:

$$
\begin{aligned}
CMSFE_{m,n} &= b_{\chi t,n,m}^2 + v_{\chi t,n,m}, \\
b_{\chi t,n,m}^2 &= \left[f_{t,m} - x_{t,m} \frac{\sum_{\tau=T-n}^{T-1} f_{\tau,m} x_{\tau,m}}{\sum_{\tau=T-n}^{T-1} x_{\tau,m}^2} \right]^2, \\
v_{\chi t,n,m} &= \sigma_u^2 + \frac{\sigma_u^2 x_{t,m}^2}{\sum_{\tau=T-n}^{T-1} x_{\tau,m}^2},
\end{aligned}
$$

where $b_{\chi t,n,m}^2$ and $v_{\chi t,n,m}$ are the conditional squared bias and conditional variance of the forecast error, respectively. For each simulation, we obtain $T-1 = 100$ values of the CMSFE. One for each value of n starting from $n = 1$ to $n = 100$. The case $n = 1$ refers to estimation of the OLS carried out with only one observation. For a particular set of parameters $\mathcal{P}$, we obtain an array of size $M \times T-1$ of CMSFEs, $\{CMSFE_{i,j}\}_{i=1,j=1}^{M,T-1}$. Finally, the benchmark MSFE for a set of parameters $\mathcal{P}$ and for an observation window of size n is given by the following:

$$
MSFE_n \approx \frac{1}{M} \sum_{i=1}^{M} CMSFE_{i,n}. \tag{5.6.4}
$$

The benchmark Monte Carlo MSFE is compared with the MSFE approximation obtained with the Taylor algorithm given by (5.6.2). The approximation (5.6.2) is constructed using the truncated central moments presented in Appendix $C.2$. Substituting the DGP in these central moments, the necessary truncated expectations are calculated by numerical integration. For example:

$$\begin{aligned}\bar{E}[(S_{1,n}-\omega_1),\mathcal{A}] &= \bar{E}[Y_t X_{t-1},\mathcal{A}] - \omega_1^2 P(X\in\mathcal{A})\\ &= \bar{E}[X_{t-1}^3,\mathcal{A}] + \bar{E}[U_t X_{t-1},\mathcal{A}] - \omega_1^2 P(X\in\mathcal{A})\\ &= \bar{E}[X_{t-1}^3,\mathcal{A}] - \omega_1^2 P(X\in\mathcal{A})\\ &= \bar{E}[X_{t-1}^3, I]P(X_{t-1}\in I)^{n-1} - \omega_1^2 P(X_{t-1}\in I)^n,\end{aligned}$$

where the probability set $\mathcal{A}$ is as defined in Appendix $C.1.2$. We note that, for this Monte Carlo experiment, knowledge of the DGP is necessary to calculate the truncated central moments. Knowledge of the DGP is not necessary in the Monte Carlo experiments in the next section or for empirical applications. To assess the robustness of the Taylor approximation, we change the size of $\mathcal{A}$ by changing the size of I_i for $i = t-n,\ldots,t-1$ by changing the size of δ_i. For $\mu_x = 1$ and $\sigma_x = .1$, the largest possible value of δ_i is $\delta_i \approx 4.21267\sigma_x$ and $P(X_i \in I_i) \approx 0.9999747$. By reducing the size of δ_i, the intervals I_i and the region $\mathcal{A}$ shrink. The other values of δ_i used are $2.8\sigma_x, 2.5\sigma_x, 2\sigma_x$, and the respective probabilities are $P(X_i \in I_i) \approx 0.99488974$, $P(X_i \in I_i) \approx 0.98758$, and $P(X_i \in I_i) \approx 0.9544979$. The resulting MSFE approximations are presented in Figure 5.2. This shows that the MSFE approximation given by (5.6.2) is not robust for large n.

Next, in what follows, we assess the robustness of the Taylor algorithm given by (5.4.11). This approximation is obtained through the assumption that the range of the explanatory random variable is contained inside a set $\mathcal{A}$, $P(X \in \mathcal{A}) = 1$, which is inside the region of convergence $\mathcal{B}$ of the Taylor series of the OLS. This results in $E[\hat{\beta}_{t,n}] = \bar{E}[\hat{\beta}_{t,n},\mathcal{A}]$ and $\bar{E}[\hat{\beta}_{t,n},\mathcal{A}^c] = 0$. We want to evaluate the performance of the MSFE approximation (5.4.11) when applied to circumstances that violate the containment assumption, i.e., when the range of the explanatory random variable goes beyond the region of convergence and $P(X \in \mathcal{A}) < 1$. We conduct two experiments with the explanatory process $\{X_\tau\} \sim IIN(\mu_x, \sigma_x)$. Clearly, this process does not satisfy the containment condition since $P(X \in \mathcal{A}) < 1$ for any compact $\mathcal{A}$. The DGP used is $Y_{t+1} = \theta_1 X_t + \theta_2 X_t^2 + U_{t+1}$ with $\{U_\tau\} \sim IIN(0, \sigma_u)$. The set of parameters investigated are given in Table 5.1. The benchmark MSFE is obtained by Monte Carlo simulations as described in the previous set of experiments. As derived in Section $C.1.1$, the radius

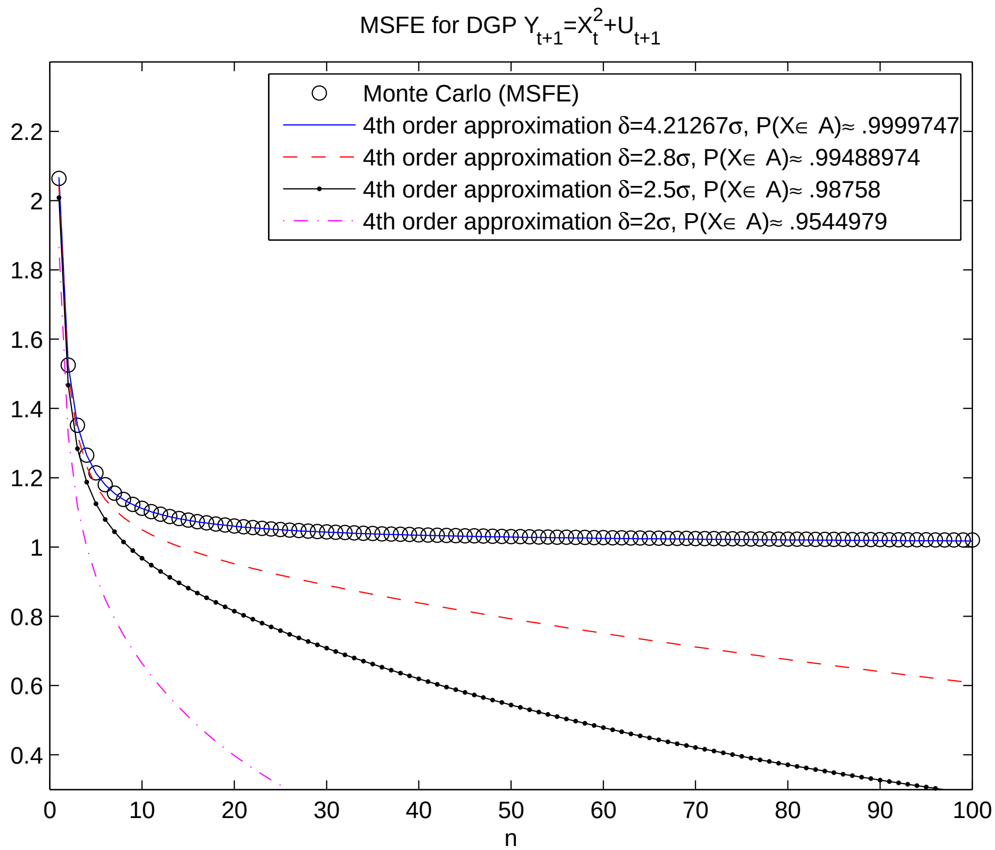


Figure 5.2: Benchmark Monte Carlo MSFE and Taylor algorithm MSFE approximation for different probability sets

of convergence in terms of the explanatory variable is $R_n = \sqrt{2n(\mu_x^2 + \sigma_x^2)}$. The experiments are designed by keeping all parameters fixed except for μ_x and σ_x. In the first experiment, $\mu_x = 10$ and $\sigma_x = 0.1$. In the second experiment $\mu_x = 0.1$ and $\sigma_x = 10$. Clearly, the radius of convergence R_n remains fixed in the two experiments by the choice of μ_x and σ_x. In the first experiment, the probability with $n = 1$ of $X \in \mathcal{B}$ is almost one, since σ_x is small. In the second experiment, this probability decreases to 0.8427. Figures 5.3 and 5.4 present the benchmark MSFE and the Taylor algorithm approximation MSFE for the two experiments. From these, we can see that the MSFE approximation (5.4.11), under violation of the containment assumption, remains robust for large values of n, but fails to replicate the benchmark MSFE for small values of n. The MSFE approximation given by (5.4.11) outperforms the MSFE approximation given by (5.6.2) and therefore validates making the containment assumption $P(X \in \mathcal{A}) \approx 1$.

$\theta_1 = 1, \theta_2 = 1,$ $\sigma_u = 1, R_1 = 14.1428$		
μ_x	10	0.1
σ_x	0.1	10
$P(X \in \mathcal{B})$	1	0.8427

Table 5.1: Set of parameters for the experiments to assess the containment condition

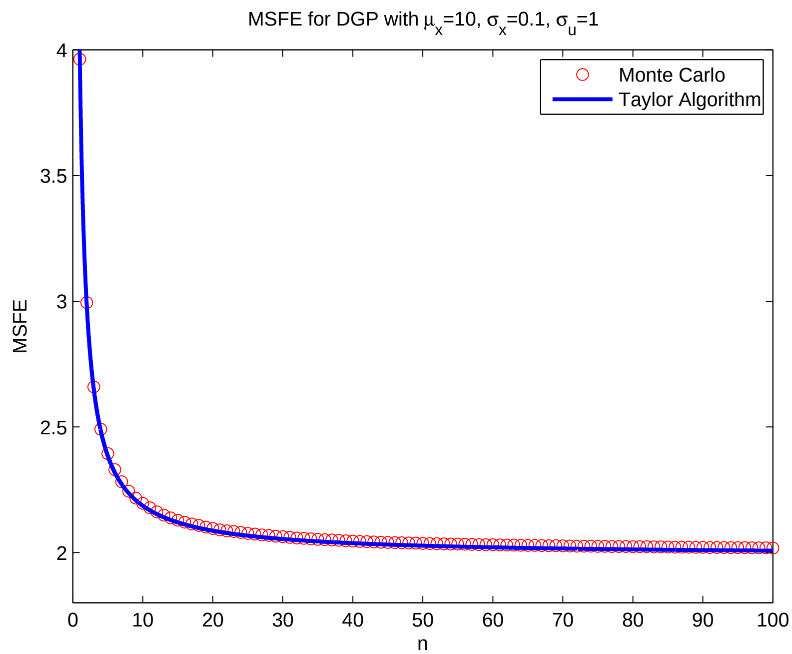


Figure 5.3: MSFE for a quadratic DGP with $\theta_1 = 1, \theta_2 = 1,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

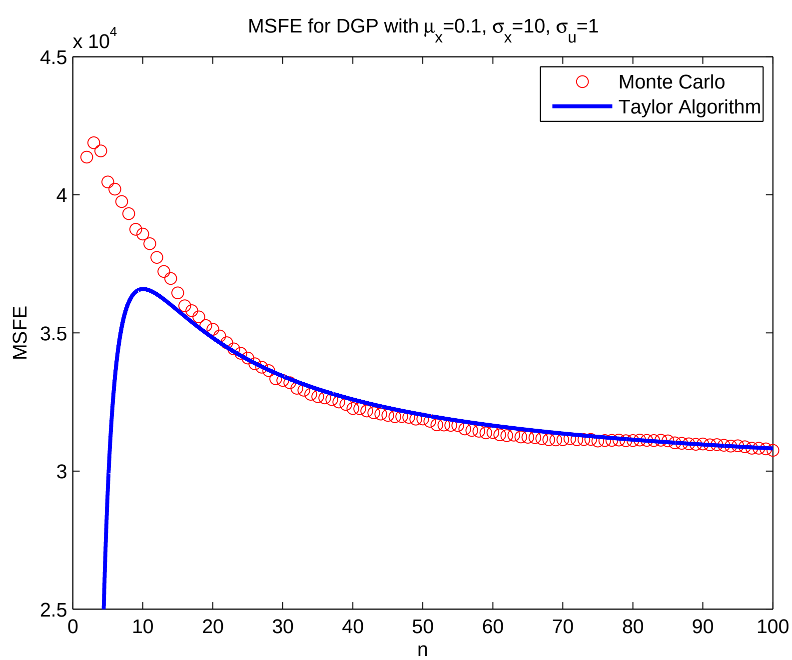


Figure 5.4: MSFE for a quadratic DGP with $\theta_1 = 1, \theta_2 = 1,\ \mu_x = 0.1, \sigma_x = 10, \sigma_u = 1$

5.6.2 Assessing misspecification

In this section, we present Monte Carlo experiments to investigate the ramifications of misspecification in the forecasting problem described in Section 2.4 with independent identically distributed processes and to evaluate the ability of the Taylor algorithm to capture these effects. The paramount assumption made in this chapter, that of independence of the explanatory variables, is imposed on the simulations that follow. To carry out this endeavor, we construct a benchmark MSFE by means of Monte Carlo simulations. This benchmark MSFE is then compared to the MSFE approximation obtained with the Taylor algorithm and given by (5.4.11). For the analysis, we consider several DGPs each of the general form

$$Y_{t+1} = \varphi(X_t, \theta) + U_{t+1},$$

where $\{U_\tau\} \sim IIN(0, \sigma_u)$ is an innovation process, $\{X_\tau\} \sim IIN(\mu_x, \sigma_x)$, and θ is a vector of parameters. The DGPs considered differ in the functional form of φ. The functions we consider are as follows:

$$\begin{aligned}
\varphi_1(X_t, \theta) &= \theta_1 X_t + \theta_2 X_t^{\theta_3}, \\
\varphi_2(X_t, \theta) &= \theta_4 - \theta_3 \log[1 + \exp(-\theta_2/\theta_3 - \theta_1 X_t/\theta_3)], \\
\varphi_3(X_t, \theta) &= \theta_1 X_t + \theta_2 (X_t + \theta_3)^2 + \sin(\pi(X_t - 1)/\theta_4), \\
\varphi_4(X_t, \theta) &= \theta_1 X_t + \theta_2 Z_t.
\end{aligned} \tag{5.6.5}$$

As described in the previous section, the MSFE cannot be evaluated analytically, so we calculate the benchmark MSFE by means of Monte Carlo simulations. The motivation behind using Monte Carlo simulations to determine a benchmark MSFE lies in the fact that the MSFE is equal to the expected value of the CMSFE. Given a realization of the processes $\{X_\tau\}_{\tau=t-n}^{t-1}$ and $\{Y_\tau\}_{\tau=t-n+1}^{t}$, it is simple to compute the CMSFE conditional on the given sample. Generating many such samples, M, by Monte Carlo simulations, we can construct M conditional mean square forecast errors, $\{CMSFE_i\}_{i=1}^{M}$, and approximate the MSFE by the sample mean of the simulations.

We now describe the details involved in the construction of the benchmark MSFE. For the given set of values of the parameters $\mathcal{P} = \{\mu_x, \sigma_x, \sigma_u, \theta\}$ and a particular functional form of φ from the given in (5.6.5), twenty thousand Monte Carlo simulations are conducted ($M = 20000$). We use the index m to denote a particular Monte-Carlo simulation. For the mth simulation, we generate the sample series $\{x_{\tau,m}\}_{\tau=1}^{T}$ of length $T = 501$ as a realization of the explanatory process $\{X_\tau\}_{\tau=t-n}^{t}$, such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $501 \leftrightarrow t$. Each x is a realization of a normally distributed random variable, $X \sim N(\mu_x, \sigma_x)$, and the population series is independent and identically distributed, $\{X_\tau\}_{\tau=t-n}^{t-1} \sim IID$. From this sample series, we calculate the sample series $\{f_{\tau,m}\}_{\tau=1}^{T}$ by means of the relation $f_{\tau,m} = \varphi_i(x_{\tau,m}, \theta)$ for each of the DGPs in (5.6.5).

Finally, with the sample series $\{x_{\tau,m}\}_{\tau=1}^{T}$, and $\{f_{\tau,m}\}_{\tau=1}^{T}$, at the forecast origin $\tau = T-1$, we construct the CMSFE as follows:

$$
\begin{aligned}
CMSFE_{m,n} &= b^2_{\chi t,n,m} + v_{\chi t,n,m},\\
b^2_{\chi t,n,m} &= \left[f_{t,m} - x_{t,m} \frac{\sum_{\tau=T-n}^{T-1} f_{\tau,m} x_{\tau,m}}{\sum_{\tau=T-n}^{T-1} x^2_{\tau,m}} \right]^2,\\
v_{\chi t,n,m} &= \sigma_u^2 + \frac{\sigma_u^2 x^2_{t,m}}{\sum_{\tau=T-n}^{T-1} x^2_{\tau,m}},
\end{aligned}
$$

where $b^2_{\chi t,n,m}$ and $v_{\chi t,n,m}$ are the conditional squared bias and conditional variance of the forecast error, respectively. For each simulation, we obtain $T-1 = 500$ values of the CMSFE. One for each value of n starting from $n = 1$ to $n = 500$. The case $n = 1$ refers to estimation of the OLS carried out with only one observation. For a particular set of parameters $\mathcal{P}$, we obtain an array of size $M \times T-1$ of CMSFEs, $\{CMSFE_{i,j}\}_{i=1,j=1}^{M,T-1}$. Finally, the benchmark MSFE for a set of parameters $\mathcal{P}$ and for an observation window of size n is given by the following:

$$
MSFE_n \approx \frac{1}{M} \sum_{i=1}^{M} CMSFE_{i,n}. \tag{5.6.6}
$$

The benchmark Monte Carlo MSFE is compared with the MSFE approximation obtained with the Taylor algorithm given by (5.4.11). The approximation (5.4.11) is constructed by use of sample moments in place of their population counterparts. For this, we generate the sample series $\{x_\tau\}_{\tau=1}^{N}$ of length $N = 5000$ as a realization of the explanatory process $\{X_\tau\}_{\tau=t-n}^{t}$ such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $5000 \leftrightarrow t$. Each x is a realization of a normally distributed random variable, $X \sim N(\mu_x, \sigma_x)$, and the population series is independent and identically distributed, $\{X_\tau\}_{\tau=t-n}^{t-1} \sim IID$. Similarly, we generate the sample series $\{u_\tau\}_{\tau=1}^{N}$ of length $N = 5000$ as a realization of the innovation process $\{U_\tau\}_{\tau=t-n}^{t}$ such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $5000 \leftrightarrow t$. Each u is a realization of a normally distributed random variable, $U \sim N(0, \sigma_u)$, and the population series is independent and identically distributed, $\{U_\tau\}_{\tau=t-n}^{t-1} \sim IID$. Finally, the sample series $\{y_\tau\}_{\tau=1}^{N}$ is generated by means of the relation $y_\tau = \varphi_i(X_\tau, \theta) + u_\tau$ for each DGP in (5.6.5).

The population moments in (5.4.11) are estimated by generating their sample counterparts. For example:

$$E[Y_t X_{t-1}^3] \approx \frac{1}{N} \sum_{\tau=1}^{N} y_\tau x_\tau^3,$$
$$E[Y_t^2 X_{t-1}^2] \approx \frac{1}{N} \sum_{\tau=1}^{N} y_\tau^2 x_\tau^2.$$

Therefore, for a given set of the parameters, $\mathcal{P} = \{\mu_x, \sigma_x, \sigma_u, \theta\}$, we can generate the necessary sample moments and ultimately evaluate (5.4.11) for different values of the observation window size n. The resulting MSFE can be compared to the benchmark MSFE (5.6.6). In the next section, we discuss results for different sets of values of the parameters involved for the four DGPs given in (5.6.5).

5.6.3 Discussion

The sets of parameter values investigated and the reference to their corresponding MSFE plots are given in tables 5.2, 5.3, 5.4, 5.5 for the four functional forms of the DGP given in (5.6.5). We first describe the results for the DGP with $\varphi_1(X_t, \theta) = \theta_1 X_t + \theta_2 X_t^{\theta_3}$. The values $\theta_1 = 2$, $\theta_2 = 0.05$, $\theta_3 = 2$ are fixed. For these parameter values, the misspecification is due to the quadratic term. The variance of the explanatory variable σ_x^2 and the variance of the innovation σ_u^2 are evaluated at different values, as shown in Table 5.2, columns one and two. For the nine experiments conducted, the extremum value is negative, $n_0 < 0$. These are given in the fourth column of the table. With this, the approximation (5.4.11), by proposition 5.4, suggests there exists no optimal observation window and all data available must be used to forecast. In all nine experiments, the benchmark MSFE monotonically decreases with minimum value at the last value of $n = 500$. In the figures, we plot both the SSD of the benchmark Monte Carlo MSFE and the SSD of the Taylor approximation MSFE for values of n from zero to one hundred. Qualitatively, in all nine experiments, the Taylor algorithm provides an MSFE approximation which replicates the form of the benchmark MSFE. For a given value of the explanatory variable variance σ_x^2, as the variance of the innovation σ_u^2 increases, the level of the benchmark MSFE and the level of the Taylor approximation MSFE increase but the SSD remains monotonic decreasing. Continuing with the DGP with $\varphi_1(X_t, \theta) = \theta_1 X_t + \theta_2 X_t^{\theta_3}$, we conduct nine more experiments with fixed new values of $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$. By increasing the value of the parameter θ_2, we increase the influence of the quadratic term in the DGP and therefore increase the misspecification of the linear forecast model employed by the Taylor algorithm. The same values of the variance of the explanatory variable σ_x^2 and the variance of the innovation σ_u^2 are used as in the previous nine experiments. For this second set of nine experiments, the extremum values are also negative, $n_0 < 0$. These values are given in the seventh column of the table. Again, the approximation (5.4.11) suggests there exists no optimal observation window and all data available must be used to forecast. As before, in all nine experiments, the benchmark MSFE monotonically decreases with minimum value at the last value of $n = 500$. Qualitatively, the results are similar to those of the previous nine experiments. The Taylor algorithm provides an MSFE approximation which replicates the form of the benchmark MSFE. Nonetheless, compared to the previous nine experiments, the Taylor approximation appears less accurate, as can be seen in Figures $5.17, 5.18, 5.19$. This should be attributed to an increase in the variance of the dependent process $\{Y_\tau\}$, rather than viewed as an effect of the "increase" in misspecification.

Similar analysis is carried out for the other three DGPs, as given in (5.6.5). For each of these three functional forms of the DGP, two sets of experiments are conducted. For $\varphi_2(X_t, \theta) = \theta_4 - \theta_3 \log[1 + \exp(-\theta_2/\theta_3 - \theta_1 X_t/\theta_3)]$, the first set of experiments has parameters $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$ and the second set of experiments has parameters $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$. For $\varphi_3(X_t, \theta) = \theta_1 X_t + \theta_2(X_t + \theta_3)^2 + \sin(\pi(X_t - 1)/\theta_4)$, the first set of experiments has parameters $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1$ and the second set of experiments has parameters $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1$. For $\varphi_4(X_t, \theta) = \theta_1 X_t + \theta_2 Z_t$, the first set of experiments has parameters $\theta_1 = 1, \theta_2 = 0.001$ and the second set of experiments has parameters $\theta_1 = 1, \theta_2 = 2$. For each of these six sets of experiments, the variance of the explanatory variable σ_x^2 is evaluated at three different values, and the variance of the innovation σ_u^2 is evaluated at nine different values, for a total of fifty-four experiments. For all experiments, as presented in the tables, the extremum values are negative, $n_0 < 0$. This implies, that for all examples studied, the approximation (5.4.11) suggests there exists no optimal observation window and all data available must be used to forecast. Furthermore, for all experiments, the benchmark MSFE monotonically decreases with minimum value at the last value of $n = 500$. Regardless of the level of misspecification achieved by the different sets of parameters, the general shape of the SSD of the benchmark MSFE, and that of the SSD of the Taylor approximation MSFE, is monotonic decreasing. The results of the experiments point to the conclusion that, when the processes involved in the forecast problem are temporally independent, there exists no optimal observation window and it is optimal to use all data available to form a forecast.

5.7 Conclusions

In this chapter, we analyze the SSD of the MSFE for a forecasting problem with a forecast model consisting of a linear regression which misspecifies the data generating problem. The observed processes are assumed to be i.i.d. As described in section 5.3, the most important result in the literature on the SSD of the MSFE under misspecification and with i.i.d. processes is given by White [150]. This result describes the behavior of the MSFE as the sample size n goes to infinity. By developing a Taylor algorithm, we formulate an approximation of the MSFE which can be used to explain the SSD of the MSFE for finite values of the sample size variable n. We evaluate this algorithm by numerical experiments and a benchmark MSFE constructed by Monte Carlo simulations. For the cases of functional misspecifications studied, the Taylor algorithm MSFE replicates the behavior of the benchmark MSFE for finite values of n. Furthermore, the experiments reveal the MSFE, for the cases studied, decreases monotonically, leading to the conclusion that no optimal observation windows of finite size exist.

		$\varphi_1,\ \mu_x = 10$					
σ_x	σ_u	$\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$			$\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$		
		Figure	n_o	$\underline{n}$	Figure	n_0	$\underline{n}$
0.1	0.01	5.5	-0.0040	NA	5.14	-0.0029	NA
	0.1	5.6	-0.0042	NA	5.15	-0.0032	NA
	1	5.7	-0.0019	NA	5.16	-0.0050	NA
1	0.1	5.8	-0.0439	NA	5.17	-0.0334	NA
	1	5.9	-0.0886	NA	5.18	-0.0356	NA
	5	5.10	-0.0824	NA	5.19	-0.0468	NA
10	1	5.11	-1.1851	NA	5.20	-1.1953	NA
	10	5.12	-1.6691	NA	5.21	-1.1913	NA
	30	5.13	-2.3033	NA	5.22	-1.1859	NA

Table 5.2: NA indicates not applicable by definition

		$\varphi_2,\ \mu_x = 10$					
σ_x	σ_u	$\theta_1 = 1, \theta_2 = 0.01,$ $\theta_3 = 1, \theta_4 = 0.001$			$\theta_1 = 1, \theta_2 = 2.5,$ $\theta_3 = 1, \theta_4 = 1.5$		
		Figure	n_o	$\underline{n}$	Figure	n_0	$\underline{n}$
0.1	0.01	5.23	-0.0016	NA	5.32	-0.0019	NA
	0.1	5.24	-0.0016	NA	5.33	-8.8952e-04	NA
	1	5.25	-0.0016	NA	5.34	-0.0016	NA
1	0.1	5.26	-0.0781	NA	5.35	-0.2303	NA
	1	5.27	-0.0781	NA	5.36	-0.0769	NA
	5	5.28	-0.0781	NA	5.37	-0.0770	NA
10	1	5.29	-6.0039	NA	5.38	-5.1178	NA
	10	5.30	-2.8170	NA	5.39	-2.6739	NA
	30	5.31	-2.5644	NA	5.40	-2.5453	NA

Table 5.3: NA indicates not applicable by definition

		φ_3, $\mu_x = 10$					
σ_x	σ_u	$\theta_1 = 1, \theta_2 = 0.001,$ $\theta_3 = 1, \theta_4 = 1$			$\theta_1 = 0.1, \theta_2 = 2,$ $\theta_3 = 0.1, \theta_4 = 1$		
		Figure	n_o	$\underline{n}$	Figure	n_0	$\underline{n}$
0.1	0.01	5.41	-0.0056	NA	5.50	-0.0024	NA
	0.1	5.42	-0.0037	NA	5.51	-0.0027	NA
	1	5.43	-4.8605e-04	NA	5.52	-0.0048	NA
1	0.1	5.44	-0.0755	NA	5.53	-0.0278	NA
	1	5.45	-0.0763	NA	5.54	-0.0300	NA
	5	5.46	-0.0777	NA	5.55	-0.0416	NA
10	1	5.47	-2.6647	NA	5.56	-1.1961	NA
	10	5.48	-2.5121	NA	5.57	-1.1920	NA
	30	5.49	-2.5160	NA	5.58	-1.1866	NA

Table 5.4: NA indicates not applicable by definition

			φ_4, $\mu_x = 10$, $\mu_z = 8$					
σ_x	σ_z	σ_u	$\theta_1 = 1, \theta_2 = 0.001$			$\theta_1 = 1, \theta_2 = 2$		
			Figure	n_o	$\underline{n}$	Figure	n_0	$\underline{n}$
0.1	0.15	0.01	5.59	-0.0012	NA	5.68	-1.0718e-04	NA
		0.1	5.60	-0.0012	NA	5.69	4.9040e-04	NA
		1	5.61	-0.0012	NA	5.70	-2.2814e-04	NA
1	1.5	0.1	5.62	-0.0916	NA	5.71	-0.1124	NA
		1	5.63	-0.0922	NA	5.72	-0.1011	NA
		5	5.64	-0.0923	NA	5.73	-0.0801	NA
10	15	1	5.65	-3.2029	NA	5.74	-2.8364	NA
		10	5.66	-3.2129	NA	5.75	-2.7043	NA
		30	5.67	-3.2137	NA	5.76	-2.6989	NA

Table 5.5: NA indicates not applicable by definition

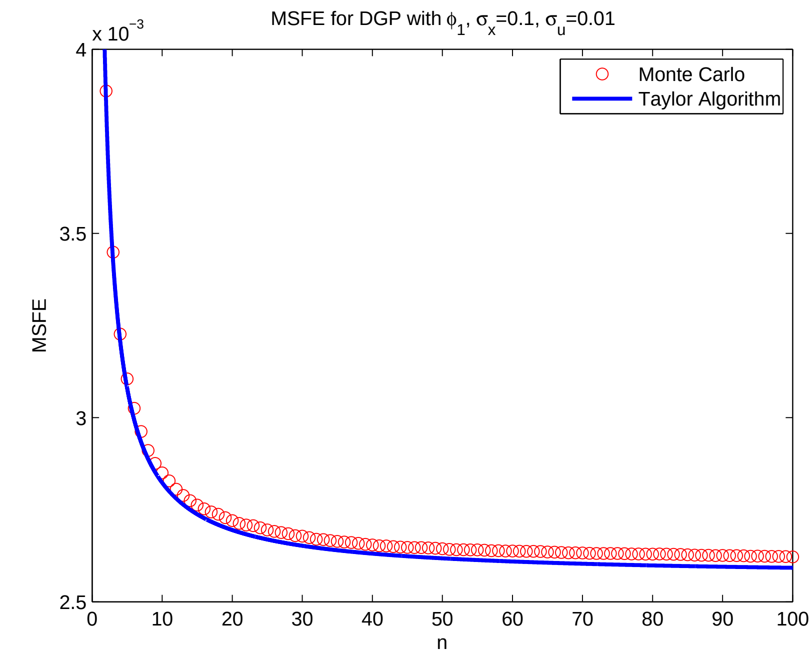


Figure 5.5: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

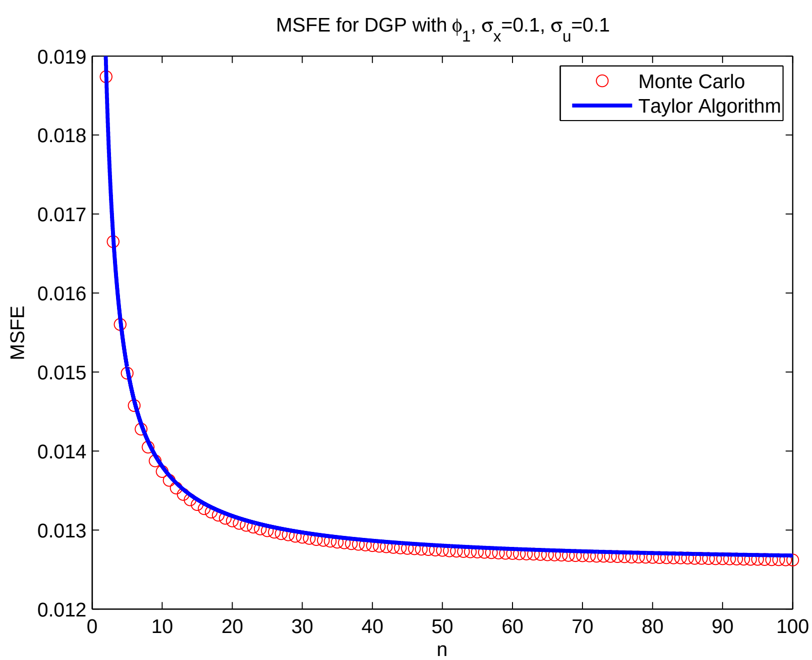


Figure 5.6: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

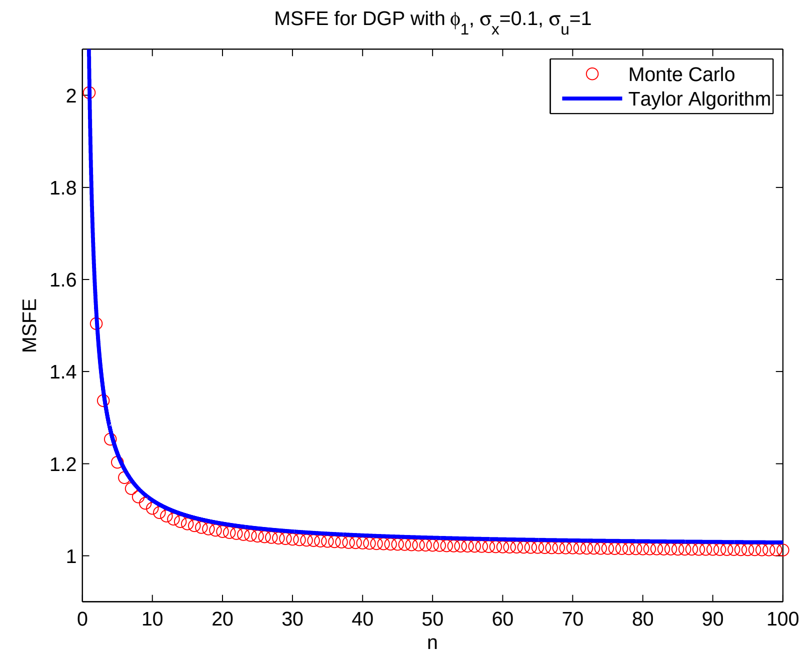


Figure 5.7: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

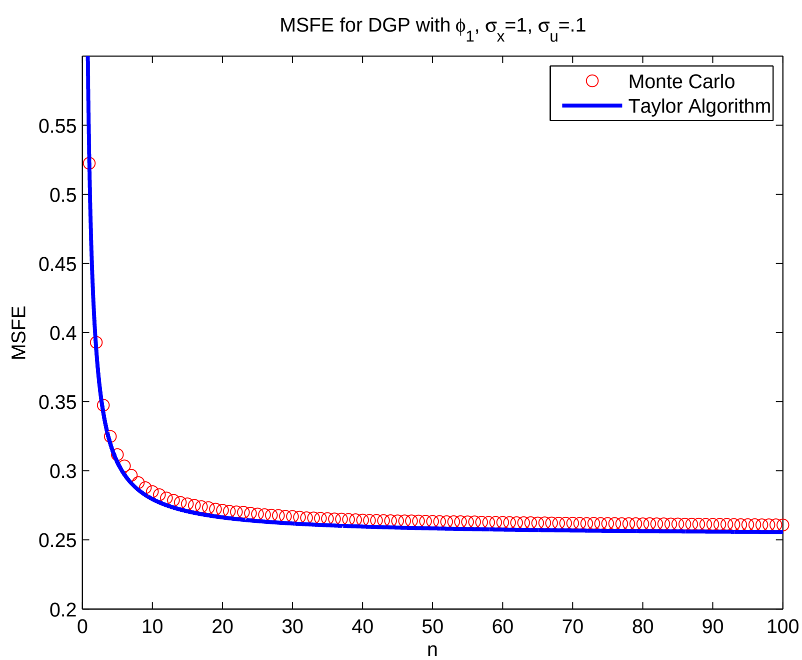


Figure 5.8: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

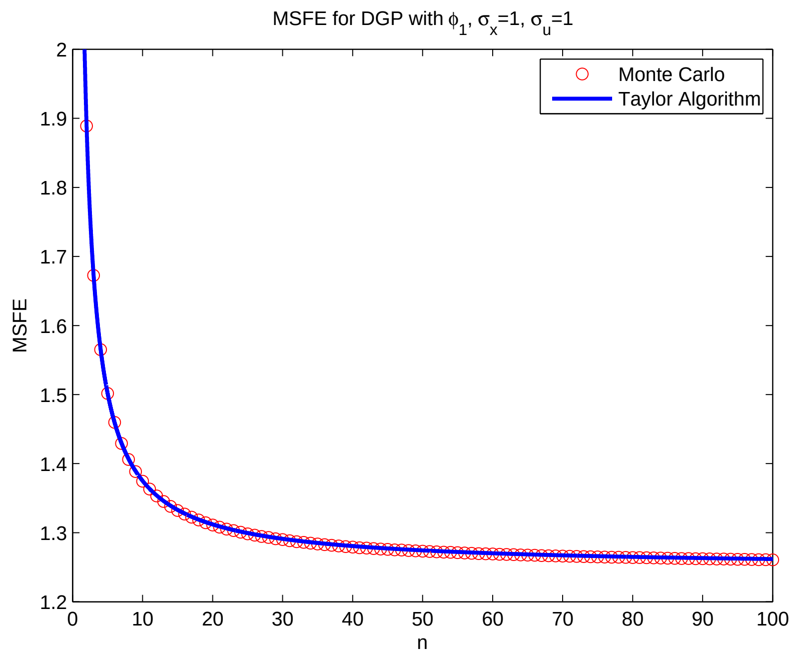


Figure 5.9: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 1$

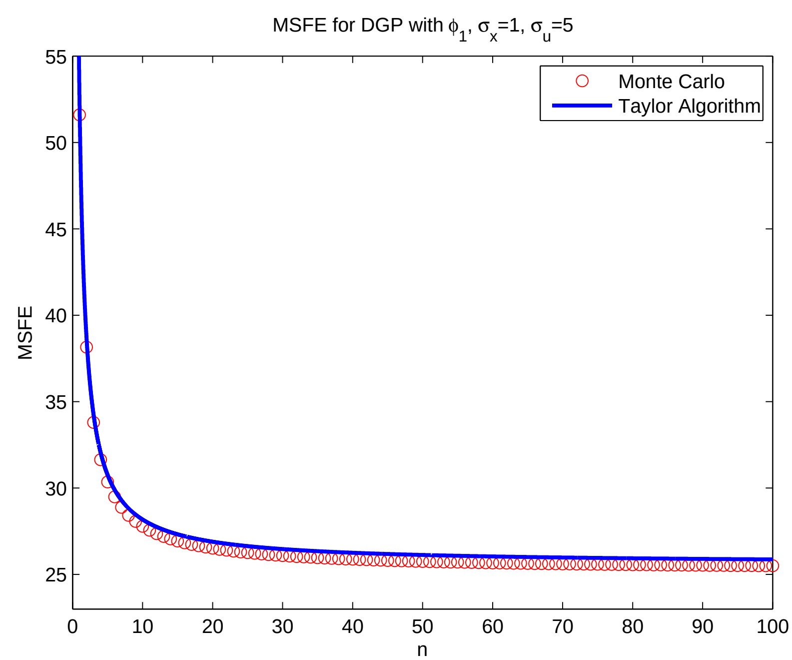


Figure 5.10: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 5$

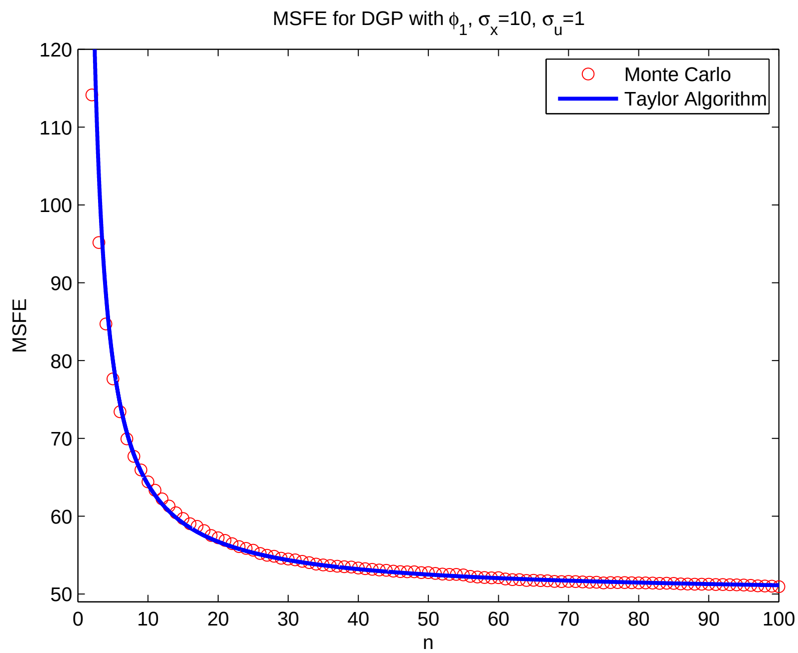


Figure 5.11: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 1$

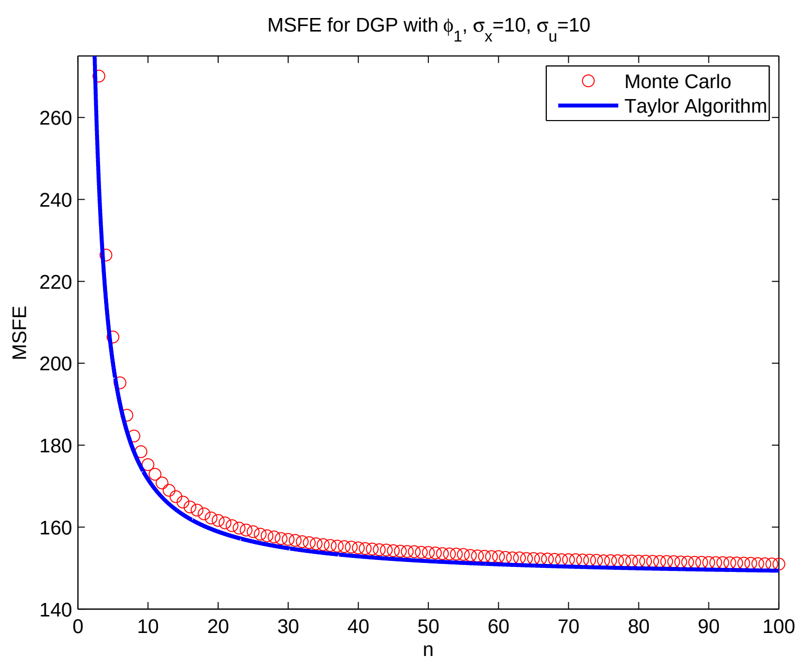


Figure 5.12: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 10$

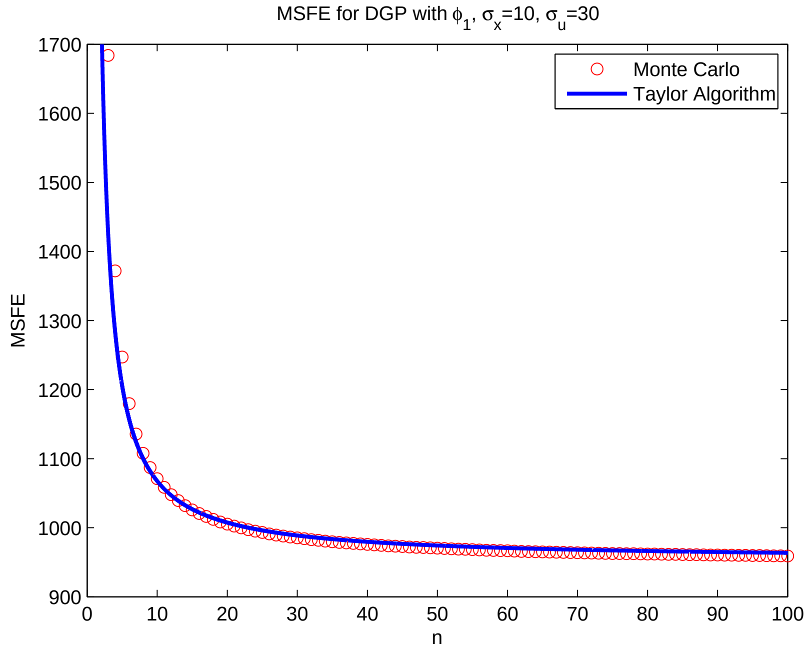


Figure 5.13: MSFE for $\varphi_1(x)$, $\theta_1 = 2, \theta_2 = 0.05, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 30$

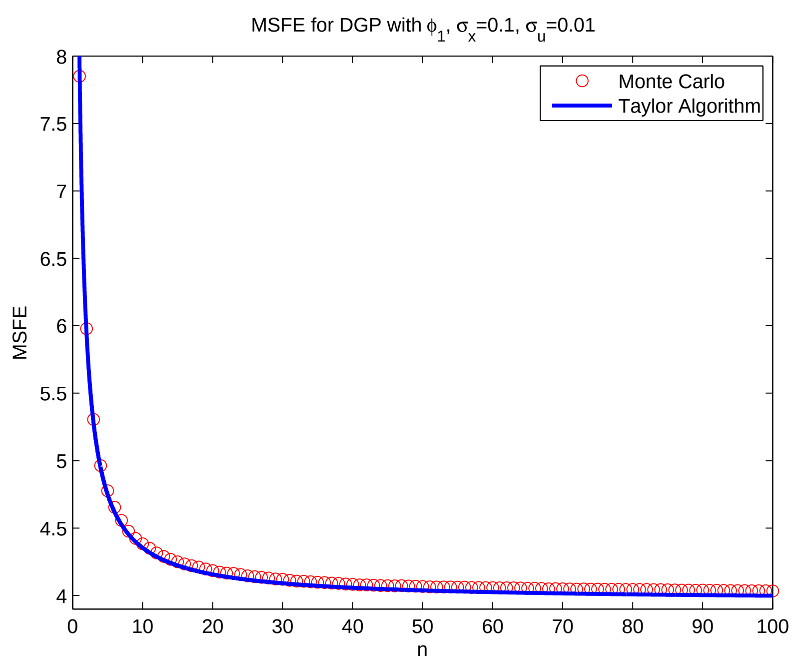


Figure 5.14: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

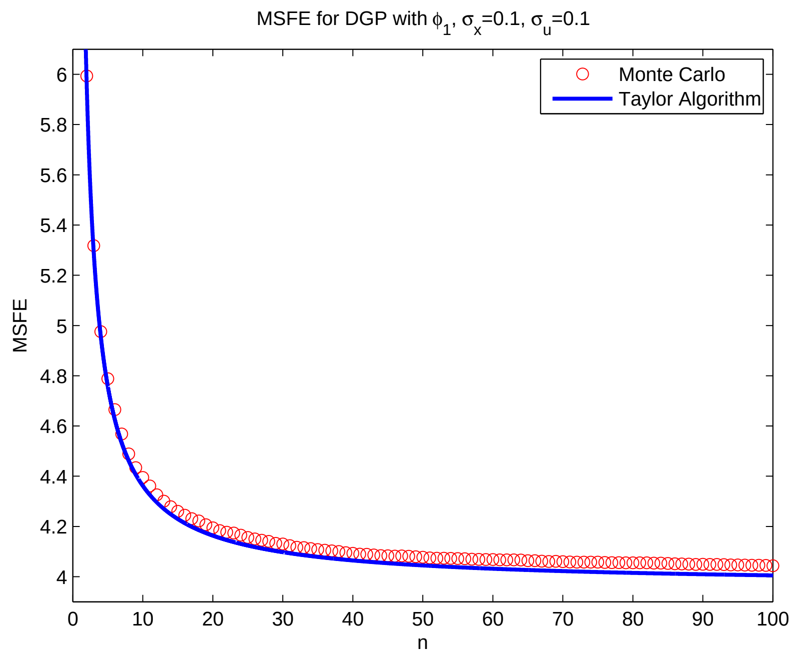


Figure 5.15: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

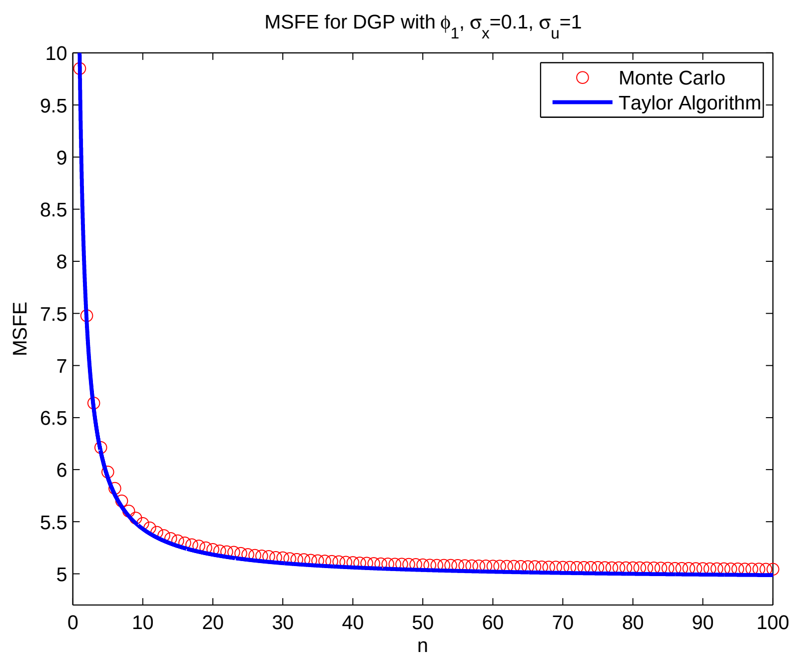


Figure 5.16: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

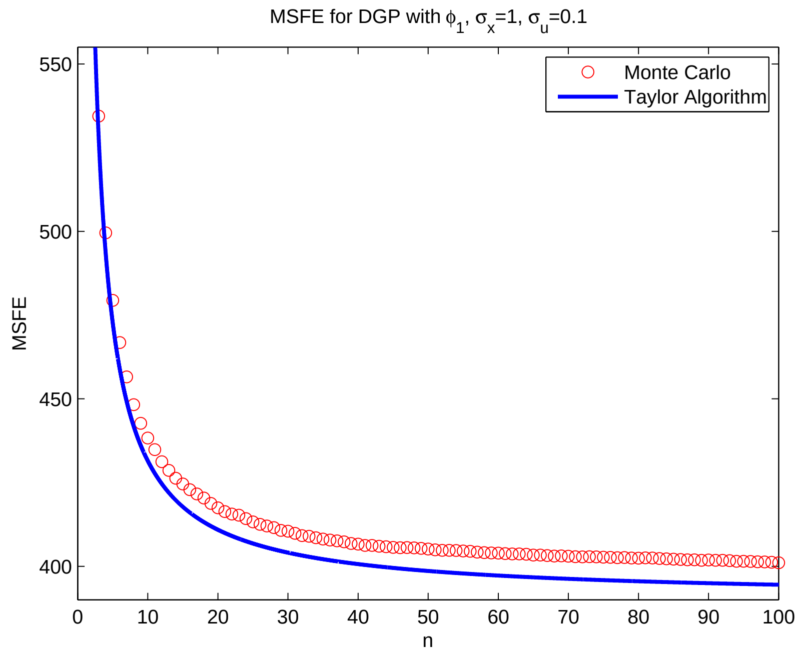


Figure 5.17: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

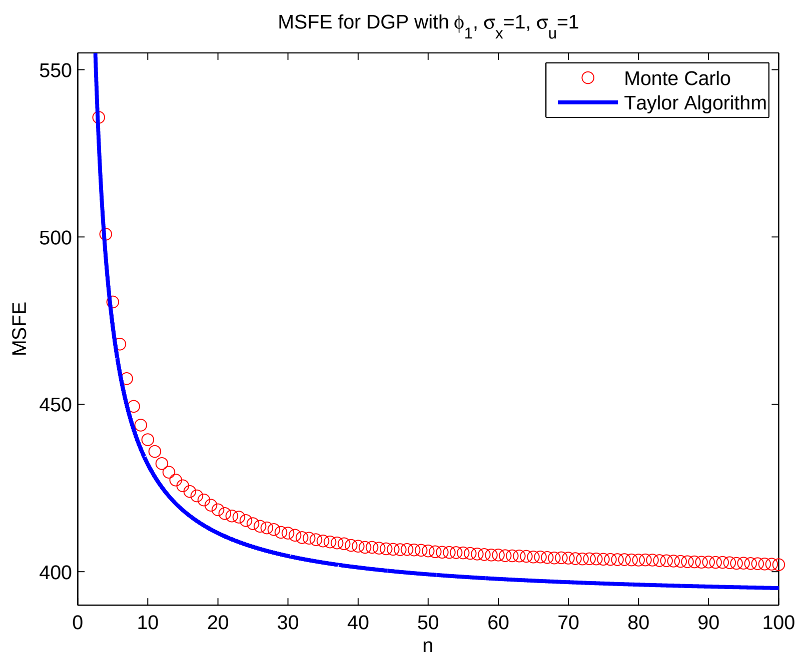


Figure 5.18: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 1$

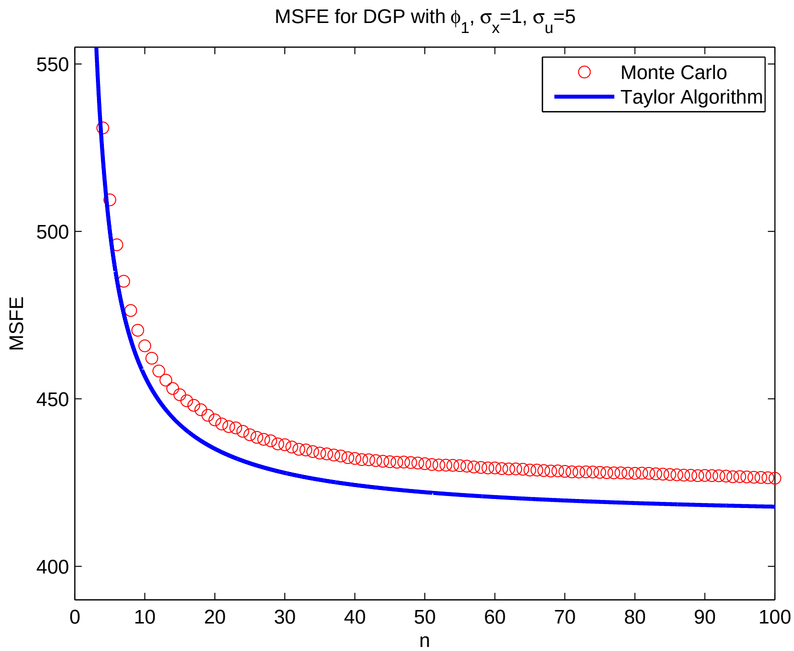


Figure 5.19: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 5$

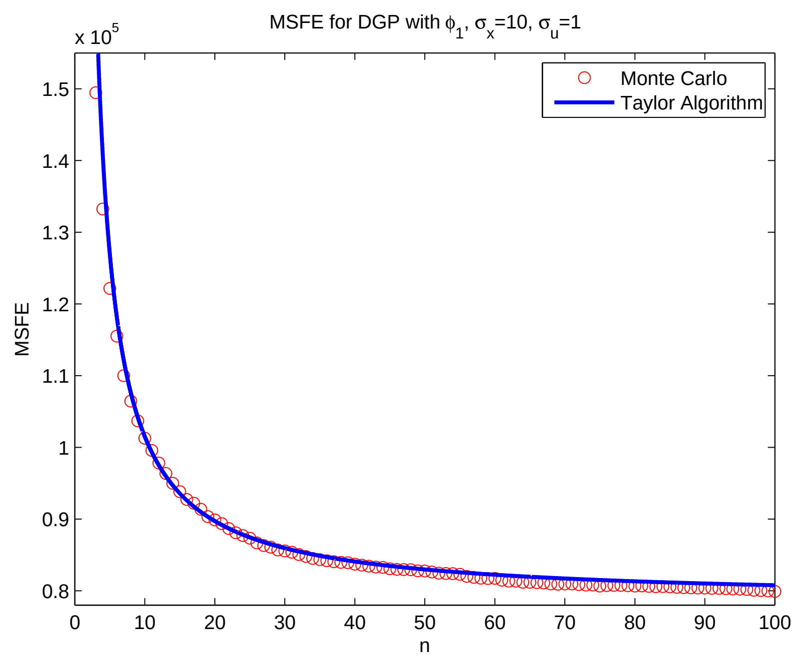


Figure 5.20: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 1$

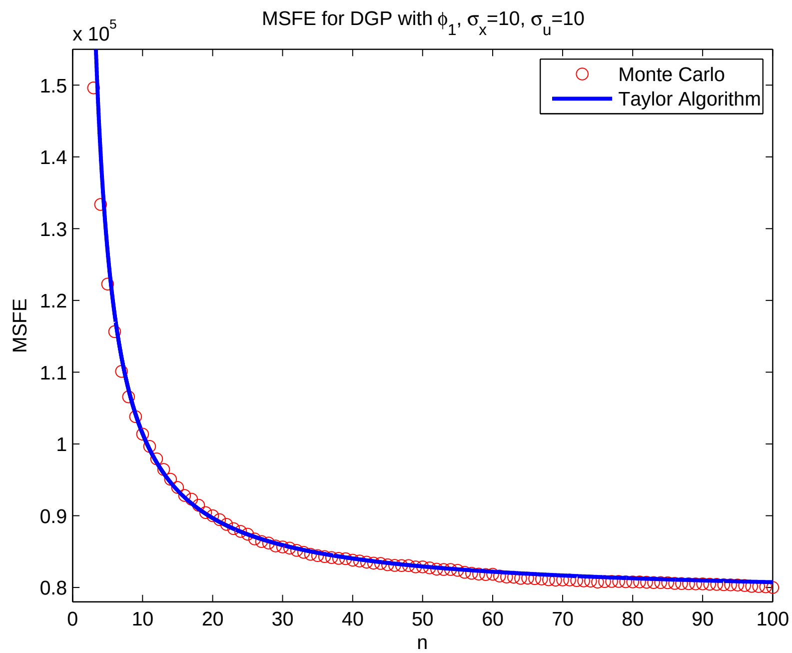


Figure 5.21: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 10$

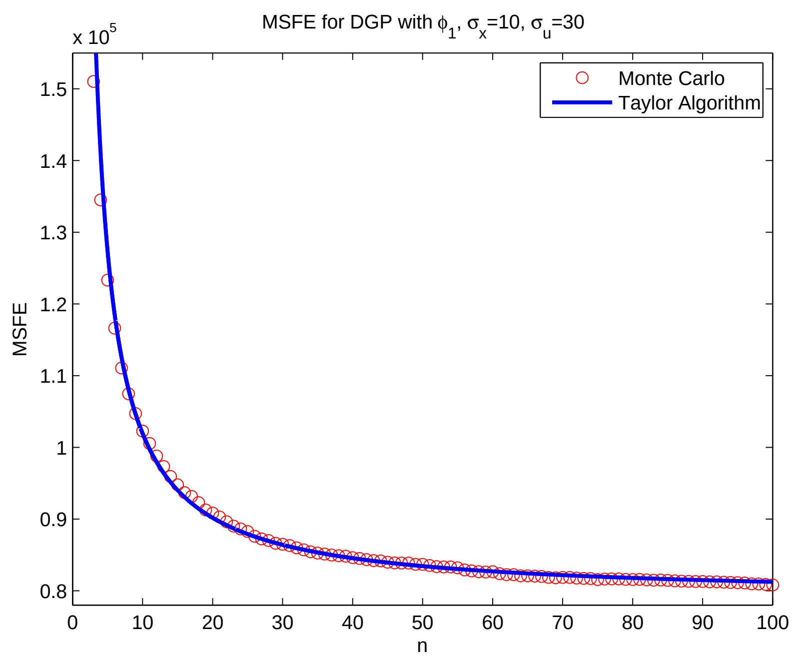


Figure 5.22: MSFE for $\varphi_1(x)$, $\theta_1 = 1, \theta_2 = 2, \theta_3 = 2$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 30$

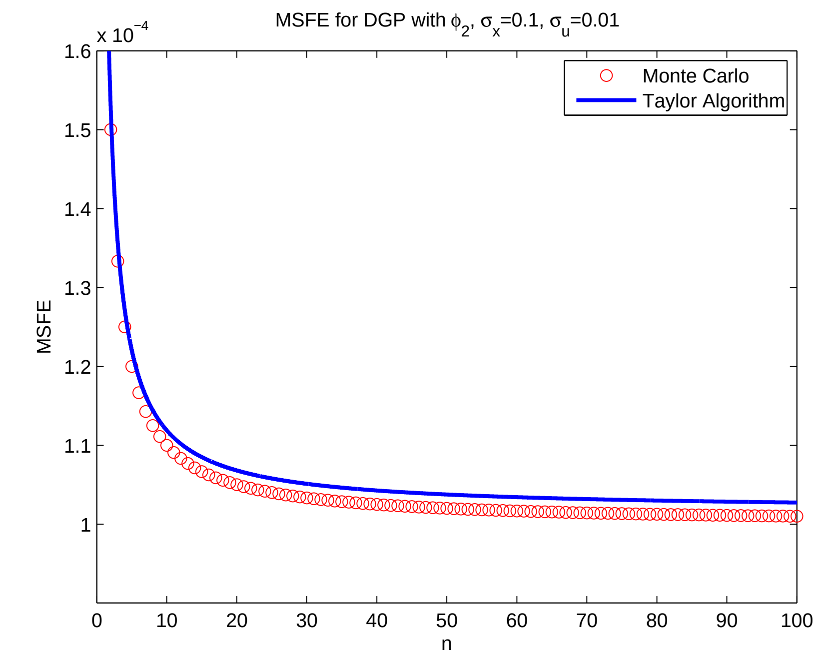


Figure 5.23: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

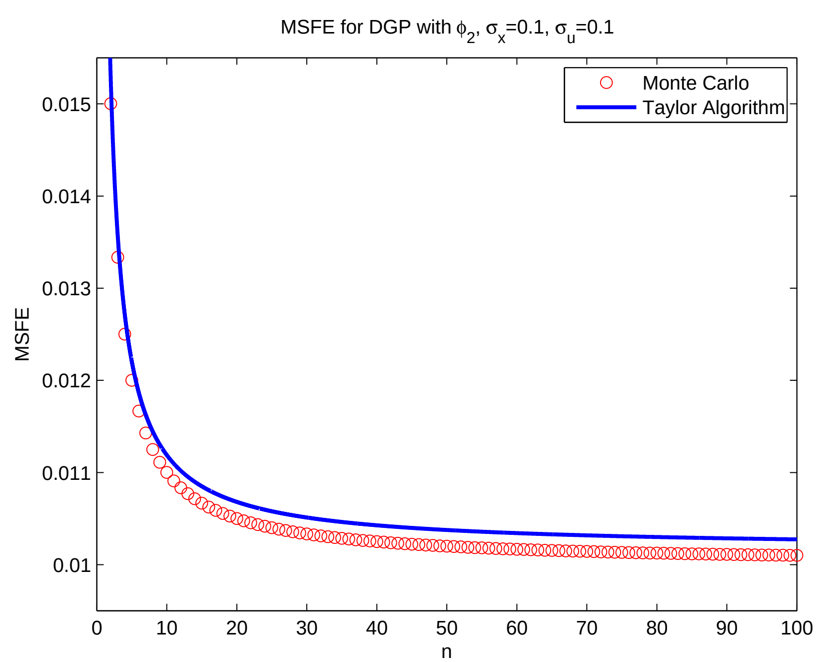


Figure 5.24: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

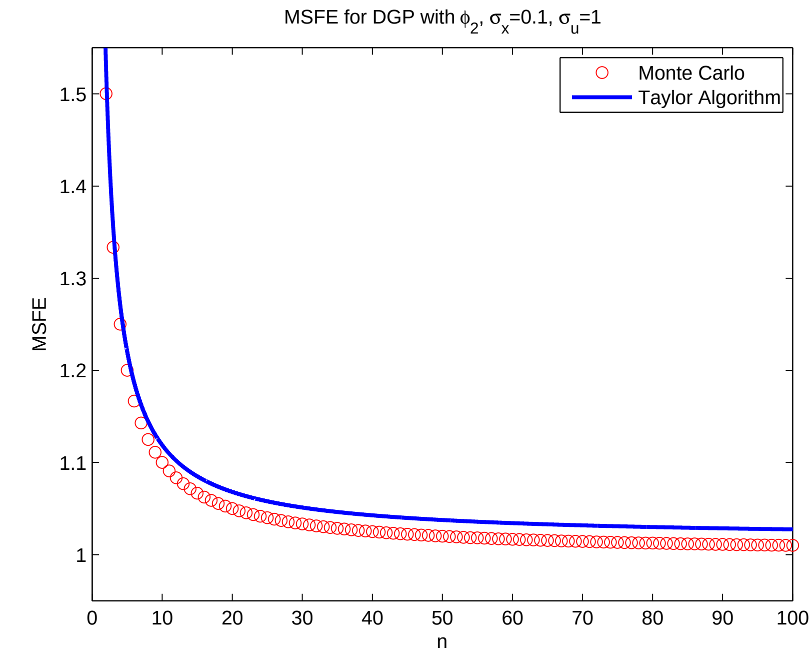


Figure 5.25: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

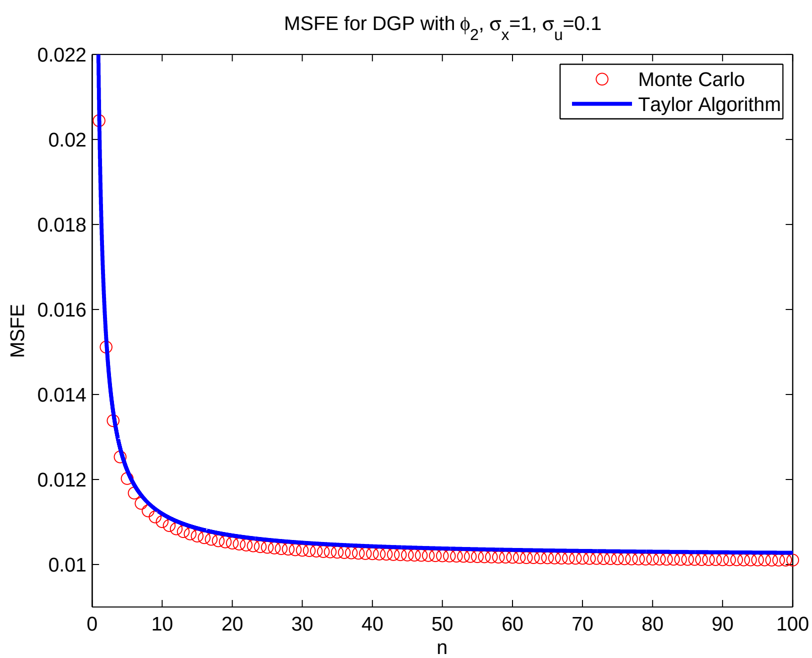


Figure 5.26: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

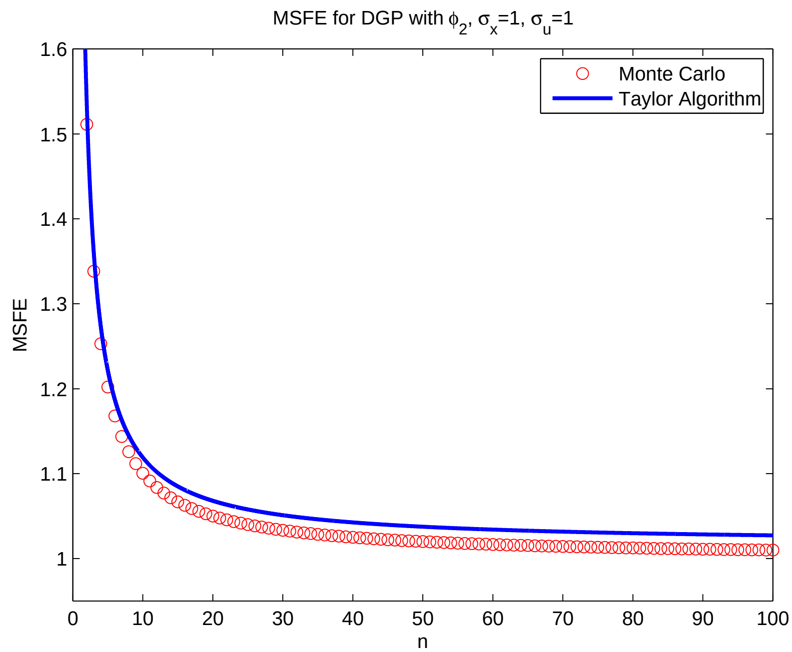


Figure 5.27: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 1$

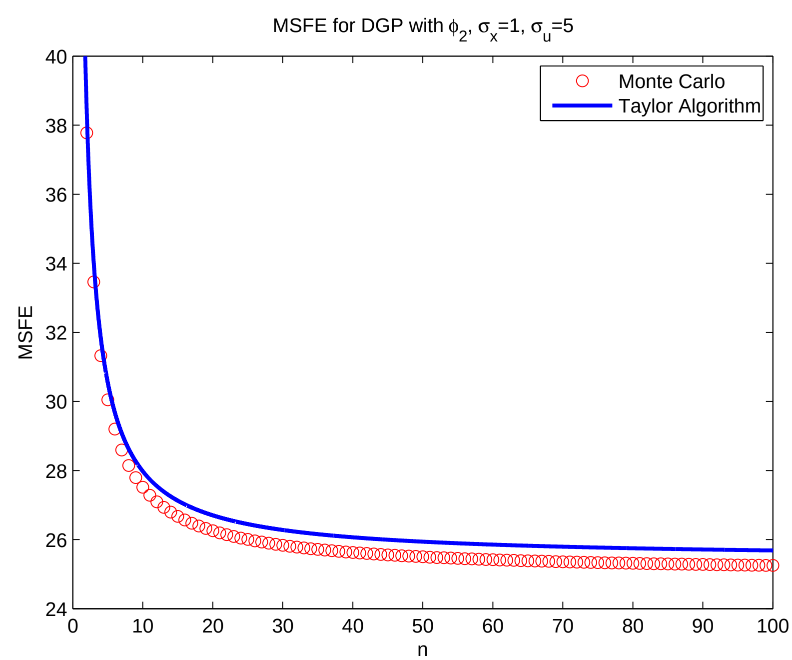


Figure 5.28: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 5$

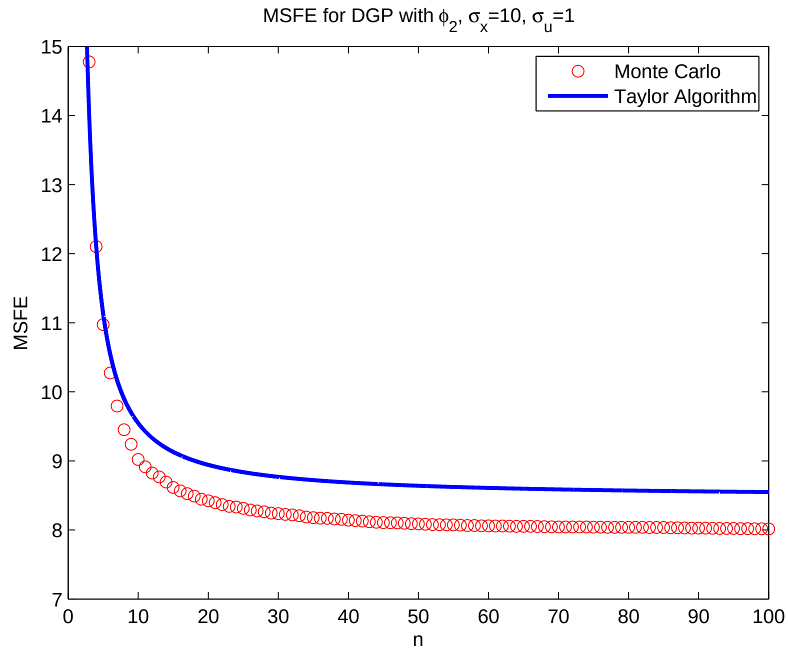


Figure 5.29: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 1$

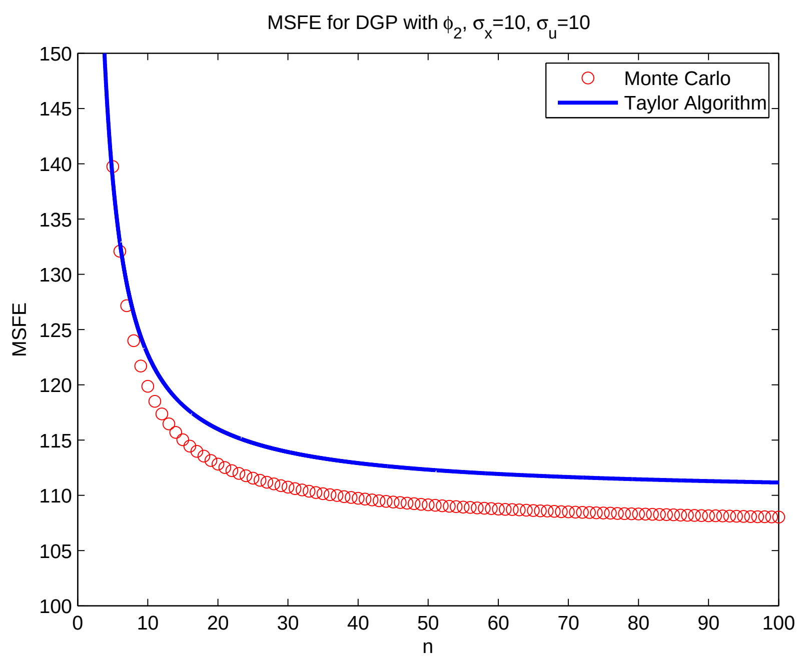


Figure 5.30: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 10$

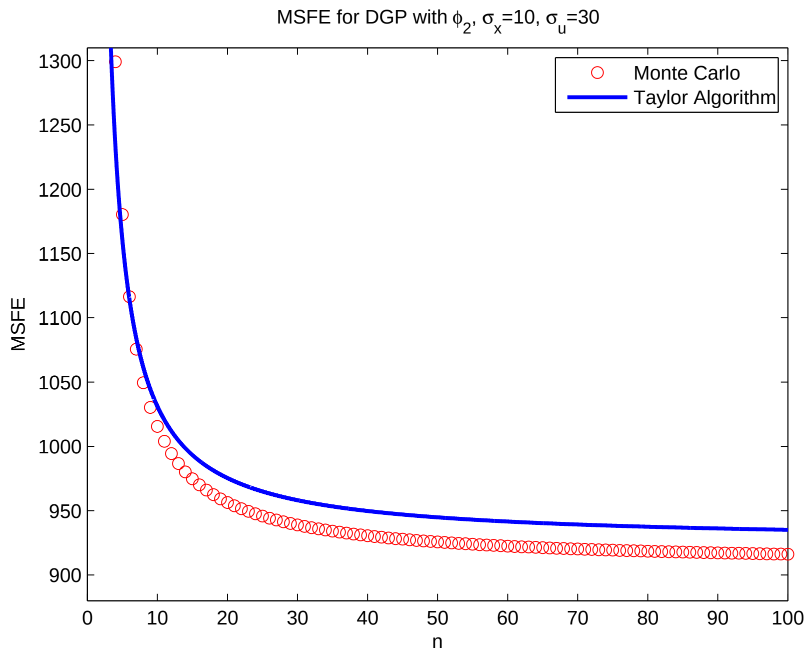


Figure 5.31: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 0.01, \theta_3 = 1, \theta_4 = 0.001$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 30$

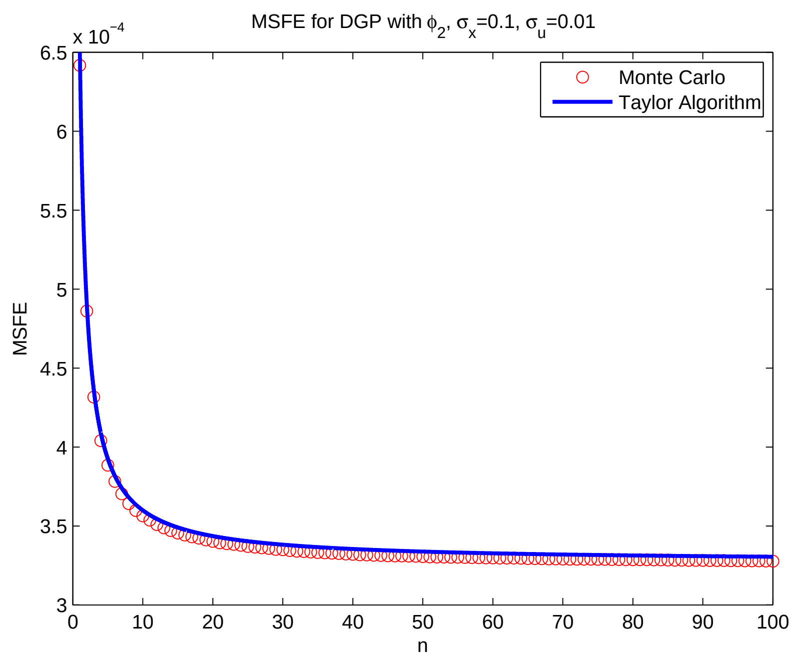


Figure 5.32: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$, $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

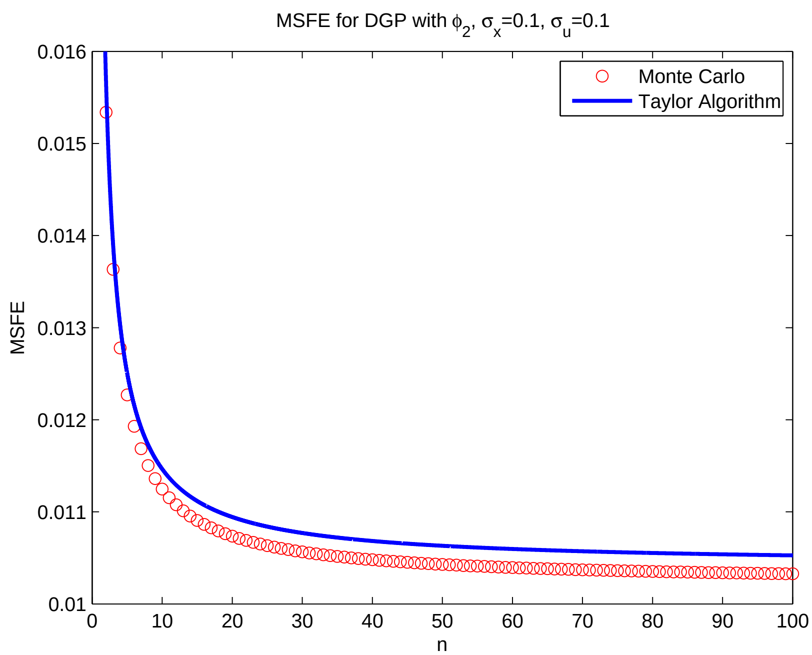


Figure 5.33: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

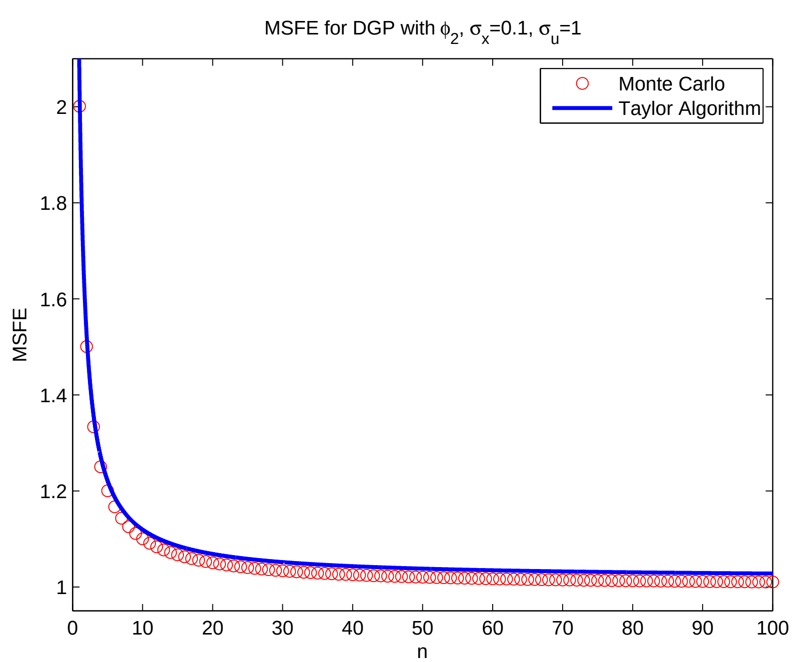


Figure 5.34: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

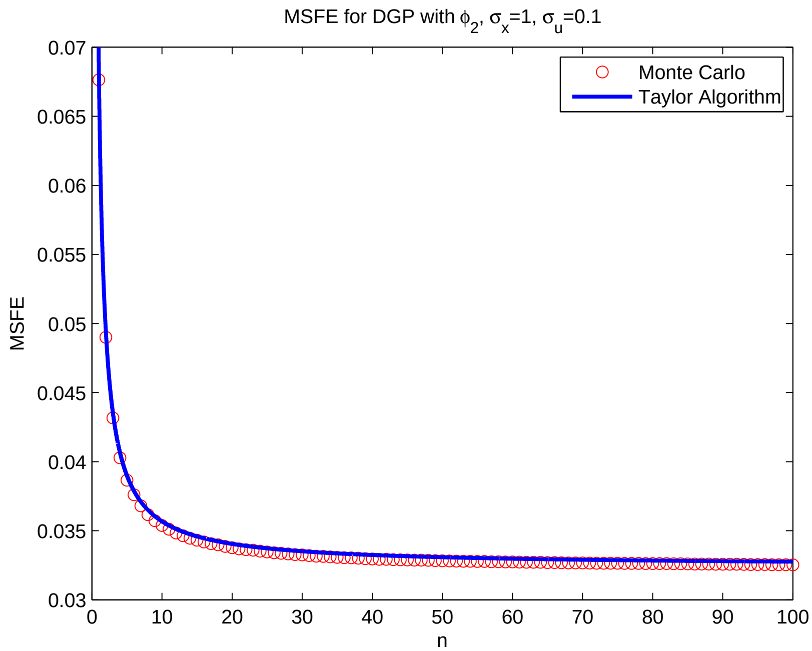


Figure 5.35: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

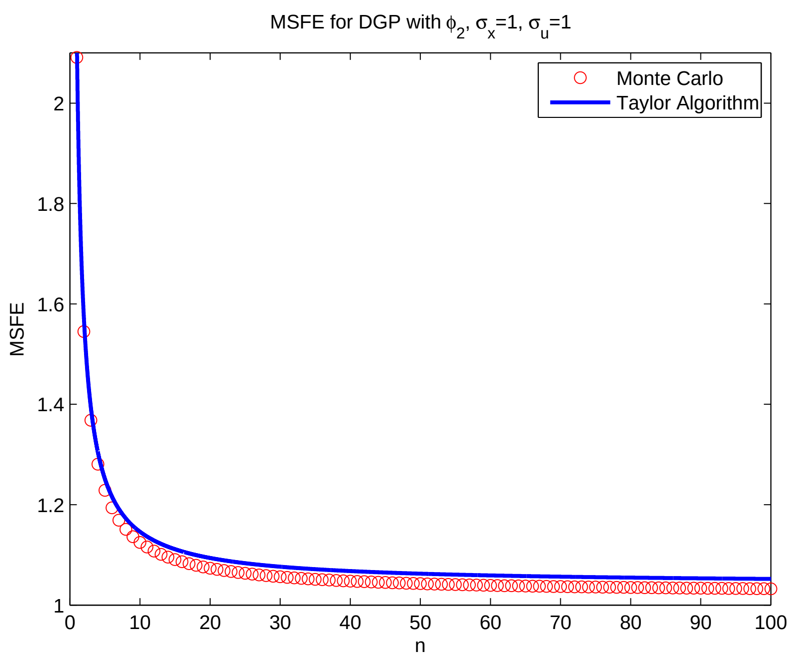


Figure 5.36: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$, $\mu_x = 10, \sigma_x = 1, \sigma_u = 1$

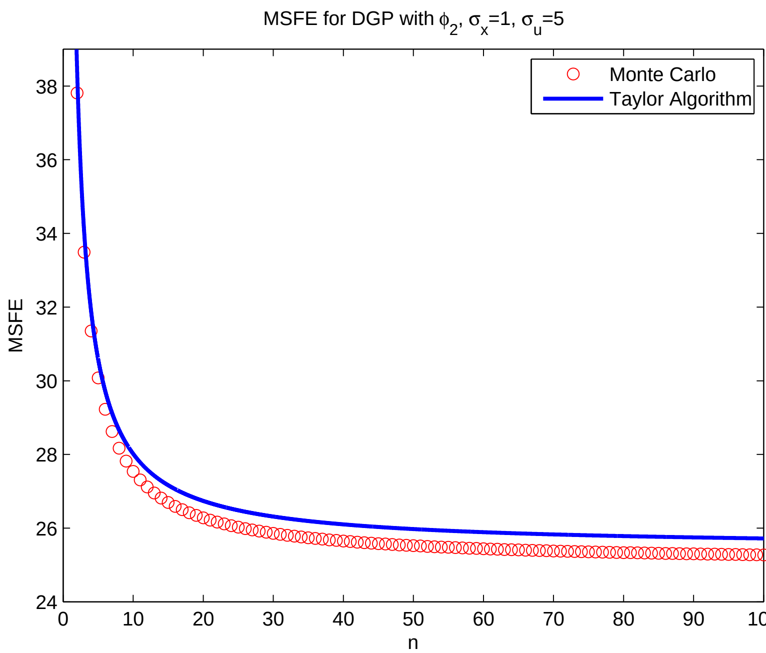


Figure 5.37: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 5$

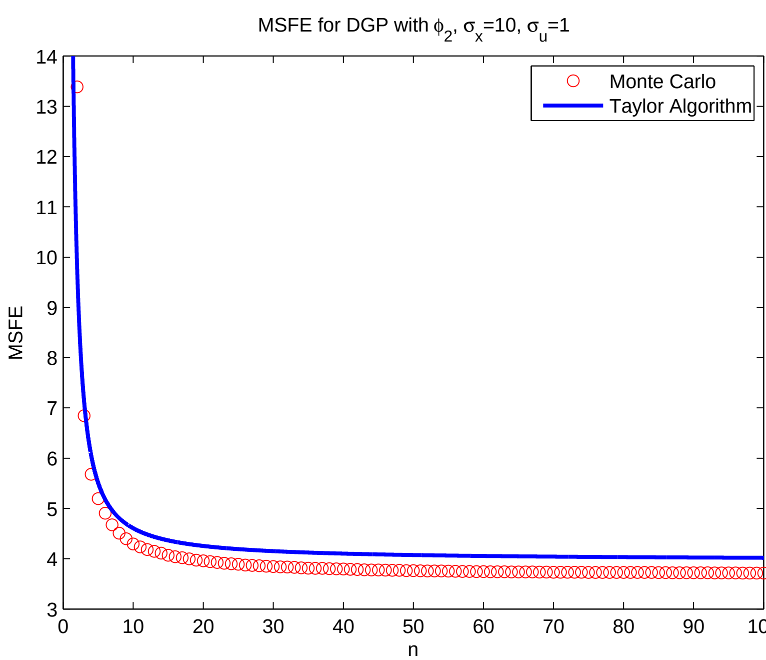


Figure 5.38: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5,\ \mu_x = 10, \sigma_x = 10, \sigma_u = 1$

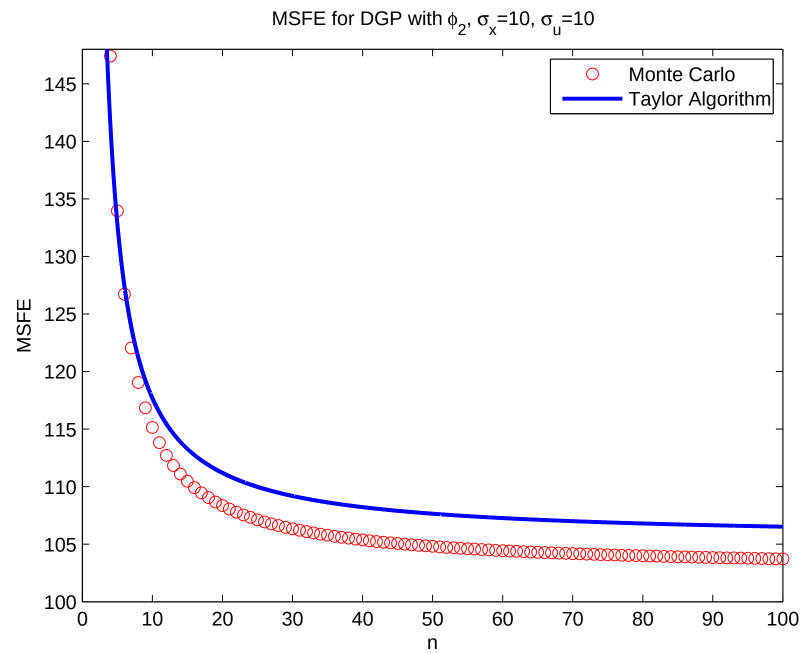


Figure 5.39: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 10$

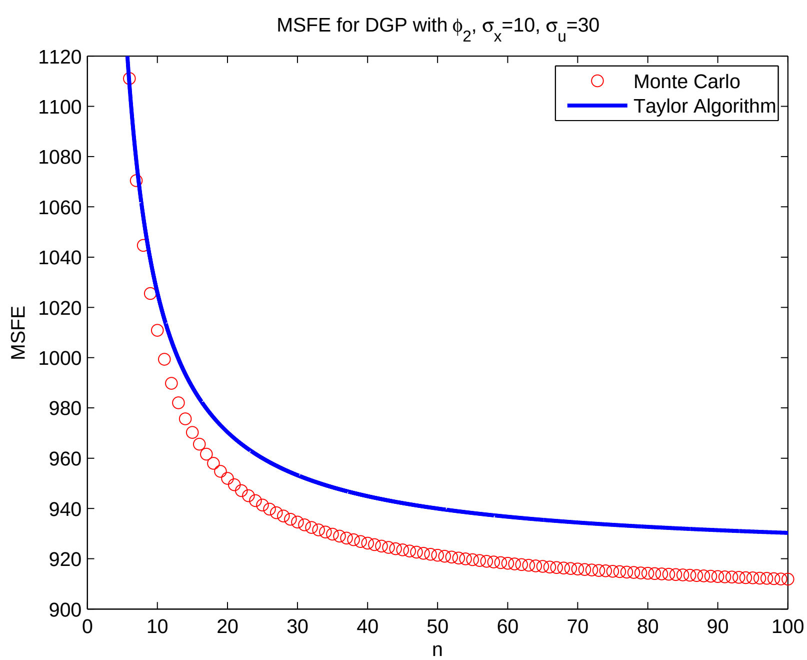


Figure 5.40: MSFE for $\varphi_2(x)$, $\theta_1 = 1, \theta_2 = 2.5, \theta_3 = 1, \theta_4 = 1.5$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 30$

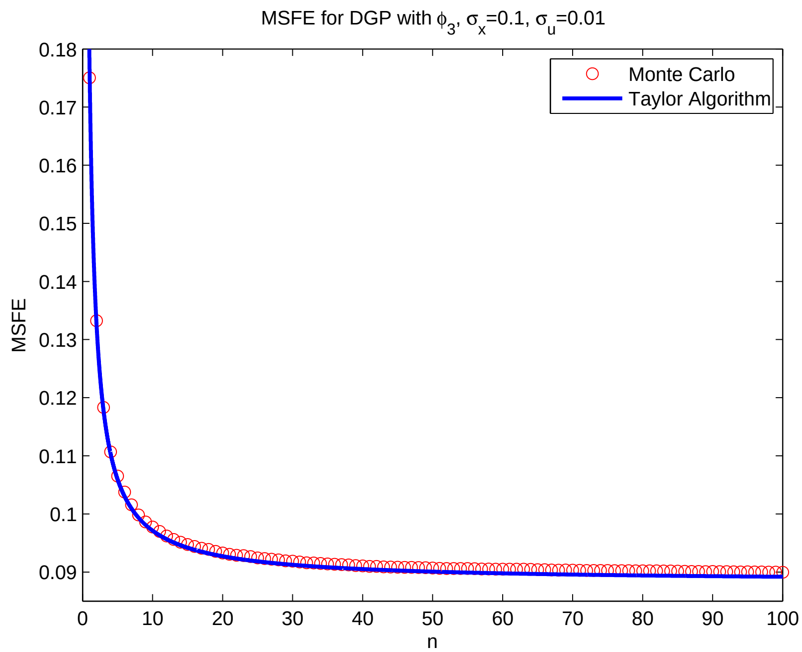


Figure 5.41: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

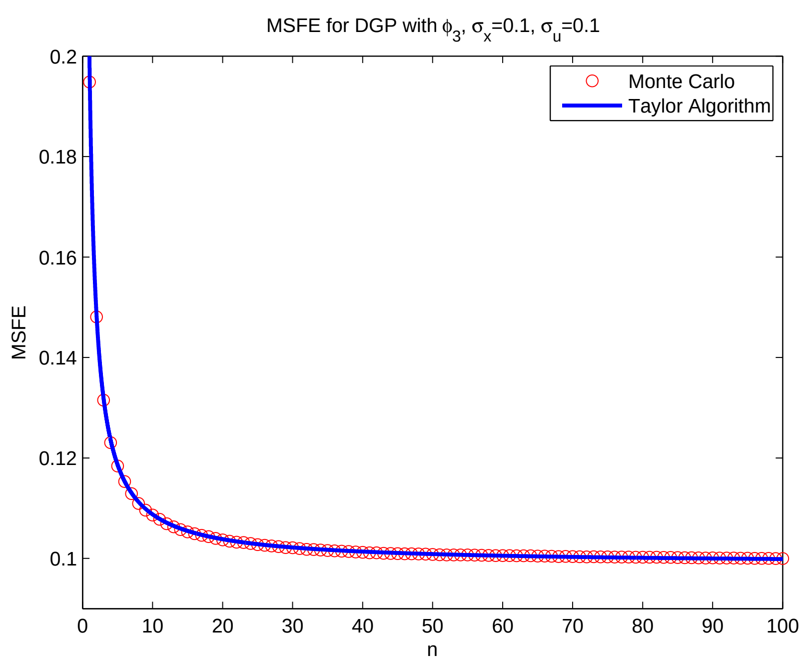


Figure 5.42: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

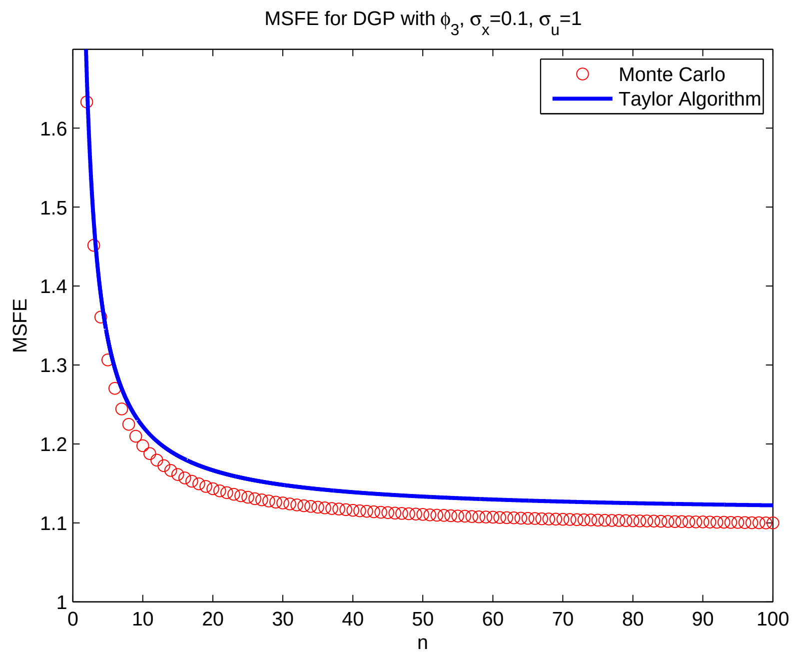


Figure 5.43: MSFE for $\varphi_3(x), \theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1, \mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

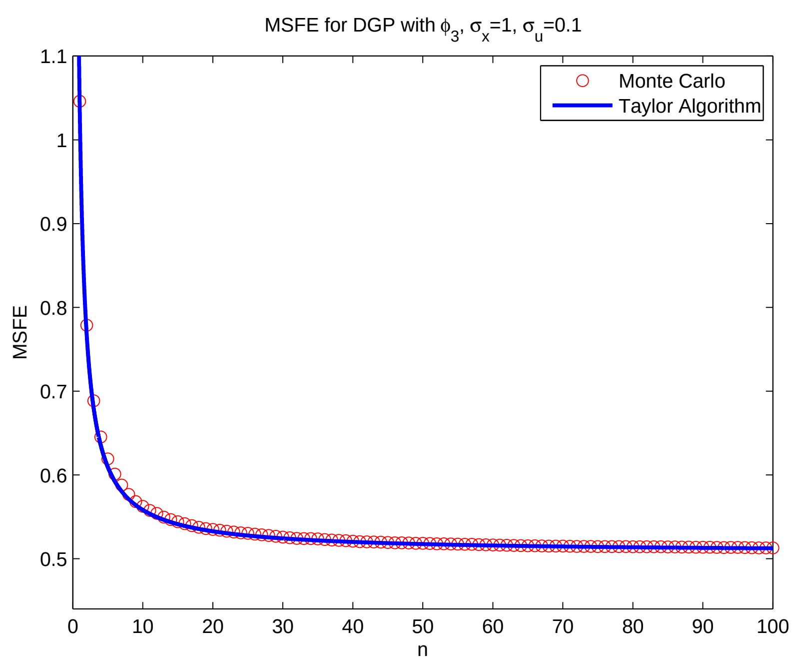


Figure 5.44: MSFE for $\varphi_3(x), \theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1, \mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

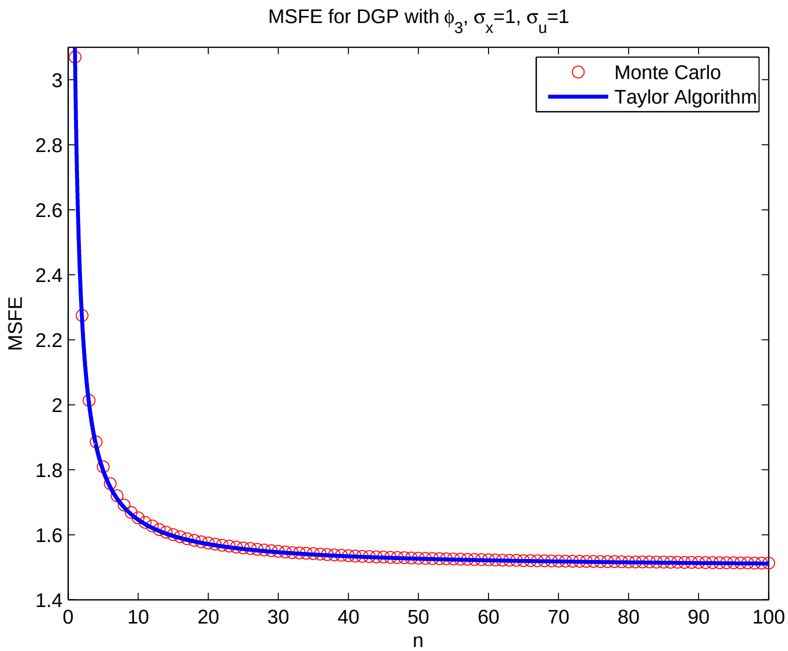


Figure 5.45: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 1$

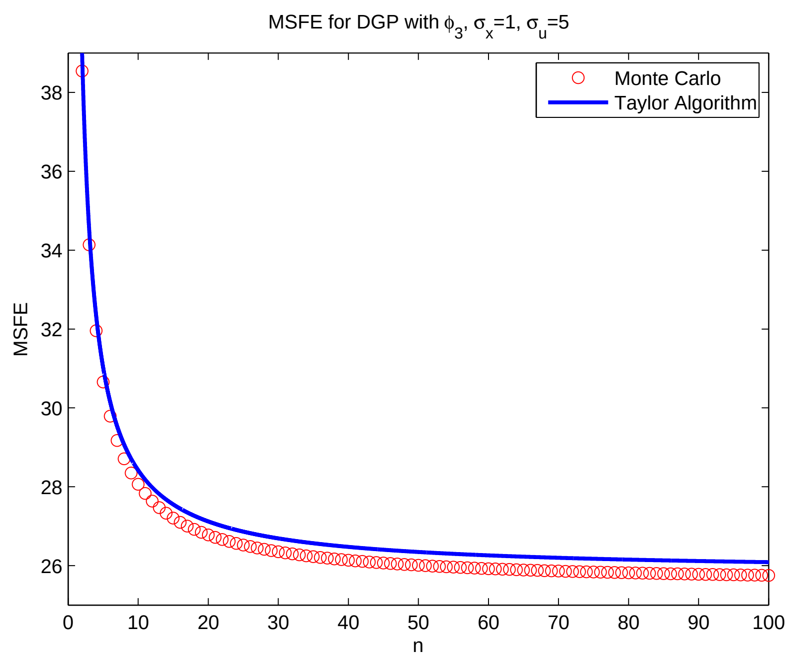


Figure 5.46: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 5$

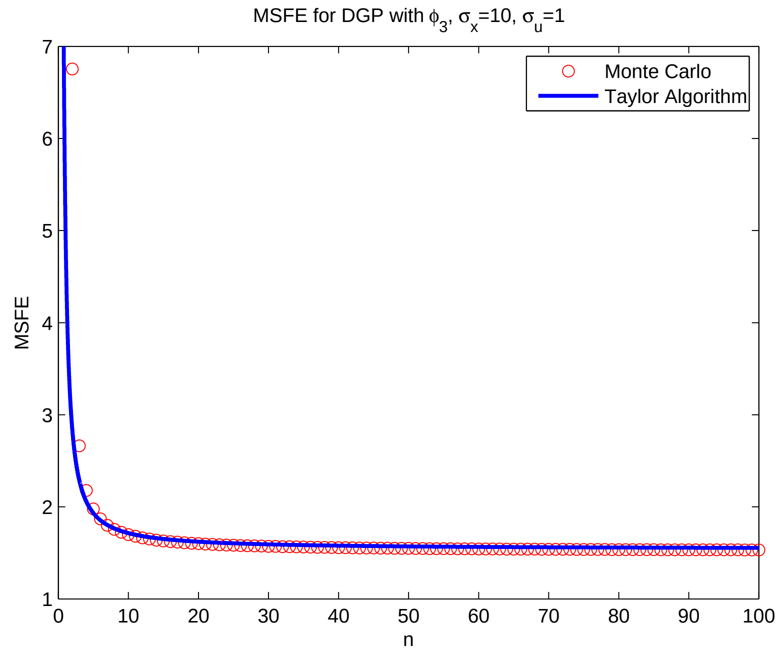


Figure 5.47: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 10, \sigma_u = 1$

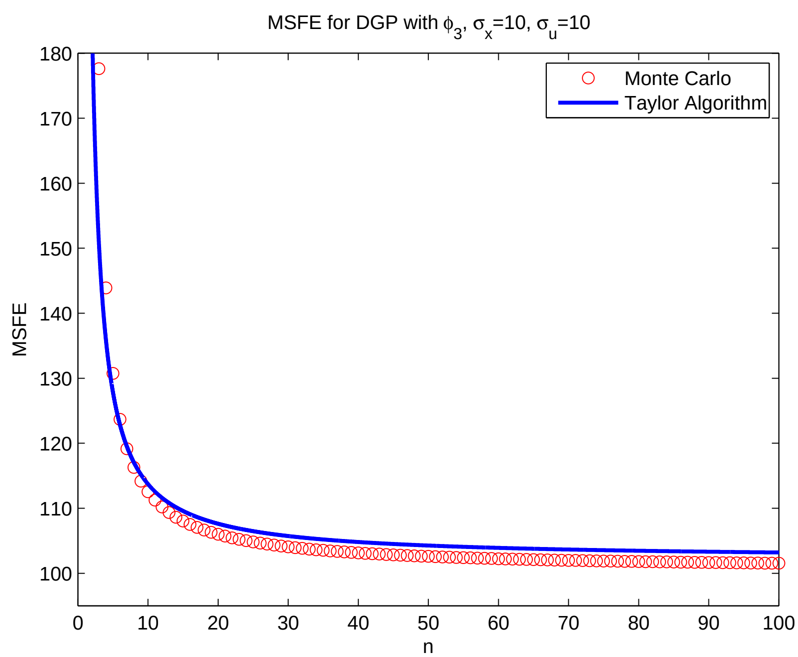


Figure 5.48: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 10, \sigma_u = 10$

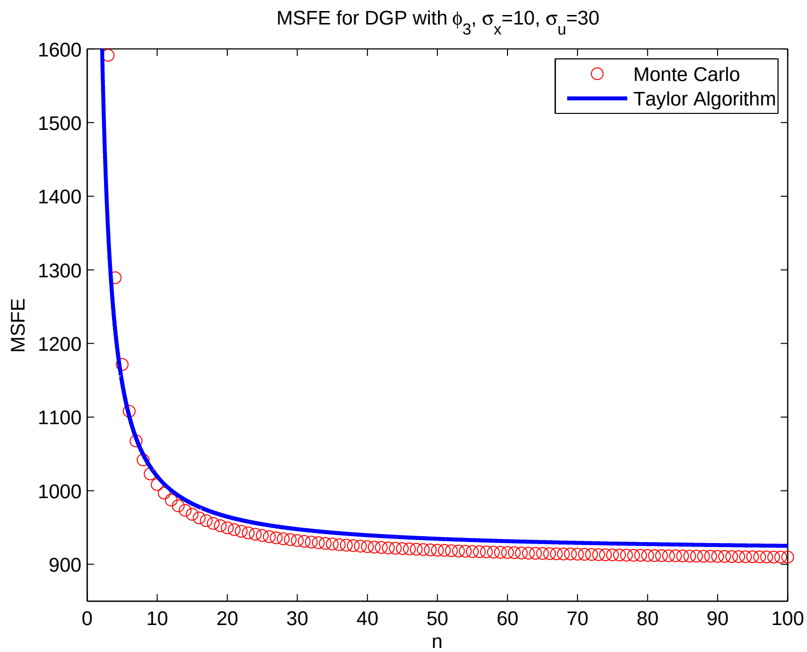


Figure 5.49: MSFE for $\varphi_3(x)$, $\theta_1 = 1, \theta_2 = 0.001, \theta_3 = 1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 10, \sigma_u = 30$

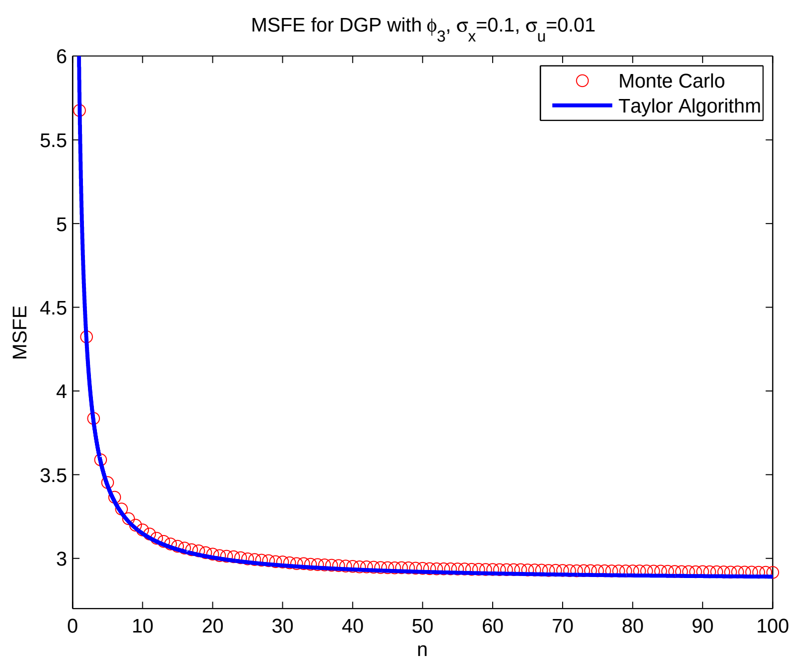


Figure 5.50: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.01$

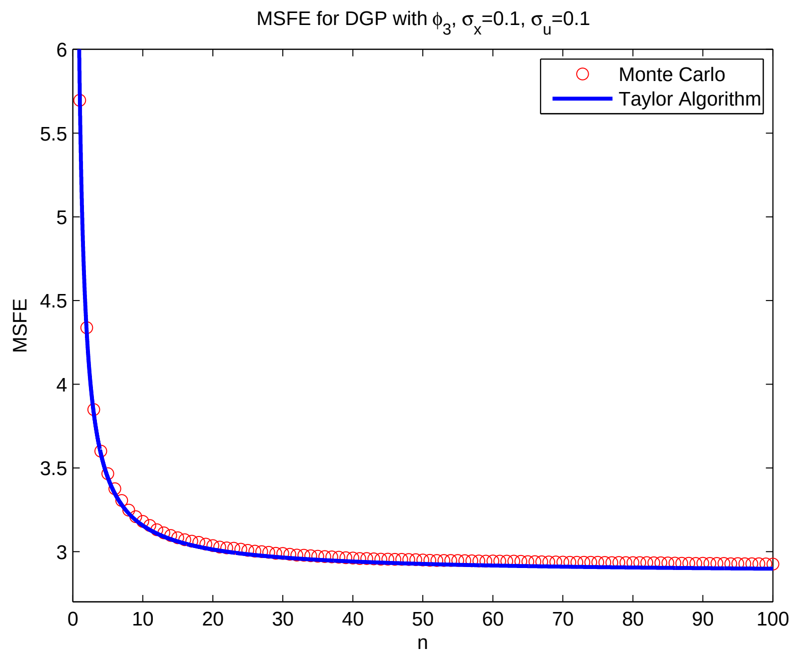


Figure 5.51: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,$ $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 0.1$

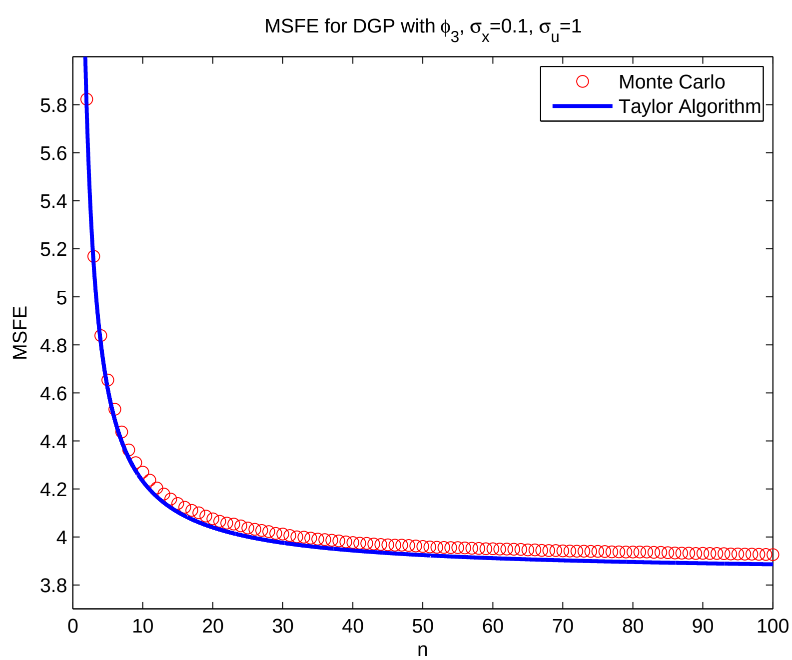


Figure 5.52: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,$ $\mu_x = 10, \sigma_x = 0.1, \sigma_u = 1$

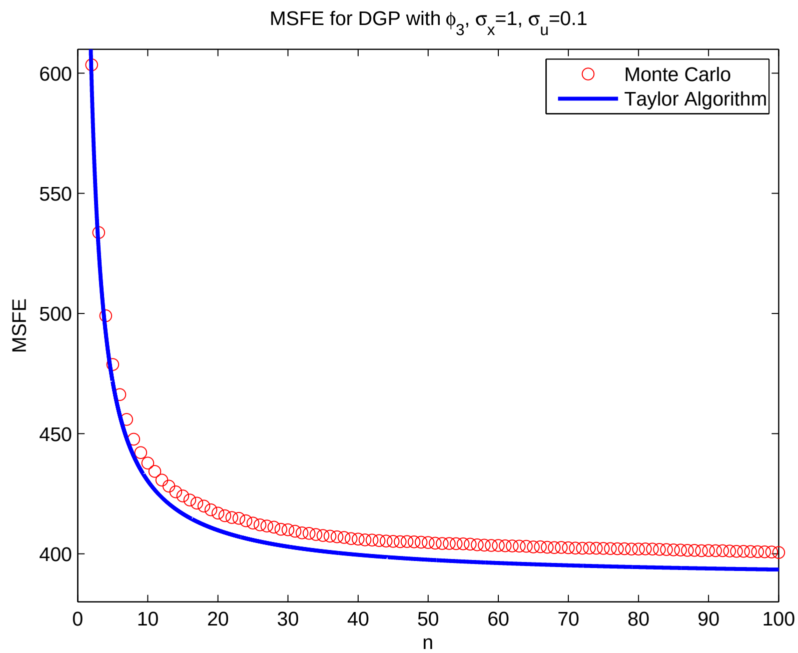


Figure 5.53: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 0.1$

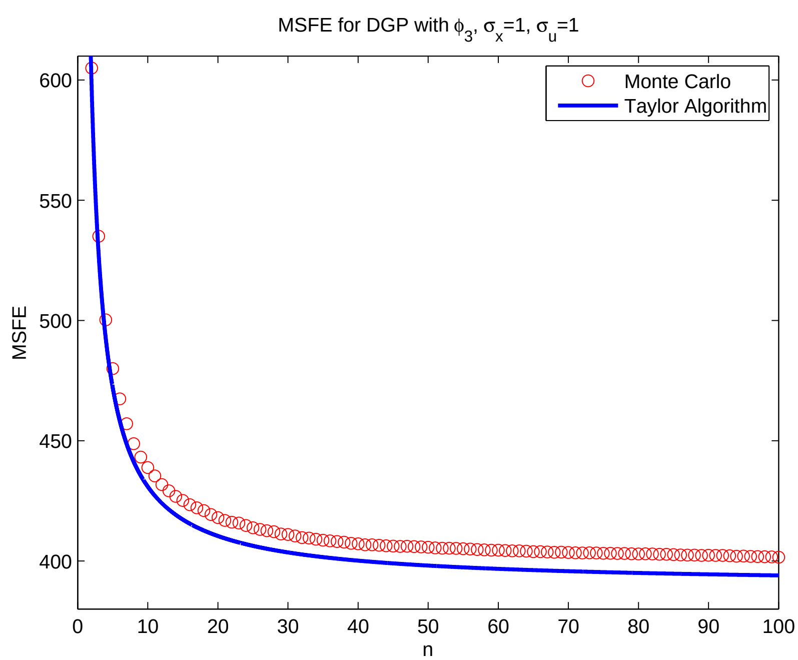


Figure 5.54: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 1$

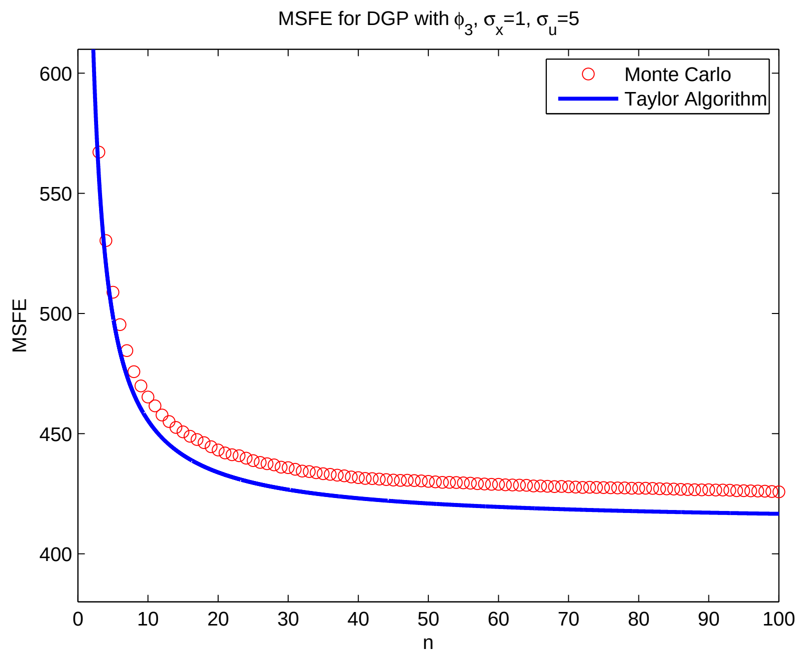


Figure 5.55: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 1, \sigma_u = 5$

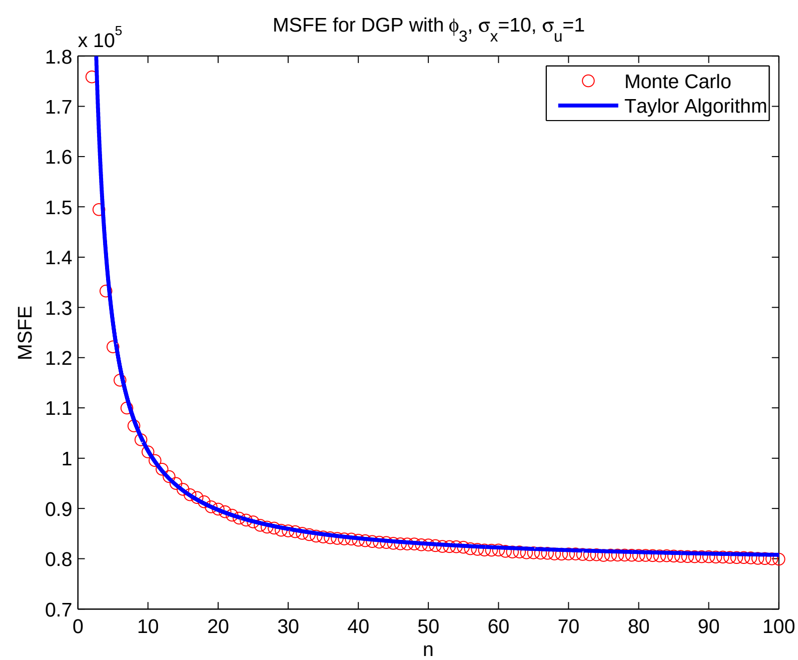


Figure 5.56: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1,\ \mu_x = 10, \sigma_x = 10, \sigma_u = 1$

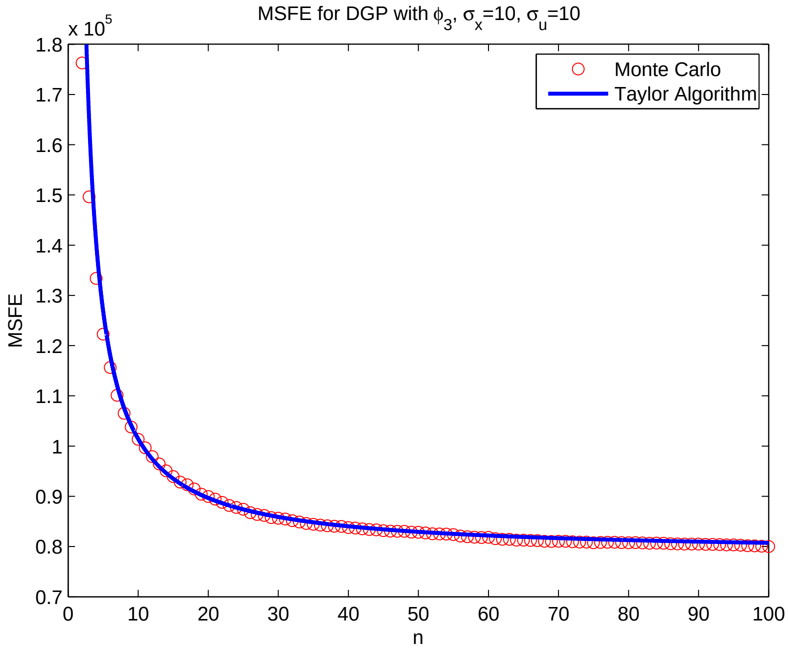


Figure 5.57: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 10$

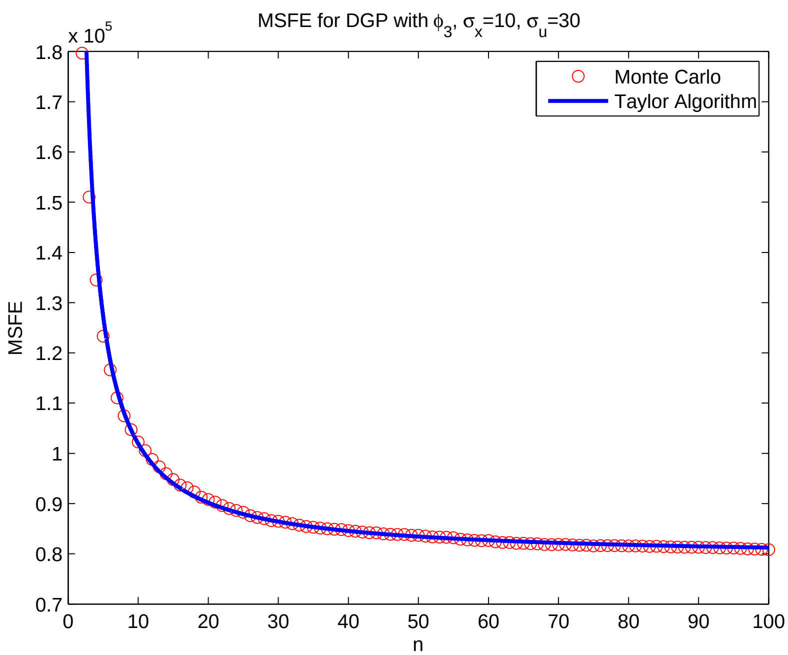


Figure 5.58: MSFE for $\varphi_3(x)$, $\theta_1 = 0.1, \theta_2 = 2, \theta_3 = 0.1, \theta_4 = 1$, $\mu_x = 10, \sigma_x = 10, \sigma_u = 30$

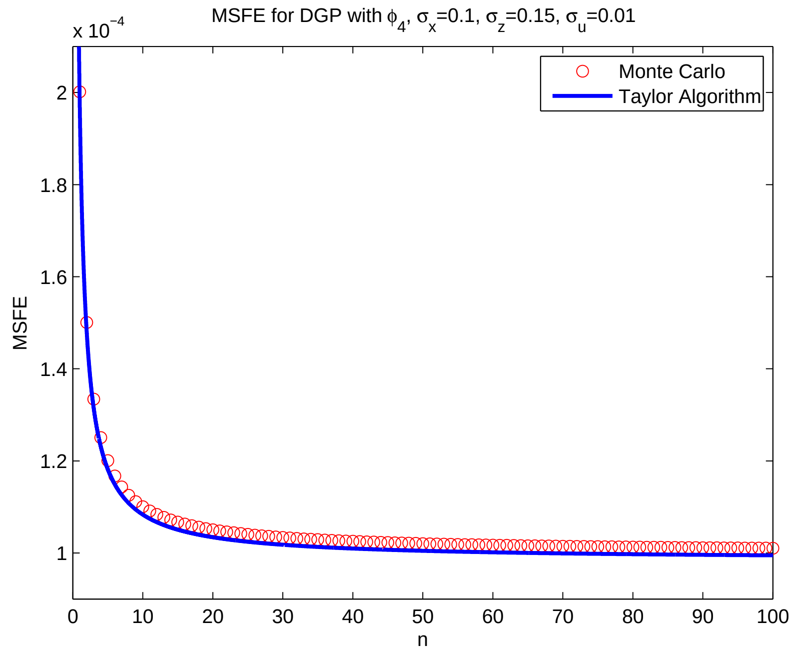


Figure 5.59: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 0.01$

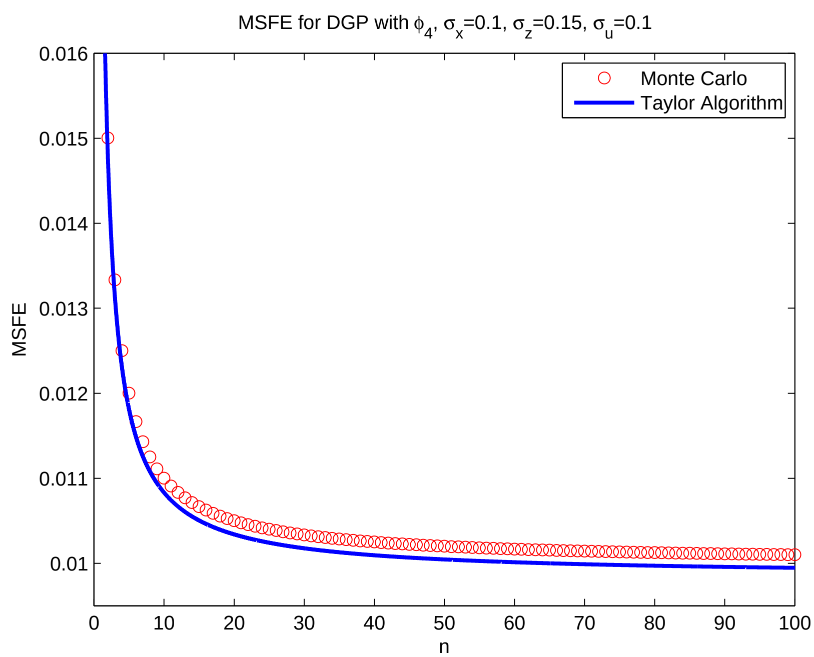


Figure 5.60: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 0.1$

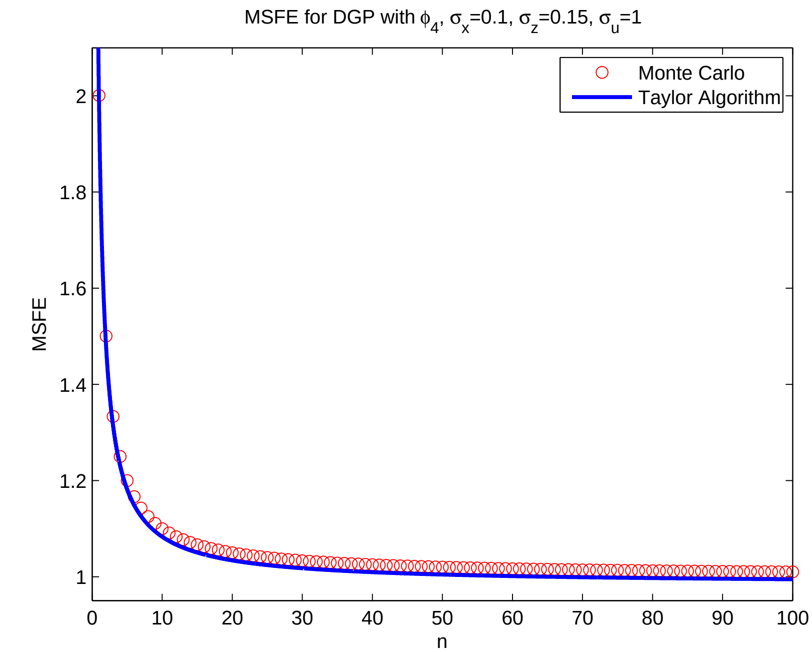


Figure 5.61: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 1$

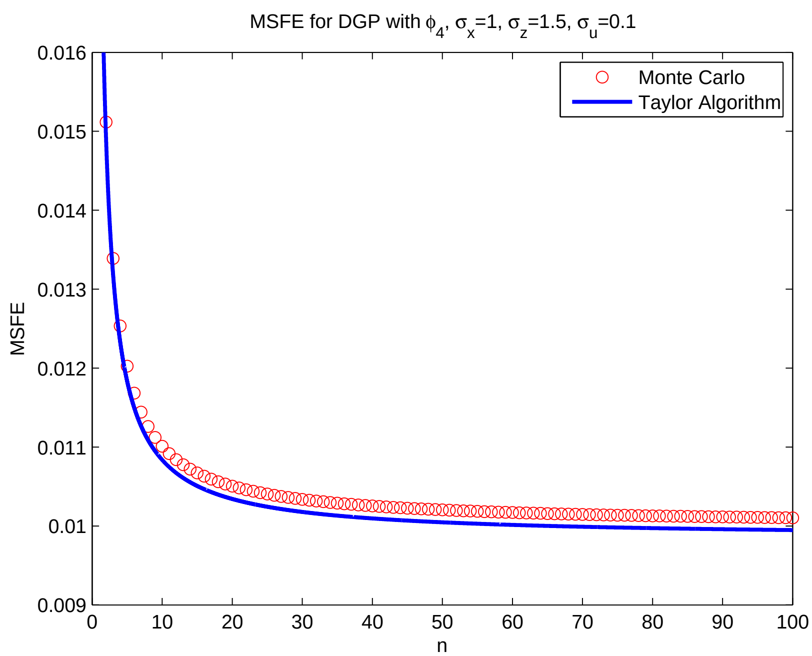


Figure 5.62: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 0.1$

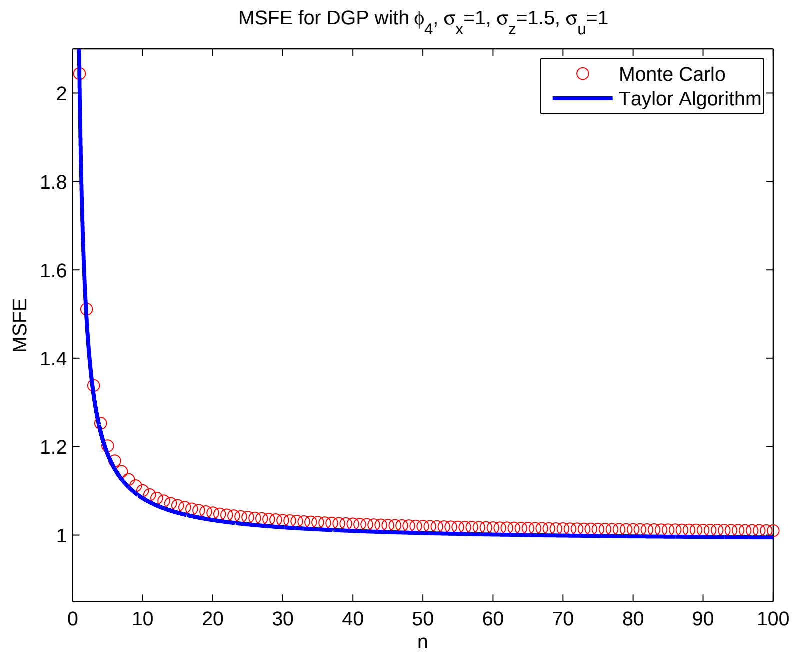


Figure 5.63: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001,\ \mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 1$

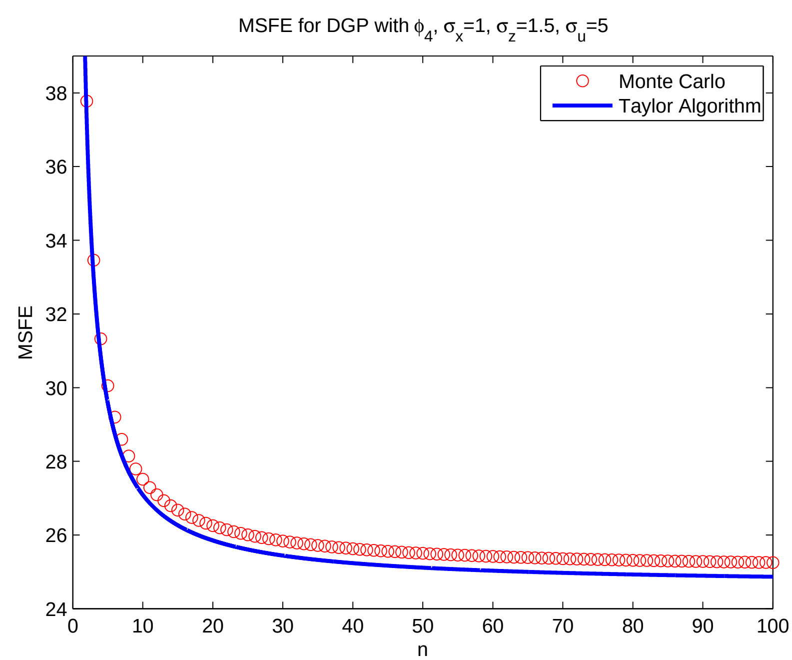


Figure 5.64: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001,\ \mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 5$

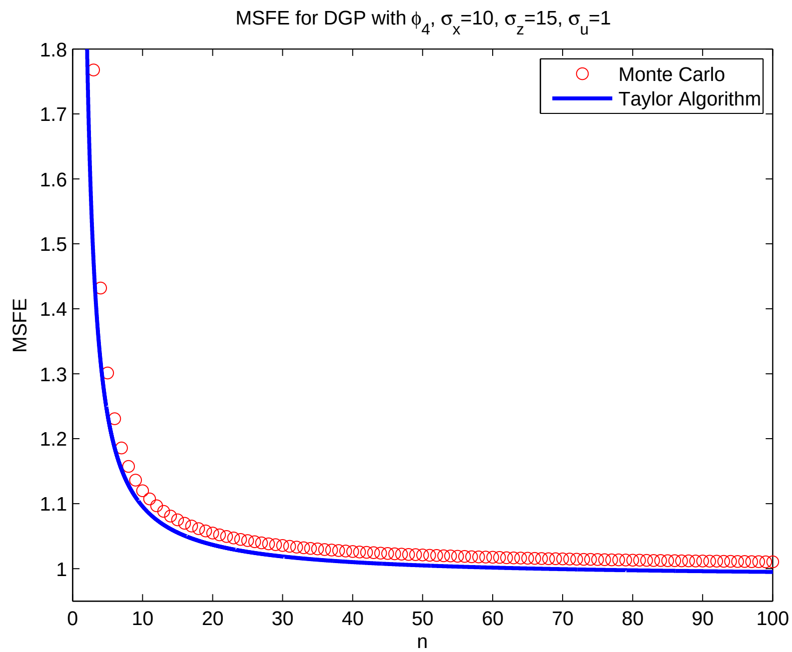


Figure 5.65: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 1$

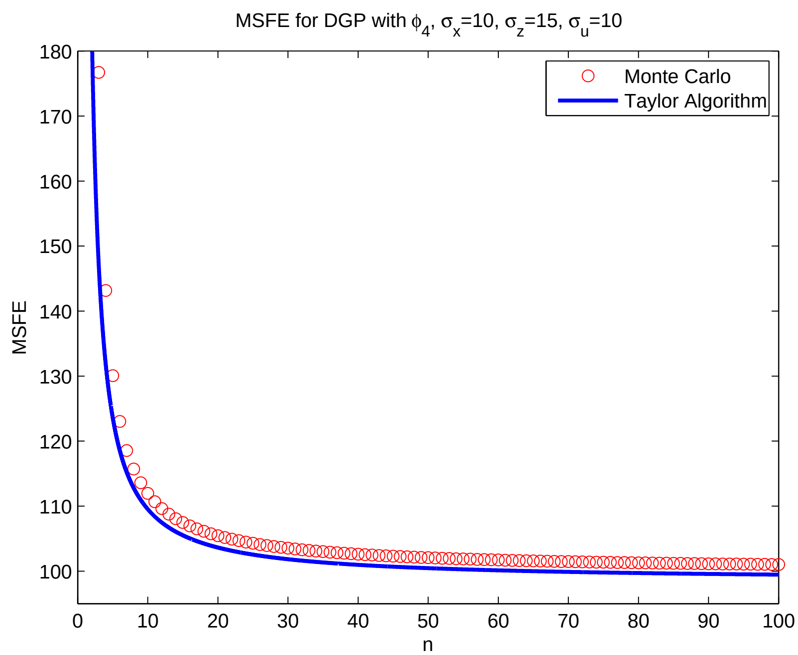


Figure 5.66: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 10$

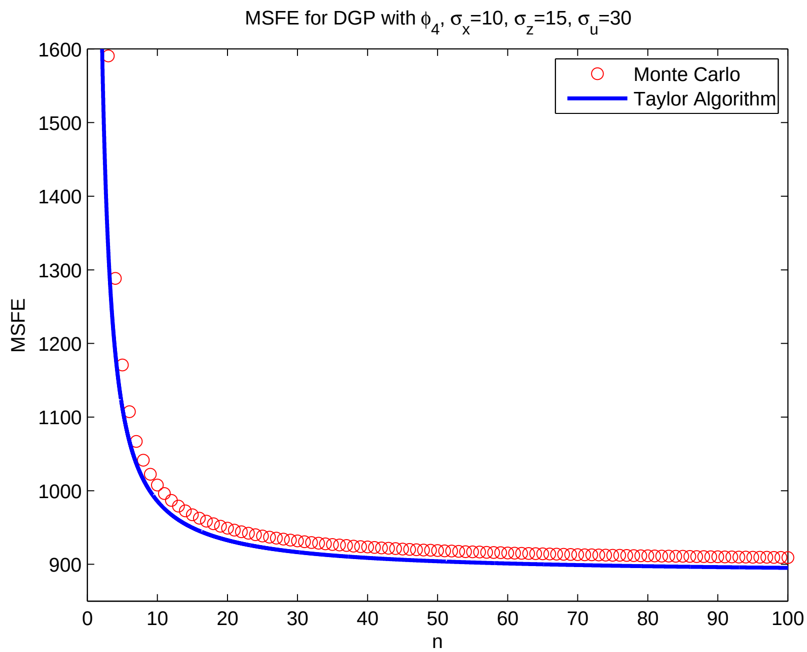


Figure 5.67: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 0.001$, $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 30$

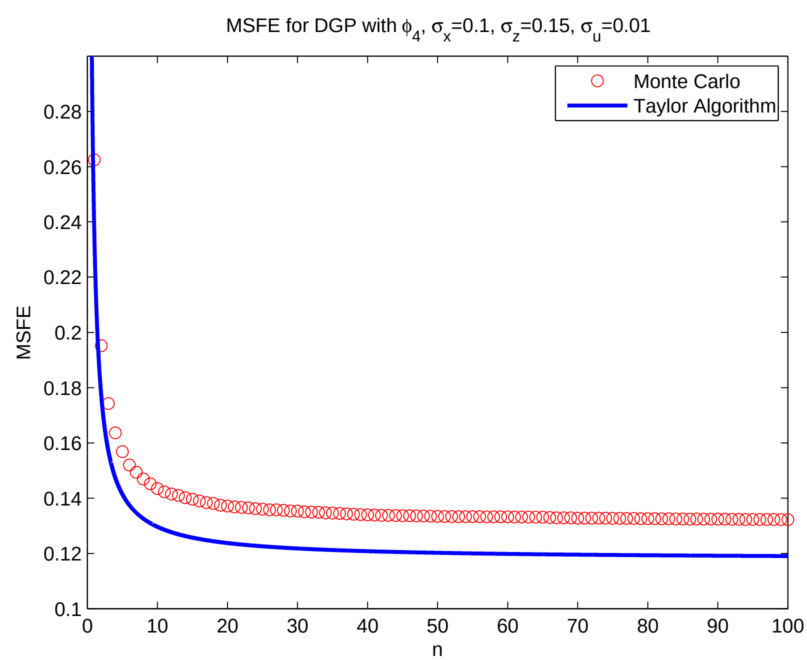


Figure 5.68: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 0.01$

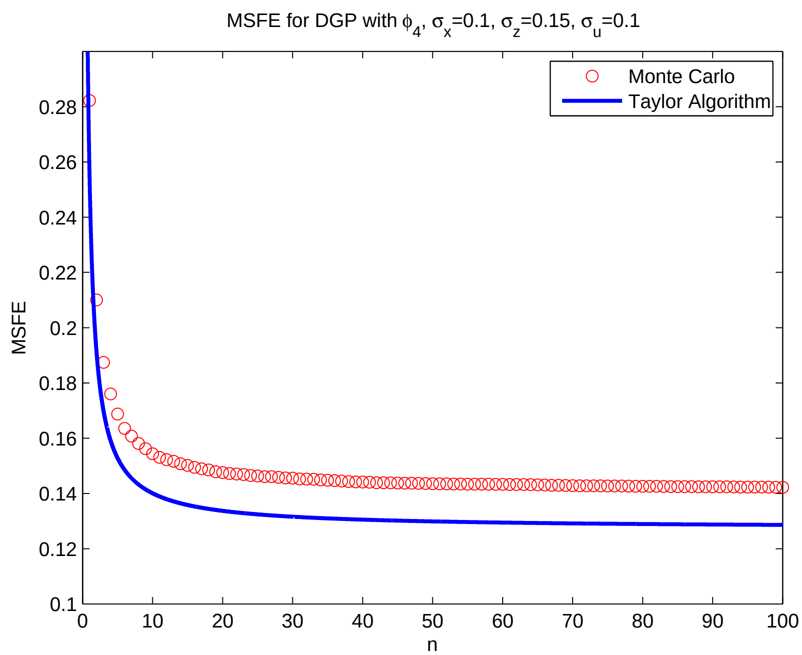


Figure 5.69: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 0.1$

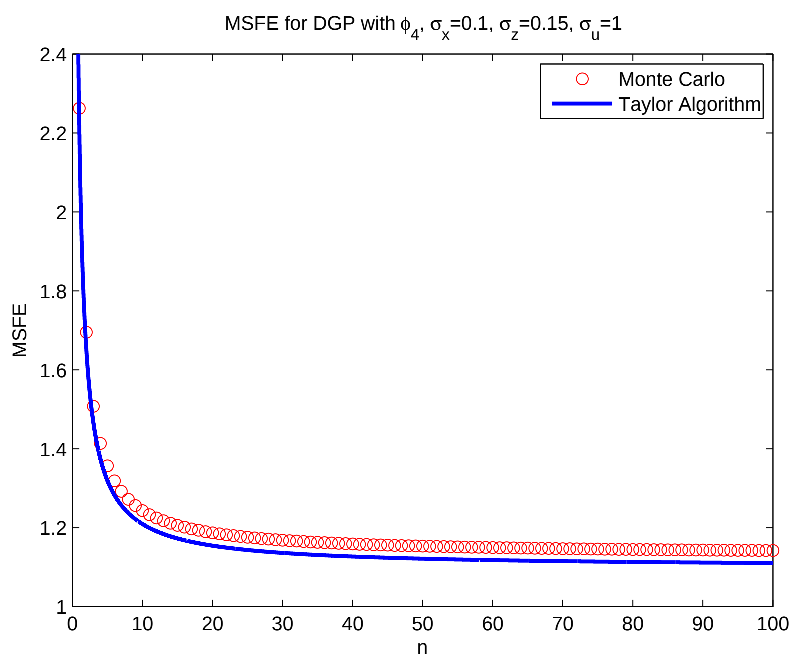


Figure 5.70: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 0.1, \sigma_z = 0.15, \sigma_u = 1$

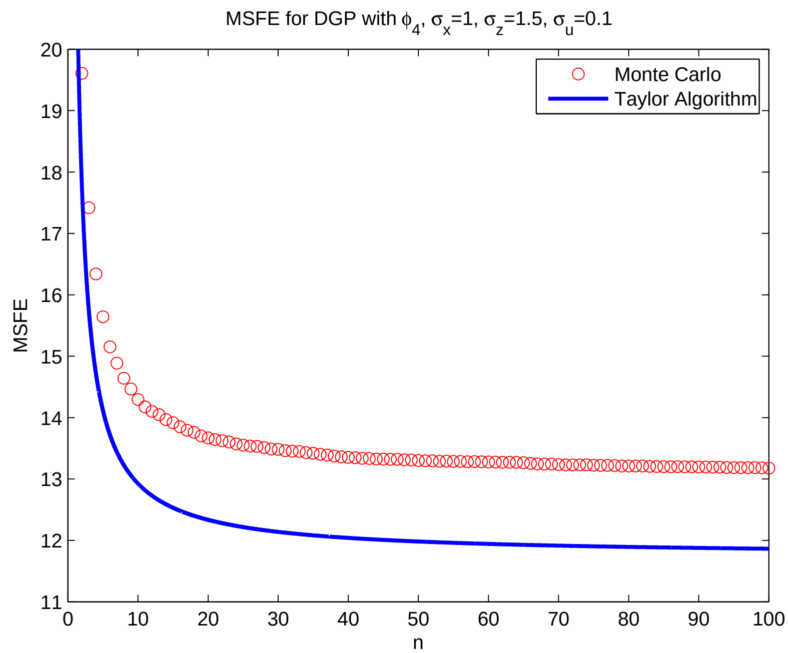


Figure 5.71: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 0.1$

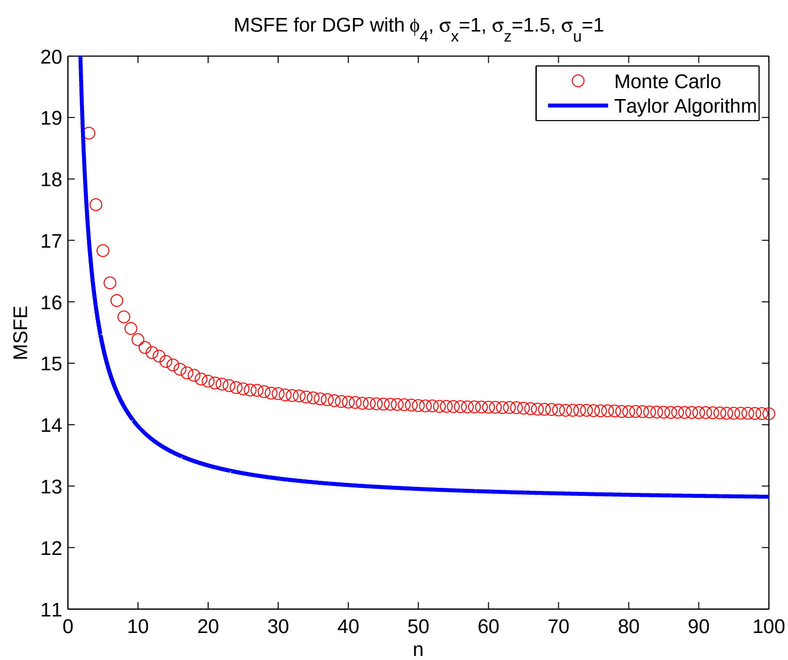


Figure 5.72: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 1$

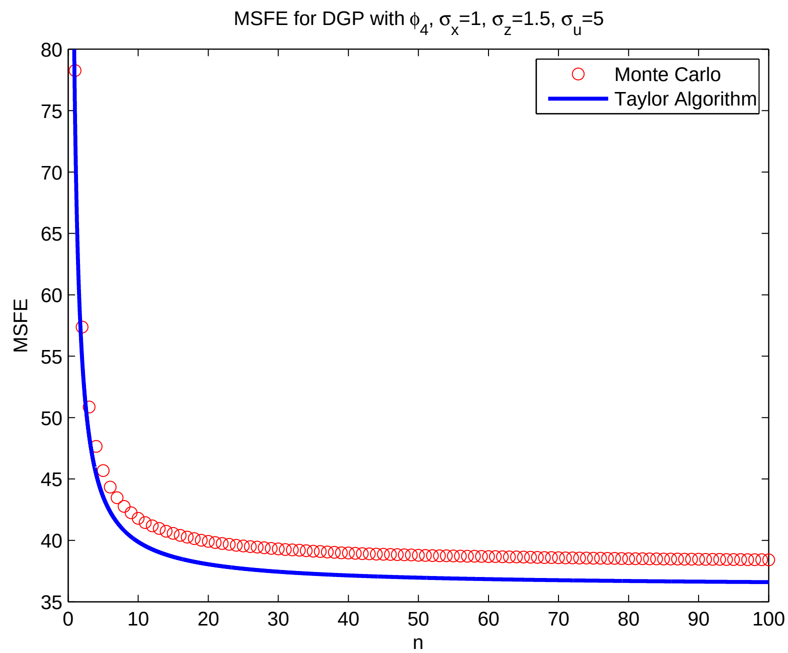


Figure 5.73: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2,$ $\mu_x = 10, \mu_z = 8, \sigma_x = 1, \sigma_z = 1.5, \sigma_u = 5$

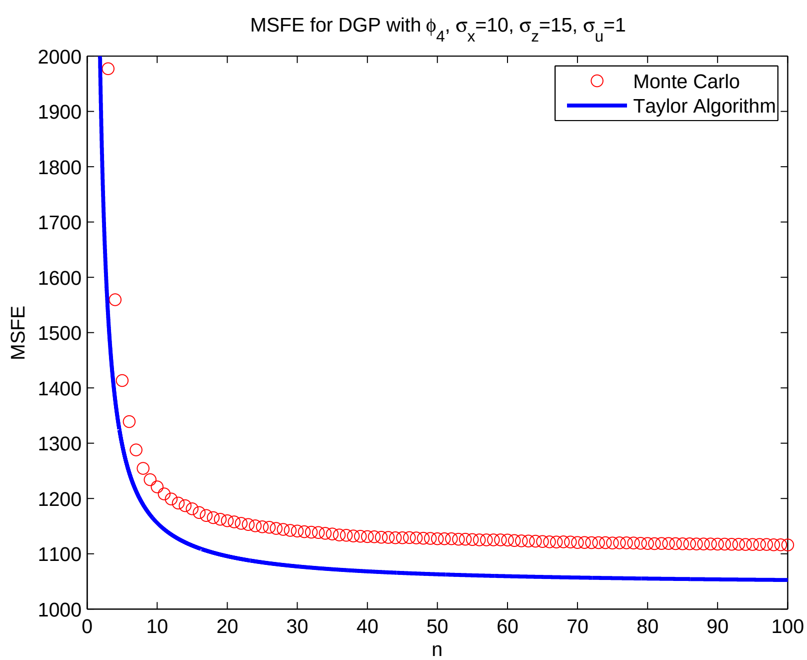


Figure 5.74: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2,$ $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 1$

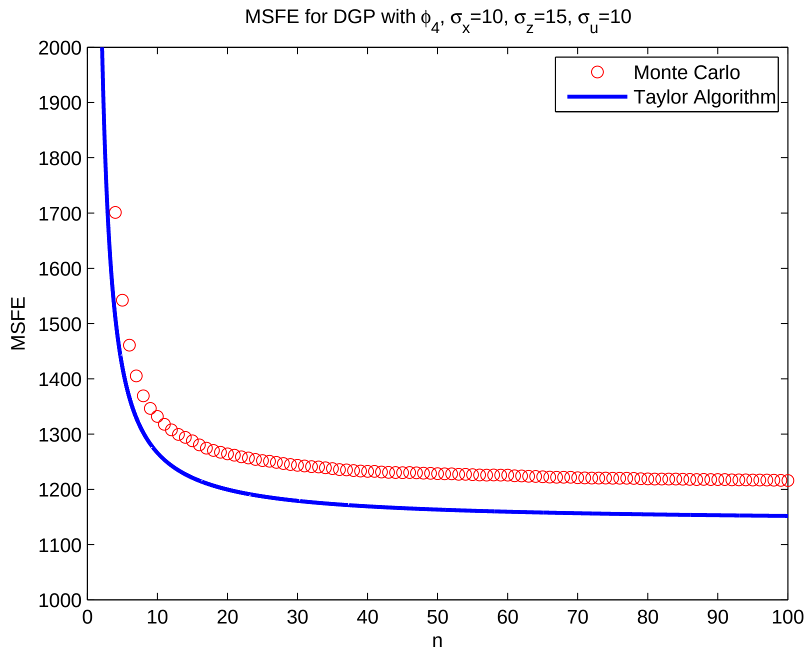


Figure 5.75: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 10$

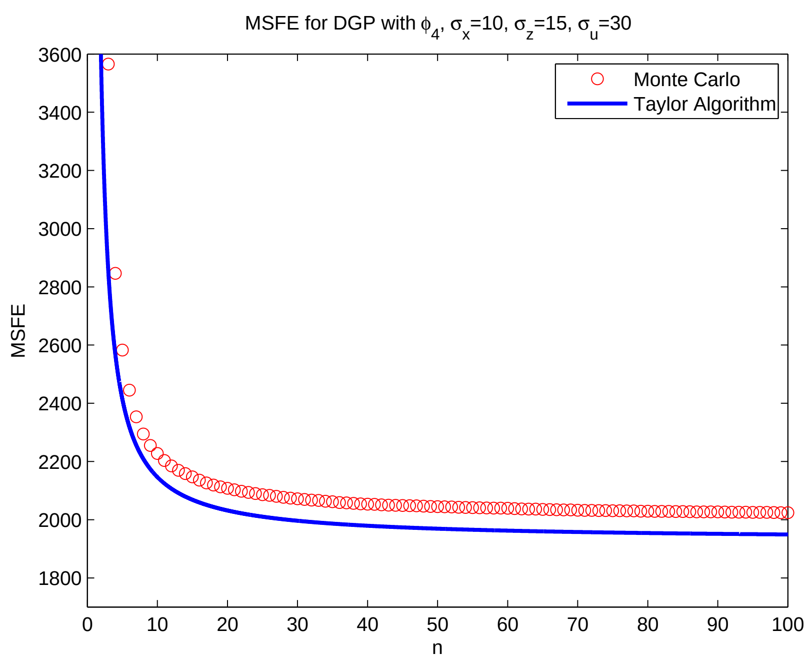


Figure 5.76: MSFE for $\varphi_4(x)$, $\theta_1 = 1, \theta_2 = 2$, $\mu_x = 10, \mu_z = 8, \sigma_x = 10, \sigma_z = 15, \sigma_u = 30$

Chapter 6

Taylor algorithm for stationary processes

6.1 Introduction

In Chapter 5, we studied a forecasting problem involving independent and identically distributed (i.i.d.) processes. In the present chapter, we allow for more general processes and construct an algorithm which yields an approximation, based on Taylor series, of the mean square forecast error (MSFE) for a forecasting problem involving stationary processes. This Taylor algorithm approximation is meant to be used as a tool to describe the sample size dependence (SSD) of the MSFE. We begin by defining two types of stationarity.

Definition 6.1 *Let G_1 be the joint distribution function of the sequence $\{Z_1, Z_2, \dots\}$, where Z_τ is a $q \times 1$ vector, and let G_{t+1} be the joint distribution function of the sequence $\{Z_{t+1}, Z_{t+2}, \dots\}$. The sequence $\{Z_\tau\}$ is strictly stationary if $G_1 = G_{t+1}$ for each $t \geq 1$.*

Definition 6.2 *If a sequence has constant variance and has covariances that depend only on the time lag between Z_t and $Z_{t+\tau}$, the sequence is said to be covariance stationary.*

Clearly, every strictly stationary process is covariance stationary but not vice versa, and an i.i.d. process is both strictly stationary and covariance stationary. To encompass as many different dependencies of stationary processes as possible, the algorithm developed in this chapter assumes covariance stationarity of the processes.

As is evident from the motivating examples given in Section 2.6, one of the possible ramifications of the presence of model misspecification is the existence of optimal observation windows for the problem of forecasting. In Chapter 5, results from the approximation of the MSFE by the Taylor algorithm and from the benchmark MSFE obtained with Monte Carlo simulations suggest no optimal observation window exists for the functional misspecifications studied. These experiments were carried out under the assumption that the processes in question were temporally independent. One can attribute the fact that no optimal observation windows exist under misspecification to the static nature of those processes.

The rest of the chapter is organized as follows. In Section 6.2, we present the only relevant results in the literature concerning estimation under misspecification with dependent observations. These consist of some large sample results for the OLS under assumptions of misspecification. In Section 6.3, we construct an algorithm to study the effects of model misspecification on the SSD of the MSFE for the forecasting problem involving covariance stationary processes. In Section 6.4, we present Monte Carlo experiments to evaluate the MSFE approximation.

6.2 Misspecification and the OLS

In Chapter 5, we analyze the SSD of the MSFE for a forecasting problem which involves i.i.d. observations. In Section 5.3, we present the most relevant result in the literature on the SSD of the MSFE for a forecasting problem with a regression model with i.i.d. observations. In this section we present the most relevant result on the properties of the OLS in a forecasting problem with dependent observations under model functional misspecification.

Domowitz and White, in [43], present large sample properties of the OLS for an estimation problem under misspecification of the DGP. We begin with a description of the DGP.

Assumption 6.1 *Let the probability space $(\Omega, \mathcal{B}, P)$ be given. A sequence of real valued responses Y_τ is generated as*

$$Y_\tau = g_\tau(Z_\tau), \quad \tau = 1, 2, \ldots, n,$$

where the g_τ are unknown measurable functions of the real valued random vector Z_τ. The vector Z_τ is finite dimensional and jointly distributed with distribution function F_τ on Ω, a Euclidean space.

Y_τ and Z_τ are not assumed to be stationary.

Assumption 6.2 *The researcher chooses a sequence of functions h_τ to approximate the data generating process. $h_\tau(z,\theta)$, $\tau = 1,2,\ldots,n$, are continuous functions of θ for each z in Ω uniformly in τ, a.s-P, and measurable functions of z for each $\theta \in \Theta$, a compact subset of a finite dimensional Euclidean space.*

The nonlinear least squares (NLS) estimator $\hat{\theta}_n$ solves the following problem

$$\min_{\theta\in\Theta} \sigma_n^2(\theta) = n^{-1}\sum_{\tau=1}^{n}(Y_\tau - h_\tau(Z_\tau,\theta))^2.$$

The OLS is obtained when the h_τ are linear. The parameter θ_n^* is defined as the vector which minimizes the average prediction mean square error

$$\bar{\sigma}_n^2 = n^{-1}\sum_{\tau=1}^{n}\int (g_\tau(z) - h_\tau(z,\theta))^2 dF_\tau. \tag{6.2.1}$$

Note the prediction mean square error is the same as the $MSFE_\tau$ evaluated at the forecast origin τ. The average given in (6.2.1) over τ is the average of the $MSFE_\tau$ evaluated at different forecast origins $\tau = 1,2,\ldots,n$. We give a definition and two assumptions needed for the main result.

Definition 6.3 *Let $\bar{Q}_n(\theta)$ be continuous on a compact set, Θ, such that $\bar{Q}_n(\theta)$ has a minimum at θ_n^*, $n = 1,2,\ldots$. Let $\mathcal{J}_n(\epsilon)$ be an open sphere centered at θ_n^* with fixed radius $\epsilon > 0$. For each $n = 1,2,\ldots$, define the neighborhood $\mathcal{N}_n = \mathcal{J}_n(\epsilon)\cap\Theta$, such that its complement in Θ, $\mathcal{N}_n^c$, is compact. The minimizer θ_n^* is said to be identifiably unique if and only if*

$$\liminf_n \left[\min_{\theta\in\mathcal{N}_n^c} \bar{Q}_n(\theta) - \bar{Q}_n(\theta_n^*)\right] > 0$$

for any fixed $\epsilon > 0$.

Assumption 6.3 *The random vectors $\{Z_\tau\}$ are either (a) ϕ-mixing, with $\phi(m)$ of size $r_1/(2r_1 - 1)$, $r_1 \geq 1$; or (b) α-mixing, with $\alpha(m)$ of size $r_1/(r_1 - 1)$, $r_1 > 1$.*

Assumption 6.4 *$\{(g_\tau(Z_\tau) - h_\tau(Z_\tau, \theta))^2\}$ is dominated by uniformly $(r_1 + \delta)$-integrable functions, $r_1 \geq 1$, $0 < \delta \leq r_1$.*

Assumption 6.5 *$\bar{\sigma}_n^2$ has a minimizer at θ_n^* which is identifiably unique.*

Theorem 6.4 (Corollary 3.1 in [43]) *Under assumptions 6.1 through 6.5, $\hat{\theta}_n - \theta_n^* \longrightarrow 0$, a.s., as $n \to \infty$.*

The theorem establishes the least squares estimator as a strongly consistent estimator of the parameter vector which minimizes the average MSE of prediction. The result of the theorem describes the behavior of the NSL or OLS as n goes to infinity.

As in Chapter 5, we are interested in the sample size dependence of the MSFE. Allowing for dependent observations, in the next section we develop an algorithm that can be used to construct an approximation of the MSFE, in order to analyze the sample size dependence for finite values of the sample size variable n and determine the possible existence of optimal observation windows of finite length.

6.3 The algorithm: scalar case

As presented in Chapter 2, the forecasting problem of interest consists of predicting the observed process $\{Y_\tau\}$ at $\tau = t + 1$, $Y_{t+1} \in \mathbb{R}$, by means of a linear regression of the $k \times 1$ column vector X_t of $\mathcal{F}_t$-measurable variables. In this section, we assume $k = 1$. The forecaster does not know the data generating process (DGP) which generates the series $\{Y_\tau\}$, and uses a linear model in X_t to approximate the conditional expectation $E_t[Y_{t+1}]$. The process $\{Y_{\tau+1}, X_\tau\}$ is assumed to be either covariance stationary or strictly stationary. We obtain the following proposition as a straight forward application of theorem $A.38$ in Appendix A.

Proposition 6.5 *Given the process $\{Y_{\tau+1}, X_\tau\}$ is strictly stationary, processes of the form $\{\prod_{j=0}^{l} Y_{\tau-j}^{i_j} X_{\tau-j-1}^{k_j}\}$, where i_j and k_j are integers, are also strictly stationary .*

The linear model used to forecast Y_{t+1} is of the form

$$Y_{t+1} = \beta X_t + V_{t+1},$$

in which the parameter β, $\beta \in B$, B compact in $\mathbb{R}$, is estimated by OLS. The estimation sample contains the n most recent observations, $\{Y_{t-n+1}, \ldots, Y_t\}$ and $\{X_{t-n}, \ldots, X_{t-1}\}$, and the OLS estimator of β has the form

$$\hat{\beta}_{t,n} = \left(\sum_{\tau=t-n}^{t-1} X_\tau X_\tau^\top \right)^{-1} \left(\sum_{\tau=t-n}^{t-1} X_\tau Y_{\tau+1} \right).$$

The OLS estimator $\hat{\beta}_{t,n}$ is used to construct the forecast of Y_{t+1}, denoted $\hat{Y}_{t+1,n}$, given by

$$\hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t.$$

The criterion used to evaluate forecast accuracy is the MSFE given by

$$MSFE_n = E[(Y_{t+1} - \hat{Y}_{t+1,n})^2] = E[Y_{t+1}^2] - 2E[Y_{t+1} X_t \hat{\beta}_{t,n}] + E[X_t^2 \hat{\beta}_{t,n}^2].$$

The MSFE is the expected value of statistics which depend on the sample size parameter n. We construct a Taylor algorithm, as developed in Chapter 4, to approximate the MSFE in order to investigate the existence of an optimal observation window. The existence of such optimal observation window can be revealed by assessing the SSD of the MSFE. For this purpose, we begin the construction of the algorithm by focusing on the expectation of the following n dependent terms

$$\Theta_{1,n} \equiv Y_{t+1} X_t \hat{\beta}_{t,n} = \frac{S_{1,n}}{S_{2,n}}, \quad \Theta_{2,n} \equiv X_t^2 \hat{\beta}_{t,n}^2 = \frac{S_{3,n}^2}{S_{2,n}^2},$$

where

$$S_{1,n} \equiv \frac{1}{n} \sum_{\tau=t-n}^{t-1} Y_{t+1} X_t Y_{\tau+1} X_\tau, \quad S_{2,n} \equiv \frac{1}{n} \sum_{\tau=t-n}^{t-1} X_\tau^2, \quad S_{3,n} \equiv \frac{1}{n} \sum_{\tau=t-n}^{t-1} X_t Y_{\tau+1} X_\tau.$$

The next step in the construction of the algorithm is to apply the techniques of Chapter 4 to find approximations of $E[\Theta_{1,n}]$ and $E[\Theta_{2,n}]$. Such approximations are conducted by means of Taylor series expansions of $\Theta_{1,n}$ and $\Theta_{2,n}$, with respect to the statistics $S_{1,n}$, $S_{2,n}$ and $S_{3,n}$ about some points $\omega_{1,n}$, ω_2 and $\omega_{3,n}$ respectively. From the theory developed in Chapter 4, we learned that approximating the expectation of a function of random variables by means of Taylor series requires one, in many instances, to approximate the expectation by a truncated expectation. Using truncated expectations is necessary because Taylor series approximations are valid only within the region of convergence and, at the same time, the random variables involved take values on a specific range. In the case of $\Theta_{1,n}$, the approximation will depend on truncated central moments of $S_{1,n}$ and $S_{2,n}$ and in the case of $\Theta_{2,n}$, the approximation will depend on truncated central moments of $S_{2,n}$ and $S_{3,n}$. Let $\mathcal{A}$ be a set inside the region of convergence $\mathcal{B}$ of the Taylor series of $\Theta_{1,n}$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$. Appendix $C.1.1$ provides details on the nature of the region of convergence of the Taylor series expansion of the OLS, and on the nature of convergence sets such as $\mathcal{A}$. We write the expectation of $\Theta_{1,n}$ and $\Theta_{2,n}$ as follows:

$$E[\Theta_{1,n}] = \bar{E}[\Theta_{1,n}, \mathcal{A}] + \bar{E}[\Theta_{1,n}, \mathcal{A}^c], \tag{6.3.1}$$

$$E[\Theta_{2,n}] = \bar{E}[\Theta_{2,n}, \mathcal{A}] + \bar{E}[\Theta_{2,n}, \mathcal{A}^c], \tag{6.3.2}$$

where Taylor series can be used in $\mathcal{A}$ to approximate $\Theta_{1,n}$ and $\Theta_{2,n}$. Within $\mathcal{A}$, we look at Taylor approximations of $\Theta_{1,n}$ with respect to $S_{1,n}$ and $S_{2,n}$ about the points $\omega_{1,n}$, ω_2, and Taylor approximations of $\Theta_{2,n}$ with respect $S_{2,n}$ and $S_{3,n}$ about the points ω_2, $\omega_{3,n}$, where

$$\begin{aligned}
\omega_{1,n} &\equiv E[S_{1,n}] = \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau],\\
\omega_2 &\equiv E[S_{2,n}] = E[X_{t-1}^2],\\
\omega_{3,n} &\equiv E[S_{3,n}] = \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau].
\end{aligned}$$

The fourth order Taylor polynomial of $\Theta_{1,n}$ is as follows:

$$\begin{aligned}
Q(\Theta_{1,n}, 4) =& \frac{\omega_{1,n}}{\omega_2} + \frac{1}{\omega_2}(S_{1,n} - \omega_{1,n}) - \frac{\omega_{1,n}}{\omega_2^2}(S_{2,n} - \omega_2) - \frac{1}{\omega_2^2}(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)\\
&+ \frac{\omega_{1,n}}{\omega_2^3}(S_{2,n} - \omega_2)^2 + \frac{1}{\omega_2^3}(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2 - \frac{\omega_1}{\omega_2^4}(S_{2,n} - \omega_2)^3
\end{aligned}$$

$$+ \frac{\omega_{1,n}}{\omega_2^5}(S_{2,n}-\omega_2)^4 - \frac{1}{\omega_2^4}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3.$$

The fourth order Taylor polynomial of $\Theta_{2,n}$ is as follows:

$$\begin{aligned}
Q(\Theta_{2,n},4) =& \frac{\omega_{3,n}^2}{\omega_2^2} + 2\frac{\omega_{3,n}}{\omega_2^2}(S_{3,n}-\omega_{3,n}) - 2\frac{\omega_{3,n}^2}{\omega_2^3}(S_{2,n}-\omega_2) + \frac{1}{\omega_2^2}(S_{3,n}-\omega_{3,n})^2 \\
&- 4\frac{\omega_{3,n}}{\omega_2^3}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2) + 3\frac{\omega_{3,n}^2}{\omega_2^4}(S_{2,n}-\omega_2)^2 \\
&- 2\frac{1}{\omega_2^3}(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2) + 6\frac{\omega_{3,n}}{\omega_2^4}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2 \\
&- 4\frac{\omega_{3,n}^2}{\omega_2^5}(S_{2,n}-\omega_2)^3 + \frac{3}{\omega_2^4}(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2 \\
&- 8\frac{\omega_{3,n}}{\omega_2^5}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3 + 5\frac{\omega_{3,n}^2}{\omega_2^6}(S_{2,n}-\omega_2)^4.
\end{aligned}$$

Using the fourth order Taylor polynomials $Q(\Theta_{1,n},4)$ and $Q(\Theta_{2,n},4)$ to approximate $\Theta_{1,n}$ and $\Theta_{2,n}$ respectively inside $\mathcal{A}$, (6.3.1) and (6.3.2) become

$$E[\Theta_{1,n}] \approx \bar{E}[Q(\Theta_{1,n},4),\mathcal{A}] + \bar{E}[\Theta_{1,n},\mathcal{A}^c], \tag{6.3.3}$$

$$E[\Theta_{2,n}] \approx \bar{E}[Q(\Theta_{2,n},4),\mathcal{A}] + \bar{E}[\Theta_{2,n},\mathcal{A}^c]. \tag{6.3.4}$$

Using these approximations, the MSFE approximation can be written as follows:

$$MSFE_n \approx E[Y_{t+1}^2] - 2(\bar{E}[Q(\Theta_{1,n},4),\mathcal{A}] + \bar{E}[\Theta_{1,n},\mathcal{A}^c]) + \bar{E}[Q(\Theta_{2,n},4),\mathcal{A}] + \bar{E}[\Theta_{2,n},\mathcal{A}^c].$$

The central moments involved in the expectation of the Taylor polynomials are expanded and simplified to derive the SSD in terms of the sample size variable n. Appendix D, Section $D.1$ presents the derivation of the central moments for the general case without assuming $P(X \in \mathcal{A}) \approx 1$. With the assumption $P(X_\tau \in \mathcal{A}) \approx 1$ for all τ, the approximations for $\Theta_{1,n}$ and $\Theta_{2,n}$ given in (6.3.3) and (6.3.4) become $E[\Theta_{1,n}] \approx E[Q(\Theta_{1,n},4)]$ and $E[\Theta_{2,n}] \approx E[Q(\Theta_{2,n},4)]$, respectively, and the MSFE approximation is as follows:

$$MSFE_n \approx E[Y_{t+1}^2] - 2E[Q(\Theta_{1,n},4)] + E[Q(\Theta_{2,n},4)]. \tag{6.3.5}$$

We write the central moments involved in the expectation of $Q(\hat{\beta}_{t,n},4)$ and $Q(\hat{\beta}^2_{t,n},4)$ under the assumptions of covariance stationarity and $P(X\in A)\approx 1$:

$$E[(S_{1,n}-\omega_{1,n})]=0,\quad E[(S_{2,n}-\omega_2)]=0,\quad E[(S_{3,n}-\omega_{3,n})]=0,$$

$$E[(S_{2,n}-\omega_2)^2]=\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_\tau^4]+\sum_{i\neq j,t-n}^{t-1}E[X_i^2X_j^2]\Big]-E^2[X_{t-1}^2],$$

$$E[(S_{3,n}-\omega_{3,n})^2]=\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_t^2Y_{\tau+1}^2X_\tau^2]+\sum_{i\neq j,t-n}^{t-1}E[X_t^2Y_{i+1}X_iY_{j+1}X_j]$$

$$-\sum_{\tau=t-n}^{t-1}E^2[X_tY_{\tau+1}X_\tau]-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\Big],$$

$$E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)]=\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau^3]$$

$$+\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^2]\Big]-\frac{1}{n}E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau],$$

$$E[(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n})]=\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau^3]+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^2]\Big]$$

$$-\frac{1}{n}E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau],$$

$$E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2]=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau^5]$$

$$+\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^4]+\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i^3X_j^2]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2]-\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E[X_\tau^4]$$

$$-\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^4]-\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E\left[X_i^2X_j^2\right]$$

$$-\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E\left[X_j^2X_k^2\right]\Big]$$

$$-\frac{2}{n^2}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1}E\left[Y_{t+1}X_tY_{\tau+1}X_\tau^3\right]+\sum_{i\neq j,t-n}^{t-1}E\left[Y_{t+1}X_tY_{i+1}X_iX_j^2\right]\Big]$$

$$+\frac{2}{n}E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau],$$

$$E[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2]=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau^5]+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^4]$$
$$+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i^3X_j^2]+\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^2X_k^2]$$
$$-\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E[X_\tau^4]-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_j^4]$$
$$-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E\left[X_i^2X_j^2\right]-\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_i]E\left[X_j^2X_k^2\right]\Big]$$
$$-\frac{2}{n^2}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1}E\left[X_tY_{\tau+1}X_\tau^3\right]+\sum_{i\neq j,t-n}^{t-1}E\left[X_tY_{i+1}X_iX_j^2\right]\Big]$$
$$+\frac{2}{n}E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau],$$

$$E[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)]=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}E[X_t^2Y_{\tau+1}^2X_\tau^4]+\sum_{i\neq j,t-n}^{t-1}E[X_t^2Y_{i+1}^2X_i^2X_j^2]$$
$$+\sum_{i\neq j,t-n}^{t-1}E[X_t^2Y_{i+1}X_iY_{j+1}X_j^3]+\sum_{i\neq j\neq k,t-n}^{t-1}E[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2]$$
$$-2\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E[X_tY_{\tau+1}X_\tau^3]-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j^3]$$
$$-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{i+1}X_iX_j^2]-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_jX_i^2]$$
$$-2\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_jX_k^2]\Big]$$
$$+\frac{1}{n^2}\Big[2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]+2E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E^2[X_tY_{\tau+1}X_\tau]$$
$$-E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_t^2Y_{\tau+1}^2X_\tau^2]-E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_t^2Y_{i+1}X_iY_{j+1}X_j]\Big],$$

$$E[(S_{2,n}-\omega_2)^3] = \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} E[X_\tau^6] + \sum_{i\neq j,t-n}^{t-1} E[X_i^4 X_j^2] + \sum_{i\neq j\neq k,t-n}^{t-1} E[X_i^2 X_j^2 X_k^2]\Big]$$

$$- \frac{3}{n^2}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} E[X_\tau^4] + \sum_{i\neq j,t-n}^{t-1} E[X_i^2 X_j^2]\Big] + 2E^3[X_{t-1}^2],$$

$$E[(S_{2,n}-\omega_2)^4] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} E[X_\tau^8] + \sum_{i\neq j,t-n}^{t-1} E[X_i^6 X_j^2] + \sum_{i\neq j,t-n}^{t-1} E[X_i^4 X_j^4]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} E[X_i^4 X_j^2 X_k^2] + \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_i^2 X_j^2 X_k^2 X_l^2]\Big]$$

$$- \frac{4}{n^3}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} E[X_\tau^6] + \sum_{i\neq j,t-n}^{t-1} E[X_i^4 X_j^2] + \sum_{i\neq j\neq k,t-n}^{t-1} E[X_i^2 X_j^2 X_k^2]\Big]$$

$$+ \frac{6}{n^2}E^2[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} E[X_\tau^4] + \sum_{i\neq j,t-n}^{t-1} E[X_i^2 X_j^2]\Big] - 5E^4[X_{t-1}^2],$$

$$E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau^7]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_iX_j^6] + \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i^5X_j^2]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i^3X_j^4] + \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_iX_j^4X_k^2]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i^3X_j^2X_k^2] + \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2X_l^2]$$

$$- \sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]E[X_\tau^6] - \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^6]$$

$$- \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_i^4X_j^2] - \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^4X_i^2]$$

$$- \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^4X_k^2] - \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_i^2X_j^2X_k^2]$$

$$- \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^2X_k^2X_l^2]\Big]$$

$$+\frac{1}{n^3}\Big[-3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau^5]-3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^4]$$

$$-3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i^3X_j^2]$$

$$-3E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2]$$

$$+3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E[X_\tau^4]$$

$$+3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^4]$$

$$+3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E[X_i^2X_j^2]$$

$$+3E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E[X_j^2X_k^2]\Big]$$

$$+\frac{1}{n^2}\Big[3E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau^3]+3E^2[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_iX_j^2]\Big]$$

$$-\frac{3}{n}E^3[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau],$$

$$E[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3]=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau^7]$$

$$+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^6]+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i^5X_j^2]$$

$$+\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i^3X_j^4]+\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^4X_k^2]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_i^3X_j^2X_k^2]+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}E[X_tY_{i+1}X_iX_j^2X_k^2X_l^2]$$

$$-\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E[X_\tau^6]-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_j^6]$$

$$-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_i^4X_j^2]-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_j^4X_i^2]$$

$$- \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_j^4X_k^2] - \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_i^2X_j^2X_k^2]$$

$$- \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_j^2X_k^2X_l^2]\Big]$$

$$+ \frac{1}{n^3}\Big[- 3E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau^5] - 3E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_iX_j^4]$$

$$- 3E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i^3X_j^2] - 3E[X_{t-1}^2] \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_iX_j^2X_k^2]$$

$$+ 3E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]E[X_\tau^4] + 3E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_j^4]$$

$$+ 3E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_i^2X_j^2]$$

$$+ 3E[X_{t-1}^2] \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_j^2X_k^2]\Big]$$

$$+ \frac{1}{n^2}\Big[3E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau^3] + 3E^2[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_iX_j^2]\Big]$$

$$- \frac{3}{n}E^3[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau],$$

$$E[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} E[X_t^2Y_{\tau+1}^2X_\tau^6] + \sum_{i\neq j,t-n}^{t-1} E[X_t^2Y_{i+1}^2X_i^2X_j^4]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[X_t^2Y_{i+1}^2X_i^4X_j^2] + \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[X_t^2Y_{i+1}X_i^5Y_{j+1}X_j] + \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3] + \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2]$$

$$+ \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2] - 2\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]E[X_tY_{\tau+1}X_\tau^5]$$

$$- 2\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j^5] - 2\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{i+1}X_iX_j^4]$$

$$- 2 \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_i^4] - 2 \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_k^4]$$

$$- 2 \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{i+1} X_i^3 X_j^2] - 2 \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j^3 X_i^2]$$

$$- 2 \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j^3 X_k^2]$$

$$- 2 \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{i+1} X_i X_j^2 X_k^2]$$

$$- 2 \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_i^2 X_k^2]$$

$$- 2 \sum_{i \neq j \neq k \neq l, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_k^2 X_l^2]$$

$$+ \sum_{\tau = t-n}^{t-1} E^2[X_t Y_{\tau+1} X_\tau] E[X_\tau^4] + \sum_{i \neq j, t-n}^{t-1} E^2[X_t Y_{i+1} X_i] E[X_j^4]$$

$$+ \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] E[X_j^4]$$

$$+ \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] E[X_k^4]$$

$$+ \sum_{i \neq j, t-n}^{t-1} E^2[X_t Y_{i+1} X_i] E[X_i^2 X_j^2] + \sum_{i \neq j \neq k, t-n}^{t-1} E^2[X_t Y_{i+1} X_i] E[X_j^2 X_k^2]$$

$$+ \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] E[X_i^2 X_j^2]$$

$$+ \sum_{i \neq j \neq k, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] E[X_j^2 X_k^2]$$

$$+ \sum_{i \neq j \neq k \neq l, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] E[X_k^2 X_l^2] \Big]$$

$$+ \frac{1}{n^3} \Big[- 2 E[X_{t-1}^2] \sum_{\tau = t-n}^{t-1} E[X_t^2 Y_{\tau+1}^2 X_\tau^4] - 2 E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[X_t^2 Y_{i+1}^2 X_i^2 X_j^2]$$

$$- 2 E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[X_t^2 Y_{i+1} X_i Y_{j+1} X_j^3]$$

$$
\begin{aligned}
&- 2E[X_{t-1}^2] \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t^2 Y_{i+1} X_i Y_{j+1} X_j X_k^2] \\
&+ 4E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] E[X_t Y_{\tau+1} X_\tau^3] \\
&+ 4E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j^3] \\
&+ 4E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{i+1} X_i X_j^2] \\
&+ 4E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_i^2] \\
&+ 4E[X_{t-1}^2] \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j X_k^2]\Big] \\
&+ \frac{1}{n^2}\Big[- 3E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E^2[X_t Y_{\tau+1} X_\tau] + E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_t^2 Y_{\tau+1}^2 X_\tau^2] \\
&+ E^2[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_t^2 Y_{i+1} X_i Y_{j+1} X_j] \\
&- 3E^2[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j]\Big].
\end{aligned}
$$

6.4 Monte Carlo evidence

6.4.1 The experiment

In this section, we present Monte Carlo experiments to investigate the ramifications of misspecification in the forecasting problem described in Chapter 2 and to evaluate the ability of the Taylor algorithm to capture these effects. In particular, we focus on the case where the explanatory and dependent variables are covariance stationary processes. To carry out this endeavor, we construct a benchmark MSFE by means of Monte Carlo simulations. This benchmark MSFE is then compared to the MSFE approximation obtained with the Taylor algorithm and given by (6.3.5). For the analysis, we consider the DGP given in the motivating example 2.16

$$Y_t = \mu + \phi Y_{t-1} + U_t,$$

where $\{U_\tau\} \sim IIN(0, \sigma_u)$ is an innovation process and ϕ is a scalar parameter. The forecasting model in the example is given by $Y_\tau = \beta + V_\tau$, so that the sequence of explanatory variables $\{X_\tau\}$ is a sequence of ones.

As described in the previous chapter, the MSFE cannot be evaluated analytically, so that we calculate the benchmark MSFE by means of Monte Carlo simulations. The motivation behind using Monte Carlo simulations to determine a benchmark MSFE lies in that the MSFE is equal to the expected value of the conditional mean square forecast error (CMSFE). Given a realization of the process $\{Y_\tau\}_{\tau=t-n+1}^{t}$, it is simple to compute the CMSFE conditional on the given sample. Generating many such samples, M, by Monte Carlo simulations, we can construct M conditional mean square forecast errors, $\{CMSFE_i\}_{i=1}^{M}$, and approximate the MSFE by the sample mean of the simulations.

We now describe the details involved in the construction of the benchmark MSFE. For the given set of values of the parameters $\mathcal{P} = \{\mu, \sigma_u, \phi\}$, one hundred thousand Monte Carlo simulations are conducted ($M = 100000$). We use the index m to denote a particular Monte Carlo simulation. For the mth simulation, we generate the sample series $\{u_{\tau,m}\}_{\tau=1}^{T}$ of length $T = 251$ as a realization of the innovation process $\{U_\tau\}_{\tau=t-n}^{t}$, such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $251 \leftrightarrow t$. Each u is a realization of a normally distributed random variable, $U \sim N(0, \sigma_u)$, and the population series is independent and identically distributed, $\{U_\tau\}_{\tau=t-n}^{t-1} \sim IID$. From this sample series of the innovation process, we calculate the sample series $\{y_{\tau,m}\}_{\tau=1}^{T}$ by means of the relation $y_{\tau,m} = \phi y_{\tau-1,m} + u_\tau$ with the starting value $y_{1,m} = 0$. The first 50 values of y are discarded.

Finally, with the sample series $\{y_{\tau,m}\}_{\tau=51}^{T}$, at the forecast origin $\tau = T-1$, we construct the CMSFE as follows:

$$CMSFE_{T-1,m,n} = y_T^2 - 2y_T\beta_{T-1,n,m} + \beta_{T-1,n,m}^2,$$
$$\beta_{T-1,n,m} = \frac{\sum_{\tau=T-n}^{T-1} y_{\tau,m}y_{\tau-1,m}}{\sum_{\tau=T-n}^{T-1} y_{\tau,m}^2}.$$

For each simulation, we obtain $T-1-50 = 200$ values of the CMSFE, one for each value of n starting from $n = 1$ to $n = 200$. The case $n = 1$ refers to estimation of the OLS carried out with only one observation. For a particular set of parameters $\mathcal{P}$, we obtain an array of size $M \times T - 51$ of CMSFEs, $\{CMSFE_{i,j}\}_{i=1,j=1}^{M,T-51}$. Finally, the benchmark MSFE for a set of parameters $\mathcal{P}$ and for an observation window of size n is given by the following:

$$MSFE_n \approx \frac{1}{M-50}\sum_{i=51}^{M} CMSFE_{i,n}. \tag{6.4.1}$$

The benchmark Monte Carlo MSFE is compared with the MSFE approximation obtained with the Taylor algorithm given by (6.3.5). The approximation (6.3.5) is constructed by use of sample moments in place of their population counterparts. For this, we generate the innovation series $\{u_\tau\}_{\tau=1}^N$ of length $N = 3100$ as a realization of the innovation process $\{U_\tau\}_{\tau=t-n}^t$, such that the first element of the series is the first observation, $1 \leftrightarrow t-n$, and the last element of the series is the last observation, $3100 \leftrightarrow t$. Each u is a realization of a normally distributed random variable, $u \sim N(0, \sigma_u)$, and the population series is independent and identically distributed, $\{U_\tau\}_{\tau=t-n}^{t-1} \sim IID$. The sample series $\{y_\tau\}_{\tau=1}^N$ is generated by means of the relation $y_\tau = \phi y_{\tau-1} + u_\tau$ with the starting value $y_1 = 0$. The first 100 values of y are discarded.

The population moments in (6.3.5) are estimated by generating their sample counterparts. For example:

$$E[X_{t-1}^2 X_{t-2}^2] \approx \frac{1}{N-101}\sum_{\tau=102}^{N} x_\tau^2 x_{\tau-1}^2,$$
$$E[X_t^2 Y_t^2 X_{t-1}^2] \approx \frac{1}{N-101}\sum_{\tau=102}^{N} x_\tau^2 y_\tau^2 x_{\tau-1}^2.$$

Therefore, for a given set of the parameters, $\mathcal{P} = \{\mu, \sigma_u, \phi\}$, we can generate the necessary sample moments and ultimately evaluate (6.3.5) for different values of the observation window size n. The resulting MSFE can be compared to the benchmark MSFE (6.4.1).

6.4.2 Discussion

The parameters $\mu = 0$, $\sigma_u = 1$ were fixed for four experiments in which we varied the value of the parameter ϕ. The values of ϕ studied were 0.1, 0.5, 0.8, and 0.95. For $\phi = 0$, the model is correctly specified so that as ϕ increases, misspecification in some sense increases. The benchmark MSFE and the Taylor approximation of the MSFE are compared for each value of ϕ in Figures 6.1, 6.2, 6.3, and 6.4. The results show that the Taylor approximation of the MSFE captures the general behavior of the benchmark MSFE, but the results are not as accurate as the results for i.i.d. processes presented in the previous chapter. The results are best for the case with $\phi = 0.1$, which is the process nearest to being i.i.d. of the processes studied. The lack of accuracy in the experiments might be attributed to the method of approximating population moments with sample moments. Future work will employ Newey-West estimators.

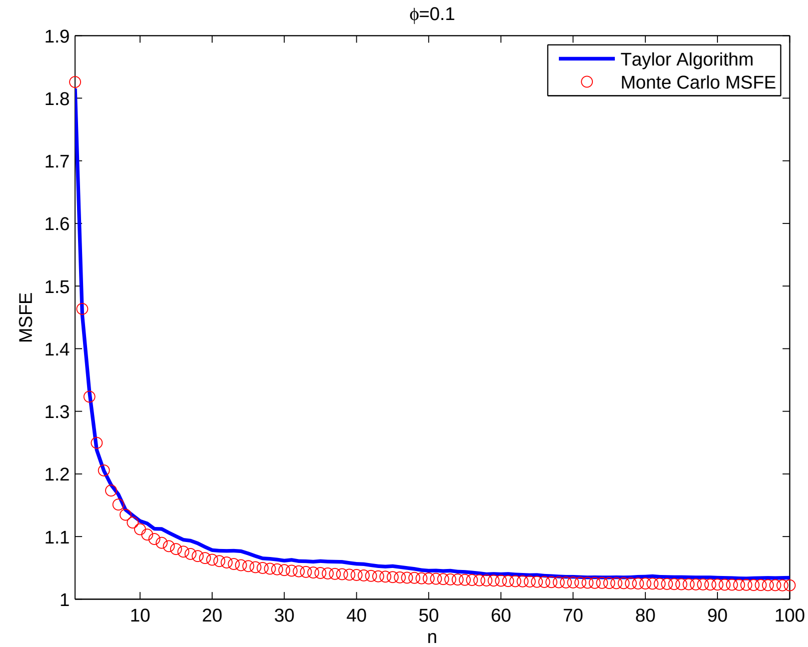


Figure 6.1: MSFE for $\sigma_u = 1, \phi = 0.1$

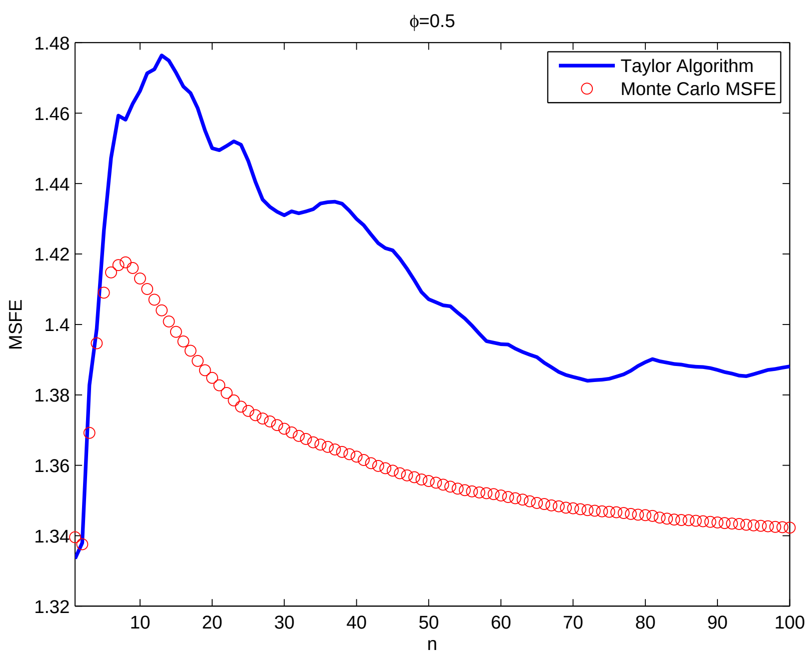


Figure 6.2: MSFE for $\sigma_u = 1, \phi = 0.5$

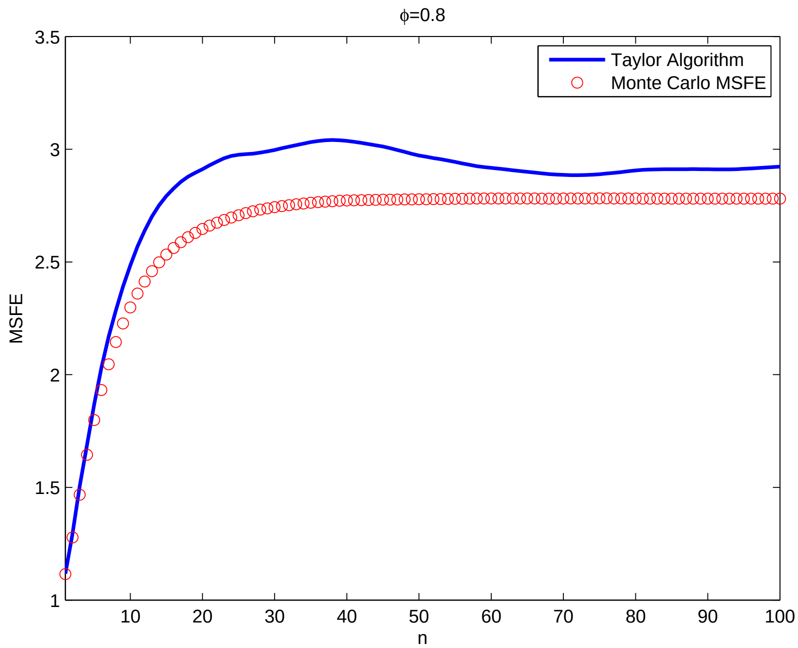


Figure 6.3: MSFE for $\sigma_u = 1, \phi = 0.8$

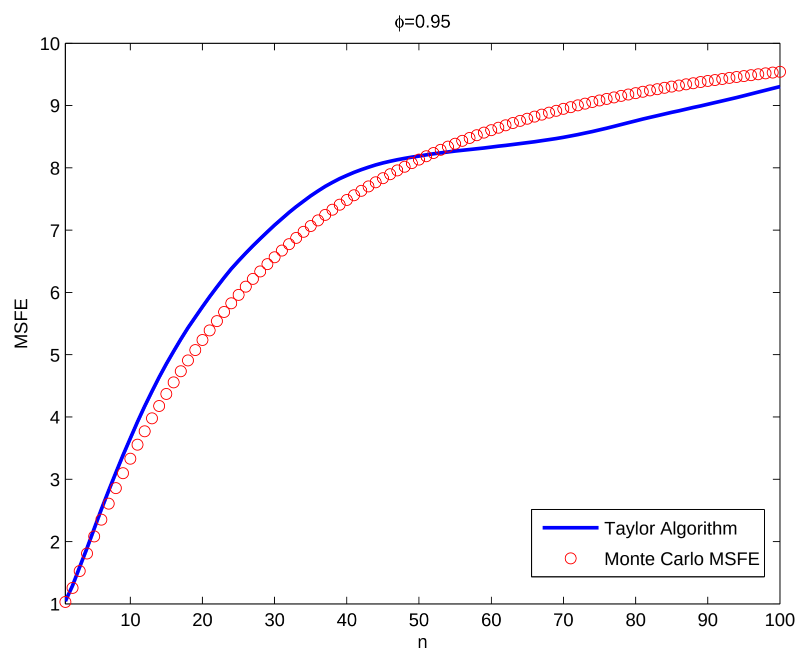


Figure 6.4: MSFE for $\sigma_u = 1, \phi = 0.95$

Chapter 7

Taylor algorithm for structural break processes

7.1 Introduction

As noted in the historical exposition of forecasting in Chapter 2, one major obstacle for the subject of forecasting to gain acceptance in the economic community has been the lack of homogeneity of economic data. Much work has been done to understand the level of regularity in economic data and, in particular, the presence of structural changes. The literature which deals with testing for structural breaks includes: the work of Chow [29], for linear regression models when the point of the break is known; the work of Brown, Durbin, and Evans [28], applicable when the point of the break is unknown; and the application of tests to dynamic models and tests for the estimation of the size and timing of the break by Plobeger, Kramer and Kontrus [116], Hansen [66], Andrews [5], Inclan and Tiao [78], Andrews and Ploberger [6], Chu, Stinchcombe and White [30] and Bai and Perron [13]. This plethora of work has led to abundant evidence of structural breaks in economic series, [3, 13, 32, 33, 53, 139].

The problem of forecasting a process which has undergone a structural change presents an ideal circumstance to address the premise of this thesis by asking the question: How much data should one use to forecast such a series. Using only post-break data for the estimation of the forecasting model would result in unbiased forecast errors. If, in addition, pre-break data is used in the estimation of the forecasting model, forecast errors would no longer be unbiased, although the variance would be lower than in the post-break-only case. In this chapter, we present a methodology to quantify this trade off and answer the question of how far back one should look when making a forecast.

Modern economies undergo major institutional, political, financial, and technological changes which manifest themselves in the data employed by econometricians. These manifestations are modeled by use of structural breaks in the form of parameter shifts. The significance of the presence of structural changes in the context of the forecasting problem has been addressed by Clements and Hendry,

> Deterministic shifts (changes in equilibrium means and steady-state trends) in the model relative to the DGP are a dominant source of forecast failure. ([33], p. 69)

> [Clements and Hendry] present taxonomies of forecast errors in both $I(0)$ and $I(1)$ systems, which suggest that structural breaks are the main culprit for systematic forecast failure. ([33], p. 36)

The most commonly used procedures developed to handle non-stationarities use a rolling window of a fixed size, an expanding window (recursive method), or apply exponentially decreasing weights. None of these schemes are likely to be optimal if the DGP undergoes a structural break. A rolling window of a short fixed size might work well immediately after the break, but valuable information will be lost as the distance from the break increases. The recursive scheme and the exponential scheme with long memory will produce significantly biased forecasts after the break until the post break information significantly outweighs the pre-break information.

Work on forecasting in the presence of structural changes has only recently began to be addressed by econometricians. Clements and Hendry [32, 33] address the analysis of forecast errors from autoregressive models subject to structural change. However, the authors assume the parameters of the AR model remain constant during the estimation period. Pesaran and Timmermann [114] develop a theoretical framework for the analysis of small-sample properties of forecasts from general autoregressive models under structural breaks. They determine conditions under which the forecast errors are unbiased and demonstrate some of their theoretical results with Monte Carlo simulations.

To our knowledge, the only work, besides our own, which explores the subject of determining quantitatively optimal observation windows for processes that undergo structural changes is that of Pesaran and Timmermann [113]. In [113], Pesaran and Timmermann analyze the sample size dependence for the conditional and unconditional MSFE when the DGP is linear with a singular structural break and the forecasting model is linear. Hence misspecification arises from not modeling the break when using pre-break data to estimate the post-break model. Under the assumption of strictly exogenous regressors, the authors obtain stylized facts describing the appropriate use of pre-break observations for the conditional MSFE. For the single regressor case, the authors apply the restrictive conditions of identically independent and jointly normally distributed disturbances and regressors to obtain an analytic expression for the unconditional MSFE.

The analyses of Clements and Hendry [32, 33], and Pesaran and Timmermann [113, 114] assume that the estimation is carried out based on a correctly specified post-break model; i.e, the functional form of the model and DGP after the break occurs are AR models with the same autoregressive parameter. The only misspecification in estimation comes from effectively "ignoring" the break when using pre-break data. Our work allows for such break misspecification but further accommodates other forms of misspecification by refraining from putting any assumptions on the DGP. We note that the work that follows focuses on the treatment of independent and identically distributed processes which undergo a structural break. However, the theory and methodology presented here can be extended to the problem of forecasting with time series models which undergo a structural break. This is subject for future research.

7.2 Forecasting a general structural break process

As presented in Chapter 2, the forecasting problem of interest consists of predicting the observed process $\{Y_\tau\}$ at $\tau = t+1$, $Y_{t+1} \in \mathbb{R}$, by means of a linear regression of the $k \times 1$ column vector X_t of $\mathcal{F}_t$-measurable variables. In this section we assume $k = 1$. We apply the techniques of Chapter 4 to approximate the optimal observation window to forecast the process $\{Y_\tau\}$ generated by a DGP with a temporal structural break. The DGP is as

follows:

$$Y_{\tau+1} = \begin{cases} Y_{1,\tau+1}, & \tau \leq t - n_b \\ Y_{2,\tau+1}, & \tau > t - n_b \end{cases}. \tag{7.2.1}$$

We assume the forecaster knows the process $\{Y_\tau\}$ undergoes a structural break at time $t-n_b$. Beyond the occurrence of a structural break at time $t-n_b$, the forecaster does not know the nature of the DGP which generates the process $\{Y_\tau\}$ and uses a model for the conditional expectation of Y_{t+1}, $E_t[Y_{t+1}]$, which is linear in X_t. The linear model used to construct the forecast of Y_{t+1} is of the form

$$Y_{t+1} = \beta^\top X_t + V_{t+1},$$

in which the parameter β, $\beta \in B$, B compact in $\mathbb{R}$, is estimated by ordinary least squares (OLS). The estimation sample contains the n most recent observations and the OLS estimator of β has the form

$$\hat{\beta}_{t,n} \equiv \left(\sum_{\tau=t-n}^{t-1} X_\tau^2 \right)^{-1} \left(\sum_{\tau=t-n}^{t-1} X_\tau Y_{\tau+1} \right).$$

7.2.1 The MSFE for $n \geq n_b$

As explained in Section 2.7, to understand the sample size dependence (SSD) of the MSFE, we seek to construct an approximation consisting of a function which depends only on moments of the explanatory and dependent variables, and on the variable n. In this way, given the necessary moments or their sample counterparts, one can compute and compare different values of the MSFE for any desired window size n. The OLS estimator has different functional forms for the two cases $n \geq n_b$ and $n < n_b$. For $n \geq n_b$, the OLS estimator can be written as the sum of two terms $\hat{\beta}_{t,n} = \Theta_{t,n} + \Lambda_{t,n}$, where

$$\Theta_{t,n} = Q^{-1} \sum_{\tau=t-n}^{t-n_b-1} X_\tau Y_{1,\tau+1}, \quad \Lambda_{t,n} = Q^{-1} \sum_{\tau=t-n_b}^{t-1} X_\tau Y_{2,\tau+1}, \quad Q = \sum_{\tau=t-n}^{t-1} X_\tau^2.$$

The above OLS estimator $\hat{\beta}_{t,n}$ is then used to construct the forecast of Y_{t+1}, denoted $\hat{Y}_{t+1,n}$, given by $\hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t = (\Theta_{t,n} + \Lambda_{t,n}) X_t$. Using as cost function the squared loss function, the criterion which provides a measure of forecast accuracy is the MSFE given by

$$MSFE_n = E[(Y_{t+1} - \hat{Y}_{t+1,n})^2] = E[Y_{t+1}^2] - 2E[Y_{t+1}\hat{Y}_{t+1,n}] + E[\hat{Y}_{t+1,n}^2]. \tag{7.2.2}$$

In this chapter we assume, for the sequence of regressors $\{X_\tau\}$, X_s and X_t to be independent and identically distributed for $s \neq t$. By independence, we can write

$$\begin{aligned} MSFE_n =& E[Y_{t+1}^2] - 2E[Y_{t+1}X_t](E[\Theta_{t,n}] + E[\Lambda_{t,n}]) \\ &+ E[X_t^2](E[\Theta_{t,n}^2] + 2E[\Theta_{t,n}\Lambda_{t,n}] + E[\Lambda_{t,n}^2]). \end{aligned}$$

The MSFE consists of the expected value of functions of statistics which depend on the parameter n. In the sections to follow, we apply Taylor algorithms developed in Chapter 4 to approximate the MSFE in order to find estimates for the optimal observation window size n. $\Theta_{t,n}$ and $\Lambda_{t,n}$ can be written as functions of three statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$ as follows

$$\Theta_{t,n} = \frac{S_{1,n}}{S_{2,n}}, \quad \Lambda_{t,n} = \frac{S_{3,n}}{S_{2,n}},$$

where

$$S_{1,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau, \quad S_{2,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2, \quad S_{3,n} = \frac{1}{n}\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau. \tag{7.2.3}$$

The objective is to apply the techniques of Chapter 4 to find approximations of $E[\Theta_{t,n}]$, $E[\Theta_{t,n}^2]$, $E[\Lambda_{t,n}]$, $E[\Lambda_{t,n}^2]$, and $E[\Theta_{t,n}\Lambda_{t,n}]$. Such approximations are conducted by means of Taylor series expansions of $\Theta_{t,n}$ and $\Theta_{t,n}^2$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$ about some points $\omega_{1,n}$ and ω_2; by means of Taylor series expansions of $\Lambda_{t,n}$ and $\Lambda_{t,n}^2$ with respect to the statistics $S_{3,n}$ and $S_{2,n}$ about some points $\omega_{3,n}$ and ω_2; and by means of Taylor series expansions of $\Theta_{t,n}\Lambda_{t,n}$ with respect to the statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$ about some points $\omega_{1,n}$, ω_2, and $\omega_{3,n}$. Once these Taylor approximations are obtained, we can approximate the expectations $E[\Theta_{t,n}]$, $E[\Theta_{t,n}^2]$, $E[\Lambda_{t,n}]$, $E[\Lambda_{t,n}^2]$, and $E[\Theta_{t,n}\Lambda_{t,n}]$. From the theory developed in Chapter 4, we learned that approximating the expectation of a function of random variables by means of Taylor series requires one, in many instances, to approximate the expectation by a truncated expectation. The need for truncated expectations arises from the fact that the Taylor approximation is valid only in the region of convergence of the Taylor series. In the case of $\Theta_{t,n}$ and $\Theta_{t,n}^2$, the approximations will depend on truncated central moments of $S_{1,n}$ and $S_{2,n}$; in the case of $\Lambda_{t,n}$ and $\Lambda_{t,n}^2$, the approximations will depend on truncated central moments of $S_{3,n}$ and $S_{2,n}$; and in the case of $\Theta_{t,n}\Lambda_{t,n}$, the approximations will depend on truncated central moments of $S_{1,n}$, $S_{3,n}$, and $S_{2,n}$. Let A be the region of convergence for the Taylor series of $\Theta_{t,n}$ with respect to the statistics $S_{1,n}$ and $S_{2,n}$, let B be the region of convergence for the Taylor series of $\Lambda_{t,n}$ with respect to the statistics $S_{3,n}$ and $S_{2,n}$, and let C be the region of convergence for the Taylor series of $\Theta_{t,n}\Lambda_{t,n}$ with respect to the statistics $S_{1,n}$, $S_{3,n}$, and $S_{2,n}$. Appendix $C.1.1$ provides details on the nature of the region of convergence for the Taylor expansion. We write the expectation of the components $\Theta_{t,n}, \Lambda_{t,n}, \Theta_{t,n}^2, \Lambda_{t,n}^2, \Theta_{t,n}\Lambda_{t,n}$ of the MSFE as follows:

$$
\begin{aligned}
&E[\Theta_{t,n}] = \bar{E}[\Theta_{t,n}, A] + \bar{E}[\Theta_{t,n}, A^c], \quad E[\Theta_{t,n}^2] = \bar{E}[\Theta_{t,n}^2, A] + \bar{E}[\Theta_{t,n}^2, A^c],\\
&E[\Lambda_{t,n}] = \bar{E}[\Lambda_{t,n}, B] + \bar{E}[\Lambda_{t,n}, B^c], \quad E[\Lambda_{t,n}^2] = \bar{E}[\Lambda_{t,n}^2, B] + \bar{E}[\Lambda_{t,n}^2, B^c],\\
&E[\Theta_{t,n}\Lambda_{t,n}] = \bar{E}[\Theta_{t,n}\Lambda_{t,n}, C] + \bar{E}[\Theta_{t,n}\Lambda_{t,n}, C^c],
\end{aligned}
$$

where A^c is the complement of A, B^c is the complement of B, and C^c is the complement of C. Taylor series can be used in A to approximate $\Theta_{t,n}$ and $\Theta_{t,n}^2$, and similarly Taylor series can be used in B to approximate $\Lambda_{t,n}$ and $\Lambda_{t,n}^2$. To obtain further analytic results, we assume $P(X \in A) \approx 1$, $P(X \in B) \approx 1$, and $P(X \in C) \approx 1$ so that $E[\Theta_{t,n}] \approx \bar{E}[\Theta_{t,n}, A]$, $E[\Theta_{t,n}^2] \approx \bar{E}[\Theta_{t,n}^2, A]$, $E[\Lambda_{t,n}] \approx \bar{E}[\Lambda_{t,n}, B]$, $E[\Lambda_{t,n}^2] \approx \bar{E}[\Lambda_{t,n}^2, B]$, and $E[\Theta_{t,n}\Lambda_{t,n}] \approx \bar{E}[\Theta_{t,n}\Lambda_{t,n}, C]$. We define the points about which to calculate the Taylor series as follows:

$$
\begin{aligned}
&\omega_{1,n} \equiv E[S_{1,n}] = (1 - \frac{n_b}{n})E[Y_{1,t-n_b}X_{t-n_b-1}],\\
&\omega_2 \equiv E[S_{2,n}] = E[X_{t-1}^2],\\
&\omega_{3,n} \equiv E[S_{3,n}] = \frac{n_b}{n}E[Y_{2,t}X_{t-1}],
\end{aligned}
$$

where the equalities follow from the i.i.d. assumption. The fourth order Taylor polynomials of $\Theta_{t,n}$ and $\Lambda_{t,n}$ about the points $\omega_{1,n}$ and ω_2 are as follows:

$$\begin{aligned}
Q(\Theta_{t,n},4) =& \frac{\omega_{1,n}}{\omega_2} + \frac{1}{\omega_2}(S_{1,n}-\omega_{1,n}) - \frac{\omega_{1,n}}{\omega_2^2}(S_{2,n}-\omega_2) - \frac{1}{\omega_2^2}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2) \\
&+ \frac{\omega_{1,n}}{\omega_2^3}(S_{2,n}-\omega_2)^2 + \frac{1}{\omega_2^3}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2 - \frac{\omega_{1,n}}{\omega_2^4}(S_{2,n}-\omega_2)^3 \\
&+ \frac{\omega_{1,n}}{\omega_2^5}(S_{2,n}-\omega_2)^4 - \frac{1}{\omega_2^4}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3, \\
Q(\Lambda_{t,n},4) =& \frac{\omega_{3,n}}{\omega_2} + \frac{1}{\omega_2}(S_{3,n}-\omega_{3,n}) - \frac{\omega_{3,n}}{\omega_2^2}(S_{2,n}-\omega_2) - \frac{1}{\omega_2^2}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2) \\
&+ \frac{\omega_{3,n}}{\omega_2^3}(S_{2,n}-\omega_2)^2 + \frac{1}{\omega_2^3}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2 - \frac{\omega_{3,n}}{\omega_2^4}(S_{2,n}-\omega_2)^3 \\
&+ \frac{\omega_{3,n}}{\omega_2^5}(S_{2,n}-\omega_2)^4 - \frac{1}{\omega_2^4}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3.
\end{aligned}$$

The fourth order Taylor polynomials of $\Theta_{t,n}^2$ and $\Lambda_{t,n}^2$ about the points $\omega_{1,n}$ and ω_2 are as follows:

$$\begin{aligned}
Q(\Theta_{t,n}^2,4) =& \frac{\omega_{1,n}^2}{\omega_2^2} + 2\frac{\omega_{1,n}}{\omega_2^2}(S_{1,n}-\omega_{1,n}) - 2\frac{\omega_{1,n}^2}{\omega_2^3}(S_{2,n}-\omega_2) + \frac{1}{\omega_2^2}(S_{1,n}-\omega_{1,n})^2 \\
&- 4\frac{\omega_{1,n}}{\omega_2^3}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2) + 3\frac{\omega_{1,n}^2}{\omega_2^4}(S_{2,n}-\omega_2)^2 \\
&- 2\frac{1}{\omega_2^3}(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2) + 6\frac{\omega_{1,n}}{\omega_2^4}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2 \\
&- 4\frac{\omega_{1,n}^2}{\omega_2^5}(S_{2,n}-\omega_2)^3 + \frac{3}{\omega_2^4}(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)^2 \\
&- 8\frac{\omega_{1,n}}{\omega_2^5}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3 + 5\frac{\omega_{1,n}^2}{\omega_2^6}(S_{2,n}-\omega_2)^4, \\
Q(\Lambda_{t,n}^2,4) =& \frac{\omega_{3,n}^2}{\omega_2^2} + 2\frac{\omega_{3,n}}{\omega_2^2}(S_{3,n}-\omega_{3,n}) - 2\frac{\omega_{3,n}^2}{\omega_2^3}(S_{2,n}-\omega_2) + \frac{1}{\omega_2^2}(S_{3,n}-\omega_{3,n})^2 \\
&- 4\frac{\omega_{3,n}}{\omega_2^3}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2) + 3\frac{\omega_{3,n}^2}{\omega_2^4}(S_{2,n}-\omega_2)^2 \\
&- 2\frac{1}{\omega_2^3}(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2) + 6\frac{\omega_{3,n}}{\omega_2^4}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2 \\
&- 4\frac{\omega_{3,n}^2}{\omega_2^5}(S_{2,n}-\omega_2)^3 + \frac{3}{\omega_2^4}(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2 \\
&- 8\frac{\omega_{3,n}}{\omega_2^5}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3 + 5\frac{\omega_{3,n}^2}{\omega_2^6}(S_{2,n}-\omega_2)^4.
\end{aligned}$$

The fourth order Taylor polynomial of $\Theta_{t,n}\Lambda_{t,n}$ about the points $\omega_{1,n}$, ω_2 and $\omega_{3,n}$ is as follows:

$$\begin{aligned}
Q(\Theta_{t,n}\Lambda_{t,n},4) =& \frac{\omega_{1,n}\omega_{3,n}}{\omega_2^2} + \frac{\omega_{3,n}}{\omega_2^2}(S_{1,n}-\omega_{1,n}) - 2\frac{\omega_{1,n}\omega_{3,n}}{\omega_2^3}(S_{2,n}-\omega_2) \\
&+ \frac{\omega_{1,n}}{\omega_2^2}(S_{3,n}-\omega_{3,n}) + \frac{1}{\omega_2^2}(S_{1,n}-\omega_{1,n})(S_{3,n}-\omega_{3,n}) \\
&- 2\frac{\omega_{3,n}}{\omega_2^3}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2) - 2\frac{\omega_{1,n}}{\omega_2^3}(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n}) \\
&+ 3\frac{\omega_{1,n}\omega_{3,n}}{\omega_2^4}(S_{2,n}-\omega_2)^2 - 4\frac{\omega_{1,n}\omega_{3,n}}{\omega_2^5}(S_{2,n}-\omega_2)^3 \\
&- \frac{2}{\omega_2^3}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n}) \\
&+ 3\frac{\omega_{3,n}}{\omega_2^4}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2 + 3\frac{\omega_{1,n}}{\omega_2^4}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2 \\
&+ 5\frac{\omega_{1,n}\omega_{3,n}}{\omega_2^6}(S_{2,n}-\omega_2)^4 + \frac{3}{\omega_2^4}(S_{1,n}-\omega_{1,n})(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2 \\
&- 4\frac{\omega_{3,n}}{\omega_2^5}(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3 - 4\frac{\omega_{1,n}}{\omega_2^5}(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3.
\end{aligned}$$

We take expectations of the fourth order polynomials to obtain the approximations

$$\begin{aligned}
&E[\Theta_{t,n}] \approx \bar{E}[\Theta_{t,n},A] \approx E[Q(\Theta_{t,n},4)], \\
&E[\Lambda_{t,n}] \approx \bar{E}[\Lambda_{t,n},B] \approx E[Q(\Lambda_{t,n},4)], \\
&E[\Theta_{t,n}^2] \approx \bar{E}[\Theta_{t,n}^2,A] \approx E[Q(\Theta_{t,n}^2,4)], \\
&E[\Lambda_{t,n}^2] \approx \bar{E}[\Lambda_{t,n}^2,B] \approx E[Q(\Lambda_{t,n}^2,4)], \\
&E[\Theta_{t,n}\Lambda_{t,n}] \approx \bar{E}[\Theta_{t,n}\Lambda_{t,n},C] \approx E[Q(\Theta_{t,n}\Lambda_{t,n},4)].
\end{aligned}$$

Using these approximations, the MSFE approximation becomes

$$\begin{aligned}
MSFE_n \approx& E[Y_{t+1}^2] - 2E[Y_{t+1}X_t]E[Q(\Theta_{t,n},4)] - 2E[Y_{t+1}X_t]E[Q(\Lambda_{t,n},4)] \\
&+ E[X_t^2]\Big(E[Q(\Theta_{t,n}^2,4)] + 2E[Q(\Theta_{t,n}\Lambda_{t,n},4)] + E[Q(\Lambda_{t,n}^2,4)]\Big).
\end{aligned}$$

The central moments involved in the expectation of the Taylor polynomials are expanded to derive the n dependence. We write the central moments involved in the expectation of $Q(\Theta_{t,n},4)$, $Q(\Lambda_{t,n},4)$, $Q(\Theta^2_{t,n},4)$, $Q(\Lambda^2_{t,n},4)$, and $Q(\Theta_{t,n}\Lambda_{t,n},4)$:

$$E[(S_{1,n}-\omega_{1,n})]=0, \quad E[(S_{2,n}-\omega_2)]=0, \quad E[(S_{3,n}-\omega_{3,n})]=0,$$

$$\begin{aligned}
E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)] &= \frac{1}{n}\Big[E[Y_{1,t-n_b}X^3_{t-n_b-1}]-E[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}]\Big]\\
&\quad+\frac{n_b}{n^2}\Big[E[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}]-E[Y_{1,t-n_b}X^3_{t-n_b-1}]\Big],
\end{aligned}$$

$$E[(S_{1,n}-\omega_{1,n})(S_{3,n}-\omega_{3,n})]=0,$$

$$E[(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n})]=\frac{n_b}{n^2}\Big[E[Y_{2,t}X^3_{t-1}]-E[Y_{2,t}X_{t-1}X^2_{t-2}]\Big],$$

$$\begin{aligned}
E[(S_{1,n}-\omega_{1,n})^2] &= \frac{1}{n}\Big[E[Y^2_{1,t-n_b}X^2_{t-n_b-1}]-E[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}]\Big]\\
&\quad+\frac{n_b}{n^2}\Big[E[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}]-E[Y^2_{1,t-n_b}X^2_{t-n_b-1}]\Big],
\end{aligned}$$

$$E[(S_{3,n}-\omega_{3,n})^2]=\frac{n_b}{n^2}\Big[E[Y^2_{2,t}X^2_{t-1}]-E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}]\Big],$$

$$\begin{aligned}
E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2] &= \frac{1}{n^2}\Big[E[Y_{1,t-n_b}X^5_{t-n_b-1}]-E[Y_{1,t-n_b}X_{t-n_b-1}]E[X^4_{t-1}]\\
&\quad-2E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[X^2_{t-1}]+2E[Y_{1,t-n_b}X_{t-n_b-1}]E^2[X^2_{t-1}]\Big]\\
&\quad+\frac{n_b}{n^3}\Big[-E[Y_{1,t-n_b}X^5_{t-n_b-1}]+E[Y_{1,t-n_b}X_{t-n_b-1}]E[X^4_{t-1}]\\
&\quad+2E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[X^2_{t-1}]-2E[Y_{1,t-n_b}X_{t-n_b-1}]E^2[X^2_{t-1}]\Big],
\end{aligned}$$

$$\begin{aligned}
E[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2] &= \frac{n_b}{n^3}\Big[E[Y_{2,t}X^5_{t-1}]-E[Y_{2,t}X_{t-1}]E[X^4_{t-2}]\\
&\quad-2E[Y_{2,t}X^3_{t-1}]E[X^2_{t-2}]+2E[Y_{2,t}X_{t-1}]E^2[X_{t-2}]\Big],
\end{aligned}$$

$$\begin{aligned}
E[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)] &= \frac{1}{n^2}\Big[E[Y^2_{1,t-n_b}X^4_{t-n_b-1}]-E[Y^2_{1,t-n_b}X^2_{t-n_b-1}]E[X^2_{t-1}]\\
&\quad-2E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[Y_{1,t-n_b-1}X_{t-n_b-2}]\\
&\quad+2E^2[Y_{1,t-n_b}X_{t-n_b-1}]E[X^2_{t-1}]\Big]\\
&\quad+\frac{n_b}{n^3}\Big[-E[Y^2_{1,t-n_b}X^4_{t-n+b-1}]+E[Y^2_{1,t-n_b}X^2_{t-n_b-1}]E[X^2_{t-1}]\\
&\quad+2E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[Y_{1,t-n_b-1}X_{t-n_b-2}]\\
&\quad-2E^2[Y_{1,t-n_b}X_{t-n_b-1}]E[X^2_{t-1}]\Big],
\end{aligned}$$

$$\begin{aligned}
E[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)] &= \frac{n_b}{n^3}\Big[E[Y^2_{2,t}X^4_{t-1}]-E[Y^2_{2,t}X^2_{t-1}]E[X^2_{t-2}]\\
&\quad-2E[Y_{2,t}X^3_{t-1}]E[Y_{2,t-1}X_{t-2}]+2E^2[Y_{2,t}X_{t-1}]E[X^2_{t-2}]\Big],
\end{aligned}$$

$$E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n})]=0,$$

$$
\begin{aligned}
E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3] = {} & \frac{3}{n^2}\Big[E[Y_{1,t-n_b}X^3_{t-n_b-1}X^4_{t-n_b-2}] \\
& - E[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}X^2_{t-n_b-3}] - \omega_2 E[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}] \\
& + \omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}]\Big] \\
& + \frac{1}{n^3}\Big[E[Y_{1,t-n_b}X^7_{t-n_b-1}] - E[Y_{1,t-n_b}X_{t-n_b-1}X^6_{t-n_b-2}] \\
& - 3E[Y_{1,t-n_b}X^5_{t-n_b-1}X^2_{t-n_b-2}] - 3(n_b+1)E[Y_{1,t-n_b}X^3_{t-n_b-1}X^4_{t-n_b-2}] \\
& + (3n_b+6)\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}] \\
& + (3n_b+6)\omega_2 E[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}] \\
& - 3(n_b+2)\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}]\Big] \\
& + \frac{n_b}{n^4}\Big[-E[Y_{1,t-n_b}X^7_{t-n_b-1}] + E[Y_{1,t-n_b}X_{t-n_b-1}X^6_{t-n_b-2}] \\
& + 3E[Y_{1,t-n_b}X^5_{t-n_b-1}X^2_{t-n_b-2}] + 3E[Y_{1,t-n_b}X^3_{t-n_b-1}X^4_{t-n_b-2}] \\
& - 6E[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}X^2_{t-n_b-3}] \\
& - 6E[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}] \\
& + 6E[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-n_b-4}]\Big],
\end{aligned}
$$

$$
\begin{aligned}
E[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3] = {} & 3\frac{n_b}{n^3}\Big[-E[Y_{2,t}X_{t-1}X^4_{t-n_b-1}X^2_{t-n_b-2}] \\
& + E[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}] + E[Y_{2,t}X^3_{t-1}X^4_{t-n_b-1}] \\
& - E[Y_{2,t}X^3_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}]\Big] \\
& + \frac{n_b}{n^4}\Big[-E[Y_{2,t}X_{t-1}X^6_{t-2}] + 6E[Y_{2,t}X_{t-1}X^4_{t-2}X^2_{t-n_b-1}] \\
& - 6E[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3}X^2_{t-n_b-1}] + E[Y_{2,t}X^7_{t-1}] - 3E[Y_{2,t}X^5_{t-1}X^2_{t-2}] \\
& - 3E[Y_{2,t}X^3_{t-1}X^4_{t-2}] + 6E[Y_{2,t}X^3_{t-1}X^2_{t-2}X^2_{t-n_b-1}]\Big],
\end{aligned}
$$

$$
\begin{aligned}
E[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)^2] = {} & \frac{1}{n^2}\Big[E[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^4_{t-n_b-2}] \\
& - E[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}] \\
& - E[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^4_{t-n_b-3}] \\
& + 2E[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{1,t-n_b-1}X^3_{t-n_b-2}] \\
& - 4E[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X^3_{t-n_b-2}X^2_{t-n_b-3}] \\
& + 3E[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-n_b-4}]\Big] \\
& + \frac{1}{n^3}\Big[E[Y^2_{1,t-n_b}X^6_{t-n_b-1}] - (n_b+1)E[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^4_{t-n_b-2}]
\end{aligned}
$$

$$
\begin{aligned}
&- 2E[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2] \\
&+ (n_b+2)E[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2] \\
&- 2E[Y_{1,t-n_b} X_{t-n_b-1}^5 Y_{1,t-n_b-1} X_{t-n_b-2}] \\
&+ (n_b+2)E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^4] \\
&- 2(2n_b+1)E[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}^3] \\
&+ 8(n_b+1)E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}^3 X_{t-n_b-3}^2] \\
&- (5n_b+6)E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-n_b-4}^2]\Big] \\
&+ \frac{n_b}{n^4}\Big[- E[Y_{1,t-n_b}^2 X_{t-n_b-1}^6] + E[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^4] \\
&+ 2E[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2] - 2E[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2] \\
&+ 2E[Y_{1,t-n_b} X_{t-n_b-1}^5 Y_{1,t-n_b-1} X_{t-n_b-2}] \\
&- 2E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^4] \\
&+ 2(n_b+1)E[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}^3] \\
&- 4(n_b+2)E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}^3 X_{t-n_b-3}^2] \\
&+ 2(n_b+3)E[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-n_b-4}^2]\Big],
\end{aligned}
$$

$$
\begin{aligned}
E[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2] = \frac{n_b}{n^3}\Big[&E[Y_{2,t}^2 X_{t-1}^2 X_{t-n_b-1}^4] - E[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2] \\
&- E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-n_b-1}^4] \\
&+ E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-n_b-1}^2 X_{t-n_b-2}^2]\Big] \\
&+ \frac{n_b}{n^4}\Big[- E[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^4] + 2E[Y_{2,t}^2 X_{t-1}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2] \\
&+ 2E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-n_b-1}^4] \\
&+ 2(n_b-3)E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-n_b-1}^2 X_{t-n_b-2}^2] + E[Y_{2,t}^2 X_{t-1}^6] \\
&- 2E[Y_{2,t}^2 X_{t-1}^4 X_{t-2}^2] - 2E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^5] \\
&+ 2(n_b-1)E[Y_{2,t} X_{t-1}^3 Y_{2,t-1} X_{t-2}^3] \\
&- 4(n_b-2)E[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^3 X_{t-3}^2]\Big],
\end{aligned}
$$

$$
\begin{aligned}
E[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2(S_{3,n}-\omega_{3,n})] = 2\frac{n_b}{n^3}\Big[&\omega_2^2 E[Y_{2,t}X_{t-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] \\
&+ E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}^3] - \omega_2 E[Y_{2,t}X_{t-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^3] \\
&- \omega_2 E[Y_{2,t}X_{t-1}^3]E[Y_{1,t-n_b}X_{t-n_b-1}]\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\frac{n_b^2}{n^4}\Big[\omega_2 E[Y_{2,t}X_{t-1}]E[Y_{1,t-n_b}X^3_{t-n_b-1}] \\
&- \omega_2^2 E[Y_{2,t}X_{t-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] \\
&+ \omega_2 E[Y_{2,t}X^3_{t-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] - E[Y_{2,t}X^3_{t-1}]E[Y_{1,t-n_b}X^3_{t-n_b-1}]\Big].
\end{aligned}
$$

Substituting the above central moments into the expressions for the expectation of the fourth order Taylor polynomials $Q(\Theta_{t,n},4)$, $Q(\Lambda_{t,n},4)$, $Q(\Theta^2_{t,n},4)$, and $Q(\Lambda^2_{t,n},4)$, and substituting the expressions for these expectations in the expression for the MSFE approximation, we obtain the following:

$$
MSFE_n \approx C + \frac{A}{n} + \frac{B}{n^2} + \frac{D}{n^3} + \frac{E}{n^4} + \frac{F}{n^5} \equiv \overline{MSFE}_n, \tag{7.2.4}
$$

where

$$
\begin{aligned}
A =&\frac{1}{\omega_2^3}\Big[- 2(E[X^4_{t-1}] - n_b\omega_2^2)E[Y_{t+1}X_t]E[Y_{1,t-n_b}X_{t-n_b-1}] \\
&+ 2\omega_2 E[Y_{t+1}X_t]E[Y_{1,t-n_b}X^3_{t-n_b-1}] - 2n_b\omega_2^2 E[Y_{t+1}X_t]E[Y_{2,t}X_{t-1}] \\
&+ \omega_2^2 E[Y^2_{1,t-n_b}X^2_{t-n_b-1}] - 2n_b\omega_2^2 E^2[Y_{1,t-n_b}X_{t-n_b-1}] + 3E[X^4_{t-1}]E^2[Y_{1,t-n_b}X_{t-n_b-1}] \\
&- 4\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X^3_{t-n_b-1}] + 2n_b\omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\Big], \\
B =&\frac{1}{\omega_2^5}\Big[(15E^2[X^4_{t-1}] - 4\omega_2 E[X^6_{t-1}] - 3\omega_2^2 E[X^4_{t-1}](1+2n_b) \\
&+ n_b\omega_2^4(n_b-1))E^2[Y_{1,t-n_b}X_{,t-n_b-1}] \\
&+ (4\omega_2^3(2n_b+1) - 24\omega_2 E[X^4_{t-1}])E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] \\
&+ 6\omega_2^2 E^2[Y_{1,t-n_b}X^3_{t-n_b-1}] + (3\omega_2^2 E[X^4_{t-1}] - \omega_2^4(n_b+1))E[Y^2_{1,t-n_b}X^2_{t-n_b-1}] \\
&+ 6\omega_2^2 E[Y_{1,t-n_b}X^5_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] - 2\omega_2^3 E[Y^2_{1,t-n_b}X^4_{t-n_b-1}] \\
&- 4n_b\omega_2^3 E[Y_{2,t}X^3_{t-1}]E[Y_{1,t-n_b}X_{t-n_b-1}] + n_b\omega_2^4 E[Y^2_{2,t}X^2_{t-1}] \\
&+ (2\omega_2 E[X^6_{t-1}] - 6E^2[X^4_{t-1}] + 2\omega_2^2 E[X^4_{t-1}] + 8n_b\omega_2^2 E[X^4_{t-1}] \\
&+ 2n_b\omega_2^4(1-n_b))E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}] \\
&+ (6\omega_2 E[X^4_{t-1}] - 2\omega_2^3(2n_b+1))E[Y_{1,t-n_b}X^3_{t-n_b-1}]E[Y_{2,t}X_{t-1}] \\
&- 2\omega_2^2 E[Y_{1,t-n_b}X^5_{t-n_b-1}]E[Y_{2,t}X_{t-1}] \\
&+ 2n_b\omega_2^3 E[Y_{2,t}X^3_{t-1}]E[Y_{2,t}X_{t-1}]
\end{aligned}
$$

$$+ (-2n_b\omega_2^2 E[X_{t-1}^4] - n_b\omega_2^4(1-n_b))E^2[Y_{2,t}X_{t-1}]\Big],$$

$$C = \frac{1}{\omega_2}\Big[E[Y_{t+1}^2]\omega_2 - 2E[Y_{t+1}X_t]E[Y_{1,t-n_b}X_{t-n_b-1}] + E^2[Y_{1,t-n_b}X_{t-n_b-1}]\Big],$$

$$D = \frac{1}{\omega_2^5}\Big[(5E[X_{t-1}^8] - 15E^2[X_{t-1}^4](2n_b+1) + \omega_2 E[X_{t-1}^6](8n_b-12)$$

$$+ 3\omega_2^2 E[X_{t-1}^4](n_b^2+n_b+6) + n_b\omega_2^4(n_b-1))E^2[Y_{1,t-n_b}X_{t-n_b-1}]$$

$$- 8\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^7]$$

$$+ (24\omega_2 E[X_{t-1}^4](2n_b+1) - 4\omega_2^3(n_b^2+n_b+6))E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^3]$$

$$- 6\omega_2^2(2n_b+1)E^2[Y_{1,t-n_b}X_{t-n_b-1}^3] + 12n_b\omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}^3]$$

$$+ 3\omega_2^2 E[Y_{1,t-n_b}^2 X_{t-n_b-1}^6] + (3\omega_2^2 E[X_{t-1}^4](1-n_b) + \omega_2^4(n_b+6))E[Y_{1,t-n_b}^2 X_{t-n_b-1}^2]$$

$$+ n_b\omega_2^2(3E[X_{t-1}^4] - \omega_2^2)E[Y_{2,t}^2 X_{t-1}^2]$$

$$+ \omega_2^2(18-12n_b)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^5]$$

$$+ 6n_b\omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^5] + 2\omega_2^3(n_b-3)E[Y_{1,t-n_b}^2 X_{t-n_b-1}^4]$$

$$- 2n_b\omega_2^3 E[Y_{2,t}^2 X_{t-1}^4] + 4n_b\omega_2(n_b\omega_2^2 - 6E[X_{t-1}^4])E[Y_{2,t}X_{t-1}^3]E[Y_{1,t-n_b}X_{t-n_b-1}]$$

$$+ (6E^2[X_{t-1}^4](6n_b+1) - 2E[X_{t-1}^8] + \omega_2 E[X_{t-1}^6](6-10n_b)$$

$$- 2\omega_2^2 E[X_{t-1}^4](3n_b^2+n_b+6) + 2n_b\omega_2^4(1-n_b))E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]$$

$$+ 2\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}^7]E[Y_{2,t}X_{t-1}]$$

$$+ (-6\omega_2 E[X_{t-1}^4](1+5n_b) + 2\omega_2^3(2n_b^2+n_b+6))E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}]$$

$$+ (6n_b\omega_2 E[X_{t-1}^4] + 2n_b\omega_2^3(1-2n_b))E[Y_{2,t}X_{t-1}^3]E[Y_{2,t}X_{t-1}]$$

$$+ \omega_2^2(8n_b-6)E[Y_{1,t-n_b}X_{t-n_b-1}^5]E[Y_{2,t}X_{t-1}] - 2n_b\omega_2^2 E[Y_{2,t}X_{t-1}^5]E[Y_{2,t}X_{t-1}]$$

$$+ (2n_b\omega_2 E[X_{t-1}^6] - 6n_b E^2[X_{t-1}^4] + n_b\omega_2^2 E[X_{t-1}^4](3n_b-1)$$

$$+ n_b\omega_2^4(n_b-1))E^2[Y_{2,t}X_{t-1}]\Big],$$

$$E = \frac{1}{\omega_2^5}\Big[(15n_b E^2[X_{t-1}^4](n_b+2) - 10n_b E[X_{t-1}^8] + 4n_b\omega_2 E[X_{t-1}^6](6-n_b)$$

$$- 30n_b\omega_2^2 E[X_{t-1}^4] + n_b\omega_2^4(n_b-18))E^2[Y_{1,t-n_b}X_{t-n_b-1}]$$

$$+ 16n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^7]$$

$$- 8n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^7]$$

$$+ (72n_b\omega_2^3 - 24n_b\omega_2 E[X_{t-1}^4](2+n_b))E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^3]$$

$$+ (24n_b\omega_2 E[X_{t-1}^4](n_b+1) - 48n_b\omega_2^3)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^3]$$

$$
\begin{aligned}
&+6n_b\omega_2^2(n_b+1)E^2[Y_{1,t-n_b}X_{t-n_b-1}^3]-12n_b^2\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}^3]\\
&+6n_b\omega_2^2(n_b-1)E^2[Y_{2,t}X_{t-1}^3]-3n_b\omega_2^2E[Y_{1,t-n_b}^2X_{t-n_b-1}^6]+3n_b\omega_2^2E[Y_{2,t}^2X_{t-1}^6]\\
&+3n_b\omega_2^2(E[X_{t-1}^4]-2\omega_2^2)E[Y_{1,t-n_b}^2X_{t-n_b-1}^2]\\
&+3n_b\omega_2^2(2\omega_2^2-E[X_{t-1}^4])E[Y_{2,t}^2X_{t-1}^2]-6n_b\omega_2^3E[Y_{2,t}^2X_{t-1}^4]\\
&+6n_b\omega_2^2(n_b-7)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^5]\\
&+6n_b\omega_2^2(4-n_b)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^5]+6n_b\omega_2^3E[Y_{1,t-n_b}^2X_{t-n_b-1}^4]\\
&+(12n_bE[X_{t-1}^8]-6n_bE^2[X_{t-1}^4](6+5n_b)+n_b\omega_2E[X_{t-1}^6](8n_b-30)\\
&+36n_b\omega_2^2(E[X_{t-1}^4]+\omega_2^2)-2n_b^2\omega_2^4)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\\
&-10n_b\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}^7]E[Y_{2,t}X_{t-1}]+2n_b\omega_2E[Y_{2,t}X_{t-1}]E[Y_{2,t}X_{t-1}^7]\\
&+(6n_b\omega_2E[X_{t-1}^4](4n_b+5)-60n_b\omega_2^3)E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}]\\
&+(36n_b\omega_2^3-6n_b\omega_2E[X_{t-1}^4](4n_b+1))E[Y_{2,t}X_{t-1}^3]E[Y_{2,t}X_{t-1}]\\
&+6n_b\omega_2^2(5-n_b)E[Y_{1,t-n_b}X_{t-n_b-1}^5]E[Y_{2,t}X_{t-1}]\\
&+6n_b\omega_2^2(n_b-2)E[Y_{2,t}X_{t-1}^5]E[Y_{2,t}X_{t-1}]\\
&+(n_bE^2[X_{t-1}^4](15n_b+6)-2n_bE[X_{t-1}^8]+n_b\omega_2E[X_{t-1}^6](6-4n_b)-6n_b\omega_2^2E[X_{t-1}^4]\\
&+n_b\omega_2^4(n_b-18))E^2[Y_{2,t}X_{t-1}]\Big],\\
F=&(5n_b^2E[X_{t-1}^8]-15n_b^2E^2[X_{t-1}^4]-12n_b^2\omega_2E[X_{t-1}^6]+12n_b^2\omega_2^2E[X_{t-1}^4]\\
&+18n_b^2\omega_2^4)E^2[Y_{1,t-n_b}X_{t-n_b-1}]-8n_b^2\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^7]\\
&+8n_b^2\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^7]\\
&+24n_b^2\omega_2(E[X_{t-1}^4]-2\omega_2^2)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^3]\\
&+24n_b^2\omega_2(2n_b\omega_2^2-E[X_{t-1}^4])E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^3]\\
&+24n_b^2\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{1,t-n_b}X_{t-n_b-1}^5]\\
&-24n_b^2\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}^5]\\
&+(30n_b^2E^2[X_{t-1}^4]-10n_b^2E[X_{t-1}^8]+24n_b^2\omega_2E[X_{t-1}^6]-24n_b^2\omega_2^2E[X_{t-1}^4]\\
&-36n_b^2\omega_2^4)E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\\
&+8n_b^2\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}^7]E[Y_{2,t}X_{t-1}]-8n_b^2\omega_2E[Y_{2,t}X_{t-1}^7]E[Y_{2,t}X_{t-1}]\\
&+24n_b^2\omega_2(2\omega_2^2-E[X_{t-1}^4])E[Y_{1,t-n_b}X_{t-n_b-1}^3]E[Y_{2,t}X_{t-1}]\\
&+24n_b^2\omega_2(E[X_{t-1}^4]-2\omega_2^2)E[Y_{2,t}X_{t-1}^3]E[Y_{2,t}X_{t-1}]
\end{aligned}
$$

$$- 24n_b^2\omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}^5]E[Y_{2,t}X_{t-1}] + 24n_b^2\omega_2^2 E[Y_{2,t}X_{t-1}^5]E[Y_{2,t}X_{t-1}]$$
$$+ (5n_b^2 E[X_{t-1}^8] - 15n_b^2 E^2[X_{t-1}^4] - 12n_b^2\omega_2 E[X_{t-1}^6] + 12n_b^2\omega_2^2 E[X_{t-1}^4]$$
$$+ 18n_b^2\omega_2^4)E^2[Y_{2,t}X_{t-1}].$$

Now we proceed to estimate the MSFE for the case when $n < n_b$.

7.2.2 The MSFE for $n < n_b$

For $n < n_b$, the OLS estimator can be written as follows

$$\hat{\beta}_{t,n} = Q^{-1}\sum_{\tau=t-n}^{t-1} X_\tau Y_{2,\tau+1}, \quad Q = \sum_{\tau=t-n}^{t-1} X_\tau^2.$$

By independence, we can write

$$MSFE_n = E[Y_{t+1}^2] - 2E[Y_{t+1}X_t]E[\hat{\beta}_{t,n}] + E[X_t^2]E[\hat{\beta}_{t,n}^2].$$

$\hat{\beta}_{t,n}$ can be written as a function of the two statistics $S_{2,n}$ and $S_{4,n}$ as follows:

$$\hat{\beta}_{t,n} = \frac{S_{4,n}}{S_{2,n}},$$

where

$$S_{2,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2, \quad S_{4,n} = \frac{1}{n}\sum_{\tau=t-n}^{t-1} Y_{2,\tau+1}X_\tau. \tag{7.2.5}$$

As before, we write the expectation of the OLS estimator and its square as follows:

$$E[\hat{\beta}_{t,n}] = \bar{E}[\hat{\beta}_{t,n}, A] + \bar{E}[\hat{\beta}_{t,n}, A^c], \quad E[\hat{\beta}_{t,n}^2] = \bar{E}[\hat{\beta}_{t,n}^2, A] + \bar{E}[\hat{\beta}_{t,n}^2, A^c],$$

where A^c is the complement of A. Taylor series can be used in A to approximate $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$. To obtain further analytic results, we assume $P(X \in A) \approx 1$, so that $E[\hat{\beta}_{t,n}] \approx \bar{E}[\hat{\beta}_{t,n}, A]$ and $E[\hat{\beta}_{t,n}^2] \approx \bar{E}[\hat{\beta}_{t,n}^2, A]$. The Taylor series of $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$ are calculated about the points ω_2 and ω_4 for the statistics $S_{2,n}$, $S_{4,n}$. We define the points about which to calculate the Taylor series as follows:

$$\omega_2 \equiv E[S_{2,n}] = E[X_{t-1}^2], \quad \omega_4 \equiv E[S_{4,n}] = E[Y_{2,t}X_{t-1}],$$

where the equalities follow from the i.i.d. assumption. The fourth order Taylor polynomials of $\hat{\beta}_{t,n}$ and $\hat{\beta}_{t,n}^2$ about the points ω_2 and ω_4 are as follows:

$$\begin{aligned}
Q(\hat{\beta}_{t,n}, 4) =& \frac{\omega_4}{\omega_2} + \frac{1}{\omega_2}(S_{4,n} - \omega_4) - \frac{\omega_4}{\omega_2^2}(S_{2,n} - \omega_2) - \frac{1}{\omega_2^2}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2) \\
&+ \frac{\omega_4}{\omega_2^3}(S_{2,n} - \omega_2)^2 + \frac{1}{\omega_2^3}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2)^2 - \frac{\omega_{4,n}}{\omega_2^4}(S_{2,n} - \omega_2)^3 \\
&+ \frac{\omega_4}{\omega_2^5}(S_{2,n} - \omega_2)^4 - \frac{1}{\omega_2^4}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2)^3, \\
Q(\hat{\beta}_{t,n}^2, 4) =& \frac{\omega_4^2}{\omega_2^2} + 2\frac{\omega_4}{\omega_2^2}(S_{4,n} - \omega_4) - 2\frac{\omega_4^2}{\omega_2^3}(S_{2,n} - \omega_2) + \frac{1}{\omega_2^2}(S_{4,n} - \omega_4)^2 \\
&- 4\frac{\omega_4}{\omega_2^3}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2) + 3\frac{\omega_4^2}{\omega_2^4}(S_{2,n} - \omega_2)^2 \\
&- 2\frac{1}{\omega_2^3}(S_{4,n} - \omega_4)^2(S_{2,n} - \omega_2) + 6\frac{\omega_4}{\omega_2^4}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2)^2 \\
&- 4\frac{\omega_4^2}{\omega_2^5}(S_{2,n} - \omega_2)^3 + \frac{3}{\omega_2^4}(S_{4,n} - \omega_4)^2(S_{2,n} - \omega_2)^2 \\
&- 8\frac{\omega_4}{\omega_2^5}(S_{4,n} - \omega_4)(S_{2,n} - \omega_2)^3 + 5\frac{\omega_4^2}{\omega_2^6}(S_{2,n} - \omega_2)^4.
\end{aligned}$$

We take expectations of the fourth order polynomials to obtain the approximations

$$\begin{aligned}
E[\hat{\beta}_{t,n}] &\approx \bar{E}[\hat{\beta}_{t,n}, A] \approx E[Q(\hat{\beta}_{t,n}, 4)], \\
E[\hat{\beta}_{t,n}^2] &\approx \bar{E}[\hat{\beta}_{t,n}^2, A] \approx E[Q(\hat{\beta}_{t,n}^2, 4)].
\end{aligned}$$

Using these approximations, the MSFE approximation for $n < n_b$ becomes

$$MSFE_n \approx E[Y_{t+1}^2] - 2E[Y_{t+1}X_t]E[Q(\hat{\beta}_{t,n}, 4)] + E[X_t^2]E[Q(\hat{\beta}_{t,n}^2, 4)].$$

The central moments involved in the expectation of the Taylor polynomials are expanded to derive the n dependence. We write the central moments involved in the expectation of $Q(\hat{\beta}_{t,n},4)$ and $Q(\hat{\beta}^2_{t,n},4)$:

$$
\begin{aligned}
&E[(S_{2,n}-\omega_2)]=0, \quad E[(S_{4,n}-\omega_4)]=0,\\
&E[(S_{2,n}-\omega_2)(S_{4,n}-\omega_4)]=\frac{1}{n}\Big[E[Y_{2,t}X^3_{t-1}]-E[Y_{2,t}X_{t-1}X^2_{t-2}]\Big],\\
&E[(S_{4,n}-\omega_4)^2]=\frac{1}{n}\Big[E[Y^2_{2,t}X^2_{t-1}]-E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}]\Big],\\
&E[(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^2]=\frac{1}{n^2}\Big[E[Y_{2,t}X^5_{t-1}]-E[Y_{2,t}X_{t-1}]E[X^4_{t-2}]\\
&\qquad -2E[Y_{2,t}X^3_{t-1}]E[X^2_{t-2}]+2E[Y_{2,t}X_{t-1}]E^2[X_{t-2}]\Big],\\
&E[(S_{4,n}-\omega_4)^2(S_{2,n}-\omega_2)]=\frac{1}{n^2}\Big[E[Y^2_{2,t}X^4_{t-1}]-E[Y^2_{2,t}X^2_{t-1}]E[X^2_{t-2}]\\
&\qquad -2E[Y_{2,t}X^3_{t-1}]E[Y_{2,t-1}X_{t-2}]+2E^2[Y_{2,t}X_{t-1}]E[X^2_{t-2}]\Big],\\
&E[(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^3]=3\frac{1}{n^2}\Big[-E[Y_{2,t}X_{t-1}X^4_{t-n_b-1}X^2_{t-n_b-2}]\\
&\qquad +E[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}]+E[Y_{2,t}X^3_{t-1}X^4_{t-n_b-1}]\\
&\qquad -E[Y_{2,t}X^3_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}]\Big]\\
&\qquad +\frac{1}{n^3}\Big[-E[Y_{2,t}X_{t-1}X^6_{t-2}]+6E[Y_{2,t}X_{t-1}X^4_{t-2}X^2_{t-n_b-1}]\\
&\qquad -6E[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3}X^2_{t-n_b-1}]+E[Y_{2,t}X^7_{t-1}]-3E[Y_{2,t}X^5_{t-1}X^2_{t-2}]\\
&\qquad -3E[Y_{2,t}X^3_{t-1}X^4_{t-2}]+6E[Y_{2,t}X^3_{t-1}X^2_{t-2}X^2_{t-n_b-1}]\Big],\\
&E[(S_{4,n}-\omega_4)^2(S_{2,n}-\omega_2)^2]=\frac{1}{n^2}\Big[E[Y^2_{2,t}X^2_{t-1}X^4_{t-n_b-1}]-E[Y^2_{2,t}X^2_{t-1}X^2_{t-2}X^2_{t-3}]\\
&\qquad -E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^4_{t-n_b-1}]\\
&\qquad +E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^2_{t-n_b-1}X^2_{t-n_b-2}]\Big]\\
&\qquad +\frac{1}{n^3}\Big[-E[Y^2_{2,t}X^2_{t-1}X^4_{t-2}]+2E[Y^2_{2,t}X^2_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}]\\
&\qquad +2E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^4_{t-n_b-1}]\\
&\qquad +2(n_b-3)E[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^2_{t-n_b-1}X^2_{t-n_b-2}]+E[Y^2_{2,t}X^6_{t-1}]\\
&\qquad -2E[Y^2_{2,t}X^4_{t-1}X^2_{t-2}]-2E[Y_{2,t}X_{t-1}Y_{2,t-1}X^5_{t-2}]\\
&\qquad +2(n_b-1)E[Y_{2,t}X^3_{t-1}Y_{2,t-1}X^3_{t-2}]\\
&\qquad -4(n_b-2)E[Y_{2,t}X_{t-1}Y_{2,t-1}X^3_{t-2}X^2_{t-3}]\Big].
\end{aligned}
$$

Substituting the above central moments into the expressions for the expectation of the fourth order Taylor polynomials $Q(\hat{\beta}_{t,n},4)$ and $Q(\hat{\beta}^2_{t,n},4)$, and substituting the expressions for these expectations in the expression for the MSFE approximation, we obtain the following:

$$MSFE_n \approx C + \frac{A}{n} + \frac{B}{n^2} + \frac{D}{n^3} \equiv \overline{MSFE}_n, \tag{7.2.6}$$

where

$$\begin{aligned}
A =& \frac{1}{\omega_2^3}\Big[\omega_2^2 E[Y_{2,t}^2 X_{t-1}^2] - 2\omega_2 E[Y_{2,t}X_{t-1}]E[Y_{2,t}X_{t-1}^3] + E[X_{t-1}^4]E^2[Y_{2,t}X_{t-1}]\Big],\\
B =& \frac{1}{\omega_2^5}\Big[6\omega_2^2 E^2[Y_{2,t}X_{t-1}^3] + (3\omega_2^2 E[X_{t-1}^4] - \omega_2^4)E[Y_{2,t}^2 X_{t-1}^2]\\
& - 2\omega_2^3 E[Y_{2,t}^2 X_{t-1}^4] + 4\omega_2^2 E[Y_{2,t}X_{t-1}^5]E[Y_{2,t}X_{t-1}]\\
& + 2\omega_2(\omega_2^2 - 9E[X_{t-1}^4])E[Y_{2,t}X_{t-1}^3]E[Y_{2,t}X_{t-1}]\\
& + (9E^2[X_{t-1}^4] - 2\omega_2 E[X_{t-1}^6] - \omega_2^2 E[X_{t-1}^4])E^2[Y_{2,t}X_{t-1}]\Big],\\
C =& E[Y_{2,t}^2] - \frac{1}{\omega_2}E^2[Y_{2,t}X_{t-1}],\\
D =& -6\omega_2^2 E^2[Y_{2,t}X_{t-1}^3] + 3\omega_2^2 E[Y_{2,t}^2 X_{t-1}^6] + 3\omega_2^2(2\omega_2^2 - E[X_{t-1}^4])E[Y_{2,t}^2 X_{t-1}^2]\\
& - 6\omega_2^3 E[Y_{2,t}^2 X_{t-1}^4] - 6\omega_2 E[Y_{2,t}X_{t-1}^7]E[Y_{2,t}X_{t-1}]\\
& + 2\omega_2(9E[X_{t-1}^4] - 6\omega_2^2)E[Y_{2,t}X_{t-1}^3]E[Y_{2,t}X_{t-1}] + 12\omega_2^2 E[Y_{2,t}X_{t-1}^5]E[Y_{2,t}X_{t-1}]\\
& + (3E[X_{t-1}^8] - 9E^2[X_{t-1}^4] - 6\omega_2 E[X_{t-1}^6] + 6\omega_2^2 E[X_{t-1}^4])E^2[Y_{2,t}X_{t-1}].
\end{aligned}$$

7.3 Monte Carlo evidence

7.3.1 The experiments

In this section, we assess the approximation of the MSFE given by the two equations (7.2.4) and (7.2.6) by means of Monte Carlo simulations. This is carried out in two sets of experiments. For the two sets of tests, we analyze robustness of the accuracy of the algorithm to the variance of the processes involved. For the exposition, we adopt a linear structural break DGP with a shift in the linear parameter and a shift in the variance of the innovation process given by the following expression:

$$Y_{\tau+1} = \begin{cases} \theta_1 X_\tau + U_{1,\tau+1}, & \tau \leq t - n_b \\ \theta_2 X_\tau + U_{2,\tau+1}, & \tau > t - n_b \end{cases}, \tag{7.3.1}$$

with $\theta_1, \theta_2 \in \mathbb{R}$, $Var(U_{1,\tau}) = \sigma_1^2$, $Var(U_{2,\tau}) = \sigma_2^2$, $E[X_\tau] = \mu_x$, and $Var(X_\tau) = \sigma_x^2$. The forecast model is given by $Y_{t+1} = \beta X_t + V_{t+1}$, the forecast is given by $\hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t$, where $\hat{\beta}_{t,n}$ is the OLS estimator, and the forecast error is $\epsilon_{t+1,n} = Y_{t+1} - \hat{Y}_{t+1,n}$. The misspecification arises from not modeling the break. We want to compare the MSFE approximation obtained with the Taylor algorithm to a benchmark MSFE determined by Monte Carlo simulations. The motivation behind using Monte Carlo simulations to determine a benchmark MSFE lies in the fact that the MSFE is equal to the expected value of the conditional mean square forecast error (CMSFE)

$$MSFE = E[CMSFE], \quad CMSFE = E_t[\epsilon_{t+1,n}^2].$$

Given a realization of the processes $\{X_\tau\}_{\tau=t-n}^{t-1}$ and $\{Y_\tau\}_{\tau=t-n+1}^{t}$, it is simple to compute the CMSFE conditional on the given sample. Generating many such samples, M, by Monte Carlo simulations, we can construct M conditional mean square forecast errors, $\{CMSFE_i\}_{i=1}^{M}$, and approximate the MSFE by the sample mean of the simulations

$$MSFE \approx \frac{1}{M} \sum_{i=1}^{M} CMSFE_i.$$

We now describe the details involved in the construction of the benchmark MSFE. For a given set of values of the parameters $\mathcal{P} = \{\mu_x, \sigma_x, \sigma_1, \sigma_2, \theta_1, \theta_2, n_b\}$, ten thousand Monte Carlo simulations are conducted ($M = 10000$). Each of the M simulations is constructed as follows. First, we generate the series $\{x_\tau\}_{\tau=1}^{N}$ of length $N = 501$ as a realization of the explanatory process $\{X_\tau\}_{\tau=t-n}^{t}$ such that the first element of the series is the first observation, $1 \leftrightarrow t - n$, and the last element of the series is the last observation, $501 \leftrightarrow t$. Each x is a realization of a normally distributed random variable, $X \sim N(\mu_x, \sigma_x)$, and the population series is independent and identically distributed, $\{X_\tau\}_{\tau=t-n}^{t-1} \sim IID$. We split the series into two, $\mathcal{S}_1 = \{x_\tau\}_{\tau=1}^{t_{n_b}}$ and $\mathcal{S}_2 = \{x_\tau\}_{\tau=t_{n_b}+1}^{501}$, with $t_{n_b} = 500 - n_b$. $\mathcal{S}_1$

includes the values of X which occur prior to the break, and $\mathcal{S}_2$ includes the values of X subsequent to the break. With $\mathcal{S}_1$ and $\mathcal{S}_2$, we construct another two series. With $\mathcal{S}_1$, we construct $\{f_{1,\tau}\}_{\tau=1}^{t_{n_b}}$ by means of the relation $f_{1,\tau} = \theta_1 x_\tau$. With $\mathcal{S}_2$, we construct $\{f_{2,\tau}\}_{\tau=t_{n_b}+1}^{501}$ by means of the relation $f_{2,\tau} = \theta_2 x_\tau$. Finally, with the sample series $\{x_\tau\}_{\tau=1}^{501}$, $\{f_{1,\tau}\}_{\tau=1}^{t_{n_b}}$ and $\{f_{2,\tau}\}_{\tau=t_{n_b}+1}^{501}$, at the forecast origin $t = N-1$, we construct the CMSFE for $n > n_b$ as follows:

$$\begin{aligned}
CMSFE_n &= b^2_{\chi t,n} + v_{\chi t,n},\\
b^2_{\chi t,n} &= \left[f_{2,t} - x_t \frac{\sum_{\tau=N-n}^{N-n_b-1} f_{1,\tau} x_\tau}{\sum_{\tau=N-n}^{N-1} x_\tau^2} - x_t \frac{\sum_{\tau=N-n_b}^{N-1} f_{2,\tau} x_\tau}{\sum_{\tau=N-n}^{N-1} x_\tau^2}\right]^2,\\
v_{\chi t,n} &= \sigma_2^2 + \sigma_1^2 x_t^2 \frac{\sum_{\tau=N-n}^{N-n_b-1} x_\tau^2}{(\sum_{\tau=N-n}^{N-1} x_\tau^2)^2} + \sigma_2^2 x_t^2 \frac{\sum_{\tau=N-n_b}^{N-1} x_\tau^2}{(\sum_{\tau=N-n}^{N-1} x_\tau^2)^2},
\end{aligned}$$

and for $n \leq n_b$, the CMSFE is as follows:

$$\begin{aligned}
CMSFE_n &= b^2_{\chi t,n} + v_{\chi t,n},\\
b^2_{\chi t,n} &= \left[f_{2,t} - x_t \frac{\sum_{\tau=N-n}^{N-1} f_{2,\tau} x_\tau}{\sum_{\tau=N-n}^{N-1} x_\tau^2}\right]^2,\\
v_{\chi t,n} &= \sigma_2^2 + \frac{\sigma_2^2 x_t^2}{\sum_{\tau=N-n}^{N-1} x_\tau^2},
\end{aligned}$$

where $b^2_{\chi t,n}$ and $v_{\chi t,n}$ are the conditional squared bias and conditional variance of the forecast error, respectively. For each simulation, we obtain $N-1 = 500$ values of the CMSFE, one for each value of n starting from $n = 1$ to $n = 500$. The case $n = 1$ refers to estimation of the OLS carried out with only one observation. For a particular set of parameters $\mathcal{P}$, we obtain an array of size $M \times N-1$ of CMSFEs, $\{CMSFE_{i,j}\}_{i=1,j=1}^{M,N-1}$. Finally, the benchmark MSFE for a set of parameters $\mathcal{P}$ and for an observation window of size n is given by the following:

$$MSFE_n \approx \frac{1}{M} \sum_{i=1}^{M} CMSFE_{i,n}. \tag{7.3.2}$$

As mentioned previously, we conduct two set tests. These two sets of tests use the same procedure to calculate the benchmark MSFE, but differ in method by which the Taylor approximation, given by (7.2.4) and (7.2.6), is computed. In the first set of experiments, the goal is to test the Taylor approximation in the best-case scenario possible. The best-case scenario would be for the forecaster to have access to the population moments and population real autocovariances involved in the expressions (7.2.4) and (7.2.6). To simulate this best-case scenario, we use Monte Carlo simulations to approximate the population moments and population real autocovariances in question with their sample counterparts. Since the goal is to obtain close representations of population moments, we use large samples of the processes. Although the legitimacy of this practice must be questioned, we recall the goal of the first set of tests is to evaluate the mathematical adequacy of the Taylor algorithm, even if it is done in an unrealistic setting. The second set of tests will evaluate the Taylor algorithm under more realistic conditions reminiscent of empirical applications.

For the first set of tests, the Taylor algorithm is constructed by first generating a realization of the explanatory process $\{X_\tau\} \sim IIN(\mu_x, \sigma_x)$. This realization is given by the series $\{x_\tau\}_{\tau=1}^{L}$ with L equal to one million. This series is divided into two series $\mathcal{X}_1 = \{x_\tau\}_{\tau=1}^{[L/2]}$ and $\mathcal{X}_2 = \{x_\tau\}_{\tau=[L/2]+1}^{L}$ ($[\,\cdot\,]$ stands for the integer part of the argument). Next, we generate a realization of the innovation processes $\{U_{1,\tau}\} \sim IIN(0, \sigma_1)$ and $\{U_{2,\tau}\} \sim IIN(0, \sigma_2)$ given by $\mathcal{U}_1 = \{u_{1,\tau}\}_{\tau=1}^{[L/2]}$ and $\mathcal{U}_2 = \{u_{2,\tau}\}_{\tau=[L/2]+1}^{L}$, respectively. Finally, we construct a realization of the dependent process $\{Y_\tau\}$ using $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{U}_1$, and $\mathcal{U}_2$. This realization of the dependent variable is given by the two series $\mathcal{Y}_1 = \{y_{1,\tau}\}_{\tau=1}^{[L/2]}$ and $\mathcal{Y}_2 = \{y_{2,\tau}\}_{\tau=[L/2]+1}^{L}$. $\mathcal{Y}_1$ is constructed by means of the relation $y_{1,\tau} = \theta_1 x_\tau + u_{1,\tau}$ for $x_\tau \in \mathcal{X}_1$, and $\mathcal{Y}_2$ is constructed by means of the relation $y_{2,\tau} = \theta_2 x_\tau + u_{2,\tau}$ for $x_\tau \in \mathcal{X}_2$. The population moments in (7.2.4) and (7.2.6) are estimated by generating their sample counterparts with $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$. For example:

$$E[Y_{1,t-n_b} X^3_{t-n_b-1}] \approx \frac{1}{[L/2]} \sum_{\tau=1}^{[L/2]} y_{1,\tau} x^3_\tau,$$

$$E[Y^2_{2,t} X^2_{t-1}] \approx \frac{1}{[L/2]} \sum_{\tau=[L/2]+1}^{L} y^2_{2,\tau} x^2_\tau.$$

Therefore, for a given set of the parameters, $\mathcal{P} = \{\mu_x, \sigma_x, \sigma_1, \sigma_2, \theta_1, \theta_2, n_b\}$, we can generate the series $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, $\mathcal{Y}_2$, the necessary sample moments, and ultimately evaluate

(7.2.4) and (7.2.6) for different values of the observation window size n. The resulting MSFE can be compared to the benchmark MSFE (7.3.2). The sets of parameter values investigated in the best-case scenario, and the reference to their corresponding MSFE, plots are given in Table 7.1.

We must clarify one issue with this procedure. By the construction of the series $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$, the structural break seems to occur in the middle of the series at $t = [L/2]$. This would seem to make $n_b = [L/2]$. This is not what we want and is not what is done in the first set of experiments. The series were taken long in each direction from the break to obtain good approximations of the population moments. We do not mean to fix $n_b = [L/2]$ but rather, the way to think of this artificial procedure is as follows: person A observed a large amount of data of size L with a structural break in the middle; person A computes sample moments as described above; person B has observed only a fraction of the data available to person A with the latest observation at time t and with the break occurring at time $t - n_b$; person A gives her sample moment calculations to person B; person B uses those sample moments together with (7.2.4) and (7.2.6) to calculate the Taylor algorithm of the MSFE. Although artificial, this procedure serves to explore the robustness of the Taylor algorithm for the MSFE in the best-case scenario when the best possible sample moments are available.

The second set of experiments also has as the main goal evaluation of the robustness of the Taylor algorithm, but under more practical considerations than the first set of tests. For these tests, we take the role of person B in the above description, without any input from person A. That is, person B must estimate sample moments with the available data as one would do in any real empirical application. The procedure is the same as previously described except for the definitions of the series. In the second set of tests, we let $\mathcal{X}_1 = \{x_\tau\}_{\tau=1}^{t_{n_b}}$, $\mathcal{X}_2 = \{x_\tau\}_{\tau=t_{n_b}+1}^{L}$, $\mathcal{U}_1 = \{u_{1,\tau}\}_{\tau=1}^{t_{n_b}}$, $\mathcal{U}_2 = \{u_{2,\tau}\}_{\tau=t_{n_b}+1}^{L}$, $\mathcal{Y}_1 = \{y_{1,\tau}\}_{\tau=1}^{t_{n_b}}$, and $\mathcal{Y}_2 = \{y_{2,\tau}\}_{\tau=t_{n_b}+1}^{L}$, so that $n_b = L - t_{n_b}$. With these definitions, the accuracy of the sample moments will depend on the size of n_b and L. This suggests that the accuracy of the Taylor algorithm will depend on the amount of post-break data available.

7.3.2 Discussion

Most of the issues we discuss regarding the Monte Carlo simulations carried out are summarized in Tables 7.1 and 7.2. As mentioned, two sets of experiments were performed. The first set of tests involved large data series to obtain accurate sample moments necessary for the calculation of the MSFE with the Taylor algorithm given by (7.2.4) and (7.2.6). For the case with $\mu_x = 10, \theta_1 = 2, \theta_2 = 2.5$, 23 experiments are performed with varying values of σ_x, σ_1, and σ_2. For all 23 experiments, the MSFE obtained with the Taylor algorithm and the benchmark MSFE seem in close agreement. Out of these 23 experiments, the benchmark MSFE has an optimal observation window in 18 of the cases. Out of these 18 cases, the optimal observation window of the Taylor algorithm MSFE agrees with the benchmark in 15 cases. In two cases, the optimal observation windows differ by one observation, and in one case by seven observations. The worst performance of the Taylor algorithm occurs for the case with the highest process variances, $\sigma_x = 10, \sigma_1 = 40, \sigma_2 = 40$.

For the case with $\mu_x = 0, \theta_1 = 2, \theta_2 = 2.5$, 23 experiments are also performed with varying values of σ_x, σ_1, and σ_2. Out of these 23 experiments, the benchmark MSFE has an optimal observation window in 14 of the cases. Out of these 14 cases, the optimal observation window of the Taylor algorithm MSFE agrees with the benchmark in 7 cases. The results seem to indicate performance worsens as σ_1 and σ_2 increase, not necessarily as σ_x increases. This can be observed by comparing the experiments with $\sigma_x = 1$ with those experiments with $\sigma_x = 10$ for the different values of σ_1 and σ_2. One can get intuition for this by examining the dependence of the MSFE for a correctly specified model on the variance of the innovation

$$MSFE = \sigma_U \left(1 + E\left[\frac{1}{\sum_{t-n}^{t-1} X_s^2}\right]\right) \to \sigma_U, \quad \text{as} \quad n \to \infty.$$

Based on this, one can understand how the Taylor algorithm MSFE can be sensitive to a volatile innovation. The worst performance occurred for the higher values of σ_1 and σ_2.

Although, for the case with $\mu_x = 10, \theta_1 = 2, \theta_2 = 2.05$, the patter of performance was similar across different values of σ_x, σ_1, and σ_2 to the other cases, the overall performance is worse than the previous two cases. Out of 17 experiments, there is agreement among the observation windows only in three tests. One explanation for this eventuality is that, when the parameter shift is small, from $\theta_1 = 2$ to $\theta_2 = 2.05$, the OLS estimator has difficulty detecting the change over the volatility of the processes. This is turn, translates to a less accurate estimate of the optimal observation window by the Taylor algorithm.

The second set of experiments makes use of smaller data samples in order to replicate an empirical setting. The performance of the Taylor algorithm is expected to worsen from that in the previous set of experiments with large data samples. Two cases are examined, the first with sample size $L = 2000$ and $n_b = 20$ and the second with sample size $L = 5000$ and $n_b = 100$. For the first case with $L = 2000$ and $n_b = 20$, the Taylor algorithm fails to identify the benchmark optimal window 15 times out of 18. In the other three tests, the Taylor algorithm misses the benchmark optimal window by 1, 2, and 5 observations. For the case with $L = 5000$ and $n_b = 100$, the results are more promising. In this case, the Taylor algorithm fails to identify the benchmark optimal window 4 times out of 18. The Taylor algorithm provides the correct optimal window in three experiments. In the other 11 experiments, the difference between the Taylor algorithm optimal window and the benchmark optimal window ranges from 1 observation to 78 observations. The performance across different values of σ_x, σ_1, and σ_2 follows the same pattern as in the first set of experiments with performance decaying with increasing values of σ_1 and σ_2.

We do not present results for Monte Carlo experiments with parameter shift in the standard deviations σ_1 and σ_2 because we do not find significant difference in the accuracy and the patterns of performance of the Taylor algorithm as presented for parameter shift in the linear parameters θ_1 and θ_2.

			$\mu_x = 10$ $\theta_1 = 2, \theta_2 = 2.5$		$\mu_x = 0$ $\theta_1 = 2, \theta_2 = 2.5$		$\mu_x = 10$ $\theta_1 = 2, \theta_2 = 2.05$	
σ_x	σ_1	σ_2	figure		figure		figure	
0.1	0	0	7.1	NA	7.24	NA	7.47	NA
	0.1	0.1	7.2	cw	7.25	cw	7.48	cw
	1	1	7.3	cw	7.26	$n > 500$	7.49	cw
1	0	0	7.4	NA	7.27	NA	7.50	NA
	0.5	0.5	7.5	cw	7.28	cw	7.51	$\Delta = 1$
	1	1	7.6	cw	7.29	cw	7.52	cw
	3	3	7.7	cw	7.30	$\Delta = 125$	7.53	$\Delta = 11$
	5	5	7.8	$\Delta = 1$	7.31	$n > 500$	7.54	$n > 500$
3	0	0	7.9	NA	7.32	NA	7.55	NA
	1	1	7.10	cw	7.33		7.56	$\Delta = 1$
	3	3	7.11	cw	7.34	cw	7.57	$\Delta = 7$
	5	5	7.12	cw	7.35	$\Delta = 1$	7.58	$n > 500$
	10	10	7.13	cw	7.36	$n > 500$	—	
5	0	0	7.14	NA	7.37	NA	7.59	NA
	1	1	7.15	cw	7.38		7.60	$\Delta = 1$
	5	5	7.16	cw	7.39	cw	7.61	$n > 500$
	10	10	7.17	cw	7.40	$\Delta = 1$	—	
	15	15	7.18	cw	7.41	$\Delta = 125$	—	
10	0	0	7.19	NA	7.42	NA	7.62	NA
	5	5	7.20	cw	7.43	cw	7.63	$n > 500$
	10	10	7.21	$\Delta = 1$	7.44	cw	—	
	20	20	7.22	cw	7.45	$\Delta = 1$	—	
	40	40	7.23	$\Delta = 7$	7.46	$n > 500$	—	

Table 7.1: Sets of parameter values for the best-case scenario experiments, $L = 1 \times 10^6$. cw indicates the optimal observation window given by the Taylor algorithm and the benchmark MSFE coincide. $\Delta = a$ indicates that the absolute difference between the optimal observation window given by the Taylor algorithm and the optimal observation window given by the benchmark MSFE is equal to the integer a. $n > 500$ indicates the optimal observation window does not occur within the observed sample of 500. NA indicates no optimal observation window exists in the benchmark MSFE.

			$L = 2000, n_b = 20$ $\mu_x = 10$ $\theta_1 = 2, \theta_2 = 2.5$		$L = 5000, n_b = 100$ $\mu_x = 10$ $\theta_1 = 2, \theta_2 = 2.5$	
σ_x	σ_1	σ_2	figure		figure	
0.1	0	0	7.64	NA	7.87	NA
	0.1	0.1	7.65	F	7.88	F
	1	1	7.66	F	7.89	cw
1	0	0	7.67	NA	7.90	NA
	0.5	0.5	7.68	F	7.91	F
	1	1	7.69	F	7.92	F
	3	3	7.70	F	7.93	cw
	5	5	7.71	$\Delta = 1$	7.94	cw
3	0	0	7.72	NA	7.95	NA
	1	1	7.73	F	7.96	$\Delta = 1$
	3	3	7.74	F	7.97	$\Delta = 1$
	5	5	7.75	F	7.98	$\Delta = 2$
	10	10	7.76	$\Delta = 2$	7.99	$\Delta = 3$
5	0	0	7.77	NA	7.100	NA
	1	1	7.78	F	7.101	$\Delta = 3$
	5	5	7.79	F	7.102	$\Delta = 5$
	10	10	7.80	F	7.103	$\Delta = 8$
	15	15	7.81	$\Delta = 5$	7.104	$\Delta = 17$
10	0	0	7.82	NA	7.105	NA
	5	5	7.83	F	7.106	$\Delta = 19$
	10	10	7.84	F	7.107	$\Delta = 19$
	20	20	7.85	F	7.108	$\Delta = 78$
	40	40	7.86	F	7.109	F

Table 7.2: Sets of parameter values for the experiments with limited samples and $n_b = 20$, $n_b = 100$. cw indicates the optimal observation window given by the Taylor algorithm and the benchmark MSFE coincide. $\Delta = a$ indicates that the absolute difference between the optimal observation window given by the Taylor algorithm and the optimal observation window given by the benchmark MSFE is equal to the integer a. NA indicates no optimal observation window exists in the benchmark MSFE. F indicates the Taylor algorithm has failed to identify an optimal observation window.

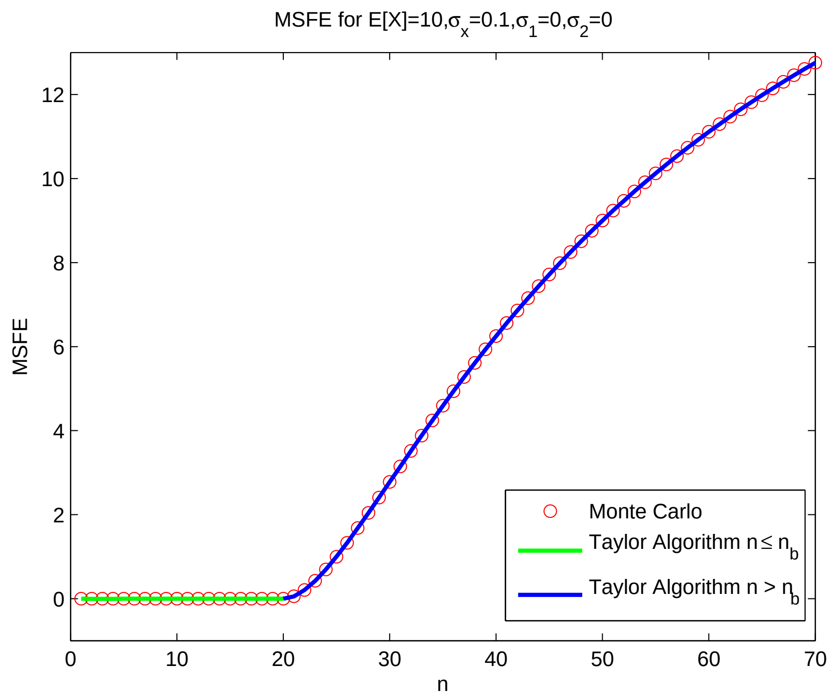


Figure 7.1: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

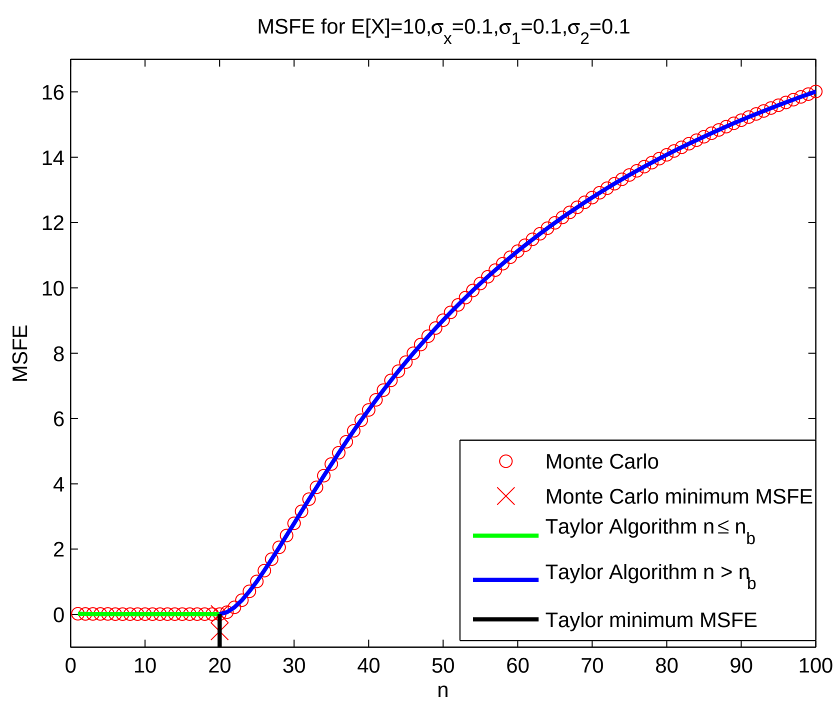


Figure 7.2: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 20$, Taylor algorithm minimum MSFE $= 20$

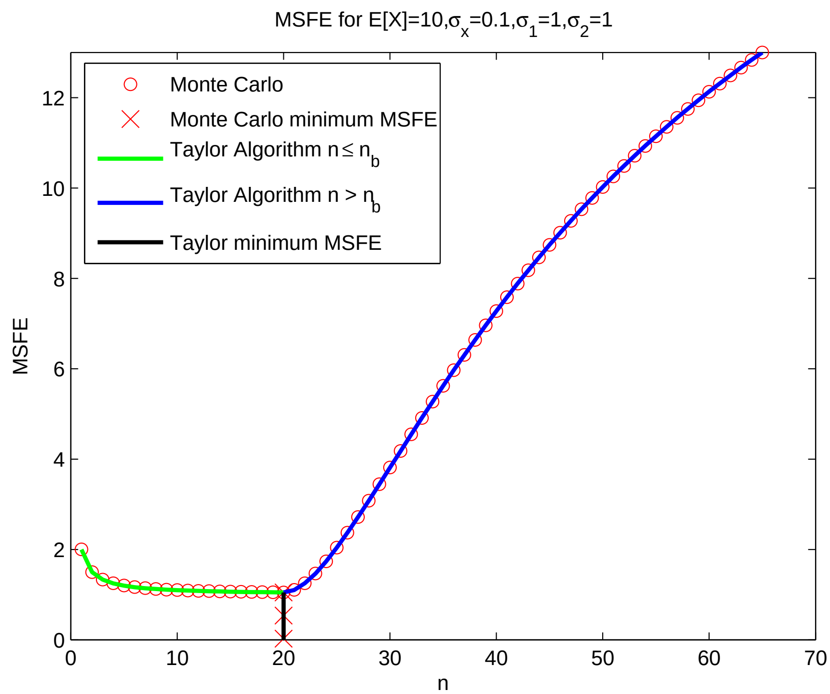


Figure 7.3: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

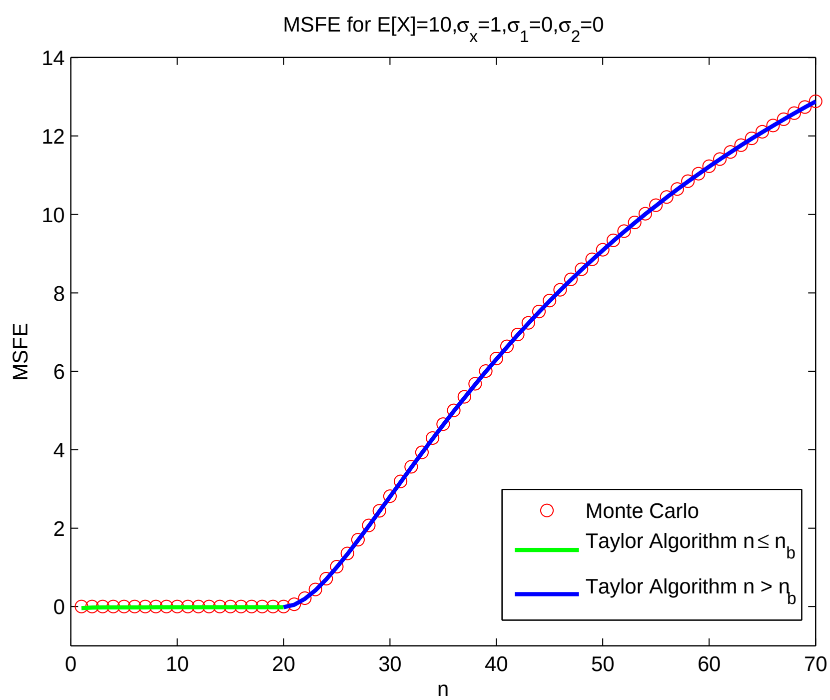


Figure 7.4: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

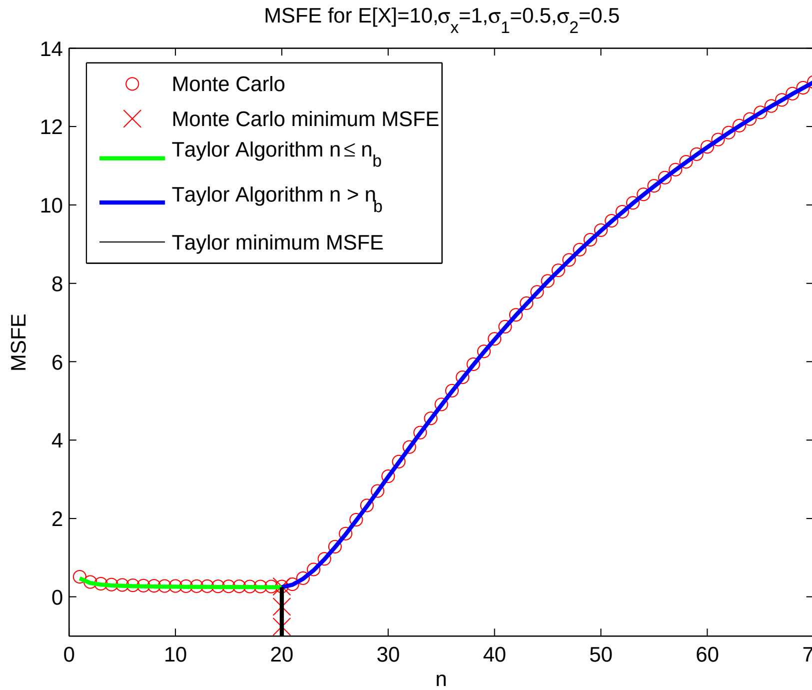


Figure 7.5: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

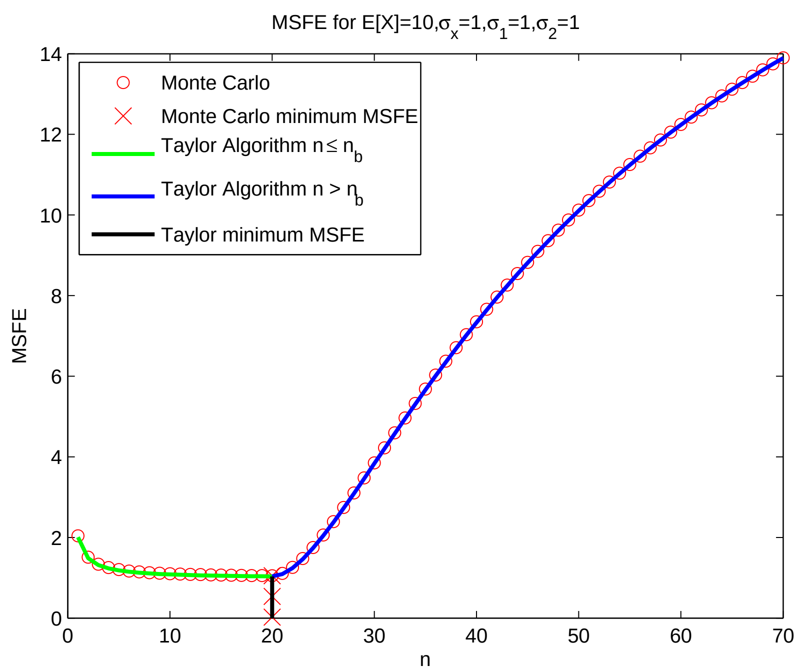


Figure 7.6: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

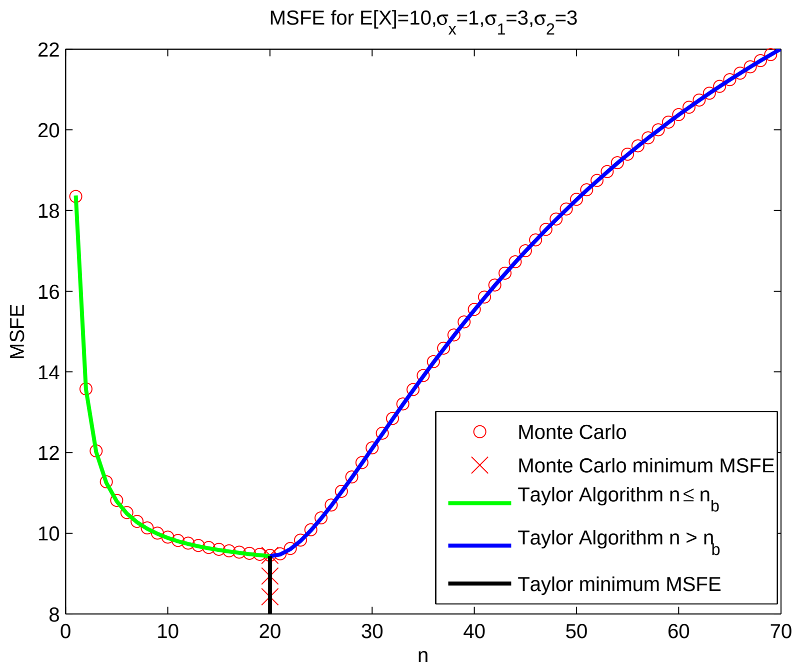


Figure 7.7: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

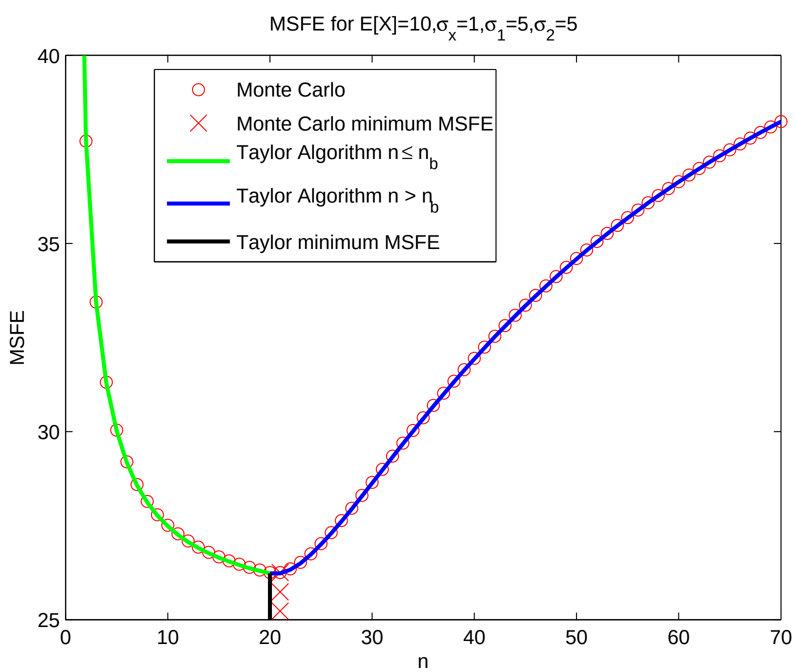


Figure 7.8: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE =21, Taylor algorithm minimum MSFE = 20

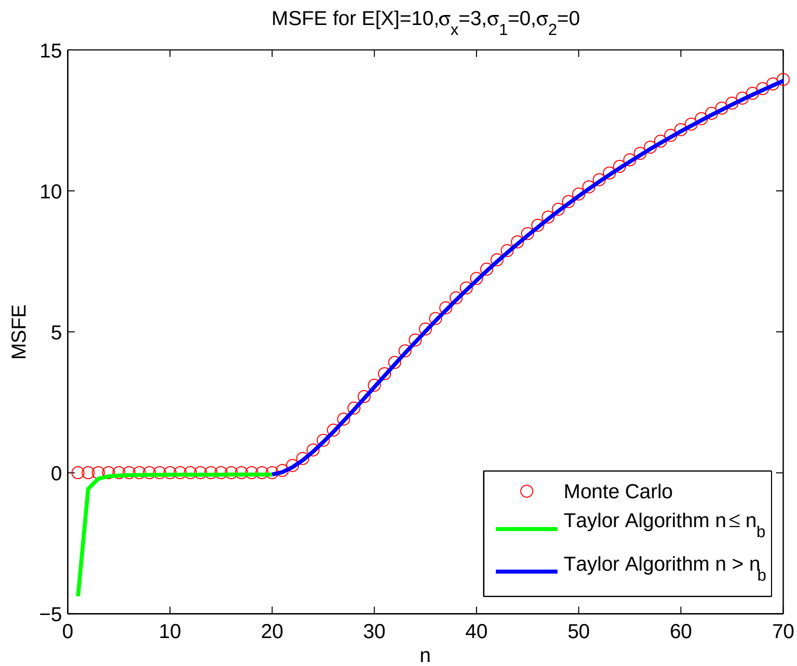


Figure 7.9: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

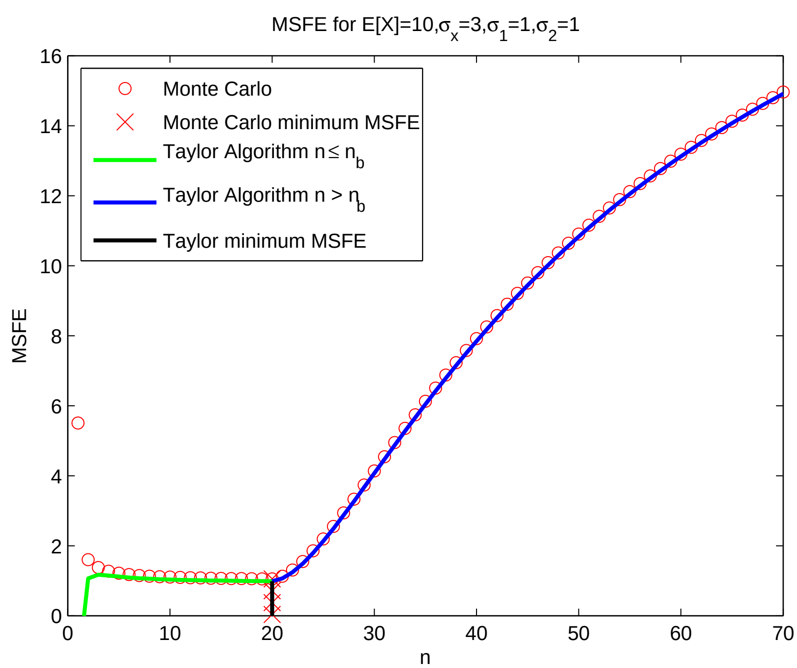


Figure 7.10: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 20$, Taylor algorithm minimum MSFE $= 20$

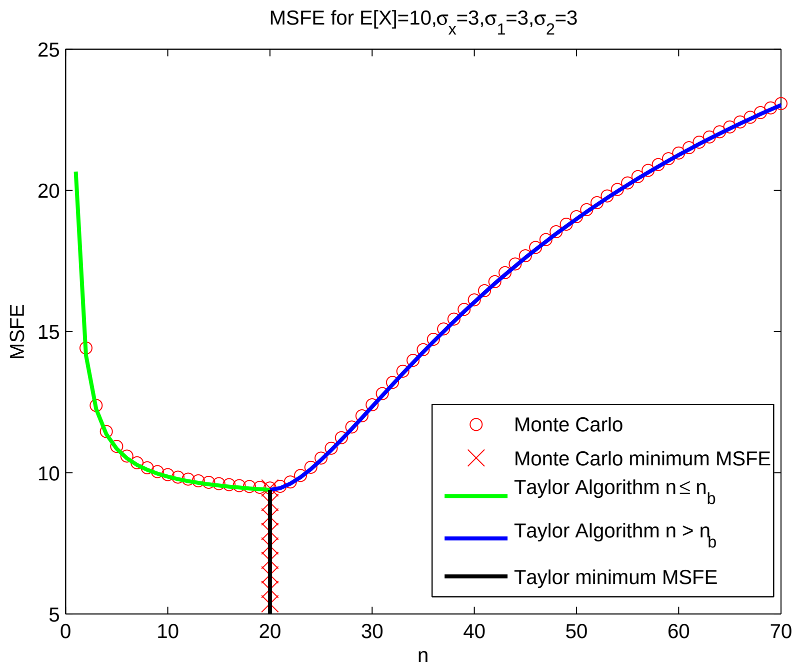


Figure 7.11: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

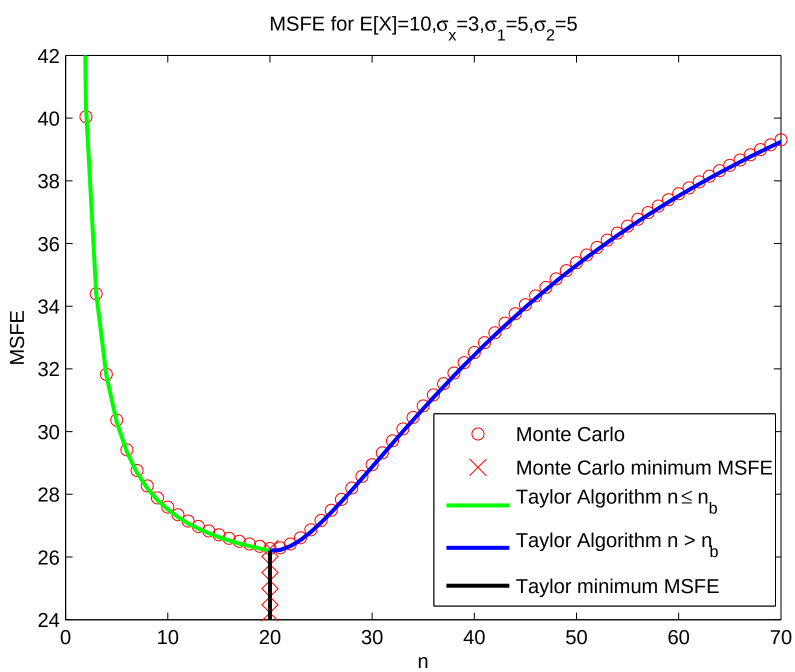


Figure 7.12: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

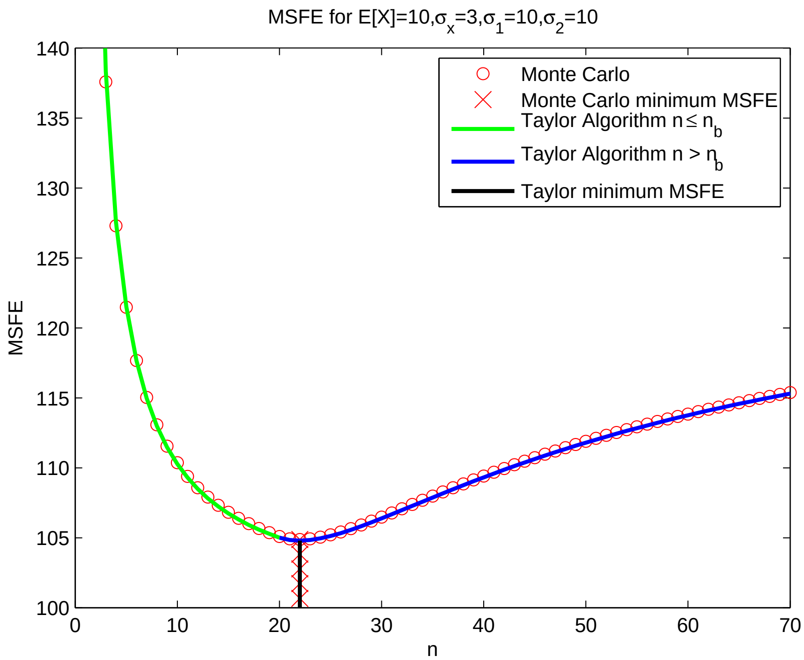


Figure 7.13: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE =22, Taylor algorithm minimum MSFE =22

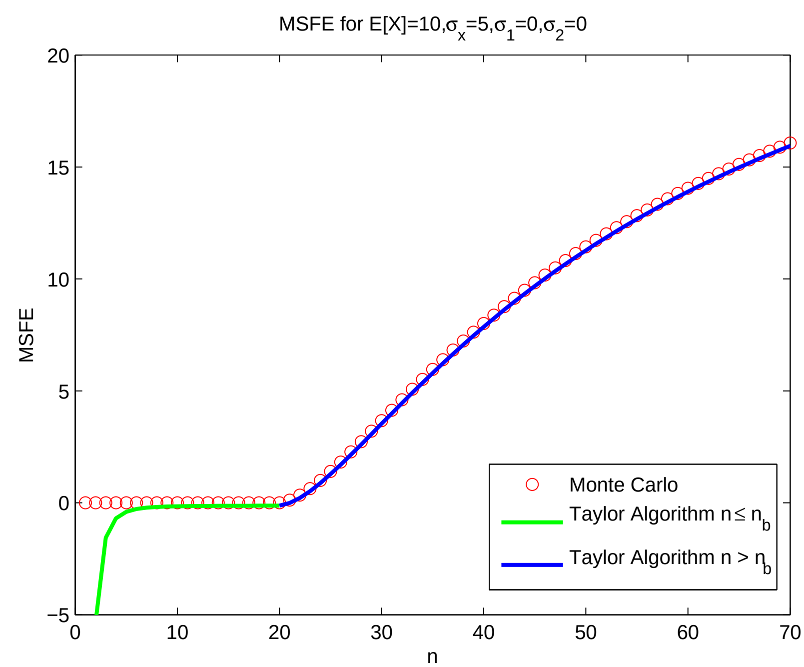


Figure 7.14: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

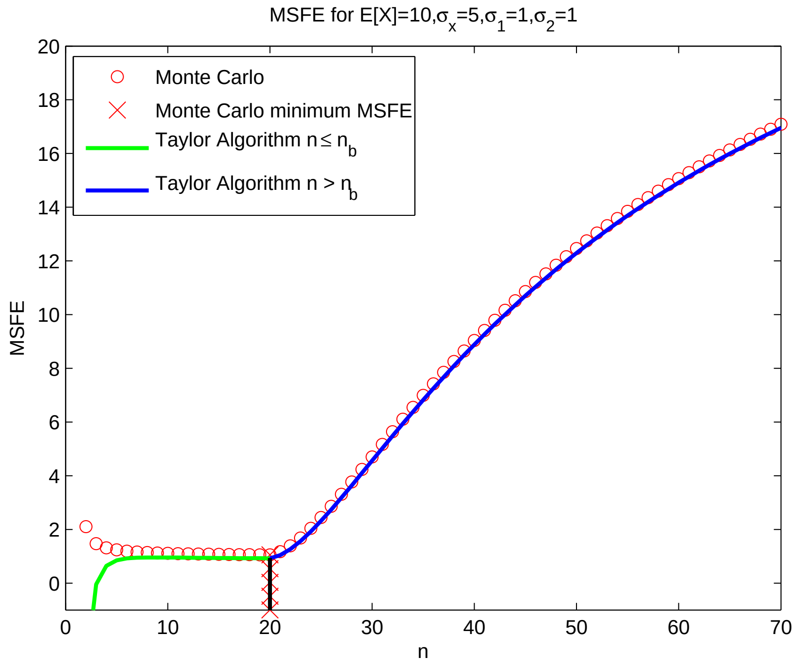


Figure 7.15: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$. Monte Carlo minimum MSFE = 20, Taylor algorithm local minimum MSFE = 20

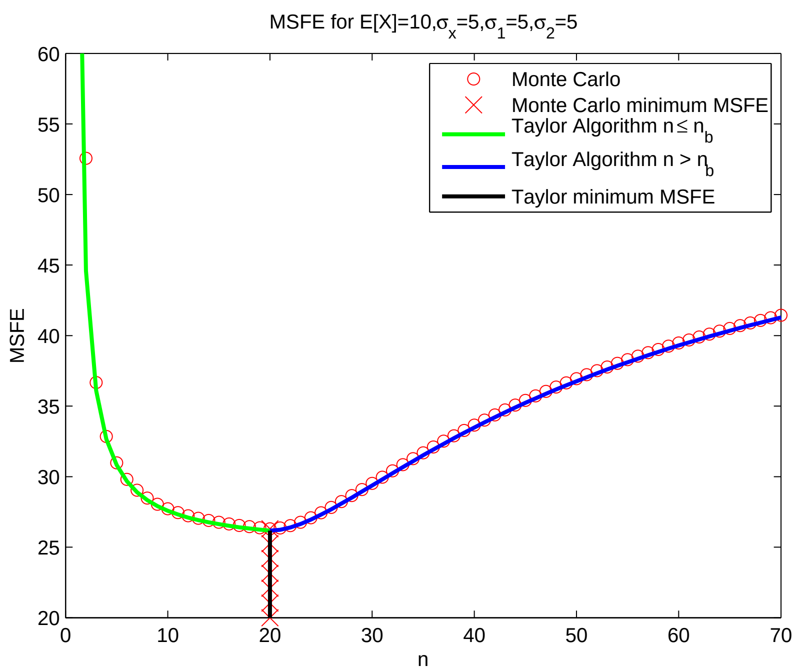


Figure 7.16: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm minimum MSFE = 20

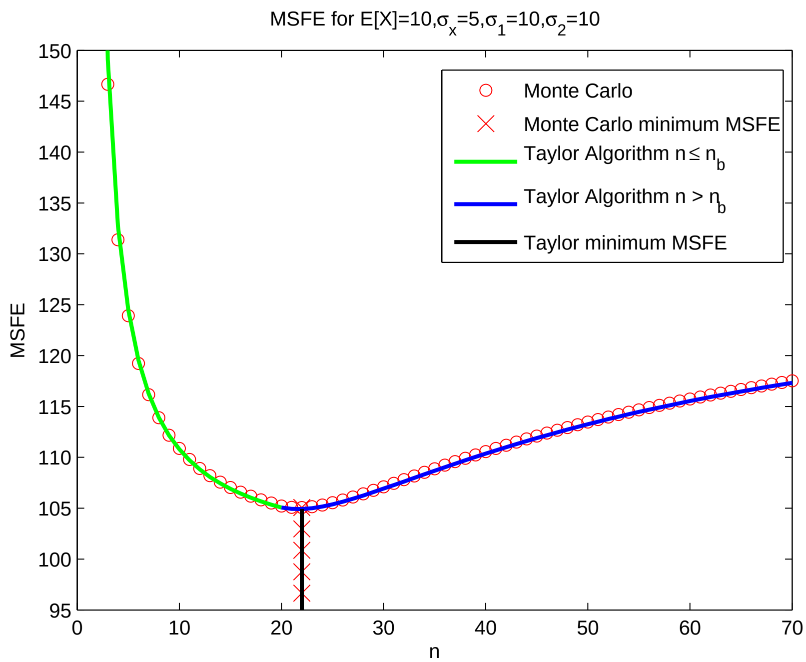


Figure 7.17: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE =22, Taylor algorithm minimum MSFE =22

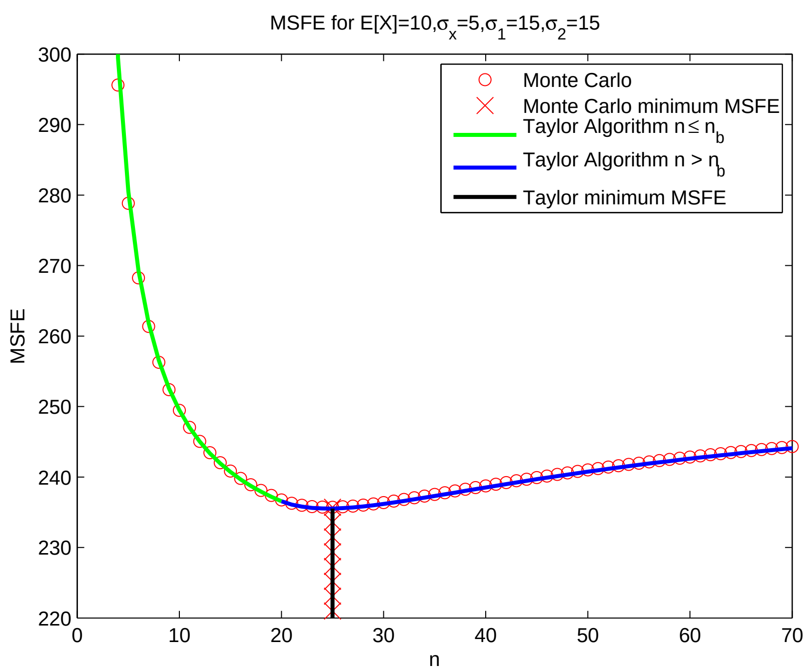


Figure 7.18: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 15$, $\sigma_2 = 15$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE =25, Taylor algorithm minimum MSFE =25

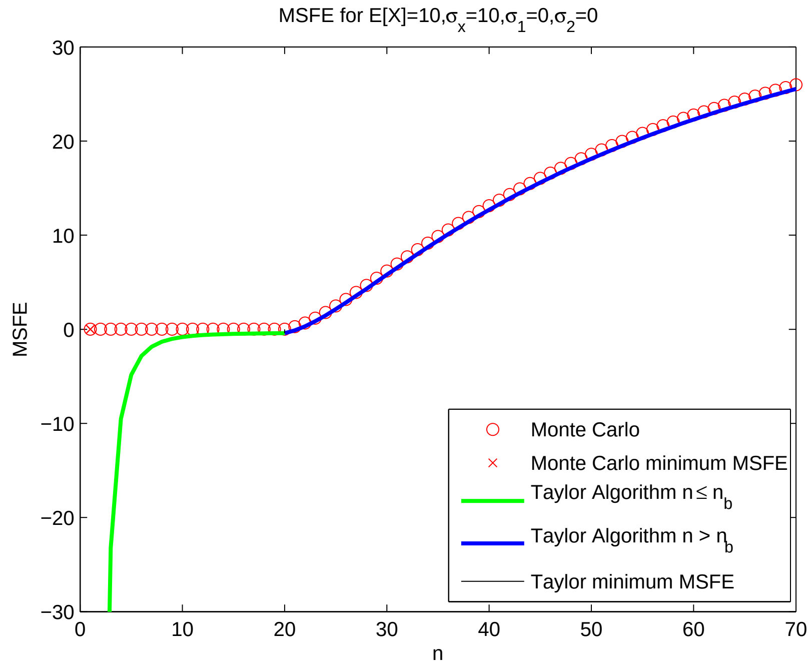


Figure 7.19: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

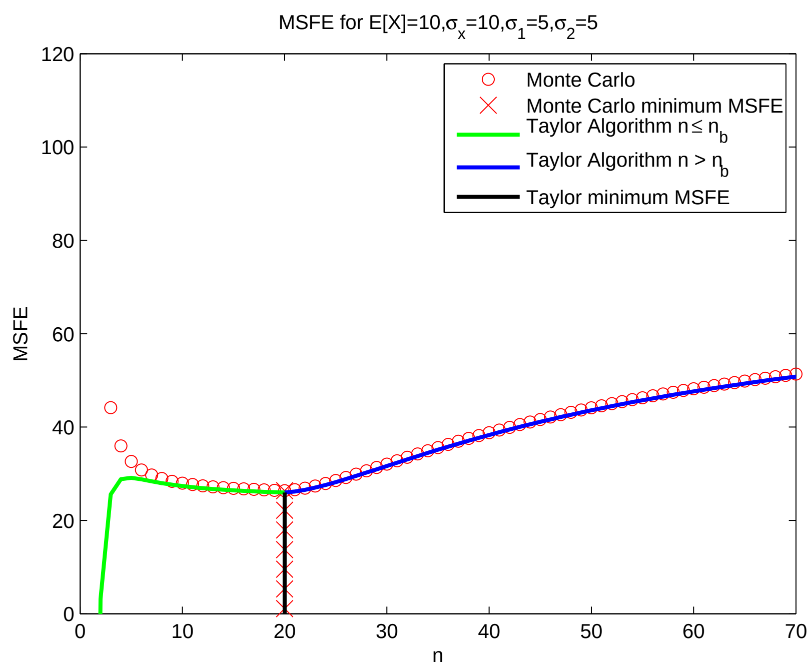


Figure 7.20: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 20$, Taylor algorithm minimum MSFE $= 20$

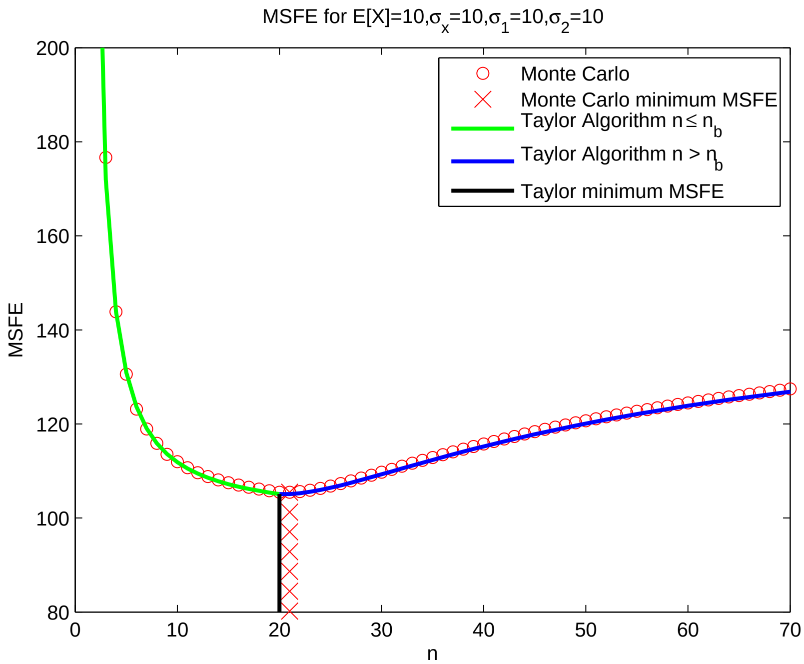


Figure 7.21: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 21$, Taylor algorithm minimum MSFE $= 20$

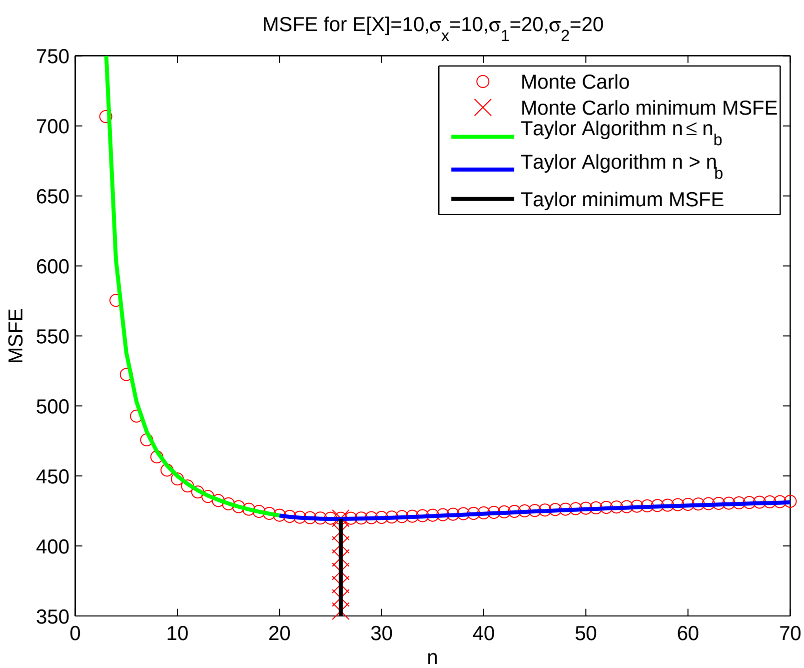


Figure 7.22: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 20$, $\sigma_2 = 20$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 26$, Taylor algorithm minimum MSFE $= 26$

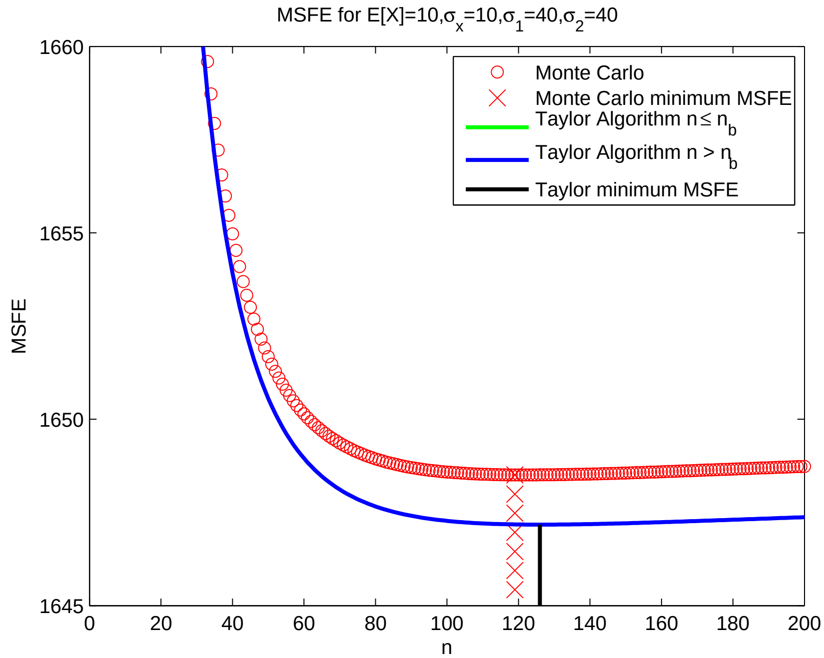


Figure 7.23: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 40$, $\sigma_2 = 40$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 119, Taylor algorithm minimum MSFE = 126

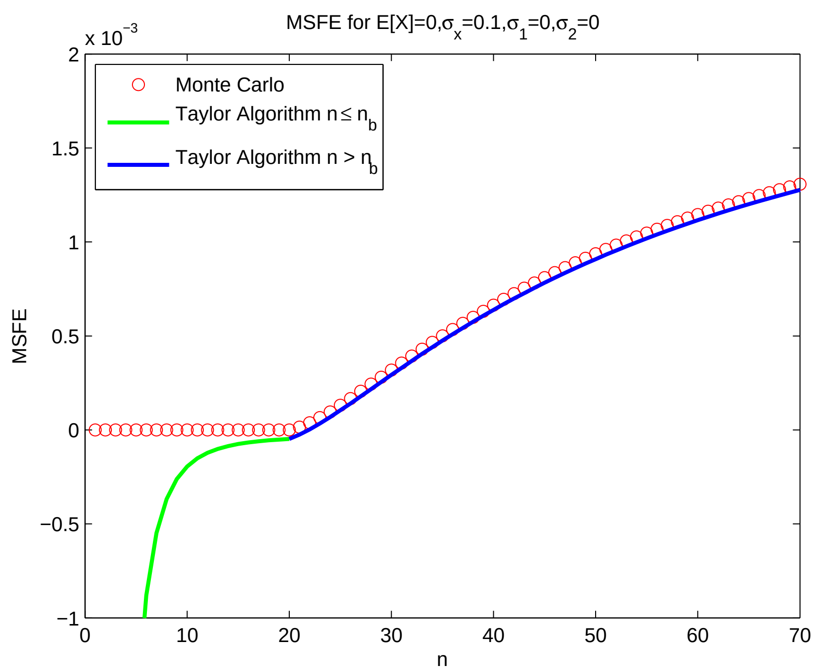


Figure 7.24: MSFE for $E[X] = 0$, $\sigma_x = 0.1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

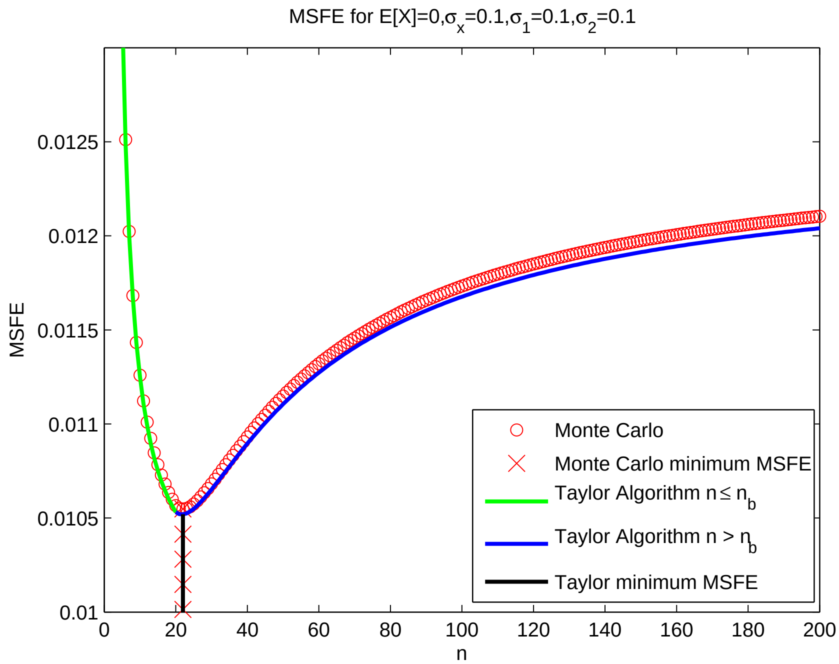


Figure 7.25: MSFE for $E[X] = 0$, $\sigma_x = 0.1$, $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 22, Taylor algorithm minimum MSFE = 22

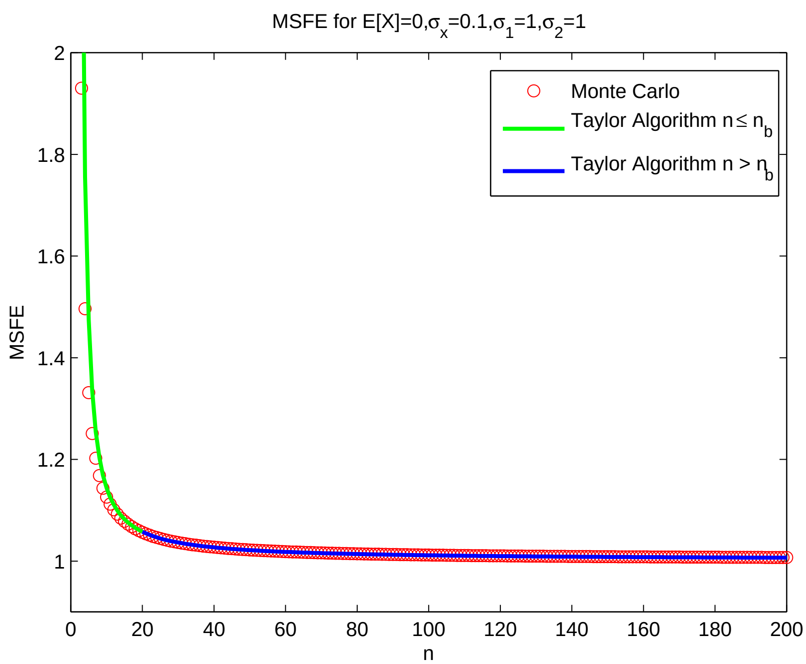


Figure 7.26: MSFE for $E[X] = 0$, $\sigma_x = 0.1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE > 500, Taylor algorithm minimum MSFE > 500

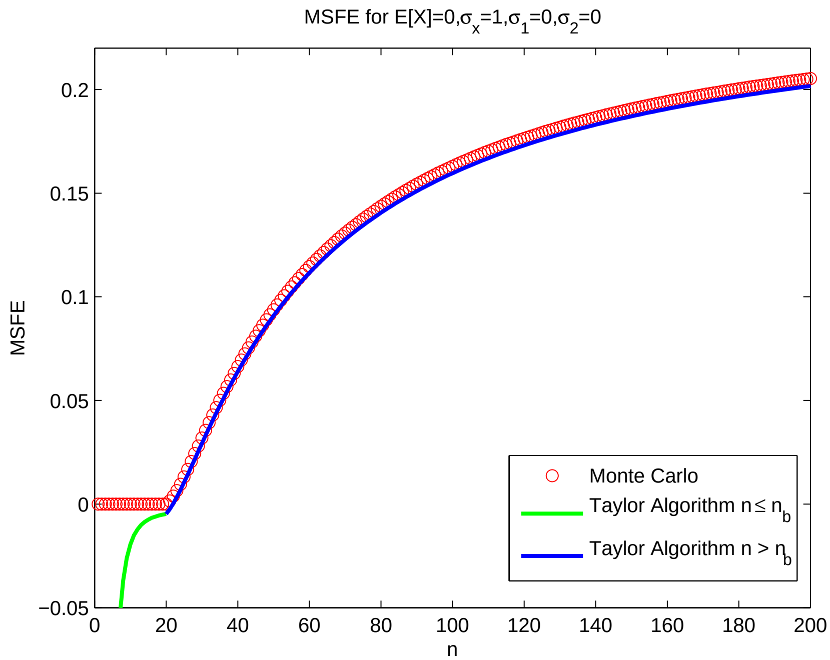


Figure 7.27: MSFE for $E[X] = 0$, $\sigma_x = 1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

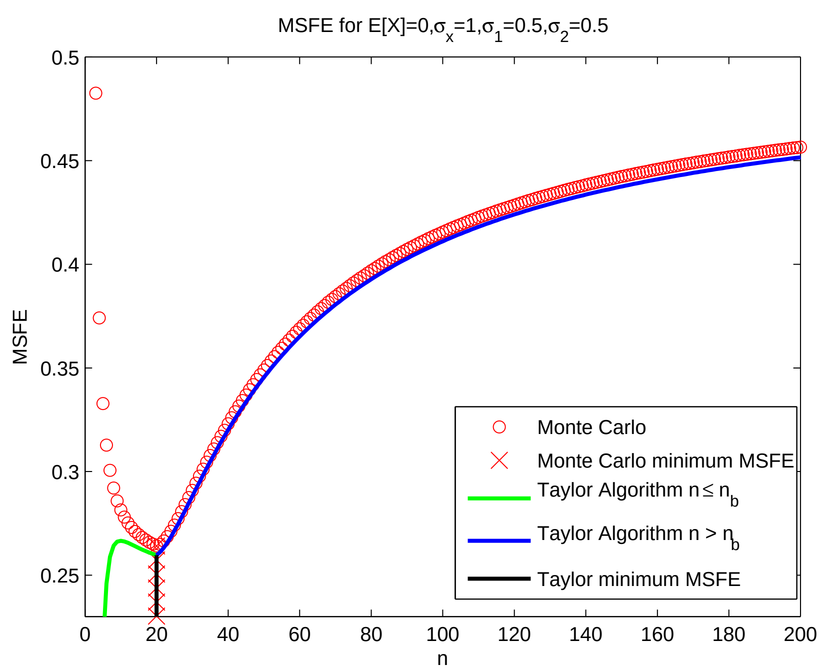


Figure 7.28: MSFE for $E[X] = 0$, $\sigma_x = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm local minimum MSFE = 20

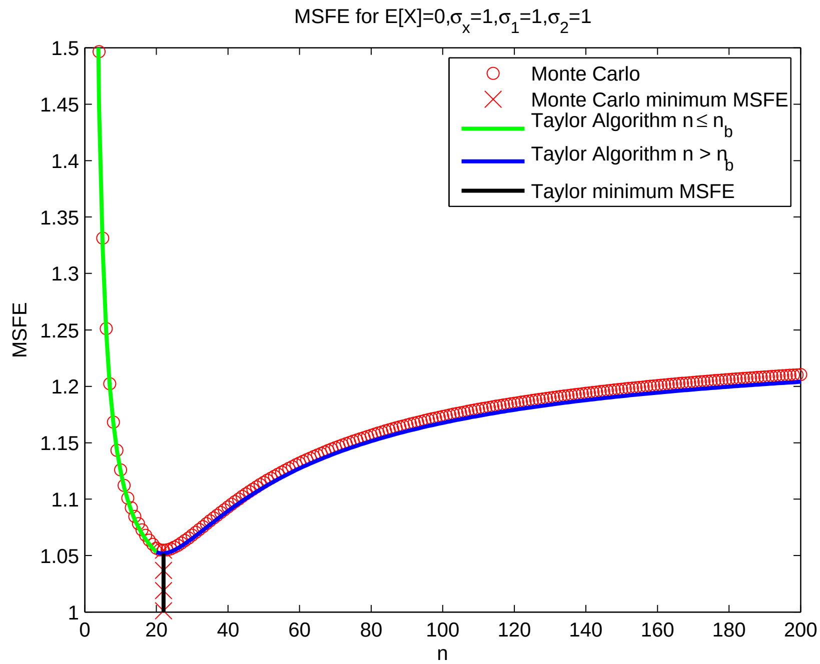


Figure 7.29: MSFE for $E[X] = 0$, $\sigma_x = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 22, Taylor algorithm minimum MSFE = 22

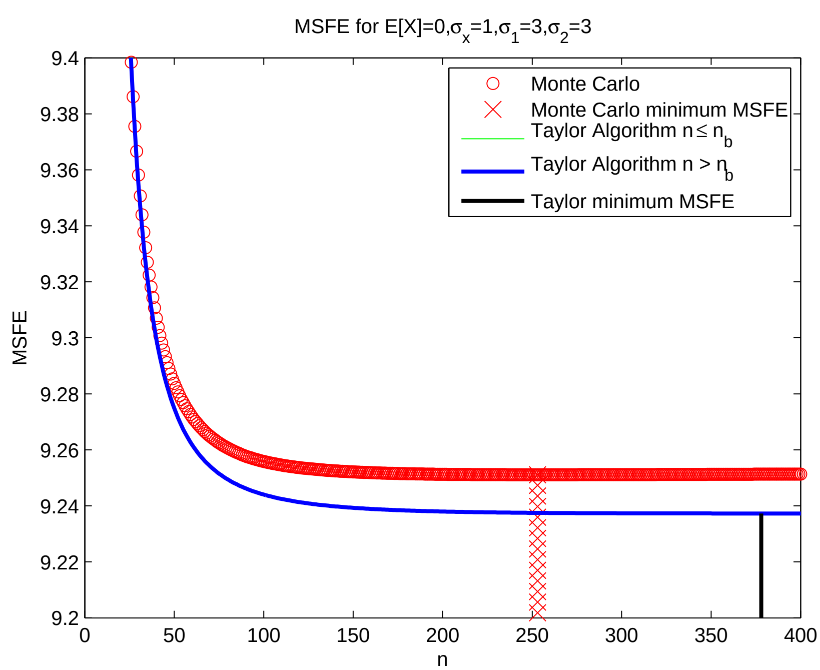


Figure 7.30: MSFE for $E[X] = 0$, $\sigma_x = 1$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 253, Taylor algorithm minimum MSFE = 378

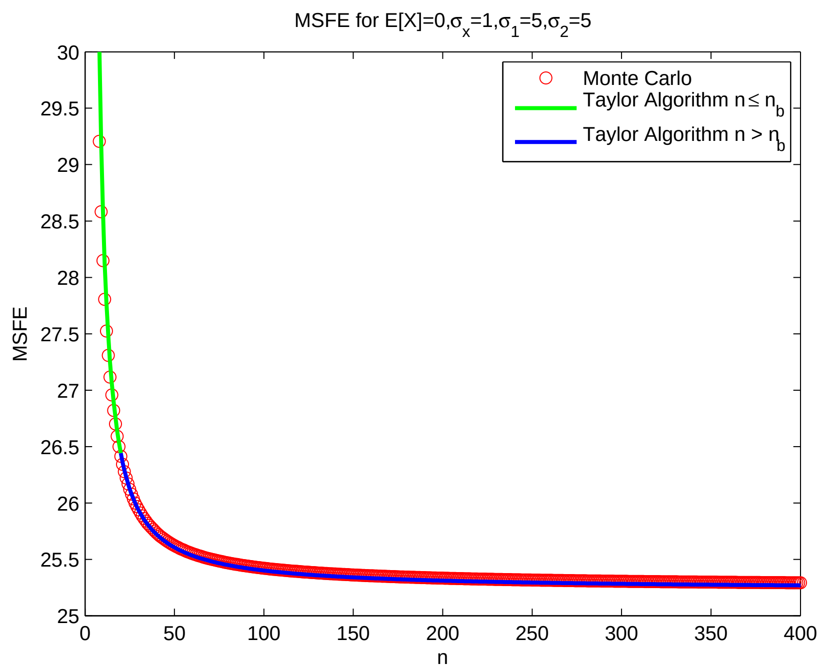


Figure 7.31: MSFE for $E[X] = 0$, $\sigma_x = 1$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE > 500, Taylor algorithm minimum MSFE > 500

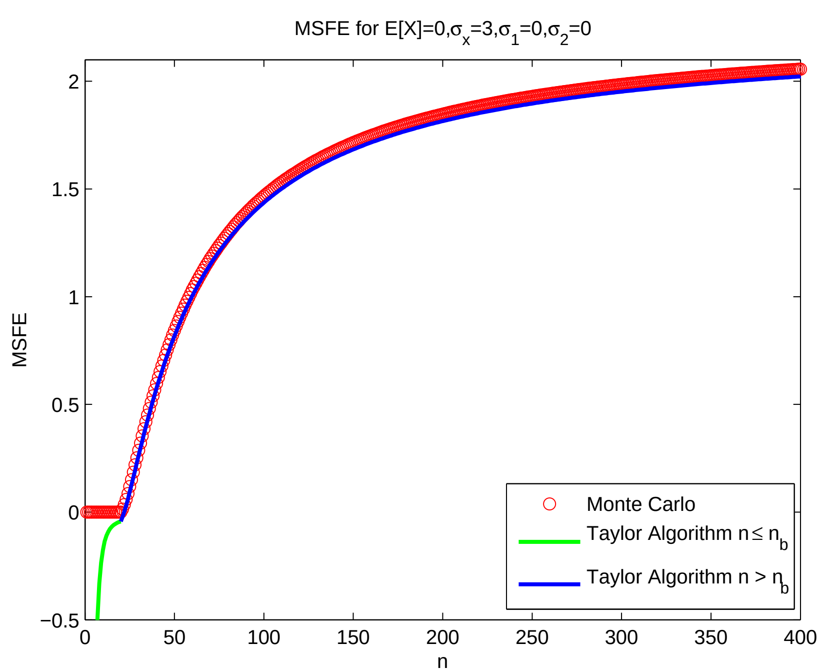


Figure 7.32: MSFE for $E[X] = 0$, $\sigma_x = 3$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

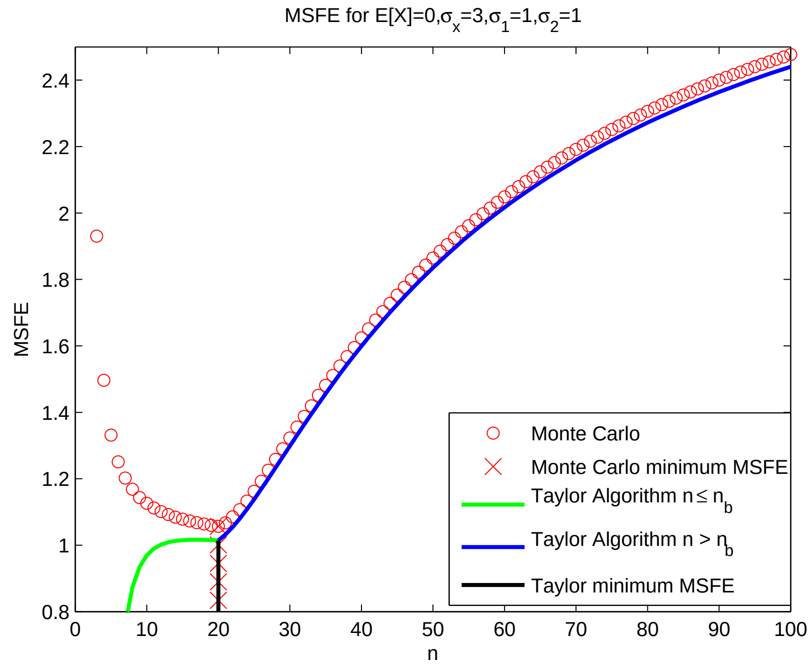


Figure 7.33: MSFE for $E[X] = 0$, $\sigma_x = 3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

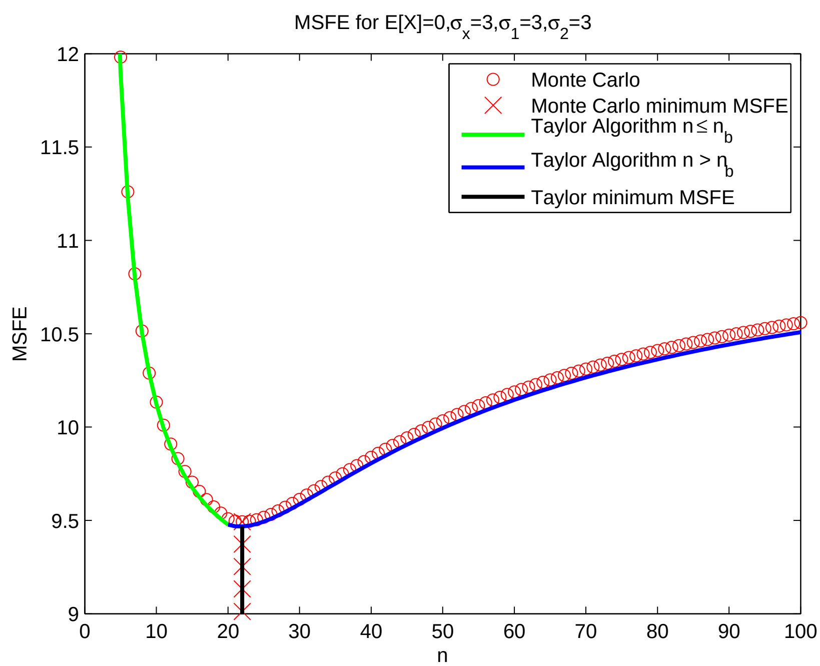


Figure 7.34: MSFE for $E[X] = 0$, $\sigma_x = 3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 22$, Taylor algorithm minimum MSFE $= 22$

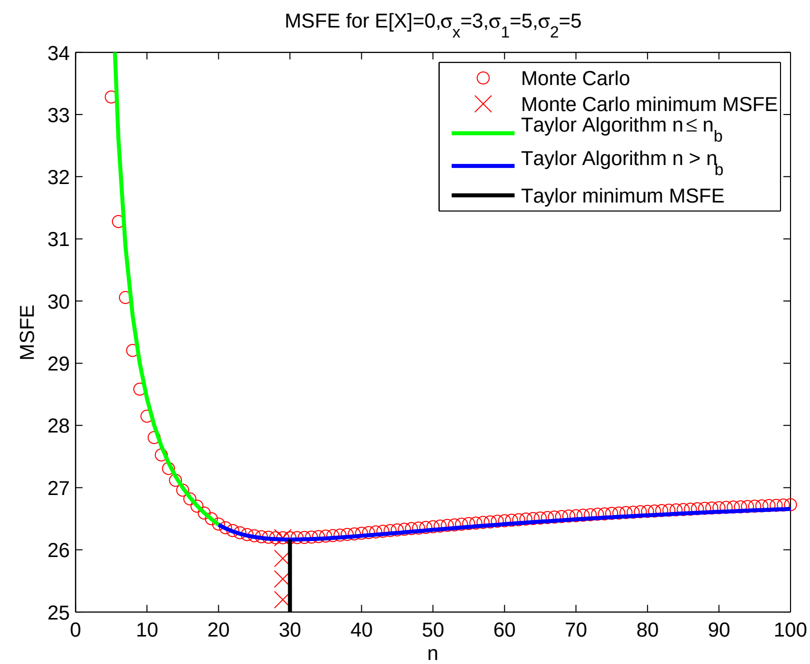


Figure 7.35: MSFE for $E[X] = 0$, $\sigma_x = 3$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 29, Taylor algorithm minimum MSFE = 30

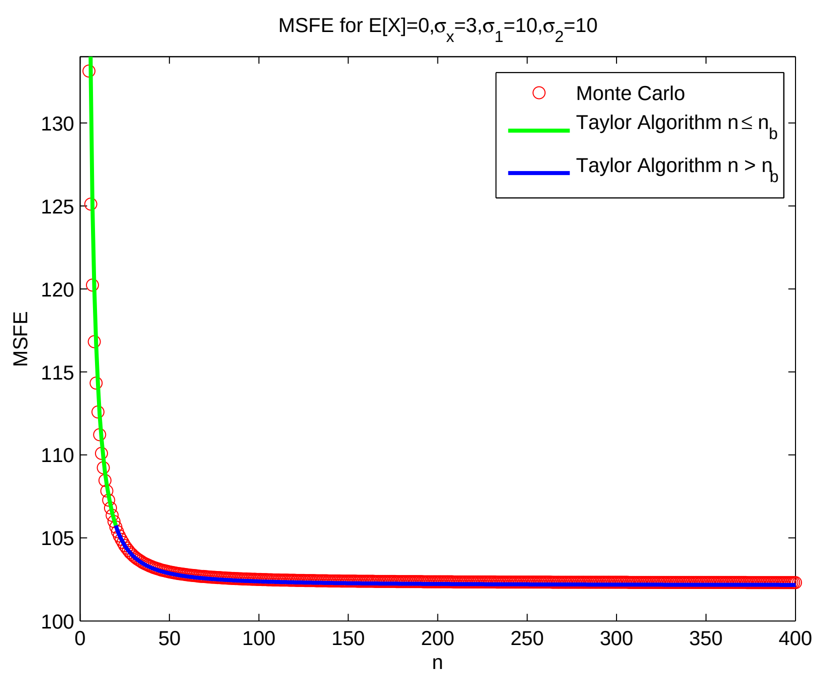


Figure 7.36: MSFE for $E[X] = 0$, $\sigma_x = 3$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE > 500, Taylor algorithm minimum MSFE > 500

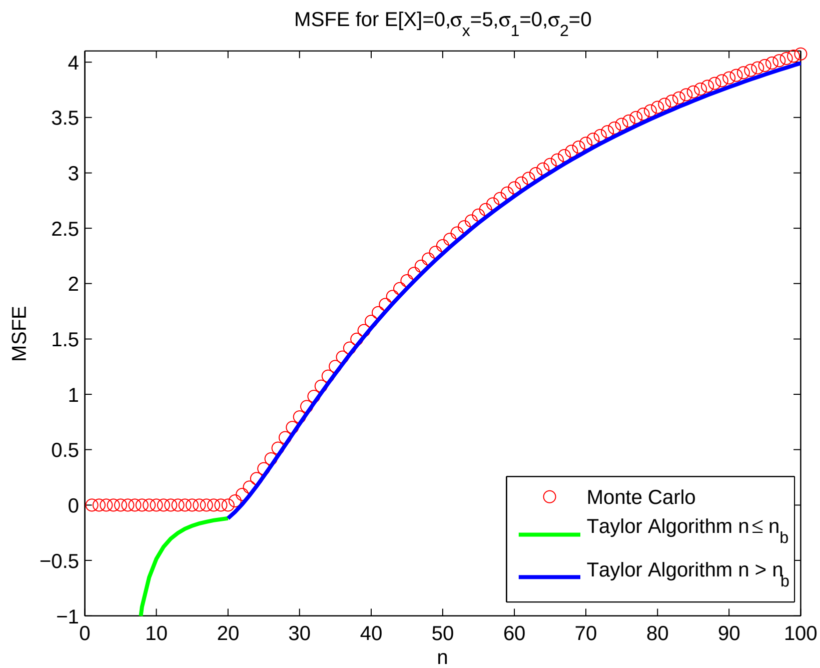


Figure 7.37: MSFE for $E[X] = 0$, $\sigma_x = 5$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

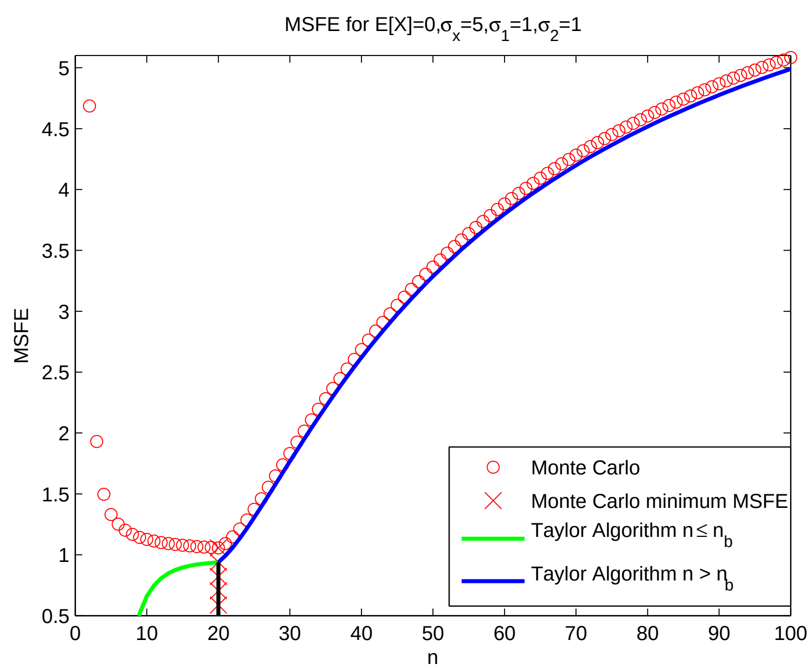


Figure 7.38: MSFE for $E[X] = 0$, $\sigma_x = 5$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

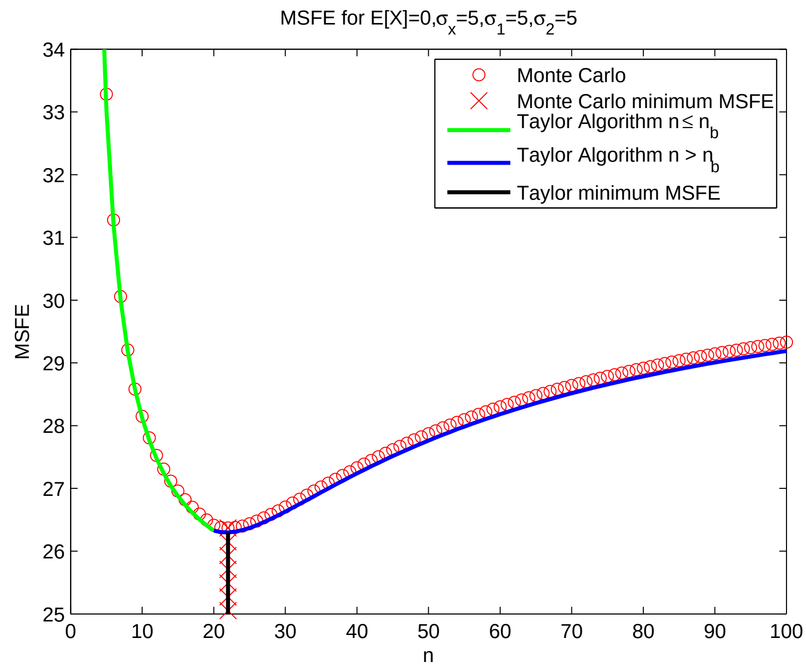


Figure 7.39: MSFE for $E[X] = 0$, $\sigma_x = 5$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 22, Taylor algorithm minimum MSFE = 22

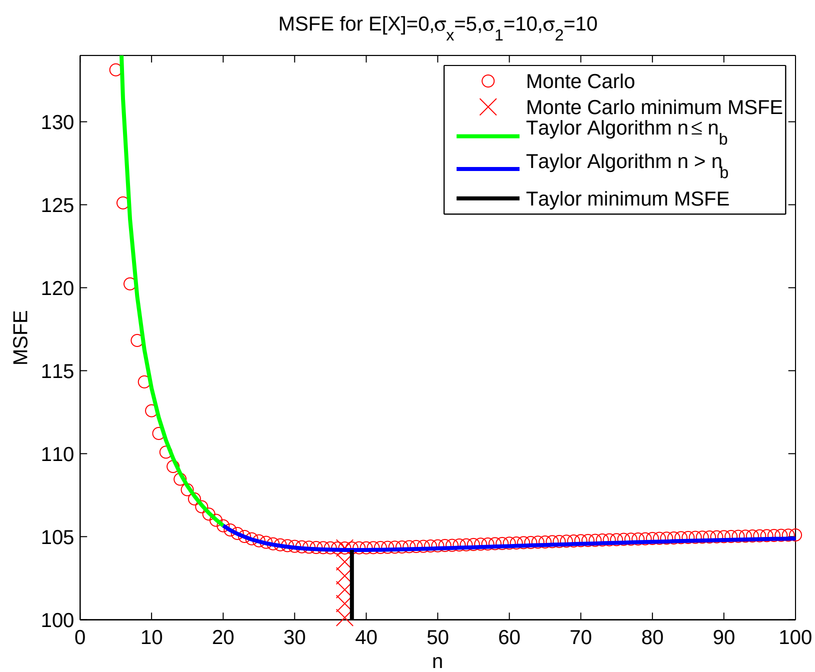


Figure 7.40: MSFE for $E[X] = 0$, $\sigma_x = 5$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 37, Taylor algorithm minimum MSFE = 38

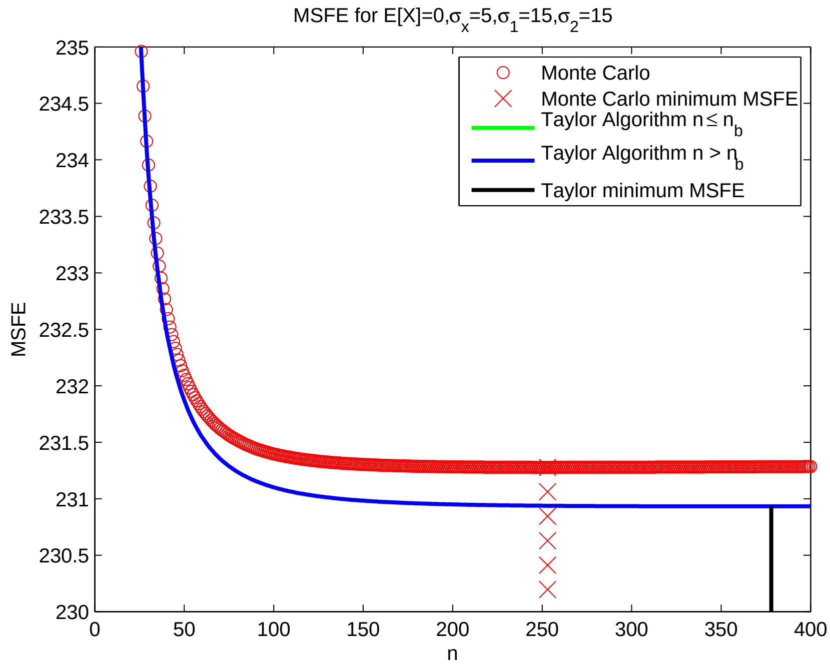


Figure 7.41: MSFE for $E[X] = 0$, $\sigma_x = 5$, $\sigma_1 = 15$, $\sigma_2 = 15$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 253, Taylor algorithm minimum MSFE = 378

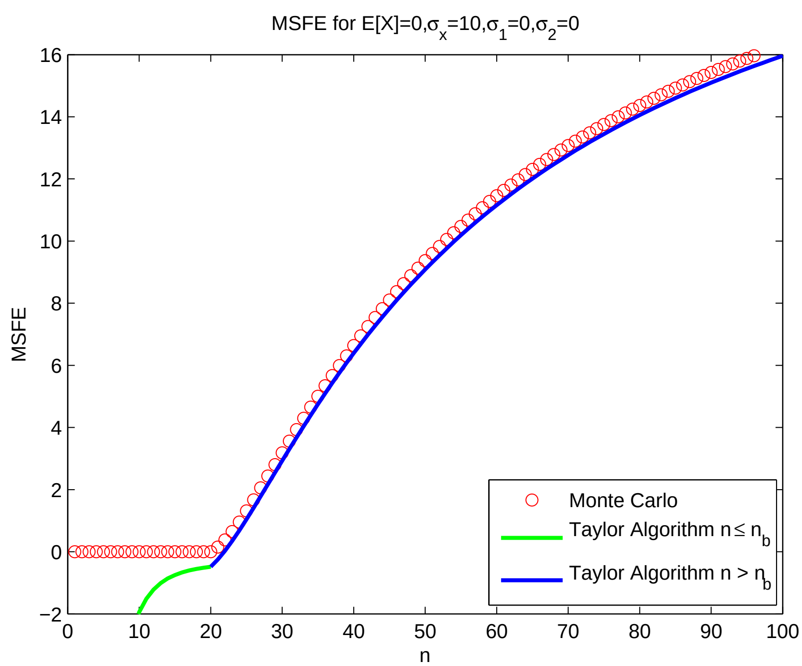


Figure 7.42: MSFE for $E[X] = 0$, $\sigma_x = 10$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$

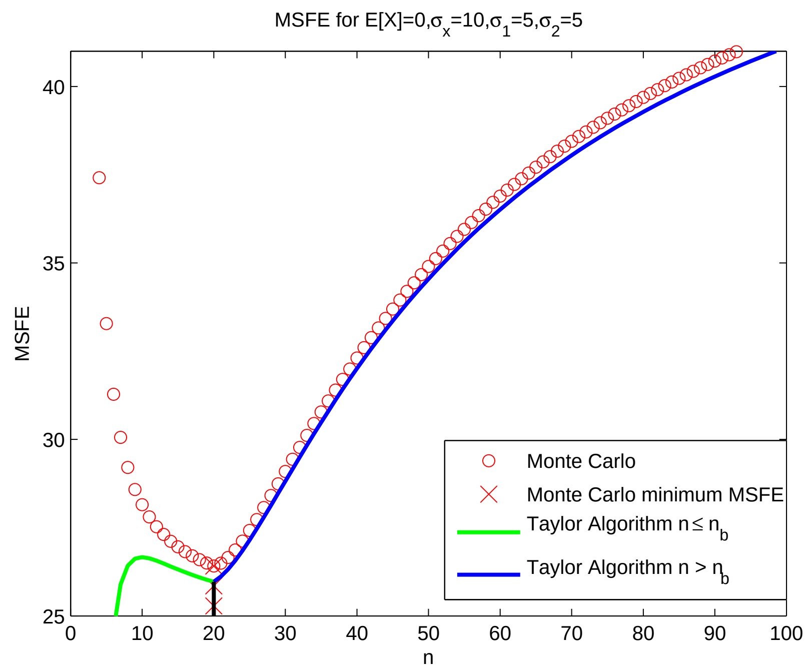


Figure 7.43: MSFE for $E[X] = 0$, $\sigma_x = 10$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 20, Taylor algorithm local minimum MSFE = 20

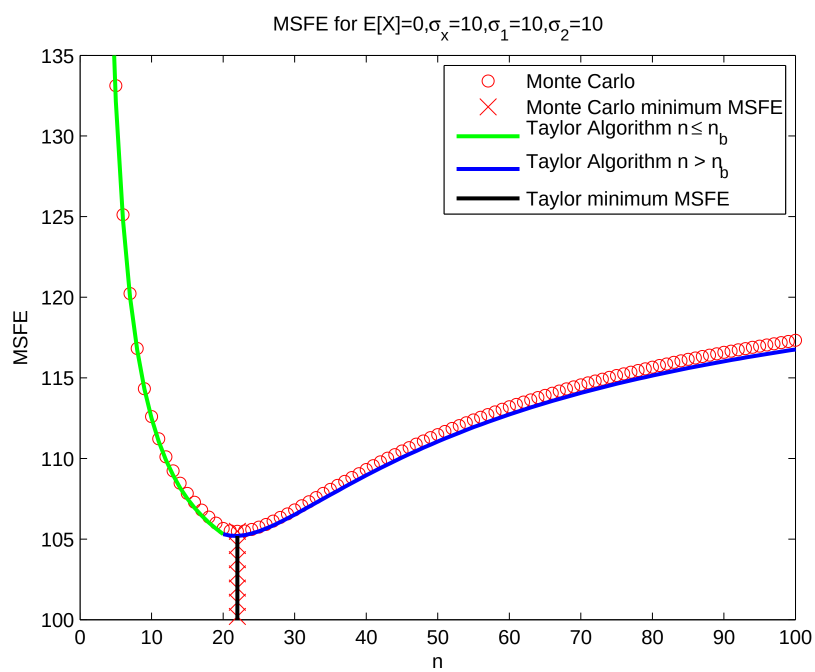


Figure 7.44: MSFE for $E[X] = 0$, $\sigma_x = 10$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE = 22, Taylor algorithm minimum MSFE = 22

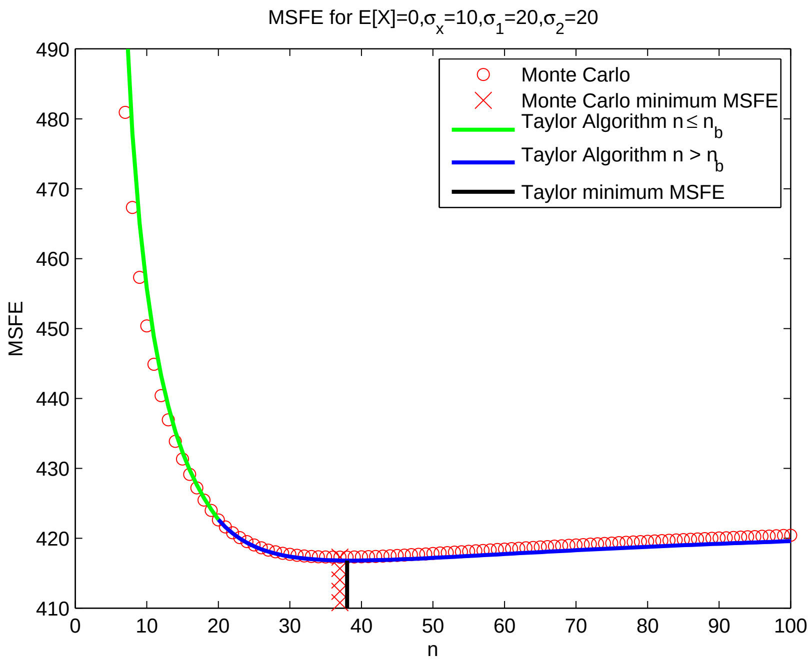


Figure 7.45: MSFE for $E[X] = 0$, $\sigma_x = 10$, $\sigma_1 = 20$, $\sigma_2 = 20$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE $= 37$, Taylor algorithm minimum MSFE $= 38$

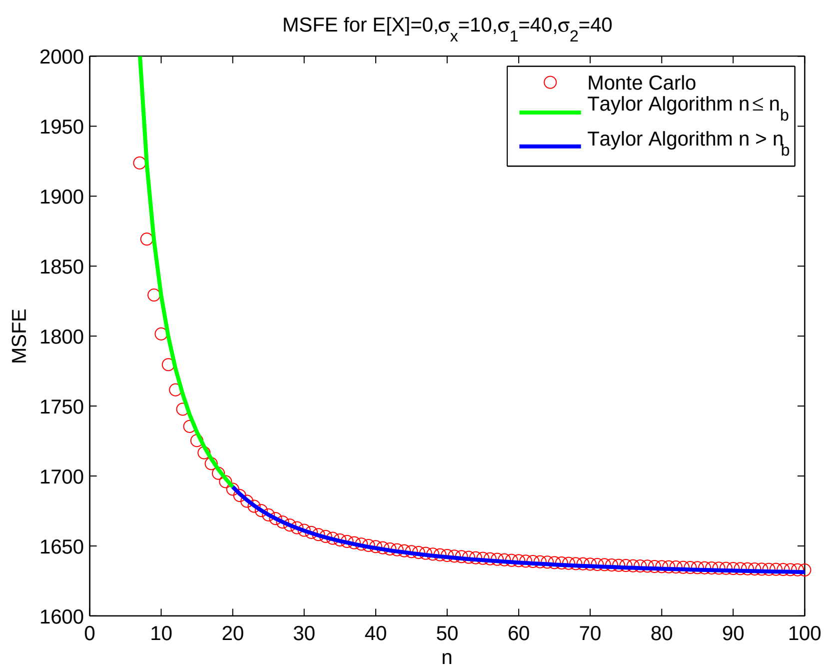


Figure 7.46: MSFE for $E[X] = 0$, $\sigma_x = 10$, $\sigma_1 = 40$, $\sigma_2 = 40$, $\theta_1 = 2$, $\theta_2 = 2.5$, $n_b = 20$, Monte Carlo minimum MSFE > 500, Taylor algorithm minimum MSFE > 500

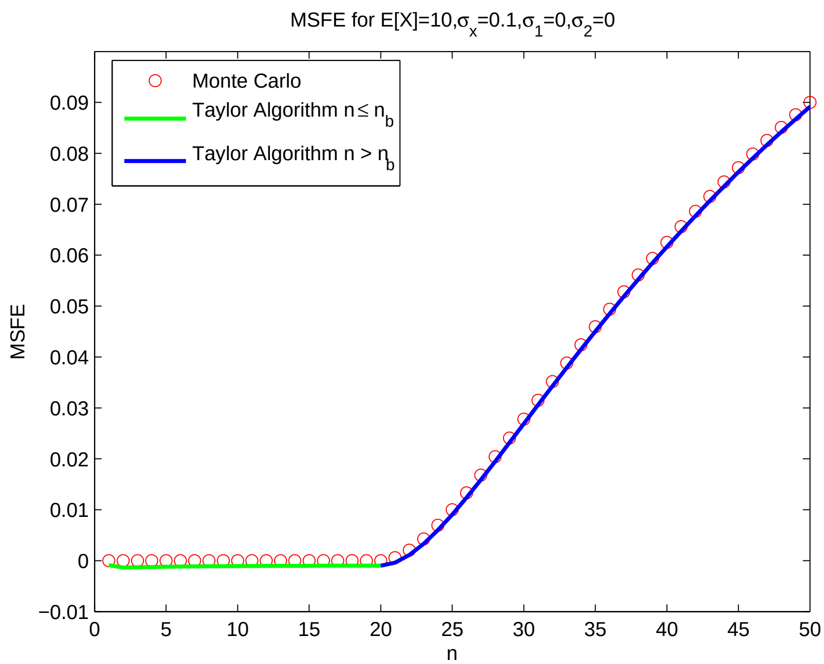


Figure 7.47: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.05$

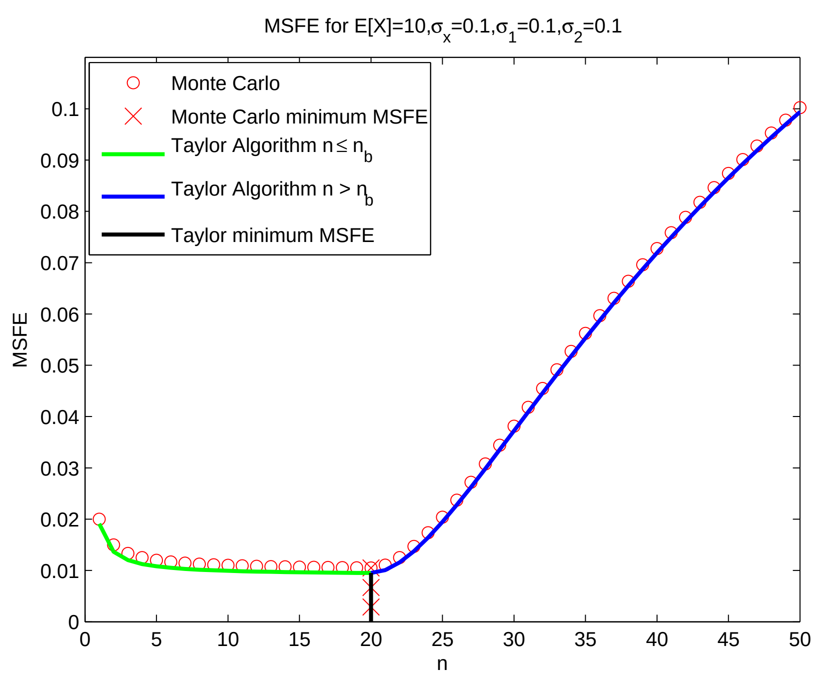


Figure 7.48: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE $= 20$,Taylor algorithm minimum MSFE $= 20$

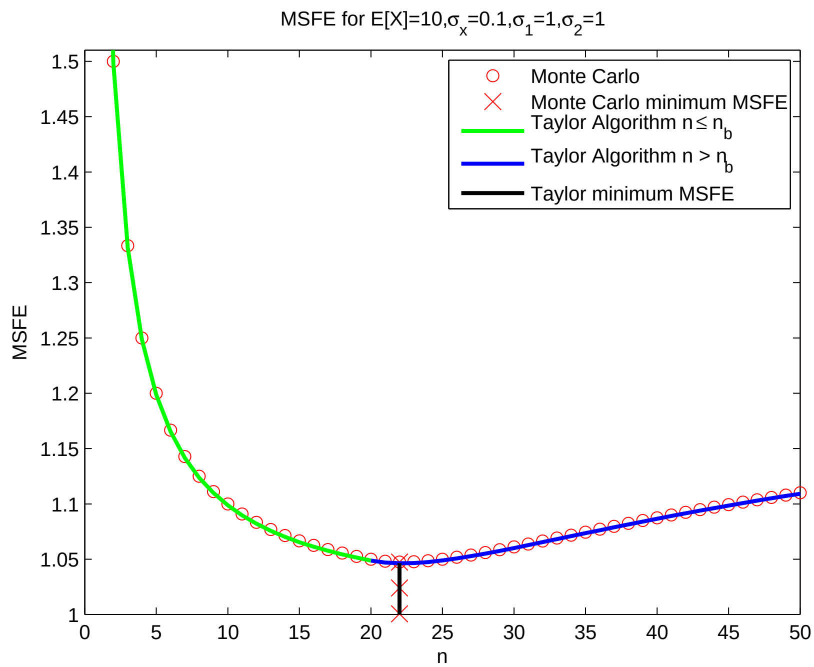


Figure 7.49: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE = 22,Taylor algorithm minimum MSFE = 22

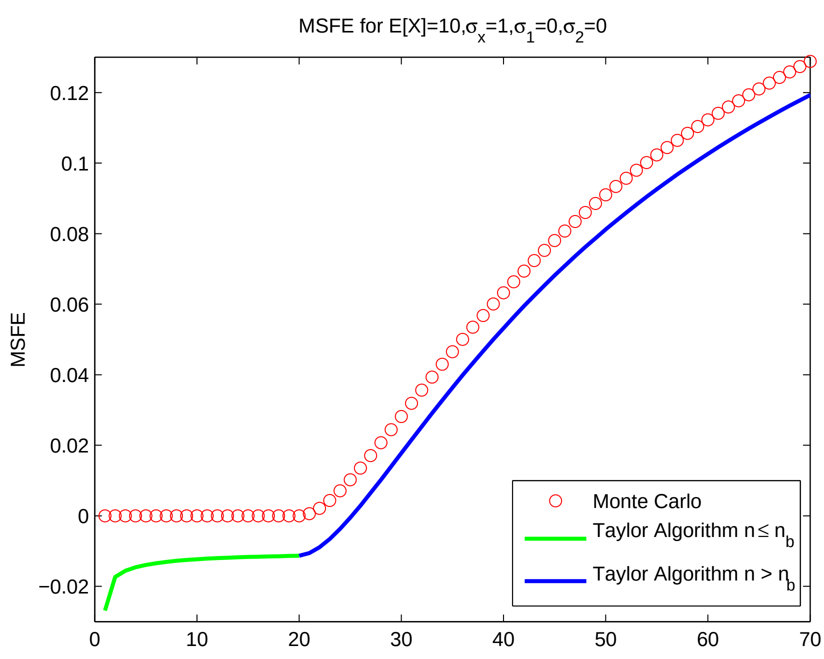


Figure 7.50: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.05$

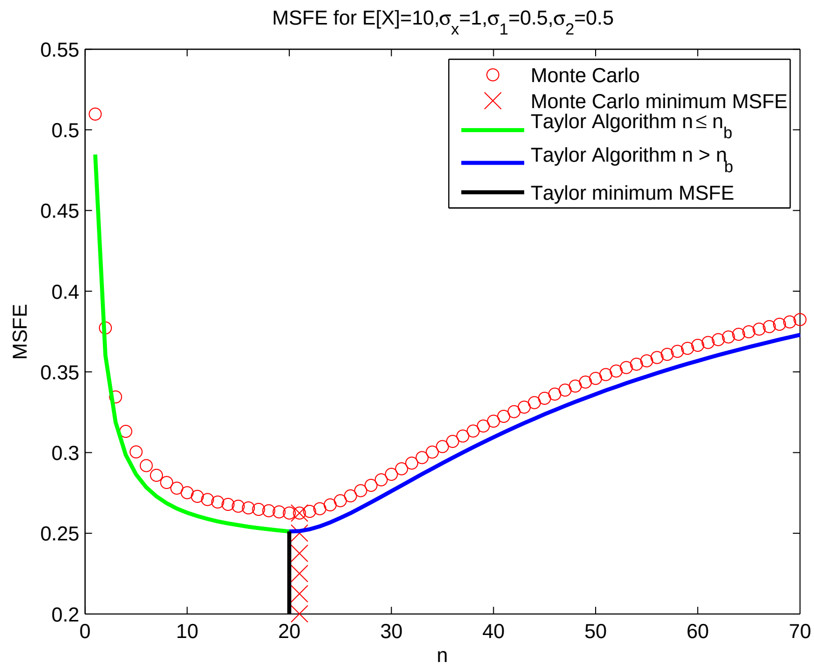


Figure 7.51: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE = 21,Taylor algorithm minimum MSFE = 20

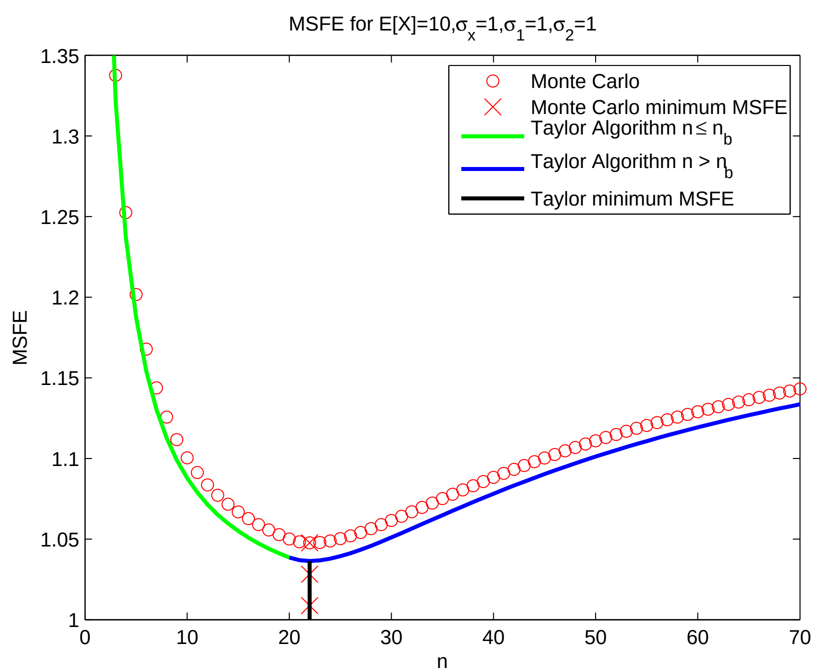


Figure 7.52: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE = 22,Taylor algorithm minimum MSFE = 22

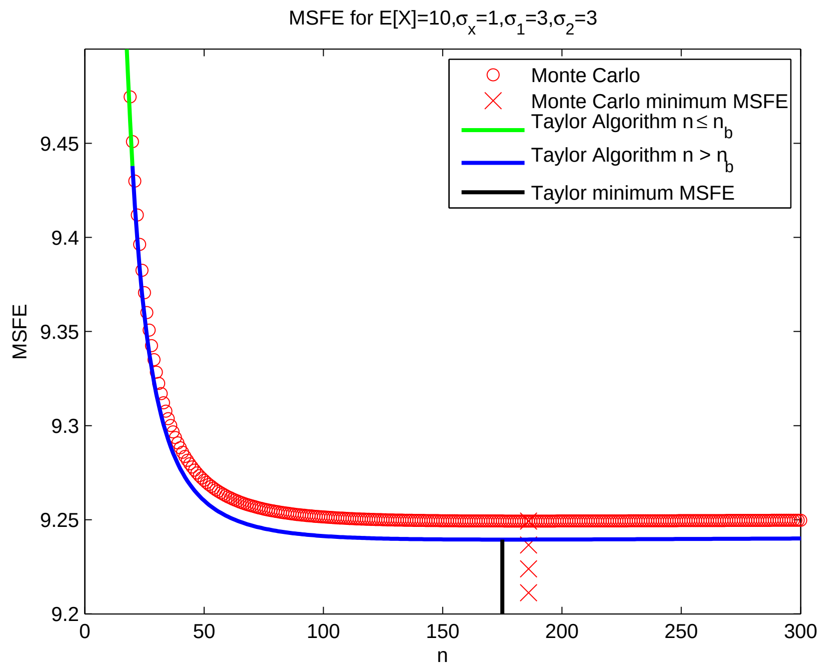


Figure 7.53: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE = 186,Taylor algorithm minimum MSFE = 175

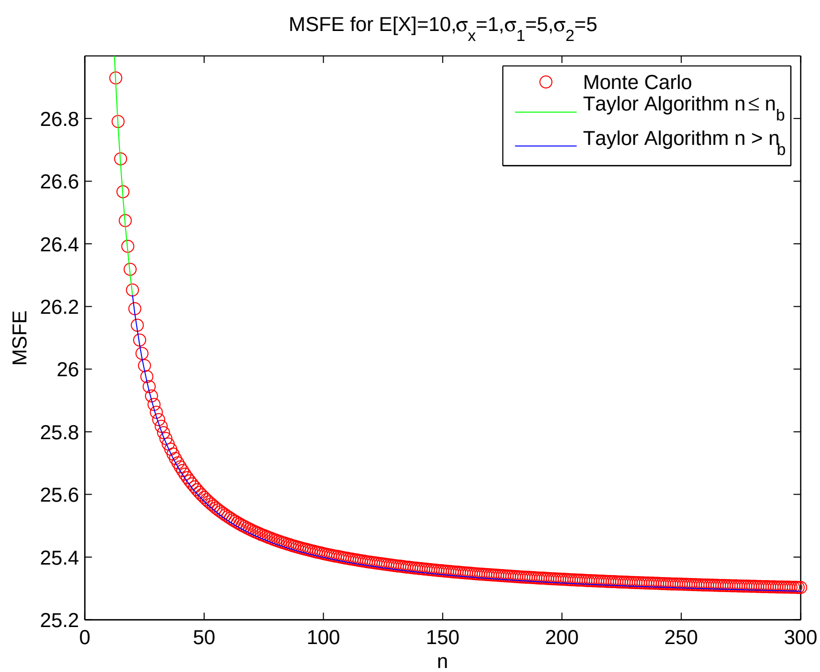


Figure 7.54: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE > 500,Taylor algorithm minimum MSFE > 500

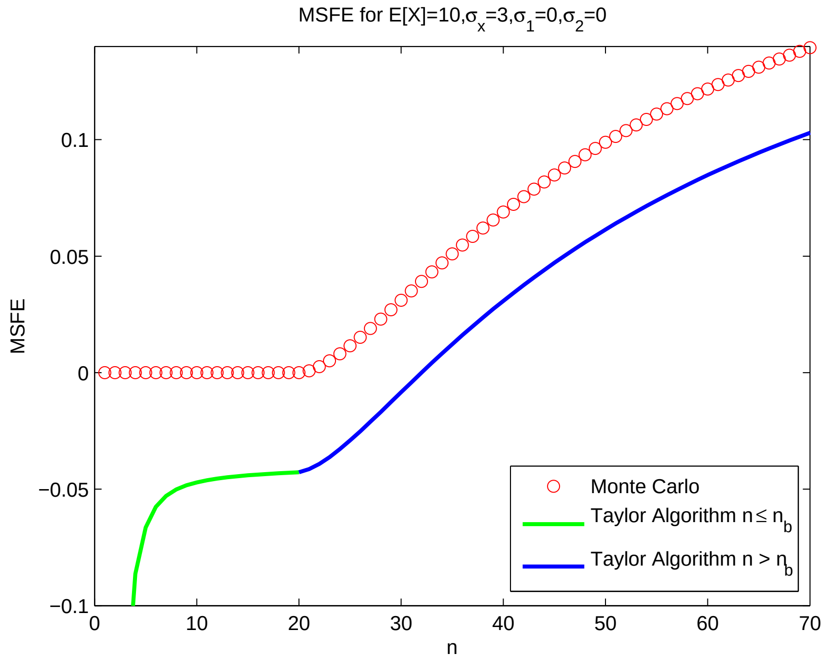


Figure 7.55: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.05$

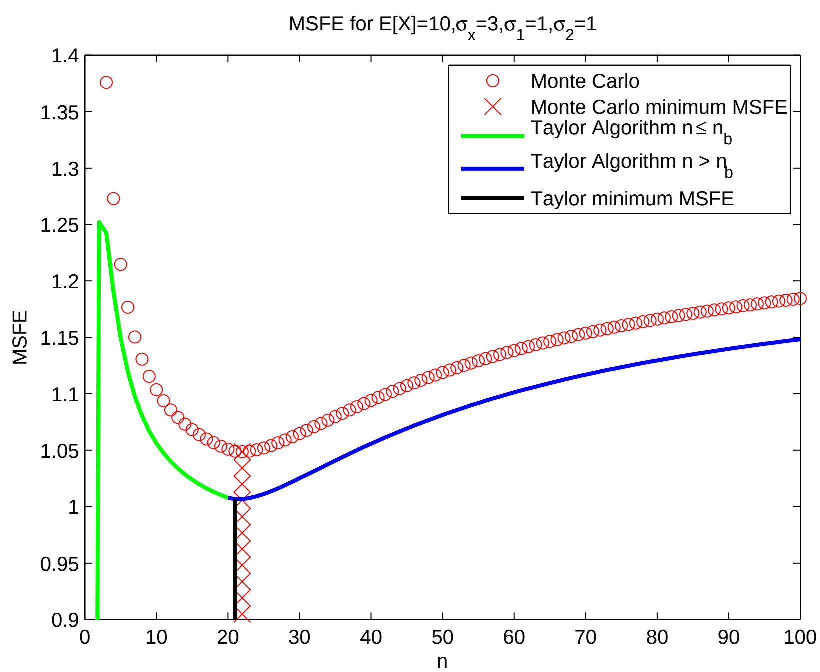


Figure 7.56: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE $= 22$,Taylor algorithm minimum MSFE $= 21$

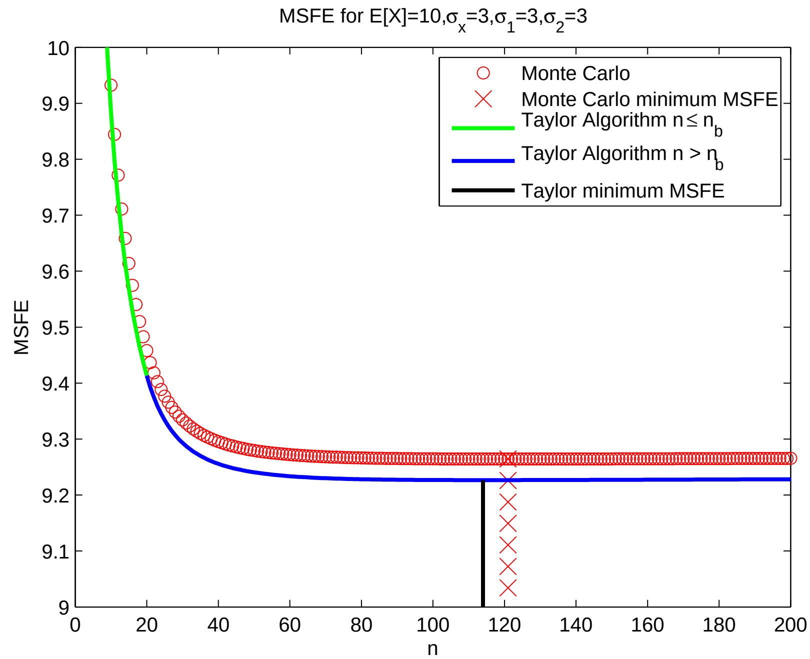


Figure 7.57: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE = 121,Taylor algorithm minimum MSFE = 114

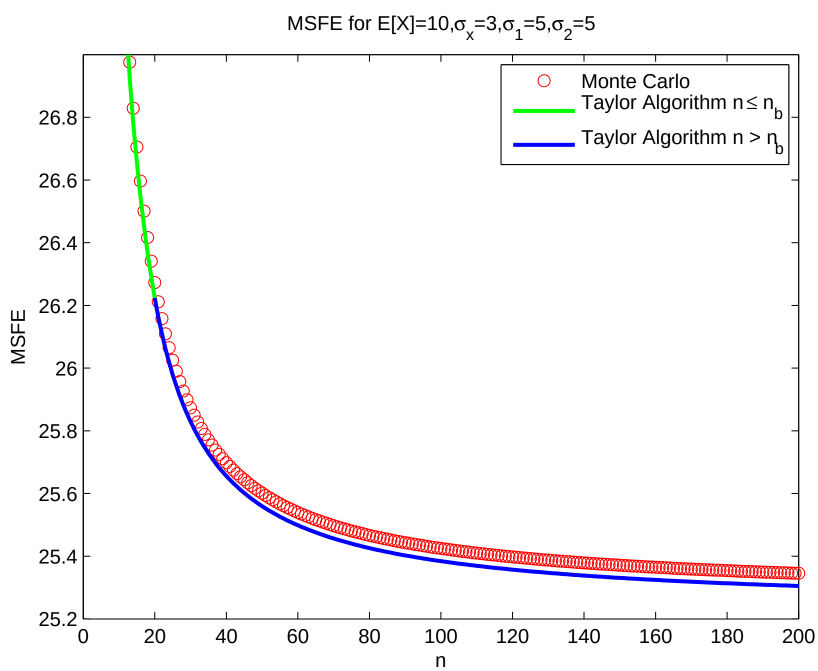


Figure 7.58: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE > 500,Taylor algorithm minimum MSFE > 500

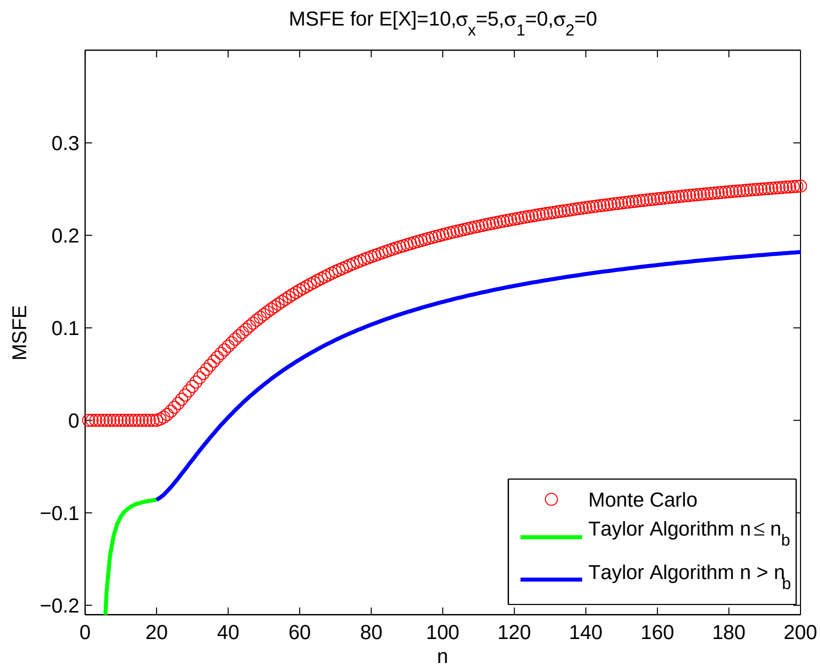


Figure 7.59: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.05$

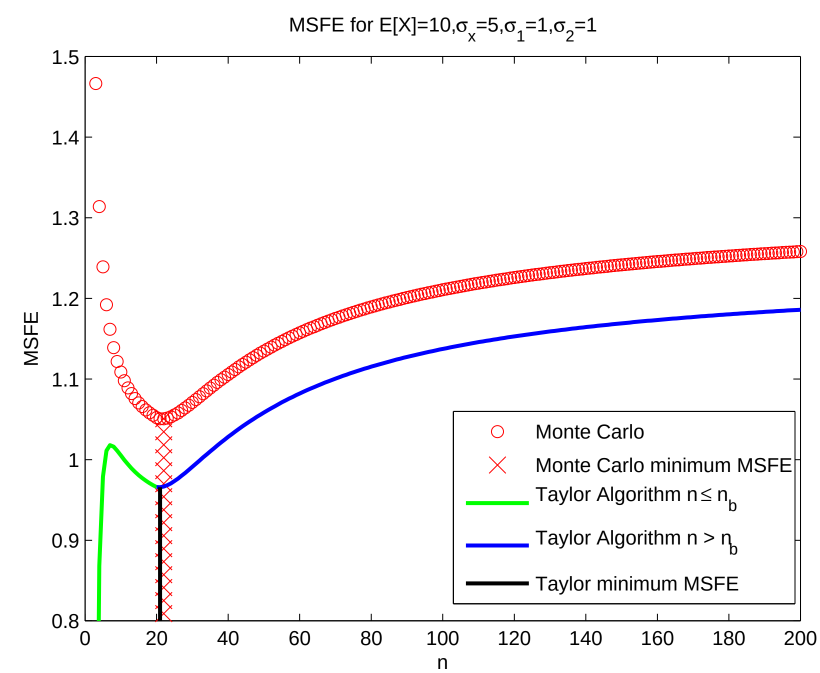


Figure 7.60: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE $= 22$,Taylor algorithm minimum MSFE $= 21$

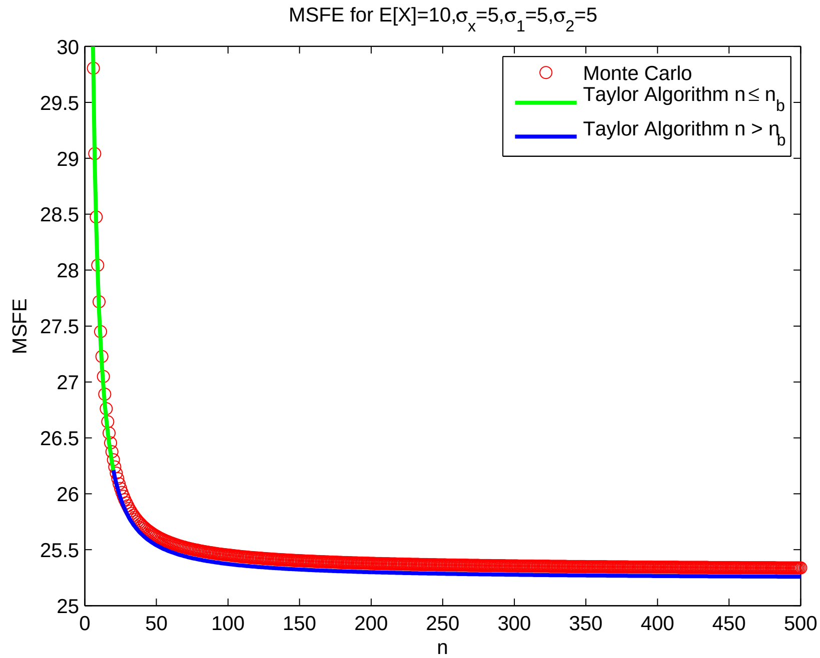


Figure 7.61: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE > 500,Taylor algorithm minimum MSFE > 500

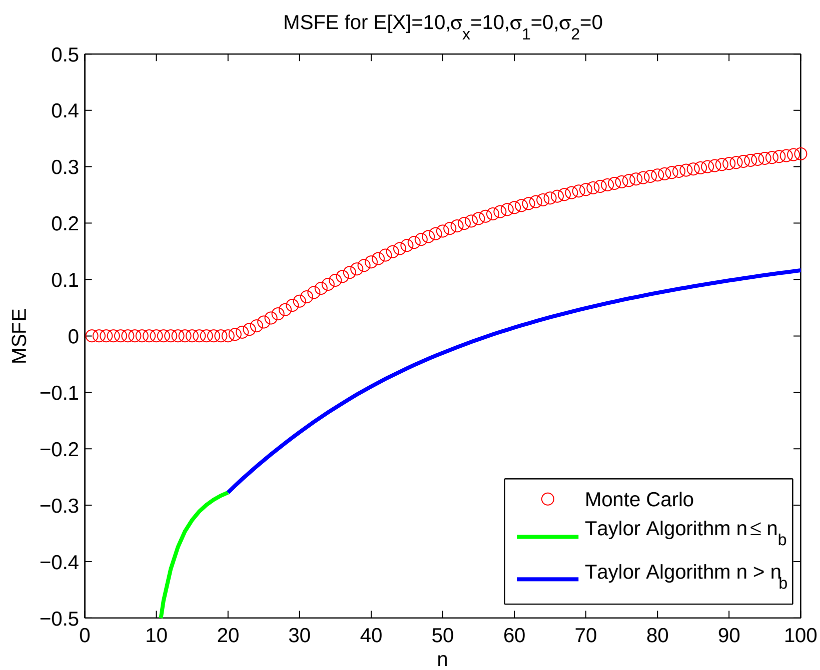


Figure 7.62: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.05$

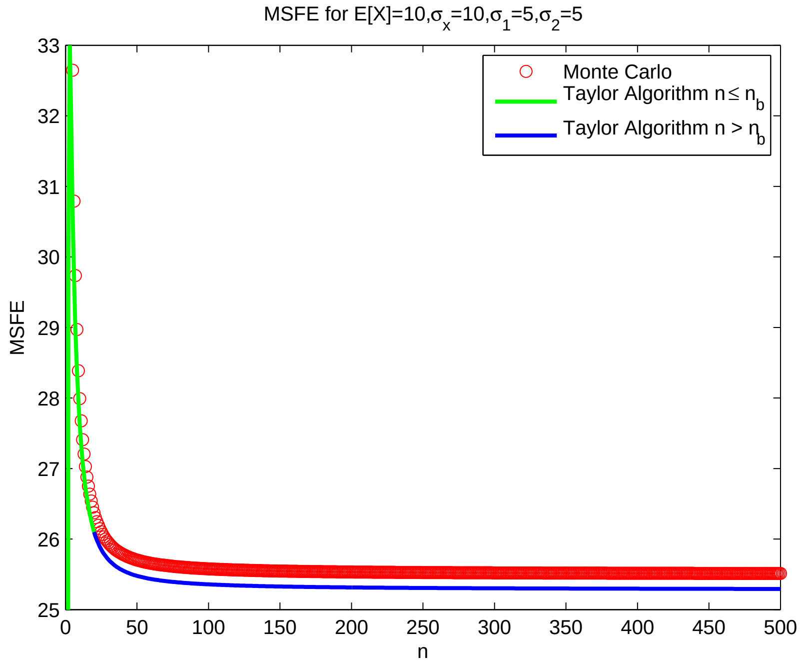


Figure 7.63: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.05$, $n_b = 20$, Monte Carlo minimum MSFE > 500,Taylor algorithm minimum MSFE > 500

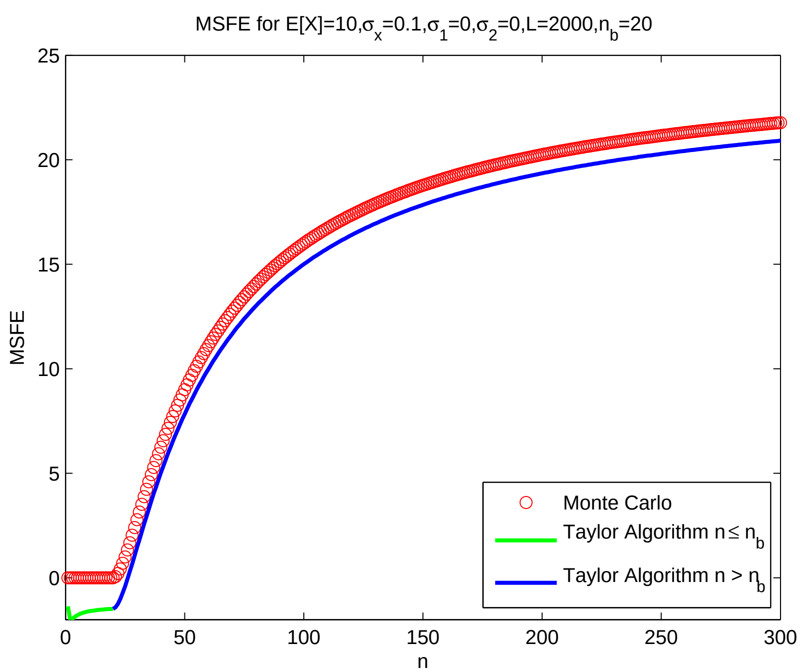


Figure 7.64: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

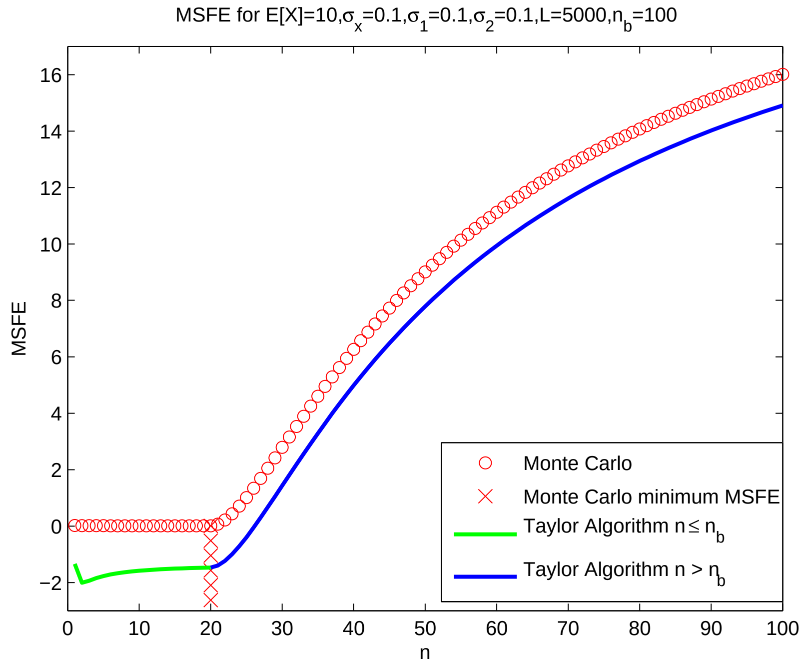


Figure 7.65: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\theta_1 = 2$, $\theta_2 = 2.5$

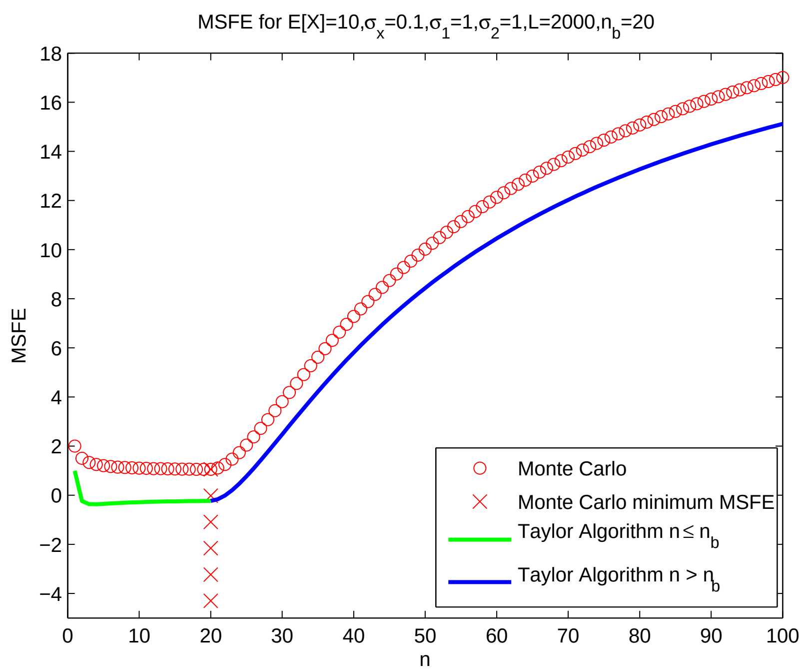


Figure 7.66: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$

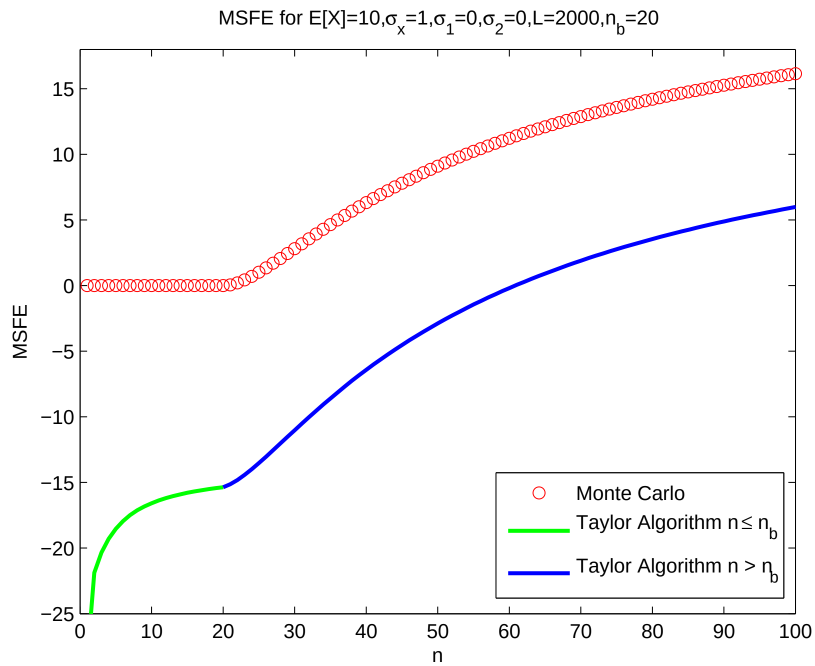


Figure 7.67: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

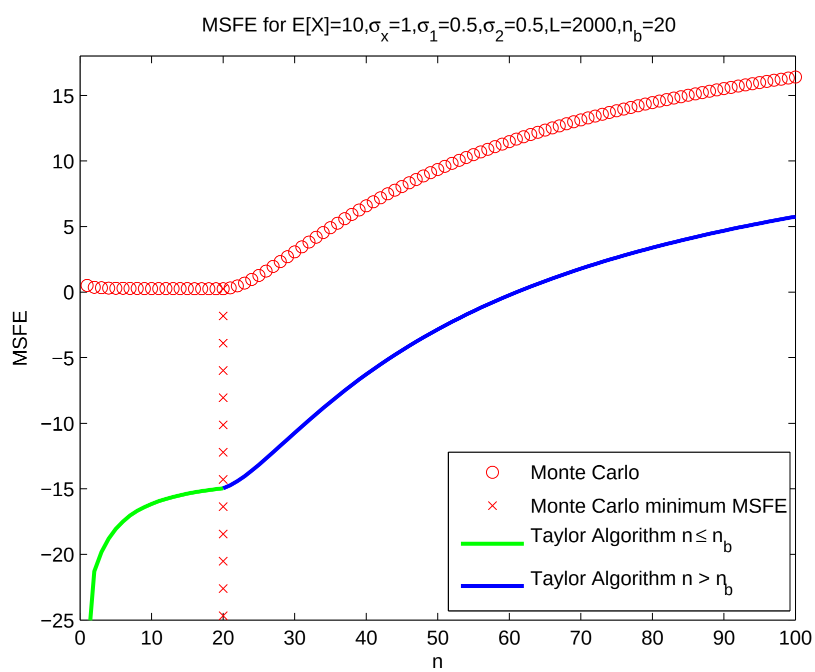


Figure 7.68: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\theta_1 = 2$, $\theta_2 = 2.5$

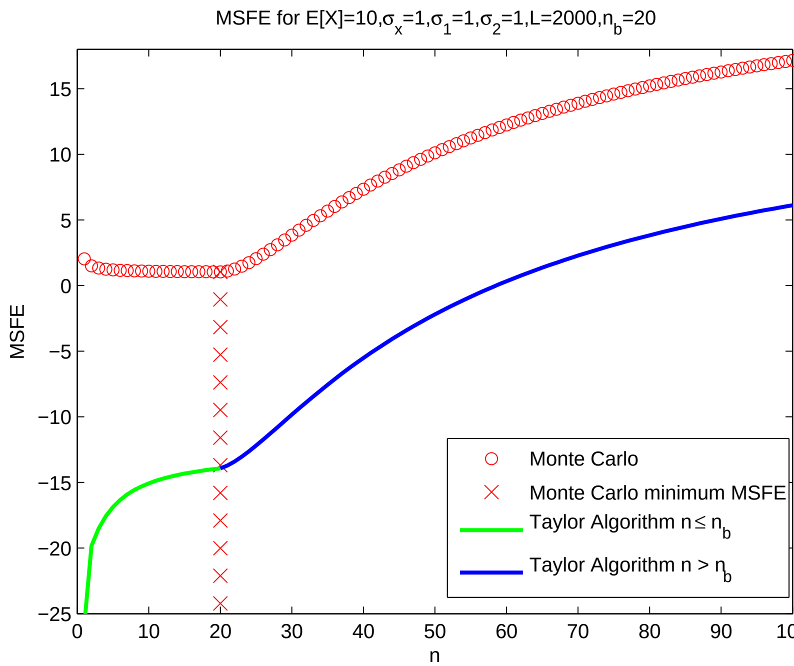


Figure 7.69: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$

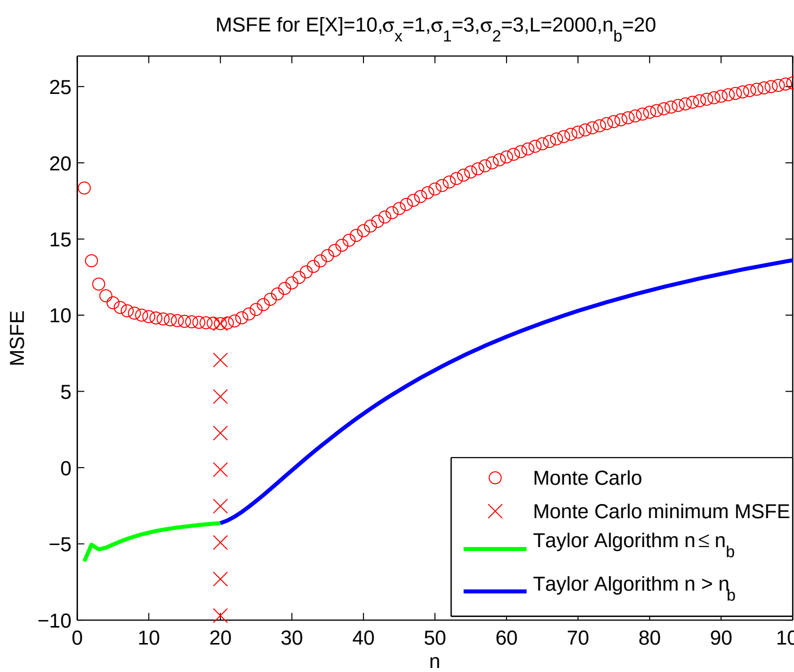


Figure 7.70: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$

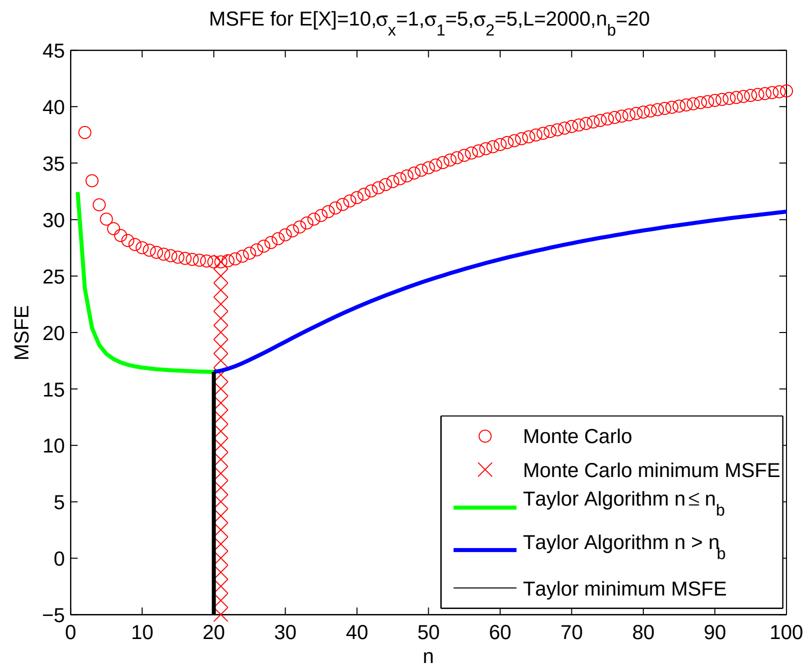


Figure 7.71: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 21,Taylor algorithm minimum MSFE = 20

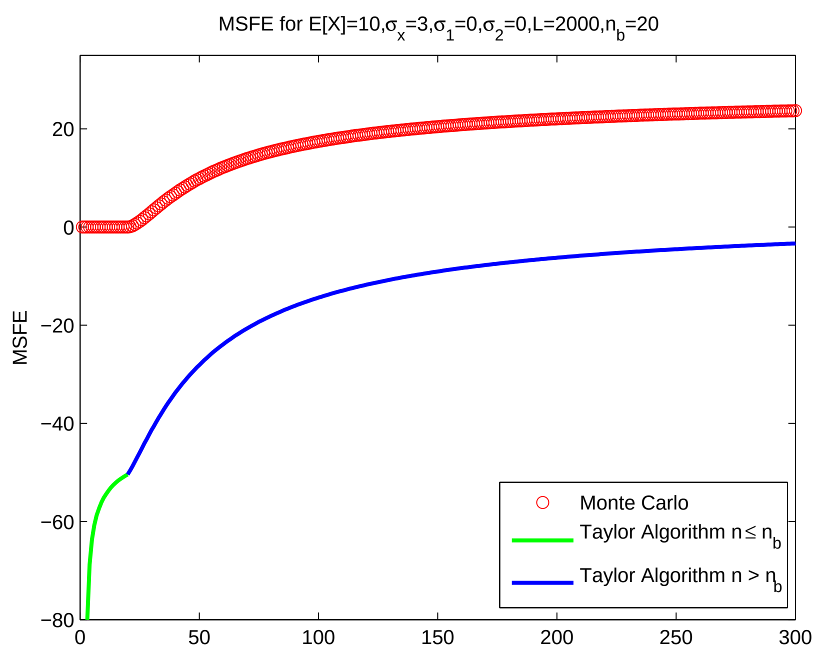


Figure 7.72: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

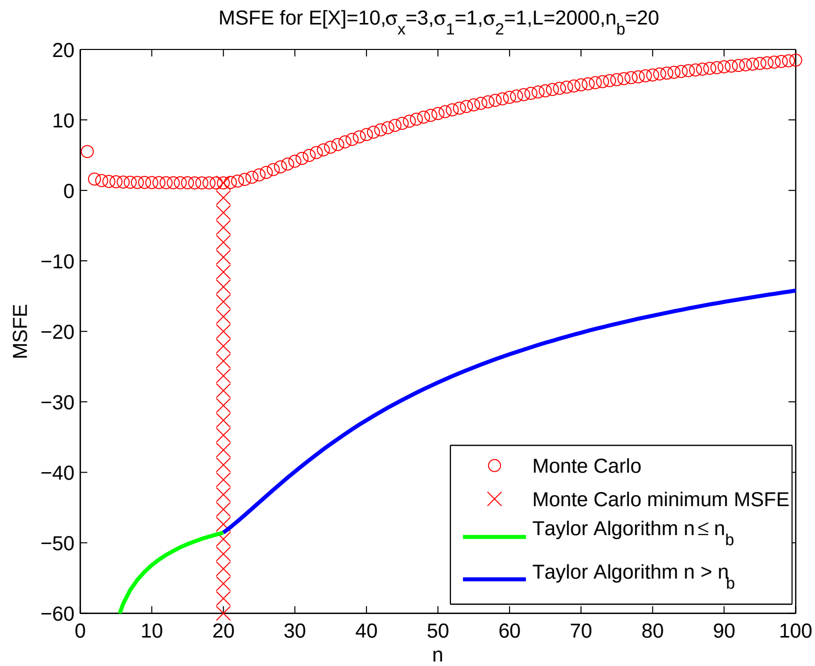


Figure 7.73: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$

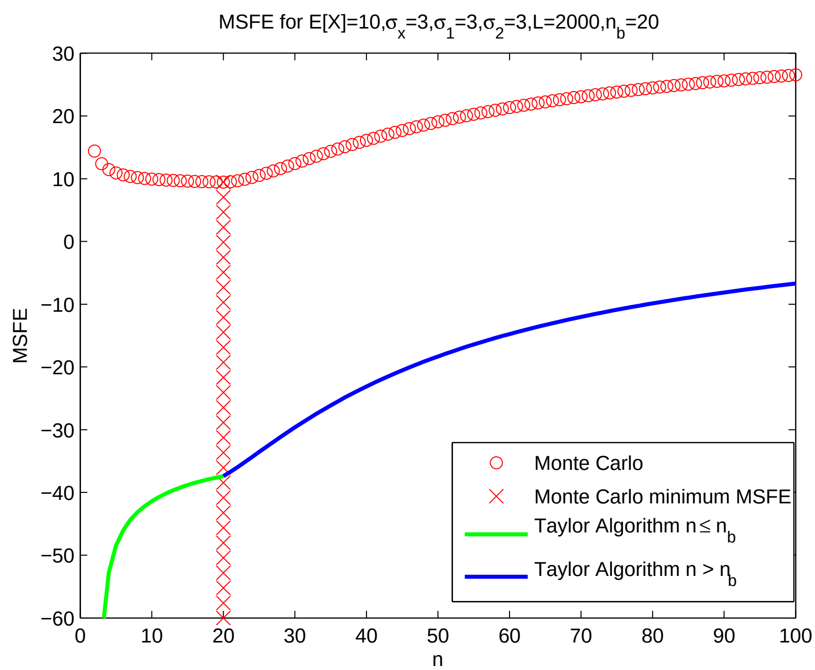


Figure 7.74: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$

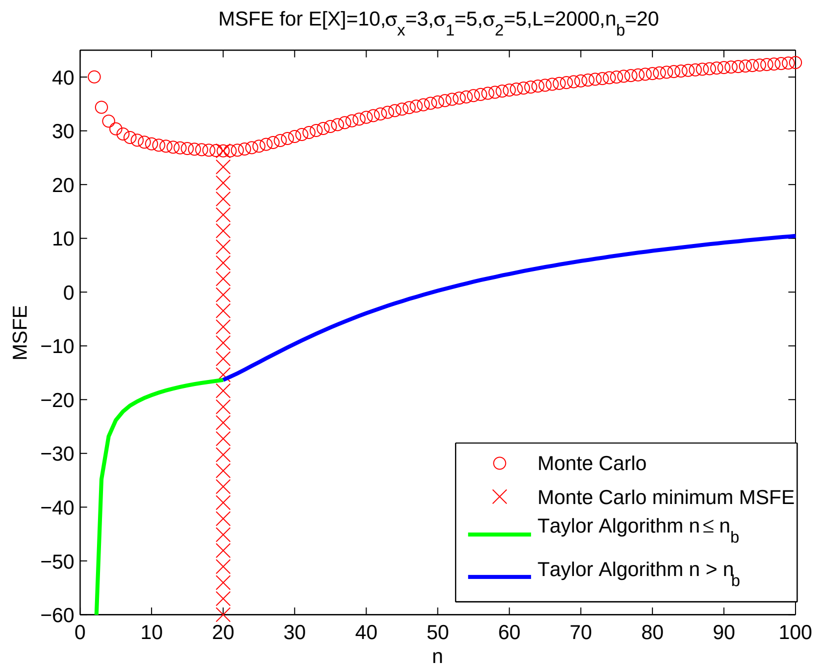


Figure 7.75: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$

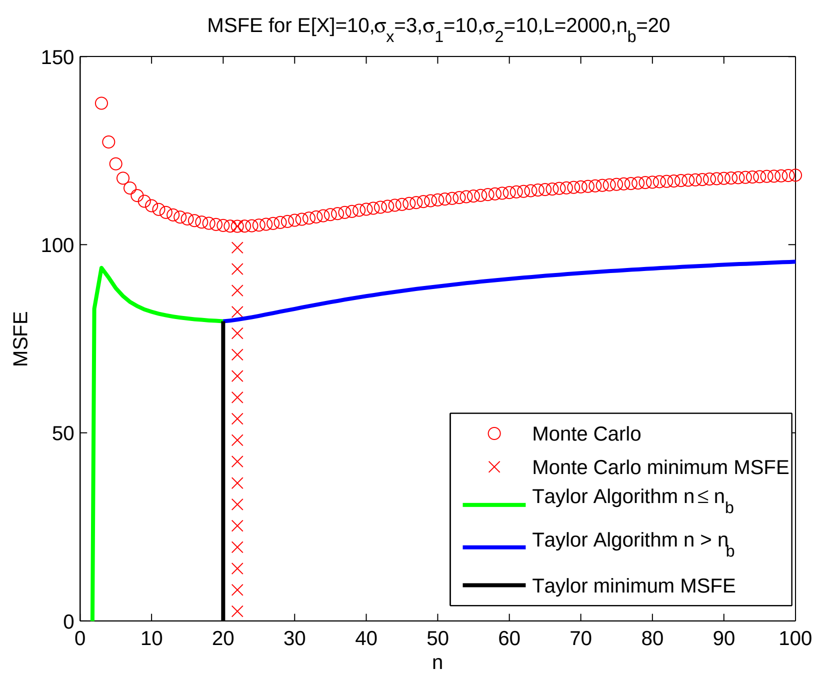


Figure 7.76: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 22$,Taylor algorithm minimum MSFE $= 20$

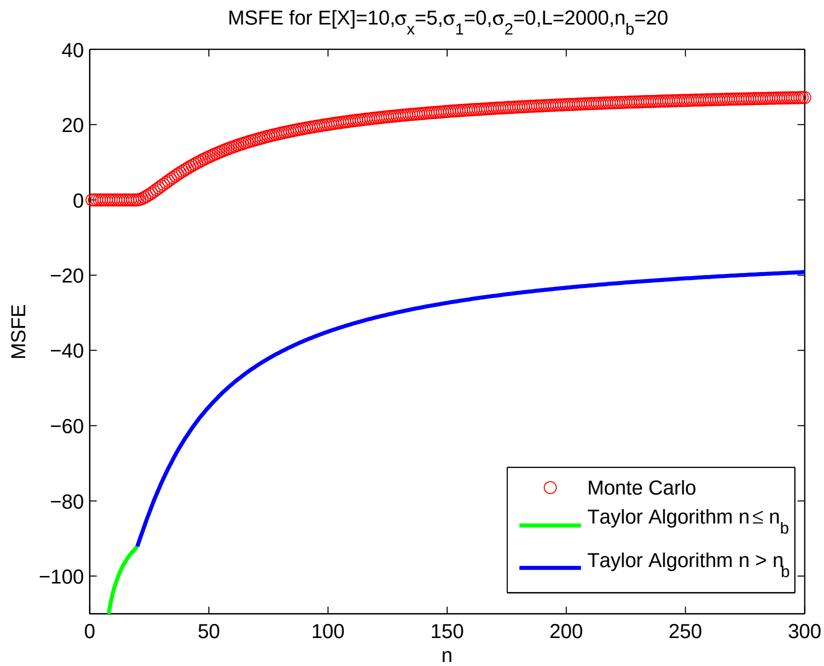


Figure 7.77: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

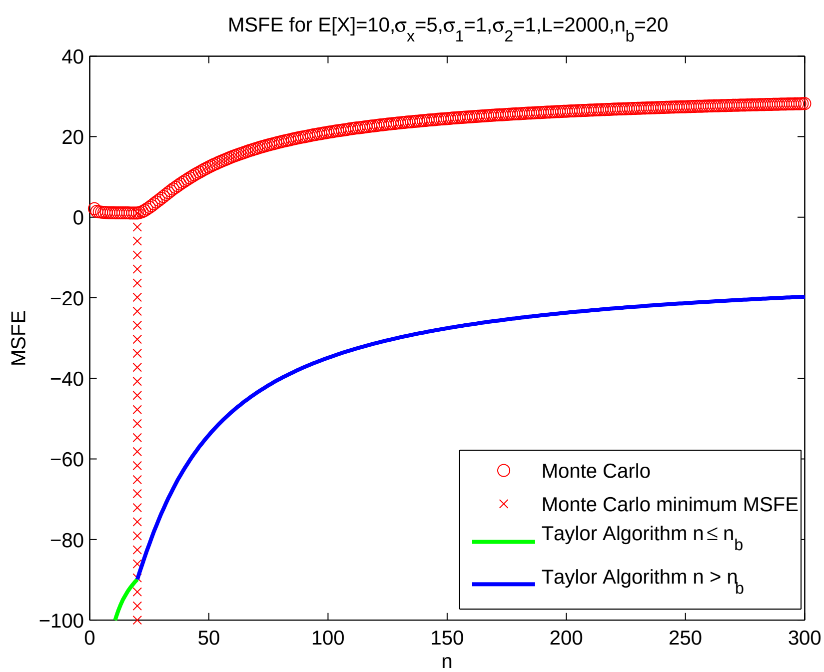


Figure 7.78: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$

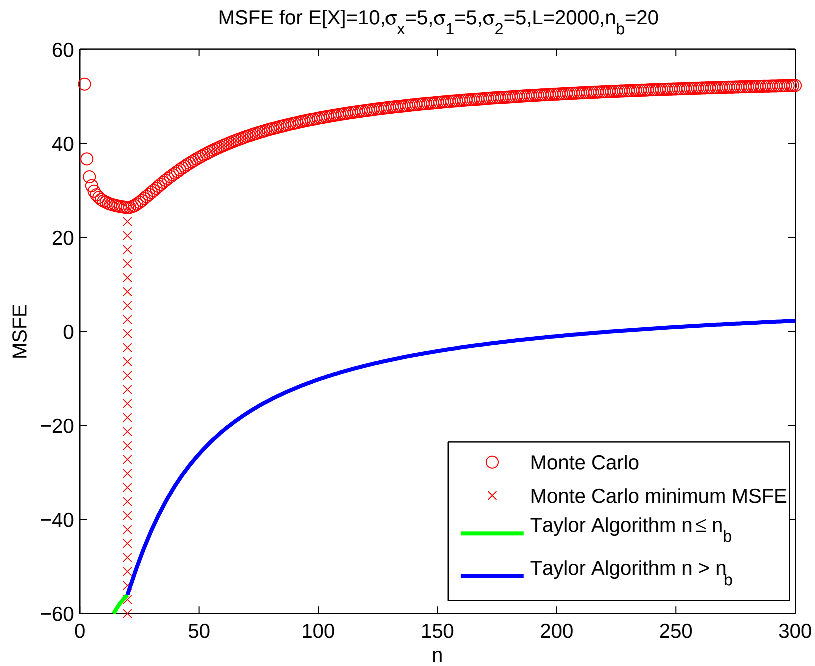


Figure 7.79: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$

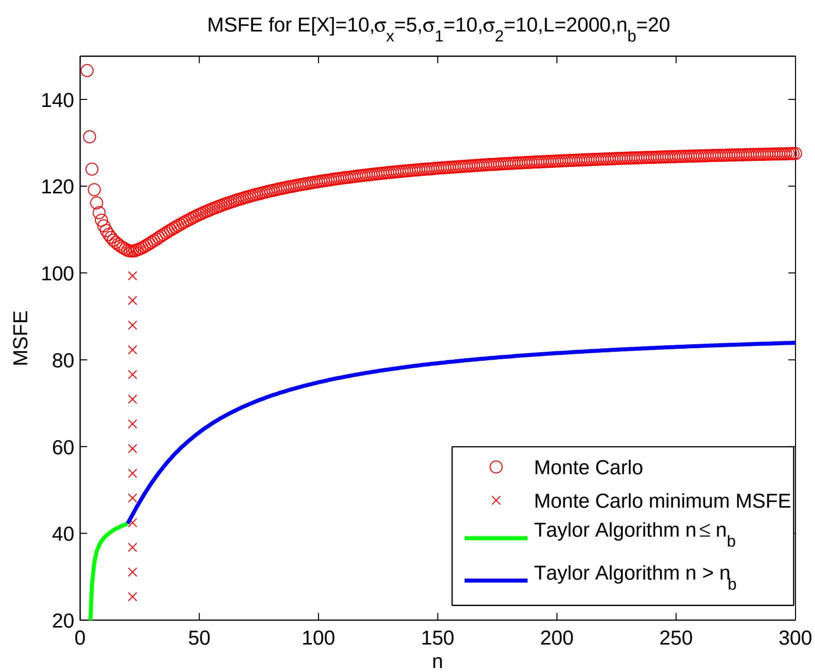


Figure 7.80: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$

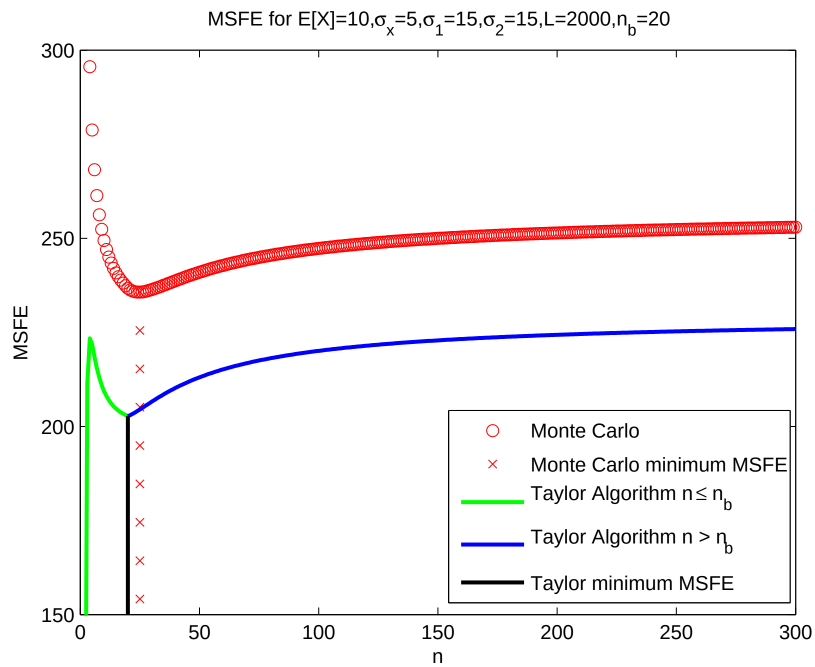


Figure 7.81: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 15$, $\sigma_2 = 15$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 25,Taylor algorithm minimum MSFE = 20

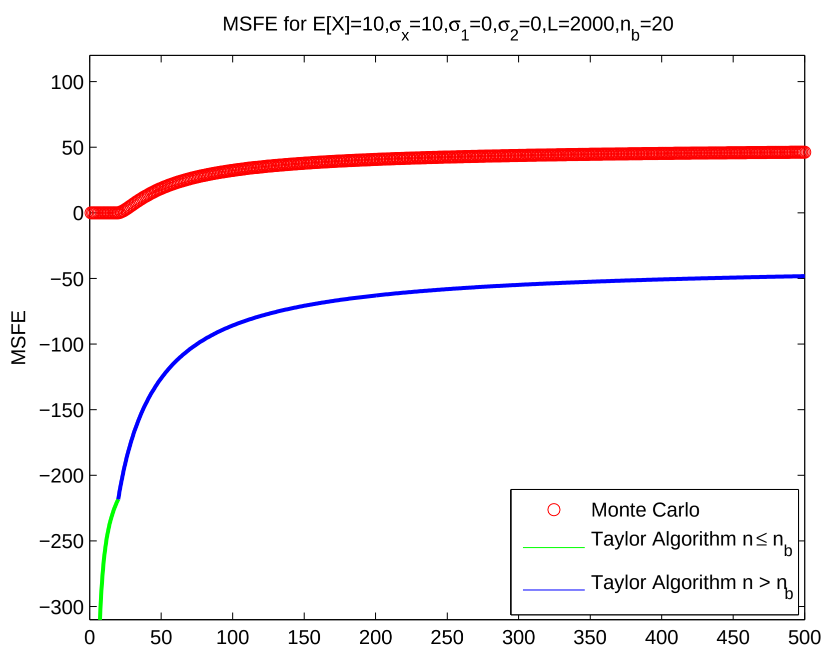


Figure 7.82: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

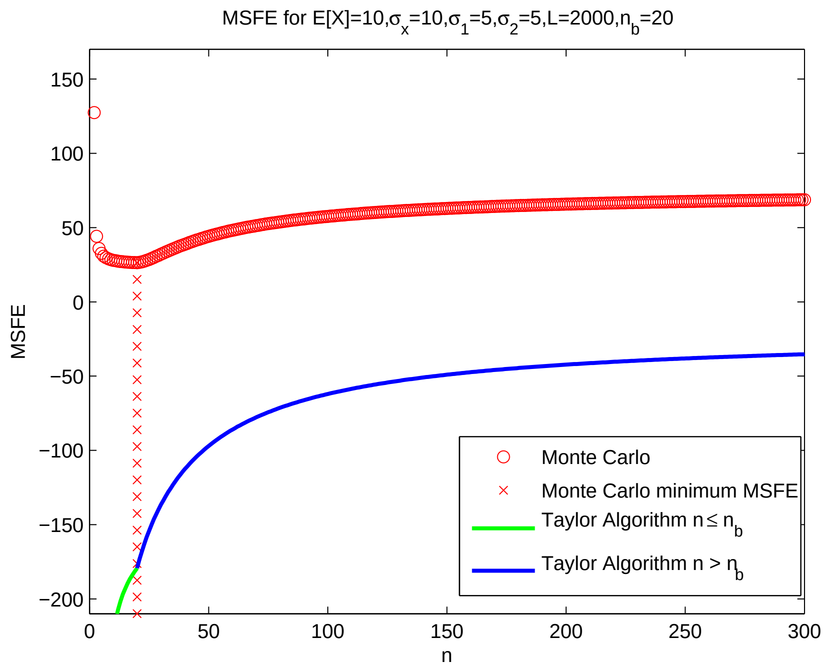


Figure 7.83: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$

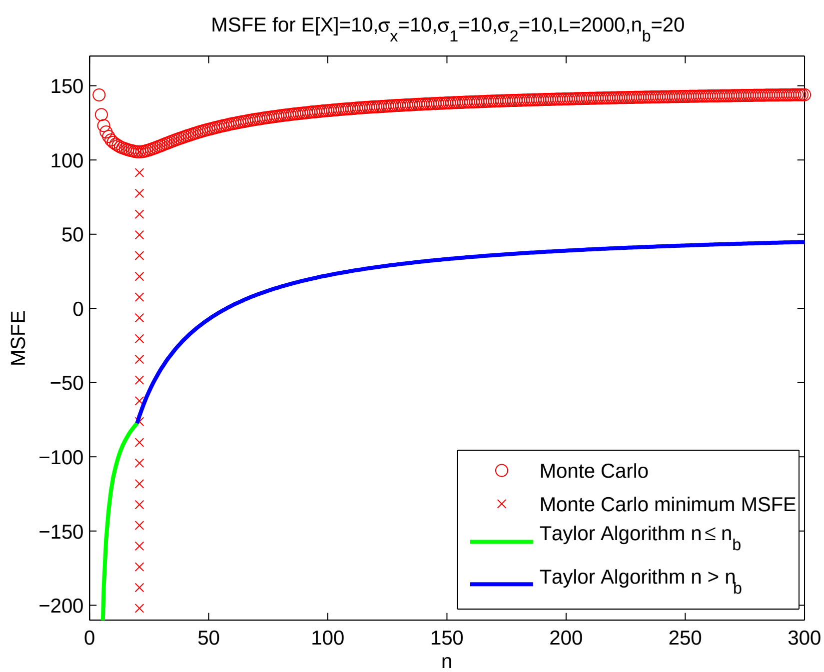


Figure 7.84: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$

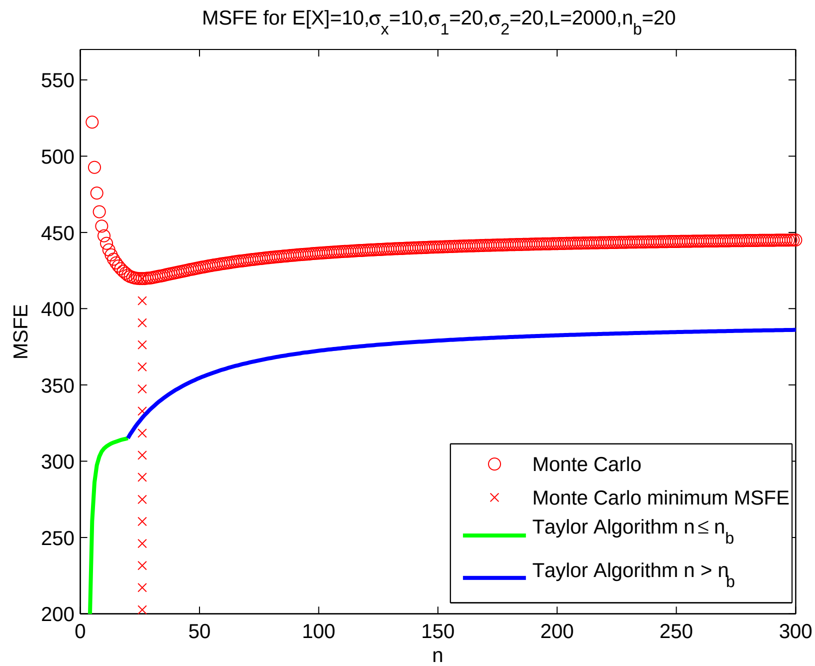


Figure 7.85: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 20$, $\sigma_2 = 20$, $\theta_1 = 2$, $\theta_2 = 2.5$

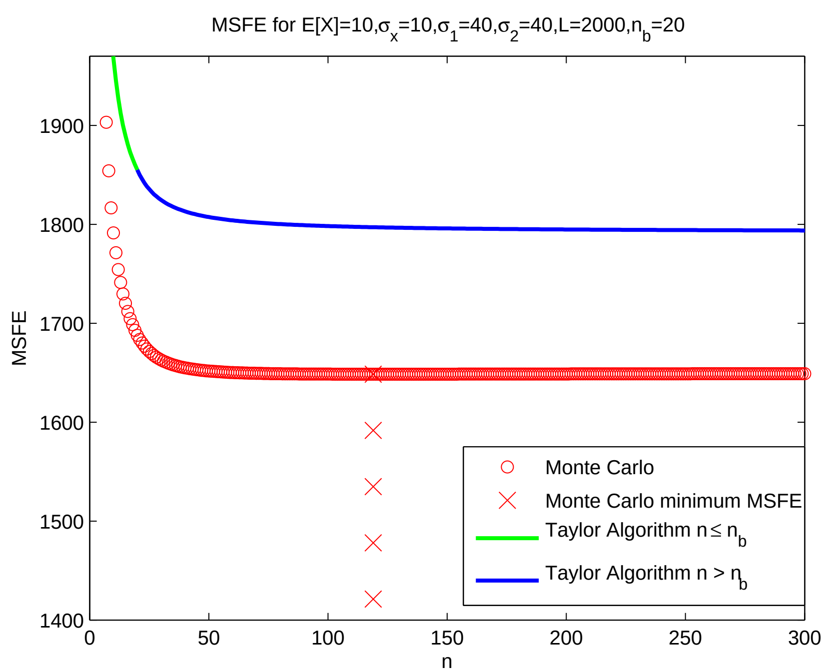


Figure 7.86: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 40$, $\sigma_2 = 40$, $\theta_1 = 2$, $\theta_2 = 2.5$

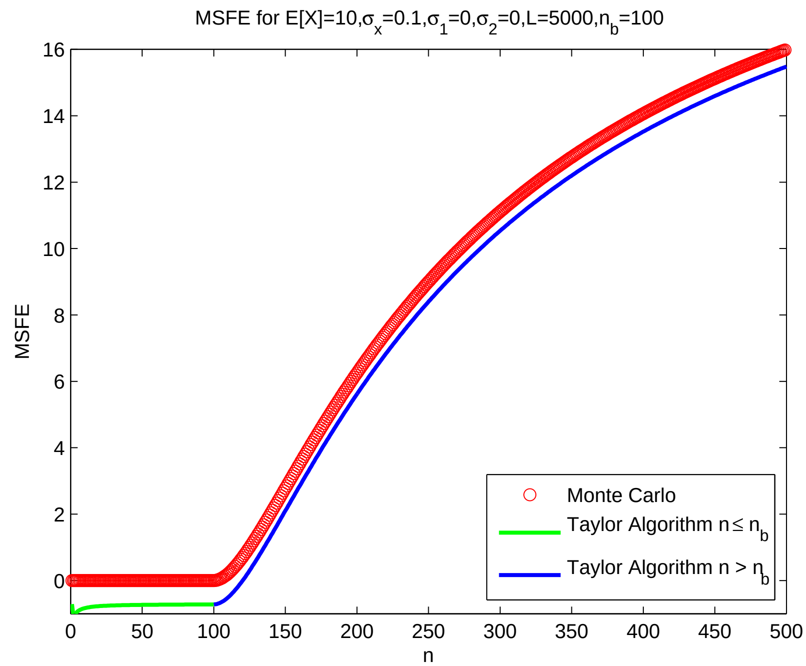


Figure 7.87: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

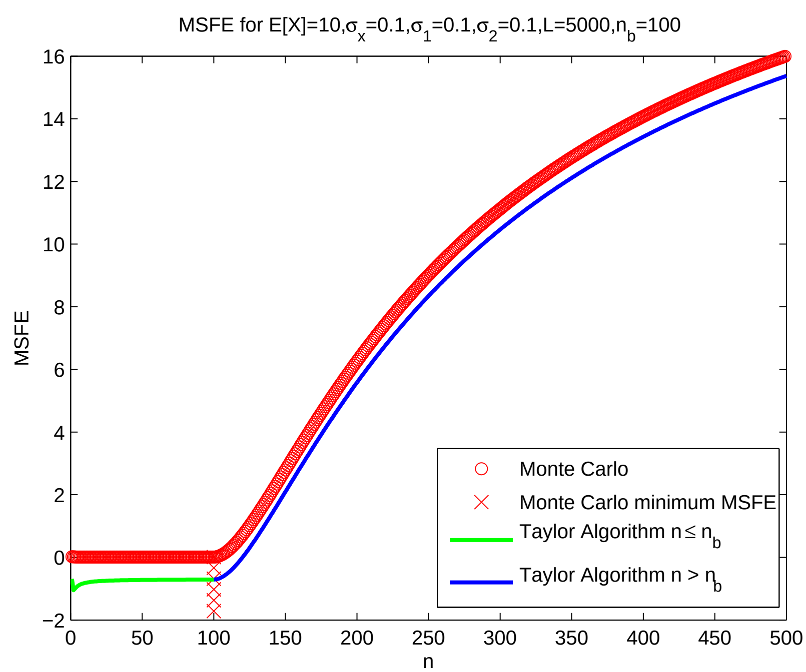


Figure 7.88: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\theta_1 = 2$, $\theta_2 = 2.5$

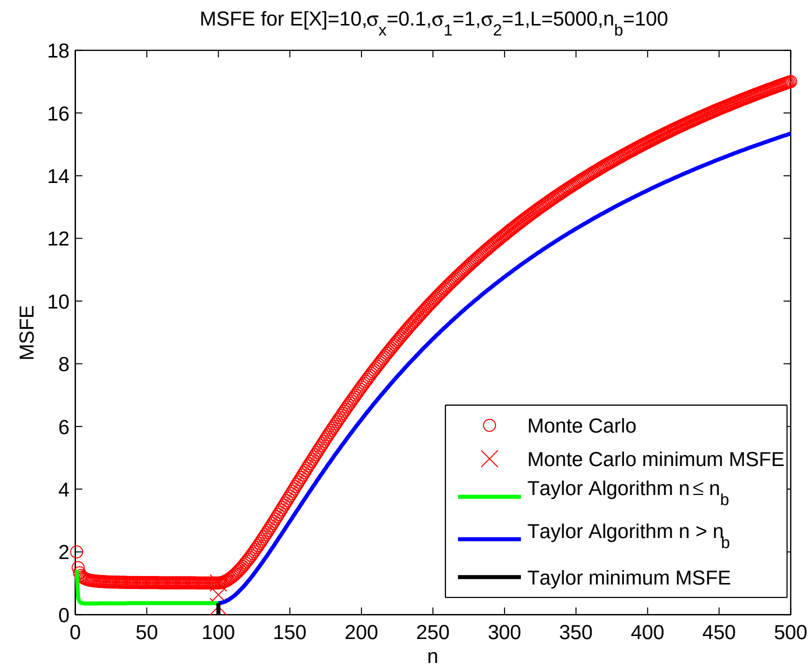


Figure 7.89: MSFE for $E[X] = 10$, $\sigma_x = 0.1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 100$

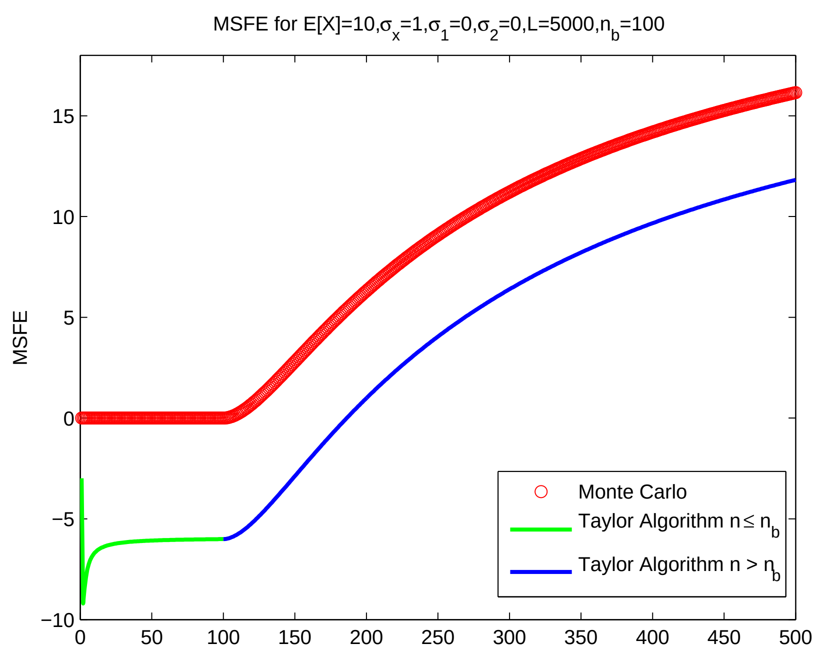


Figure 7.90: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

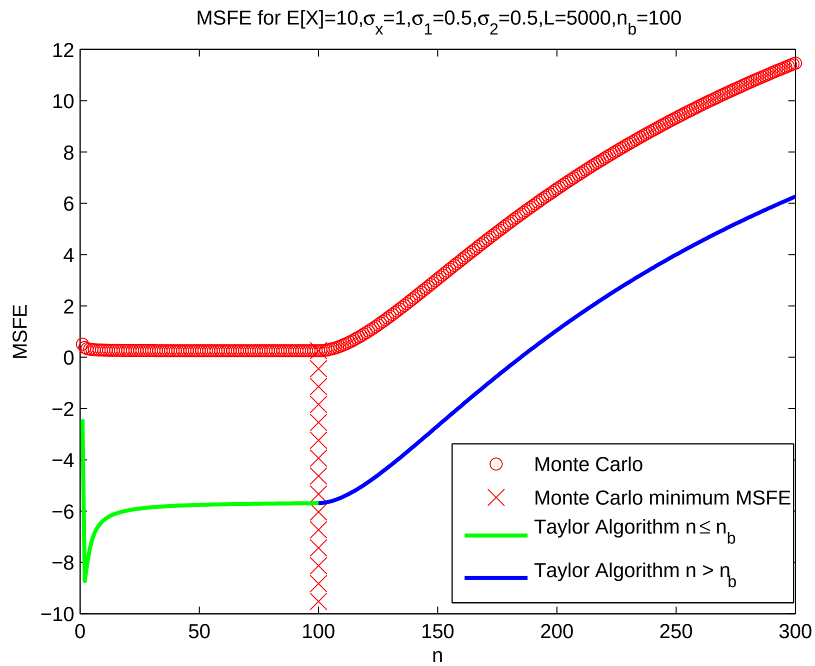


Figure 7.91: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\theta_1 = 2$, $\theta_2 = 2.5$

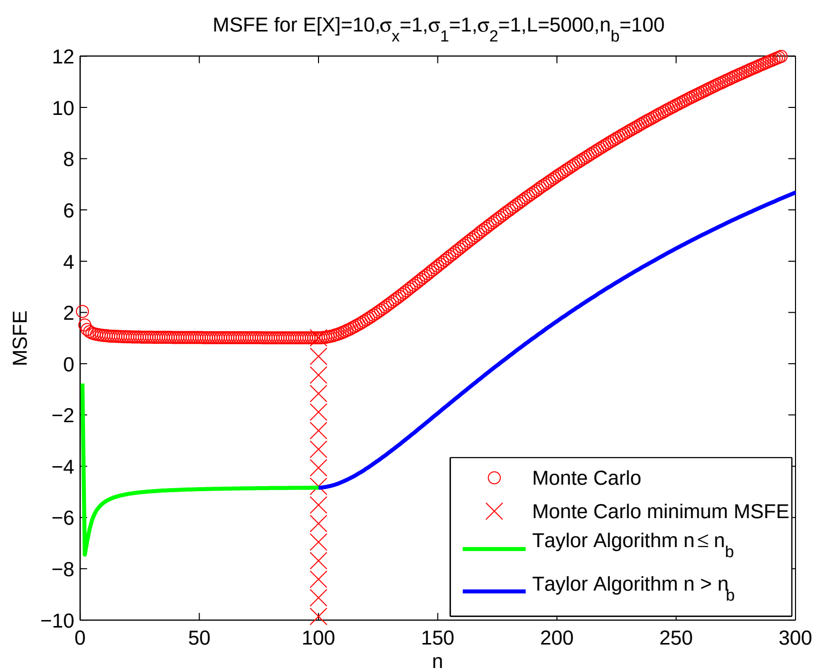


Figure 7.92: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$

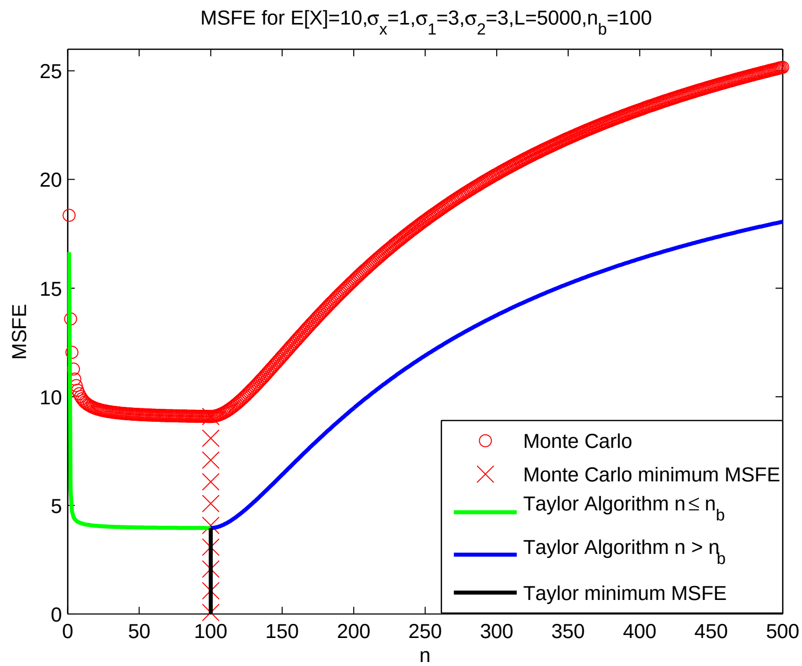


Figure 7.93: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 100$

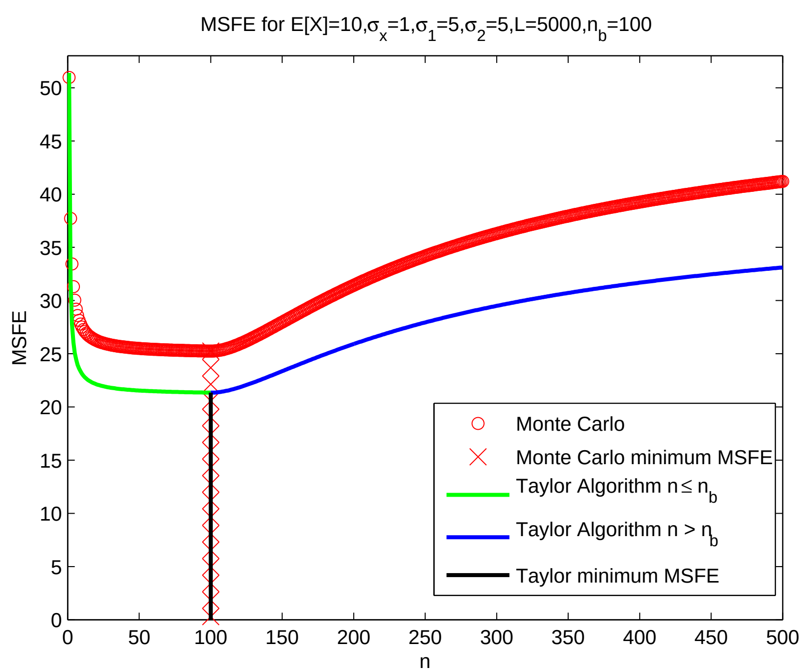


Figure 7.94: MSFE for $E[X] = 10$, $\sigma_x = 1$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 100$

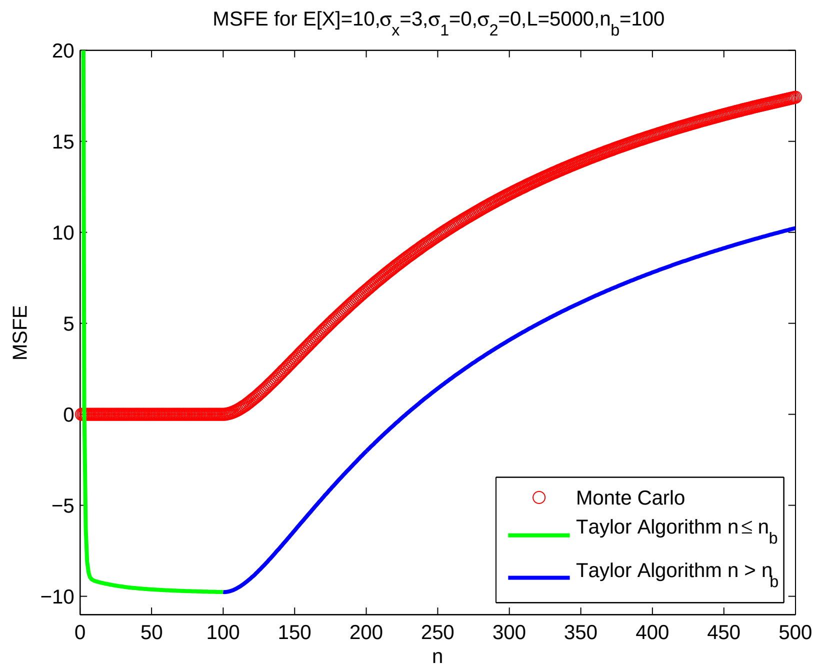


Figure 7.95: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

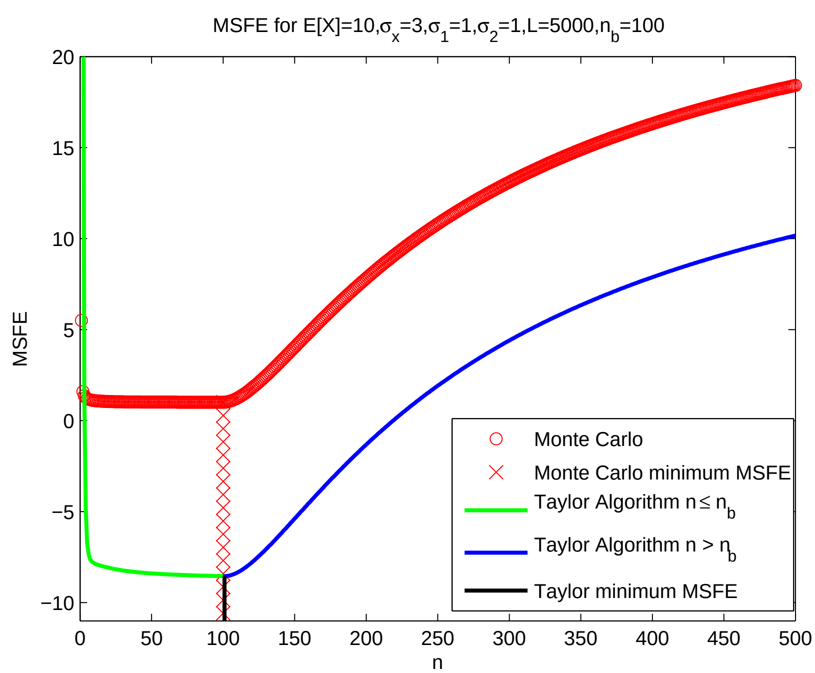


Figure 7.96: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 101$

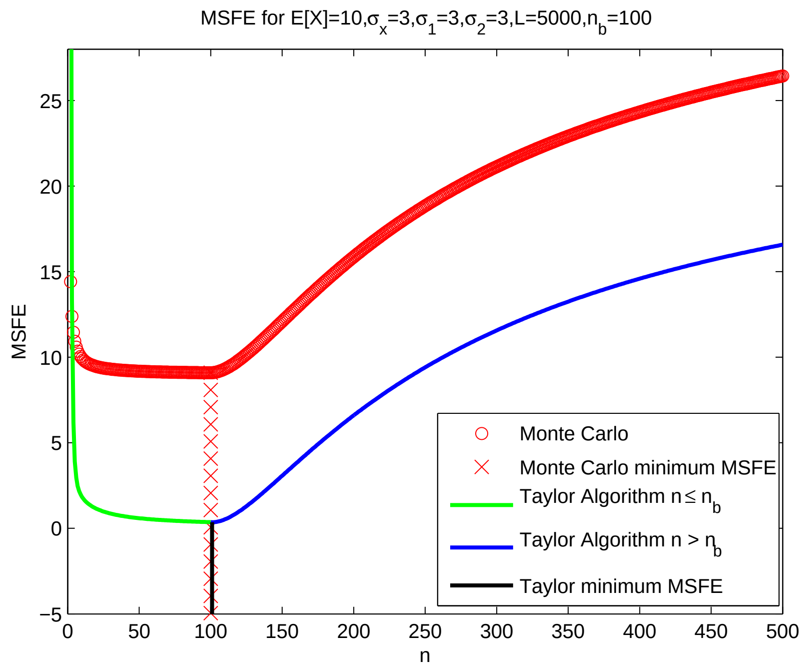


Figure 7.97: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 100,Taylor algorithm minimum MSFE = 101

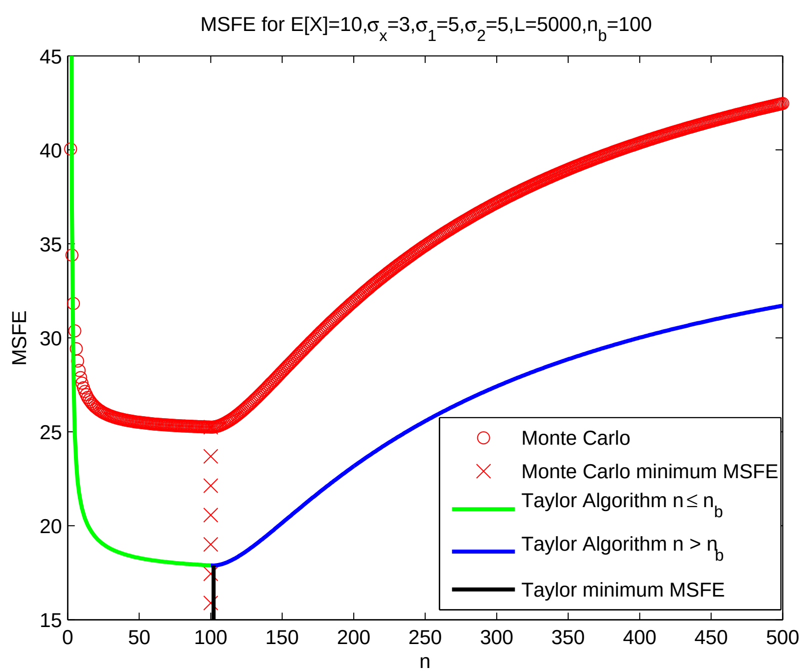


Figure 7.98: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 100,Taylor algorithm minimum MSFE = 102

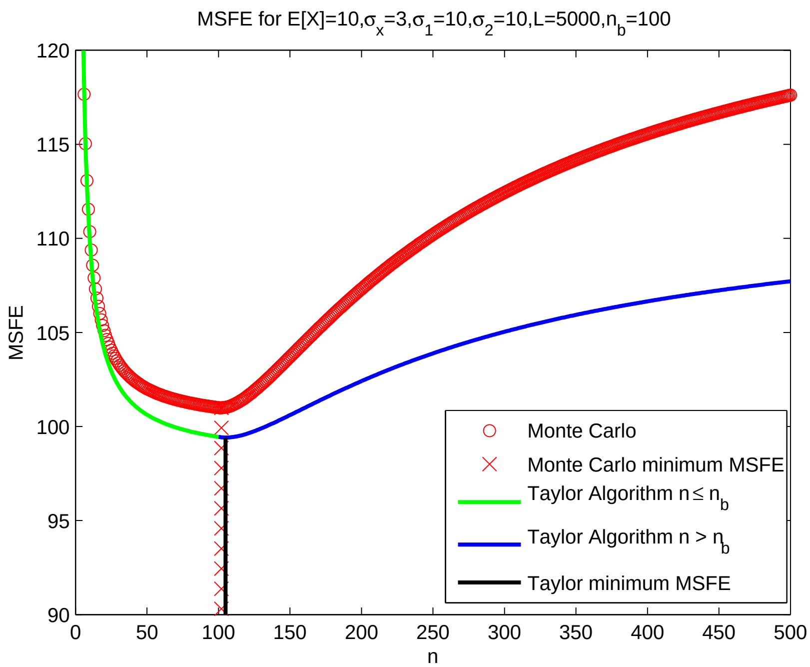


Figure 7.99: MSFE for $E[X] = 10$, $\sigma_x = 3$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 102,Taylor algorithm minimum MSFE = 105

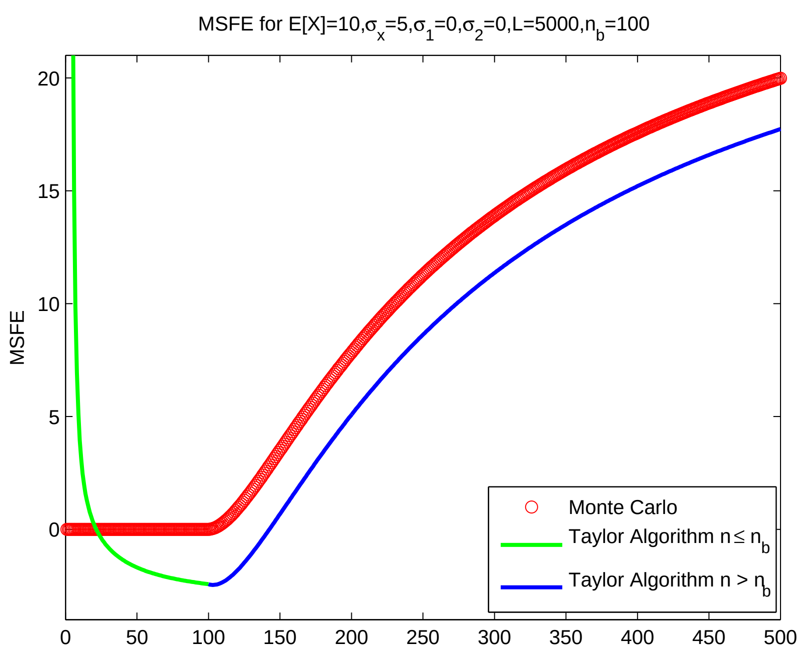


Figure 7.100: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$

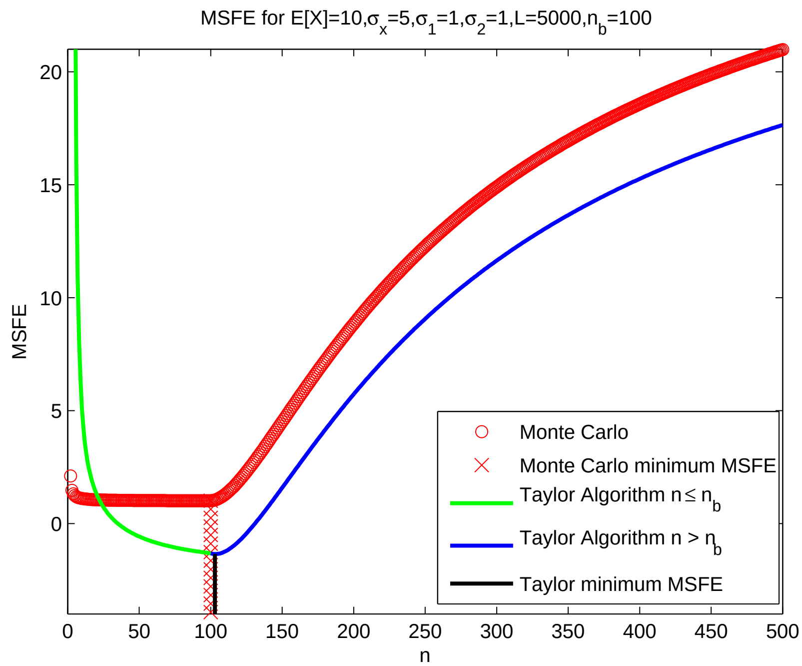


Figure 7.101: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 100,Taylor algorithm minimum MSFE = 103

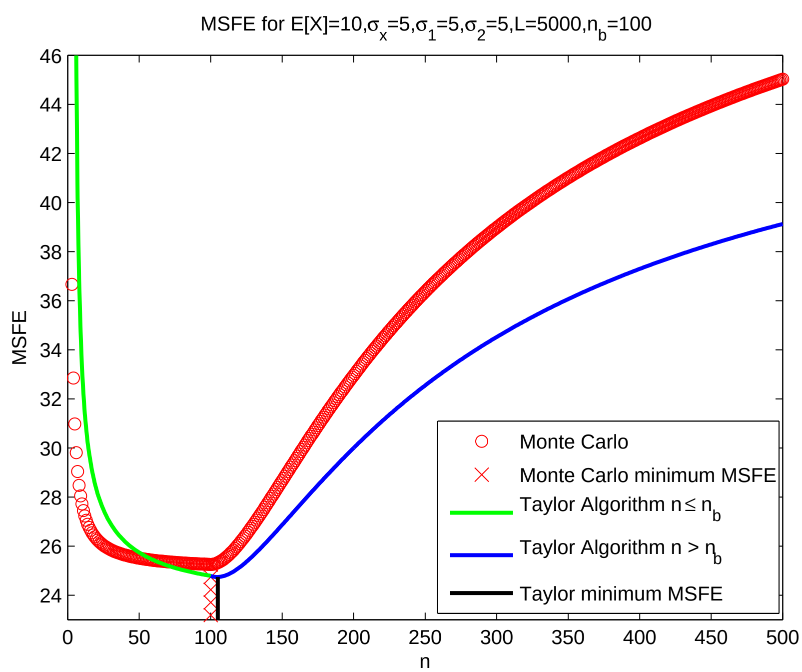


Figure 7.102: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE = 100,Taylor algorithm minimum MSFE = 105

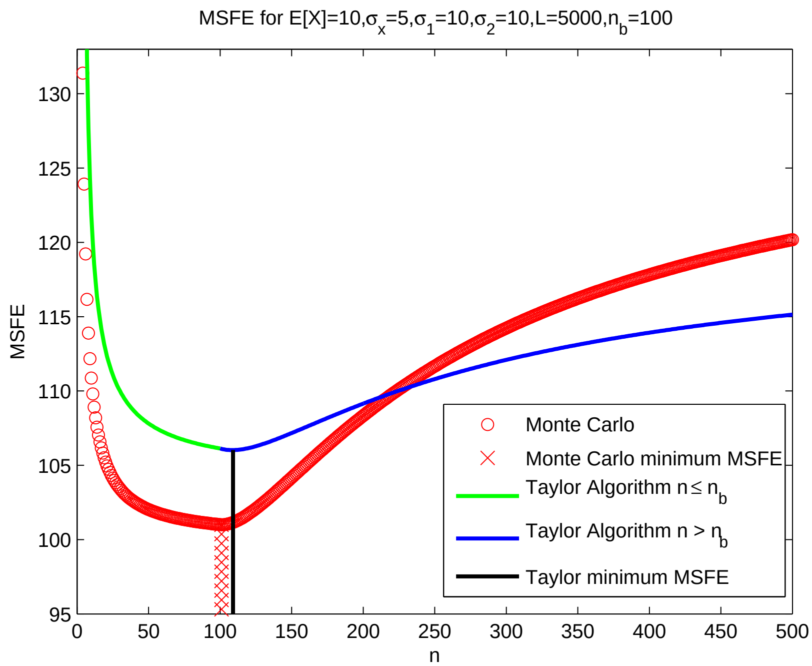


Figure 7.103: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 101$,Taylor algorithm minimum MSFE $= 109$

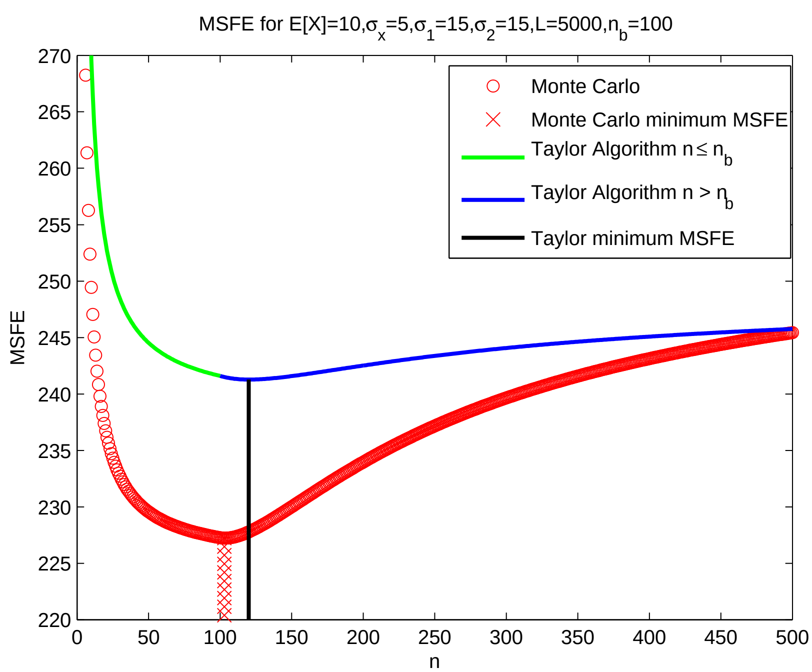


Figure 7.104: MSFE for $E[X] = 10$, $\sigma_x = 5$, $\sigma_1 = 15$, $\sigma_2 = 15$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 103$,Taylor algorithm minimum MSFE $= 120$

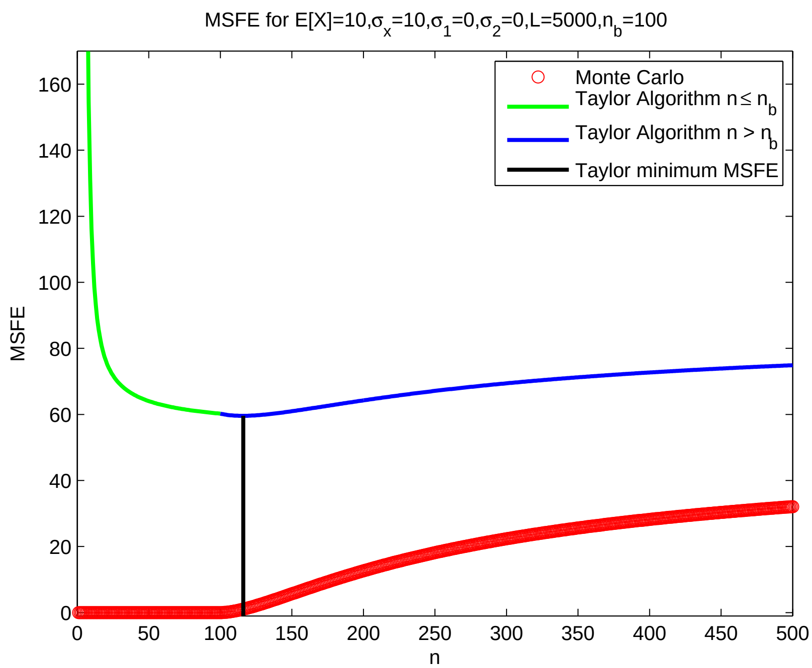


Figure 7.105: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 0$, $\sigma_2 = 0$, $\theta_1 = 2$, $\theta_2 = 2.5$.Taylor algorithm minimum MSFE $= 116$

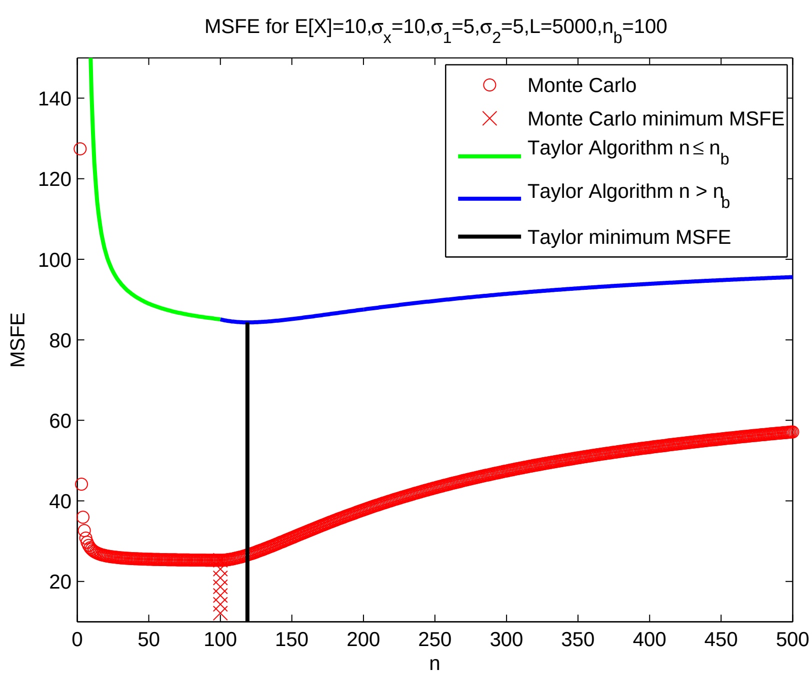


Figure 7.106: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 5$, $\sigma_2 = 5$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 119$

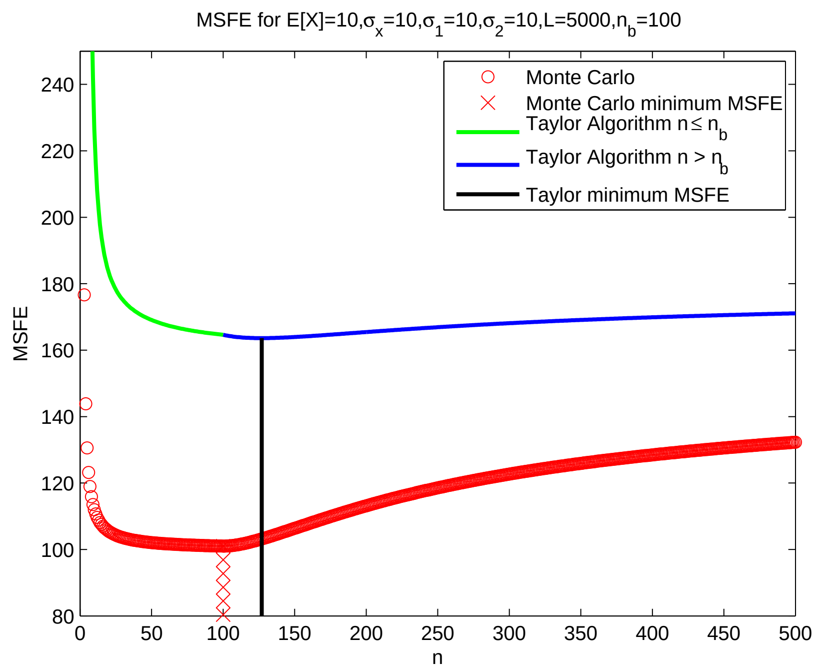


Figure 7.107: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 10$, $\sigma_2 = 10$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 100$,Taylor algorithm minimum MSFE $= 119$

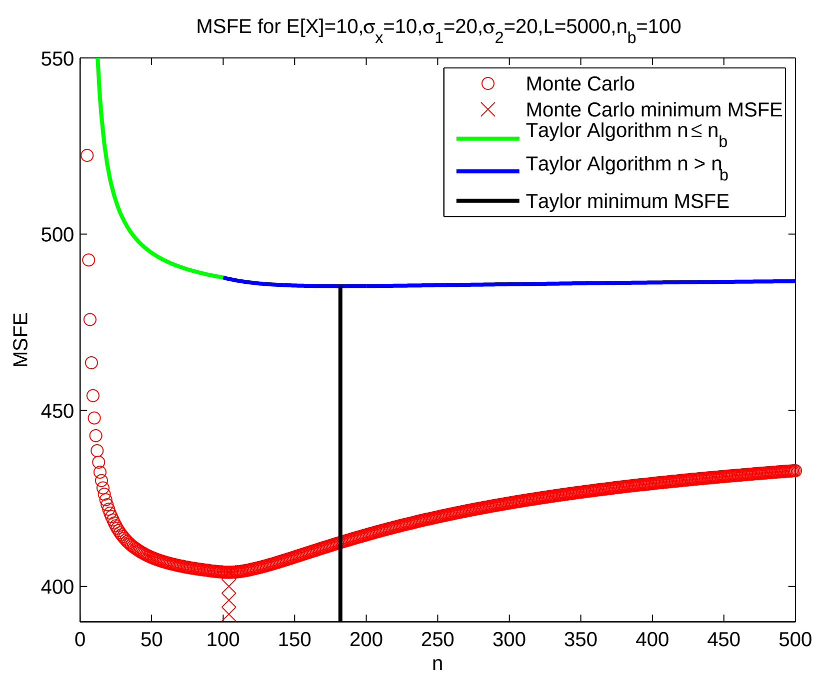


Figure 7.108: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 20$, $\sigma_2 = 20$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 104$,Taylor algorithm minimum MSFE $= 182$

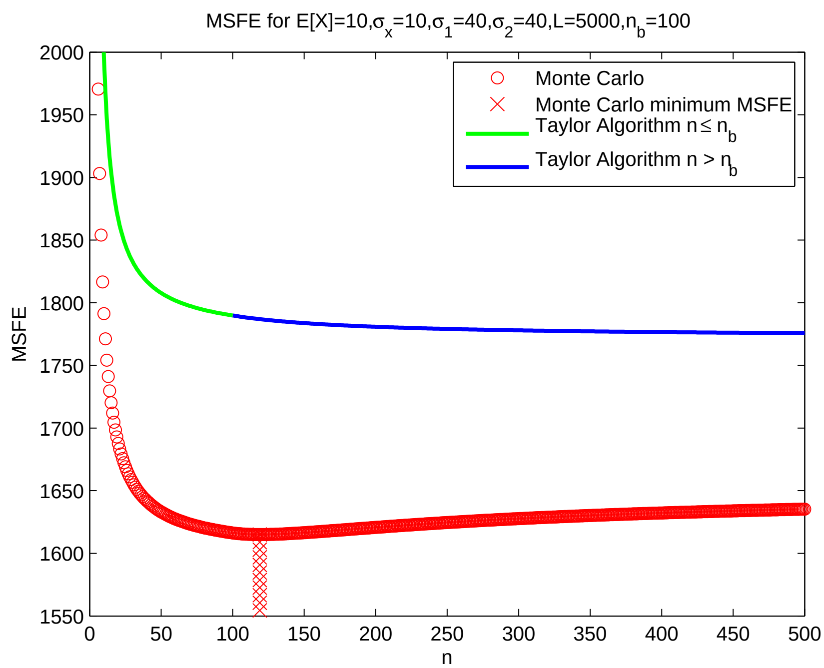


Figure 7.109: MSFE for $E[X] = 10$, $\sigma_x = 10$, $\sigma_1 = 40$, $\sigma_2 = 40$, $\theta_1 = 2$, $\theta_2 = 2.5$. Monte Carlo minimum MSFE $= 104$,Taylor algorithm minimum MSFE $= 182$

Chapter 8

The Delta Method

8.1 Introduction

When studying random processes, whether continuous or discrete, scalar or multivariate, all information concerning the process is contained in the distribution or joint distribution functions. Distribution functions can be functionally quite complicated. The expectation is the main tool which provides quantitative measures of different characteristics of the distribution and density functions. For example, in the case of a normally distributed random variable the expectation provides the center value around which observations occur. Similarly, the variance provides a measure of the dispersion of the events around the mean. In general, moments and central moments of random processes provide criteria by which one can understand the occurrences of random events.

Functions of random variables are ubiquitous in economics, econometrics, and finance, and therefore it becomes critical to understand the distribution of functions of random variables. The Delta method is a tool used in statistics to approximate the moments of a function of random variables. In this chapter, we begin by exploring the underlying tool used by the Delta method, the Taylor approximation. We follow with a literature overview of the different results that fall under the title of the Delta Method, including the conditions for their application.

The Delta method provides an approximation to the expectation of a function φ of random variables by taking expectation of a polynomial approximation to φ. This polynomial approximation is usually a truncated Taylor series centered at the population mean $E[X]$, and the convergence depends on the smoothness and boundedness of φ as well as the moments of X.

8.2 Delta Method for bounded functions

The first version of the Delta method was presented by Cramér [36, p.353]. Cramér uses the Delta method to approximate the mean and variance of some function of sample moments. We present the theorem and its proof to illustrate the methods and assumptions. We begin with a random variable X with distribution F and $X_1, \ldots, X_n$ an i.i.d. random sample from F. $x_1, \cdots, x_n$ is a realization of the random sample. μ_j is the jth population moment of X, $\mu_j = E[X^j]$, and $\bar{\mu}_j$ is the jth population central moment, $\bar{\mu}_j = E[(X-\mu)^j]$. Denote the jth sample moment by $m_{j,n} = \sum_{i=1}^n x_i^j/n$ and the jth sample central moment by $\bar{m}_{j,n} = \sum_{i=1}^n (x_i - \bar{x})^j/n$. Both $m_{j,n}$ and $\bar{m}_{j,n}$ are functions from $\mathbb{R}^n$ into $\mathbb{R}$ and therefore a function depending on these moments is a function on $\mathbb{R}^n$. Given a function φ of two sample central moments $\bar{m}_{i,n}, \bar{m}_{j,n}$, the mean and variance of φ can be estimated as follows:

Theorem 8.1 (Cramér) *Suppose:*

1. In some neighborhood of the point $\bar{m}_{i,n} = \bar{\mu}_i$, $\bar{m}_{j,n} = \bar{\mu}_j$ the function φ is continuous and has continuous derivatives of the first and second order with respect to the arguments m_v and m_ρ.

2. For all possible values of x_i, it follows $|\varphi| < Cn^p$, where C and p are non-negative constants.

Denoting $\varphi_0 = \varphi(\bar{\mu}_i, \bar{\mu}_j)$, $\varphi_1 = \partial\varphi/\partial\bar{m}_i\,(\bar{\mu}_i, \bar{\mu}_j)$ and $\varphi_2 = \partial\varphi/\partial\bar{m}_j\,(\bar{\mu}_i, \bar{\mu}_j)$, the mean and variance of the random variable $\varphi(\bar{m}_i, \bar{m}_j)$ are :

$$
\begin{aligned}
&E[\varphi(\bar{m}_i, \bar{m}_j)] = \varphi_0 + O(n^{-1}),\\
&Var(\varphi) = \mu_2(\bar{m}_i)\varphi_1^2 + 2\mu_{11}(\bar{m}_i, \bar{m}_j)\varphi_1\varphi_2 + \mu_2(\bar{m}_2)\varphi_2^2 + O(n^{-3/2}).
\end{aligned}
$$

Proof. Let $P(S)$ be the probability function of the joint distribution of $X_1, \cdots, X_n$. $P(S)$ is a set function in $\mathbb{R}^n$. Since $X_1, \cdots, X_n$ is a random sample of X, we know from the characteristics of sampling distributions

$$E[(\bar{m}_{i,n} - \bar{\mu}_i)^{2k}] = O(n^{-k}). \tag{8.2.1}$$

Using this, and by Tchebycheff's theorem, it follows that

$$P[(\bar{m}_{i,n} - \bar{\mu}_i)^{2k} \geq \epsilon^{2k}] < \frac{E[(\bar{m}_{i,n} - \bar{\mu}_i)^{2k}]}{\epsilon^{2k}} = \frac{A}{\epsilon^{2k} n^k},$$

or

$$P[|\bar{m}_{i,n} - \bar{\mu}_i| \geq \epsilon] < \frac{A}{\epsilon^{2k} n^k}, \tag{8.2.2}$$

for some constant A independent of ϵ and n. The corresponding inequalities hold for $\bar{m}_{j,n}$. Define the set $Z = \{(x_1, \cdots, x_n) : |\bar{m}_{i,n} - \bar{\mu}_i| < \epsilon, |\bar{m}_{j,n} - \bar{\mu}_j| < \epsilon\}$ and denote by Z^c the complement of Z. It follows from (8.2.2) that

$$P(Z^c) < \frac{2A}{\epsilon^{2k} n^k}, \quad P(Z) > 1 - \frac{2A}{\epsilon^{2k} n^k}. \tag{8.2.3}$$

It follows that $E[\varphi] = \int_Z \varphi dP + \int_{Z^c} \varphi dP$. By condition 2), (8.2.3) and choosing $k > p+1$ $\left|\int_{Z^c} \varphi dP\right| < 2ACn^p/\epsilon^{2k} n^k = O(n^{-1})$. For ϵ small enough, it follows from condition 1) that for any point in Z

$$\begin{aligned} &\varphi(\bar{m}_i, \bar{m}_j) = \varphi_0 + (\bar{m}_i - \bar{\mu}_i)\varphi_1 + (\bar{m}_j - \bar{\mu}_j)\varphi_2 + R, \\ &R = \frac{1}{2}\left[(\bar{m}_i - \bar{\mu}_i)^2 \varphi'_{11} + 2(\bar{m}_i - \bar{\mu}_i)(\bar{m}_j - \bar{\mu}_j)\varphi'_{12} + (\bar{m}_j - \bar{\mu}_j)^2 \varphi'_{22}\right], \end{aligned}$$

where φ'_{ij} denotes second order derivatives evaluated at a point between $(\bar{m}_i, \bar{m}_j)$ and $(\bar{\mu}_i, \bar{\mu}_j)$. It follows

$$\int_Z \varphi dP = \varphi_0 P(Z) + \varphi_1 \int_Z (\bar{m}_i - \bar{\mu}_i) dP + \varphi_2 \int_Z (\bar{m}_j - \bar{\mu}_j) dP + \int_Z R dP. \tag{8.2.4}$$

By (8.2.3), the first term on the right of the equality differs from φ_0 by a quantity of order n^{-k} which is smaller than n^{-1} by our choice of k. For the other two terms we first note φ_1 and φ_2 are independent of n. Furthermore, given that for sample distributions

$$E[\bar{m}_{i,n}] = \bar{\mu}_i + O(n^{-1}) \quad \text{and} \quad E[(\bar{m}_{i,n} - \bar{\mu}_i)^{2k}] = O(n^{-k}) \tag{8.2.5}$$

and applying Schwarz inequality it follows

$$\begin{aligned}
\int_Z (\bar{m}_{i,n} - \bar{\mu}_i) dP &= E[\bar{m}_{i,n} - \bar{\mu}_i] - \int_{Z^c} (\bar{m}_{i,n} - \bar{\mu}_i) dP \\
&= O(n^{-1}) - \int_{Z^c} (\bar{m}_{i,n} - \bar{\mu}_i) dP, \\
\left| \int_{Z^c} (\bar{m}_{i,n} - \bar{\mu}_i) dP \right| &\leq \left[\int_{Z^c} (\bar{m}_{i,n} - \bar{\mu}_i)^2 dP \int_{Z^c} dP \right]^{1/2} \\
&\leq \left[E[(\bar{m}_{i,n} - \bar{\mu}_i)^2] P(Z^c) \right]^{1/2} = O(n^{-(k+1)/2}),
\end{aligned}$$

and similarly for $\bar{m}_{j,n}$. The derivatives φ'_{ij} are bounded for sufficiently small ϵ by condition 1), and it follows that the last term in (8.2.4) is of order n^{-1}. Hence the right hand side of (8.2.4) differs from φ_0 by a quantity of order n^{-1}, and this proves the first relation of the theorem. We omit the proof of the variance term and direct the reader to the original text. ∎

In summary, Cramér proves a Delta method for a function of two central moments which depends on the sample size n only through the sample moments. The same proof can be extended for functions of any number of central moments. The main assumptions on the function φ are first that φ is bounded by Cn^p for positive constants C, p and second that φ is twice continuously differentiable in a neighborhood of the population moments $\bar{\mu}_i$ and $\bar{\mu}_j$. The process X is assumed to have sufficient finite moments. The fact that the function φ has as its arguments sample moments, makes Cramér's result rather restrictive. This can be seen from the required bounds (8.2.1) and (8.2.5), which are derived for characteristics of sampling distributions.

Hurt, in [76], expands the application of the Delta method by allowing more general random variables as arguments of the function φ, by allowing the function φ to depend explicitly on the sample size n, and by taking more terms of the Taylor series expansion in the approximation. Specifically, he derives asymptotic formulas for $E[\varphi(T_n, n)]$ and $Var(\varphi(T_n, n))$ where T_n is a possibly multi-dimensional statistic. The order of the remainder depends on the smoothness of the function φ and on the size of the moments of T_n. The main theorem when T_n is a one dimensional statistic is as follows:

Theorem 8.2 (Theorem 1 in [76], Hurt) *Let $\varphi = \varphi(t,n)$ be a function defined on $\mathbb{R}^1 \times N$. Assume, for all n and some $q \geq 1$, φ admits the continuous $(q+1)st$ derivative for $t \in [\theta-\delta, \theta+\delta]$ where $\delta > 0$ is independent of n. Suppose φ is bounded on $\mathbb{R}^1 \times N$ and all derivatives $\varphi', \cdots, \varphi^{(q+1)}$ are bounded on $[\theta-\delta, \theta+\delta] \times N$. Let $\{T_n\}$ be a sequence of statistics with finite moments up to order $2(q+1)$ such that $E|T_n - \theta|^{2(q+1} = O(n^{-(q+1)})$. Then*

$$E[\varphi(T_n, n) - \varphi(\theta, n)] = \sum_{j=1}^{q} \frac{1}{j!} \left(\frac{\partial^j \varphi}{\partial t^j}\right)_{t=\theta} E[(T_n - \theta)^j] + O(n^{-(q+1)/2}),$$

$$Var[\varphi(T_n, n) - \varphi(\theta, n)] = \sum_{j=1}^{q} \sum_{k=1}^{q} \frac{1}{j!} \frac{1}{k!} \left(\frac{\partial^j \varphi}{\partial t^j}\right)_{t=\theta} \left(\frac{\partial^k \varphi}{\partial t^k}\right)_{t=\theta}$$

$$\cdot\, cov[(T_n - \theta)^j, (T_n - \theta)^k] + O(n^{-(q+2)/2}).$$

The theorem for the multi-dimensional case follows:

Theorem 8.3 (Theorem 2 in [76],Hurt) *Let $\varphi(t_1, \ldots, t_r, n)$ be a function defined on $\mathbb{R}^r \times N$. Assume:*

1) for all n, φ is $(q+1)$ times totally differentiable with respect to t_i's in the interval $K = X_{i=1}^r [\theta_i - \delta_i, \theta_i + \delta_i]$, $\delta_1 > 0$, δ_i independent of n,

2) φ is bounded on $\mathbb{R}^r \times N$,

3) all the derivatives up to the order $q+1$ are bounded on $K \times N$,

4) $\{(T_{1n}, \cdots, T_{rn})\}_{n=1}^{\infty}$ is a sequence of multidimensional statistics such that

5) there exists absolute moments of T_{in} up to order $2(q+1)$

6) for $i = 1, \cdots, r$ $E|T_{in} - \theta_i|^{2(q+1)} = O(n^{-(q+1)})$

Then with $i_1 + \cdots + i_r = j$:

$$E[\varphi(T_{1n}, \cdots, T_{rn}, n) - \varphi(\theta_1, \cdots, \theta_r, n)] =$$
$$\sum_{j=1}^{q} \frac{1}{j!} \sum_{i_1} \cdots \sum_{i_r} \left[\frac{\partial^j \varphi}{\partial t_1^{i_1} \ldots \partial t_r^{i_r}}\right]_{t=\theta} E[(T_{1n} - \theta_1)^{i_1} \ldots (T_{rn} - \theta_r)^{ir}] + O(n^{-(q+1)/2}),$$

and with $i_1 + \cdots + i_r = j,\ m_1 + \cdots + m_r = k$

$$Var[\varphi(T_{1n}, \cdots, T_{rn}, n) - \varphi(\theta_1, \cdots, \theta_r, n)] =$$
$$\sum_{j=1}^{q}\sum_{k=1}^{q}\frac{1}{j!}\frac{1}{k!}\sum_{i_1}\cdots\sum_{i_r}\sum_{m_1}\cdots\sum_{m_r}\left[\frac{\partial^j\varphi}{\partial t_1^{i_1}\dots\partial t_r^{i_r}}\right]_{t=\theta}\left[\frac{\partial^k\varphi}{\partial t_1^{m_1}\dots\partial t_r^{m_r}}\right]_{t=\theta}$$
$$\cdot\, Cov\left[(T_{1n}-\theta_1)^{i_1}\dots(T_{rn}-\theta_r)^{ir}, (T_{1n}-\theta_1)^{m_1}\dots(T_{rn}-\theta_r)^{mr}\right] + O(n^{-(q+2)/2}),$$

where $t = (t_1, \cdots, t_r), \theta = (\theta_1, \cdots, \theta_r)$.

The proof for theorems 8.2 and 8.3 follow similarly as the proof by Cramér in that the expected value is split into an integral on a neighborhood around the corresponding population statistic θ, $Z = \{t : |t - \theta| < \varepsilon\}$ for the one dimensional case, and an integral on the complement of said neighborhood, Z^c. The dependence of the size of the remainder on the sample size n follows from the assumption that the sequence $\{T_n\}$ of statistics has finite moments up to order $2(q + 1)$ such that $E|T_n - \theta|^{2(q+1)}$.

The Delta method theorems up to this point assume boundedness of the function φ. There are many unbounded functions that are of interest, such as the squared error function. In the next section, we examine Delta method results for some classes of unbounded functions.

8.3 Delta Method for polynomial bounded functions

Lehmann [92] presents a Delta method for the special case where φ does not need to be bounded as long as the derivatives of φ up to some order exist and are bounded.

Theorem 8.4 (Theorem 5.1 in [92], Lehmann) *Let* $X_1, \cdots, X_n$ *be i.i.d. with* $E[X_1] = \xi$, $Var(X_1) = \sigma^2$ *and finite fourth moment. Suppose* φ *is a function of a real variable whose first four derivatives* $\varphi'(x), \varphi''(x), \varphi'''(x), \varphi^{(iv)}(x)$ *exist for all* $x \in I$ *where* I *is an interval with* $P(X_1 \in I) = 1$, *and such that* $|\varphi^{(iv)}(x)| \leq M$ *for all* $x \in I$, *for some* $M < \infty$. *Then*

$$E[\varphi(\bar{X})] = \varphi(\xi) + \frac{\sigma^2}{2n}\varphi''(\xi) + R_n,$$

and if, in addition, the fourth derivative of φ^2 is also bounded,

$$Var(\varphi(\bar{X})) = \frac{\sigma^2}{n}(\varphi'(\xi))^2 + R_n,$$

where the remainder R_n in both cases is $O(n^{-2})$.

Proof. The result follows from the strong assumptions on the function φ and the fact that $E[(\bar{X}-\xi)^{2k-1}]$ and $E[(\bar{X}-\xi)^{2k}]$, if they exist, are of order $1/n^k$ for $k \geq 1$. First, we make note of the following relations:

$$E[\bar{X}-\xi] = 0, \quad E[(\bar{X}-\xi)^2] = \frac{\sigma^2}{n}, \quad E[(\bar{X}-\xi)^3] = O(n^{-2}), \quad E[(\bar{X}-\xi)^4] = O(n^{-2}).$$

If for all x, the fourth derivative $\varphi^{(iv)}$ exists and satisfies $|\varphi^{(iv)}(x)| \leq M$ for some $M < \infty$, then

$$\varphi(\bar{x}) = \varphi(\xi) + \varphi'(\xi)(\bar{x}-\xi) + \frac{1}{2}\varphi''(\xi)(\bar{x}-\xi)^2 + \frac{1}{6}\varphi'''(\xi)(\bar{x}-\xi)^3 + R(\bar{x},\xi), \quad (8.3.1)$$

where $|R(\bar{x},\xi)| \leq M(\bar{x}-\xi)^4/24$. Taking expectations of (8.3.1) the result follows. ∎

Theorem 8.5 (Theorem 5.1a in [92], Lehmann) *The results in theorem* 8.4 *remain valid if for some $k \geq 3$ the function φ has k derivatives, the kth derivative is bounded, and the first k moments of the X's exists.*

The assumptions of bounded derivatives of the function φ up to some order k are equivalent to polynomial boundedness of φ by a polynomial of order of at least k.

In [107], Oehlert attempts extend previous Delta method theorems in that the approximating polynomial does not need to be a truncated Taylor series and that the function in question needs to be only polynomially bounded in its arguments. We give some notation and state the theorem. The theorem is proven for functions of the normalized sample moments $u_{j,n} = \sum_{i=1}^{n}(x_i^j - \mu_j)/\sqrt{n}$. For polynomials in the first J normalized sample moments, let $p = (p_1, p_2, \ldots, p_J)^T$ be a vector of powers and $u^p = u_n^p = u_{1,n}^{p_1} u_{2,n}^{p_2} \cdots, u_{J,n}^{p_J}$. The sets P_A and P_B are finite sets of powers that define the approximating and bounding polynomials.

Theorem 8.6 (Oehlert) *Let the random variables $u_{i,n}$ be the normalized sample moments of an i.i.d. sample of size n from a distribution with finite tth moment. Suppose that there are approximating and bounding polynomials*

$$A_n(u_n) = \sum_{p \in P_A} a_{n,p} u^p, \quad B_n(u_n) = \sum_{p \in P_B} b_{n,p} u^p,$$

such that

$$n^\beta |\varphi(n, u_n) - A_n(u_n)| \xrightarrow{P} 0, \tag{8.3.2}$$

$$n^\beta |\varphi(n, u_n) - A_n(u_n)| \leq B(u_n), \tag{8.3.3}$$

for all n sufficiently large. If $t > 2J$ and $t > \max_{p \in P_B \cup P_A} \sum_{j=1}^{J} jp_j$, then $n^\beta E|\varphi(n, u_n) - A_n(u_n)| \to 0$, and consequently, $E[\varphi(n, u_n)] = E[A_n(u_n)] + o(n^{-\beta})$.

Assumption (8.3.2) of this theorem is quite strong and limits its applicability in very important situations.

8.4 Delta Method for exponentially bounded functions

In this section, we present Delta method results for a class of functions which might grow faster than a polynomial function but can be bounded by an exponential function.

In [84], Khan applies stronger conditions on the random variables than those in [36] and [92] in order to obtain a Delta method theorem that applies to a larger family of functions. Consider the i.i.d. random variables $X_1, X_2, \cdots, X_n$ with mean μ, variance σ^2 and $\bar{X}_n = \sum_{i=1}^{n} X_i/n$. Let $\mathcal{A} \subset \mathbb{R}$ be an interval such that $P(X_1 \in \mathcal{A}) = 1$. Define $\mathcal{F}$ as the class of functions, continuous on $\mathcal{A}$, such that $\varphi \in \mathcal{F}$ implies

$$|\varphi(x)| = O(e^{\alpha|x|}) \text{ as } |x| \to \infty, \text{ for some } \alpha > 0. \tag{8.4.1}$$

It follows that bounded functions and polynomially bounded functions belong to $\mathcal{F}$.

Theorem 8.7 (Theorem 1 in [84], Khan) *Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with mean μ and variance σ^2, and assume that X_1 has a finite moment generating function (m.g.f.). Let φ be a continuous function on $\mathcal{A}$ with $\varphi \in \mathcal{F}$ where $\mathcal{A}$ is an interval such that $P(X_1 \in \mathcal{A}) = 1$. Suppose that the first four derivatives of f are continuous in $(\mu - \delta, \mu + \delta)$ for some $\delta > 0$. Then*

$$\begin{aligned} E[\varphi(\bar{X}_n)] &= \varphi(\mu) + \frac{\sigma^2}{2n}\varphi''(\mu) + O(n^{-2}), \\ var(\varphi(\bar{X}_n)) &= \frac{\sigma^2}{n}(\varphi'(\mu))^2 + O(n^{-2}). \end{aligned}$$

The following two Lemmas are required for the proof of the theorem.

Lemma 8.8 (Chernoff) *Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with mean μ, and assume X_1 has a finite m.g.f. $\phi(\theta)$ for $\theta \in J$ containing zero. Then, for any $\delta > 0$, there exist numbers ρ and ρ_1 $(0 < \rho, \rho_1 < 1)$ such that*

$$P(\bar{X}_n - \mu \geq \delta) \leq \rho_1^n, \quad P(|\bar{X}_n - \mu| \geq \delta) \leq 2\rho^n.$$

Lemma 8.9 (Khan) *Let $\varphi \in \mathcal{F}$,and let $E|\varphi(\bar{X}_n)| < \infty$. Then under the conditions of Lemma* 8.8

$$E[\varphi(\bar{X}_n)]I\{|\bar{X}_n - \mu| \geq \delta\} = O(1)(\rho^n + \rho_1^n) = O(n^{-2}).$$

We now present the proof of the theorem.

Proof. (Theorem 8.7) Let $Q(x) = \sum_{k=0}^{4}((x-\mu)^k/k!)\varphi^{(k)}(\mu)$ be the Taylor polynomial, and consider the Taylor expansion of φ in $(\mu - \delta, \mu + \delta)$ as

$$\begin{aligned} \varphi(x) &= Q(x) + \frac{(x-\mu)^4}{4!}(\varphi^{(4)}(\mu + \eta(x-\mu)) - \varphi^{(4)}(\mu)) \\ &= Q(x) + R(x), \quad 0 \leq \eta \leq 1. \end{aligned}$$

It is well known that $E[(\bar{X}_n - \mu)^3] = O(n^{-2})$, $\quad E[(\bar{X}_n - \mu)^4] = O(n^{-2})$, and therefore it follows $E[Q(\bar{X}_n)] = \varphi(\mu) + \frac{\sigma^2}{2n}\varphi''(\mu) + O(n^{-2})$. Let $0 < \delta_1 < \delta$ and set $T_n = E[\varphi(\bar{X}_n)]I\{|\bar{X}_n - \mu| < \delta_1\}$ By lemma 8.9 we have $E[\varphi(\bar{X}_n)] = T_n + E[\varphi(\bar{X}_n)I\{|\bar{X}_n - \theta| \geq$

$\delta_1\}] = T_n + O(n^{-2})$. Clearly

$$T_n = E[Q(\bar{X}_n)I\{|\bar{X}_n - \theta| < \delta_1\}] + E[R(\bar{X}_n)I\{|\bar{X}_n - \theta| < \delta_1\}]$$
$$= E[Q(\bar{X}_n)] - E[Q(\bar{X}_n)I\{|\bar{X}_n - \theta| \geq \delta_1\}] + E[R(\bar{X}_n)I\{|\bar{X}_n - \theta| < \delta_1\}]$$

Since $Q(x) \in \mathcal{F}$, by lemma 8.9 we have

$$T_n = E[Q(\bar{X}_n)] + O(n^{-2}) + E[R(\bar{X}_n)I\{|\bar{X}_n - \theta| < \delta_1\}].$$

Now consider the remainder term. Let $Z_n = \mu + \eta(\bar{X}_n - \mu)$, $0 \leq \eta = \eta_n \leq 1$. $|\bar{X}_n - \mu| < \delta_1$ implies $|Z_n - \mu| < \delta_1$, and Z_n is in the closed interval $[\mu - \delta_1, \mu + \delta_1]$. Since $\varphi^{(4)}(x)$ is continuous in $[\mu - \delta_1, \mu + \delta_1]$, hence $(1/4!)|\varphi^{(4)}(Z_n) - \varphi^{(4)}(\mu)|$ remains bounded by some constant K. Thus we have

$$E[R(\bar{X}_n)I\{|\bar{X}_n - \theta| < \delta_1\}] \leq KE[(\bar{X}_n - \mu)^4] = O(n^{-2}).$$

■

Khan's theorem has many weaknesses. To begin with, the theorem only applies to functions with one argument consisting of a sample mean of a random sample. The proof is not general enough to be extended to functions of other sample statistics or functions with more general dependence on several random variables. This weakness can be traced to lemma 8.9 which is an application of the Law of Large Numbers.

The following theorem extends the work of Khan [84] and Hurt [76]. The theorem replaces the need for finite m.g.f.'s with a more general condition. The statistic S_n is allowed to be arbitrary as opposed to being the sample mean of r.v.'s $X_1, \ldots, X_n$. The condition in [76] that φ must be bounded is relaxed to the condition given in [84] for bounding φ with an exponential function. We consider a continuous function $\varphi(S_n) : \mathcal{A} \subset \mathbb{R} \to \mathbb{R}$. S_n is a one dimensional statistic with $P(S_n \in \mathcal{A}) = 1$ and which itself can be a function of n random variables $X_1, \ldots, X_n$, i.e., $S_n(X_1, \ldots, X_n)$. $\varphi \in \mathcal{F}_\alpha$ implies condition (8.4.1) holds. First, given some assumptions, we prove a lemma.

Assumption 8.1 *$1/p_1 + 1/p_2 = 1$ with $p_1 > 1$ $p_2 > 1$.*

Assumption 8.2 *$\{S_n\}$ is a sequence of one dimensional statistics with finite moments up to order $p_2(q+1)$.*

Assumption 8.3 $S_n \xrightarrow{P} \theta$.

Assumption 8.4 $E|S_n - \theta|^{p_2(q+1)} = O(n^{-(q+1)})$.

Assumption 8.5 $E\exp(p_1\alpha|S_n|) < \infty$.

Assumption 8.6 *$\varphi \in \mathcal{F}_\alpha$ and $E|\varphi(S_n)| < \infty$.*

Assumption 8.1 is the condition required by Hölder's inequality, which is used in the lemma to follow. Parameters p_1, p_2 make the result of the lemma and the theorem more general than the results in [84]. In fact, there is no reason to use Schwarz inequality instead of the more general Hölder's inequality in lemma 2 of [84]. Assumption 8.3 has implications regarding the dynamic nature of the process $\{X_\tau\}$. For example, if the statistic S_n is the sample mean of $X_1, \ldots, X_n$, assumption 8.3 implies the r.v.'s of the process $\{X_\tau\}$ must be identically distributed, (see Chapter 3 in [153]). Assumption 8.5 is weaker than the assumption of finite m.g.f's used in [84]. Assumption 8.6 characterizes the growth nature of the function $\varphi(x)$ as $|x| \to \infty$ and establishes the existence of the expected value we attempt to approximate.

Lemma 8.10 (Martinez) *Under assumptions* 8.1 *through* 8.6

$$E[\varphi(S_n)I\{|S_n - \theta| \geq \delta\}] = O(n^{-(q+1)/p_2}).$$

Proof. $\varphi \in \mathcal{F}_\alpha$ implies there exists a finite N and a constant C, both independent of x, such that $|\varphi(x)| \leq C\exp(\alpha|x|)\ \forall\ x$ with $|x-\theta| \geq N$. Let $B(x) = \{x : \delta \leq |x-\theta| \leq N\}$ and $\bar{B}(x) = \{x : |x-\theta| > N\}$ where $I\{\cdot\}$ is the indicator function. It follows

$$E[\varphi(S_n)I\{|S_n - \theta| \geq \delta\}] = E[\varphi(S_n)I\{B(S_n)\}] + E[\varphi(S_n)I\{\bar{B}(S_n)\}]. \tag{8.4.2}$$

By continuity of φ, $|\varphi(S_n)| \leq M$ in $B(S_n)$ for some constant M independent of n. By

Markov's inequality and the assumption on the moments of S_n

$$E[|\varphi(S_n)|I\{B(S_n)\}] \leq MP(S_n \in B(S_n)) \leq MP(|S_n - \theta| \geq \delta)$$
$$\leq \frac{M}{\delta^{p_2(q+1)}} E[\,|S_n - \theta|^{p_2(q+1)}] = O(n^{-(q+1)}). \quad (8.4.3)$$

Let us comment on (8.4.3). The exponent involved in Markov's inequality can be set to any finite number. In the above, we set this exponent equal to $p_2(q+1)$. The reason for this lies in the use of this lemma in theorem 8.11. The present lemma is used in the said theorem to bound a Taylor expansion in a neighborhood of θ. In theory, the exponent in Markov's inequality can be set equal to any number greater or equal to $p_2(q+1)$. On $\bar{B}(S_n)$, $|\varphi(S_n)| \leq C\exp(\alpha|S_n|)$, and it follows that

$$E[|\varphi(S_n)|I\{\bar{B}(S_n)\}] \leq CE[\exp(\alpha|S_n|)I\{\bar{B}(S_n)\}].$$

By Hölder's inequality,

$$E[\exp(\alpha|S_n|)I\{\bar{B}(S_n)\}] \leq E^{1/p_1}[\exp(p_1\alpha|S_n|)](P(|S_n - \theta| \geq N))^{1/p_2}.$$

Since $S_n \xrightarrow{P} \theta$ and $\exp(x)$ is continuous on $\mathcal{A}$, it follows that $\exp(p_1\alpha|S_n|) \xrightarrow{P} \exp(p_1\alpha|\theta|)$ (see proposition $A.18$ in Appendix B). Furthermore, assumption 8.5 implies, by the dominated convergence theorem (see Appendix F), $E[\exp(p_1\alpha|S_n|)]$ converges to $E[\exp(p_1\alpha|\theta|)]$ as $n \to \infty$ and therefore $E[\exp(p_1\alpha|S_n|)] = O(1)$. It follows that

$$E[|\varphi(S_n)|I\{\bar{B}(S_n)\}] \leq CE^{1/p_1}[\exp(p_1\alpha|S_n|)](P(|S_n - \theta| \geq N))^{1/p_2}$$
$$= O(1)O(n^{-(q+1)/p_2}). \quad (8.4.4)$$

(8.4.2), (8.4.3), and (8.4.4) give the result. ■

We give one assumption and state the theorem.

Assumption 8.7 : *For some $q > 1$, $\varphi(x)$ has finite and continuous derivatives up to order $q+1$ in $(\theta - \delta, \theta + \delta)$ for some $\delta > 0$.*

Assumption 8.7 is needed to write the Taylor expansion of φ in the interval $(\theta-\delta, \theta+\delta)$

in the following theorem.

Theorem 8.11 (Martinez) *Under assumptions* 8.1 *through* 8.7

$$E[\varphi(S_n)] = \varphi(\theta) + \sum_{j=1}^{q} \frac{1}{j!} \left(\frac{\partial^j \varphi}{\partial s^j} \right)_{s=\theta} E[(S_n - \theta)^j] + O(n^{-(q+1)/p_2}).$$

Proof. Let $\varphi^{(k)}(x)$ denote the kth derivative of φ with respect to x. Let $Q_q(x) = \sum_{k=0}^{q} \varphi^{(k)}(\theta)(x-\theta)^k/k!$. The Taylor expansion of φ in $(\theta - \delta, \theta + \delta)$ is

$$\begin{aligned} \varphi(x) = Q_q(x) + \frac{1}{(q+1)!}\varphi^{(q+1)}(\theta + \eta(x-\theta))(x-\theta)^{q+1} \\ = Q_q(x) + R_q(x), \quad 0 \le \eta \le 1. \quad (8.4.5) \end{aligned}$$

It follows

$$E[Q_q(S_n)] = \varphi(\theta) + \sum_{k=1}^{q} \frac{1}{k!} \varphi^{(k)}(\theta) E[(S_n - \theta)^k]. \tag{8.4.6}$$

Let $0 < \delta_1 < \delta$ and set $T_n = E[\varphi(S_n) I\{|S_n - \theta| < \delta_1\}]$. By lemma 8.10 it follows

$$E[\varphi(S_n)] = T_n + E[\varphi(S_n) I\{|S_n - \theta| \ge \delta_1\}] = T_n + O(n^{-(q+1)/p_2}). \tag{8.4.7}$$

One can write T_n as follows:

$$\begin{aligned} T_n &= E[Q_q(S_n) I\{|S_n - \theta| < \delta_1\}] + E[R_q(S_n) I\{|S_n - \theta| < \delta_1\}] \\ &= E[Q_q(S_n)] - E[Q_q(S_n) I\{|S_n - \theta| \ge \delta_1\}] + E[R_q(S_n) I\{|S_n - \theta| < \delta_1\}]. \end{aligned}$$

Since $Q_q(x) \in \mathcal{F}_\alpha$, by lemma 8.10 it follows

$$T_n = E[Q_q(S_n)] + O(n^{-(q+1)/p_2}) + E[R_q(S_n) I\{|S_n - \theta| < \delta_1\}]. \tag{8.4.8}$$

To understand the order of the remainder term we first note

$$E[R_q(S_n) I\{|S_n - \theta| < \delta_1\}] = E[R_q(S_n) I\{|S_n - \theta| \le \delta_1\}],$$

and $Z_n = \theta + \eta(S_n - \theta) \in [\theta - \delta_1, \theta + \delta_1]$ for $0 \le \eta \le 1$. It follows, since $\varphi^{(q+1)}$ is continuous in $[\theta - \delta_1, \theta + \delta_1]$, $\varphi^{(q+1)}(Z_n)$ is bounded and we have

$$E[\,|R_q(S_n)|I\{|S_n - \theta| < \delta_1\}] \le KE[\,|S_n - \theta|^{(q+1)}] = O(n^{-(q+1)/p_2}). \tag{8.4.9}$$

The result follows from (8.4.6), (8.4.7), (8.4.8) and (8.4.9). ■

The previous theorems apply to functions, $\varphi(x)$, which become unbounded as $|x| \to \infty$ but do not apply to functions which become unbounded at a finite point in $\mathcal{A}$. Next, we consider the case of a function $\varphi(x) : \mathcal{A} \subset \mathbb{R} \to \mathbb{R}$ with an essential discontinuity at a point x_0. Define subintervals $A_1(x) = \{x \in \mathbb{R} : |x-\theta| \le \delta\}$, $A_2(x) = \{x \in \mathbb{R} : |x-x_0| \le \delta_1\}$ and $A_3(x) = (A_1(x) \cup A_2(x))^c$ for $\delta > 0, \delta_1 > 0$ such that $x_0 + \delta_1 = \theta - \delta$ and $\mathcal{A} = A_1 \cup A_2 \cup A_3$. Let $\mathcal{G}_{(\alpha,\beta)}$ denote the class of functions on $\mathcal{A}$ such that

$$\begin{aligned}
&|\varphi(x)| = O(e^{\alpha|x|}), \quad \text{as} \quad |x| \to \infty \quad \text{and} \\
&|\varphi(x)| = O(e^{\beta/|x-x_0|}), \quad \text{as} \quad x \to x_0, \quad \text{for some} \quad \alpha, \beta > 0.
\end{aligned}$$

As before, S_n is a one dimensional statistic with $P(S_n \in \mathcal{A}) = 1$. We give some assumptions and prove two lemmas.

Assumption 8.8 $\theta \ne x_0$ *and* $E[\exp(p_1\beta/|S_n - x_0|)] < \infty$.

Assumption 8.9 $\varphi \in \mathcal{G}_{(\alpha,\beta)}$ *and* $E|\varphi(S_n)| < \infty$.

The following lemma is similar to lemma 8.10 except that assumption 8.6 is replaced by assumption 8.9.

Lemma 8.12 (Martinez) *Under assumptions* 8.1, 8.2, 8.3, 8.4, 8.5, *and* 8.9

$$E[\varphi(S_n)I\{S_n \in A_3\}] = O(n^{-(q+1)/p_2})$$

Proof. The proof of this lemma follows similarly as the proof of lemma 8.10. Without loss of generality we take $x_0 = 0$, $\theta > 0$. $\varphi \in \mathcal{G}_{\alpha,\beta}$ implies there exists a finite C and a constant N, both independent of x, such that $|\varphi(x)| \le C\exp(\alpha|x|)\ \forall x$ with $|x - \theta| \ge N$. Let $B_1(x) = \{x \in \mathbb{R} : \delta < x - \theta \le N\}$, $B_2(x) = \{x \in \mathbb{R} : \theta - N \le x \le -\delta_1\}$ and

$\bar{B}(x) = \{x : |x - \theta| > N\}$. It follows

$$
\begin{aligned}
E[\varphi(S_n)I\{S_n \in A_3\}] =& E[\varphi(S_n)I\{S_n \in B_1(S_n)\}] + E[\varphi(S_n)I\{S_n \in B_2(S_n)\}] \\
&+ E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}].
\end{aligned}
$$

Following the same arguments of lemma 8.10, $E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}] = O(n^{-(q+1)/p_2})$. By continuity of φ in $A_1 \cup A_3$, $|\varphi(S_n)| \le M$ on $B_1(S_n)$ and $B_2(S_n)$ for some constant M independent of n and

$$
\begin{aligned}
E[\varphi(S_n)I\{S_n \in B_1(S_n)\}] &\le MP(S_n \in B_1(S_n)) \le MP(|S_n - \theta| \ge \delta) \\
&\le \frac{M}{\delta^{p_2(q+1)}} E[|S_n - \theta|^{p_2(q+1)}] = O(n^{-(q+1)}).
\end{aligned}
$$

Similarly, $E[\varphi(S_n)I\{S_n \in B_2(S_n)\}] = O(n^{-(q+1)})$ and the result follows. ■

Lemma 8.13 (Martinez) *Under assumptions* 8.1, 8.2, 8.3, 8.4, 8.8 *and* 8.9

$$
E[\varphi(S_n)I\{|S_n \in A_2\}] = O(n^{-(q+1)/p_2})
$$

Proof. Without loss of generality we take $x_0 = 0$, $\theta > 0$. Since $\varphi \in \mathcal{G}_{\alpha,\beta}$, $\exists$ a finite C and a $\beta > 0$ such that $|\varphi(S_n)| \le C\exp(\beta/|S_n|)$ on A_2 and we have

$$
\begin{aligned}
E[|\varphi(S_n)|I\{S_n \in A_2\}] &\le CE[\exp(\beta/|S_n|)I\{S_n \in A_2\}] \\
&\le CE^{1/p_1}[\exp(p_1\beta/|S_n|)](P(I\{S_n \in A_2\}))^{1/p_2} \\
&\le CE^{1/p_1}[\exp(p_1\beta/|S_n|)](P(I\{|S_n - \theta| > \delta\}))^{1/p_2},
\end{aligned}
$$

where the second inequality follows from Hölder's inequality. Since $S_n \xrightarrow{P} \theta$ and the expression $\exp(p_1\beta/|x|)$ is continuous at θ, it follows by proposition $A.18$, $\exp(p_1\beta/|S_n|) \xrightarrow{P} \exp(p_1\beta/|\theta|)$. Furthermore, assumption 8.8 implies, by the dominated convergence theorem, that $E[\exp(p_1\beta/|S_n|)] \to E[\exp(p_1\beta/|\theta|)]$ as $n \to \infty$ and therefore $E[\exp(p_1\beta/|S_n|)] = O(1)$. It follows that

$$
E[|\varphi(S_n)|I\{S_n \in A_2\}] \le O(1)\left(\frac{E[|S_n - \theta|^{p_2(q+1)}}{\delta^{p_2(q+1)}}\right)^{1/p_2} = O(n^{-(q+1)/p_2}), \quad (8.4.10)
$$

and (8.4.10) gives the result. ■

Theorem 8.14 (Martinez) *Under assumptions* 8.1, 8.2, 8.3, 8.4, 8.5, 8.7, 8.8, 8.9

$$E[\varphi(S_n)] = \varphi(\theta) + \sum_{j=1}^{q} \frac{1}{j!} \left(\frac{\partial^j \varphi}{\partial s^j} \right)_{s=\theta} E[(S_n - \theta)^j] + O(n^{-(q+1)/p_2}).$$

Proof. As before, let $Q_q(x) = \sum_{k=0}^{q} \varphi^{(k)}(\theta)(x-\theta)^k/k!$. The Taylor expansion of φ in $(\theta-\delta, \theta+\delta)$ is given by (8.4.5) and the expected value of $Q_q(S_n)$ is given by (8.4.6). The expected value of $\varphi(S_n)$ can be written as follows

$$E[\varphi(S_n)] = E[\varphi(S_n)I\{S_n \in A_1\}] + E[\varphi(S_n)I\{S_n \in A_2\}] + E[\varphi(S_n)I\{S_n \in A_3\}]. \quad (8.4.11)$$

From lemma 8.12, it follows $E[\varphi(S_n)I\{S_n \in A_3\}] = O(n^{-(q+1)/p_2})$. Denote $T_n = E[\varphi(S_n)I\{S_n \in A_1\}]$ and

$$\begin{aligned} T_n &= E[Q_q(S_n)I\{S_n \in A_1\}] + E[R_q(S_n)I\{S_n \in A_1\}] \\ &= E[Q_q(S_n)] - E[Q_q(S_n)I\{S_n \in A_2\}] \\ &\quad - E[Q_q(S_n)I\{S_n \in A_3\}] + E[R_q(S_n)I\{S_n \in A_1\}]. \end{aligned}$$

Since $Q_q(S_n) \in \mathcal{G}_{(\alpha,\beta)}$, by lemma 8.12, lemma 8.13 and (8.4.9) $T_n = E[Q_q(S_n)] + O(n^{-(q+1)/p_2})$. (8.4.11) becomes $E[\varphi(S_n)] = E[Q_q(S_n)] + E[\varphi(S_n)I\{S_n \in A_2\}] + O(n^{-(q+1)/p_2})$ and the result follows by applying lemma 8.13 again . ■

The multivariate version of the previous theorem can be formulated by considering a function $\varphi(x) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}$ with an essential discontinuity at a point $x_0 \equiv (x_{10}, \cdots, x_{r0})$. Using the Euclidean norm $||\cdot||_2$, we define subsets $A_1(x) = \{x \in \mathbb{R}^r : ||x-\theta||_2 \leq \delta\}$, $A_2(x) = \{x \in \mathbb{R}^r : ||x - x_0||_2 < \delta_1\}$ and $A_3(x) = (A_1(x) \cup A_2(x))^c$ with $\delta > 0$, $\delta_1 > 0$, $||\theta - x_0||_2 = \delta_1 + \delta$ such that $\mathcal{A} = A_1 \cup A_2 \cup A_3$. $\mathcal{H}_{\alpha,\beta}$ denotes the class of functions on $\mathcal{A}$ such that $|\varphi(x)| = O(e^{\alpha||x||})$ as $||x|| \to \infty$ and $|\varphi(x)| = O(e^{\beta/||x-x_0||})$ as $||x - x_0|| \to 0$ for some $\alpha, \beta > 0$, where $||\cdot||$ is the one norm $||x|| = \sum_{i=1}^{n} |x_i|$. S_{in} is a one dimensional statistic for $i = 1, \ldots, r$ with $P((S_{1n}, \cdots, S_{rn}) \in \mathcal{A}) = 1$.

Assumption 8.10 *$\{S_n \equiv (S_{1n}, \cdots, S_{rn})\}$ is a sequence of multidimensional statistics with finite absolute moments of S_{in} up to order $p_2(q+1)$.*

Assumption 8.11 *$S_{in} \xrightarrow{P} \theta_i$ for $i = 1, \ldots, r$.*

Assumption 8.12 *$E|S_{in} - \theta_i|^{p_2(q+1)} = O(n^{-(q+1)})$ for $i = 1, \ldots, r$.*

Assumption 8.13 *$E[\exp(p_1\alpha||S_n||)] < \infty$.*

Assumption 8.14 *$\theta \neq x_0$ and $E[\exp(p_1\beta/||S_n - x_0||)] < \infty$.*

Assumption 8.15 *$\varphi \in \mathcal{G}_{\alpha,\beta}$ and $E|\varphi(S_n)| < \infty$.*

Assumption 8.16 *φ has finite and continuous partial derivatives up to order $q+1$ in A_1.*

We note, for the case of a multivariate function $\varphi(x) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}^s$, it is sufficient to check the assumptions above for each $\varphi_i(x) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}$, $i = 1, \ldots, s$ where $\varphi(x) \equiv (\varphi_1(x), \ldots, \varphi_s(x))^\top$.

Lemma 8.15 (Martinez) *Under assumptions 8.1, 8.10,8.11,8.12,8.13 and 8.15*

$$E[\varphi(S_n)I\{S_n \in A_3(S_n)\}] = O(n^{-(q+1)/p_2})$$

Proof. $\varphi \in \mathcal{H}_{\alpha,\beta}$ implies $\exists$ a finite N, $N > \delta + 2\delta_1$, and a constant C, both independent of x, such that $|\varphi(x)| \leq C\exp(\alpha||x||)$ $\forall x$ with $||x - \theta||_2 > N$. Let $\bar{B}(x) = \{x : ||x - \theta||_2 > N\}$ and $B(x) = \{x : \bar{B}^c - A_1 - A_2\}$. It follows

$$E[\varphi(S_n)I\{S_n \in A_3(S_n)\}] = E[\varphi(S_n)I\{S_n \in B(S_n)\}] + E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}].$$

By continuity of φ on $A_3(x)$, $|\varphi(S_n)| \leq M$ on $B(S_n)$ for some constant M independent of n. It follows by Markov's inequality

$$\begin{aligned} E[\varphi(S_n)I\{S_n \in B(S_n)\}] &\leq MP(S_n \in B(S_n)) \leq MP(||S_n - \theta||_2 \geq \delta) \\ &\leq \frac{M}{\delta^{p_2(q+1)}} E[\,||S_n - \theta||_2^{p_2(q+1)}]. \end{aligned}$$

Furthermore,

$$E[||S_n - \theta||_2^{p_2(q+1)}] = E\Big[\Big(\sum_{i=1}^{r}(S_{in} - \theta_i)^2\Big)^{p_2(q+1)/2}\Big] \leq E\Big[\Big(\sum_{i=1}^{r}|S_{in} - \theta_i|\Big)^{p_2(q+1)}\Big]$$
$$\leq \Big\{\sum_{i=1}^{r}\Big(E^{1/p_2(q+1)}[\,|S_{in} - \theta_i|^{p_2(q+1)}]\Big)\Big\}^{p_2(q+1)}$$
$$= \Big\{\sum_{i=1}^{r}\Big(O(n^{-(q+1)})\Big)^{1/p_2(q+1)}\Big\}^{p_2(q+1)} = O(n^{-(q+1)}), \quad (8.4.12)$$

where the second inequality is due to Minkowski's inequality and the second equality follows from the assumption on the moments of S_n. On $\bar{B}(S_n)$, $|\varphi(S_n)| \leq C\exp(\alpha||S_n||)$ and it follows

$$E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}] \leq CE[\exp(\alpha||S_n||)I\{S_n \in \bar{B}(S_n)\}]$$
$$\leq CE^{1/p_1}[\exp(p_1\alpha||S_n||)]P(||S_n - \theta||_2 > N))^{1/p_2},$$

where the second inequality follows by Hölder's inequality. Since $S_{in} \xrightarrow{P} \theta_i$ and $\exp(x)$ is continuous on $\mathcal{A}$, it follows $\exp(p_1\alpha|S_{in}|) \xrightarrow{P} \exp(p_1\alpha|\theta_i|)$ and $\exp(p_1\alpha||S_n||) \xrightarrow{P} \exp(p_1\alpha||\theta||)$. Furthermore, the assumption $E|\exp(p_1\alpha||S_n||)| < \infty$ implies, by the dominated convergence theorem, that $E[\exp(p_1\alpha||S_n||)] \to E[\exp(p_1\alpha||\theta||)]$ as $n \to \infty$ and therefore $E[\exp(p_1\alpha||S_n||)] = O(1)$. It follows

$$E[|\varphi(S_n)|I\{\bar{B}(S_n)\}] \leq CE^{1/p_1}[\exp(p_1\alpha||S_n||)](P(|S_n - \theta| > N))^{1/p_2}$$
$$= O(1)O(n^{-(q+1)/p_2}). \quad (8.4.13)$$

The result follows from 8.4.12 and 8.4.13. ■

Lemma 8.16 (Martinez) *Under assumptions 8.1,8.10,8.11,8.12, 8.14 and 8.15*

$$E[\varphi(S_n)I\{S_n \in A_2(S_n)\}] = O(n^{-(q+1)/p_2})$$

Proof. Since $\varphi \in \mathcal{H}_{\alpha,\beta}$, $\exists$ a finite C such that $|\varphi(S_n)| \leq C\exp(\beta/||S_n||)$ on A_2 and we have

$$\begin{aligned}E[|\varphi(S_n)|I\{S_n \in A_2\}] &\leq CE[\exp(\beta/||S_n||)I\{S_n \in A_2\}]\\ &\leq CE^{1/p_1}[\exp(p_1\beta/||S_n||)](P(I\{S_n \in A_2\}))^{1/p_2}\\ &\leq CE^{1/p_1}[\exp(p_1\beta/||S_n||)](P(I\{||S_n-\theta||_2 \geq \delta\}))^{1/p_2},\end{aligned}$$

where the second inequality follows from Hölder's inequality. Since $S_{in} \xrightarrow{P} \theta_i$, $||S_n|| \xrightarrow{P} ||\theta||$. By continuity of the expression $\exp(p_1\beta/||x||)$ at θ, it follows by proposition $A.18$, $\exp(p_1\beta/||S_n||) \xrightarrow{P} \exp(p_1\beta/||\theta||)$. Furthermore, assumption 8.14 implies, by the dominated convergence theorem, $E[\exp(p_1\beta/||S_n||)] \to E[\exp(p_1\beta/||\theta||)]$ as $n \to \infty$ and therefore $E[\exp(p_1\beta/||S_n||)] = O(1)$. It follows

$$\begin{aligned}E[|\varphi(S_n)|I\{S_n \in A_2\}] \leq CE^{1/p_1}[\exp(p_1\beta/||S_n||)]\left(\frac{E[\,||S_n-\theta||_2^{p_2(q+1)}]}{\delta^{p_2(q+1)}}\right)^{1/p_2}&\\ = O(1)O(n^{-(q+1)/p_2}).& \quad (8.4.14)\end{aligned}$$

and (8.4.14) gives the result. ■

Theorem 8.17 (Martinez) *Given $\varphi(S_{1n}, \cdots, S_{rn}) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}$, under assumptions* 8.1 *and* 8.10 *through* 8.16, *it follows*

$$\begin{aligned}&E[\varphi(S_{1n}, \cdots, S_{rn})] = \varphi(\theta_1, \cdots, \theta_r)+\\ &\sum_{j=1}^{q}\frac{1}{j!}\sum_{i_1}\cdots\sum_{i_r}\left[\frac{\partial^j\varphi}{\partial s_1^{i_1}\ldots\partial s_r^{i_r}}\right]_{s=\theta} E[(S_{1n}-\theta_1)^{i_1}\ldots(S_{rn}-\theta_r)^{ir}] + O(n^{-(q+1)/p_2}),\end{aligned}$$

with $i_1 + \cdots + i_r = j$, $s = (s_1, \cdots, s_r)$ and $\theta = (\theta_1, \cdots, \theta_r)$.

Proof. The proof is a generalization of the one dimensional theorem. The multivariate Taylor expansion of $\varphi(S_n)$ in A_1 is given by $\varphi(S_n) = Q_q(S_n) + R_q(S_n)$, where

$$\begin{aligned}Q_q(S_n) = \varphi(\theta_1, \cdots, \theta_r)+&\\ \sum_{j=1}^{q}\frac{1}{j!}\sum_{i_1}\cdots\sum_{i_r}&\left[\frac{\partial^j\varphi}{\partial s_1^{i_1}\ldots\partial s_r^{i_r}}\right]_{s=\theta}(S_{1n}-\theta_1)^{i_1}\ldots(S_{rn}-\theta_r)^{ir},\end{aligned}$$

for $i_1 + \cdots + i_r = j$ and

$$R_q(S_n) = \frac{1}{(q+1)!} \sum_{i_1} \cdots \sum_{i_r} \left[\frac{\partial^j \varphi}{\partial s_1^{i_1} \ldots \partial s_r^{i_r}} \right]_{s=\theta+\eta(S_n-\theta)} (S_{1n} - \theta_1)^{i_1} \ldots (S_{rn} - \theta_r)^{ir},$$

for $i_1 + \cdots + i_r = q+1$, $0 \le \eta \le 1$. The expected value of $\varphi(S_n)$ can be written as follows

$$E[\varphi(S_n)] = E[\varphi(S_n) I\{S_n \in A_1\}] + E[\varphi(S_n) I\{S_n \in A_2\}] + E[\varphi(S_n) I\{S_n \in A_3\}]. \quad (8.4.15)$$

From lemmas 8.15 and 8.16, $E[\varphi(S_n) I\{S_n \in A_2\}] = O(n^{-(q+1)/p_2})$ and $E[\varphi(S_n) I\{S_n \in A_3\}] = O(n^{-(q+1)/p_2})$, respectively. Denote $T_n = E[\varphi(S_n) I\{S_n \in A_1\}]$ and

$$\begin{aligned} T_n =& E[Q_q(S_n) I\{S_n \in A_1\}] + E[R_q(S_n) I\{S_n \in A_1\}] \\ =& E[Q_q(S_n)] - E[Q_q(S_n) I\{S_n \in A_2\}] \\ &- E[Q_q(S_n) I\{S_n \in A_3\}] + E[R_q(S_n) I\{S_n \in A_1\}]. \end{aligned}$$

Given $Z_n = \theta + \eta(S_n - \theta) \in A_1$, and since all partial and total derivatives of order $q+1$ are continuous and bounded,

$$\begin{aligned} E[|R_q(S_n)| I\{S_n \in A_1\}] &\le K E[|(S_{1n} - \theta)^{i_1} \cdots (S_{rn} - \theta)^{i_r}|] \le K E[|S_{1n} - \theta|^{i_1} \cdots |S_{rn} - \theta|^{i_r}] \\ &\le K \Big\{ [E|S_{1n} - \theta|^{q+1}]^{i_1} \cdots [E|S_{rn} - \theta|^{q+1}]^{i_r} \Big\}^{1/(q+1)} = O(n^{-(q+1)/p_2}), \quad (8.4.16) \end{aligned}$$

where the third inequality follows from lemma $F.2$. By lemma 8.15, lemma 8.16 and (8.4.16), $T_n = E[Q_q(S_n)] + O(n^{-(q+1)/p_2})$, and (8.4.15) becomes $E[\varphi(S_n)] = E[Q_q(S_n)] + O(n^{-(q+1)/p_2})$. ■

We next consider a rational function $\varphi(x) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}$ of the form $\varphi(x) = Q_1(x)/Q_2(x)$ where Q_1 and Q_2 are polynomials. Using the Euclidean norm $|| \cdot ||_2$, we define subsets $A_1(x) = \{x \in \mathbb{R}^r : ||x - \theta||_2 \le \delta\}$, $A_2(x) = \{x \in \mathbb{R}^r : ||Q_2(x)||_2 < \delta_1\}$ and $A_3(x) = (A_1(x) \cup A_2(x))^c$ with $\delta > 0$, $\delta_1 > 0$ such that $A_1 \cap A_2 = \emptyset$, $\mathcal{A} = A_1 \cup A_2 \cup A_3$.

$\mathcal{J}_{\alpha,\beta}$ denotes the class of rational functions on $\mathcal{A}$ such that

$$|\varphi(x)| = O(e^{\alpha||x||}) \quad \text{as} \quad ||x|| \to \infty.$$
$$|\varphi(x)| = O(e^{\beta/||Q_2(x)||}) \quad \text{as} \quad ||Q_2(x)|| \to 0.$$

S_{in} is a one dimensional statistic for $i = 1, \ldots, r$ and $P(S_n \equiv (S_{1n}, \cdots, S_{rn}) \in \mathcal{A}) = 1$.

Assumption 8.17 *$E[\exp(p_1\alpha||S_n||)] < \infty$ and $E[\exp(p_1\beta/||Q_2(S_n)||)] < \infty$.*

Assumption 8.18 *$\varphi \in \mathcal{J}_{\alpha,\beta}$ and $E|\varphi(S_n)| < \infty$.*

Lemma 8.18 (Martinez) *Under assumptions 8.1, 8.10,8.11,8.12,8.17 and 8.18*

$$E[\varphi(S_n)I\{S_n \in A_3(S_n)\}] = O(n^{-(q+1)/p_2}).$$

Proof. $\varphi \in \mathcal{J}_{\alpha,\beta}$ implies there exists a constant C, independent of x, such that $|\varphi(x)| \leq C\exp(\alpha||x|| + \beta/||Q_2(x)||)$ $\forall x \in A_3$. Let $N > \delta$ and define sets $B(x) = \{x : ||x - \theta||_2 < N\} - (A_1 \cup A_2)$ and $\bar{B}(x) = \{x : ||x - \theta||_2 \geq N\} - A_2$. It follows

$$E[\varphi(S_n)I\{S_n \in A_3(S_n)\}] = E[\varphi(S_n)I\{S_n \in B(S_n)\}] + E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}].$$

By continuity of φ on $A_3(x)$, $|\varphi(S_n)| \leq M$ on $B(S_n)$ for some constant M independent of n. It follows by Markov's inequality

$$\begin{aligned} E[\varphi(S_n)I\{S_n \in B(S_n)\}] &\leq MP(S_n \in B(S_n)) \leq MP(||S_n - \theta||_2 \geq \delta) \\ &\leq \frac{M}{\delta^{p_2(q+1)}} E[\, ||S_n - \theta||_2^{p_2(q+1)}] = O(n^{-(q+1)}), \quad (8.4.17) \end{aligned}$$

where (8.4.12) is used in the last equality. On $\bar{B}(S_n)$, $|\varphi(S_n)| \leq C\exp(\alpha||S_n|| + \beta/||Q_2(x)||)$ and it follows

$$\begin{aligned} E[\varphi(S_n)I\{S_n \in \bar{B}(S_n)\}] &\leq CE[\exp(\alpha||S_n|| + \beta/||Q_2(S_n)||)I\{S_n \in \bar{B}(S_n)\}] \\ &\leq CE^{1/p_1}[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)]P(||S_n - \theta||_2 > N))^{1/p_2}, \end{aligned}$$

where the second inequality is by Hölder's inequality. Since $S_{in} \xrightarrow{P} \theta_i$, $||S_n|| \xrightarrow{P} ||\theta||$ and

$||Q_2(S_n)|| \xrightarrow{P} ||Q_2(\theta)||$. By continuity of $\exp(p_1\beta/||Q_2(x)||)$ and $\exp(p_1\alpha||x||)$ at θ, it follows by proposition $A.18$, $\exp(p_1\alpha||S_n||+p_1\beta/||Q_2(S_n)||) \xrightarrow{P} \exp(p_1\alpha||\theta||+p_1\beta/||Q_2(\theta)||)$. Furthermore, assumption 8.17 implies, by the dominated convergence theorem,

$$E[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)] \rightarrow E[\exp(p_1\alpha||\theta|| + p_1\beta/||Q_2(\theta)||)] \quad \text{as} \quad n \rightarrow \infty,$$

and $E[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)] = O(1)$. It follows

$$E[|\varphi(S_n)|I\{\bar{B}(S_n)\}] \leq O(1)(P(|S_n - \theta| > N))^{1/p_2} = O(n^{-(q+1)/p_2}). \tag{8.4.18}$$

The result follows from 8.4.17 and 8.4.18. ■

Lemma 8.19 (Martinez) *Under assumptions 8.1,8.10,8.11,8.12, 8.17 and 8.18*

$$E[\varphi(S_n)I\{S_n \in A_2(S_n)\}] = O(n^{-(q+1)/p_2}).$$

Proof. Since $\varphi \in \mathcal{J}_{\alpha,\beta}$, $\exists$ a finite C such that $|\varphi(S_n)| \leq C\exp(\alpha||S_n|| + \beta/||Q_2(S_n)||)$ on A_2 and we have

$$\begin{aligned} E[|\varphi(S_n)|I\{S_n \in A_2\}] &\leq CE[\exp(\alpha||S_n|| + \beta/||Q_2(S_n)||)I\{S_n \in A_2\}] \\ &\leq CE^{1/p_1}[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)](P(I\{S_n \in A_2\}))^{1/p_2} \\ &\leq CE^{1/p_1}[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)](P(I\{||S_n - \theta||_2 \geq \delta\}))^{1/p_2}, \end{aligned}$$

where the second inequality follows from Hölder's inequality. By the same arguments given in lemma 8.18, $E[\exp(p_1\alpha||S_n|| + p_1\beta/||Q_2(S_n)||)] = O(1)$ and it follows

$$E[|\varphi(S_n)|I\{S_n \in A_2\}] \leq O(1)\left(\frac{E[\,||S_n - \theta||_2^{p_2(q+1)}]}{\delta^{p_2(q+1)}}\right)^{1/p_2} = O(n^{-(q+1)/p_2}). \tag{8.4.19}$$

■

Theorem 8.20 (Martinez) *Given*

$$\varphi(S_{1n}, \cdots, S_{rn}) = Q_1(S_{1n}, \cdots, S_{rn})/Q_2(S_{1n}, \cdots, S_{rn}) : \mathcal{A} \subset \mathbb{R}^r \rightarrow \mathbb{R},$$

where Q_1 and Q_2 are polynomials, under assumptions 8.1, 8.10,8.11,8.12,8.16, 8.17 and 8.18, it follows

$$\begin{aligned} E[\varphi(S_{1n}, \cdots, S_{rn})] = \varphi(\theta_1, \cdots, \theta_r) \\ + \sum_{j=1}^{q} \frac{1}{j!} \sum_{i_1} \cdots \sum_{i_r} \left[\frac{\partial^j \varphi}{\partial s_1^{i_1} \dots \partial s_r^{i_r}} \right]_{s=\theta} E[(S_{1n} - \theta_1)^{i_1} \dots (S_{rn} - \theta_r)^{ir}] \\ + O(n^{-(q+1)/p_2}), \end{aligned}$$

with $i_1 + \cdots + i_r = j$, $s = (s_1, \cdots, s_r)$ and $\theta = (\theta_1, \cdots, \theta_r)$.

Proof. The proof follows identical to that of theorem 8.17 with lemmas 8.15 and 8.16 replaced by lemmas 8.18 and 8.19, respectively. ■

The moment conditions given by assumptions 8.4 and 8.12 can be quite restrictive and might not be satisfied, for example, in situations where strong dependencies between the variables exist. The following theorems are versions of theorems 8.17 and 8.20 with the moment conditions removed.

Theorem 8.21 (Martinez) *Given $\varphi(S_{1n}, \cdots, S_{rn}) : \mathcal{A} \subset \mathbb{R}^r \to \mathbb{R}$, under assumptions* 8.1*,* 8.10*,* 8.11*,and* 8.13 *through* 8.16*, it follows*

$$\begin{aligned} E[\varphi(S_{1n}, \cdots, S_{rn})] = \varphi(\theta_1, \cdots, \theta_r) \\ + \sum_{j=1}^{q} \frac{1}{j!} \sum_{i_1} \cdots \sum_{i_r} \left[\frac{\partial^j \varphi}{\partial s_1^{i_1} \dots \partial s_r^{i_r}} \right]_{s=\theta} E[(S_{1n} - \theta_1)^{i_1} \dots (S_{rn} - \theta_r)^{ir}] \\ + O\left(E^{1/p_2} \left[||S_n - \theta||_2^{p_2(q+1)} \right] \right), \end{aligned}$$

with $i_1 + \cdots + i_r = j$, $s = (s_1, \cdots, s_r)$ and $\theta = (\theta_1, \cdots, \theta_r)$.

As can be seen from the statement of the theorem, the price paid for removing the moment conditions from the assumptions is an order condition in the approximation of the expected value which depends non-trivially on central moments of the statistics. In the next chapter, we will take a closer look at these order expressions to understand their dependence on the data size n and the correlation of the statistics.

Chapter 9

Satisfying the conditions of the Delta Method

In Chapter 8, we discussed algorithms for obtaining approximations of the expected value of a function φ of r statistics $S_{1n}, \cdots, S_{rn}$. Under certain conditions, we showed such an approximation consists of the Taylor polynomial of degree q, plus terms of order $O(n^{-(q+1)})$.

Now, consider the scalar case, $k = 1$, of the forecasting problem described in Chapter 6 with processes $\{X_\tau\}$ and $\{Y_\tau\}$. $\Pi_{1,n}$ and $\Pi_{2,n}$ are given by

$$\Pi_{1,n} = Y_{t+1} X_t \Big(\sum_{\tau=t-n}^{t-1} X_\tau^2 \Big)^{-1} \sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau,$$
$$\Pi_{2,n} = \Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2 \Big)^{-1} X_t \sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau \Big]^2.$$

Define the statistics

$$S_{1,n} = \frac{1}{n} \sum_{\tau=t-n}^{t-1} Y_{t+1} X_t Y_{\tau+1} X_\tau, \quad S_{2,n} = \frac{1}{n} \sum_{\tau=t-n}^{t-1} X_\tau^2, \quad S_{3,n} = \frac{1}{n} \sum_{\tau=t-n}^{t-1} X_t Y_{\tau+1} X_\tau. \tag{9.0.1}$$

It follows $\Pi_{1,n} = S_{1,n}/S_{2,n}$ and $\Pi_{2,n} = (S_{3,n}/S_{2,n})^2$. The objective is to apply theorem 8.20 to find approximations to $E[S_{1,n}/S_{2,n}]$ and $E[(S_{3,n}/S_{2,n})^2]$. In the sections to follow, we examine conditions necessary for the Delta method theorems of Chapter 8 to hold.

9.1 Laws of large numbers

In this section we examine conditions on the processes $\{X_i\}$ and $\{Y_i\}$ necessary for assumptions 8.3 and 8.11 of Chapter 8 to hold.

Assumptions 8.3 and 8.11 can be satisfied with appropriate consistency results. For this, we consider the most general framework consisting of dependent heterogeneously distributed observations. To obtain the adequate laws of large numbers, we require conditions on the dependence of a sequence known as mixing conditions. We begin with some definitions.

Definition 9.1 *The Borel σ-field generated by $\{Z_t, t = n, \ldots, n+m\}$, denoted $\mathcal{B}_n^{n+m} = \sigma(Z_n, \ldots, Z_{n+m})$ is the smallest σ-algebra of Ω that includes*

- *all sets of the form $\times_{i=1}^{n-1}\mathbb{R}^q \times_{i=n}^{n+m} B_i \times_{i=n+m+1}^{\infty} \mathbb{R}^q$, where each $B_i \in \mathcal{B}^q$;*
- *the complement A^c of any set A in $\mathcal{B}_n^{n+m}$;*
- *the union $\cup_{i=1}^{\infty} A_i$ of any sequence $\{A_i\}$ in $\mathcal{B}_n^{n+m}$.*

Definition 9.2 *Let $\mathcal{B}_{-\infty}^n \equiv \sigma(\ldots, Z_n)$ be the smallest collection of subsets of Ω that contains the union of the σ-fields $\mathcal{B}_a^n$ as $a \to -\infty$; let $\mathcal{B}_{n+m}^{\infty} = \sigma(Z_{n+m}, \ldots)$ be the smallest collection of subsets of Ω that contains the union of the σ-fields $\mathcal{B}_{n+m}^a$ as $a \to \infty$.*

Intuitively, $\mathcal{B}_{-\infty}^n$ can be viewed as representing all the information contained in the past of the sequence $\{Z_s\}$ up to time n and $\mathcal{B}_{n+m}^{\infty}$ represent all the information contained in the future of the sequence $\{Z_s\}$ beginning from time $n+m$.

The following definition from [162] presents measures which describe weak dependence or asymptotic independence of a sequence $\{X_\tau\}$.

Definition 9.3 *Let $\mathcal{G}$ and $\mathcal{H}$ be σ-fields and define*

$$\alpha(\mathcal{G}, \mathcal{H}) \equiv \sup_{\{G \in \mathcal{G}, H \in \mathcal{H}\}} |P(GH) - P(G)P(H)|,$$

$$\rho(\mathcal{G}, \mathcal{H}) \equiv \sup_{X \in L_2(\mathcal{G}), Y \in L_2(\mathcal{H})} \frac{|EXY - EXEY|}{\sqrt{VarXVarY}},$$

$$\varphi(\mathcal{G},\mathcal{H}) \equiv \sup_{\{G\in\mathcal{G},H\in\mathcal{H}:P(G)>0\}} |P(H|G) - P(H)|,$$

$$\psi(\mathcal{G},\mathcal{H}) \equiv \sup_{\{G\in\mathcal{G},H\in\mathcal{H}:P(G)P(H)>0\}} \frac{|P(GH) - P(G)P(H)|}{P(G)P(H)},$$

$$\beta(\mathcal{G},\mathcal{H}) \equiv E(tvar_{G\in\mathcal{G}}|P(G|\mathcal{H}) - P(G)|),$$

$$\lambda(\mathcal{G},\mathcal{H}) \equiv \sup_{X\in L_{1/\alpha}(\mathcal{G}),Y\in L_{1/\beta}(\mathcal{H})} \frac{|EXY - EXEY|}{||X||_{1/\alpha}||Y||_{1/\beta}},$$

where tvar is total variation and $||X||_p = (E|X|^p)^{1/p}$.

The following definition provides two quantities which measure the dependence existing between two events separated by at least m time periods.

Definition 9.4 *A sequence of random vectors* $\{Z_s\}$, *with* $\mathcal{B}^n_{-\infty}$ *and* $\mathcal{B}^\infty_{n+m}$ *as above, is*

1. *α-mixing or strong mixing if* $\alpha(m) \equiv \sup_n \alpha(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$,

2. *ρ-mixing if* $\rho(m) \equiv \sup_n \rho(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$,

3. *φ-mixing or uniformly strong mixing if* $\varphi(m) \equiv \sup_n \varphi(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$,

4. *ψ-mixing if* $\psi(m) \equiv \sup_n \psi(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$,

5. *absolutely regular if* $\beta(m) \equiv \sup_n \beta(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$,

6. *(α,β)-mixing if* $\lambda(m) \equiv \sup_n \lambda(\mathcal{B}^n_{-\infty}, \mathcal{B}^\infty_{n+m}) \to 0$ *as* $m \to \infty$.

The following definition is required to state the law of large numbers for mixing sequences.

Definition 9.5 *Let* $a \in \mathbb{R}$. *(i) If* $\varphi(m) = O(m^{-a-\epsilon})$ *for some* $\epsilon > 0$, *then* φ *is of size* $-a$. *(ii) If* $\alpha(m) = O(m^{-a-\epsilon})$ *for some* $\epsilon > 0$, *then* α *is of size* $-a$.

The following law of large numbers, based on the concept of mixing, applies to heterogeneously dependent sequences.

Theorem 9.6 (McLeish) *Let $\{Z_t\}$ be a sequence of scalars with finite means $\mu_t \equiv E[Z_t]$ and suppose that $\sum_{t=1}^{\infty}(E|Z_t - \mu_t|^{1+\delta})/t^{1+\delta} < \infty$ for some $0 < \delta \leq r$ where $r \geq 1$. If φ is of size $-r/(2r-1)$ or α is of size $-r/(r-1)$, $r > 1$, then $\bar{Z}_n - \bar{\mu}_n \overset{a.s.}{\to} 0$.*

Proof. See [100] (Theorem 2.10). ■

Corollary 9.7 (White) *Let $\{Z_\tau\}$ be a sequence with φ of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$, such that $E|Z_\tau|^{r+\delta} < \Delta < \infty$ for some $\delta > 0$ and all s. Then $\bar{Z}_n - \bar{\mu}_n \overset{a.s.}{\to} 0$.*

Proof. See [153] (Corollary 3.48). ■

We next apply corollary 9.7 to the sequences of statistics S_{1n}, S_{2n}, and S_{3n} of the forecasting problem to obtain the consistency required by the Delta method theorems.

Proposition 9.8 (Martinez) *Let $\{X_\tau\}$ be a sequence of scalars with φ being of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$, such that $\mu_\tau \equiv E[X_\tau^2] < \infty$ and $E|X_\tau^2|^{r+\delta} < \Delta < \infty$ for some $\delta > 0$ and all s. Then $S_{2n} - \bar{\mu}_n \overset{a.s.}{\to} 0$.*

Proof. Given the sequence $\{X_\tau\}$ is φ-mixing of size $-r/(2r-1)$, $r \geq 1$, or α-mixing of size $-r/(r-1)$, $r > 1$, by theorem $A.40$ $\{X_s^2\}$ is a sequence with φ of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$. The result follows applying corollary 9.7 with $Z_\tau = X_s^2$. ■

Proposition 9.9 (Martinez) *Let $\{X_\tau\}$ and $\{Y_\tau\}$ be sequences of scalars with φ of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$, such that $\mu_\tau \equiv E[Y_{t+1}X_tY_\tau X_{s-1}] < \infty$ and $E|Y_{t+1}X_tY_\tau X_{s-1}|^{r+\delta} < \Delta < \infty$ for some $\delta > 0$ and all s. Then $S_{1n} - \bar{\mu}_n \overset{a.s.}{\to} 0$.*

Proof. Applying theorem $A.40$, $\{Y_{t+1}X_tY_\tau X_{s-1}\}$ is a sequence with φ of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$. The result follows applying corollary 9.7 with $Z_\tau = Y_{t+1}X_tY_\tau X_{s-1}$. ■

Proposition 9.10 (Martinez) *Let $\{X_\tau\}$ and $\{Y_\tau\}$ be sequences of scalars with φ of size $-r/(2r-1)$, $r \geq 1$, or α of size $-r/(r-1)$, $r > 1$, such that $\mu_\tau \equiv E[X_tY_\tau X_{s-1}] < \infty$ and $E|X_tY_\tau X_{s-1}|^{r+\delta} < \Delta < \infty$ for some $\delta > 0$ and all s. Then $S_{3n} - \bar{\mu}_n \overset{a.s.}{\to} 0$.*

Proof. The proof follows as that of proposition 9.9. ■

9.2 Moment inequalities for sums of random variables

For the application of the Delta method theorems of Chapter 8, certain moment conditions must be satisfied. In the previous section, we examined laws of large numbers for the statistics involved in the forecasting problem of Chapter 2. These laws of large numbers essentially warranty some level of stochastic convergence of a sample mean of statistics of the sequences $\{X_\tau\}$ and $\{Y_\tau\}$ to population means. In this section, we examine further conditions to determine rates at which the sample means converges to the respective population means. These rates of convergence are expressed by assumptions 8.4 and 8.12 of Chapter 8.

Consider a sequence of statistics $\{a_\tau\}$ and write $S_n = n^{-1}\sum a_\tau$ for the sample mean. For identically distributed sequences, we want to understand the n dependence of the central moments $E(S_n - \theta)^k$ where $\theta = E[a_\tau]$. For heterogeneously distributed sequences we study the central moments $E(S_n - \theta_n)^k$ where $\theta_\tau = E[a_\tau]$.

We now present the most significant moment inequality, results in the literature beginning with some covariance inequalities.

Theorem 9.11 (Theorem 17.2.3 in [77]) *Suppose the strictly stationary process $\{X_\tau\}$ satisfies the φ-mixing condition, and let the random variables ξ and η, respectively, be measurable with respect to $\mathcal{B}^n_{-\infty}$ and $\mathcal{B}^\infty_{n+m}$. If $E|\xi|^p < \infty$ and $E|\eta|^q < \infty$ with $p > 1, q > 1, 1/p + 1/q = 1$, then*

$$|E\xi\eta - E\xi E\eta| \leq 2\varphi(m)^{1/p} E^{1/p}|\xi|^p E^{1/q}|\eta|^q.$$

Theorem 9.12 (Lemma 2.1 in [38]) *Let the strictly stationary process $\{X_\tau\}$ satisfy the strong mixing condition, and let the random variables ξ and η, respectively, be measurable with respect to $\mathcal{B}^n_{-\infty}$ and $\mathcal{B}^\infty_{n+m}$; moreover, assume $E|\xi|^p < \infty$ for $p > 1$ and $|\eta| < C$ a.s. Then*

$$|E\xi\eta - E\xi E\eta| \leq 6CE^{1/p}|\xi|^p \alpha(m)^{1/q},$$

where q is such that $1/q + 1/p = 1$.

Corollary 9.13 (Corollary in [38]) *Under the assumptions of theorem* 9.12*, let the moments* $E|\xi|^p$ *and* $E|\eta|^q$ *exist with* $1/q + 1/p < 1$*. Then*

$$|E\xi\eta - E\xi E\eta| \leq 12E^{1/p}|\xi|^p E^{1/q}|\eta|^q \alpha(m)^{1-1/q-1/p}.$$

Other covariance inequalities for (α, β)-mixing sequences, ρ-mixing sequences and ψ-mixing sequences can be found in [162]. We are mainly interested in results concerning moment inequalities of partial sums. The next section presents results for sums of independent random variables.

9.2.1 Inequalities for moments of sums of independent random variables

Given an arbitrary sequence of random variables $\{X_\tau\}$, the following inequalities hold

$$E|S_n|^p \leq \sum_{\tau=1}^{n} E|X_\tau|^p, \quad 0 < p \leq 1, \tag{9.2.1}$$

$$E|S_n|^p \leq \sum_{\tau=1}^{n} n^{p-1} E|X_\tau|^p, \quad p > 1, \tag{9.2.2}$$

where $S_n = \sum_{\tau=1}^{n} X_\tau$. Inequalities (9.2.1) and (9.2.2) follow from the elementary inequalities

$$\Big|\sum_{\tau=1}^{n} a_\tau\Big|^p \leq \sum_{\tau=1}^{n} |a_\tau|^p, \quad 0 < p \leq 1$$

$$\Big|\sum_{\tau=1}^{n} a_\tau\Big|^p \leq n^{p-1} \sum_{\tau=1}^{n} |a_\tau|^p, \quad p > 1,$$

for every positive integer n and real numbers $a_1, \cdots, a_n$. Inequalities (9.2.1) and (9.2.2) can be strengthened with additional assumptions, as the following theorems demonstrate.

Theorem 9.14 (theorem 2.9 in [115]) *Let the sequence* $X_1, \cdots, X_n$ *be independent random variables with* $E[X_\tau] = 0$, $\tau = 1, \cdots, n$, $S_n = \sum_{\tau=1}^{n} X_\tau$ *and let* $p \geq 2$*. Define*

$$M_{p,n} = \sum_{\tau=1}^{n} E|X_\tau|^p, \quad B_n = \sum_{\tau=1}^{n} E[X_\tau^2].$$

Then

$$E|S_n|^p \le c(p)(M_{p,n} + B_n^{p/2}). \tag{9.2.3}$$

Inequality (9.2.3) is called the Rosenthal inequality.

Theorem 9.15 (theorem 2.10 in [115]) *Let $X_1, \cdots, X_n$ be a sequence of independent random variables with $E[X_\tau] = 0$, $\tau = 1, \cdots, n$, and let $p \ge 2$. Then*

$$E|S_n|^p \le C(p) n^{p/2-1} M_{p,n}, \tag{9.2.4}$$

where $C(p)$ is a positive constant depending only on p.

Theorem 9.16 (theorem 2.11 in [115]) *Let $X_1, \cdots, X_n$ be a sequence of independent random variables with $E[X_\tau] = 0$, $\tau = 1, \cdots, n$, $S_n = \sum_{\tau=1}^n X_\tau$ and let $p \ge 2$. Then*

$$E|S_n|^p \le c(p) \left[1 + \left(\sum_{\tau=1}^n P(X_\tau \ne 0) \right)^{p/2-1} \right] M_{p,n}. \tag{9.2.5}$$

If the sum $\sum_{\tau=1}^n P(X_\tau \ne 0)$ grows slower than n, then (9.2.5) is a better estimate than (9.2.4). The following theorems generalize the previous theorems by assuming $p > 1$ instead of $p \ge 2$.

Theorem 9.17 (theorem 2.12 in [115]) *Let $X_1, \cdots, X_n$ be a sequence of independent random variables and let $p > 1$. Define*

$$M_{p,n} = \sum_{\tau=1}^n E|X_\tau|^p, \quad D_n = \sum_{\tau=1}^n E|X_\tau|.$$

Then

$$E|S_n|^p \le c(p)(M_{p,n} + D_n^p), \tag{9.2.6}$$

where $S_n = \sum_{\tau=1}^n X_\tau$ and $c(p)$ is a positive constant depending only on p.

Theorem 9.18 (theorem 2.13 in [115]) *Let $X_1, \cdots, X_n$ be a sequence of independent random variables, $S_n = \sum_{\tau=1}^{n} X_\tau$ and let $p > 1$. Then*

$$E|S_n|^p \leq c(p)\left[1 + \left(\sum_{\tau=1}^{n} P(X_\tau \neq 0)\right)^{p-1}\right] M_{p,n}. \tag{9.2.7}$$

Another type of inequality called the Marcinkiewics-Zygmund inequality is of importance. Brillinger in [26] gives a Marcinkiewics-Zygmund inequality for a sequence of i.i.d. random variables.

Theorem 9.19 ([26]) *Let $X_1, \ldots, X_n$ be a sample from a distribution with cdf $F(x)$ having mean zero. If there exists m, $m \geq 2$, such that $E|X|^m < \infty$, then there exists n_0 such that $E|X_1 + \cdots + X_n|^m < Kn^{m/2}$ for all $n > n_0$ and some positive K.*

9.2.2 Inequalities for moments of sums of dependent random variables

Doob in [44] presents a moment inequality for a stationary Markov sequence satisfying Doeblin's condition.

Theorem 9.20 (Lemma 7.4 in [44]) *Let $\{X_\tau\}$ be a stationary aperiodic Markov sequence which is Markov ergodic and satisfies Doeblin's condition and $E|X_\tau|^v \leq C$ for all $s \geq 1$, some $v > 2$, and some $C < \infty$. Then $E|\sum_{i=a+1}^{a+n} X_i|^v \leq Kn^{v/2}$ for all $a \geq 0$, all $n \geq 1$ and some $K < \infty$.*

Stout in [136] obtains the same moment inequality as Doob for a martingale difference sequence. Yoshihara in [160] provides even order moment inequalities for weighted partial sums of φ-mixing processes.

Theorem 9.21 (Theorem 1 in [160]) *Let $\{\xi_\tau\}$ be φ-mixing. We assume that for an even integer $m \geq 2$, $E[\xi_\tau] = 0$ and $E|\xi_\tau|^m \leq M$ $\tau = 1, 2, \cdots$, and*

$$\sum_{i=1}^{\infty} (i+1)^{m/2-1} \varphi(i)^{1/m} < \infty.$$

Then, for every sequence $\{a_s\}$ and for every integer n

$$E\Big[\Big(\sum_{i=b+1}^{b+n} a_i\xi_i\Big)^m\Big] \leq c_m A_{b,n}^m,$$

for all $b \geq 0, n \geq 1$ where c_m is an absolute constant depending only on m and $A_{b,n}^2 = \sum_{i=b+1}^{b+n} a_i^2$.

Theorem 9.22 (theorem 3 in [46]) *Let $\{X_\tau\}$ be a sequence of centered φ-mixing random variables with $|X_\tau| \leq 1$ a.s., $E[X_\tau^2] \leq M$ $\forall n$, $\exists q \in \mathbb{N}$, $q \geq 2$ then*

$$|E[S_n^q]| \leq K(\varphi, q) \sum_{i=1}^{[q/2]} n^i M^i,$$

where $S_n = \sum_{\tau=1}^{n} X_\tau$ and $K(\varphi, q)$ is a constant polynomial of $(\Phi_0(1/2), \cdots, \Phi_{q-1}(1/2))$ and

$$\Phi_a(b) = \sum_{i=0}^{\infty}(i+1)^a \varphi_i^b.$$

The first results for strong mixing processes were given by Yoshihara in [160]. The following result is for strong mixing processes which need not be strictly stationary.

Theorem 9.23 (Theorem 3 in [160]) *Let $\{\xi_\tau\}$ be a strong mixing sequence with coefficient $\alpha(n)$. We assume that for some $\delta > 0$ and for an even integer $m \geq 2$, $E[\xi_\tau] = 0$, $E|\xi_\tau|^{m+\delta} \leq M < \infty$ and $\sum_{i=1}^{\infty}(i+1)^{m/2-1}\alpha(i)^{\delta/(m+\delta)} < \infty$. Then, for every sequence $\{a_\tau\}$ and for every integer n*

$$E\Big[\Big(\sum_{i=b+1}^{b+n} a_i\xi_i\Big)^m\Big] \leq c'_m A_{b,n}^m,$$

with $A_{b,n}^m = \sum_{i=b+1}^{b+n} a_i^m$ for all $b \geq 0, n \geq 1$ where c'_m is n absolute constant depending only on m.

Yokoyama in [159] presents moment bounds for a stationary strong mixing sequence.

Theorem 9.24 (Theorem 1 in [159]) *Let $\{X_\tau\}$ be a strictly stationary strong mixing sequence with $E[X_\tau] = 0$ and $E|X_1|^{r+\delta} < \infty$ for some $r > 2$ and $\delta > 0$. If $\sum_{i=0}^{\infty}(i+$*

$1)^{r/2-1}\alpha(i)^{\delta/(r+\delta)} < \infty$, *then there exists a constant* K *such that*

$$E|S_n|^r \le Kn^{r/2}, \tag{9.2.8}$$

with $n \ge 1$ *and* $S_n = \sum_{i=1}^n X_i$.

Theorem 9.25 (Theorem 2 in [159]) *Let* $\{X_\tau\}$ *be a strictly stationary strong mixing sequence with* $E[X_\tau] = 0$ *and* $|X_1| \le C < \infty$ *a.s.. If* $\sum_{i=1}^\infty (i+1)^{r/2-1}\alpha(i) < \infty$, *then* (9.2.8) *holds.*

Theorem 9.26 (theorem 10 in [46]) *Let* $\{X_\tau\}$ *be a strong mixing sequence with* $|X_\tau| \le 1$ *a.s.,* $E[X_\tau^2] \le M$ $\forall n$, $\exists \delta > 0, q \in \mathbb{N}, q \ge 2$, $S_n = \sum_{\tau=1}^n X_\tau$, $A_{q-2}(1/2) < \infty$ *with* $A_a(b) = \sum_{i=0}^\infty (i+1)^a \alpha^b$, *then*

$$|E[S_n^q]| \le k(q,\alpha) \sum_{i=1}^{[q/2]} n^i M^i.$$

Theorem 9.27 (theorem 11 in [46]) *Let* $\{X_\tau\}$ *be a centered strong mixing sequence of random variables and* $S_n = \sum_{\tau=1}^n X_\tau$ *such that*

$$M_h = \sup\{||X_\tau||_{h+\delta}, n \ge 0\} < \infty \text{ and } A_{h-2}(\delta/(h+\delta)) < \infty,$$

with $A_a(b) = \sum_{i=0}^\infty (i+1)^a \alpha^b$. *Then*

$$|E[S_n^q]| \le k'(q,\alpha) \sum_{i=1}^{[q/2]} n^i M_{q-2i+2}^q.$$

For moment inequalities of ρ-mixing sequences, early work is due to Peligrad [109, 110] which was later improved and generalized by Shao in [128, 129]. Some of the proofs to the following theorems can be found in [162].

Theorem 9.28 *Let* $\{X_\tau\}$ *be a* ρ*-mixing sequence with* $E[X_\tau] = 0, E[X_\tau^2] < \infty$ *for each* $\tau \ge 1$. *Then for any* $\epsilon > 0$, *there exists a* $C = C(\epsilon) > 0$ *such that*

$$E[S_k^2(n)] \le Cn \exp\left\{(1+\epsilon)\sum_{i=0}^{[\log n]} \rho(2^i)\right\} \max_{k<i\le k+n} E[X_i^2],$$

for each $k \geq 1$ and $n \geq 1$, where $S_k(n) = \sum_{i=k+1}^{k+n} X_i$.

Theorem 9.29 *Let $\{X_\tau\}$ be a ρ-mixing sequence with $E[X_\tau] = 0, \sum_\tau E|X_\tau|^{2+\delta} < \infty$ for some $\delta \geq 0$ and $S_k(n) = \sum_{i=k+1}^{k+n} X_i$. Then, for any $\epsilon > 0$, there exists a $C = C(\delta, \rho(\dots), \epsilon) > 0$, such that for each $n \geq 2$*

$$E|S_k(n)|^{2+\delta} \leq C\Big\{\Big(n \exp\Big\{(1+\epsilon)\sum_{i=0}^{[\log n]} \rho(2^i)\Big\} \max_{k<i\leq k+n} EX_i^2\Big)^{1+\delta/2} + n \exp\Big\{C \sum_{i=0}^{[\log n]} \rho^{2/(2+\delta)}(2^i)\Big\} \max_{k<i\leq k+n} E|X_i|^{2+\delta}\Big\}.$$

Theorem 9.30 *Let $\{X_\tau\}$ be a ρ-mixing sequence with $E[X_\tau] = 0, E|X_\tau|^q < \infty$, $q \geq 2$, $S_k(n) = \sum_{i=k+1}^{k+n} X_i$ and $E[S_k^2(n)] \leq nh(n) \max_{k<i\leq k+n} E_i^2$. Suppose there exists a function $h(n)$ and there exists a positive integer n_0 and a constant $0 < \theta < 2^{1-2/(q\wedge 3)}$ such that*

$$\max(h([n/2]), h(n-[n/2])) \leq \theta h(n),$$

for $n \geq n_0$. Furthermore, when $q > 3$ assume that there exists a $C > 0$ such that

$$h(n) \geq \frac{1}{C} \exp\Big\{-C \sum_{i=0}^{[\log n]} \rho^{2/q}(2^i)\Big\}.$$

Then there exists a constant $K = K(q, n_0, \theta, C, \rho(\cdot))$, such that for every $k \geq 0, n \geq 1$

$$E|S_k(n)|^q \leq K\Big\{(nh(n) \max_{k<i\leq k+n} EX_i^2)^{q/2} + n \exp\Big\{K \sum_{i=0}^{[\log n]} \rho^{2/q}(2^i) \max_{k<i\leq k+n} E|X_i|^q\Big\}.$$

A finite family $\{X_1, \cdots, X_m\}$ of r.v.'s is associated if for any two coordinate-wise none-decreasing functions f, g on $\mathbb{R}^m$, $Cov(f(X_1, \cdots, X_m), g(X_1, \cdots, X_m)) > 0$ holds whenever the covariance is defined. An infinite family is associated if every finite sub-family is associated.

Theorem 9.31 (theorem 1 in [21]) *Let $\{X_\tau\}$ be a sequence of associated random variables with $E[X_\tau] = 0$ and $\sup_{j\in\mathbb{N}} E|X_j|^{r+\delta} < \infty$ for some $r > 2$ and $\delta > 0$ and*

$S_n = \sum_{\tau=1}^{n} X_\tau$. *Assume* $u(n) = O(n^{-(r-2)(r+\delta)/2\delta})$. *Then there is a constant* B *not depending on* n *such that for all* $n \in \mathbb{N}$

$$\sup_{m \in \mathbb{N} \cup \{0\}} E|S_{n+m} - S_m| \leq B n^{r/2}.$$

For the following theorem, we define, for $q > 1$ and any $A > q$, the class $\Phi_{2,A}$ of Orlicz functions as follows:

$$\Phi_{2,A} = \{\phi : \mathbb{R}^+ \to \mathbb{R}^+ ; \phi \quad \text{convex}, \phi(x)/x^q \quad \text{increasing}, \phi(x)/x^A \quad \text{decreasing}\}.$$

Theorem 9.32 (theorem 1 in [122]) *Assume* $M_{2,\alpha} < \infty$ *and let* ϕ *be some element of* $\Phi_{2,A}$ *such that* $E[\phi(|X_0|)] < \infty$. *Let* $S_n^* = \sup_{j \leq n} |S_j|$ *with* $S_n = \sum_{\tau=1}^{n} X_\tau$. *Then there exists some positive constant* C_A *depending only on* A *such that*

$$E[\phi(S_n^*)] \leq C_A(\phi(\sqrt{n M_{2,\alpha}}) + n M_{\phi,\alpha,n}).$$

Theorem 9.33 (theorem 1 in [45]) *Let* $\{X_\tau\}$ *be a sequence of centered random variables fulfilling for some fixed* $q \in \mathbb{N}, q \geq 2$ $C_{r,q} = O(r^{-q/2})$ *as* $r \to \infty$ *with* $C_{r,q} \equiv \sup |Cov(X_{t_1} \cdots X_{t_m}, X_{t_{m+1}} \cdots X_{t_q})|$ *where the sumpremum is taken over all* $\{t_1, \ldots, t_q\}$ *such that* $1 \leq t_1 \leq \cdots \leq t_q$ *and* m, r *satisfy* $t_{m+1} - t_m = r$. *Then there exists a positive constant* B *not depending on* n *for which* $|E[S_n^q]| \leq B n^{q/2}$ *where* $S_n = \sum_{\tau=1}^{n} X_\tau$.

For the theorem that follows, we define an AG sequence $\{X_\tau\}$ as a sequence fulfilling the following inequality:

$$|Cov(H(X_i, i \in A), K(X_j, j \in B))| \leq \sum_{i \in A} \sum_{j \in B} \left\| \frac{\partial H}{\partial x_i} \right\|_\infty \left\| \frac{\partial K}{\partial x_i} \right\|_\infty |Cov(X_i, X_j)|,$$

where A and B are arbitrary finite disjoint subsets of $\mathbb{N}$, and H and K are real valued functions having uniformly bounded first derivatives.

Theorem 9.34 (theorem 1 in [96]) *Let* r *be a fixed real number,* $r > 2$. *Let* $\{X_\tau\}$ *be a strictly stationary sequence of centered and* AG *random variables. Suppose, moreover, this sequence is bounded by* M. *Then there exists a positive constant* C_r *depending only* *on r, such that*

$$E|S_n|^r \leq C_r\Big[s_n^r + \sum_{k=1}^{n}\sum_{i=0}^{k-1} M^{r-2}(i+1)^{r-2}|Cov(X_1, X_{1+i})|\Big],$$

where $s_n^2 := n\sum_{i=0}^{n} |Cov(X_1, X_{1+i})|$.

9.3 Finite moment generating functions

In Chapter 8, we encounter moment generating functions which are required to be finite for the Delta method theorems to hold. These assumptions, which include 8.5, 8.8, 8.13, 8.14, 8.17, depend only on the nature of the processes involved. In this section, we present theorems providing conditions under which different moment generating functions are finite.

Proposition 9.35 (Martinez) *Consider a sequence $\{X_\tau\}$ of r.v's with $E[X_\tau] = \theta_\tau$, $\bar{X}_n = 1/n\sum_{\tau=1}^{n} X_\tau$, and $E(X_\tau - \theta_\tau)^k < \infty$ for $\tau = 1, \ldots, n$ and all k. If, for $c > 0$ a finite constant, either of the two following conditions hold*

1. $\sum_{k=0}^{\infty} c^k/k!E(X_\tau - \theta_\tau)^k < \infty$ *and* $\sum_{k=0}^{\infty}(-1)^k c^k/k!E(X_\tau - \theta_\tau)^k < \infty$ *for* $\tau = 1, \ldots, n$,
2. $\sum_{k=0}^{\infty} c^k/k!E|X_\tau - \theta_\tau|^k < \infty$,

then $E[\exp(c|\bar{X}_n|)] < \infty$.

Proof. First, we use the fact $\exp(c|\bar{X}_n|) \leq \exp(c\bar{X}_n) + \exp(-c\bar{X}_n)$. We show under condition 1, $E[\exp(c\bar{X}_n)] < \infty$. By Hölder's inequality

$$E[\exp(c\bar{X}_n)] \leq E^{1/n}[\exp(cX_1)] \cdots E^{1/n}[\exp(cX_n)].$$

We can now prove $E[\exp(cX_\tau)] < \infty$ for $\tau = 1, \ldots, n$. Condition 1, $\sum_{k=0}^{\infty} c^k/k!E(X_\tau - \theta_\tau)^k < \infty$, implies $\sum_{k=1}^{\infty} c^k/k!(X_\tau - \theta_\tau)^k$ converges absolutely a.e. and $\exp(cX_\tau) = \sum_{k=1}^{\infty} c^k/k!(X_\tau - \theta_\tau)^k$ a.e. and $E[\exp(cX_\tau)] = \sum_{k=0}^{\infty} c^k/k!E(X_\tau - \theta_\tau)^k$. Therefore, we obtain the bound $E[\exp(cX_\tau)] < \infty$ for $\tau = 1, \ldots, n$. Similarly, condition 1 implies

$E[\exp(-cX_\tau)] < \infty$ for $\tau = 1, \ldots, n$ and result is obtained. Condition 2 implies condition 1 so no proof is required. ■

Note, even though $c > 0$, $E[\exp(c|\bar{X}_n|)] < \infty$ implies $E[\exp(-c|\bar{X}_n|)] < \infty$. Also, condition 2 is included in the proposition only for completeness since condition 2 implies condition 1. Condition 2 is weaker but more difficult to verify than condition 1. The condition of the proposition states that if the central moments of the r.v's $X_1, \ldots, X_n$ decay with k or grow at a rate which makes the appropriate series converge, then $E[\exp(c|\bar{X}_n|)] < \infty$.

Example 9.36 *Let $\{X_\tau\}$ be a sequence of i.i.d. r.v's with $\{X_\tau\} \sim N(\theta, \sigma)$. We know the central moments are*

$$E(X-\theta)^k = \frac{k!\sigma^k}{2^{k/2}(k/2)!} \quad \textit{for} \quad k \quad \textit{even},$$
$$E(X-\theta)^k = 0 \quad \textit{for} \quad k \quad \textit{odd}.$$

It follows

$$\sum_{k=0}^{\infty} \frac{c^k}{k!} E(X-\theta)^k = \sum_{k=0}^{\infty} \frac{(c\sigma)^{2k}}{2^k k!},$$

converges by the ratio test for any c and σ. Similarly, the second series of condition 1 of proposition 9.35 *converges and $E[\exp(c|\bar{X}_n|)] < \infty$ for any c, θ, σ.* □

The result of proposition 9.35 can be extended to the multivariate case.

Proposition 9.37 (Martinez) *Let each of $\{Z_{1\tau}\}, \ldots, \{Z_{m\tau}\}$ be a strictly stationary strong mixing sequence with $E[Z_{j\tau}] = \theta_{j\tau}$, $E|Z_{j\tau} - \theta_{j\tau}|^{r+\delta} < \infty$ for $r > 2$ and $\delta > 0$ and satisfying $\sum_{i=0}^{\infty} (i+1)^{r/2-1} \alpha_j(i)^{\delta/(r+\delta)} < \infty$ for $j = 1, \ldots, m$. Then $E[\exp(c||\bar{Z}_\tau||)] < \infty$ for some finite constant $c > 0$ with $||\bar{Z}_\tau|| = \bar{Z}_{1\tau} + \cdots + \bar{Z}_{m\tau}$ for any $\tau > 0$.*

Proof. By proposition 9.35, $E[\exp(c_j|\bar{Z}_{j\tau}|)] < \infty$, $j = 1, \ldots, m$ for finite constants $c_1, \ldots c_m$. The result follows from an application of Hölder's inequality. ■

Chapter 10

Concluding remarks

10.1 Summary

In this thesis, we address the combined effects of misspecification and stochastic dynamics on the forecasts of time series. The problem consists of using a linear regression model in conjunction with the OLS estimator to form a forecast of a dependent variable whose data generating process is unknown to the practitioner. The MSFE is the forecast evaluation criterion of choice. The main consequence of interest is the existence of optimal observation windows as a result of model misspecification. To determine the existence of optimal observation windows, we need to understand the behavior of the MSFE for finite values of the sample size variable n. The sample size dependence of the square forecast error is implicit through the OLS and understanding the sample size dependence of the MSFE has to be done through an approximation. To obtain an approximation of the MSFE, we construct an algorithm based on Taylor expansions of the MSFE which do not require knowledge of the functional form of the DGP of the dependent process. Three type of stochastic dynamics are studied: independent and identically distributed processes, covariance stationary processes, and independent and identically distributed processes which undergo a structural break at point in time $t - n_b$. An approximation for each of the three stochastic dynamics is constructed which exploits their particular characteristics. For the independent and identically distributed processes, the MSFE approximation depends explicitly on the sample size variable n and on population moments of the explanatory and dependent processes. For practical applications, the population moments can be replaced by sample moments. Numerical experiments are carried out under the assumption that the range of the random variables be mostly contained inside the region of convergence of the Taylor expansions. The approximation performs well in replicating the benchmark MSFE, even when the condition that the range of the random variables be mostly contained inside the region of convergence is violated. Several examples of functional misspecification are explored with all resulting in no optimal observation windows. For the covariance stationary processes, the MSFE approximation depends implicitly on the sample size variable n through summations of population moments of the explanatory and dependent processes. The implicit dependence on the sample size variable n complicates the analysis of the SSD, but practical applications are still feasible with sample moments. Finally, for the independent and identically distributed processes which undergo a temporal structural break, the MSFE approximation depends explicitly on the sample size variable n, the known variable n_b, and on population moments of the explanatory and dependent processes. In numerical experiments, the approximation performs well in replicating the benchmark MSFE even when the condition that the range of the random variables be mostly contained inside the region of convergence is violated.

10.2 Some remarks

10.2.1 Monte Carlo simulations for the MSFE and OLS process

10.2.1.1 General Principles

Monte Carlo simulation are methods to estimate the expected value of a process based on observations of the process or to estimate the expected value of functions of processes based on observations of the processes or on observations of the functions of the processes. At the core of Monte Carlo simulations is the idea that as the number of observations increases we can expect stochastic convergence. In this sense, Monte Carlo simulations rely on the concept known as a law of large numbers. Laws of large numbers have the general form given by the following proposition.

Proposition 10.1 *Given restrictions on the dependence, heterogeneity, and moments of a sequence of random variables* $\{Z_\tau\}$, $\bar{Z}_m - \bar{\mu}_m \overset{a.s.}{\to} 0$, *where* $\bar{Z}_m \equiv m^{-1}\sum_{\tau=1}^{m} Z_\tau$ *and*

$\bar{\mu}_m \equiv E[\bar{Z}_m]$.

Four different cases can be outlined based on dependence, heterogeneity, and moments of the processes. The four cases are: independent identically distributed observations; independent heterogeneously distributed observations; dependent identically distributed observations; and dependent heterogeneously distributed observations. The following four theorems state the conditions necessary for stochastic convergence for the four cases outlined above.

(1) Independent identically distributed observations:

Theorem 10.2 (Kolmogorov) *Let $\{Z_\tau\}$ be a sequence of i.i.d. random variables. Then $\bar{Z}_n \overset{a.s.}{\to} \mu$ if and only if $E|Z_\tau| < \infty$ and $E[Z_\tau] = \mu$ where $\bar{Z}_m \equiv m^{-1}\sum_{\tau=1}^{m} Z_\tau$.*

Proof. [119], p. 115. ■ (2) Independent heterogeneously distributed observations:

Theorem 10.3 (Markov) *Let $\{Z_\tau\}$ be a sequence of independent random variables, with finite means $\mu_\tau \equiv E[Z_\tau]$. If for some $\delta > 0$, $\sum_{\tau=1}^{\infty}(E|Z_\tau - \mu_\tau|^{1+\delta})/\tau^{1+\delta} < \infty$, then $\bar{Z}_m - \bar{\mu}_m \overset{a.s.}{\to} 0$.*

Proof. [31], pp. 125-126. ■ (3) Dependent identically distributed observations:

Theorem 10.4 (Ergodic theorem) *Let $\{Z_\tau\}$ be a stationary ergodic scalar sequence with $E|Z_\tau| < \infty$. Then $\bar{Z}_n \overset{a.s.}{\to} \mu \equiv E[Z_t]$*

Proof. [136], p. 181. ■ (4) Dependent heterogeneously distributed observations:

Theorem 10.5 (McLeish) *Let $\{Z_\tau\}$ be a sequence of scalars with finite means $\mu_\tau \equiv E[Z_\tau]$ and suppose that $\sum_{\tau=1}^{\infty}(E|Z_\tau - \mu_\tau|^{1+\delta})/\tau^{1+\delta} < \infty$ for some δ, $0 < \delta \leq r$ where $r \geq 1$. If ϕ is of size $-r/(2r-1)$ or α is of size $-r/(r-1)$, $r > 1$, then $\bar{Z}_m - \bar{\mu}_m \overset{a.s.}{\to} 0$. ϕ and α are the uniform mixing and strong mixing parameters respectively.*

Proof. [100], Theorem 2.10. ■ The second and fourth cases concern covariance stationary processes as well as non-stationary (evolutionary) processes. The laws of large numbers for these cases establish the convergence of the average of the process realizations to the average of the population means. If the means μ_τ are a constant μ, the average of process realizations converge simply to μ. The first and third cases concern a particular subset of

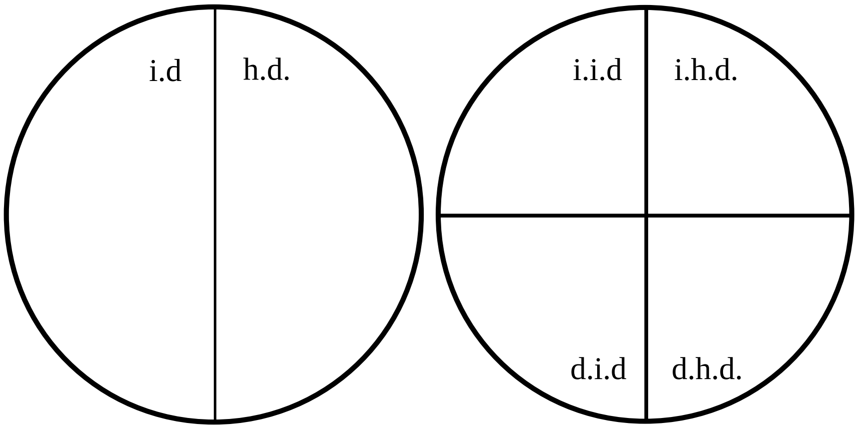


Figure 10.1: Identically distributed, heterogeneously distributed

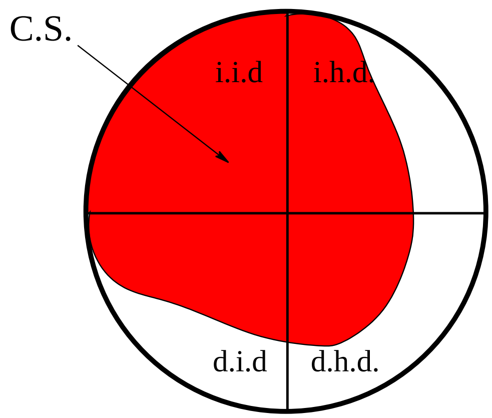


Figure 10.2: Covariance Stationary

covariance stationary processes as well as strictly stationary processes. Depending on the nature of the process at hand, one must chose a Monte Carlo method which appropriately applies one of the laws of large numbers described above. In what follows, we describe two commonly used methods to build Monte Carlo simulations. The following propositions will be useful.

Proposition 10.6 (White) *Let $g : \mathbb{R}^k \to \mathbb{R}^l$ be a measurable function. (i) Let Z_τ and Z_t be identically distributed. Then $g(Z_\tau)$ and $g(Z_t)$ are identically distributed. (ii) Let Z_τ and Z_t be independent. Then $g(Z_\tau)$ and $g(Z_t)$ are independent.*

Proof. [153], p.32 ■

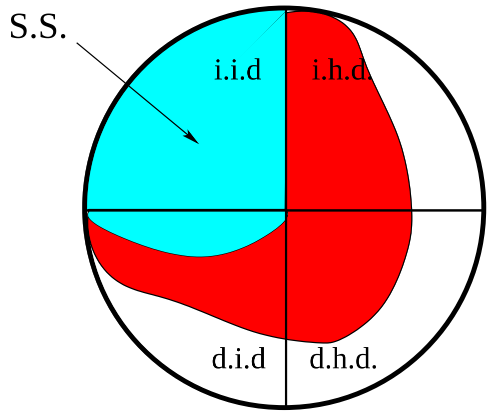


Figure 10.3: Strictly Stationary

Proposition 10.7 (White) *Let g be an $\mathcal{F}$ measurable function into $\mathbb{R}^k$ and define $Y_t \equiv g(\ldots, Z_{t-1}, Z_t, Z_{t+1}, \ldots)$, where Z_t is $q \times 1$. (i) If $\{Z_\tau\}_\tau$ is stationary, then $\{Y_\tau\}_\tau$ is stationary. (ii) If $\{Z_\tau\}_\tau$ is stationary and ergodic, then $\{Y_\tau\}_\tau$ is stationary and ergodic.*

Proof. [153], p. 44. ■

10.2.1.2 Method 1

To describe the first method, we consider the two cases of independent and identically distributed processes and heterogeneously distributed processes. Furthermore, we will illustrate the method for the sum statistic of the process and for measurable functions of the sum statistic of the process.

Identically distributed processes

We construct a Monte Carlo method to estimate $E[S_n]$ and $E[g(S_n)]$ where $S_n = \sum_{i=1}^n X_i$ and g is a measurable function. The method is constructed by generating a single i.i.d. series $\{X_1, X_2, \ldots, X_{n+m-1}\}$. From this series, we construct the following vectors:

$$\begin{aligned}
Z_1 &= (X_1, X_2, \ldots, X_n),\\
Z_2 &= (X_2, X_3, \ldots, X_{n+1}),\\
&\vdots\\
Z_m &= (X_m, X_{m+1}, \ldots, X_{n+m-1}).
\end{aligned}$$

The sequence $\{Z_\tau\}_\tau$ is i.i.d. m sum statistics can be constructed from these vectors as follows:

$$S_{n,1} = \sum_{i=1}^{n} X_i, \quad S_{n,2} = \sum_{i=2}^{n+1} X_i, \quad \dots, \quad S_{n,m} = \sum_{i=m}^{n+m-1} X_i. \tag{10.2.1}$$

It follows by proposition 10.6, the process $\{S_{n,\tau}\}_\tau$ is i.d.d. Defining $\bar{S}_{n,m} = 1/m \sum_{i=1}^{m} S_{n,i}$, from the law of large numbers, theorem 10.2, it follows:

$$\bar{S}_{n,m} \xrightarrow{a.s.} E[S_{n,\tau}] \equiv \mu.$$

Next, setting $Y_{n,\tau} = g(S_{n,\tau})$ with g a measurable function, it follows $\{Y_{n,\tau}\}_\tau$ is an i.i.d. process. With $\bar{Y}_{n,m} = 1/m \sum_{i=1}^{m} Y_{n,i}$, it follows $\bar{Y}_{n,m} \xrightarrow{a.s.} E[Y_{n,\tau}] = E[g(S_{n,\tau})] \equiv \nu$ by the law of large numbers, theorem 10.2.

Heterogeneously distributed processes

The method is constructed by generating a single heterogeneously distributed series $\{X_1, X_2, \dots, X_{n+m-1}\}$. m sum statistics can be constructed from this series as follows:

$$S_{n,1} = \sum_{i=1}^{n} X_i, \quad S_{n,2} = \sum_{i=2}^{n+1} X_i, \quad \dots, \quad S_{n,m} = \sum_{i=m}^{n+m-1} X_i. \tag{10.2.2}$$

It follows that the process $\{S_{n,\tau}\}_\tau$ is heterogeneously distributed with $E[S_{n,\tau}] \equiv \mu_\tau$. Defining $\bar{S}_{n,m} = 1/m \sum_{i=1}^{m} S_{n,i}$ and $\bar{\mu}_m = 1/m \sum_{i=1}^{m} \mu_i$, from the law of large numbers, theorem 10.3 and theorem 10.5, it follows:

$$\bar{S}_{n,m} - \bar{\mu}_m \xrightarrow{a.s.} 0.$$

Next, setting $Y_{n,\tau} = g(S_{n,\tau})$ with g a measurable function, it follows $\{Y_{n,\tau}\}_\tau$ is heterogeneously distributed with $\nu_\tau \equiv E[Y_{n,\tau}]$. With $\bar{Y}_{n,m} = 1/m \sum_{i=1}^{m} Y_{n,i}$ and $\bar{\nu}_m = 1/m \sum_{i=1}^{m} \nu_i$, it follows $\bar{Y}_{n,m} - \bar{\nu}_m \xrightarrow{a.s.} 0$ by the law of large numbers, theorem 10.3 and theorem 10.5. Now suppose $\{X_\tau\}_\tau$ is heterogeneously distributed and in addition

$$E[X_\tau] = E[X_t] \equiv \alpha, \quad \text{for any} \quad \tau \quad \text{and} \quad t, \tag{10.2.3}$$

all variances are constant and finite and covariances depend only on the time lag between X_τ and X_t. This includes a large set of weakly stationary processes. As before, the process $\{S_{n,\tau}\}_\tau$ is heterogeneously distributed but (10.2.3) implies

$$\mu_\tau \equiv E[S_{n,\tau}] = E[S_{n,t}] \equiv \mu_t \quad \text{for any} \quad t, \tau, \quad \mu_\tau \equiv \mu = n\alpha \quad \text{and} \quad \bar{\mu}_m = \mu. \quad (10.2.4)$$

It follows $\bar{S}_{n,m} \xrightarrow{a.s.} \mu$ by the LLN. As before, setting $Y_{n,\tau} = g(S_{n,\tau})$ with g a measurable function, it follows $\{Y_{n,\tau}\}_\tau$ is heterogeneously distributed with $\nu_\tau \equiv E[Y_{n,\tau}] = E[g(S_{n,\tau})]$. It is important to note (10.2.4) does not imply $E[g(S_{n,\tau})] = E[g(S_{n,t})]$ for $t \neq \tau$. In fact, the equality $E[g(S_{n,\tau})] = E[g(S_{n,t})]$ is unlikely to hold since the expectations depend on the distributions of $S_{n,\tau}$ and $S_{n,t}$ which are heterogeneous. This construction makes the use of method 2 and brute force methods inappropriate to estimate the expected value of functions of weakly stationary processes.

10.2.1.3 Method 2

To describe the second method, we again consider the two cases of identically distributed processes and heterogeneously distributed processes. We will illustrate the method for the sum statistic of the process and for measurable functions of the sum statistic of the process.

Identically distributed processes

We construct a Monte Carlo method to estimate $E[S_n]$ and $E[g(S_n)]$ where $S_n = \sum_{i=1}^{n} X_i$ and g is a measurable function. The method is constructed by generating in an identical manner m independent series of length n

$$\begin{gathered} \{X_{1,1}, X_{1,2}, \ldots, X_{1,n}\}, \\ \vdots \\ \{X_{m,1}, X_{m,2}, \ldots, X_{m,n}\}, \end{gathered}$$

where $\{X_{i,j}\}_j$ is identically distributed for each fixed i. From these series, we construct the following vectors:

$$Z_1 = (X_{1,1}, X_{1,2}, \ldots, X_{1,n}),$$
$$\vdots$$
$$Z_m = (X_{m,1}, X_{m,2}, \ldots, X_{m,n}).$$

By construction, the sequence $\{Z_\tau\}_\tau$ is i.i.d. m sum statistics can be constructed from these vectors as follows:

$$S_{n,1} = \sum_{i=1}^{n} X_{1,i}, \quad S_{n,2} = \sum_{i=1}^{n} X_{2,i}, \quad \ldots, \quad S_{n,m} = \sum_{i=1}^{n} X_{m,i}. \tag{10.2.5}$$

It follows from proposition 10.6, $\{S_{n,\tau}\}_\tau$ is an i.i.d. process. Setting $\bar{S}_{n,m} = 1/m \sum_{i=1}^{m} S_{n,i}$, it follows by the law of large numbers, theorem 10.2, $\bar{S}_{n,m} \xrightarrow{a.s.} E[S_{n,\tau}] \equiv \mu$. Next, setting $Y_{n,\tau} = g(S_{n,\tau})$ with g a measurable function, from proposition 10.6, it follows $\{Y_{n,\tau}\}_\tau$ is an i.i.d process. With $\bar{Y}_{n,m} = 1/m \sum_{i=1}^{m} Y_{n,i}$, it follows $\bar{Y}_{n,m} \xrightarrow{a.s.} E[Y_{n,\tau}] = E[g(S_{n,\tau})] \equiv \nu$ by the law of large numbers, theorem 10.2.

Heterogeneously distributed processes

We construct a Monte Carlo method to estimate $E[S_n]$ and $E[g(S_n)]$ where $S_n = \sum_{i=1}^{n} X_i$ and g is a measurable function. As before, the method is constructed by generating in the same manner, m independent series of length n

$$\{X_{1,1}, X_{1,2}, \ldots, X_{1,n}\},$$
$$\vdots$$
$$\{X_{m,1}, X_{m,2}, \ldots, X_{m,n}\}.$$

The following properties hold:

- For a fixed i, $\{X_{i,j}\}_j$ is heterogeneously distributed.
- For a fixed j, $\{X_{i,j}\}_i$ is identically distributed.
- For any j and k, $\{X_{i,j}\}_j$ and $\{X_{l,k}\}_k$ are independent for $i \neq l$.

From these series, we construct the following vectors:

$$Z_1 = (X_{1,1}, X_{1,2}, \ldots, X_{1,n}),$$
$$\vdots$$
$$Z_m = (X_{m,1}, X_{m,2}, \ldots, X_{m,n}).$$

The sequence $\{Z_\tau\}_\tau$ is i.i.d. m sum statistics can be constructed from these vectors as follows:

$$S_{n,1} = \sum_{i=1}^{n} X_{1,i}, \quad S_{n,2} = \sum_{i=1}^{n} X_{2,i}, \quad \ldots, \quad S_{n,m} = \sum_{i=1}^{n} X_{m,i}. \tag{10.2.6}$$

By proposition 10.6, $\{S_{n,\tau}\}_\tau$ is i.i.d. and for $\tau \neq t$, $E[S_{n,\tau}] = E[S_{n,t}] \equiv \mu$. Setting $\bar{S}_{n,m} = 1/m \sum_{i=1}^{m} S_{n,i}$, it follows by the law of large numbers, theorem 10.2, $\bar{S}_{n,m} \xrightarrow{a.s.} E[S_{n,\tau}] \equiv \mu$. Next, setting $Y_{n,\tau} = g(S_{n,\tau})$ with g a measurable function, from proposition 10.6, it follows $\{Y_{n,\tau}\}_\tau$ is an i.i.d process. With $\bar{Y}_{n,m} = 1/m \sum_{i=1}^{m} Y_{n,i}$, it follows $\bar{Y}_{n,m} \xrightarrow{a.s.} E[Y_{n,\tau}] = E[g(S_{n,\tau})] \equiv \nu$ by the law of large numbers, theorem 10.2.

10.2.1.4 Heterogeneity of the OLS and MSFE processes

In this section, we first describe the heterogeneity (i.e. the extend to which the distributions of a process X_τ may differ across τ) of the OLS and MSFE as processes with respect to the forecast origin. Second, we describe the construction of Monte Carlo simulations according to the second method described in the previous section to estimate the expected value of the OLS estimator and the MSFE. The OLS is given by $\hat{\beta}_{t,n} = (\sum_{s=t-n}^{t-1} X_s^2)^{-1} \sum_{s=t-n}^{t-1} Y_{s+1} X_s$ and the squared forecast error is given by $SFE_{t,n} = (Y_{t+1} - \hat{\beta}_{t,n} X_t)^2$.

Let $Z_t = (X_{t-n}, \ldots, X_t, Y_{t-n+1}, \ldots, Y_{t+1})$ and consider the sequence $\{Z_\tau\}_\tau$. If $\{Z_\tau\}_\tau$ is i.i.d., by proposition 10.6, $\{\hat{\beta}_{\tau,n}\}_\tau$ is i.i.d. and $\{SFE_{\tau,n}\}_\tau$ is i.i.d. By the LLN, theorem 10.2, $1/m \sum_{\tau=t}^{m} \hat{\beta}_{\tau,n} \xrightarrow{a.s.} E[\hat{\beta}_n]$ and $1/m \sum_{\tau=t}^{m} SFE_{\tau,n} \xrightarrow{a.s.} MSFE_n$. If $\{Z_\tau\}_\tau$ is stationary and ergodic, by proposition 10.7, $\{\hat{\beta}_{\tau,n}\}_\tau$ is stationary and ergodic and $\{SFE_{\tau,n}\}_\tau$ is stationary and ergodic. By the LLN theorem, 10.4, $1/m \sum_{\tau=t}^{m} \hat{\beta}_{\tau,n} \xrightarrow{a.s.} E[\hat{\beta}_n]$ and $1/m \sum_{\tau=t}^{m} SFE_{\tau,n} \xrightarrow{a.s.} MSFE_n$.

10.2.1.5 Monte Carlo simulations for the OLS and MSFE processes

Identically distributed observations

We construct Monte Carlo simulations to estimate the expected value of the OLS estimator $\hat{\beta}_{t,n}$ and the expected value of the squared forecast error; i.e. the MSFE. We employ the second method described in the previous section. The method is constructed by generating, in an identical manner, m independent series of length $n+1$ with forecast origin t for the X process and the Y process:

$$\begin{matrix} \{X_{1,t-n}, X_{1,t-n+1}, \dots, X_{1,t-1}, X_{1,t}\} \\ \vdots \\ \{X_{m,t-n}, X_{m,t-n+1}, \dots, X_{m,t-1}, X_{m,t}\} \end{matrix} \quad , \quad \begin{matrix} \{Y_{1,t-n+1}, Y_{1,t-n+2}, \dots, Y_{1,t}, Y_{1,t+1}\} \\ \vdots \\ \{Y_{m,t-n+1}, Y_{m,t-n+2}, \dots, Y_{m,t}, Y_{m,t+1}\} \end{matrix}$$

$\{X_{i,j}\}_j$ is identically distributed for $i = 1, \dots, m$ and $\{Y_{i,j}\}_j$ is identically distributed for $i = 1, \dots, m$. From these series, we construct the following vectors:

$$\begin{aligned} Z_{t,n,1} &= (X_{1,t-n}, X_{1,t-n+1}, \dots, X_{1,t-1}, X_{1,t}, Y_{1,t-n+1}, Y_{1,t-n+2}, \dots, Y_{1,t}, Y_{1,t+1}), \\ Z_{t,n,2} &= (X_{2,t-n}, X_{2,t-n+1}, \dots, X_{2,t-1}, X_{2,t}, Y_{2,t-n+1}, Y_{2,t-n+2}, \dots, Y_{2,t}, Y_{2,t+1}), \\ &\vdots \\ Z_{t,n,m} &= (X_{m,t-n}, X_{m,t-n+1}, \dots, X_{m,t-1}, X_{m,t}, Y_{m,t-n+1}, Y_{m,t-n+2}, \dots, Y_{m,t}, Y_{m,t+1}). \end{aligned}$$

The sequence of vectors $\{Z_{t,n,\tau}\}_\tau$ is i.i.d. The OLS estimator $\hat{\beta}_{t,n,\tau}$ is constructed as a measurable function from the elements of the vector $Z_{t,n,\tau}$. It follows from proposition 10.6 the sequence $\{\hat{\beta}_{t,n,\tau}\}_\tau$ is i.i.d. and by the law of large numbers $1/m \sum_{\tau=1}^{m} \hat{\beta}_{t,n,\tau} \xrightarrow{a.s.} E[\hat{\beta}_{t,n}]$. As shown in section 10.2.1.4, the expected value of the OLS is independent of the forecast origin $E[\hat{\beta}_{t,n}] = \mu_n$. In a similar manner as done for the OLS process, we construct the i.i.d. process $\{Y_{\tau,t+1} X_{\tau,t} \hat{\beta}_{t,n,\tau}\}_\tau$, the i.i.d. process $\{X_{\tau,t}^2 \hat{\beta}_{t,n,\tau}^2\}_\tau$, and the i.i.d. process $\{Y_{\tau,t+1}^2\}_\tau$. From these processes, we form the i.i.d. SFE process $\{SFE_{t,n,\tau}\}_\tau$ and from the law of large numbers $1/m \sum_{\tau=1}^{m} SFE_{t,n,\tau} \xrightarrow{a.s.} E[SFE_{t,n}] \equiv MSFE_n$ which, as shown in section 10.2.1.4, is independent of the forecast origin.

Heterogeneously distributed observations

We construct Monte Carlo simulations to estimate the expected value of the OLS estimator $\hat{\beta}_{t,n}$ and the expected value of the squared forecast error; i.e. the MSFE. We employ the second method described in section 10.2.1.3. The method is constructed by generating, in an identical manner, m independent series of length $n+1$ with forecast origin t for the X process and the Y process:

$$\begin{array}{ccc} \{X_{1,t-n}, X_{1,t-n+1}, \ldots, X_{1,t-1}, X_{1,t}\} & & \{Y_{1,t-n+1}, Y_{1,t-n+2}, \ldots, Y_{1,t}, Y_{1,t+1}\} \\ \vdots & , & \vdots \\ \{X_{m,t-n}, X_{m,t-n+1}, \ldots, X_{m,t-1}, X_{m,t}\} & & \{Y_{m,t-n+1}, Y_{m,t-n+2}, \ldots, Y_{m,t}, Y_{m,t+1}\} \end{array}$$

The following properties hold:

- For a fixed i, $\{X_{i,j}\}_j$ is heterogeneously distributed and $\{Y_{i,j}\}_j$ is heterogeneously distributed.
- For a fixed j, $\{X_{i,j}\}_i$ is identically distributed and $\{Y_{i,j}\}_i$ is identically distributed.
- For any j and k, $\{X_{i,j}\}_j$ and $\{X_{l,k}\}_k$ are independent for $i \neq l$ and $\{Y_{i,j}\}_j$ and $\{Y_{l,k}\}_k$ are independent for $i \neq l$.

From these series, we construct the following vectors:

$$\begin{gathered} Z_{t,n,1} = (X_{1,t-n}, X_{1,t-n+1}, \ldots, X_{1,t-1}, X_{1,t}, Y_{1,t-n+1}, Y_{1,t-n+2}, \ldots, Y_{1,t}, Y_{1,t+1}), \\ Z_{t,n,2} = (X_{2,t-n}, X_{2,t-n+1}, \ldots, X_{2,t-1}, X_{2,t}, Y_{2,t-n+1}, Y_{2,t-n+2}, \ldots, Y_{2,t}, Y_{2,t+1}), \\ \vdots \\ Z_{t,n,m} = (X_{m,t-n}, X_{m,t-n+1}, \ldots, X_{m,t-1}, X_{m,t}, Y_{m,t-n+1}, Y_{m,t-n+2}, \ldots, Y_{m,t}, Y_{m,t+1}). \end{gathered}$$

The sequence of vectors $\{Z_{t,n,\tau}\}_\tau$ is i.i.d. The OLS estimator $\hat{\beta}_{t,n,\tau}$ is constructed as a measurable function from the elements of the vector $Z_{t,n,\tau}$. The process $\{\hat{\beta}_{t,n,\tau}\}_\tau$ is i.i.d. and by the law of large numbers $1/m \sum_{\tau=1}^{m} \hat{\beta}_{t,n,\tau} \xrightarrow{a.s.} E[\hat{\beta}_{t,n}]$. The expected value of the OLS depends on the forecast origin and the sample size $E[\hat{\beta}_{t,n}] = \mu_{t,n}$. In a similar manner as done for the OLS process, we construct the i.i.d. process $\{Y_{\tau,t+1} X_{\tau,t} \hat{\beta}_{t,n,\tau}\}_\tau$, the i.i.d. process $\{X^2_{\tau,t} \hat{\beta}^2_{t,n,\tau}\}_\tau$, and the i.i.d. process $\{Y^2_{\tau,t+1}\}_\tau$. From these processes, we form the i.i.d. SFE process $\{SFE_{t,n,\tau}\}_\tau$ and from the law of large numbers $1/m \sum_{\tau=1}^{m} SFE_{t,n,\tau} \xrightarrow{a.s.} E[SFE_{t,n}] \equiv MSFE_{t,n}$ which depends on the forecast origin.

Example 10.1 *We consider two processes $\{X_\tau\}_\tau$ and $\{Y_\tau\}_\tau$ and construct the OLS* *estimators* $\hat{\beta}_{t,20}$ *and* $\hat{\beta}_{t,60}$ *at the forecast origin* t *as follows:*

$$\hat{\beta}_{t,20} = \Big(\sum_{\tau=t-20}^{t-1} X_\tau^2 \Big)^{-1} \sum_{\tau=t-20}^{t-1} Y_{\tau+1} X_\tau, \quad \hat{\beta}_{t,60} = \Big(\sum_{\tau=t-60}^{t-1} X_\tau^2 \Big)^{-1} \sum_{\tau=t-60}^{t-1} Y_{\tau+1} X_\tau.$$

For a fixed t, $\hat{\beta}_{t,20}$ *and* $\hat{\beta}_{t,60}$ *are two random variables. Set* $Z_{t,20} = (X_{t-20}, \ldots, X_t, Y_{t-19}, \ldots, Y_{t+1})$ *and* $Z_{t,60} = (X_{t-60}, \ldots, X_t, Y_{t-59}, \ldots, Y_{t+1})$. *If* $\{Z_{\tau,20}\}_\tau$ *is i.i.d, identically distributed or stationary,* $\{\hat{\beta}_{\tau,20}\}_\tau$ *is i.i.d, identically distributed or stationary respectively. Similarly, if* $\{Z_{\tau,60}\}_\tau$ *is i.i.d, identically distributed or stationary,* $\{\hat{\beta}_{\tau,60}\}_\tau$ *is i.i.d, identically distributed or stationary respectively. It follows that either of the two Monte Carlo methods described in section* 10.2.1.1 *can be used to estimate* $E[\hat{\beta}_{t,20}] \equiv \mu_{20}$ *or* $E[\hat{\beta}_{t,60}] \equiv \mu_{60}$.

Method 1, $\hat{\beta}_{t,20}$, $\hat{\beta}_{t,60}$ *:*

We begin by generating a series of the X *process of length* $20m$ *for* m *an integer and a series of the* Y *process of length* $20m$ *and constructing* m *OLS estimators as follows:*

$$\{X_1, \ldots, X_{20}\}, \quad \{Y_2, \ldots, Y_{21}\}, \quad \hat{\beta}_{21,20} = \Big(\sum_{\tau=1}^{20} X_\tau^2 \Big)^{-1} \sum_{\tau=1}^{20} Y_{\tau+1} X_\tau,$$

$$\{X_{21}, \ldots, X_{40}\}, \quad \{Y_{22}, \ldots, Y_{41}\}, \quad \hat{\beta}_{41,20} = \Big(\sum_{\tau=21}^{40} X_\tau^2 \Big)^{-1} \sum_{\tau=21}^{40} Y_{\tau+1} X_\tau,$$

$$\vdots$$

$$\{X_{20(m-1)+1}, \ldots, X_{20m}\}, \quad \{Y_{20(m-1)+2}, \ldots, Y_{20m+1}\},$$

$$\hat{\beta}_{20m+1,20} = \Big(\sum_{\tau=20(m-1)+1}^{20m} X_\tau^2 \Big)^{-1} \sum_{\tau=20(m-1)+1}^{20m} Y_{\tau+1} X_\tau.$$

By the law of large numbers $1/m \sum_{\tau=1}^{m} \hat{\beta}_{20\tau+1,20} \xrightarrow{a.s.} \mu_{20}$.

For $\hat{\beta}_{t,60}$, *we begin by generating a series of the* X *process of length* $60m$ *for* m *an integer* *and a series of the Y process of length $60m$ and constructing m OLS estimators as follows:*

$$
\begin{aligned}
&\{X_1, \ldots, X_{60}\}, \quad \{Y_2, \ldots, Y_{61}\}, \quad \hat{\beta}_{61,60} = \Big(\sum_{\tau=1}^{60} X_\tau^2\Big)^{-1} \sum_{\tau=1}^{60} Y_{\tau+1} X_\tau, \\
&\{X_{61}, \ldots, X_{120}\}, \quad \{Y_{62}, \ldots, Y_{121}\}, \quad \hat{\beta}_{121,60} = \Big(\sum_{\tau=61}^{120} X_\tau^2\Big)^{-1} \sum_{\tau=61}^{120} Y_{\tau+1} X_\tau, \\
&\vdots \\
&\{X_{60(m-1)+1}, \ldots, X_{60m}\}, \quad \{Y_{60(m-1)+2}, \ldots, Y_{60m+1}\}, \\
&\hat{\beta}_{60m+1,60} = \Big(\sum_{\tau=60(m-1)+1}^{60m} X_\tau^2\Big)^{-1} \sum_{\tau=60(m-1)+1}^{60m} Y_{\tau+1} X_\tau.
\end{aligned}
$$

By the law of large numbers $1/m \sum_{\tau=1}^{m} \hat{\beta}_{60\tau+1,60} \xrightarrow{a.s.} \mu_{60}$. *It is important to note* $\hat{\beta}_{21,20} + \hat{\beta}_{41,20} + \hat{\beta}_{61,20} \neq \hat{\beta}_{61,60}$ *or more generally* $\hat{\beta}_{21+60(m-1),20} + \hat{\beta}_{41+60(m-1),20} + \hat{\beta}_{60m+1,20} \neq \hat{\beta}_{60m+1,60}$. *By the law of large numbers*

$$
\begin{aligned}
&1/m \sum_{\tau=1}^{m} \hat{\beta}_{21+60(\tau-1),20} + 1/m \sum_{\tau=1}^{m} \hat{\beta}_{41+60(\tau-1),20} + 1/m \sum_{\tau=1}^{m} \hat{\beta}_{60\tau+1,20} \xrightarrow{a.s.} 3\mu_{20}, \\
&1/m \sum_{\tau=1}^{m} \hat{\beta}_{60\tau+1,60} \xrightarrow{a.s.} \mu_{60}.
\end{aligned}
$$

□

Example 10.2 *Let the forecaster observe a process* $\{Y_\tau\}_\tau$ *such that the DGP, model and forecast are as follow:*

$$
\begin{aligned}
&DGP\ : Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + U_t, \quad \{U_\tau\} \sim IIN(0,1), \\
&Model\ : Y_t = \beta Y_{t-1} + V_t, \\
&Forecast\ : \hat{Y}_{t+1,n} = \hat{\beta}_{t,n} Y_t, \quad \hat{\beta}_{t,n} = \Big(\sum_{\tau=t-n}^{t-1} Y_\tau^2\Big)^{-1} \sum_{\tau=t-n}^{t-1} Y_{\tau+1} Y_\tau.
\end{aligned}
$$

The SFE is given by $SFE_{t,n} = (Y_{t+1} - \hat{\beta}_{t,n} Y_t)^2$. *The goal is to estimate the MSFE,* $MSFE_{t,n} = E[SFE_{t,n}]$, *by the second Monte Carlo method described in section* 10.2.1.1 *and extensively described in section* 10.2.1.5 *for the OLS and the MSFE processes. Since the DGP is a strong autoregressive process, we only need to generate the following set* m

of independent series:

$$\{Y_{1,t-n}, Y_{1,t-n+1}, \ldots, Y_{1,t}, Y_{1,t+1}\},$$
$$\vdots$$
$$\{Y_{m,t-n}, Y_{m,t-n+1}, \ldots, Y_{m,t}, Y_{m,t+1}\}.$$

We form the i.i.d. sequence $\{Z_\tau\}_\tau$ *by defining the vectors:*

$$Z_1 = (Y_{1,t-n}, Y_{1,t-n+1}, \ldots, Y_{1,t}, Y_{1,t+1}),$$
$$\vdots$$
$$Z_m = (Y_{m,t-n}, Y_{m,t-n+1}, \ldots, Y_{m,t}, Y_{m,t+1}).$$

From this sequence of vectors, we construct the i.i.d. sequence $\{SFE_{t,n,\tau}\}_\tau$ *with* $SFE_{t,n,\tau} = (Y_{\tau,t+1} - \hat{\beta}_{t,n,\tau} Y_{\tau,t})^2$. *It follows* $1/m \sum_{\tau=1}^{m} SFE_{t,n,\tau} \xrightarrow{a.s.} MSFE_{t,n}$. *Since a strong autoregressive process is strictly stationary, the* $MSFE_{t,n}$ *is independent of the forecast origin* t. *Figure* 10.4 *shows the* $MSFE_{t,n}$ *as a function of the sample size* n *for different forecast origins* t *and with* $AR(3)$ *parameters* $\phi_1 = 0.1, \phi_2 = 0.3, \phi_3 = 0.5$. $\square$

10.2.2 Taylor algorithm v.s. brute force methods

In this section, we compare the performance of the Taylor algorithm to the performance of brute force methods for estimating the MSFE. As with Monte Carlo simulations, the law of large numbers is the property upon which brute force methods are constructed and can be use to approximate population quantities from sample data. Brute force methods are constructed from a single times series of data which represents one realization of the process under consideration. The brute force method is constructed from a single observed series of the explanatory variable $\{X_{t-n-m+1}, \ldots, X_t\}$, and a single observed series of the dependent variable $\{Y_{t-n-m+2}, \ldots, Y_{t+1}\}$. From these series, we construct

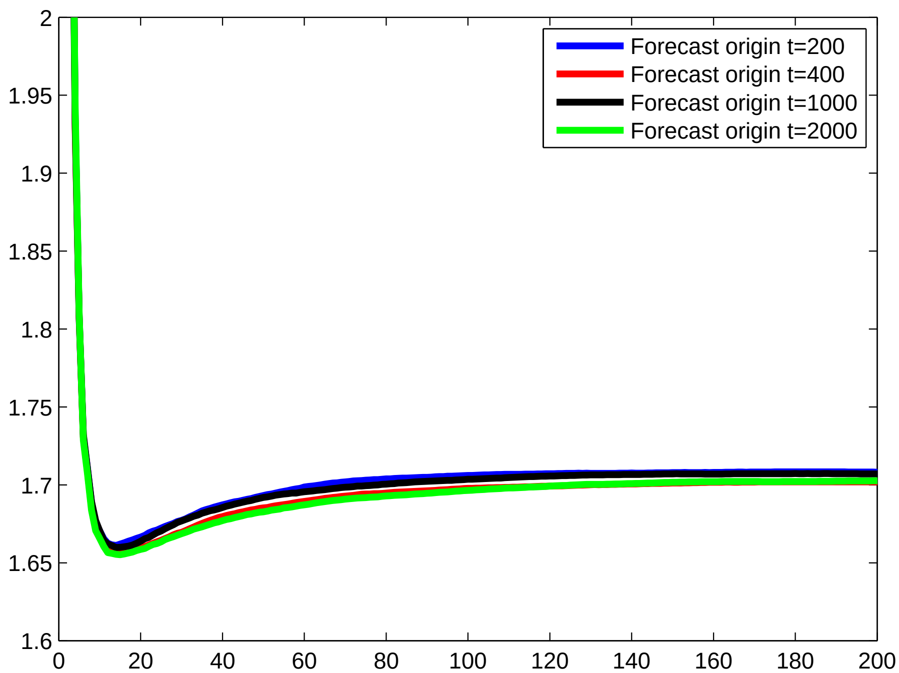


Figure 10.4: MSFE for AR(3) DGP and AR(1) model

the following m vectors:

$$\begin{aligned}
Z_1 &= (X_{t-n}, \dots, X_t, Y_{t-n+1}, \dots, Y_{t+1}), \\
Z_2 &= (X_{t-n-1}, \dots, X_{t-1}, Y_{t-n}, \dots, Y_t), \\
&\vdots \\
Z_m &= (X_{t-n-m+1}, \dots, X_{t-m+1}, Y_{t-n-m+2}, \dots, Y_{t-m+2}).
\end{aligned}$$

From these vectors, we can construct m realizations of the SFE as follows:

$$\begin{aligned}
SFE_{n,i} &= Y_{t-i+2}^2 - 2Y_{t-i+2}X_{t-i+1}\hat{\beta}_{n,i} + X_{t-i+1}^2\hat{\beta}_{n,i}^2, \quad i = 1, \dots, m, \\
\hat{\beta}_{n,i} &= \left(\sum_{\tau=t-n-i+1}^{t-i} X_\tau^2\right)^{-1} \sum_{\tau=t-n-i+1}^{t-i} Y_{\tau+1}X_\tau \quad i = 1, \dots, m.
\end{aligned}$$

n is the sample size of data used in the estimation of the OLS estimator $\hat{\beta}_{n,i}$. $n+1$ is the sample size of data used to form each of the m realizations of the SFE, $SFE_{n,i}$. If the total length of each of the series of the explanatory variables and dependent variables is N, it follows $N = n + m$ and for a given n, $m = N - m$ realizations of the SFE can be constructed. From this construction, we want to form estimates of the expected value $E[SFE_n]$, for a given n, based on the single series of data. From the law of large numbers it follows:

$$1/m \sum_{i=1}^{m} SFE_{n,i} \xrightarrow{a.s.} E[SFE_n] \equiv MSFE_n.$$

This law of large numbers holds only if the sequence $\{Z_\tau\}_\tau$ is i.i.d or strictly stationary. For this reason, brute force methods cannot be applied to covariance stationary processes or to non-stationary processes such as structural break processes. The Taylor algorithm has been developed for structural break processes with a known break time. The MSFE approximation from the brute force method is given by

$$MSFE_n \approx 1/m \sum_{i=1}^{m} SFE_{n,i}.$$

For each n, the accuracy of the approximation depends crucially on the amount of data available. For a fixed data series length N, where $N = m + n$, m is the number of SFE realizations available for averaging, $m = N - n$. Therefore, fixing N fixes and limits the largest value of n for which an estimate of the MSFE can be obtained. This value of n is $N - 1$ and $m = 1$ and the approximation consists of only one realization of the SFE. In general, as n increases, m decreases and the accuracy of the MSFE approximation worsens. In what follows, we compare the Taylor algorithm to brute force methods for forecasting problems involving i.i.d process and for forecasting problems involving strictly stationary process.

Independent identically distributed processes

We compare the Taylor algorithm and the brute force method by qualitatively analyzing the sources of error in both estimation procedures. For the case of i.i.d. process, the first source of error for the Taylor algorithm comes from the fact that the Taylor expansion is valid only inside a convergence region A but the MSFE is as follows:

$$MSFE_n = \bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A] + \bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A^c].$$

The first approximation and source of error comes from assuming the term $\bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A^c]$ is negligible and that the truncated expectation $\bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A]$ can be replaced by the expectation $E[(Y_{t+1} - \hat{Y}_{t+1})^2]$. The second source of error in the Taylor approximation comes from the remainder of the Taylor series. From numerical examples, the fourth order Taylor polynomial appears to be a very good approximation, i.e., the difference between the third order MSFE Taylor approximation and the fourth order MSFE Taylor approximation is practically zero. Therefore, the source of error from the Taylor remainder, given that the first approximation dealing with the region of convergence is acceptable, will be negligible. In numerical and empirical applications, the third source of error comes from estimating population moments with sample moments. The MSFE Taylor approximation up to fourth order was found to be as follows:

$$MSFE_n \approx \frac{1}{\omega_2^5}\left[C + \frac{A}{n} - \frac{\Delta}{n^2} + \frac{\Omega}{n^3}\right], \tag{10.2.7}$$

with $\Delta = A + 2B - D$, $\Omega = 6A - 6B - D + E$, $\omega_1 = E[Y_{t+1}X_t]$, $\omega_2 = E[X_t^2]$,

$$
\begin{aligned}
A &= \omega_1^2\omega_2^2 E[X_{t-1}^4] - 2\omega_1\omega_2^3 E[Y_t X_{t-1}^3] + \omega_2^4 E[Y_t^2 X_{t-1}^2],\\
B &= \omega_1^2\omega_2 E[X_{t-1}^6] - 2\omega_1\omega_2^2 E[Y_t X_{t-1}^5] + \omega_2^3 E[Y_t^2 X_{t-1}^4],\\
C &= E[Y_{t+1}^2]\omega_2^5 - \omega_1^2\omega_2^4,\\
D &= 9\omega_1^2 E^2[X_{t-1}^4] - 18\omega_1\omega_2 E[Y_t X_{t-1}^3]E[X_{t-1}^4] + 3\omega_2^2 E[Y_t^2 X_{t-1}^2]E[X_{t-1}^4]\\
&\quad + 6\omega_2^2 E^2[Y_t X_{t-1}^3],\\
E &= 3\omega_1^2 E[X_{t-1}^8] - 6\omega_1\omega_2 E[Y_t X_{t-1}^7] + 3\omega_2^2 E[Y_t^2 X_{t-1}^6].
\end{aligned}
$$

Approximating the MSFE by a fourth order Taylor expansion requires approximating twelve population moments, $E[Y_{t+1}^2]$, $E[X_t^2]$, $E[X_{t-1}^4]$, $E[X_{t-1}^6]$, $E[X_{t-1}^8]$, $E[Y_t X_{t-1}]$, $E[Y_t X_{t-1}^3]$, $E[Y_t^2 X_{t-1}^2]$, $E[Y_t X_{t-1}^5]$, $E[Y_t^2 X_{t-1}^4]$, $E[Y_t X_{t-1}^7]$, $E[Y_t^2 X_{t-1}^6]$, with their sample counterparts. Once the twelve approximations of the population moments have been obtained, the MSFE approximation can be given for any values of the sample size n. The brute force method, on the other hand, requires one approximation of the MSFE for every value n of the sample size. For example, if one requires approximations of the MSFE for $n = 1, 2, \ldots, 500$, the Taylor algorithm requires twelve approximations of sample moments necessary in the expression 10.2.7. The brute force method requires 500 individual approximations of the MSFE. Furthermore, approximations for the brute force method can have great deviations for different values of n as consequence of realizations resulting in small denominators of the OLS. This type of errors are not encountered in the Taylor algorithm. In the example that follows, we illustrate the trade-offs between the sources of error for the Taylor algorithm and the brute force method.

Example 10.3 *Let the forecaster observe a dependent process $\{Y_\tau\}_\tau$ and an explanatory process $\{X_\tau\}_\tau$ such that the DGP, model and forecast are as follow:*

$$
\begin{aligned}
&DGP\ : Y_{t+1} = \phi_1 X_t + \phi_2 X_t^2 + U_{t+1}, \quad \{X_\tau\} \sim IIN(10,1), \quad \{U_\tau\} \sim IIN(0,1),\\
&Model\ : Y_t = \beta X_{t-1} + V_t,\\
&Forecast\ : \hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t, \quad \hat{\beta}_{t,n} = \Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^{-1} \sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau.
\end{aligned}
$$

The SFE is given by $SFE_{t,n} = (Y_{t+1} - \hat{\beta}_{t,n}Y_t)^2$. *We generate a series of explanatory data and a series of dependent data, each of length* $N = 500$. *Figure* 10.5 *presents the benchmark MSFE generated with Monte Carlo simulations, the Taylor algorithm approximation, and the brute force method approximation. This example illustrates that the brute force method lacks robustness to the data as can be seen from the jaggedness of the MSFE and the fact that the approximation worsens as* n *increases as* m *decreases. The error of the Taylor algorithm is manifested in a shift from the Monte Carlo MSFE.* □

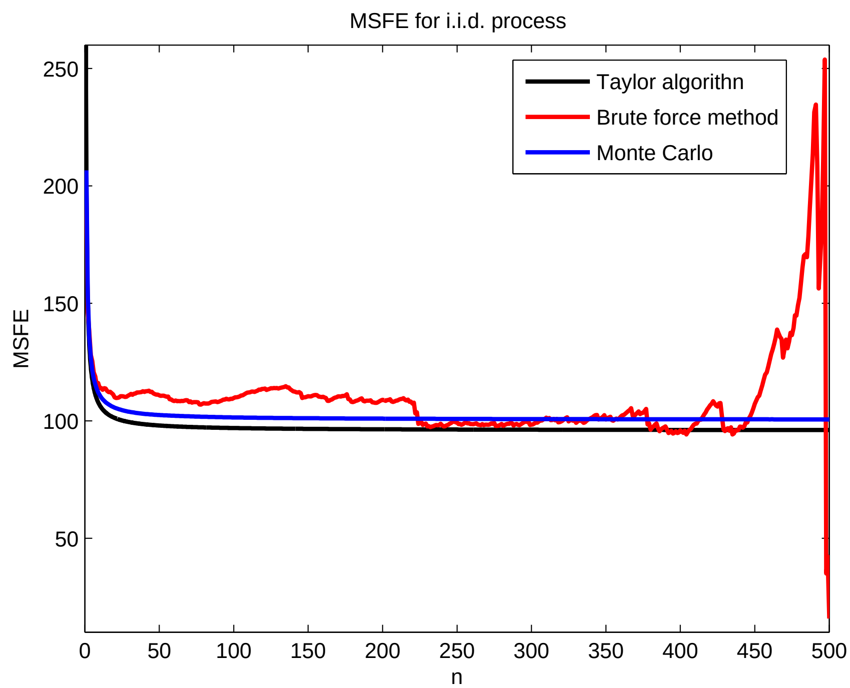


Figure 10.5: MSFE with DGP $Y_{t+1} = X_t + X_t^2 + U_{t+1}$

Stationary processes

As before, we compare the Taylor algorithm and the brute force method by qualitatively analyzing the sources of error in both estimation procedures. For the case of stationary process, the first source of error for the Taylor algorithm comes from the fact that the Taylor expansion is valid only inside a convergence region A but the MSFE is as follows:

$$MSFE_n = \bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A] + \bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A^c].$$

The first approximation and source of error for the Taylor algorithm comes from assuming the term $\bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A^c]$ is negligible and that the truncated expectation $\bar{E}[(Y_{t+1} - \hat{Y}_{t+1})^2|A]$ can be replaced by the expectation $E[(Y_{t+1} - \hat{Y}_{t+1})^2]$. As for the i.i.d. case, the second source of error in the Taylor approximation comes from the remainder of the Taylor series. In numerical examples, we investigate a second order Taylor polynomial approximation. As for the i.i.d. case, in numerical and empirical applications, the third source of error in the Taylor algorithm comes from estimating population moments with sample moments. This source of error can be more severe for the general stationary case than in the i.i.d. case due to the fact that a larger number of covariances need to be estimated. For a given n, $4n^2 + 2n + 2$ moments must be estimated. This makes the Taylor approximation computationally expensive compared to the brute force method. The brute force method requires one approximation of the MSFE for every value n of the sample size. The only advantage of the Taylor algorithm over the brute force method is that, the brute force method lacks robustness to realizations of the denominator of the OLS being close to zero. The following example illustrates the performance of both methods.

Example 10.4 *We consider the forecast problem where the DGP is generated by an* $AR(1)$ *process of the form:*

$$Y_t = \mu + \phi Y_{t-1} + U_t.$$

The forecaster applies a white noise model of the form $Y_t = \beta + V_t$*, resulting in the forecast* $\hat{Y}_{t+1} = \hat{\beta}_{t,n}$*. This problem has been shown to have the following analytic solution for the MSFE:*

$$MSFE = Var(Y_t)\left[1 + \left(1 + \frac{2\phi^{n+1}}{1-\phi}\right)\frac{1}{n} - 2\phi\left(\frac{1-\phi^n}{(1-\phi)^2}\right)\frac{1}{n^2}\right].$$

Figures 10.6*,* 10.7*, and* 10.8 *present results for values of the autoregressive parameter of* 0.1*,* 0.49*, and* 0.95*, respectively. The figures show both the approximation from the Taylor algorithm and the approximation from the brute force method. The approximation from the brute force method, as in the i.i.d. case, worsens as* n *increases because* m*, the*

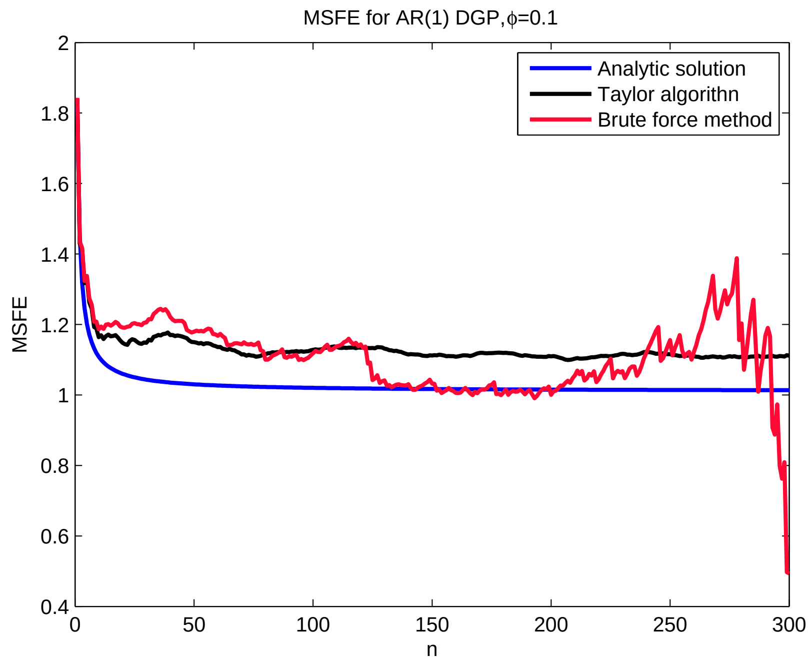


Figure 10.6: MSFE with AR(1) DGP, $\phi = 0.1$

number of SFE realizations available for averaging, decreases. $\square$

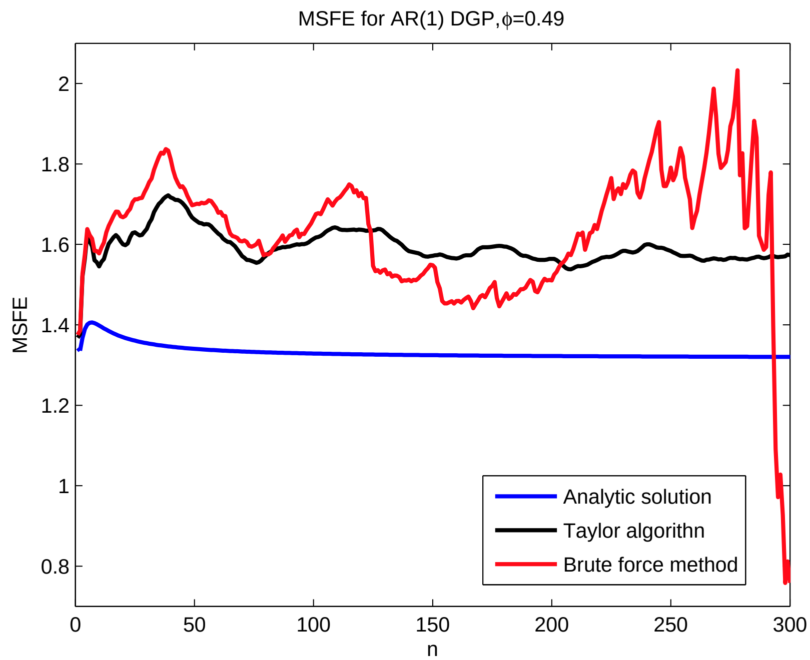


Figure 10.7: MSFE with AR(1) DGP, $\phi = 0.49$

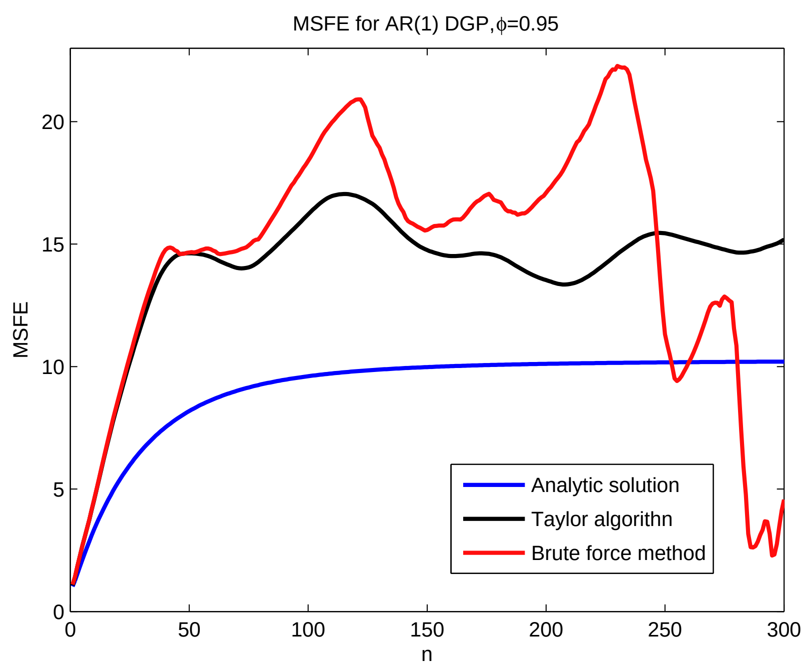


Figure 10.8: MSFE with AR(1) DGP, $\phi = 0.95$

10.2.3 Further topics on forecasting structural break processes

The Taylor algorithm to estimate the MSFE of a structural break process was developed under the assumption that the time of the break is known to the forecaster. In this section, we try to relax this assumption by investigating the situation in which the forecaster believes a break has occurred at a time $t-n_b$ but in reality no break has occurred. To do this, we compare, in the following example, the change of the the MSFE approximation as the size of the break decreases to zero.

Example 10.5 *We consider a DGP consisting of a structural break process as follows:*

$$Y_{\tau+1} = \begin{cases} \beta_1 X_\tau + U_{1,\tau+1}, & \tau \leq t-n_b \\ \beta_2 X_\tau + U_{2,\tau+1}, & \tau > t-n_b \end{cases}, \tag{10.2.8}$$

with $\beta_1, \beta_2 \in \mathbb{R}$, $Var(U_{1,\tau}) = 1$, $Var(U_{2,\tau}) = 1$, $\{X_\tau\}_\tau = IIN(10,1)$. *The forecast model is given by* $Y_{t+1} = \beta X_t + V_{t+1}$, *the forecast is given by* $\hat{Y}_{t+1,n} = \hat{\beta}_{t,n} X_t$, *where* $\hat{\beta}_{t,n}$ *is the OLS estimator of* β_2. *In this example, we examine the Taylor approximation of the MSFE for varying size of the break . The break occurs* 500 *time units in the past from the forecast origin. The moments in the Taylor approximation after the break are estimated with* 500 *data points and the moments in the Taylor approximation previous to the break are estimated with* 2500 *data points. The small amount of data used to estimate moments contributes to the error in the approximation. Figure* 10.9 *presents the results for four cases: (1)* $\beta_1 = 2.5$ *changes to* $\beta_2 = 2$, *(2)* $\beta_1 = 2.5$ *changes to* $\beta_2 = 2.3$, *(3)* $\beta_1 = 2.5$ *changes to* $\beta_2 = 2.5$, *(4)* $\beta_1 = 2.5$ *changes to* $\beta_2 = 2.8$. *The important case is (3). It represents what happens when the forecaster believes a break occurred at* $t-500$ *but in reality no break occurred. The resulting Taylor approximation of the MSFE decreases monotonically. Figure* 10.9 *also shows the benchmark MSFE if no break occurs. The difference between the benchmark MSFE and the Taylor approximation of the MSFE in case (3) is in the level of the MSFE but the shape of the MSFEs, which decrease monotonically, are similar. From this example we can conclude that, when the forecaster believes a break occurred but in reality no break occurred, the resulting MSFE Taylor approximation would decrease monotonically, i.e., the bias of the forecast error will not increase and the variance of the forecast error will decrease. Given this information, the*

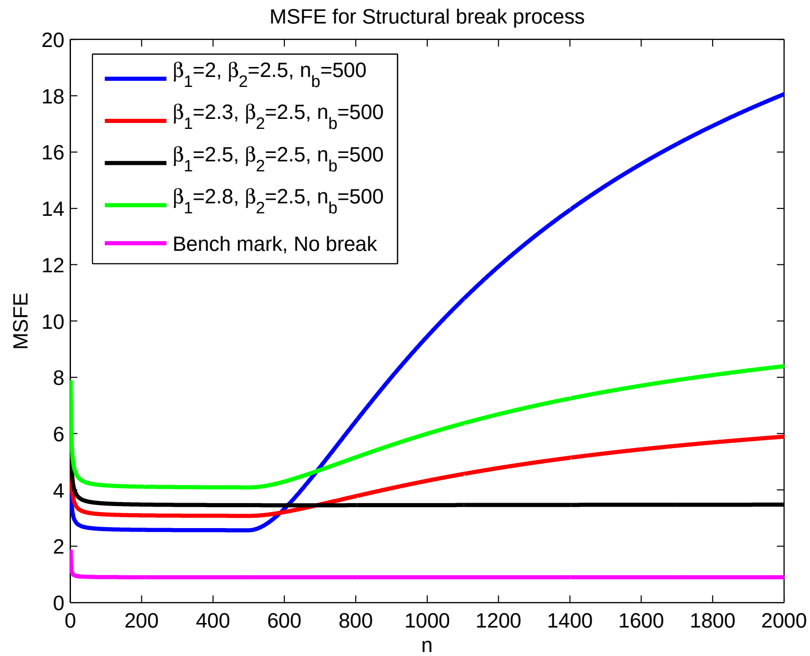


Figure 10.9: MSFE for different break sizes

forecaster can then re-evaluate her prediction of the occurrence of a structural break. □

The Taylor algorithm does not work well in the situation where the forecaster thinks no break has occurred but in reality a break has occurred. This is illustrated in Figure 10.10. The figure shows the MSFE which should result from correct prediction of the break time as well as the MSFE Taylor approximation resulting from the erroneous prediction of the break time.

10.3 Future directions

Many questions and problems are left open. For the forecasting problem with independent identically distributed processes, we described the multivariate algorithm and numerical experiments can be constructed as done for the univariate case. Furthermore, it would be important to conduct empirical studies to verify the MSFE approximation. For the forecasting problem with covariance stationary processes and with structural break processes, one would require a multivariate algorithm with corresponding numeri-

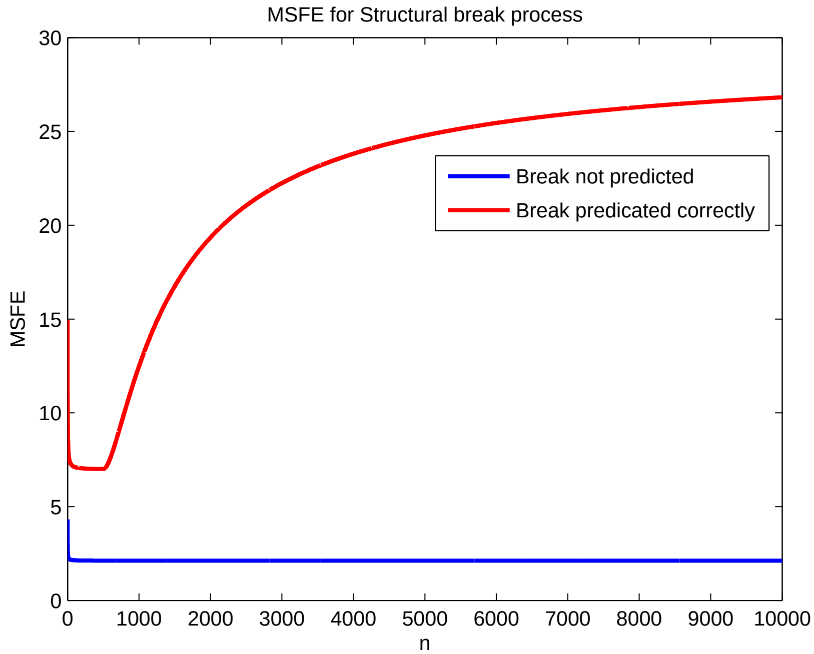


Figure 10.10: MSFE for structural break DGP with no break predicted

cal experiments and empirical studies. Another interesting problem would be to develop a similar Taylor algorithm for a forecasting problem involving covariance stationary processes that undergo a structural break. Finally, for empirical studies, the problem of forecasting volatility under misspecification can be of great interest for the finance community.

Appendix A

A.1 Identities

Many of these identities were obtained from [19].

Identity A.1 *For a nonnegative random variable X and a positive number α*

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}.$$

Identity A.2 (Markov's Inequality) *For a random variable X and a positive number α*

$$P(|X| \geq \alpha) \leq \frac{E[|X|^k]}{\alpha^k}.$$

Identity A.3 (Chebyshev-Bienaymé Inequality) *For a random variable X with $m = E[X]$ and a positive number α*

$$P(|X - m| \geq \alpha) \leq \frac{Var[X]}{\alpha^2}.$$

Identity A.4 (Jensen's Inequality) *For a random variable X with $m = E[X]$ and a convex function ϕ*

$$\phi(E[X]) \leq E[\phi(X)].$$

Identity A.5 (Hölder's Inequality) *Given*

$$\frac{1}{p} + \frac{1}{q} = 1, \quad p > 1, q > 1,$$

it follows $E[|XY|] \leq E^{1/p}[\,|X|^p]E^{1/q}[\,|Y|^q]$.

Identity A.6 (Schwarz's Inequality)

$$E[|XY|] \leq E^{1/2}[X^2]E^{1/2}[Y^2].$$

Identity A.7 (Lyapounov's Inequality)

$$E^{1/\alpha}[\,|X|^\alpha] \leq E^{1/\beta}[\,|X|^\beta], \quad 0 < \alpha \leq \beta.$$

Identity A.8 (Minkowski's Inequality) *For* $p \geq 1$,

$$E^{1/p}[\,|X+Y|^p] \leq E^{1/p}[\,|X|^p] + E^{1/p}[\,|Y|^p].$$

A.2 Asymptotic theory

The most fundamental concept for the study of non-random sequences and series is the limit.

Definition A.9 *Let* $\{a_n\}$ *be a sequence of real numbers. The number* a *is called the limit of the sequence* $\{a_n\}$ *if for every* $\delta > 0$ *there exists an integer* $N(\delta)$ *such that for all* $n \geq N(\delta)$, $|a_n - a| < \delta$.

When the limit exists, we say the sequence $\{a_n\}$ converges to a as n tends to ∞, $a = \lim_{n\to\infty} a_n$. We refer the reader to [86, 87] for a comprehensive look at deterministic sequences and series.

When considering sequences and series of random variables, there are several concepts of stochastic convergence. The setting for defining any stochastic convergence consists of a probability space $(\Omega, \mathcal{F}, P)$ and a sequence of random variables $\{X_i, i \geq 1\}$ defined on $(\Omega, \mathcal{F}, P)$. The modes of stochastic convergence which we will discuss include almost sure convergence, convergence in probability, convergence in rth mean, and convergence in distribution. We first present definitions.

Definition A.10 *Let the sequence* $\{X_n\}$ *and* X *be real valued random variables on the* *probability space* $(\Omega, \mathcal{F}, P)$. X_n converges almost surely *to* X, $X_n \stackrel{a.s.}{\rightarrow} X$, *if* $P\{\omega : X_n(\omega) \rightarrow X(\omega)\} = 1$.

Other terminology used for almost sure convergence includes convergence with probability 1, convergence almost everywhere, and strong consistency.

Definition A.11 *Let the sequence* $\{X_n\}$ *and* X *be real valued random variables on the probability space* $(\Omega, \mathcal{F}, P)$. X_n converges in probability *to* X, $X_n \stackrel{P}{\rightarrow} X$, *if* $P\{\omega : |X_n(\omega) - X(\omega)| < \varepsilon\} \rightarrow 1$ *as* $n \rightarrow \infty$.

Convergence in probability is also referred to as weak consistency or convergence in measure.

Definition A.12 *An estimator* $\hat{\theta}_n$ *of a parameter* θ *is a* consistent *estimator if and only if* $\hat{\theta}_n \stackrel{P}{\rightarrow} \theta$.

Theorem A.13 *The mean of a random sample from any population with finite population mean* μ *and finite population variance is a consistent estimator of* μ.

Proof. See [61] p.112. ■

We denote by $L^p(\Omega)$ the class of all measurable functions $f(\omega)$ such that $\int_\Omega |f(\omega)|^p dP < \infty$, $p > 0$.

Definition A.14 *Given* $X_n \in L^p(\Omega)$, $n = 1, 2, \ldots$, $p > 0$ *and* $X \in L^p(\Omega)$, X_n *converges in* L^p *to* X, $X_n \stackrel{p.m.}{\rightarrow} X$, *if* $E|X_n - X|^p \rightarrow 0$ *as* $n \rightarrow \infty$.

L^p convergence is also known as convergence in L^p-norm or convergence in pth mean. When $p = 2$ this is known as convergence in mean square and is denoted by $X_n \stackrel{m.s.}{\rightarrow} X$.

Definition A.15 *Let* $\{X_n\}$ *be a sequence of random finite-dimensional vectors with joint distribution functions* $\{F_n\}$. *If* $F_n(z) \rightarrow F(z)$ *as* $n \rightarrow \infty$ *for every continuity point* z, *where* F *is the distribution function of a random variable* Z, *then* X_n *converges in distribution to the random variable* Z, $X_n \stackrel{d}{\rightarrow} Z$.

Convergence in distribution is also known as convergence in law, $X_n \stackrel{L}{\rightarrow} Z$, or that X_n is asymptotically distributed as F, $X_n \stackrel{A}{\sim} F$.

We now present some important theorems.

Theorem A.16 *Let X_n, X and Y_n be random vectors. Then*

1. *If $X_n \xrightarrow{a.s.} X$ then $X_n \xrightarrow{P} X$.*
2. *If $X_n \xrightarrow{L_p} X$ then $X_n \xrightarrow{P} X$.*
3. *If $X_n \xrightarrow{P} X$, and $\{|X_n|^p\}_1^\infty$ is uniformly integrable, then $X_n \xrightarrow{L_p} X$.*
4. *If $X_n \xrightarrow{P} X$ then $X_n \xrightarrow{d} X$.*
5. *$X_n \xrightarrow{P} c$ for a constant c if and only if $X_n \xrightarrow{d} X$.*
6. *if $X_n \xrightarrow{d} X$ and $d(X_n, Y_n) \xrightarrow{P} 0$, then $Y_n \xrightarrow{d} X$.*
7. *if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$ for a constant c, then $(X_n, Y_n) \xrightarrow{d} (X, c)$.*
8. *if $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, then $(X_n, Y_n) \xrightarrow{P} (X, Y)$.*

Proof. See [37], p.284, 287 and [39], p. 10. ■

Theorem A.17 (Cramér's Theorem) *If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} a$ for a a constant, then*

1. *$X_n + Y_n \xrightarrow{d} X + a$.*
2. *$X_n Y_n \xrightarrow{d} aX$.*
3. *$X_n / Y_n \xrightarrow{d} X/a$, for $a \neq 0$.*

Proof. See [37], p. 355. ■

Theorem A.18 *Let $g : \mathbb{R}^k \to \mathbb{R}$ be a Borel function, let $C_g \subseteq \mathbb{R}^k$ be the set of continuity points of g, and assume $P(X \in C_g) = 1$.*

1. *If $X_n \xrightarrow{a.s.} X$ then $g(X_n) \xrightarrow{a.s.} g(X)$.*
2. *If $X_n \xrightarrow{P} X$ then $g(X_n) \xrightarrow{P} g(X)$.*
3. *If $X_n \xrightarrow{d} X$ then $g(X_n) \xrightarrow{d} g(X)$.*

Proof. See [37], p.286, 355. ■

Theorem A.19 *Let $\{X_\tau\}$ denote a sequence of $(n \times 1)$ random vectors with plim c, and let $g(c)$ be a vector-valued function, $g : \mathbb{R}^n \to \mathbb{R}^m$, where $g(\cdot)$ is continuous at c and does not depend on τ. Then $g(X_\tau) \xrightarrow{P} g(c)$.*

Proposition A.20 *Let $\{X_{1\tau}\}$ be a sequence of $(n \times n)$ random matrices with $X_{1\tau} \xrightarrow{P} C_1$, a nonsingular matrix. Let $X_{2\tau}$ denote a sequence of $(n \times 1)$ random vectors with $X_{2\tau} \xrightarrow{P} c_2$. Then $[X_{1\tau}]^{-1} X_{2\tau} \xrightarrow{P} [C_1]^{-1} c_2$.*

Proof. See [64], p. 182. ■

Theorem A.21 *Let $\{X_n\}$ and $\{Z_n\}$ be sequences of k-vectors (not necessarily converging) and g the function defined in theorem A.18, and let $P(X_n \in C_g) = P(Z_n \in C_g) = 1$ for every n.*

1. *If $||X_n - Z_n|| \xrightarrow{a.s.} 0$ then $||g(X_n) - g(Z_n)|| \xrightarrow{a.s.} 0$.*
2. *If $||X_n - Z_n|| \xrightarrow{P} 0$ then $||g(X_n) - g(Z_n)|| \xrightarrow{P} 0$.*

Proof. See [37], p. 286. ■

Theorem A.22 *Let $g : \mathbb{R}^k \to \mathbb{R}$ be a Borel function, continuous at a.*

1. *If $X_n \xrightarrow{a.s.} a$ then $g(X)_n \xrightarrow{a.s.} g(a)$.*
2. *If $X_n \xrightarrow{P} a$ then $g(X)_n \xrightarrow{P} g(a)$.*

Proof. See [37], p. 286. ■

Theorem A.23 *Let a sequence $\{Y_n\}_1^\infty$ be bounded in probability (i.e., $O_p(1)$ as $n \to \infty$); if $X_n \xrightarrow{P} 0$, then $X_n Y_n \xrightarrow{P} 0$.*

Proof. See [37], p. 287. ■

Theorem A.24 *Let $\{X_n\}_1^\infty$ be a uniformly integrable sequence. If $X_n \xrightarrow{a.s.} X$, then $EX_n \to EX$.*

Proof. See [37], p. 188. ■

Theorem A.25 *If $X_n \xrightarrow{d} X$ and $\{X_n\}$ is uniformly integrable, then $E|X| < \infty$ and $EX_n \to EX$.*

Proof. See [37], p. 357. ■

Given a sequence of random variables $\{X_i, i \geq 1\}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ and setting $S_n = \sum_{i=1}^{n} X_i$ for $i \geq 1$, the sequence $\{S_n, n \geq 1\}$ is referred to as the sequence of partial sums. Convergence almost surely of the series $\sum_{i=1}^{\infty} X_i$ is equivalent to the convergence almost surely of the sequence of partial sums

$$\sum_{i=1}^{\infty} X_i = S < \infty \quad \text{a.s.} \quad \Longleftrightarrow \quad S_n \xrightarrow{a.s.} S.$$

[136], presents results for almost sure convergence of the basic sequence $\{X_i\}$ for a variety of dependence structures.

Definition A.26 (martigale difference sequence) *A sequence of scalars $\{Y_\tau\}_{\tau=1}^{\infty}$ satisfying $E[Y_\tau] = 0$ for all τ and $E[Y_\tau | Y_{\tau-1}, Y_{\tau-2}, \ldots, Y_1] = 0$, for $\tau = 2, 3, \ldots$ is said to be a martigale difference sequence.*

Definition A.27 (L^1-Mixingale) *Consider a sequence of random variables $\{Y_\tau\}_{\tau=1}^{\infty}$ with $E[Y_\tau] = 0$ for $t = 1, 2, \cdots$ Let Ω_τ denote information available at time τ. Let $\{c_\tau\}_{\tau=1}^{\infty}$ and $\{\xi_\tau\}_{\tau=1}^{\infty}$ be sequences of nonegative deterministic constants such that $\lim_{m\to\infty} \xi_m = 0$ and $E|E[Y_\tau | \Omega_{\tau-m}]| \leq c_\tau \xi_m$ for all $t \geq 1$ and all $m \geq 0$. $\{Y_\tau\}$ is said to follow an L^1-mixingale with respect to $\{\Omega_\tau\}$. A zero-mean process for which the m-period ahead forecast $E[Y_\tau | \Omega_{\tau-m}]$ converges to the unconditional mean of zero is an L^1-mixingale.*

Proposition A.28 *Let $\{Y_\tau\}$ be a martigale difference sequence. Let $c_\tau = E|Y_\tau|$, and choose $\xi_0 = 1$ and $\xi_m = 0$ for $m = 1, 2, \ldots$. Then $\{Y_\tau\}$ is an L^1-mixingale sequence.*

Proof. See [64], p. 190. ■

Definition A.29 (uniformly integrale) *A sequence $\{Y_\tau\}$ is said to be uniformly integrale if for every $\epsilon > 0$ there exists a number $c > 0$ such that $E\left[|Y_\tau| \delta_{[|Y_\tau| \geq c]}\right] < \epsilon$ for all t, where $\delta_{[|Y_\tau| \geq c]} = 1$ if $|Y_\tau| \geq c$ and 0 otherwise.*

Proposition A.30 *Let $\{Y_\tau\}$ be an L^1-mixingale. If $\{Y_\tau\}$ is uniformly integrale and there exists a choice for $\{c_\tau\}$ such that $\lim_{T\to\infty}(1/T)\sum_{\tau=1}^{T} c_\tau < \infty$ then $(1/T)\sum_{\tau=1}^{T} Y_\tau \xrightarrow{P} 0$.*

Proof. See [4]. ■

Proposition A.31 *Let $\overline{Y}_T$ be the sample mean from a martigale difference sequence, $\overline{Y}_T = (1/T)\sum_{\tau=1}^{T} Y_\tau$ with $E|Y_\tau|^r < M'$ for some $r > 1$ and $M' < \infty$. Then $\overline{Y}_T \xrightarrow{P} 0$.*

Proof. See [64], p.191. ■

A.3 Laws of large numbers

Theorem A.32 (Kolmogorov) *Let $\{Z_s\}$ be a sequence of independent identically distributed random variables. Then $\bar{Z}_n \xrightarrow{a.s.} \mu$ if and only if $E|Z_s| < \infty$ and $E[Z_s] = \mu$.*

Proof. See [119], p. 115. ■

Proposition A.33 *Let $\mathbf{g} : \mathbb{R}^k \to \mathbb{R}^l$ be a continuous function. (i) Let $\mathbf{Z}_t$ and $\mathbf{Z}_\tau$ be identically distributed. Then $\mathbf{g}(\mathbf{Z}_t)$ and $\mathbf{g}(\mathbf{Z}_\tau)$ are identically distributed. (ii) Let $\mathbf{Z}_t$ and $\mathbf{Z}_\tau$ be independent. Then $\mathbf{g}(\mathbf{Z}_t)$ and $\mathbf{g}(\mathbf{Z}_\tau)$ are independent.*

Proposition A.34 *If $\{(\mathbf{Z}_t^\top, \mathbf{X}_t^\top, \epsilon_t)\}$ is an independent identically distributed random sequence, then $\{\mathbf{X}_t\mathbf{X}_t^\top\}$, $\{\mathbf{X}_t\epsilon_t\}$, $\{\mathbf{Z}_t\mathbf{X}_t^\top\}$, $\{\mathbf{Z}_t\epsilon_t\}$, and $\{\mathbf{Z}_t\mathbf{Z}_t^\top\}$ are also independent identically distributed sequences.*

Theorem A.35 (Markov) *Let $\{Z_t\}$ be a sequence of independent random variables, with finite means $\mu_t \equiv E[Z_t]$. If for some $\sigma > 0$, $\sum_{t=1}^{\infty}(E|Z_t - \mu_t|^{1+\delta})/t^{1+\delta} < \infty$, then $\bar{Z}_n - \bar{\mu}_n \xrightarrow{a.s.} 0$.*

Proof. See [31], pp. 125-126. ■

Corollary A.36 *Let $\{Z_t\}$ be a sequence of independent random variables such that $E|Z_t|^{1+\delta} < \Delta < \infty$ for some $\delta > 0$ and all t. Then $\bar{Z}_n - \bar{\mu}_n \xrightarrow{a.s.} 0$.*

Theorem A.37 (Ergodic theorem) *Let $\{Z_t\}$ be a stationary ergodic scalar sequence with $E|Z_t| < \infty$. Then $\bar{Z}_n \xrightarrow{a.s.} \mu \equiv E[Z_t]$.*

Proof. See [136], p. 181. ■

Theorem A.38 *Let* $\mathbf{g}$ *be a* $\mathcal{F}$*-measurable function into* $\mathbb{R}^k$ *and define* $\mathcal{Y}_t \equiv \mathbf{g}(\ldots, \mathbf{Z}_{t-1}, \mathbf{Z}_t, \mathbf{Z}_{t+1}, \ldots)$*, where* $\mathbf{Z}_t$ *is* $q \times 1$*. (i) If* $\{\mathbf{Z}_t\}$ *is stationary, then* $\{\mathcal{Y}_t\}$ *is stationary. (ii)If* $\{\mathbf{Z}_t\}$ *is stationary and ergodic, then* $\{\mathcal{Y}_t\}$ *is stationary and ergodic.*

Proof. See [136], p. 170, p. 182. ■

Proposition A.39 *If* $\{(\mathbf{Z}_t^\top, \mathbf{X}_t^\top, \epsilon_t)\}$ *is a stationary ergodic sequence, then* $\{\mathbf{X}_t\mathbf{X}_t^\top\}$*,* $\{\mathbf{X}_t\epsilon_t\}$*,* $\{\mathbf{Z}_t\mathbf{X}_t^\top\}$*,* $\{\mathbf{Z}_t\epsilon_t\}$*, and* $\{\mathbf{Z}_t\mathbf{Z}_t^\top\}$ *are stationary ergodic sequences.*

Theorem A.40 *Let* $\mathbf{g}$ *be a measurable function into* $\mathbb{R}^k$ *and define*

$$\mathcal{Y}_t \equiv \mathbf{g}(\mathbf{Z}_t, \mathbf{Z}_{t+1}, \ldots, \mathbf{Z}_{t+\tau}),$$

where τ *is finite. If the sequence of* $q \times 1$ *vectors* $\{\mathbf{Z}_t\}$ *is* ϕ*-mixing (*α*-mixing) of size* $-a$*,* $a > 0$*, then* $\{\mathcal{Y}_t\}$ ϕ*-mixing (*α*-mixing) of size* $-a$*,* $a > 0$*.*

Proof. See [151] (Lemma 2.1). ■

Appendix B

Appendix for Chapter 4

The following theorems, corollaries, propositions, and their proofs can be found in [31, 119, 153].

B.1 Random power series

A random power series is a power series with some of its components represented by random variables. In the literature, much attention has been given to the scenario with a probability space $(\Omega, \mathcal{F}, P)$ and an arbitrary sequence $\{a_n(\omega)\}_{n=0}^{\infty}$ of complex-valued random variables defined on it such that the series

$$\sum_{n=0}^{\infty} a_n(\omega) z^n, \tag{B.1.1}$$

with z an element of the complex plane C, is called a random power series. These are not the series of interest to us but for the interested reader we refer to the many expositions on the subject, [12, 81, 105].

We are interested in the setting consisting of a probability space $(\Omega, \mathcal{F}, P)$ and an arbitrary sequence $\{X_n(\omega)\}_{n=0}^{\infty}$ of random variables defined on it. We look at the power series given by

$$\sum_{n=0}^{\infty} c_n X_n(\omega)^n, \tag{B.1.2}$$

with $\{c_n\}$ a sequence of real constants.

The methods needed to study the convergence properties of $(B.1.1)$ and $(B.1.2)$ are quite different. This difference originates from the well known fact that the convergence of a power series depends on the limit of the coefficient series $\{a_n(\omega)\}$ in the case of $(B.1.1)$, and $\{c_n\}$ in the case of $(B.1.2)$. Since $\{c_n\}$ is a sequence of deterministic constants, the convergence of $(B.1.2)$ can be studied as its deterministic counterpart. We present the most important theorems for power series.

Theorem B.1 *Let $\sum c_k X^k$ be an arbitrary power series, and set $\limsup \sqrt[k]{c_k} = \alpha$. Then*

1. *for $\alpha = 0$, the series converges for all X.*
2. *for $\alpha = +\infty$, the series is divergent for every $z \neq 0$.*
3. *for $0 < \alpha < +\infty$ the series is absolutely convergent for every X with $|X| < r = 1/\alpha$, divergent for every X with $|X| > r$.*

Proof. See [86], p. 99. ■

When considering a power series $\sum c_k (X - X_0)^k$ with radius of convergence not equal to 0, the series is absolutely convergent for every X with $|X - X_0| < r$. Its value is a function of X and denoted $\phi(X)$, and we say the power series represents the function $\phi(X)$, or conversely, that the function $\phi(X)$ is expanded in a power series. We now present some theorems regarding such functions.

Theorem B.2 *The function represented by a power series is continuous at the center X_0 of its circle of convergence.*

Proof. See [86], p. 102. ■

Theorem B.3 *Let $\sum c_k X^k$ be a power series with positive radius r. If X_1 is an interior point of its circle of convergence, then the function $\phi(X)$ represented by this series can also be expanded in a power series*

$$\phi(X) = \sum_{k=0}^{\infty} b_k (X - X_1)^k, \tag{B.1.3}$$

in a neighborhood of X_1. Every coefficient b_k is represented by the absolutely convergent *series*

$$b_k = \sum_{v=0}^{\infty} \begin{pmatrix} k+v \\ v \end{pmatrix} c_{k+v} X_1^v, \quad k = 0, 1, \ldots,$$

which, regarded as a power series, again has the exact radius r*. Furthermore, the radius* r_1 *of* $(B.1.3)$ *is at least equal to* $r - |X_1|$.

Proof. See [86], p. 105. ■

Theorem B.4 *A function represented by a power series* $\sum c_k X^k$ *is continuous at every interior point of its circle of convergence.*

Proof. See [86], p. 107. ■

Theorem B.5 *A function represented by a power series is differentiable arbitrarily often at every interior point of its circle of convergence, and its derivatives may be obtained by term-by-term differentiation.*

Proof. See [86], p. 107. ■

Corollary B.6 *Given a function represented by a power series with a radius of convergence* r, $\phi(X) = \sum_{k=0}^{\infty} c_k (X - X_0)^k$*, then* $c_k = \frac{1}{k!}\phi^{(k)}(X_0)$.

Proof. See [86], p. 108. ■

B.2 Theorems

Theorem B.7 (Ratio Test) *Given a series* $\sum a_n$ *of nonzero complex terms, let*

$$r = \lim_{n\to\infty} \inf \left| \frac{a_{n+1}}{a_n} \right|, \quad R = \lim_{n\to\infty} \sup \left| \frac{a_{n+1}}{a_n} \right|.$$

1. *The series* $\sum a_n$ *converges absolutely if* $R < 1$.

2. *The series* $\sum a_n$ *diverges if* $r > 1$.

3. *The test is inconclusive if* $r \leq 1 \leq R$.

Proof. See [7], p. 193. ■

Theorem B.8 (Comparison Test) *If $a_n > 0$ and $b_n > 0$ for $n = 1, 2, \ldots,$ and if there exists positive constants c and N such that $a_n < cb_n, \quad for \quad n \geq N$, then convergence of $\sum b_n$ implies convergence of $\sum a_n$.*

Proof. See [7], p. 190. ■

Theorem B.9 *If $\sum_n f_n$ converges almost everywhere and $|\sum_{k=1}^n f_k| \leq g$ almost everywhere, where g is integrable, then $\sum_n f_n$ and the f_n are integrable and $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$.*

Theorem B.10 *If $\sum_n \int |f_n| d\mu < \infty$, then $\sum_n f_n$ converges absolutely almost everywhere and is integrable, and $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$.*

Proof. See [19], corollary to theorem 16.7, p. 211. ■

Theorem B.11 *Let $\{X_n\}_1^\infty$ be a uniformly integrable sequence. If $X_n \overset{a.s.}{\to} X$, then $EX_n \to EX$.*

Proof. See [37], theorem 12.8, p. 188. ■

Appendix C

Appendix for Chapter 5

C.1 Convergence and probability sets

When applying the Taylor series approximation method developed in Chapters 5, 6, and 7, two sets are of importance; a convergence set and a probability set. For the approximation of the expectation of a function of random variables or statistic by means of a Taylor series, the convergence set describes the region where the Taylor series converges. If B is such a convergence set, the expectation of a function f of n random variables $X_1, \ldots, X_n$, with domain $\mathbb{R}^n$, can be written with truncated expectations $E[f(X_1, \ldots, X_n)] = \bar{E}[f(X_1, \ldots, X_n), B] + \bar{E}[f(X_1, \ldots, X_n), B^c]$, where $B \cup B^c = \mathbb{R}^n$. The approximation of interest is $E[f(X_1, \ldots, X_n)] \approx \bar{E}[f(X_1, \ldots, X_n), B] \approx \bar{E}[Q(f, m), B]$ where $Q(f, m)$ is the mth order Taylor polynomial of f. There might be situations in which the convergence set B is such that the truncated expectations are difficult to calculate. In such cases, we are interested in defining a probability set A. The probability set A is a region of the domain of the random variables $\mathbb{R}^n$ chosen to ease the calculation of truncated expectations. For the Taylor series approximation method to work, the probability set must be a subset of the convergence set $A \subseteq B$.

C.1.1 Convergence set for the approximation of the OLS

We begin by assuming $|E[\hat{\beta}_{t,n}]| < \infty$ and consider the scalar case $k = 1$. Recall the OLS for the forecasting problem described in Chapter 5 as given by (5.4.2) and (5.4.6) is as

follows:

$$\hat{\beta}_{t,n} = \left(\sum_{\tau=t-n}^{t-1} X_\tau X_\tau' \right)^{-1} \left(\sum_{\tau=t-n}^{t-1} X_\tau Y_{\tau+1} \right) = S_{2,n}^{-1} S_{1,n}.$$

In Chapter 5 we described the approximation of the expectation of $\hat{\beta}_{t,n}$ by means of a Taylor series with respect to the variables $S_{1,n}$ and $S_{2,n}$ about the points ω_1 and ω_2. The approximation given is $E[\hat{\beta}_{t,n}] \approx \bar{E}[Q(\hat{\beta}_{t,n}, M), A]$, where $Q(\hat{\beta}_{t,n}, M)$ is the Mth order Taylor polynomial of $\hat{\beta}_{t,n}$. Presently, we are interested in determining the set A of the truncated expectation involved in the approximation. To do so, we assume the random variable Y_t depends on the mutually independent processes $\{X_\tau\}$ and $\{U_\tau\}$ for $\tau < t$. It follows $\hat{\beta}_{t,n}$ is a function of the random variables $X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}$. Next, we write the decomposition (5.4.8) as follows:

$$\begin{aligned}
E[\hat{\beta}_{t,n}] = & \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \hat{\beta}_{t,n}(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) \\
& \cdot f(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) dX_{t-n} \ldots dX_{t-1} dU_{t-n} \ldots dU_{t-1} \\
= & \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \Big[\int_{I_1} \cdots \int_{I_n} \hat{\beta}_{t,n}(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) f_1(X_{t-n}, \ldots, X_{t-1}) \\
& \cdot dX_{t-n} \ldots dX_{t-1} + \int_{I_1^c} \cdots \int_{I_n^c} \hat{\beta}_{t,n}(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) \\
& \cdot f_1(X_{t-n}, \ldots, X_{t-1}) dX_{t-n} \ldots dX_{t-1} \Big] f_2(U_{t-n}, \ldots, U_{t-1}) dU_{t-n} \ldots dU_{t-1} \\
= & \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{I_1} \cdots \int_{I_n} \hat{\beta}_{t,n}(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) f_1(X_{t-n}, \ldots, X_{t-1}) \\
& \cdot f_2(U_{t-n}, \ldots, U_{t-1}) dX_{t-n} \ldots dX_{t-1} dU_{t-n} \ldots dU_{t-1} \\
& + \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{I_1^c} \cdots \int_{I_n^c} \hat{\beta}_{t,n}(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) f_1(X_{t-n}, \ldots, X_{t-1}) \\
& \cdot f_2(U_{t-n}, \ldots, U_{t-1}) dX_{t-n} \ldots dX_{t-1} dU_{t-n} \ldots dU_{t-1} \\
= & \bar{E}[\hat{\beta}_{t,n}, A] + \bar{E}[\hat{\beta}_{t,n}, A^c],
\end{aligned}$$

where f is the joint distribution of the random variables $X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}$, f_1 is the joint distribution of the random variables $X_{t-n}, \ldots, X_{t-1}$, f_2 is the joint distribution of the random variables $U_{t-n}, \ldots, U_{t-1}$, I_i is an interval in $\mathbb{R}$ for $i = 1, \ldots, n$, and I_i^c is the respective complement. From the above development, it follows $A = \mathbb{R}^n \times I_1 \times \cdots \times I_n$

and $A^c = \mathbb{R}^n \times I_1^c \times \cdots \times I_n^c$. The objective is to specify the intervals I_i for $i = 1, \ldots, n$ in a manner such that the Taylor series of $\hat{\beta}_{t,n}$ converges in the region $I_1 \times \cdots \times I_n$. The Taylor series of $\hat{\beta}_{t,n}$ will converge in the set

$$B = \{(S_{1,n}, S_{2,n}) : 0 < S_2 < 2\omega_2\}.$$

Since $S_{1,n}$ is a function of the random variables $X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}$, and $S_{2,n}$ is a function of the random variables $X_{t-n}, \ldots, X_{t-1}$, it follows the set B can be rewritten as follows:

$$B = \{(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) \in \mathbb{R}^{2n} : 0 < \sum_{\tau=t-n}^{t-1} X_\tau^2 < 2n\omega_2\}. \tag{C.1.1}$$

The integrals involving the intervals I_i for $i = 1, \ldots, n$ are parametrized such that the volume enclosed coincides with the hyper-sphere $\sum_{\tau=t-n}^{t-1} X_\tau^2 = 2n\omega_2$ less the origin. For this, we make use of the following hyper-spherical coordinates:

$$\begin{pmatrix} X_{t-1} \\ X_{t-2} \\ \vdots \\ X_{t-k} \\ \vdots \\ X_{t-n-1} \\ X_{t-n} \end{pmatrix} = \begin{pmatrix} r\cos\phi_1 \\ r\sin\phi_1\cos\phi_2 \\ \vdots \\ r\left(\prod_{i=1}^{k-1}\sin\phi_i\right)\cos\phi_k \\ \vdots \\ r\sin\phi_1\sin\phi_2\cdots\sin\phi_{n-2}\cos\theta \\ r\sin\phi_1\sin\phi_2\cdots\sin\phi_{n-2}\sin\theta \end{pmatrix},$$

where $\phi_i \in [0, \pi]$ for $i = 1, \ldots, n-2$ are polar angles and $\theta \in [0, 2\pi)$ is the azimuthal angle. The transformation can be carried out with the following differential relations:

$$\begin{aligned}
dX_{t-1} &= \cos\phi_1 dr - r\sin\phi_1 d\phi_1, \\
dX_{t-2} &= \sin\phi_1\cos\phi_2 dr + r\cos\phi_1\cos\phi_2 d\phi_1 - r\sin\phi_1\sin\phi_2 d\phi_2, \\
dX_{t-k} &= \left(\prod_{i=1}^{k-1}\sin\phi_i\right)\cos\phi_k dr + r\sum_{i=1}^{k-1}\cos\phi_i\left(\prod_{j\neq i}\sin\phi_j\right)\cos\phi_k d\phi_i
\end{aligned}$$

$$
\begin{aligned}
&\quad - r\left(\prod_{i=1}^{k-1}\sin\phi_i\right)\sin\phi_k d\phi_k,\\
dX_{t-n-1} &= \left(\prod_{i=1}^{n-2}\sin\phi_i\right)\cos\theta dr + r\sum_{i=1}^{n-2}\cos\phi_i\left(\prod_{j\neq i}\sin\phi_j\right)\cos\theta d\phi_i\\
&\quad - r\left(\prod_{i=1}^{n-2}\sin\phi_i\right)\sin\theta d\theta,\\
dX_{t-n} &= \left(\prod_{i=1}^{n-2}\sin\phi_i\right)\sin\theta dr + r\sum_{i=1}^{n-2}\cos\phi_i\left(\prod_{j\neq i}\sin\phi_j\right)\sin\theta d\phi_i\\
&\quad + r\left(\prod_{i=1}^{n-2}\sin\phi_i\right)\cos\theta d\theta.
\end{aligned}
$$

The integrals involving the intervals I_i for $i = 1, \ldots, n$ are replaced by integrals involving the variables $\phi_1, \ldots, \phi_{n-2}, \theta, r$ with respective intervals $\phi_i \in [0, \pi]$ for $i = 1, \ldots, n-2$, $\theta \in [0, 2\pi)$ and $r \in [\delta, \sqrt{2n\omega_2}\,]$. The above analysis follows exactly for the approximation of $E[\hat{\beta}^2_{t,n}]$ by $\bar{E}[\hat{\beta}^2_{t,n}, A]$ where A is the same set.

C.1.2 Probability sets for the approximation of the OLS

Of particular interest is the probability set A, defined as follows:

$$
\begin{aligned}
&A \equiv \{(X_{t-n}, \ldots, X_{t-1}, U_{t-n}, \ldots, U_{t-1}) \in \mathbb{R}^{2n} : X_{t-n} \in I_{t-n}, \ldots, X_{t-1} \in I_{t-1}\},\\
&I_i = [E[X_i] - \delta_i, E[X_i] + \delta_i] \in \mathbb{R} \quad \text{for} \quad i = t-n, \ldots, t-1,
\end{aligned}
$$

such that $A \subset B$ where B is as defined in $(C.1.1)$. To determine if $A \subset B$, it suffices to show $\bar{A} \subset \bar{B}$ where

$$
\begin{aligned}
\bar{A} &= \{(X_{t-n}, \ldots, X_{t-1}) \in \mathbb{R}^n : X_{t-n} \in I_{t-n}, \ldots, X_{t-1} \in I_{t-1}\},\\
\bar{B} &= \{(X_{t-n}, \ldots, X_{t-1}) \in \mathbb{R}^n : \sum_{\tau=t-n}^{t-1} X_\tau^2 \leq 2n\omega_2\}.
\end{aligned}
$$

The center of the polytope $\bar{A}$ is the point $\mu = (\mu_{t-n}, \ldots, \mu_{t-1})$ with $\mu_i = E[X_i]$ for $i = t-n, \ldots, t-1$. The distance between the origin and μ is $r = \sqrt{\mu_{t-n}^2 + \cdots + \mu_{t-1}^2}$. The radius of the hypersphere $\bar{B}$ is $R = \sqrt{2n\omega_2}$. Clearly $\mu \in \bar{B}$, since $r = \sqrt{\omega_2} < R$.

Without loss of generality, we assume $\mu > 0$. We are interested in giving conditions on δ_i to ensure the polytope $\bar{A}$ is the largest polytope completely contained in the hyper-sphere $\bar{B}$. The square distance from μ to the closest point on the hyper-sphere is given by the following optimization problem:

$$s = \min_{X \in \bar{B}} \sum_{i=t-n}^{t-1} (X_i - \mu_i)^2,$$

and the point on $\bar{B}$ nearest to μ is as follows:

$$v = \left(\frac{\mu_{t-n} R}{r}, \ldots, \frac{\mu_{t-1} R}{r} \right).$$

The largest polytope centered at μ completely contained in the hyper-sphere $\bar{B}$ is the polytope with $\sqrt{s}$ as the largest distance from its center. For the case $n = 2$, the largest polytopes are rectangles, as shown in Figure $C.1$. For a general n and the case $\mu_{t-n} = \cdots = \mu_{t-1}$, the polytope is a hyper-cube and the nearest point to μ on $\bar{B}$ is

$$v = (\bar{x}_{t-n}, \ldots, \bar{x}_{t-1}), \quad \bar{x}_i = \sqrt{2(\mu_{t-1}^2 + \sigma_{t-1}^2)}, \quad \text{for} \quad i = t-n, \ldots, t-1,$$

and $\delta_i = \bar{x}_i - \mu_i$ for $i = t-n, \ldots, t-1$. The interval I_i is centered at the mean μ_i and has width $2\delta_i$. We are interested in understanding the probability $P(X_i \in I_i)$. To do this, we examine the following ratio:

$$\frac{\delta_i}{\sigma_i} = \sqrt{2\left(\frac{\mu_i^2}{\sigma_i^2} + 1\right)} - \frac{\mu_i}{\sigma_i}.$$

The limit of δ_i/σ_i as $\sigma_i \to \infty$ is $\sqrt{2}$, and the minimum is $\delta_i/\sigma_i = 1$, which occurs at $\sigma_i = \mu_i$. Figure $C.2$ shows δ_i/σ_i as a function of σ_i for three different values of μ_i. Figure $C.2$ also shows the probability $P(X_i \in I_i)$ as a function of σ_i for three different values of μ_i.

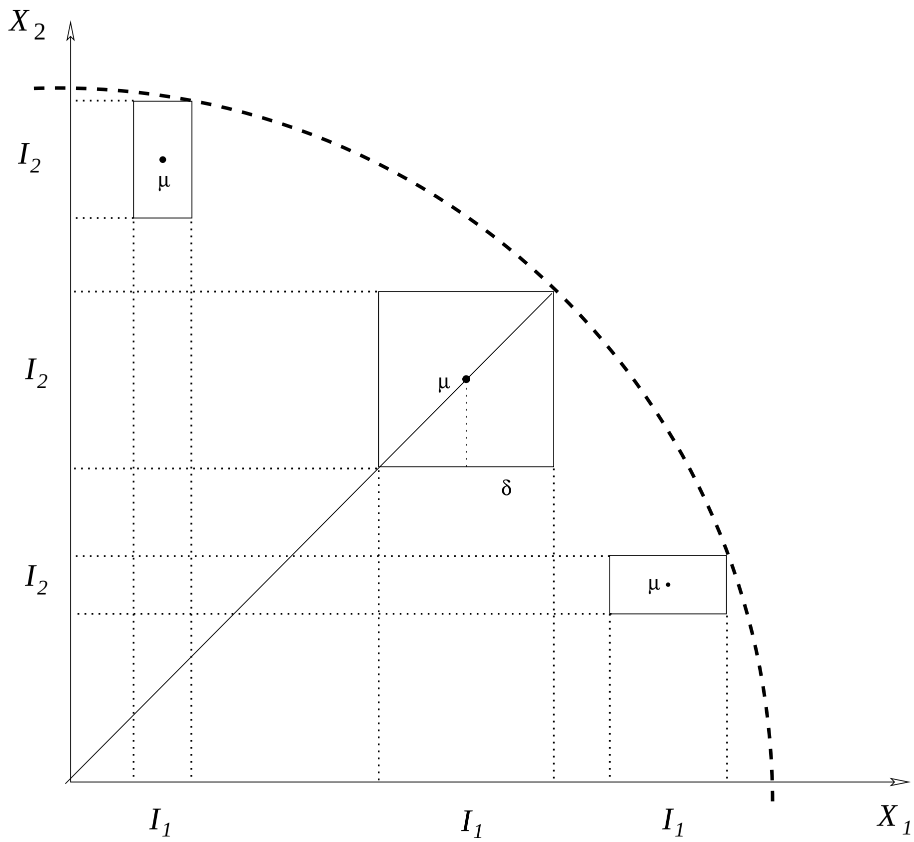
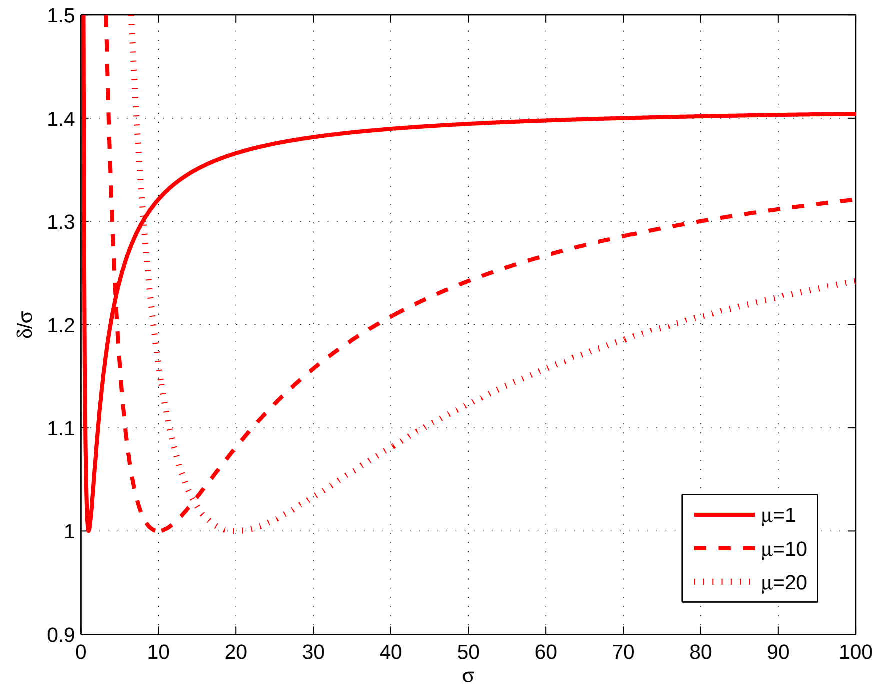


Figure C.1: Probability sets for $n = 2$



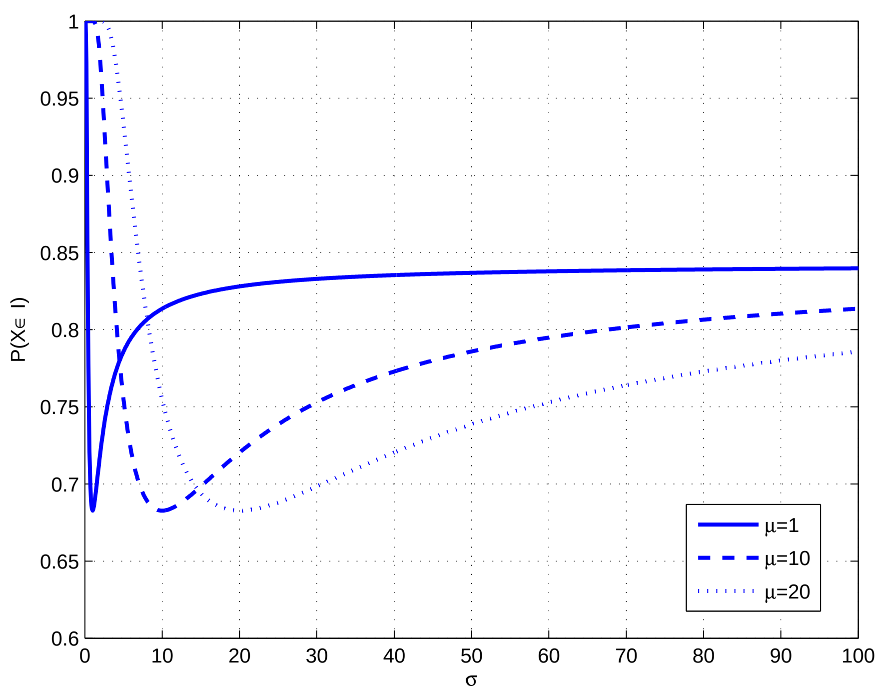


Figure C.2: The case $\mu_{t-n} = \cdots = \mu_{t-1}$

C.2 Expansion of truncate central moments for the scalar problem

We begin by expanding powers and products of the statistics $S_{1,n}$ and $S_{2,n}$ and the corresponding truncated expectations.

- $\bar{E}[S_{1,n}, A] = \frac{1}{n}\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau, A\Big] = \bar{E}\left[Y_t X_{t-1}, A\right]$
- $\bar{E}[S_{2,n}, A] = \frac{1}{n}\bar{E}\Big[\sum_{\tau=t-n}^{t-1} X_\tau^2, A\Big] = \bar{E}\left[X_{t-1}^2, A\right].$

Next we expand $\bar{E}[S_{1,n}^2, A]$:

$$S_{1,n}^2 = \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)^2 = \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^2 + \sum_{i\neq j} Y_{i+1}X_i Y_{j+1}X_j\right).$$

The truncated expectation of the two terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^2, A\Big] = n\bar{E}\left[Y_t^2 X_{t-1}^2, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i Y_{j+1}X_j, A\Big] = (n^2-n)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}, A\right].$$

The truncated expectation of $S_{1,n}^2$ is as follows:

- $\bar{E}[S_{1,n}^2, A] = \frac{1}{n}\bar{E}\left[Y_t^2 X_{t-1}^2, A\right] + \left(1-\frac{1}{n}\right)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}, A\right].$

Next we expand $\bar{E}[S_{2,n}^2, A]$:

$$S_{2,n}^2 = \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^2 = \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} X_\tau^4 + \sum_{i\neq j} X_i^2 X_j^2\right).$$

The truncated expectation of the two terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} X_\tau^4, A\Big] = n\bar{E}\left[X_{t-1}^4, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} X_i^2 Y_{j+1} X_j^2, A\Big] = (n^2-n)\bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right].$$

The truncated expectation of $S_{2,n}^2$ is as follows:

- $\bar{E}[S_{2,n}^2, A] = \frac{1}{n}\bar{E}\left[X_{t-1}^4, A\right] + \left(1-\frac{1}{n}\right)\bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right].$

Next we expand $\bar{E}[S_{1,n}S_{2,n}, A]$:

$$\begin{aligned} S_{1,n}S_{2,n} =& \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right) \\ =& \frac{1}{n^2}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^3 + \sum_{i\neq j} Y_{i+1}X_i X_j^2\right). \end{aligned}$$

The truncated expectation of the two terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^3, A\Big] = n\bar{E}\left[Y_t X_{t-1}^3, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i X_j^2, A\Big] = (n^2-n)\bar{E}\left[Y_t X_{t-1} X_{t-2}^2, A\right].$$

The truncated expectation of $S_{1,n}S_{2,n}$ is as follows:

- $\bar{E}[S_{1,n}S_{2,n}, A] = \frac{1}{n}\bar{E}\left[Y_t X_{t-1}^3, A\right] + \left(1-\frac{1}{n}\right)\bar{E}\left[Y_t X_{t-1} X_{t-2}^2, A\right].$

Next we expand $\bar{E}[S_{2,n}^3, A]$:

$$S_{2,n}^3 = \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^3 = \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} X_\tau^6 + \sum_{i\neq j} X_i^4 X_j^2 + \sum_{i\neq j\neq k} X_i^2 X_j^2 X_k^2\right).$$

The truncated expectation of the three terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} X_\tau^6, A\Big] = n\bar{E}\left[X_{t-1}^6, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} X_i^4 X_j^2, A\Big] = 3(n^2-n)\bar{E}\left[X_{t-1}^4 X_{t-2}^2, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k} X_i^2 X_j^2 X_k^2, A\Big] = (n^3-3n^2+2n)\bar{E}\left[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A\right].$$

The truncated expectation of $S_{2,n}^3$ is as follows:

- $$\bar{E}[S_{2,n}^3, A] = \frac{1}{n^2}\bar{E}\left[X_{t-1}^6, A\right] + 3\left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}\left[X_{t-1}^4 X_{t-2}^2, A\right] + \left(1-3\frac{1}{n}+2\frac{1}{n^2}\right)\bar{E}\left[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A\right].$$

Next we expand $\bar{E}[S_{1,n}S_{2,n}^2, A]$:

$$\begin{aligned} S_{1,n}S_{2,n}^2 =& \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^2 \\ =& \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^5 + \sum_{i\neq j} Y_{i+1}X_iX_j^4 + \sum_{i\neq j} Y_{i+1}X_i^3X_j^2 + \sum_{i\neq j\neq k} Y_{i+1}X_iX_j^2X_k^2\right). \end{aligned}$$

The truncated expectation of the four terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^5, A\Big] = n\bar{E}\left[Y_tX_{t-1}^5, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_iX_j^4, A\Big] = (n^2-n)\bar{E}\left[Y_tX_{t-1}X_{t-2}^4, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i^3X_j^2, A\Big] = 2(n^2-n)\bar{E}\left[Y_tX_{t-1}^3X_{t-2}^2, A\right],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_iX_j^2X_k^2, A\Big] = (n^3-3n^2+2n)\bar{E}\left[Y_tX_{t-1}X_{t-2}^2X_{t-3}^2, A\right].$$

The truncated expectation of $S_{1,n}S_{2,n}^2$ is as follows:

- $$\bar{E}[S_{1,n}S_{2,n}^2, A] = \frac{1}{n^2}\bar{E}\left[Y_tX_{t-1}^5, A\right] + \left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}\left[Y_tX_{t-1}X_{t-2}^4, A\right]$$

$$+ 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1}^3 X_{t-2}^2, A\right]$$
$$+ \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A\right].$$

Next we expand $\bar{E}[S_{1,n}^2 S_{2,n}, A]$:

$$\begin{aligned} S_{1,n}^2 S_{2,n} =& \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)^2\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right) \\ =& \frac{1}{n^3}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^4 + \sum_{i\neq j} Y_{i+1}^2 X_i^2 X_j^2 + \sum_{i\neq j} Y_{i+1}X_i Y_{j+1} X_j^3 \right. \\ & \left. + \sum_{i\neq j\neq k} Y_{i+1}X_i Y_{j+1} X_j X_k^2\right). \end{aligned}$$

The truncated expectation of the four terms are as follows:

$$\begin{aligned} &\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^4, A\Big] = n\bar{E}\left[Y_t^2 X_{t-1}^4, A\right], \\ &\bar{E}\Big[\sum_{i\neq j} Y_{i+1}^2 X_i^2 X_j^2, A\Big] = (n^2-n)\bar{E}\left[Y_t^2 X_{t-1}^2 X_{t-2}^2, A\right], \\ &\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i Y_{j+1} X_j^3, A\Big] = 2(n^2-n)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}^3, A\right], \\ &\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_i Y_{j+1} X_j X_k^2, A\Big] = (n^3-3n^2+2n)\bar{E}\Big[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2, A\Big]. \end{aligned}$$

The truncated expectation of $S_{1,n}^2 S_{2,n}$ is as follows:

- $$\begin{aligned} \bar{E}[S_{1,n}^2 S_{2,n}, A] =& \frac{1}{n^2}\bar{E}\left[Y_t^2 X_{t-1}^4, A\right] + \left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t^2 X_{t-1}^2 X_{t-2}^2, A\right] \\ & + 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}^3, A\right] \\ & + \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2, A\right]. \end{aligned}$$

Next we expand $\bar{E}[S_{2,n}^4, A]$:

$$S_{2,n}^4 = \frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^4$$

$$
\begin{aligned}
=&\frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} X_\tau^8 + \sum_{i\neq j} X_i^6 X_j^2 + \sum_{i\neq j} X_i^4 X_j^4 + \sum_{i\neq j\neq k} X_i^4 X_j^2 X_k^2\right.\\
&\left.+ \sum_{i\neq j\neq k\neq l} X_i^2 X_j^2 X_k^2 X_l^2\right).
\end{aligned}
$$

The truncated expectation of the five terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n}^{t-1} X_\tau^8, A\Big] = n\bar{E}[X_{t-1}^8, A],\\
&\bar{E}\Big[\sum_{i\neq j} X_i^6 X_j^2, A\Big] = 4n(n-1)\bar{E}[X_{t-1}^6 X_{t-2}^2, A],\\
&\bar{E}\Big[\sum_{i\neq j} X_i^4 X_j^4, A\Big] = 3n(n-1)\bar{E}[X_{t-1}^4 X_{t-2}^4, A],\\
&\bar{E}\Big[\sum_{i\neq j\neq k} X_i^4 X_j^2 X_k^2, A\Big] = 6(n^3-3n^2+2n))\bar{E}[X_{t-1}^4 X_{t-2}^2 X_{t-3}^2, A],\\
&\bar{E}\Big[\sum_{i\neq j\neq k\neq l} X_i^2 X_j^2 X_k^2 X_l^2, A\Big] = (n^4-6n^3+11n^2-6n)\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A].
\end{aligned}
$$

The truncated expectation of $S_{2,n}^4$ is as follows:

- $$
\begin{aligned}
\bar{E}[S_{2,n}^4, A] =&\frac{1}{n^3}\bar{E}[X_{t-1}^8, A] + 4\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[X_{t-1}^6 X_{t-2}^2, A]\\
&+ 3\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[X_{t-1}^4 X_{t-2}^4, A]\\
&+ 6\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[X_{t-1}^4 X_{t-2}^2 X_{t-3}^2, A]\\
&+ \left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A].
\end{aligned}
$$

Next we expand $\bar{E}[S_{1,n}S_{2,n}^3, A]$:

$$
\begin{aligned}
S_{1,n}S_{2,n}^3 =&\frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)\left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^3\\
=&\frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^7 + \sum_{i\neq j} Y_{i+1}X_i X_j^6 + \sum_{i\neq j} Y_{i+1}X_i^5 X_j^2 + \sum_{i\neq j} Y_{i+1}X_i^3 X_j^4\right.
\end{aligned}
$$

$$+\sum_{i\neq j\neq k} Y_{i+1}X_iX_j^4X_k^2+\sum_{i\neq j\neq k} Y_{i+1}X_i^3X_j^2X_k^2+\sum_{i\neq j\neq k\neq l} Y_{i+1}X_iX_j^2X_k^2X_l^2\Bigg).$$

The truncated expectation of the seven terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau^7, A\Big] = n\bar{E}[Y_tX_{t-1}^7, A],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_iX_j^6, A\Big] = n(n-1)\bar{E}[Y_tX_{t-1}X_{t-2}^6, A],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i^5X_j^2, A\Big] = 3(n^2-n)\bar{E}[Y_tX_{t-1}^5X_{t-2}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i^3X_j^4, A\Big] = 3(n^2-n)\bar{E}[Y_tX_{t-1}^3X_{t-2}^4, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_iX_j^4X_k^2, A\Big] = 3(n^3-3n^2+2n)\bar{E}[Y_tX_{t-1}X_{t-2}^4X_{t-3}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_i^3X_j^2X_k^2, A\Big] = 3(n^3-3n^2+2n)\bar{E}[Y_tX_{t-1}^3X_{t-2}^2X_{t-3}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k\neq l} Y_{i+1}X_iX_j^2X_k^2X_l^2, A] = (n^4-6n^3+11n^2-6n)\bar{E}[Y_tX_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A].$$

The truncated expectation of $S_{1,n}S_{2,n}^3$ is as follows:

- $$\begin{aligned}\bar{E}[S_{1,n}S_{2,n}^3, A] =& \frac{1}{n^3}\bar{E}[Y_tX_{t-1}^7, A] + \left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^6, A]\\ &+3\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}^5X_{t-2}^2, A]\\ &+3\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}^3X_{t-2}^4, A]\\ &+3\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^4X_{t-3}^2, A]\\ &+3\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[Y_tX_{t-1}^3X_{t-2}^2X_{t-3}^2, A]\\ &+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A].\end{aligned}$$

Next we expand $\bar{E}[S_{1,n}^2 S_{2,n}^2, A]$:

$$\begin{aligned}
S_{1,n}^2 S_{2,n}^2 =& \frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}X_\tau\right)^2 \left(\sum_{\tau=t-n}^{t-1} X_\tau^2\right)^2 \\
=& \frac{1}{n^4}\left(\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^6 + \sum_{i\neq j} Y_{i+1}^2 X_i^2 X_j^4 + \sum_{i\neq j} Y_{i+1}^2 X_i^4 X_j^2 + \sum_{i\neq j\neq k} Y_{i+1}^2 X_i^2 X_j^2 X_k^2\right. \\
&+ \sum_{i\neq j} Y_{i+1}X_i^5 Y_{j+1}X_j + \sum_{i\neq j\neq k} Y_{i+1}X_i Y_{j+1}X_j X_k^4 + \sum_{i\neq j\neq k} Y_{i+1}X_i^3 Y_{j+1}X_j X_k^2 \\
&\left.+ \sum_{i\neq j} Y_{i+1}X_i^3 Y_{j+1}X_j^3 + \sum_{i\neq j\neq k\neq l} Y_{i+1}X_i Y_{j+1}X_j X_k^2 X_l^2\right).
\end{aligned}$$

The truncated expectation of the nine terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^2 X_\tau^6, A\Big] = n\bar{E}[Y_t^2 X_{t-1}^6, A],$$

$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}^2 X_i^2 X_j^4, A\Big] = (n^2-n)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}^2 X_i^4 X_j^2, A\Big] = 2(n^2-n)\bar{E}[Y_t^2 X_{t-1}^4 X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}^2 X_i^2 X_j^2 X_k^2, A\Big] = (n^3-3n^2+2n)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i^5 Y_j X_j, A\Big] = 2n(n-1)\bar{E}[Y_t X_{t-1}^5 Y_{t-1} X_{t-2}, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_i Y_{j+1}X_j X_k^4, A\Big] = (n^3-3n^2+2n)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k} Y_{i+1}X_i^3 Y_{j+1}X_j X_k^2, A\Big] = 4(n^3-3n^2+2n)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j} Y_{i+1}X_i^3 Y_{j+1}X_j^3, A\Big] = 2n(n-1)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2}^3, A],$$

$$\begin{aligned}
&\bar{E}\Big[\sum_{i\neq j\neq k\neq l} Y_{i+1}X_i Y_{j+1}X_j X_k^2 X_l^2, A\Big] \\
&\qquad = (n^4-6n^3+11n^2-6n)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A].
\end{aligned}$$

The truncated expectation of $S_{1,n}^2 S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{1,n}^2 S_{2,n}^2, A] =& \frac{1}{n^3}\bar{E}[Y_t^2 X_{t-1}^6, A] + \left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^4, A] \\ &+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t^2 X_{t-1}^4 X_{t-2}^2, A] \\ &+ \left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\ &+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t X_{t-1}^5 Y_{t-1} X_{t-2}, A] \\ &+ \left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^4, A] \\ &+ 4\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A] \\ &+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2}^3, A] \\ &+ \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3}\right)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A].\end{aligned}$$

The expressions for powers and products of the statistics $S_{1,n}$ and $S_{2,n}$ given above are used to expand truncated central moments of first, second, and third order. We expand $\bar{E}[(S_{1,n} - \omega_1), A]$:

- $\bar{E}[(S_{1,n} - \omega_1), A] = \bar{E}\left[Y_t X_{t-1}, A\right] - \omega_1 P(X \in A).$

The truncated expectation of $(S_{2,n} - \omega_2)$ is as follows:

- $\bar{E}[(S_{2,n} - \omega_2), A] = \bar{E}\left[X_{t-1}^2, A\right] - \omega_2 P(X \in A).$

We expand $\bar{E}[(S_{1,n} - \omega_1)^2, A]$:

$$\begin{aligned}\bar{E}[(S_{1,n} - \omega_1)^2, A] =& \bar{E}[S_{1,n}^2, A] - 2\omega_1\bar{E}[S_{1,n}, A] + \omega_1^2 P(X \in A) \\ =& \frac{1}{n}\bar{E}\left[Y_t^2 X_{t-1}^2, A\right] + \left(1 - \frac{1}{n}\right)\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}, A\right] \\ &- 2\omega_1\bar{E}\left[Y_t X_{t-1}, A\right] + \omega_1^2 P(X \in A).\end{aligned}$$

The truncated expectation of $(S_{1,n}-\omega_1)^2$ is as follows:

- $$\bar{E}[(S_{1,n}-\omega_1)^2, A] = \Big[\bar{E}[Y_tX_{t-1}Y_{t-1}X_{t-2}, A] - 2\omega_1\bar{E}\left[Y_tX_{t-1}, A\right] + \omega_1^2 P(X\in A)\Big] + \frac{1}{n}\Big[\bar{E}\left[Y_t^2X_{t-1}^2, A\right] - \bar{E}[Y_tX_{t-1}Y_{t-1}X_{t-2}, A]\Big]$$

We expand $\bar{E}[(S_{2,n}-\omega_2)^2, A]$:

$$\begin{aligned}\bar{E}[(S_{2,n}-\omega_2)^2, A] =& \bar{E}[S_{2,n}^2, A] - 2\omega_2\bar{E}[S_{2,n}, A] + \omega_2^2P(X\in A)\\ =& \frac{1}{n}\bar{E}\left[X_{t-1}^4, A\right] + \left(1-\frac{1}{n}\right)\bar{E}\left[X_{t-1}^2X_{t-2}^2, A\right] - 2\omega_2\bar{E}\left[X_{t-1}^2, A\right]\\ &+ \omega_2^2P(X\in A).\end{aligned}$$

The truncated expectation of $(S_{2,n}-\omega_2)^2$ is as follows:

- $$\bar{E}[(S_{2,n}-\omega_2)^2, A] = \Big[\bar{E}\left[X_{t-1}^2X_{t-2}^2, A\right] - 2\omega_2\bar{E}\left[X_{t-1}^2, A\right] + \omega_2^2P(X\in A)\Big] + \frac{1}{n}\Big[\bar{E}\left[X_{t-1}^4, A\right] - \bar{E}\left[X_{t-1}^2X_{t-2}^2, A\right]\Big].$$

We expand $\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2), A]$:

$$\begin{aligned}\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2), A] =& \bar{E}[S_{1,n}S_{2,n}, A] - \omega_1\bar{E}[S_{2,n}, A] - \omega_2\bar{E}[S_{1,n}, A]\\ &+ \omega_1\omega_2P(X\in A)\\ =& \frac{1}{n}\bar{E}\left[Y_tX_{t-1}^3, A\right] + \left(1-\frac{1}{n}\right)\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]\\ &- \omega_1\bar{E}\left[X_{t-1}^2, A\right] - \omega_2\bar{E}\left[Y_tX_{t-1}, A\right] + \omega_1\omega_2P(X\in A).\end{aligned}$$

The truncated expectation of $(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2), A] =& \Big[\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right] - \omega_1\bar{E}\left[X_{t-1}^2, A\right]\\ &- \omega_2\bar{E}\left[Y_tX_{t-1}, A\right] + \omega_1\omega_2P(X\in A)\Big]\\ &+ \frac{1}{n}\Big[\bar{E}\left[Y_tX_{t-1}^3, A\right] - \bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]\Big].\end{aligned}$$

We expand $\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)^2, A]$:

$$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)^2, A] &= \bar{E}[S_{1,n}S_{2,n}^2, A] - 2\omega_2\bar{E}[S_{1,n}S_{2,n}, A] - \omega_1\bar{E}[S_{2,n}^2, A] \\
&\quad + \omega_2^2\bar{E}[S_{1,n}, A] + 2\omega_1\omega_2\bar{E}[S_{2,n}, A] - \omega_1\omega_2^2 P(X \in A) \\
&= \frac{1}{n^2}\bar{E}\left[Y_t X_{t-1}^5, A\right] + \left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1} X_{t-2}^4, A\right] \\
&\quad + 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1}^3 X_{t-2}^2, A\right] \\
&\quad + \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\bar{E}\left[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A\right] \\
&\quad - 2\omega_2\left(\frac{1}{n}\bar{E}\left[Y_t X_{t-1}^3, A\right] + \left(1 - \frac{1}{n}\right)\bar{E}\left[Y_t X_{t-1} X_{t-2}^2, A\right]\right) \\
&\quad - \omega_1\left(\frac{1}{n}\bar{E}\left[X_{t-1}^4, A\right] + \left(1 - \frac{1}{n}\right)\bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right]\right) \\
&\quad + \omega_2^2\bar{E}\left[Y_t X_{t-1}, A\right] + 2\omega_1\omega_2\bar{E}\left[X_{t-1}^2, A\right] - \omega_1\omega_2^2 P(X \in A).
\end{aligned}$$

The truncated expectation of $(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)^2$ is as follows:

- $$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_1)(S_{2,n}-\omega_2)^2, A] &= \Big[\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] - 2\omega_2\bar{E}[Y_t X_{t-1} X_{t-2}^2, A] \\
&\quad - \omega_1\bar{E}[X_{t-1}^2 X_{t-2}^2, A] + \omega_2^2\bar{E}[Y_t X_{t-1}, A] + 2\omega_1\omega_2\bar{E}[X_{t-1}^2, A] \\
&\quad - \omega_1\omega_2^2 P(X \in A)\Big] \\
&+ \frac{1}{n}\Big[\bar{E}[Y_t X_{t-1} X_{t-2}^4, A] + 2\bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] \\
&\quad - 3\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] - 2\omega_2\bar{E}[Y_t X_{t-1}^3, A] \\
&\quad + 2\omega_2\bar{E}[Y_t X_{t-1} X_{t-2}^2, A] - \omega_1\bar{E}[X_{t-1}^4, A] + \omega_1\bar{E}[X_{t-1}^2 X_{t-2}^2, A]\Big] \\
&+ \frac{1}{n^2}\Big[\bar{E}[Y_t X_{t-1}^5, A] - \bar{E}[Y_t X_{t-1} X_{t-2}^4, A] \\
&\quad - 2\bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] + 2\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A]\Big].
\end{aligned}$$

We expand $\bar{E}[(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2), A]$:

$$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2), A] &= \bar{E}[S_{1,n}^2 S_{2,n}, A] - 2\omega_1\bar{E}[S_{1,n}S_{2,n}, A] - \omega_2\bar{E}[S_{1,n}^2, A] \\
&\quad + \omega_1^2\bar{E}[S_{2,n}, A] + 2\omega_1\omega_2\bar{E}[S_{1,n}, A] - \omega_1^2\omega_2 P(X \in A) \\
&= \frac{1}{n^2}\bar{E}\left[Y_t^2 X_{t-1}^4, A\right] + \left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}\left[Y_t^2 X_{t-1}^2 X_{t-2}^2, A\right]
\end{aligned}$$

$$
\begin{aligned}
&+2\left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}^3, A\right]\\
&+\left(1-3\frac{1}{n}+2\frac{1}{n^2}\right)\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}X_{t-3}^2, A\right]\\
&-2\omega_1\Big(\frac{1}{n}\bar{E}\left[Y_tX_{t-1}^3, A\right]+\left(1-\frac{1}{n}\right)\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]\Big)\\
&-\omega_2\Big(\frac{1}{n}\bar{E}\left[Y_t^2X_{t-1}^2, A\right]+\left(1-\frac{1}{n}\right)\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}, A\right]\Big)\\
&+\omega_1^2\bar{E}\left[X_{t-1}^2, A\right]+2\omega_1\omega_2\bar{E}\left[Y_tX_{t-1}, A\right]-\omega_1^2\omega_2P(X\in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2)$ is as follows:

- $\bar{E}[(S_{1,n}-\omega_1)^2(S_{2,n}-\omega_2), A]$

$$
\begin{aligned}
=&\Big[\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}X_{t-3}^2, A\right]-2\omega_1\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]\\
&-\omega_2\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}, A\right]+\omega_1^2\bar{E}\left[X_{t-1}^2, A\right]+2\omega_1\omega_2\bar{E}\left[Y_tX_{t-1}, A\right]\\
&-\omega_1^2\omega_2P(X\in A)\Big]\\
+&\frac{1}{n}\Big[\bar{E}\left[Y_t^2X_{t-1}^2X_{t-2}^2, A\right]+2\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}^3, A\right]\\
&-3\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}X_{t-3}^2, A\right]-2\omega_1\bar{E}\left[Y_tX_{t-1}^3, A\right]\\
&+2\omega_1\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]-\omega_2\bar{E}\left[Y_t^2X_{t-1}^2, A\right]\\
&+\omega_2\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}, A\right])\Big]\\
+&\frac{1}{n^2}\Big[\bar{E}\left[Y_t^2X_{t-1}^4, A\right]-\bar{E}\left[Y_t^2X_{t-1}^2X_{t-2}^2, A\right]-2\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}^3, A\right]\\
&+2\bar{E}\left[Y_tX_{t-1}Y_{t-1}X_{t-2}X_{t-3}^2, A\right]\Big].
\end{aligned}
$$

We expand $\bar{E}[(S_{2,n}-\omega_2)^3, A]$:

$$
\begin{aligned}
\bar{E}[(S_{2,n}-\omega_2)^3, A]=&\bar{E}[S_{2,n}^3, A]-3\omega_2\bar{E}[S_{2,n}^2, A]+3\omega_2^2\bar{E}[S_{2,n}, A]-\omega_2^3P(X\in A)\\
=&\frac{1}{n^2}\bar{E}\left[X_{t-1}^6, A\right]+3\left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}\left[X_{t-1}^4X_{t-2}^2, A\right]\\
&+\left(1-3\frac{1}{n}+2\frac{1}{n^2}\right)\bar{E}\left[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A\right]\\
&-3\omega_2\Big(\frac{1}{n}\bar{E}\left[X_{t-1}^4, A\right]+\left(1-\frac{1}{n}\right)\bar{E}\left[X_{t-1}^2X_{t-2}^2, A\right]\Big)\\
&+3\omega_2^2\bar{E}\left[X_{t-1}^2, A\right]-\omega_2^3P(X\in A).
\end{aligned}
$$

The truncated expectation of $(S_{2,n}-\omega_2)^3$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{2,n}-\omega_2)^3, A] = &\Big[\bar{E}\left[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A\right] - 3\omega_2 \bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right] \\ &+ 3\omega_2^2 \bar{E}\left[X_{t-1}^2, A\right] - \omega_2^3 P(X \in A)\Big] \\ &+\frac{1}{n}\Big[3\bar{E}\left[X_{t-1}^4 X_{t-2}^2, A\right] - 3\bar{E}\left[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A\right] \\ &- 3\omega_2 \bar{E}\left[X_{t-1}^4, A\right] + 3\omega_2 \bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right]\Big] \\ &+\frac{1}{n^2}\Big[\bar{E}\left[X_{t-1}^6, A\right] - 3\bar{E}\left[X_{t-1}^4 X_{t-2}^2, A\right] + 2\bar{E}\left[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A\right]\Big].\end{aligned}$$

We expand $\bar{E}[(S_{2,n}-\omega_2)^4, A]$:

$$\begin{aligned}\bar{E}[(S_{2,n}-\omega_2)^4, A] =& \bar{E}[S_{2,n}^4, A] - 4\omega_2 \bar{E}[S_{2,n}^3, A] + 6\omega_2^2 \bar{E}[S_{2,n}^2, A] - 4\omega_2^3 \bar{E}[S_{2,n}, A] \\ &+ \omega_2^4 P(X \in A) \\ =& \frac{1}{n^3}\bar{E}[X_{t-1}^8, A] + 4\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[X_{t-1}^6 X_{t-2}^2, A] \\ &+ 3\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[X_{t-1}^4 X_{t-2}^4, A] \\ &+ 6\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^3}\right)\bar{E}[X_{t-1}^4 X_{t-2}^2 X_{t-3}^2, A] \\ &+ \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3}\right)\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] \\ &- \frac{4}{n^2}\omega_2 \bar{E}[X_{t-1}^6, A] - 12\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_2 \bar{E}[X_{t-1}^4 X_{t-2}^2, A] \\ &- 4\left(1 - \frac{3}{n} + \frac{2}{n^2}\right)\omega_2 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + \frac{6}{n}\omega_2^2 \bar{E}[X_{t-1}^4, A] \\ &+ 6\left(1 - \frac{1}{n}\right)\omega_2^2 \bar{E}[X_{t-1}^2 X_{t-2}^2, A] - 4\omega_2^3 \bar{E}[X_{t-1}^2, A] + \omega_2^4 P(X \in A).\end{aligned}$$

The truncated expectation of $(S_{2,n}-\omega_2)^4$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{2,n}-\omega_2)^4, A] = &\Big[\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] - 4\omega_2 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\ &+ 6\omega_2^2 \bar{E}[X_{t-1}^2 X_{t-2}^2, A] - 4\omega_2^3 \bar{E}[X_{t-1}^2, A] + \omega_2^4 P(X \in A)\Big] \\ &+\frac{1}{n}\Big[6\bar{E}[X_{t-1}^4 X_{t-2}^2 X_{t-3}^2, A] - 6\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] \\ &- 12\omega_2 \bar{E}[X_{t-1}^4 X_{t-2}^2, A] + 12\omega_2 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + 6\omega_2^2 \bar{E}[X_{t-1}^4, A]\end{aligned}$$

$$
\begin{aligned}
&- 6\omega_2^2\bar{E}[X_{t-1}^2X_{t-2}^2, A]\Big] \\
&+\frac{1}{n^2}\Big[4\bar{E}[X_{t-1}^6X_{t-2}^2, A] + 3\bar{E}[X_{t-1}^4X_{t-2}^4, A] - 18\bar{E}[X_{t-1}^4X_{t-2}^2X_{t-3}^2, A] \\
&+ 11\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2X_{t-4}^2, A] - 4\omega_2\bar{E}[X_{t-1}^6, A] \\
&+ 12\omega_2\bar{E}[X_{t-1}^4X_{t-2}^2, A] - 8\omega_2\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]\Big] \\
&+\frac{1}{n^3}\Big[\bar{E}[X_{t-1}^8, A] - 4\bar{E}[X_{t-1}^6X_{t-2}^2, A] - 3\bar{E}[X_{t-1}^4X_{t-2}^4, A] \\
&+ 12\bar{E}[X_{t-1}^4X_{t-2}^2X_{t-3}^2, A] - 6\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]\Big].
\end{aligned}
$$

We expand $\bar{E}[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3, A]$:

$$
\begin{aligned}
\bar{E}[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3, A] &= \bar{E}[S_{1,n}S_{2,n}^3, A] - \omega_1\bar{E}[S_{2,n}^3, A] - 3\omega_2\bar{E}[S_{1,n}S_{2,n}^2, A] \\
&+ 3\omega_1\omega_2\bar{E}[S_{2,n}^2, A] + 3\omega_2^2\bar{E}[S_{1,n}S_{2,n}, A] - 3\omega_1\omega_2^2\bar{E}[S_{2,n}, A] - \omega_2^3\bar{E}[S_{1,n}, A] \\
&+ \omega_1\omega_2^3P(X \in A) \\
=&\frac{1}{n^3}\bar{E}[Y_tX_{t-1}^7, A] + \left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^6, A] \\
&+ 3\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}^5X_{t-2}^2, A] + 3\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_tX_{t-1}^3X_{t-2}^4, A] \\
&+ 3\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^4X_{t-3}^2, A] \\
&+ 3\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^3}\right)\bar{E}[Y_tX_{t-1}^3X_{t-2}^2X_{t-3}^2, A] \\
&+ \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3}\right)\bar{E}[Y_tX_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A] \\
&- \frac{\omega_1}{n^2}\bar{E}\left[X_{t-1}^6, A\right] - 3\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_1\bar{E}[X_{t-1}^4X_{t-2}^2, A] \\
&- \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\omega_1\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A] - 3\omega_2\frac{1}{n^2}\bar{E}\left[Y_tX_{t-1}^5, A\right] \\
&- 3\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_2\bar{E}[Y_tX_{t-1}X_{t-2}^4, A] \\
&- 6\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_2\bar{E}[Y_tX_{t-1}^3X_{t-2}^2, A] \\
&- 3\left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\omega_2\bar{E}[Y_tX_{t-1}X_{t-1}^2X_{t-2}^2, A] \\
&+ \frac{3}{n}\omega_1\omega_2\bar{E}\left[X_{t-1}^4, A\right] + 3\left(1 - \frac{1}{n}\right)\omega_1\omega_2\bar{E}\left[X_{t-1}^2X_{t-2}^2, A\right] \\
&+ \frac{3}{n}\omega_2^2\bar{E}\left[Y_tX_{t-1}^3, A\right] + 3\left(1 - \frac{1}{n}\right)\omega_2^2\bar{E}\left[Y_tX_{t-1}X_{t-2}^2, A\right]
\end{aligned}
$$

$$- 3\omega_1\omega_2^2\bar{E}[S_{2,n}, A] - \omega_2^3\bar{E}[S_{1,n}, A] + \omega_1\omega_2^3 P(X \in A).$$

The truncated expectation of $(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3$ is as follows:

- $$\begin{aligned}
\bar{E}[(S_{1,n} - \omega_1)(S_{2,n} - \omega_2)^3, A] = \Big[&\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] \\
&- \omega_1 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] - 3\omega_2 \bar{E}[Y_t X_{t-1} X_{t-1}^2 X_{t-2}^2, A] \\
&+ 3\omega_1\omega_2 \bar{E}[X_{t-1}^2 X_{t-2}^2, A] + 3\omega_2^2 \bar{E}[Y_t X_{t-1} X_{t-2}^2, A] \\
&- 3\omega_1\omega_2^2 \bar{E}[X_{t-1}^2, A] - \omega_3^2 \bar{E}[Y_t X_{t-1}, A] + \omega_1\omega_3^2 P(X \in A)\Big] \\
+\frac{1}{n}\Big[&3\bar{E}[Y_t X_{t-1} X_{t-2}^4 X_{t-3}^2, A] + 3\bar{E}[Y_t X_{t-1}^3 X_{t-2}^2 X_{t-3}^2, A] \\
&- 6\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] - 3\omega_1 \bar{E}[X_{t-1}^4 X_{t-2}^2, A] \\
&+ 3\omega_1 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] - 3\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^4, A] \\
&- 6\omega_2 \bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] + 9\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] \\
&+ 3\omega_1\omega_2 \bar{E}[X_{t-1}^4, A] - 3\omega_1\omega_2 \bar{E}[X_{t-1}^2 X_{t-2}^2, A] \\
&+ 3\omega_2^2 \bar{E}[Y_t X_{t-1}^3, A] - 3\omega_2^2 \bar{E}[Y_t X_{t-1} X_{t-2}^2, A]\Big] \\
+\frac{1}{n^2}\Big[&\bar{E}[Y_t X_{t-1} X_{t-2}^6, A] + 3\bar{E}[Y_t X_{t-1}^5 X_{t-2}^2, A] \\
&+ 3\bar{E}[Y_t X_{t-1}^3 X_{t-2}^4, A] - 9\bar{E}[Y_t X_{t-1} X_{t-2}^4 X_{t-3}^2, A] \\
&- 9\bar{E}[Y_t X_{t-1}^3 X_{t-2}^2 X_{t-3}^2, A] + 11\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A] \\
&- \omega_1 \bar{E}[X_{t-1}^6, A] + 3\omega_1 \bar{E}[X_{t-1}^4 X_{t-2}^2, A] - 2\omega_1 \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\
&- 3\omega_2 \bar{E}[Y_t X_{t-1}^5, A] + 3\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^4, A] \\
&+ 6\omega_2 \bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] - 6\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A]\Big] \\
+\frac{1}{n^3}\Big[&\bar{E}[Y_t X_{t-1}^7, A] - \bar{E}[Y_t X_{t-1} X_{t-2}^6, A] \\
&- 3\bar{E}[Y_t X_{t-1}^5 X_{t-2}^2, A] - 3\bar{E}[Y_t X_{t-1}^3 X_{t-2}^4, A] \\
&+ 6\bar{E}[Y_t X_{t-1} X_{t-2}^4 X_{t-3}^2, A] + 6\bar{E}[Y_t X_{t-1}^3 X_{t-2}^2 X_{t-3}^2, A] \\
&- 6\bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2 X_{t-4}^2, A]\Big].
\end{aligned}$$

We expand $\bar{E}[(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^2, A]$:

$$\bar{E}[(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^2, A] = \bar{E}[S_{1,n}^2 S_{2,n}^2, A] - 2\omega_1\bar{E}[S_{1,n}S_{2,n}^2, A] + \omega_1^2\bar{E}[S_{2,n}^2, A]$$

$$
\begin{aligned}
&- 2\omega_2\bar{E}[S_{1,n}^2 S_{2,n}, A] + 4\omega_1\omega_2\bar{E}[S_{1,n}S_{2,n}, A] - 2\omega_1^2\omega_2\bar{E}[S_{2,n}, A] \\
&+ \omega_2^2\bar{E}[S_{1,n}^2, A] - 2\omega_1\omega_2^2\bar{E}[S_{1,n}, A] + \omega_1^2\omega_2^2 P(X \in A) \\
=&\frac{1}{n^3}\bar{E}[Y_t^2 X_{t-1}^6, A] + \left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^4, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t^2 X_{t-1}^4 X_{t-2}^2, A] + \left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t X_{t-1}^5 Y_{t-1} X_{t-2}, A] \\
&+ \left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^4, A] \\
&+ 4\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^2}\right)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2}^3, A] \\
&+ \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3}\right)\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A] \\
&- \frac{2}{n^2}\omega_1\bar{E}\left[Y_t X_{t-1}^5, A\right] - 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_1\bar{E}\left[Y_t X_{t-1} X_{t-2}^4, A\right] \\
&- 4\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_1\bar{E}\left[Y_t X_{t-1}^3 X_{t-2}^2, A\right] \\
&- 2\left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\omega_1\bar{E}\left[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A\right] \\
&+ \frac{1}{n}\omega_1^2\bar{E}\left[X_{t-1}^4, A\right] + \left(1 - \frac{1}{n}\right)\omega_1^2\bar{E}\left[X_{t-1}^2 X_{t-2}^2, A\right] \\
&- \frac{2}{n^2}\omega_2\bar{E}\left[Y_t^2 X_{t-1}^4, A\right] - 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_2\bar{E}\left[Y_t^2 X_{t-1}^2 X_{t-2}^2, A\right] \\
&- 4\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_2\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}^3, A\right] \\
&- 2\left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\omega_2\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2, A\right] \\
&+ 4\omega_1\omega_2\frac{1}{n}\bar{E}[Y_t X_{t-1}^3, A] + 4\left(1 - \frac{1}{n}\right)\omega_1\omega_2\bar{E}[Y_t X_{t-1} X_{t-2}^2, A] \\
&- 2\omega_1^2\omega_2\bar{E}[X_{t-2}^2, A] + \frac{1}{n}\omega_2^2\bar{E}[Y_t^2 X_{t-1}^2, A] \\
&+ \left(1 - \frac{1}{n}\right)\omega_2^2\bar{E}\left[Y_t X_{t-1} Y_{t-1} X_{t-2}, A\right] - 2\omega_1\omega_2^2\bar{E}[Y_t X_{t-1}, A] + \omega_1^2\omega_2^2 P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^2$ is as follows:

- $\bar{E}[(S_{1,n} - \omega_1)^2(S_{2,n} - \omega_2)^2, A] = \Big[\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-2}^2 X_{t-3}^2, A]$

$$
\begin{aligned}
&- 2\omega_1 \bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] + \omega_1^2 \bar{E}[X_{t-2}^2 X_{t-3}^2, A] \\
&- 2\omega_2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2, A] + 4\omega_1\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^2, A] \\
&- 2\omega_1^2\omega_2 \bar{E}[X_{t-2}^2, A] + \omega_2^2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2}, A] - 2\omega_1\omega_2^2 \bar{E}[Y_t X_{t-1}, A] \\
&+ \omega_1^2\omega_2^2 P(X \in A)\Big] \\
+&\frac{1}{n}\Big[\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^4, A] \\
&+ 4\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A] - 6\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-2}^2 X_{t-3}^2, A] \\
&- 2\omega_1 \bar{E}[Y_t X_{t-1} X_{t-2}^4, A] - 4\omega_1 \bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] \\
&+ 6\omega_1 \bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] + \omega_1^2 \bar{E}[X_{t-2}^4, A] - \omega_1^2 \bar{E}[X_{t-2}^2 X_{t-3}^2, A] \\
&- 2\omega_2 \bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2, A] - 4\omega_2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2}^3, A] \\
&+ 6\omega_2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-2}^2, A] + 4\omega_1\omega_2 \bar{E}[Y_t X_{t-1}^3, A] \\
&- 4\omega_1\omega_2 \bar{E}[Y_t X_{t-1} X_{t-2}^2, A] + \omega_2^2 \bar{E}[Y_t^2 X_{t-1}^2, A] \\
&- \omega_2^2 \bar{E}^2[Y_t X_{t-1} Y_{t-1} X_{t-2}, A]\Big] \\
+&\frac{1}{n^2}\Big[\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^4, A] + 2\bar{E}[Y_t^2 X_{t-1}^4 X_{t-2}^2, A] \\
&- 3\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + 2\bar{E}[Y_t X_{t-1}^5 Y_{t-1} X_{t-2}, A] \\
&- 3\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-2}^4, A] - 12\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A] \\
&+ 2\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2}^3, A] + 11\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A] \\
&- 2\omega_1 \bar{E}[Y_t X_{t-1}^5, A] + 2\omega_1 \bar{E}[Y_t X_{t-1} X_{t-2}^4, A] + 4\omega_1 \bar{E}[Y_t X_{t-1}^3 X_{t-2}^2, A] \\
&- 4\omega_1 \bar{E}[Y_t X_{t-1} X_{t-2}^2 X_{t-3}^2, A] - 2\omega_2 \bar{E}[Y_t^2 X_{t-1}^4, A] \\
&+ 2\omega_2 \bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2, A] + 4\omega_2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2}^3, A] \\
&- 4\omega_2 \bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2, A]\Big] \\
+&\frac{1}{n^3}\Big[\bar{E}[Y_t^2 X_{t-1}^6, A] - \bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^4, A] \\
&- 2\bar{E}[Y_t^2 X_{t-1}^4 X_{t-2}^2, A] + 2\bar{E}[Y_t^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\
&- 2\bar{E}[Y_t X_{t-1}^5 Y_{t-1} X_{t-2}, A] + 2\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^4, A] \\
&+ 8\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2} X_{t-3}^2, A] - 2\bar{E}[Y_t X_{t-1}^3 Y_{t-1} X_{t-2}^3, A] \\
&- 6\bar{E}[Y_t X_{t-1} Y_{t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A]\Big].
\end{aligned}
$$

C.3 Proof of proposition 5.3

We begin with some lemmas.

Lemma C.1 *Given x, y real and $x \geq 0$, $x \geq |y|$ implies $x^2 \geq y^2$ or $x^2 \geq |y|^2$.*

Proof. Draw a picture of x^2. ■

Lemma C.2 $E^2[|Y_{t+1}X_t|] \geq E^2[Y_{t+1}X_t]$.

Proof. Follows by the lemma above setting $x = E[|Y_{t+1}X_t|]$ and $y = E[Y_{t+1}X_t]$. ■

Lemma C.3 $(E^{1/2}[Y_{t+1}^2]E^{1/2}[X_t^2])^2 \geq E^2[|Y_{t+1}X_t|]$.

Proof. Schwarz's inequality is $E^{1/2}[Y_{t+1}^2]E^{1/2}[X_t^2] \geq E[|Y_{t+1}X_t|] = |E[|Y_{t+1}X_t|]|$. Now apply the first lemma. ■

Proof of Proposition 5.3

Proof. We write

$$C = E[Y_{t+1}^2]\omega_2^4 - \omega_1^2\omega_2^3 = \omega_2^3(E[Y_{t+1}^2]\omega_2 - \omega_1^2),$$

and since $\omega_2 > 0$ it is enough to prove $E[Y_{t+1}^2]\omega_2 - \omega_1^2 \geq 0$. It follows

$$\begin{aligned} E[Y_{t+1}^2]\omega_2 - \omega_1^2 &= E[Y_{t+1}^2]E[X_t^2] - E^2[Y_{t+1}X_t] \\ &\geq E[Y_{t+1}^2]E[X_t^2] - E^2[|Y_{t+1}X_t|] \\ &\geq E[Y_{t+1}^2]E[X_t^2] - (E^{1/2}[Y_{t+1}^2]E^{1/2}[X_t^2])^2 = 0, \end{aligned}$$

where the inequalities follow from the lemmas above and Schwarz's inequality. ■

C.4 Matrix calculus

Most of the definitions that follow can be found in [148]. We begin with some general notation.

1. A - a general matrix $\underset{p\times q}{A} \equiv [a_{ij}]$
2. $\underset{p}{I}$ - identity matrix, p-dimensioned

3. $\underset{p}{e}^k$ - the kth elementary vector, p-dimensioned, all zeros except for a 1 in the kth position

4. $\underset{p\times q}{E}^{kl}$ - the klth elementary matrix, $p \times q$-dimensioned, all zeros except 1 in the klth position

5. $E_{q\times p}^{p\times q}$ - a permutation matrix, $pq \times pq$-dimensioned, consisting of a $q \times p$ array of $q \times p$-dimensioned elementary submatrices

$$E_{q\times p}^{p\times q} \equiv \begin{bmatrix} \underset{p\times q}{E}^{11} & \underset{p\times q}{E}^{21} & \cdots & \underset{p\times q}{E}^{p1} \\ \underset{p\times q}{E}^{21} & \underset{p\times q}{E}^{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ \underset{p\times q}{E}^{q1} & \underset{p\times q}{E}^{q2} & \cdots & \underset{p\times q}{E}^{qp} \end{bmatrix}$$

6. $A \otimes B$ - Kronecker, direct, or tensor product of two matrices $\underset{p\times q}{A}$ and $\underset{s\times t}{B}$, $ps \times qt$-dimensioned

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1q}B \\ a_{21}B & a_{22}B & \cdots & \\ \vdots & \vdots & \ddots & \\ a_{p1}B & a_{p2}B & & a_{pq}B \end{bmatrix}$$

7. $A^{\otimes k}$ - the kth Kronecker power of A

$$A^{\otimes k} \equiv \underbrace{A \otimes A \otimes \cdots \otimes A}_{k \text{ factors}}$$

$$A^{\otimes 0} \equiv 1, \quad A^{\otimes 1} \equiv A$$

8. csA - the column string of A, the column sequenced vector structure of the elements of $\underset{p\times q}{A}$

$$\underset{pq\times 1}{csA} \equiv \sum_{j=1}^{q} (\underset{q}{e}^j \otimes \underset{p}{I}) \underset{p\times q}{A} \, \underset{q}{e}^j$$

9. rsA - the row string of A, the row sequenced vector structure of the elements of $\underset{p\times q}{A}$

$$\underset{1\times pq}{rsA} \equiv \sum_{j=1}^{p} \underset{p}{e}^{j\top} \underset{p\times q}{A} (\underset{p}{e}^{j\top} \otimes \underset{q}{I})$$

10. The derivative of a matrix-valued function $\underset{p\times q}{A}(B)$ with respect to a scalar b_{kl}:

$$\mathcal{D}_{b_{kl}} A(B) \equiv \begin{bmatrix} \frac{\partial a_{11}}{\partial b_{kl}} & \frac{\partial a_{12}}{\partial b_{kl}} & \cdots & \frac{\partial a_{1q}}{\partial b_{kl}} \\ \frac{\partial a_{21}}{\partial b_{kl}} & \frac{\partial a_{22}}{\partial b_{kl}} & \cdots & \\ \vdots & \vdots & \ddots & \\ \frac{\partial a_{p1}}{\partial b_{kl}} & \frac{\partial a_{q2}}{\partial b_{kl}} & \cdots & \frac{\partial a_{pq}}{\partial b_{kl}} \end{bmatrix},$$

11. The derivative of a matrix-valued function $\underset{p\times q}{A}(B)$ with respect to a matrix $\underset{s\times t}{B}$:

$$\mathcal{D}_B A(B) \equiv \sum_{ij} \underset{s\times t}{E}^{ij} \otimes \mathcal{D}_{b_{ij}} A = \begin{bmatrix} \mathcal{D}_{b_{11}} A & \mathcal{D}_{b_{12}} A & \cdots & \mathcal{D}_{b_{1t}} A \\ \mathcal{D}_{b_{21}} A & \mathcal{D}_{b_{22}} A & \cdots & \\ \vdots & \vdots & \ddots & \\ \mathcal{D}_{b_{s1}} A & \mathcal{D}_{b_{p2}} A & \cdots & \mathcal{D}_{b_{st}} A \end{bmatrix},$$

12. Matrix derivative composition:

$$\begin{aligned} \mathcal{D}^n_{B^n} A(B) &\equiv \underbrace{\mathcal{D}_B(\mathcal{D}_B(\cdots(\mathcal{D}_B}_{n \text{ derivatives}} A(B))\cdots)), \\ \mathcal{D}^3_{CBB^\top} A(B,C) &\equiv \mathcal{D}_C(\mathcal{D}_B(\mathcal{D}_{B^\top} A)), \\ \mathcal{D}^5_{B(B^\top B)^2} A(B) &\equiv \mathcal{D}_B(\mathcal{D}_{B^\top}(\mathcal{D}_B(\mathcal{D}_{B^\top}(\mathcal{D}_B A)))), \end{aligned}$$

13. Matrix Taylor expansion: The Taylor expansion for a matrix-valued function $A(b)$ of a vector b, where b may be the row string or column string of a matrix B:

$$\begin{aligned} \underset{p\times q}{A}(\underset{s}{b}) &= A(\bar{b}) + \sum_{m=1}^{M} \frac{1}{m!} \left(\mathcal{D}^m_{b^{\top m}} A(b) \right)_{b=\bar{b}} ((b-\bar{b})^{\otimes m} \otimes \underset{q}{I}) + R_{M+1}(\bar{b}, b), \\ R_{m+1}(\bar{b}, b) &= \frac{1}{m!} \int_{\xi=\bar{b}}^{b} \left(\mathcal{D}^{m+1}_{\xi^{\top m+1}} A(\xi) \right) (\underset{s}{I} \otimes (b-\xi)^{\otimes m} \otimes \underset{q}{I})(d\xi \otimes \underset{q}{I}). \end{aligned}$$

C.5 The multi-variate problem

In this section we present the calculations needed to re-express the term $(b - \bar{b})^{\otimes i}$ for $i = 2, 3, 4$. We begin by writing the terms needed for the second order expansion.

C.5.1 Expansion of the central moments for the multi-variate problem

- $E[S_{1i,n}] = \frac{1}{n} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i Y_{\tau+1}\right] = E\left[X_{t-1}^i Y_t\right] = \omega_{1i},$
- $E[S_{2ij,n}] = \frac{1}{n} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i X_\tau\right] = E\left[X_{t-1}^i X_{t-1}\right] = \omega_{2ij},$

- $$\begin{aligned} E[S_{1i,n} S_{1j,n}] =& \frac{1}{n^2} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i Y_{\tau+1} \sum_{\tau=t-n}^{t-1} X_\tau^j Y_{\tau+1}\right] \\ =& \frac{1}{n^2} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i Y_{\tau+1} X_\tau^j Y_{\tau+1} + \sum_{\tau_1 \neq \tau_2} X_{\tau_1}^i Y_{\tau_1+1} X_{\tau_2}^j Y_{\tau_2+1}\right] \\ =& \frac{1}{n} E\left[Y_t^2 X_{t-1}^i X_{t-1}^j\right] + \left(1 - \frac{1}{n}\right) \omega_{1i} \omega_{1j} \end{aligned}$$

- $$\begin{aligned} E[S_{1i,n} S_{2jk,n}] =& \frac{1}{n^2} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i Y_{\tau+1} \sum_{\tau=t-n}^{t-1} X_\tau^j X_\tau^k\right] \\ =& \frac{1}{n^2} E\left[\sum_{\tau=t-n}^{t-1} Y_{\tau+1} X_\tau^i X_\tau^j X_\tau^k + \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X_{\tau_1}^i X_{\tau_2}^j X_{\tau_2}^k\right] \\ =& \frac{1}{n} E\left[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k\right] + \left(1 - \frac{1}{n}\right) \omega_{1i} \omega_{2jk} \end{aligned}$$

- $$E[S_{2ij,n} S_{2kl,n}] = \frac{1}{n^2} E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i X_\tau^j \sum_{\tau=t-n}^{t-1} X_\tau^k X_\tau^l\right]$$

$$=\frac{1}{n^2}E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i X_\tau^j X_\tau^k X_\tau^l + \sum_{\tau_1\neq\tau_2} X_{\tau_1}^i X_{\tau_1}^j X_{\tau_2}^k X_{\tau_2}^l\right]$$
$$=\frac{1}{n}E\left[X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^k\right] + \left(1-\frac{1}{n}\right)\omega_{2ij}\omega_{2kl}$$

We can now expand the second order central moments and with a superscript 2 to indicate these second order terms, we define the quantities $V_{1,ij}^2$, $V_{2,ijk}^2$, and $V_{3,ijkl}^2$:

- $$\begin{aligned} E\left[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})\right] =& E\left[S_{1i,n}S_{1j,n}-\omega_{1i}S_{1j,n}-\omega_{1j}S_{1i,n}+\omega_{1i}\omega_{1j}\right] \\ =& \frac{1}{n}\left(E[Y_t^2 X_{t-1}^i X_{t-1}^j]-\omega_{1i}\omega_{1j}\right) \\ \equiv& \frac{1}{n}V_{1,ij}^2 \end{aligned}$$

- $$\begin{aligned} E\left[(S_{1i,n}-\omega_{1i})(S_{2jk,n}-\omega_{2jk})\right] =& E\left[S_{1i,n}S_{2jk,n}-\omega_{1i}S_{2jk,n}-\omega_{2jk}S_{1i,n}+\omega_{1i}\omega_{2jk}\right] \\ =& \frac{1}{n}\left(E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]-\omega_{1i}\omega_{2jk}\right) \\ \equiv& \frac{1}{n}V_{2,ijk}^2 \end{aligned}$$

- $$\begin{aligned} E\left[(S_{2ij,n}-\omega_{2ij})(S_{2kl,n}-\omega_{2kl})\right] =& E\left[S_{2ij,n}S_{2kl,n}-\omega_{2ij}S_{2kl,n}-\omega_{2kl}S_{2ij,n}+\omega_{2ij}\omega_{2kl}\right] \\ =& \frac{1}{n}\left(E[X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l]-\omega_{2ij}\omega_{2kl}\right) \\ \equiv& \frac{1}{n}V_{3,ijkl}^2 \end{aligned}$$

We proceed with the terms needed for the third order term:

- $$E[S_{1i,n}S_{1j,n}S_{1k,n}] = \frac{1}{n^3}E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i Y_{\tau+1}\sum_{\tau=t-n}^{t-1} X_\tau^j Y_{\tau+1}\sum_{\tau=t-n}^{t-1} X_\tau^k Y_{\tau+1}\right]$$
$$=\frac{1}{n^3}E\left[\sum_{\tau=t-n}^{t-1} Y_{\tau+1}^3 X_\tau^i X_\tau^j X_\tau^k + \sum_{\tau_1\neq\tau_2} Y_{\tau_1+1}X_{\tau_1}^i Y_{\tau_2+1}^2 X_{\tau_2}^j X_{\tau_2}^k\right.$$
$$\left. + \sum_{\tau_1\neq\tau_2} Y_{\tau_1+1}^2 X_{\tau_1}^i X_{\tau_1}^k Y_{\tau_2+1}X_{\tau_2}^j + \sum_{\tau_1\neq\tau_2} Y_{\tau_1+1}^2 X_{\tau_1}^i X_{\tau_1}^j Y_{\tau_2+1}X_{\tau_2}^k\right.$$

$$
\begin{aligned}
&+ \sum_{\tau_1 \neq \tau_2 \neq \tau_3} Y_{\tau_1+1} X^i_{\tau_1} Y_{\tau_2+1} X^j_{\tau_2} Y_{\tau_3+1} X^k_{\tau_3} \Bigg] \\
=& \frac{1}{n^2} E\left[Y_t^3 X^i_{t-1} X^j_{t-1} X^k_{t-1} \right] + \left(1 - \frac{1}{n^2}\right) E[Y_t^2 X^j_{t-1} X^k_{t-1}] \omega_{1i} \\
&+ \left(1 - \frac{1}{n^2}\right) E[Y_t^2 X^i_{t-1} X^k_{t-1}] \omega_{1j} + \left(1 - \frac{1}{n^2}\right) E[Y_t^2 X^i_{t-1} X^j_{t-1}] \omega_{1k} \\
&+ \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) \omega_{1i} \omega_{1j} \omega_{1k}
\end{aligned}
$$

- $E[S_{1i,n} S_{1j,n} S_{2kl,n}] = \dfrac{1}{n^3} E\left[\sum_{\tau=t-n}^{t-1} X^i_\tau Y_{\tau+1} \sum_{\tau=t-n}^{t-1} X^j_\tau Y_{\tau+1} \sum_{\tau=t-n}^{t-1} X^k_\tau X^l_\tau \right]$

$$
\begin{aligned}
=& \frac{1}{n^3} E\Bigg[\sum_{\tau=t-n}^{t-1} Y^2_{\tau+1} X^i_\tau X^j_\tau X^k_\tau X^l_\tau + \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X^i_{\tau_1} Y_{\tau_2+1} X^j_{\tau_2} X^k_{\tau_2} X^l_{\tau_2} \\
&+ \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X^i_{\tau_1} X^k_{\tau_1} X^l_{\tau_1} Y_{\tau_2+1} X^j_{\tau_2} + \sum_{\tau_1 \neq \tau_2} Y^2_{\tau_1+1} X^i_{\tau_1} X^j_{\tau_1} X^k_{\tau_2} X^l_{\tau_2} \\
&+ \sum_{\tau_1 \neq \tau_2 \neq \tau_3} Y_{\tau_1+1} X^i_{\tau_1} Y_{\tau_2+1} X^j_{\tau_2} X^k_{\tau_3} X^l_{\tau_3} \Bigg] \\
=& \frac{1}{n^2} E\left[Y_t^2 X^i_{t-1} X^j_{t-1} X^k_{t-1} X^l_{t-1} \right] + \left(1 - \frac{1}{n^2}\right) E[Y_t X^j_{t-1} X^k_{t-1} X^l_{t-1}] \omega_{1i} \\
&+ \left(1 - \frac{1}{n^2}\right) E[Y_t X^i_{t-1} X^k_{t-1} X^l_{t-1}] \omega_{1j} + \left(1 - \frac{1}{n^2}\right) E[Y_t^2 X^i_{t-1} X^j_{t-1}] \omega_{2kl} \\
&+ \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) \omega_{1i} \omega_{1j} \omega_{2kl}
\end{aligned}
$$

- $E[S_{1i,n} S_{2jk,n} S_{2lo,n}] = \dfrac{1}{n^3} E\left[\sum_{\tau=t-n}^{t-1} X^i_\tau Y_{\tau+1} \sum_{\tau=t-n}^{t-1} X^j_\tau X^k_\tau \sum_{\tau=t-n}^{t-1} X^l_\tau X^o_\tau \right]$

$$
\begin{aligned}
=& \frac{1}{n^3} E\Bigg[\sum_{\tau=t-n}^{t-1} Y_{\tau+1} X^i_\tau X^j_\tau X^k_\tau X^l_\tau X^o_\tau + \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X^i_{\tau_1} X^j_{\tau_2} X^k_{\tau_2} X^l_{\tau_2} X^o_{\tau_2} \\
&+ \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X^i_{\tau_1} X^l_{\tau_1} X^o_{\tau_1} X^j_{\tau_2} X^k_{\tau_2} + \sum_{\tau_1 \neq \tau_2} Y_{\tau_1+1} X^i_{\tau_1} X^j_{\tau_1} X^k_{\tau_1} X^l_{\tau_2} X^o_{\tau_2} \\
&+ \sum_{\tau_1 \neq \tau_2 \neq \tau_3} Y_{\tau_1+1} X^i_{\tau_1} X^j_{\tau_2} X^k_{\tau_2} X^l_{\tau_3} X^o_{\tau_3} \Bigg]
\end{aligned}
$$

$$=\frac{1}{n^2}E\left[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o\right] + \left(\frac{1}{n}-\frac{1}{n^2}\right)E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i}$$
$$+\left(\frac{1}{n}-\frac{1}{n^2}\right)E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]\omega_{2jk} + \left(\frac{1}{n}-\frac{1}{n^2}\right)E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]\omega_{2lo}$$
$$+\left(1-\frac{3}{n}+\frac{2}{n^2}\right)\omega_{1i}\omega_{2jk}\omega_{2lo}$$

- $E[S_{2ij,n}S_{2kl,n}S_{2op,n}] = \frac{1}{n^3}E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i X_\tau^j \sum_{\tau=t-n}^{t-1} X_\tau^k X_\tau^l \sum_{\tau=t-n}^{t-1} X_\tau^o X_\tau^p\right]$

$$=\frac{1}{n^3}E\left[\sum_{\tau=t-n}^{t-1} X_\tau^i X_\tau^j X_\tau^k X_\tau^l X_\tau^o X_\tau^p + \sum_{\tau_1\neq\tau_2} X_{\tau_1}^i X_{\tau_1}^j X_{\tau_2}^k X_{\tau_2}^l X_{\tau_2}^o X_{\tau_2}^p\right.$$
$$+\sum_{\tau_1\neq\tau_2} X_{\tau_1}^i X_{\tau_1}^j X_{\tau_1}^o X_{\tau_1}^p X_{\tau_2}^k X_{\tau_2}^l + \sum_{\tau_1\neq\tau_2} X_{\tau_1}^i X_{\tau_1}^j X_{\tau_1}^k X_{\tau_1}^l X_{\tau_2}^o X_{\tau_2}^p$$
$$\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3} X_{\tau_1}^i X_{\tau_1}^j X_{\tau_2}^k X_{\tau_2}^l X_{\tau_3}^o X_{\tau_3}^p\right]$$
$$=\frac{1}{n^2}E\left[X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p\right] + \left(\frac{1}{n}-\frac{1}{n^2}\right)E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{2ij}$$
$$+\left(\frac{1}{n}-\frac{1}{n^2}\right)E[X_{t-1}^i X_{t-1}^j X_{t-1}^o X_{t-1}^p]\omega_{2kl} + \left(\frac{1}{n}-\frac{1}{n^2}\right)E[X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l]\omega_{2op}$$
$$+\left(1-\frac{3}{n}+\frac{2}{n^2}\right)\omega_{2ij}\omega_{2kl}\omega_{2op}$$

We can now expand the third order central moments and with a superscript 3 to indicate these third order terms, we define the quantities $V_{1,ijk}^3$, $V_{2,ijkl}^3$, $V_{3,ijklo}^3$, and $V_{3,ijklop}^3$:

- $E\left[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{1k,n}-\omega_{1k})\right]$

$$=E[S_{1i,n}S_{1j,n}S_{1k,n}] - \omega_{1i}E[S_{1j,n}S_{1k,n}] - \omega_{1j}E[S_{1i,n}S_{1k,n}] - \omega_{1k}E[S_{1i,n}S_{1j,n}]$$
$$+\omega_{1i}\omega_{1j}E[S_{1k,n}] + \omega_{1i}\omega_{1k}E[S_{1j,n}] + \omega_{1j}\omega_{1k}E[S_{1i,n}] + \omega_{1i}\omega_{1j}\omega_{1k}$$
$$=\frac{1}{n^2}\left[E[Y_t^3 X_{t-1}^i X_{t-1}^j X_{t-1}^k] - E[Y_t^2 X_{t-1}^j X_{t-1}^k]\omega_{1i} - E[Y_t^2 X_{t-1}^i X_{t-1}^k]\omega_{1j}\right.$$
$$\left.-E[Y_t^2 X_{t-1}^i X_{t-1}^j]\omega_{1k} + 2\omega_{1i}\omega_{1j}\omega_{1k}\right]$$
$$\equiv\frac{1}{n^2}V_{1,ijk}^3$$

- $E\left[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{2kl,n}-\omega_{2kl})\right]$

$$
\begin{aligned}
=&E[S_{1i,n}S_{1j,n}S_{2kl,n}]-\omega_{1i}E[S_{1j,n}S_{2kl,n}]-\omega_{1j}E[S_{1i,n}S_{2kl,n}]-\omega_{2kl}E[S_{1i,n}S_{1j,n}]\\
&+\omega_{1i}\omega_{1j}E[S_{2kl,n}]+\omega_{1i}\omega_{2kl}E[S_{1j,n}]+\omega_{1j}\omega_{2kl}E[S_{1i,n}]+\omega_{1i}\omega_{1j}\omega_{2kl}\\
=&\frac{1}{n^2}\Big[E[Y_t^2X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^l]-E[Y_tX_{t-1}^jX_{t-1}^kX_{t-1}^l]\omega_{1i}\\
&-E[Y_tX_{t-1}^iX_{t-1}^kX_{t-1}^l]\omega_{1j}-E[Y_t^2X_{t-1}^iX_{t-1}^j]\omega_{2kl}+2\omega_{1i}\omega_{1j}\omega_{2kl}\Big]\\
\equiv&\frac{1}{n^2}V^3_{2,ijkl}
\end{aligned}
$$

- $E\left[(S_{1i,n}-\omega_{1i})(S_{2jk,n}-\omega_{2jk})(S_{2lo,n}-\omega_{2lo})\right]$

$$
\begin{aligned}
=&E[S_{1i,n}S_{2jk,n}S_{2lo,n}]-\omega_{1i}E[S_{2jk,n}S_{2lo,n}]-\omega_{2jk}E[S_{1i,n}S_{2lo,n}]-\omega_{2lo}E[S_{1i,n}S_{2jk,n}]\\
&+\omega_{1i}\omega_{2jk}E[S_{2lo,n}]+\omega_{1i}\omega_{2lo}E[S_{2jk,n}]+\omega_{2jk}\omega_{2lo}E[S_{1i,n}]+\omega_{1i}\omega_{2jk}\omega_{2lo}\\
=&\frac{1}{n^2}\Big[E[Y_tX_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o]-E[X_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o]\omega_{1i}\\
&-E[Y_tX_{t-1}^iX_{t-1}^lX_{t-1}^o]\omega_{2jk}-E[Y_tX_{t-1}^iX_{t-1}^jX_{t-1}^k]\omega_{2lo}+2\omega_{1i}\omega_{2jk}\omega_{2lo}\Big]\\
\equiv&\frac{1}{n^2}V^3_{3,ijklo}
\end{aligned}
$$

- $E\left[(S_{2ij,n}-\omega_{2ij})(S_{2kl,n}-\omega_{2kl})(S_{2op,n}-\omega_{2op})\right]$

$$
\begin{aligned}
=&E[S_{2ij,n}S_{2kl,n}S_{2op,n}]-\omega_{2ij}E[S_{2kl,n}S_{2op,n}]-\omega_{2kl}E[S_{2ij,n}S_{2op,n}]-\omega_{2op}E[S_{2ij,n}S_{2kl,n}]\\
&+\omega_{2ij}\omega_{2kl}E[S_{2op,n}]+\omega_{2ij}\omega_{2op}E[S_{2kl,n}]+\omega_{2kl}\omega_{2op}E[S_{2ij,n}]+\omega_{2ij}\omega_{2kl}\omega_{2op}\\
=&\frac{1}{n^2}\Big[E[X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^oX_{t-1}^p]-E[X_{t-1}^kX_{t-1}^lX_{t-1}^oX_{t-1}^p]\omega_{2ij}\\
&-E[X_{t-1}^iX_{t-1}^jX_{t-1}^oX_{t-1}^p]\omega_{2kl}-E[X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^l]\omega_{2op}+2\omega_{2ij}\omega_{2kl}\omega_{2op}\Big]\\
\equiv&\frac{1}{n^2}V^3_{4,ijklop}
\end{aligned}
$$

We proceed with the terms needed for the fourth order term:

- $E[S_{1i,n}S_{1j,n}S_{1k,n}S_{1l,n}]$

$$
\begin{aligned}
=&\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X_\tau^iY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^jY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^kY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^lY_{\tau+1}\right]\\
=&\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}Y_{\tau+1}^4X_\tau^iX_\tau^jX_\tau^kX_\tau^l+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jY_{\tau_2+1}^2X_{\tau_2}^kX_{\tau_2}^l\right.\\
&+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^kY_{\tau_2+1}^2X_{\tau_2}^jX_{\tau_2}^l+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^lY_{\tau_2+1}^2X_{\tau_2}^jX_{\tau_2}^k\\
&+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iY_{\tau_2+1}^3X_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^l+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^jY_{\tau_2+1}^3X_{\tau_2}^iX_{\tau_2}^kX_{\tau_2}^l\\
&+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^kY_{\tau_2+1}^3X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^l+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^lY_{\tau_2+1}^3X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^k\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jY_{\tau_2+1}X_{\tau_2}^kY_{\tau_3+1}X_{\tau_3}^l+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^jY_{\tau_3+1}X_{\tau_3}^l\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^jX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^iY_{\tau_3+1}X_{\tau_3}^l+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^jY_{\tau_3+1}X_{\tau_3}^k\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^jX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^iY_{\tau_3+1}X_{\tau_3}^k+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^iY_{\tau_3+1}X_{\tau_3}^j\\
&\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3\neq\tau_4}Y_{\tau_1+1}X_{\tau_1}^iY_{\tau_2+1}X_{\tau_2}^jY_{\tau_3+1}X_{\tau_3}^kY_{\tau_4+1}X_{\tau_4}^l\right]\\
=&\frac{1}{n^3}E[Y_t^4X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^l]+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^j]E[Y_t^2X_{t-1}^kX_{t-1}^l]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^k]E[Y_t^2X_{t-1}^jX_{t-1}^l]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^l]E[Y_t^2X_{t-1}^jX_{t-1}^k]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^3X_{t-1}^jX_{t-1}^kX_{t-1}^l]\omega_{1i}+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^3X_{t-1}^iX_{t-1}^kX_{t-1}^l]\omega_{1j}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^l]\omega_{1k}+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^k]\omega_{1l}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^j]\omega_{1k}\omega_{1l}+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^k]\omega_{1j}\omega_{1l}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^k]\omega_{1i}\omega_{1l}+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^l]\omega_{1j}\omega_{1k}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^l]\omega_{1i}\omega_{1k}+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^kX_{t-1}^l]\omega_{1i}\omega_{1j}
\end{aligned}
$$

$$+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\omega_{1i}\omega_{1j}\omega_{1k}\omega_{1l}$$

- $E[S_{1i,n}S_{1j,n}S_{1k,n}S_{2lo,n}]$

$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X_\tau^iY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^jY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^kY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^lX_\tau^o\right]$$

$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}Y_{\tau+1}^3X_\tau^iX_\tau^jX_\tau^kX_\tau^lX_\tau^o+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jY_{\tau_2+1}X_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^o\right.$$
$$+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^jX_{\tau_2}^lX_{\tau_2}^o+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^jX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^lX_{\tau_2}^o$$
$$+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iY_{\tau_2+1}^2X_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^o+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^jY_{\tau_2+1}^2X_{\tau_2}^iX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^o$$
$$+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^kY_{\tau_2+1}^2X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^lX_{\tau_2}^o+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}^3X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^k$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jY_{\tau_2+1}X_{\tau_2}^kX_{\tau_3}^lX_{\tau_3}^o+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^jX_{\tau_3}^lX_{\tau_3}^o$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^jX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^iX_{\tau_3}^lX_{\tau_3}^o$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^jY_{\tau_3+1}X_{\tau_3}^k$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^jX_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^iY_{\tau_3+1}X_{\tau_3}^k$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^kX_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^iY_{\tau_3+1}X_{\tau_3}^j$$
$$\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3\neq\tau_4}Y_{\tau_1+1}X_{\tau_1}^iY_{\tau_2+1}X_{\tau_2}^jY_{\tau_3+1}X_{\tau_3}^kX_{\tau_4}^lX_{\tau_4}^o\right]$$
$$=\frac{1}{n^3}E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o]$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^j]E[Y_tX_{t-1}^kX_{t-1}^lX_{t-1}^o]$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^k]E[Y_tX_{t-1}^jX_{t-1}^lX_{t-1}^o]$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^k]E[Y_tX_{t-1}^iX_{t-1}^lX_{t-1}^o]$$

$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o]\omega_{1i}$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^kX_{t-1}^lX_{t-1}^o]\omega_{1j}$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^jX_{t-1}^lX_{t-1}^o]\omega_{1k}$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^k]\omega_{2lo}+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^j]\omega_{1k}\omega_{2lo}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^k]\omega_{1j}\omega_{2lo}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^k]\omega_{1i}\omega_{2lo}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^iX_{t-1}^lX_{t-1}^o]\omega_{1j}\omega_{1k}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^jX_{t-1}^lX_{t-1}^o]\omega_{1i}\omega_{1k}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X_{t-1}^kX_{t-1}^lX_{t-1}^o]\omega_{1i}\omega_{1j}$$
$$+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\omega_{1i}\omega_{1j}\omega_{1k}\omega_{2lo}$$

- $E[S_{1i,n}S_{1j,n}S_{2kl,n}S_{2op,n}]$

$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X_\tau^iY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^jY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^kX_\tau^l\sum_{\tau=t-n}^{t-1}X_\tau^oX_\tau^p\right]$$
$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}Y_{\tau+1}^2X_\tau^iX_\tau^jX_\tau^kX_\tau^lX_\tau^oX_\tau^p+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^p\right.$$
$$+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^jX_{\tau_2}^oX_{\tau_2}^p+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^jX_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^oX_{\tau_2}^p$$
$$+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iY_{\tau_2+1}X_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^p+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^jY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^p$$
$$+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}^2X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^oX_{\tau_2}^p+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^oX_{\tau_1}^pY_{\tau_2+1}^2X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^l$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}^2X_{\tau_1}^iX_{\tau_1}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_3}^oX_{\tau_3}^p+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^jX_{\tau_3}^oX_{\tau_3}^p$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^jX_{\tau_1}^kX_{\tau_1}^lY_{\tau_2+1}X_{\tau_2}^iX_{\tau_3}^oX_{\tau_3}^p$$

$$
\begin{aligned}
&+\sum_{\tau_1\neq\tau_2\neq\tau_3} Y_{\tau_1+1}X^i_{\tau_1}X^o_{\tau_1}X^p_{\tau_1}Y_{\tau_2+1}X^j_{\tau_2}X^k_{\tau_3}X^l_{\tau_3}\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3} Y_{\tau_1+1}X^j_{\tau_1}X^o_{\tau_1}X^p_{\tau_1}Y_{\tau_2+1}X^i_{\tau_2}X^k_{\tau_3}X^l_{\tau_3}\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3} X^k_{\tau_1}X^l_{\tau_1}X^o_{\tau_1}X^p_{\tau_1}Y_{\tau_2+1}X^i_{\tau_2}Y_{\tau_3+1}X^j_{\tau_3}\\
&\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3\neq\tau_4} Y_{\tau_1+1}X^i_{\tau_1}Y_{\tau_2+1}X^j_{\tau_2}X^k_{\tau_3}X^l_{\tau_3}X^o_{\tau_4}X^p_{\tau_4}\right]\\
=&\frac{1}{n^3}E[Y_t^2X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X^i_{t-1}X^j_{t-1}]E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^i_{t-1}X^k_{t-1}X^l_{t-1}]E[Y_tX^j_{t-1}X^o_{t-1}X^p_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^i_{t-1}X^o_{t-1}X^p_{t-1}]E[Y_tX^j_{t-1}X^k_{t-1}X^l_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{1i}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^i_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{1j}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2kl}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_t^2X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}]\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_t^2X^i_{t-1}X^j_{t-1}]\omega_{2kl}\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^i_{t-1}X^k_{t-1}X^l_{t-1}]\omega_{1j}\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^j_{t-1}X^k_{t-1}X^l_{t-1}]\omega_{1i}\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^i_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{1j}\omega_{2kl}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^j_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{1i}\omega_{2kl}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{1i}\omega_{1j}\\
&+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\omega_{1i}\omega_{1j}\omega_{2kl}\omega_{2op}
\end{aligned}
$$

- $E[S_{1i,n}S_{2jk,n}S_{2lo,n}S_{2pq,n}]$

$$
\begin{aligned}
=&\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X_\tau^iY_{\tau+1}\sum_{\tau=t-n}^{t-1}X_\tau^jX_\tau^k\sum_{\tau=t-n}^{t-1}X_\tau^lX_\tau^o\sum_{\tau=t-n}^{t-1}X_\tau^pX_\tau^q\right]\\
=&\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}Y_{\tau+1}X_\tau^iX_\tau^jX_\tau^kX_\tau^lX_\tau^oX_\tau^pX_\tau^q+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^jX_{\tau_1}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^pX_{\tau_2}^q\right.\\
&+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^lX_{\tau_1}^oX_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^pX_{\tau_2}^q+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^jX_{\tau_1}^kX_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^pX_{\tau_2}^q\\
&+\sum_{\tau_1\neq\tau_2}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^pX_{\tau_2}^q+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^jX_{\tau_1}^kY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^lX_{\tau_2}^oX_{\tau_2}^pX_{\tau_2}^q\\
&+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^pX_{\tau_2}^q+\sum_{\tau_1\neq\tau_2}X_{\tau_1}^pX_{\tau_1}^qY_{\tau_2+1}X_{\tau_2}^iX_{\tau_2}^jX_{\tau_2}^kX_{\tau_2}^lX_{\tau_2}^o\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^jX_{\tau_1}^kX_{\tau_2}^lX_{\tau_2}^oX_{\tau_3}^pX_{\tau_3}^q+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^lX_{\tau_1}^oX_{\tau_2}^jX_{\tau_2}^kX_{\tau_3}^pX_{\tau_3}^q\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}X_{\tau_1}^jX_{\tau_1}^kX_{\tau_1}^lX_{\tau_1}^oY_{\tau_2+1}X_{\tau_2}^iX_{\tau_3}^pX_{\tau_3}^q\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_1}^pX_{\tau_1}^qX_{\tau_2}^jX_{\tau_2}^kX_{\tau_3}^lX_{\tau_3}^o\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}X_{\tau_1}^jX_{\tau_1}^kX_{\tau_1}^pX_{\tau_1}^qY_{\tau_2+1}X_{\tau_2}^iX_{\tau_3}^lX_{\tau_3}^o\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3}X_{\tau_1}^lX_{\tau_1}^oX_{\tau_1}^pX_{\tau_1}^qY_{\tau_2+1}X_{\tau_2}^iX_{\tau_3}^jX_{\tau_3}^k\\
&\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3\neq\tau_4}Y_{\tau_1+1}X_{\tau_1}^iX_{\tau_2}^jX_{\tau_2}^kX_{\tau_3}^lX_{\tau_3}^oX_{\tau_4}^pX_{\tau_4}^q\right]\\
=&\frac{1}{n^3}E[Y_tX_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^oX_{t-1}^pX_{t-1}^q]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX_{t-1}^iX_{t-1}^jX_{t-1}^k]E[X_{t-1}^lX_{t-1}^oX_{t-1}^pX_{t-1}^q]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX_{t-1}^iX_{t-1}^lX_{t-1}^o]E[X_{t-1}^jX_{t-1}^kX_{t-1}^pX_{t-1}^q]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX_{t-1}^iX_{t-1}^pX_{t-1}^q]E[X_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^oX_{t-1}^pX_{t-1}^q]\omega_{1i}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX_{t-1}^iX_{t-1}^lX_{t-1}^oX_{t-1}^pX_{t-1}^q]\omega_{2jk}
\end{aligned}
$$

$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^i_{t-1}X^j_{t-1}X^k_{t-1}X^p_{t-1}X^q_{t-1}]\omega_{2lo}$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[Y_tX^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}]\omega_{2pq}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^i_{t-1}X^j_{t-1}X^k_{t-1}]\omega_{2lo}\omega_{2pq}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^i_{t-1}X^l_{t-1}X^o_{t-1}]\omega_{2jk}\omega_{2pq}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}]\omega_{1i}\omega_{2pq}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[Y_tX^i_{t-1}X^p_{t-1}X^q_{t-1}]\omega_{2jk}\omega_{2lo}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^j_{t-1}X^k_{t-1}X^p_{t-1}X^q_{t-1}]\omega_{1i}\omega_{2lo}$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^l_{t-1}X^o_{t-1}X^p_{t-1}X^q_{t-1}]\omega_{1i}\omega_{2jk}$$
$$+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\omega_{1i}\omega_{2jk}\omega_{2lo}\omega_{2pq}$$

- $E[S_{2ij,n}S_{2kl,n}S_{2op,n}S_{2qr,n}]$

$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X^i_\tau X^j_\tau\sum_{\tau=t-n}^{t-1}X^k_\tau X^l_\tau\sum_{\tau=t-n}^{t-1}X^o_\tau X^p_\tau\sum_{\tau=t-n}^{t-1}X^q_\tau X^r_\tau\right]$$
$$=\frac{1}{n^4}E\left[\sum_{\tau=t-n}^{t-1}X^i_\tau X^j_\tau X^k_\tau X^l_\tau X^o_\tau X^p_\tau X^q_\tau X^r_\tau+\sum_{\tau_1\neq\tau_2}X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_1}X^l_{\tau_1}X^o_{\tau_2}X^p_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}\right.$$
$$+\sum_{\tau_1\neq\tau_2}X^i_{\tau_1}X^j_{\tau_1}X^q_{\tau_1}X^r_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_2}X^p_{\tau_2}+\sum_{\tau_1\neq\tau_2}X^i_{\tau_1}X^j_{\tau_1}X^o_{\tau_1}X^p_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}$$
$$+\sum_{\tau_1\neq\tau_2}X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_2}X^p_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}+\sum_{\tau_1\neq\tau_2}X^k_{\tau_1}X^l_{\tau_1}X^i_{\tau_2}X^j_{\tau_2}X^o_{\tau_2}X^p_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}$$
$$+\sum_{\tau_1\neq\tau_2}X^q_{\tau_1}X^r_{\tau_1}X^i_{\tau_2}X^j_{\tau_2}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_2}X^p_{\tau_2}+\sum_{\tau_1\neq\tau_2}X^o_{\tau_1}X^p_{\tau_1}X^i_{\tau_2}X^j_{\tau_2}X^k_{\tau_2}X^l_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_1}X^l_{\tau_1}X^o_{\tau_2}X^p_{\tau_2}X^q_{\tau_3}X^r_{\tau_3}+\sum_{\tau_1\neq\tau_2\neq\tau_3}X^i_{\tau_1}X^j_{\tau_1}X^q_{\tau_1}X^r_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_3}X^p_{\tau_3}$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^q_{\tau_2}X^r_{\tau_2}X^o_{\tau_3}X^p_{\tau_3}$$
$$+\sum_{\tau_1\neq\tau_2\neq\tau_3}X^i_{\tau_1}X^j_{\tau_1}X^o_{\tau_1}X^p_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^q_{\tau_3}X^r_{\tau_3}$$

$$
\begin{aligned}
&+\sum_{\tau_1\neq\tau_2\neq\tau_3} X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_2}X^p_{\tau_2}X^q_{\tau_3}X^r_{\tau_3}\\
&+\sum_{\tau_1\neq\tau_2\neq\tau_3} X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_3}X^p_{\tau_3}X^q_{\tau_3}X^r_{\tau_3}\\
&\left.+\sum_{\tau_1\neq\tau_2\neq\tau_3\neq\tau_4} X^i_{\tau_1}X^j_{\tau_1}X^k_{\tau_2}X^l_{\tau_2}X^o_{\tau_3}X^p_{\tau_3}X^q_{\tau_4}X^r_{\tau_4}\right]\\
=&\frac{1}{n^3}E[X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}]E[X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^q_{t-1}X^r_{t-1}]E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}]E[X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}]\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2kl}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2rs}\\
&+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}]\omega_{2op}\omega_{2qr}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2kl}\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\omega_{2op}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2jk}\omega_{2qr}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2ij}\omega_{2qr}\\
&+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)E[X^q_{t-1}X^r_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2ij}\omega_{2kl}\\
&+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\omega_{2ij}\omega_{2kl}\omega_{2op}\omega_{2qr}
\end{aligned}
$$

We can now expand the fourth order central moments:

- $E[(S_{1i,n}-\omega_{1i})(S_{1j,n}-\omega_{1j})(S_{1k,n}-\omega_{1k})(S_{1l,n}-\omega_{1l})]$

$$
\begin{aligned}
=&E[S_{1i,n}S_{1j,n}S_{1k,n}S_{1l,n}] - \omega_{1i}E[S_{1j,n}S_{1k,n}S_{1l,n}] - \omega_{1j}E[S_{1i,n}S_{1k,n}S_{1l,n}] \\
&- \omega_{1k}E[S_{1i,n}S_{1j,n}S_{1l,n}] - \omega_{1l}E[S_{1i,n}S_{1j,n}S_{1k,n}] + \omega_{1i}\omega_{1j}E[S_{1k,n}S_{1l,n}] \\
&+ \omega_{1i}\omega_{1k}E[S_{1j,n}S_{1l,n}] + \omega_{1j}\omega_{1k}E[S_{1i,n}S_{1l,n}] + \omega_{1i}\omega_{1l}E[S_{1j,n}S_{1k,n}] \\
&+ \omega_{1j}\omega_{1l}E[S_{1i,n}S_{1k,n}] + \omega_{1k}\omega_{1l}E[S_{1i,n}S_{1j,n}] - \omega_{1i}\omega_{1j}\omega_{1k}E[S_{1l,n}] \\
&- \omega_{1i}\omega_{1j}\omega_{1l}E[S_{1k,n}] - \omega_{1i}\omega_{1k}\omega_{1l}E[S_{1j,n}] - \omega_{1j}\omega_{1k}\omega_{1l}E[S_{1i,n}] + \omega_{1i}\omega_{1j}\omega_{1k}\omega_{1l} \\
=&\frac{1}{n^3}\Big[E[Y_t^4X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^l] - E[Y_t^2X_{t-1}^iX_{t-1}^j]E[Y_t^2X_{t-1}^kX_{t-1}^l] \\
&- E[Y_t^2X_{t-1}^iX_{t-1}^k]E[Y_t^2X_{t-1}^jX_{t-1}^l] - E[Y_t^2X_{t-1}^iX_{t-1}^l]E[Y_t^2X_{t-1}^jX_{t-1}^k] \\
&- E[Y_t^3X_{t-1}^jX_{t-1}^kX_{t-1}^l]\omega_{1i} - E[Y_t^3X_{t-1}^iX_{t-1}^kX_{t-1}^l]\omega_{1j} \\
&- E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^l]\omega_{1k} - E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^k]\omega_{1l} \\
&+ 2E[Y_t^2X_{t-1}^iX_{t-1}^j]\omega_{1k}\omega_{1l} + 2E[Y_t^2X_{t-1}^iX_{t-1}^k]\omega_{1j}\omega_{1l} + 2E[Y_t^2X_{t-1}^jX_{t-1}^k]\omega_{1i}\omega_{1l} \\
&+ 2E[Y_t^2X_{t-1}^iX_{t-1}^l]\omega_{1j}\omega_{1k} + 2E[Y_t^2X_{t-1}^jX_{t-1}^l]\omega_{1i}\omega_{1k} + 2E[Y_t^2X_{t-1}^kX_{t-1}^l]\omega_{1i}\omega_{1j} \\
&- 6\omega_{1i}\omega_{1j}\omega_{1k}\omega_{1l}\Big] \\
&+ \frac{1}{n^2}\Big[E[Y_t^2X_{t-1}^iX_{t-1}^j]E[Y_t^2X_{t-1}^kX_{t-1}^l] + E[Y_t^2X_{t-1}^iX_{t-1}^k]E[Y_t^2X_{t-1}^jX_{t-1}^l] \\
&+ E[Y_t^2X_{t-1}^iX_{t-1}^l]E[Y_t^2X_{t-1}^jX_{t-1}^k] - E[Y_t^2X_{t-1}^iX_{t-1}^j]\omega_{1k}\omega_{1l} \\
&- E[Y_t^2X_{t-1}^iX_{t-1}^k]\omega_{1j}\omega_{1l} - E[Y_t^2X_{t-1}^jX_{t-1}^k]\omega_{1i}\omega_{1l} - E[Y_t^2X_{t-1}^iX_{t-1}^l]\omega_{1j}\omega_{1k} \\
&- E[Y_t^2X_{t-1}^jX_{t-1}^l]\omega_{1i}\omega_{1k} - E[Y_t^2X_{t-1}^kX_{t-1}^l]\omega_{1i}\omega_{1j} + 3\omega_{1i}\omega_{1j}\omega_{1k}\omega_{1l}\Big] \\
&\equiv \frac{1}{n^3}U^4_{1,ijkl} + \frac{1}{n^2}V^4_{1,ijkl}
\end{aligned}
$$

- $E[(S_{1i,n} - \omega_{1i})(S_{1j,n} - \omega_{1j})(S_{1k,n} - \omega_{1k})(S_{2lo,n} - \omega_{2lo})]$

$$
\begin{aligned}
=&E[S_{1i,n}S_{1j,n}S_{1k,n}S_{2lo,n}] - \omega_{1i}E[S_{1j,n}S_{1k,n}S_{2lo,n}] - \omega_{1j}E[S_{1i,n}S_{1k,n}S_{2lo,n}] \\
&- \omega_{1k}E[S_{1i,n}S_{1j,n}S_{2lo,n}] - \omega_{2lo}E[S_{1i,n}S_{1j,n}S_{1k,n}] + \omega_{1i}\omega_{1j}E[S_{1k,n}S_{2lo,n}] \\
&+ \omega_{1i}\omega_{1k}E[S_{1j,n}S_{2lo,n}] + \omega_{1j}\omega_{1k}E[S_{1i,n}S_{2lo,n}] + \omega_{1i}\omega_{2lo}E[S_{1j,n}S_{1k,n}] \\
&+ \omega_{1j}\omega_{2lo}E[S_{1i,n}S_{1k,n}] + \omega_{1k}\omega_{2lo}E[S_{1i,n}S_{1j,n}] - \omega_{1i}\omega_{1j}\omega_{1k}E[S_{2lo,n}] \\
&- \omega_{1i}\omega_{1j}\omega_{2lo}E[S_{1k,n}] - \omega_{1i}\omega_{1k}\omega_{2lo}E[S_{1j,n}] - \omega_{1j}\omega_{1k}\omega_{2lo}E[S_{1i,n}] + \omega_{1i}\omega_{1j}\omega_{1k}\omega_{2lo} \\
=&\frac{1}{n^3}\Big[E[Y_t^3X_{t-1}^iX_{t-1}^jX_{t-1}^kX_{t-1}^lX_{t-1}^o] - E[Y_t^2X_{t-1}^iX_{t-1}^j]E[Y_tX_{t-1}^kX_{t-1}^lX_{t-1}^o]
\end{aligned}
$$

$$
\begin{aligned}
&- E[Y_t^2 X_{t-1}^i X_{t-1}^k] E[Y_t X_{t-1}^j X_{t-1}^l X_{t-1}^o] - E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o] E[Y_t^2 X_{t-1}^j X_{t-1}^k] \\
&- E[Y_t^2 X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i} - E[Y_t^2 X_{t-1}^i X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1j} \\
&- E[Y_t^2 X_{t-1}^i X_{t-1}^j X_{t-1}^l X_{t-1}^o]\omega_{1k} - E[Y_t^3 X_{t-1}^i X_{t-1}^j X_{t-1}^k]\omega_{2lo} \\
&+ 2E[Y_t^2 X_{t-1}^i X_{t-1}^j]\omega_{1k}\omega_{2lo} + 2E[Y_t^2 X_{t-1}^i X_{t-1}^k]\omega_{1j}\omega_{2lo} + 2E[Y_t^2 X_{t-1}^j X_{t-1}^k]\omega_{1i}\omega_{2lo} \\
&+ 2E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]\omega_{1j}\omega_{1k} + 2E[Y_t X_{t-1}^j X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{1k} \\
&+ 2E[Y_t X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{1j} - 6\omega_{1i}\omega_{1j}\omega_{1k}\omega_{2lo}\Big] \\
&+ \frac{1}{n^2}\Big[E[Y_t^2 X_{t-1}^i X_{t-1}^j] E[Y_t X_{t-1}^k X_{t-1}^l X_{t-1}^o] + E[Y_t^2 X_{t-1}^i X_{t-1}^k] E[Y_t X_{t-1}^j X_{t-1}^l X_{t-1}^o] \\
&+ E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o] E[Y_t^2 X_{t-1}^j X_{t-1}^k] - E[Y_t^2 X_{t-1}^i X_{t-1}^j]\omega_{1k}\omega_{2lo} \\
&- E[Y_t^2 X_{t-1}^i X_{t-1}^k]\omega_{1j}\omega_{2lo} - E[Y_t^2 X_{t-1}^j X_{t-1}^k]\omega_{1i}\omega_{2lo} - E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]\omega_{1j}\omega_{1k} \\
&- E[Y_t X_{t-1}^j X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{1k} - E[Y_t X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{1j} + 3\omega_{1i}\omega_{1j}\omega_{1k}\omega_{2lo}\Big] \\
&\equiv \frac{1}{n^3} U_{2,ijklo}^4 + \frac{1}{n^2} V_{2,ijklo}^4
\end{aligned}
$$

- $E[(S_{1i,n} - \omega_{1i})(S_{1j,n} - \omega_{1j})(S_{2kl,n} - \omega_{2kl})(S_{2op,n} - \omega_{op})]$

$$
\begin{aligned}
=&E[S_{1i,n}S_{1j,n}S_{2kl,n}S_{2op,n}] - \omega_{1i}E[S_{1j,n}S_{2kl,n}S_{2op,n}] - \omega_{1j}E[S_{1i,n}S_{2kl,n}S_{2op,n}] \\
&- \omega_{2kl}E[S_{1i,n}S_{1j,n}S_{2op,n}] - \omega_{op}E[S_{1i,n}S_{1j,n}S_{2kl,n}] + \omega_{1i}\omega_{1j}E[S_{2kl,n}S_{2op,n}] \\
&+ \omega_{1i}\omega_{2kl}E[S_{1j,n}S_{2op,n}] + \omega_{1j}\omega_{2kl}E[S_{1i,n}S_{2op,n}] + \omega_{1i}\omega_{op}E[S_{1j,n}S_{2kl,n}] \\
&+ \omega_{1j}\omega_{op}E[S_{1i,n}S_{2kl,n}] + \omega_{2kl}\omega_{op}E[S_{1i,n}S_{1j,n}] - \omega_{1i}\omega_{1j}\omega_{2kl}E[S_{2op,n}] \\
&- \omega_{1i}\omega_{1j}\omega_{op}E[S_{2kl,n}] - \omega_{1i}\omega_{2kl}\omega_{op}E[S_{1j,n}] - \omega_{1j}\omega_{2kl}\omega_{op}E[S_{1i,n}] + \omega_{1i}\omega_{1j}\omega_{2kl}\omega_{op} \\
=&\frac{1}{n^3}\Big[E[Y_t^2 X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p] - E[Y_t^2 X_{t-1}^i X_{t-1}^j] E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p] \\
&- E[Y_t X_{t-1}^i X_{t-1}^k X_{t-1}^l] E[Y_t X_{t-1}^j X_{t-1}^o X_{t-1}^p] \\
&- E[Y_t X_{t-1}^i X_{t-1}^o X_{t-1}^p] E[Y_t X_{t-1}^j X_{t-1}^k X_{t-1}^l] \\
&- E[Y_t X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{1i} - E[Y_t X_{t-1}^i X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{1j} \\
&- E[Y_t^2 X_{t-1}^i X_{t-1}^j X_{t-1}^o X_{t-1}^p]\omega_{2kl} - E[Y_t^2 X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l]\omega_{2op} \\
&+ 2E[Y_t^2 X_{t-1}^i X_{t-1}^j]\omega_{2kl}\omega_{2op} + 2E[Y_t X_{t-1}^i X_{t-1}^k X_{t-1}^l]\omega_{1j}\omega_{2op} \\
&+ 2E[Y_t X_{t-1}^j X_{t-1}^k X_{t-1}^l]\omega_{1i}\omega_{2op} + 2E[Y_t X_{t-1}^i X_{t-1}^o X_{t-1}^p]\omega_{1j}\omega_{2kl}
\end{aligned}
$$

$$+ 2E[Y_t X_{t-1}^j X_{t-1}^o X_{t-1}^p]\omega_{1i}\omega_{2kl} + 2E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{1i}\omega_{1j}$$
$$- 6\omega_{1i}\omega_{1j}\omega_{2kl}\omega_{2op}\Big]$$
$$+ \frac{1}{n^2}\Big[E[Y_t^2 X_{t-1}^i X_{t-1}^j]E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]$$
$$+ E[Y_t X_{t-1}^i X_{t-1}^k X_{t-1}^l]E[Y_t X_{t-1}^j X_{t-1}^o X_{t-1}^p]$$
$$+ E[Y_t X_{t-1}^i X_{t-1}^o X_{t-1}^p]E[Y_t X_{t-1}^j X_{t-1}^k X_{t-1}^l] - E[Y_t^2 X_{t-1}^i X_{t-1}^j]\omega_{2kl}\omega_{2op}$$
$$- E[Y_t X_{t-1}^i X_{t-1}^k X_{t-1}^l]\omega_{1j}\omega_{2op} - E[Y_t X_{t-1}^j X_{t-1}^k X_{t-1}^l]\omega_{1i}\omega_{2op}$$
$$- E[Y_t X_{t-1}^i X_{t-1}^o X_{t-1}^p]\omega_{1j}\omega_{2kl} - E[Y_t X_{t-1}^j X_{t-1}^o X_{t-1}^p]\omega_{1i}\omega_{2kl}$$
$$- E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{1i}\omega_{1j} + 3\omega_{1i}\omega_{1j}\omega_{2kl}\omega_{2op}\Big]$$
$$\equiv \frac{1}{n^3}U_{3,ijklop}^4 + \frac{1}{n^2}V_{3,ijklop}^4$$

- $E[(S_{1i,n} - \omega_{1i})(S_{2jk,n} - \omega_{2jk})(S_{2lo,n} - \omega_{2lo})(S_{2pq,n} - \omega_{2pq})]$

$$=E[S_{1i,n}S_{2jk,n}S_{2lo,n}S_{2pq,n}] - \omega_{1i}E[S_{2jk,n}S_{2lo,n}S_{2pq,n}] - \omega_{2jk}E[S_{1i,n}S_{2lo,n}S_{2pq,n}]$$
$$- \omega_{2lo}E[S_{1i,n}S_{2jk,n}S_{2pq,n}] - \omega_{2pq}E[S_{1i,n}S_{2jk,n}S_{2lo,n}] + \omega_{1i}\omega_{2jk}E[S_{2lo,n}S_{2pq,n}]$$
$$+ \omega_{1i}\omega_{2lo}E[S_{2jk,n}S_{2pq,n}] + \omega_{2jk}\omega_{2lo}E[S_{1i,n}S_{2pq,n}] + \omega_{1i}\omega_{2pq}E[S_{2jk,n}S_{2lo,n}]$$
$$+ \omega_{2jk}\omega_{2pq}E[S_{1i,n}S_{2lo,n}] + \omega_{2lo}\omega_{2pq}E[S_{1i,n}S_{2jk,n}] - \omega_{1i}\omega_{2jk}\omega_{2lo}E[S_{2pq,n}]$$
$$- \omega_{1i}\omega_{2jk}\omega_{2pq}E[S_{2lo,n}] - \omega_{1i}\omega_{2lo}\omega_{2pq}E[S_{2jk,n}] - \omega_{2jk}\omega_{2lo}\omega_{2pq}E[S_{1i,n}]$$
$$+ \omega_{1i}\omega_{2jk}\omega_{2lo}\omega_{2pq}$$
$$=\frac{1}{n^3}\Big[E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]$$
$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]E[X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]$$
$$- E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]E[X_{t-1}^j X_{t-1}^k X_{t-1}^p X_{t-1}^q]$$
$$- E[Y_t X_{t-1}^i X_{t-1}^p X_{t-1}^q]E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]$$
$$- E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]\omega_{1i} - E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]\omega_{2jk}$$
$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^p X_{t-1}^q]\omega_{2lo} - E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{2pq}$$
$$+ 2E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]\omega_{2lo}\omega_{2pq} + 2E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]\omega_{2jk}\omega_{2pq}$$
$$+ 2E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{2pq} + 2E[Y_t X_{t-1}^i X_{t-1}^p X_{t-1}^q]\omega_{2jk}\omega_{2lo}$$

$$+ 2E[X_{t-1}^j X_{t-1}^k X_{t-1}^p X_{t-1}^q]\omega_{1i}\omega_{2lo} + 2E[X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]\omega_{1i}\omega_{2jk}$$

$$- 6\omega_{1i}\omega_{2jk}\omega_{2lo}\omega_{2pq}\Big]$$

$$+ \frac{1}{n^2}\Big[E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]E[X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]$$

$$+ E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]E[X_{t-1}^j X_{t-1}^k X_{t-1}^p X_{t-1}^q]$$

$$+ E[Y_t X_{t-1}^i X_{t-1}^p X_{t-1}^q]E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o] - E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k]\omega_{2lo}\omega_{2pq}$$

$$- E[Y_t X_{t-1}^i X_{t-1}^l X_{t-1}^o]\omega_{2jk}\omega_{2pq} - E[X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o]\omega_{1i}\omega_{2pq}$$

$$- E[Y_t X_{t-1}^i X_{t-1}^p X_{t-1}^q]\omega_{2jk}\omega_{2lo} - E[X_{t-1}^j X_{t-1}^k X_{t-1}^p X_{t-1}^q]\omega_{1i}\omega_{2lo}$$

$$- E[X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q]\omega_{1i}\omega_{2jk} + 3\omega_{1i}\omega_{2jk}\omega_{2lo}\omega_{2pq}\Big]$$

$$\equiv \frac{1}{n^3}U_{4,ijklopq}^4 + \frac{1}{n^2}V_{4,ijklopq}^4$$

- $E[(S_{2ij,n} - \omega_{2ij})(S_{2kl,n} - \omega_{2kl})(S_{2op,n} - \omega_{2op})(S_{2qr,n} - \omega_{2qr})]$

$$=E[S_{2ij,n}S_{2kl,n}S_{2op,n}S_{2qr,n}] - \omega_{2ij}E[S_{2kl,n}S_{2op,n}S_{2qr,n}] - \omega_{2kl}E[S_{2ij,n}S_{2op,n}S_{2qr,n}]$$

$$- \omega_{2op}E[S_{2ij,n}S_{2kl,n}S_{2qr,n}] - \omega_{2qr}E[S_{2ij,n}S_{2kl,n}S_{2op,n}] + \omega_{2ij}\omega_{2kl}E[S_{2op,n}S_{2qr,n}]$$

$$+ \omega_{2ij}\omega_{2op}E[S_{2kl,n}S_{2qr,n}] + \omega_{2kl}\omega_{2op}E[S_{2ij,n}S_{2qr,n}] + \omega_{2ij}\omega_{2qr}E[S_{2kl,n}S_{2op,n}]$$

$$+ \omega_{2kl}\omega_{2qr}E[S_{2ij,n}S_{2op,n}] + \omega_{2op}\omega_{2qr}E[S_{2ij,n}S_{2kl,n}] - \omega_{2ij}\omega_{2kl}\omega_{2op}E[S_{2qr,n}]$$

$$- \omega_{2ij}\omega_{2kl}\omega_{2qr}E[S_{2op,n}] - \omega_{2ij}\omega_{2op}\omega_{2qr}E[S_{2kl,n}] - \omega_{2kl}\omega_{2op}\omega_{2qr}E[S_{2ij,n}]$$

$$+ \omega_{2ij}\omega_{2kl}\omega_{2op}\omega_{2qr}$$

$$=\frac{1}{n^3}\Big[E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q X_{t-1}^r]$$

$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l]E[X_{t-1}^o X_{t-1}^p X_{t-1}^q X_{t-1}^r]$$

$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^o X_{t-1}^p]E[X_{t-1}^k X_{t-1}^l X_{t-1}^q X_{t-1}^r]$$

$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^q X_{t-1}^r]E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]$$

$$- E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p X_{t-1}^q X_{t-1}^r]\omega_{2ij} - E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^o X_{t-1}^p X_{t-1}^q X_{t-1}^r]\omega_{2kl}$$

$$- E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^q X_{t-1}^r]\omega_{2op} - E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{2qr}$$

$$+ 2E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^k X_{t-1}^l]\omega_{2op}\omega_{2qr} + 2E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^o X_{t-1}^p]\omega_{2kl}\omega_{2qr}$$

$$+ 2E[X_{t-1}^k X_{t-1}^l X_{t-1}^o X_{t-1}^p]\omega_{2ij}\omega_{2qr} + 2E[Y_t X_{t-1}^i X_{t-1}^j X_{t-1}^q X_{t-1}^r]\omega_{2kl}\omega_{2op}$$

$$
\begin{aligned}
&+ 2E[X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\omega_{2op} + 2E[X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\omega_{2kl} \\
&- 6\omega_{2ij}\omega_{2kl}\omega_{2op}\omega_{2qr}\Big] \\
&+ \frac{1}{n^2}\Big[E[Y_tX^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}]E[X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}] \\
&+ E[Y_tX^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}]E[X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}] \\
&+ E[Y_tX^i_{t-1}X^j_{t-1}X^q_{t-1}X^r_{t-1}]E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}] \\
&- E[Y_tX^i_{t-1}X^j_{t-1}X^k_{t-1}X^l_{t-1}]\omega_{2op}\omega_{2qr} \\
&- E[Y_tX^i_{t-1}X^j_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2kl}\omega_{2qr} - E[X^k_{t-1}X^l_{t-1}X^o_{t-1}X^p_{t-1}]\omega_{2ij}\omega_{2qr} \\
&- E[Y_tX^i_{t-1}X^j_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2kl}\omega_{2op} - E[X^k_{t-1}X^l_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\omega_{2op} \\
&- E[X^o_{t-1}X^p_{t-1}X^q_{t-1}X^r_{t-1}]\omega_{2ij}\omega_{2kl} + 3\omega_{2ij}\omega_{2kl}\omega_{2op}\omega_{2qr}\Big] \\
&\equiv \frac{1}{n^3}U^4_{5,ijklopqr} + \frac{1}{n^2}V^4_{5,ijklopqr}.
\end{aligned}
$$

C.5.2 Re-expressing $E[(b_n - \bar{b})^{\otimes i}]$

In this section, we derive an expression for the term $E[(b_n - \bar{b})^{\otimes i}]$ for $i = 2, 3, 4$ with an explicit dependence on the sample size n. To obtain this expression, we make use of the central moments derived in Section C.5.1. We assume $k = m$, the number of independent variables, $X^1_t, \dots, X^m_t$. Recall the statistics $S_{1,n}$ and $S_{2,n}$ defined as follows:

$$
S_{1,n} \equiv \frac{1}{n}X^\top_{t,n}Y_{t,n} \in \mathbb{R}^{m\times 1}, \quad S_{2,n} \equiv \frac{1}{n}X^\top_{t,n}X_{t,n} \in \mathbb{R}^{m\times m}. \tag{C.5.1}
$$

These can be expressed as follows:

$$
S_{1,n} = \begin{bmatrix} \frac{1}{n}\sum_{\tau=t-n}^{t-1} X^1_\tau Y_{\tau+1} \\ \vdots \\ \frac{1}{n}\sum_{\tau=t-n}^{t-1} X^m_\tau Y_{\tau+1} \end{bmatrix} \equiv \begin{bmatrix} S_{11,n} \\ \cdots \\ S_{1m,n} \end{bmatrix},
$$

$$S_{2,n} = \begin{bmatrix} \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^1 X_\tau^1 & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^1 X_\tau^2 & \cdots & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^1 X_\tau^m \\ \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2 X_\tau^1 & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2 X_\tau^2 & \cdots & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^2 X_\tau^m \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^m X_\tau^1 & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^m X_\tau^2 & \cdots & \frac{1}{n}\sum_{\tau=t-n}^{t-1} X_\tau^m X_\tau^m \end{bmatrix}$$

$$\equiv \begin{bmatrix} S_{211,n} & S_{212,n} & \cdots & S_{21m,n} \\ S_{221,n} & S_{222,n} & \cdots & S_{22m,n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{2m1,n} & S_{2m2,n} & \cdots & S_{2mm,n} \end{bmatrix}.$$

One can now express the vectors b_n and $\bar{b}$ as follows:

$$b_n = \left(\frac{S_{1,n}}{\text{cs } S_{2,n}} \right) = \begin{pmatrix} S_{11,n} \\ \vdots \\ S_{1m,n} \\ S_{211,n} \\ \vdots \\ S_{2mm,n} \end{pmatrix}, \quad \bar{b} = \left(\frac{\omega_1}{\text{cs } \omega_2} \right) = \begin{pmatrix} \omega_{11} \\ \vdots \\ \omega_{1m} \\ \omega_{211} \\ \vdots \\ \omega_{2mm} \end{pmatrix}.$$

To find an expression for $(b_n - \bar{b})^{\otimes 2}$, we define $\delta_1 = b_n - \bar{b}$ and it follows:

$$(b_n - \bar{b})^{\otimes 2} = \begin{pmatrix} (S_{11,n} - \omega_{11})\delta_1 \\ \hline \vdots \\ \hline (S_{1m,n} - \omega_{1m})\delta_1 \\ \hline (S_{211,n} - \omega_{211})\delta_1 \\ \hline \vdots \\ \hline (S_{2mm,n} - \omega_{2mm})\delta_1 \end{pmatrix} \equiv \delta_2 \in \mathbb{R}^{m^2(m+1)^2 \times 1}.$$

In Section $C.5.1$, we defined the terms $V^2_{1,ij}$, $V^2_{2,ijk}$, and $V^2_{3,ijkl}$. In what follows, we present notation to express $E[(b_n - \bar{b})^{\otimes 2}]$ in terms of $V^2_{1,ij}$, $V^2_{2,ijk}$, and $V^2_{3,ijkl}$. To begin, we note, given $i, j, k, l = 1, \ldots, m$, that $V^2_{1,ij}$ represents m^2 elements, $V^2_{2,ijk}$ represents m^3 elements and $V^2_{3,ijkl}$ represents m^4 elements. Our notation is meant to manipulate the the different elements of $V^2_{1,ij}$, $V^2_{2,ijk}$, and $V^2_{3,ijkl}$ into matrices and vectors of different shapes and sizes.

To illustrate the notation, we define the following m $m \times 1$ vectors and one $m \times m$ matrix as follows:

$$V^2_{1,i[j]} \equiv \begin{pmatrix} V^2_{1,i1} \\ \vdots \\ V^2_{1,im} \end{pmatrix} \in \mathbb{R}^{m\times 1}, \quad i = 1, \ldots, m,$$

$$V^2_{1,[ij]} \equiv \begin{pmatrix} V^2_{1,11} & \cdots & V^2_{1,1m} \\ \cdots & \ddots & \cdots \\ V^2_{1,m1} & \cdots & V^2_{1,mm} \end{pmatrix} \in \mathbb{R}^{m\times m}.$$

In this notation, the index within the bracket runs from 1 to m. Nested brackets are evaluated from the outside in as in the following case:

$$V^2_{1,[i[j]]} \equiv \begin{pmatrix} V^2_{1,1[i]} \\ \hline V^2_{1,2[i]} \\ \hline \vdots \\ \hline V^2_{1,m[i]} \end{pmatrix} \in \mathbb{R}^{m^2\times 1}.$$

cs $[ij]$ indicates the column string of the matrix indexed by ij, as follows:

$$V^2_{1,\mathrm{cs}\ [ij]} \equiv \begin{pmatrix} V^2_{1,11} \\ V^2_{1,21} \\ \vdots \\ V^2_{1,mm} \end{pmatrix} \in \mathbb{R}^{m^2\times 1}.$$

Given the symmetry $V^2_{1,ij} = V^2_{1,ji}$, it follows $V^2_{1,[i[j]]} = V^2_{1,\text{cs }[ij]}$. We define a set of matrices:

$$E_{1,1} \equiv \begin{pmatrix} \underset{m\times m}{I} & 0 & \cdots & 0 \\ \underset{m^2\times m}{0} & 0 & \cdots & 0 \\ \underset{m\times m}{0} & \underset{m\times m}{I} & \cdots & 0 \\ \underset{m^2\times m}{0} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underset{m\times m}{0} & 0 & \cdots & \underset{m\times m}{I} \\ \underset{m^2\times m}{0} & 0 & \cdots & \underset{m^2\times m}{0} \end{pmatrix} \in \mathbb{R}^{m^2(m+1)\times m^2},$$

$$Z_{11} \equiv \underset{m^3(m+1)\times m^2}{0}, \quad E_{2,1} \equiv \left(\frac{E_{1,1}}{Z_{11}}\right) \in \mathbb{R}^{m^2(m+1)^2\times m^2},$$

$$E_{1,2} \equiv \begin{pmatrix} \underset{m\times m^2}{0} & 0 & \cdots & 0 \\ \underset{m^2\times m^2}{I} & 0 & \cdots & 0 \\ \underset{m\times m^2}{0} & \underset{m\times m^2}{0} & \cdots & 0 \\ \underset{m^2\times m^2}{0} & \underset{m^2\times m^2}{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underset{m\times m^2}{0} & 0 & \cdots & \underset{m\times m^2}{0} \\ \underset{m^2\times m^2}{0} & 0 & \cdots & \underset{m^2\times m^2}{I} \end{pmatrix} \in \mathbb{R}^{m^2(m+1)\times m^3},$$

$$Z_{12} \equiv \underset{m^3(m+1)\times m^3}{0}, \quad E_{2,2} \equiv \left(\frac{E_{1,2}}{Z_{12}}\right) \in \mathbb{R}^{m^2(m+1)^2\times m^3},$$

$$E_{1,3} \equiv \begin{pmatrix} \underset{m\times m}{I} & 0 & \cdots & 0 \\ \underset{m^2\times m}{0} & 0 & \cdots & 0 \\ \underset{m\times m}{0} & \underset{m\times m}{I} & \cdots & 0 \\ \underset{m^2\times m}{0} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underset{m\times m}{0} & 0 & \cdots & \underset{m\times m}{I} \\ \underset{m^2\times m}{0} & 0 & \cdots & \underset{m^2\times m}{0} \end{pmatrix} \in \mathbb{R}^{m^3(m+1)\times m^3},$$

$$Z_{13} \equiv \underset{m^2(m+1)\times m^3}{0}, \quad E_{2,3} \equiv \begin{pmatrix} Z_{13} \\ \hline E_{1,3} \end{pmatrix} \in \mathbb{R}^{m^2(m+1)^2\times m^3},$$

$$E_{1,4} \equiv \begin{pmatrix} \underset{m\times m^2}{0} & 0 & \cdots & 0 \\ \underset{m^2\times m^2}{I} & 0 & \cdots & 0 \\ \underset{m\times m^2}{0} & \underset{m\times m^2}{0} & \cdots & 0 \\ \underset{m^2\times m^2}{0} & \underset{m^2\times m^2}{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \underset{m\times m^2}{0} & 0 & \cdots & \underset{m\times m^2}{0} \\ \underset{m^2\times m^2}{0} & 0 & \cdots & \underset{m^2\times m^2}{I} \end{pmatrix} \in \mathbb{R}^{m^3(m+1)\times m^4},$$

$$Z_{14} \equiv \underset{m^2(m+1)\times m^4}{0}, \quad E_{2,4} \equiv \begin{pmatrix} Z_{14} \\ \hline E_{1,4} \end{pmatrix} \in \mathbb{R}^{m^2(m+1)^2\times m^4}.$$

With these matrices we rewrite the term $E[(b_n - \bar{b})^{\otimes 2}]$ as follows:

$$E[(b_n - \bar{b})^{\otimes 2}] = \frac{1}{n}\Big[E_{2,1}\, V^2_{1,[i[j]]} + E_{2,2}\, V^2_{2,[i \text{ cs } [jk]]} + E_{2,3}\, V^2_{2, \text{ cs } [[i]jk]} \\ + E_{2,4}\, V^2_{3, \text{ cs } [ij \text{ cs } [kl]]}\Big].$$

Next, to find an expression for $(b_n - \bar{b})^{\otimes 3}$ we have

$$(b_n - \bar{b})^{\otimes 3} = \left(\begin{array}{c} (S_{11,n} - \omega_{11})\delta_2 \\ \hline \vdots \\ \hline (S_{1m,n} - \omega_{1m})\delta_2 \\ \hline (S_{211,n} - \omega_{211})\delta_2 \\ \hline \vdots \\ \hline (S_{2mm,n} - \omega_{2mm})\delta_2 \end{array} \right) \equiv \delta_3 \in \mathbb{R}^{m^3(m+1)^3 \times 1}.$$

In Section $C.5.1$, we defined the terms $V^3_{1,ijk}$, $V^3_{2,ijkl}$, $V^3_{3,ijklo}$, and $V^3_{4,ijklop}$. Given each of the index i, j, k, l, o, p run from 1 to m, it follows $V^3_{1,ijk}$ represents m^3 elements, $V^3_{2,ijkl}$ represents m^4 elements, $V^3_{3,ijklo}$ represents m^5 elements, and $V^3_{4,ijklop}$ represents m^6 elements. We use the previous subscript notation on the index elements to form matrices and vectors of different sizes. We define another set of matrices:

$$E_{3,1} \equiv Diag[E_{2,1} \cdots E_{2,1}] \in \mathbb{R}^{m^3(m+1)^2 \times m^3}, \quad Z_{21} \equiv \underset{m^4(m+1)^2 \times m^3}{0},$$

$$E_{4,1} \equiv \left(\begin{array}{c} E_{3,1} \\ \hline Z_{21} \end{array} \right) \in \mathbb{R}^{m^3(m+1)^3 \times m^3},$$

$$E_{3,2} \equiv Diag[E_{2,2} \cdots E_{2,2}] \in \mathbb{R}^{m^3(m+1)^2 \times m^4}, \quad Z_{22} \equiv \underset{m^4(m+1)^2 \times m^4}{0},$$

$$E_{4,2} \equiv \left(\begin{array}{c} E_{3,2} \\ \hline Z_{22} \end{array} \right) \in \mathbb{R}^{m^4(m+1)^2 \times m^4},$$

$$E_{3,3} \equiv Diag[E_{2,3} \cdots E_{2,3}] \in \mathbb{R}^{m^3(m+1)^2 \times m^4},$$

$$E_{4,3} \equiv \left(\begin{array}{c} E_{3,3} \\ \hline Z_{22} \end{array} \right) \in \mathbb{R}^{m^3(m+1)^3 \times m^4},$$

$$E_{3,4} \equiv Diag[E_{2,4} \cdots E_{2,4}] \in \mathbb{R}^{m^3(m+1)^2 \times m^5}, \quad Z_{23} \equiv \underset{m^4(m+1)^2 \times m^5}{0},$$

$$E_{4,4} \equiv \left(\begin{array}{c} E_{3,4} \\ \hline Z_{23} \end{array} \right) \in \mathbb{R}^{m^3(m+1)^3 \times m^5},$$

$$E_{3,5} \equiv Diag[E_{2,1} \cdots E_{2,1}] \in \mathbb{R}^{m^4(m+1)^2 \times m^4}, \quad Z_{24} \equiv \underset{m^3(m+1)^2 \times m^4}{0},$$

$$E_{4,5} \equiv \left(\frac{Z_{24}}{E_{3,5}}\right) \in \mathbb{R}^{m^3(m+1)^3 \times m^4},$$

$$E_{3,6} \equiv Diag[E_{2,2} \cdots E_{2,2}] \in \mathbb{R}^{m^4(m+1)^2 \times m^5}, \quad Z_{25} \equiv \underset{m^3(m+1)^2 \times m^5}{0},$$

$$E_{4,6} \equiv \left(\frac{Z_{25}}{E_{3,6}}\right) \in \mathbb{R}^{m^3(m+1)^3 \times m^5},$$

$$E_{3,7} \equiv Diag[E_{2,3} \cdots E_{2,3}] \in \mathbb{R}^{m^4(m+1)^2 \times m^5},$$

$$E_{4,7} \equiv \left(\frac{Z_{25}}{E_{3,7}}\right) \in \mathbb{R}^{m^3(m+1)^3 \times m^5},$$

$$E_{3,8} \equiv Diag[E_{2,4} \cdots E_{2,4}] \in \mathbb{R}^{m^4(m+1)^2 \times m^6}, \quad Z_{26} \equiv \underset{m^3(m+1)^2 \times m^6}{0},$$

$$E_{4,8} \equiv \left(\frac{Z_{26}}{E_{3,8}}\right) \in \mathbb{R}^{m^3(m+1)^3 \times m^6}.$$

With all the previously defined matrices, we rewrite the term $E[(b_n - \bar{b})^{\otimes 3}]$ as follows:

$$\begin{aligned} E[(b_n - \bar{b})^{\otimes 3}] =& \frac{1}{n^2}\Big[E_{4,1}\, V^3_{1,[i[j[k]]]} + E_{4,2}\, V^3_{2,[i[j\ \mathrm{cs}[kl]]]} + E_{4,3}\, V^3_{2,[i\ \mathrm{cs}[[j]kl]]} \\ & + E_{4,4}\, V^3_{3,[i\ \mathrm{cs}[jk\ \mathrm{cs}[lo]]]} + E_{4,5}\, V^3_{2,[[i[j]]kl]} + E_{4,6}\, V^3_{3,\ \mathrm{cs}[[i\ \mathrm{cs}[jk]]lo]} \\ & + E_{4,7}\, V^3_{3,\ \mathrm{cs}[\ \mathrm{cs}[[i]jk]lo]} + E_{4,8}\, V^3_{4,\ \mathrm{cs}[ij\ \mathrm{cs}[kl\ \mathrm{cs}[op]]]}\Big]. \end{aligned}$$

Next, to find an expression for $(b_n - \bar{b})^{\otimes 4}$, we have

$$(b_n - \bar{b})^{\otimes 4} = \begin{pmatrix} (S_{11,n} - \omega_{11})\delta_3 \\ \hline \vdots \\ \hline (S_{1m,n} - \omega_{1m})\delta_3 \\ \hline (S_{211,n} - \omega_{211})\delta_3 \\ \hline \vdots \\ \hline (S_{2mm,n} - \omega_{2mm})\delta_3 \end{pmatrix} \equiv \delta_4 \in \mathbb{R}^{m^4(m+1)^4 \times 1}.$$

In section $C.5.1$, we defined the terms $V^4_{1,ijkl}$, $V^4_{2,ijklo}$, $V^4_{3,ijklop}$, $V^4_{4,ijklopq}$, and $V^4_{5,ijklopqr}$. Given each of the index i, j, k, l, o, p, q, r run from 1 to m, it follows $V^4_{1,ijkl}$ represents m^4 elements, $V^4_{2,ijklo}$ represents m^5 elements, $V^4_{3,ijklop}$ represents m^6 elements, $V^4_{4,ijklopq}$

represents m^7 elements, and $V^4_{5,ijklopqr}$ represents m^8 elements. We use the previous subscript notation on the index elements to form matrices and vectors of different sizes. We define another set of matrices:

$$E_{5,1} \equiv Diag[E_{4,1} \cdots E_{4,1}] \in \mathbb{R}^{m^4(m+1)^3 \times m^4}, \quad Z_{31} \equiv \underset{m^5(m+1)^3 \times m^4}{0},$$

$$E_{6,1} \equiv \left(\frac{E_{5,1}}{Z_{31}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^4},$$

$$E_{5,2} \equiv Diag[E_{4,2} \cdots E_{4,2}] \in \mathbb{R}^{m^4(m+1)^3 \times m^5}, \quad Z_{32} \equiv \underset{m^5(m+1)^3 \times m^5}{0},$$

$$E_{6,2} \equiv \left(\frac{E_{5,2}}{Z_{32}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^5},$$

$$E_{5,3} \equiv Diag[E_{4,3} \cdots E_{4,3}] \in \mathbb{R}^{m^4(m+1)^3 \times m^5},$$

$$E_{6,3} \equiv \left(\frac{E_{5,3}}{Z_{32}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^5},$$

$$E_{5,4} \equiv Diag[E_{4,4} \cdots E_{4,4}] \in \mathbb{R}^{m^4(m+1)^3 \times m^6}, \quad Z_{33} \equiv \underset{m^5(m+1)^3 \times m^6}{0},$$

$$E_{6,4} \equiv \left(\frac{E_{5,4}}{Z_{33}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,5} \equiv Diag[E_{4,5} \cdots E_{4,5}] \in \mathbb{R}^{m^4(m+1)^3 \times m^5},$$

$$E_{6,5} \equiv \left(\frac{E_{5,5}}{Z_{32}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^5},$$

$$E_{5,6} \equiv Diag[E_{4,6} \cdots E_{4,6}] \in \mathbb{R}^{m^4(m+1)^3 \times m^6},$$

$$E_{6,6} \equiv \left(\frac{E_{5,6}}{Z_{33}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,7} \equiv Diag[E_{4,7} \cdots E_{4,7}] \in \mathbb{R}^{m^4(m+1)^3 \times m^6},$$

$$E_{6,7} \equiv \left(\frac{E_{5,7}}{Z_{33}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,8} \equiv Diag[E_{4,8} \cdots E_{4,8}] \in \mathbb{R}^{m^4(m+1)^3 \times m^7}, \quad Z_{34} \equiv \underset{m^5(m+1)^3 \times m^7}{0},$$

$$E_{6,8} \equiv \left(\frac{E_{5,8}}{Z_{34}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^7},$$

$$E_{5,9} \equiv Diag[E_{4,1} \cdots E_{4,1}] \in \mathbb{R}^{m^5(m+1)^3 \times m^5}, \quad Z_{35} \equiv \underset{m^4(m+1)^3 \times m^5}{0},$$

$$E_{6,9} \equiv \left(\frac{Z_{35}}{E_{5,9}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^5},$$

$$E_{5,10} \equiv Diag[E_{4,2} \cdots E_{4,2}] \in \mathbb{R}^{m^5(m+1)^3 \times m^6}, \quad Z_{36} \equiv \underset{m^4(m+1)^3 \times m^6}{0},$$

$$E_{6,10} \equiv \left(\frac{Z_{36}}{E_{5,10}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,11} \equiv Diag[E_{4,3} \cdots E_{4,3}] \in \mathbb{R}^{m^5(m+1)^3 \times m^6},$$

$$E_{6,11} \equiv \left(\frac{Z_{36}}{E_{5,11}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,12} \equiv Diag[E_{4,4} \cdots E_{4,4}] \in \mathbb{R}^{m^5(m+1)^3 \times m^7}, \quad Z_{37} \equiv \underset{m^4(m+1)^3 \times m^7}{0},$$

$$E_{6,12} \equiv \left(\frac{Z_{37}}{E_{5,12}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^7},$$

$$E_{5,13} \equiv Diag[E_{4,5} \cdots E_{4,5}] \in \mathbb{R}^{m^5(m+1)^3 \times m^6},$$

$$E_{6,13} \equiv \left(\frac{Z_{36}}{E_{5,13}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^6},$$

$$E_{5,14} \equiv Diag[E_{4,6} \cdots E_{4,6}] \in \mathbb{R}^{m^5(m+1)^3 \times m^7},$$

$$E_{6,14} \equiv \left(\frac{Z_{37}}{E_{5,14}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^7},$$

$$E_{5,15} \equiv Diag[E_{4,7} \cdots E_{4,7}] \in \mathbb{R}^{m^5(m+1)^3 \times m^7},$$

$$E_{6,15} \equiv \left(\frac{Z_{37}}{E_{5,15}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^7},$$

$$E_{5,16} \equiv Diag[E_{4,8} \cdots E_{4,8}] \in \mathbb{R}^{m^5(m+1)^3 \times m^8}, \quad Z_{38} \equiv \underset{m^4(m+1)^3 \times m^8}{0},$$

$$E_{6,16} \equiv \left(\frac{Z_{38}}{E_{5,16}} \right) \in \mathbb{R}^{m^4(m+1)^4 \times m^8}.$$

With all the previously defined matrices, we rewrite the second order term of $E[(b_n - \bar{b})^{\otimes 4}]$ as follows:

$$
\begin{aligned}
E[(b_n - \bar{b})^{\otimes 4}|_2] =& \frac{1}{n^2}\Big[E_{6,1}\, V^4_{1,[i[j[k[l]]]]} + E_{6,2}\, V^4_{2,[i[j[k\ \mathrm{cs}[lo]]]]} + E_{6,3}\, V^4_{2,[i[j\ \mathrm{cs}[[k]lo]]]} \\
&+ E_{6,4}\, V^4_{3,[i[j\ \mathrm{cs}[kl\ \mathrm{cs}[op]]]]} + E_{6,5}\, V^4_{2,[i\ \mathrm{cs}[[j[k]]lo]]} + E_{6,6}\, V^4_{3,[i\ \mathrm{cs}[[j\ \mathrm{cs}[kl]]lo]]} \\
&+ E_{6,7}\, V^4_{3,[i\ \mathrm{cs}[\ \mathrm{cs}[[j]kl]op]]} + E_{6,8}\, V^4_{4,[i\ \mathrm{cs}[jk\ \mathrm{cs}[lo\ \mathrm{cs}[pq]]]]} \\
&+ E_{6,9}\, V^4_{2,\ \mathrm{cs}[[i[j[k]]]lo]} + E_{6,10}\, V^4_{3,\ \mathrm{cs}[[i[j\ \mathrm{cs}[kl]]]op]} \\
&+ E_{6,11}\, V^4_{3,\ \mathrm{cs}[[i\ \mathrm{cs}[[j]kl]]op]} + E_{6,12}\, V^4_{4,\ \mathrm{cs}[[i\ \mathrm{cs}[jk\ \mathrm{cs}[lo]]]pq]} \\
&+ E_{6,13}\, V^4_{3,\ \mathrm{cs}[\ \mathrm{cs}[[i[j]]kl]op]} + E_{6,14}\, V^4_{4,\ \mathrm{cs}[\ \mathrm{cs}[[i\ \mathrm{cs}[jk]]lo]pq]} \\
&+ E_{6,15}\, V^4_{4,\ \mathrm{cs}[\ \mathrm{cs}[\ \mathrm{cs}[[i]jk]lo]pq]} + E_{6,16}\, V^4_{5,\ \mathrm{cs}[ij\ \mathrm{cs}[kl\ \mathrm{cs}[op\ \mathrm{cs}[pq]]]]}\Big] \\
&+ O\left(\frac{1}{n^3}\right).
\end{aligned}
$$

Appendix D

Appendix for Chapter 6

D.1 Expansion of central moments for the scalar problem

We present expressions for powers and products of the statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$, and the corresponding expectations. The expectation of $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$ are as follows:

- $\bar{E}[S_{1,n}, A] = \frac{1}{n} \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau, A]$
- $\bar{E}[S_{2,n}, A] = \frac{1}{n} \bar{E}\Big[\sum_{s=t-n}^{t-1} X_s^2, A\Big] = \bar{E}[X_{t-1}^2, A]$
- $\bar{E}[S_{3,n}, A] = \frac{1}{n} \sum_{\tau=t-n}^{t-1} \bar{E}[X_t Y_{\tau+1} X_\tau, A]$.

The expectation of $S_{1,n}^2$ is as follows:

- $$\begin{aligned} \bar{E}[S_{1,n}^2, A] =& \frac{1}{n^2} \bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1} X_t Y_{\tau+1} X_\tau \Big)^2, A\Big] \\ =& \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{\tau+1}^2 X_\tau^2, A] + \sum_{i\neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A]\Big]. \end{aligned}$$

The expectation of $S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{2,n}^2, A] =& \frac{1}{n^2}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^2, A\Big] \\ =& \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^4, A] + \sum_{i\neq j, t-n}^{t-1} \bar{E}\left[X_i^2 X_j^2, A\right]\Big].\end{aligned}$$

The truncated expectation of $S_{3,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{3,n}^2, A] =& \frac{1}{n^2}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} X_t Y_{\tau+1} X_\tau\Big)^2, A\Big] \\ =& \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[X_t^2 Y_t^2 X_{t-1}^2, A\right] + \sum_{i\neq j, t-n}^{t-1} \bar{E}\left[X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A\right]\Big].\end{aligned}$$

The truncated expectation of $S_{1,n}S_{2,n}$ is as follows:

- $$\begin{aligned}\bar{E}[S_{1,n}S_{2,n}, A] =& \frac{1}{n^2}\bar{E}\left[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1} X_t Y_{\tau+1} X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big), A\right] \\ =& \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{\tau+1} X_\tau^3, A\right] + \sum_{i\neq j, t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{i+1} X_i X_j^2, A\right]\Big].\end{aligned}$$

The truncated expectation of $S_{2,n}S_{3,n}$ is as follows:

- $$\begin{aligned}\bar{E}[S_{2,n}S_{3,n}, A] =& \frac{1}{n^2}\bar{E}\left[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)\Big(\sum_{\tau=t-n}^{t-1} X_t Y_{\tau+1} X_\tau\Big), A\right] \\ =& \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[X_t Y_{\tau+1} X_\tau^3, A\right] + \sum_{i\neq j, t-n}^{t-1} \bar{E}\left[X_t Y_{i+1} X_i X_j^2, A\right]\Big].\end{aligned}$$

The truncated expectation of $S_{2,n}^3$ is as follows:

- $$\begin{aligned}\bar{E}[S_{2,n}^3, A] =& \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^3, A\Big] \\ =& \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^6, A] + \sum_{i\neq j, t-n}^{t-1} \bar{E}[X_i^4 X_j^2, A] + \sum_{i\neq j\neq k, t-n}^{t-1} \bar{E}[X_i^2 X_j^2 X_k^2, A]\Big].\end{aligned}$$

The truncated expectation of $S_{1,n}S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{1,n}S_{2,n}^2, A] =& \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^2, A\Big] \\ =& \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1} X_\tau^4 + \sum_{i\neq j,t-n}^{t-1} X_i^2X_j^2\Big), A\Big] \\ =& \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^5, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4, A] \\ &+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2, A] \\ &+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2, A]\Big].\end{aligned}$$

The truncated expectation of $S_{2,n}^2S_{3,n}$ is as follows:

- $$\begin{aligned}\bar{E}[S_{2,n}^2S_{3,n}, A] =& \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big)^2\Big(\sum_{\tau=t-n}^{t-1} X_tY_{\tau+1}X_\tau\Big), A\Big] \\ =& \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} X_\tau^4 + \sum_{i\neq j,t-n}^{t-1} X_i^2X_j^2\Big)\Big(\sum_{\tau=t-n}^{t-1} X_tY_{\tau+1}X_\tau\Big), A\Big] \\ =& \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_tY_{\tau+1}X_\tau^5, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_iX_j^4, A] \\ &+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_i^3X_j^2, A] \\ &+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_iX_j^2X_k^2, A]\Big].\end{aligned}$$

The truncated expectation of $S_{1,n}^2S_{2,n}$ is as follows:

- $$\begin{aligned}&\bar{E}[S_{1,n}^2S_{2,n}, A] = \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1}X_tY_{\tau+1}X_\tau\Big)^2\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big), A\Big] \\ &= \frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1} Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^2 + \sum_{i\neq j,t-n}^{t-1} Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j\Big)\Big(\sum_{\tau=t-n}^{t-1} X_\tau^2\Big), A\Big] \\ &= \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2, A]\end{aligned}$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j^3,A]$$
$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2,A]\Big].$$

The truncated expectation of $S_{3,n}^2S_{2,n}$ is as follows:

- $$\bar{E}[S_{3,n}^2S_{2,n},A]=\frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_tY_{\tau+1}X_\tau\Big)^2\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big),A\Big]$$
$$=\frac{1}{n^3}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_t^2Y_{\tau+1}^2X_\tau^2+\sum_{i\neq j,t-n}^{t-1}X_t^2Y_{i+1}X_iY_{j+1}X_j\Big)\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big),A\Big]$$
$$=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j^3,A]$$
$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2,A]\Big].$$

The truncated expectation of $S_{2,n}^4$ is as follows:

- $$\bar{E}[S_{2,n}^4,A]=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big)^4,A\Big]$$
$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^8,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^6X_j^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^4X_j^4,A]$$
$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_i^4X_j^2X_k^2,A]+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_i^2X_j^2X_k^2X_l^2,A]\Big].$$

The truncated expectation of $S_{1,n}S_{2,n}^3$ is as follows:

- $$\bar{E}[S_{1,n}S_{2,n}^3,A]=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}Y_{t+1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big)^3,A\Big]$$
$$=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}Y_{t+1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1}X_\tau^6+\sum_{i\neq j,t-n}^{t-1}X_i^4X_j^2$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1} X_i^2X_j^2X_k^2\Big),A\Big]$$

$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^7,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^6,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^5X_j^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^4,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4X_k^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2X_k^2,A]$$

$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2X_l^2,A]\Big].$$

The truncated expectation of $S_{3,n}S_{2,n}^3$ is as follows:

- $\bar{E}[S_{3,n}S_{2,n}^3,A]=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big)^3,A\Big]$

$$=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_tY_{\tau+1}X_\tau\Big)\Big(\sum_{\tau=t-n}^{t-1}X_\tau^6+\sum_{i\neq j,t-n}^{t-1}X_i^4X_j^2+\sum_{i\neq j\neq k,t-n}^{t-1}X_i^2X_j^2X_k^2\Big),A\Big]$$

$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^7,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^6,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^5X_j^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^4,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4X_k^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2X_k^2,A]$$

$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2X_l^2,A]\Big].$$

The truncated expectation of $S_{1,n}^2S_{2,n}^2$ is as follows:

- $\bar{E}[S_{1,n}^2S_{2,n}^2,A]=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}Y_{t+1}X_tY_{\tau+1}X_\tau\Big)^2\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big)^2,A\Big]$

$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^6,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^4,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^4X_j^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^5Y_{j+1}X_j,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2,A]$$
$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2,A]\Big].$$

The truncated expectation of $S_{3,n}^2S_{2,n}^2$ is as follows:

- $$\bar{E}[S_{3,n}^2S_{2,n}^2,A]=\frac{1}{n^4}\bar{E}\Big[\Big(\sum_{\tau=t-n}^{t-1}X_tY_{\tau+1}X_\tau\Big)^2\Big(\sum_{\tau=t-n}^{t-1}X_\tau^2\Big)^2,A\Big]$$
$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^6,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^4,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^4X_j^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_i^5Y_{j+1}X_j,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4,A]$$
$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2,A]$$
$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2,A]\Big].$$

The expressions for powers and products of the statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$ given above are used to expand truncated central moments of first, second, and third order. The truncated expectation of $(S_{1,n}-\omega_{1,n})$ is as follows:

- $$\bar{E}[(S_{1,n}-\omega_{1,n}),A]=\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau,A]-\omega_{1,n}P(X\in A).$$

The truncated expectation of $(S_{2,n}-\omega_2)$ is as follows:

- $$\bar{E}[(S_{2,n}-\omega_2),A]=\bar{E}\left[X_{t-1}^2,A\right]-\omega_2P(X\in A).$$

The truncated expectation of $(S_{3,n}-\omega_{3,n})$ is as follows:

- $$\bar{E}[(S_{3,n}-\omega_{3,n}),A]=\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau,A]-\omega_{3,n}P(X\in A).$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})^2$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{1,n}-\omega_{1,n})^2,A]=&\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^2,A]\\&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j,A]\\&-2\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau,A]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\\&+\Big(\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\Big)^2P(X\in A)\Big]\\=&\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j,A]\\&-2\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau,A]E[Y_{t+1}X_tY_{\tau+1}X_\tau]\\&-2\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i,A]E[Y_{t+1}X_tY_{j+1}X_j]\\&+\sum_{\tau=t-n}^{t-1}E^2[Y_{t+1}X_tY_{\tau+1}X_\tau]P(X\in A)\\&+\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]P(X\in A)\Big].\end{aligned}$$

The truncated expectation of $(S_{2,n}-\omega_2)^2$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{2,n}-\omega_2)^2,A]=&\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^2X_j^2,A]\Big]\\&-2E[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]+E^2[X_{t-1}^2]P(X\in A).\end{aligned}$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})^2$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{3,n}-\omega_{3,n})^2, A] = \frac{1}{n^2}\Big[&\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2 Y_{\tau+1}^2 X_\tau^2, A] \\ &+ \sum_{i\neq j, t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A] \\ &- 2\sum_{\tau=t-n}^{t-1} \bar{E}[X_t Y_{\tau+1} X_\tau, A] \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] \\ &+ \Big(\sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau]\Big)^2 P(X \in A)\Big] \\ =\frac{1}{n^2}\Big[&\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2 Y_{\tau+1}^2 X_\tau^2, A] + \sum_{i\neq j, t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A] \\ &- 2\sum_{\tau=t-n}^{t-1} \bar{E}[X_t Y_{\tau+1} X_\tau, A] E[X_t Y_{\tau+1} X_\tau] \\ &- 2\sum_{i\neq j, t-n}^{t-1} \bar{E}[X_t Y_{i+1} X_i, A] E[X_t Y_{j+1} X_j] \\ &+ \sum_{\tau=t-n}^{t-1} E^2[X_t Y_{\tau+1} X_\tau] P(X \in A) \\ &+ \sum_{i\neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] P(X \in A)\Big].\end{aligned}$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n} - \omega_2), A] = \frac{1}{n^2}\Big[&\sum_{\tau=t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{\tau+1} X_\tau^3, A\right] \\ &+ \sum_{i\neq j, t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{i+1} X_i X_j^2, A\right]\Big] \\ &+ \frac{1}{n}\Big[- \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] \bar{E}[X_{t-1}^2, A] \\ &- E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau, A] \\ &+ \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] E[X_{t-1}^2] P(X \in A)\Big].\end{aligned}$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2)$ is as follows:

- $$\bar{E}[(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2), A] = \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[X_t Y_{\tau+1} X_\tau^3, A\right] + \sum_{i \neq j, t-n}^{t-1} \bar{E}\left[X_t Y_{i+1} X_i X_j^2, A\right] \Big] + \frac{1}{n}\Big[- \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] \bar{E}[X_{t-1}^2, A] - E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[X_t Y_{\tau+1} X_\tau, A] + \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] E[X_{t-1}^2] P(X \in A) \Big].$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2$ is as follows:

- $$\begin{aligned} \bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2, A] = \frac{1}{n^3}\Big[& \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau^5, A] \\ & + \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i X_j^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i^3 X_j^2, A] \\ & + \sum_{i \neq j \neq k, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i X_j^2 X_k^2, A] \Big] \\ & - \frac{2}{n^2} E[X_{t-1}^2] \Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{\tau+1} X_\tau^3, A\right] + \sum_{i \neq j, t-n}^{t-1} \bar{E}\left[Y_{t+1} X_t Y_{i+1} X_i X_j^2, A\right] \Big] \\ & - \frac{1}{n} \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] \frac{1}{n^2} \Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}\left[X_i^2 X_j^2, A\right] \Big] \\ & + E^2[X_{t-1}^2] \frac{1}{n} \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau, A] \\ & + 2\frac{1}{n} \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] E[X_{t-1}^2] \bar{E}[X_{t-1}^2, A] \\ & - \frac{1}{n} \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] E^2[X_{t-1}^2] P(X \in A) \\ = \frac{1}{n^3}\Big[& \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau^5, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i X_j^4, A] \\ & + \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i^3 X_j^2, A] + \sum_{i \neq j \neq k, t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{i+1} X_i X_j^2 X_k^2, A] \end{aligned}$$

$$
\begin{aligned}
&- \sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4, A] \\
&- \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^4, A] \\
&- \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}\left[X_i^2X_j^2, A\right] \\
&- \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}\left[X_j^2X_k^2, A\right]\Big] \\
-\frac{1}{n^2}\Big[&2E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}\left[Y_{t+1}X_tY_{\tau+1}X_\tau^3, A\right] \\
&+ 2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}\left[Y_{t+1}X_tY_{i+1}X_iX_j^2, A\right]\Big] \\
+\frac{1}{n}\Big[&E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau, A] \\
&+ 2E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A]\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau] \\
&- E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]P(X\in A)\Big].
\end{aligned}
$$

The truncated expectation of $(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2$ is as follows:

- $\bar{E}[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2, A] = \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^5, A]$

$$
\begin{aligned}
&+ \sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4, A] + \sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2, A] \\
&+ \sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2, A]\Big] \\
&- \frac{2}{n^2}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1}\bar{E}\left[X_tY_{\tau+1}X_\tau^3, A\right] + \sum_{i\neq j,t-n}^{t-1}\bar{E}\left[X_tY_{i+1}X_iX_j^2, A\right]\Big] \\
&- \frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1}\bar{E}\left[X_i^2X_j^2, A\right]\Big] \\
&+ E^2[X_{t-1}^2]\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau, A]
\end{aligned}
$$

$$+ 2\frac{1}{n}\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A]$$
$$- \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]E^2[X_{t-1}^2]P(X \in A)$$
$$=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_tY_{\tau+1}X_\tau^5, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_iX_j^4, A]$$
$$+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_i^3X_j^2, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_tY_{i+1}X_iX_j^2X_k^2, A]$$
$$- \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4, A] - \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_j^4, A]$$
$$- \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}\left[X_i^2X_j^2, A\right] - \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}\left[X_j^2X_k^2, A\right]\Big]$$
$$-\frac{1}{n^2}\Big[2E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1} \bar{E}\left[X_tY_{\tau+1}X_\tau^3, A\right] + 2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1} \bar{E}\left[X_tY_{i+1}X_iX_j^2, A\right]\Big]$$
$$+\frac{1}{n}\Big[E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1} \bar{E}[X_tY_{\tau+1}X_\tau, A] + 2E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A]\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]$$
$$- E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]P(X \in A)\Big].$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2)$ is as follows:

- $\bar{E}[(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2), A]$

$$=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2, A]$$
$$+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j^3, A]$$
$$+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2, A]\Big]$$
$$- \frac{2}{n^3}\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[Y_{t+1}X_tY_{\tau+1}X_\tau^3, A\right]$$
$$+ \sum_{i\neq j,t-n}^{t-1} \bar{E}\left[Y_{t+1}X_tY_{i+1}X_iX_j^2, A\right]\Big]$$

$$
\begin{aligned}
&- \frac{1}{n^2} E[X_{t-1}^2] \Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{\tau+1}^2 X_\tau^2, A] \\
&+ \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A] \Big] \\
&+ \frac{1}{n^2} \Big[\sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] \Big]^2 \bar{E}[X_{t-1}^2, A] \\
&+ \frac{2}{n^2} \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1} X_t Y_{\tau+1} X_\tau, A] \\
&- \frac{1}{n^2} \Big[\sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] \Big]^2 E[X_{t-1}^2] P(X \in A) \\
= & \frac{1}{n^3} \Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{\tau+1}^2 X_\tau^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1}^2 X_i^2 X_j^2, A] \\
&+ \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1} X_i Y_{j+1} X_j^3, A] \\
&+ \sum_{i \neq j \neq k, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1} X_i Y_{j+1} X_j X_k^2, A] \\
&- 2 \sum_{\tau=t-n}^{t-1} E[Y_{t+1} X_t Y_{\tau+1} X_\tau] \bar{E} \left[Y_{t+1} X_t Y_{\tau+1} X_\tau^3, A \right] \\
&- 2 \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1} X_t Y_{i+1} X_i] \bar{E} \left[Y_{t+1} X_t Y_{j+1} X_j^3, A \right] \\
&- 2 \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1} X_t Y_{i+1} X_i] \bar{E}[Y_{t+1} X_t Y_{i+1} X_i X_j^2, A] \\
&- 2 \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1} X_t Y_{i+1} X_i] \bar{E}[Y_{t+1} X_t Y_{j+1} X_j X_i^2, A] \\
&- 2 \sum_{i \neq j \neq k, t-n}^{t-1} E[Y_{t+1} X_t Y_{i+1} X_i] \bar{E}[Y_{t+1} X_t Y_{j+1} X_j X_k^2, A] \Big] \\
+ & \frac{1}{n^2} \Big[- E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{\tau+1}^2 X_\tau^2, A] \\
&- E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2 X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A] \\
&+ \bar{E}[X_{t-1}^2, A] \sum_{\tau=t-n}^{t-1} E^2[Y_{t+1} X_t Y_{\tau+1} X_\tau]
\end{aligned}
$$

$$+ \bar{E}[X_{t-1}^2, A] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]$$

$$+ 2E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau, A]$$

$$+ 2E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j, A]$$

$$- E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E^2[Y_{t+1}X_tY_{\tau+1}X_\tau]P(X \in A)$$

$$- E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]P(X \in A)\Big].$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2)$ is as follows:

- $\bar{E}[(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2), A]$

$$=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2Y_{\tau+1}^2X_\tau^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2, A]$$

$$+ \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j^3, A] + \sum_{i \neq j \neq k, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2, A]\Big]$$

$$- \frac{2}{n^3} \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]\Big[\sum_{\tau=t-n}^{t-1} \bar{E}\left[X_tY_{\tau+1}X_\tau^3, A\right]$$

$$+ \sum_{i \neq j, t-n}^{t-1} \bar{E}\left[X_tY_{i+1}X_iX_j^2, A\right] \Big]$$

$$- \frac{1}{n^2}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2Y_{\tau+1}^2X_\tau^2, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j, A]\Big]$$

$$+ \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]\Big]^2 \bar{E}[X_{t-1}^2, A]$$

$$+ \frac{2}{n^2} \sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[X_tY_{\tau+1}X_\tau, A]$$

$$- \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]\Big]^2 E[X_{t-1}^2]P(X \in A)$$

$$=\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2Y_{\tau+1}^2X_\tau^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2, A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2 Y_{i+1}X_iY_{j+1}X_j^3, A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2 Y_{i+1}X_iY_{j+1}X_jX_k^2, A] - 2\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\bar{E}\left[X_tY_{\tau+1}X_\tau^3, A\right]$$

$$-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}\left[X_tY_{j+1}X_j^3, A\right]$$

$$-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_tY_{i+1}X_iX_j^2, A]$$

$$-2\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_jX_i^2, A]$$

$$-2\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_jX_k^2, A]\Big]$$

$$+\frac{1}{n^2}\Big[-E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^2, A]$$

$$-E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j, A] + \bar{E}[X_{t-1}^2, A]\sum_{\tau=t-n}^{t-1}E^2[X_tY_{\tau+1}X_\tau]$$

$$+\bar{E}[X_{t-1}^2, A]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]$$

$$+2E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\bar{E}[X_tY_{\tau+1}X_\tau, A]$$

$$+2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_j, A]$$

$$-E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E^2[X_tY_{\tau+1}X_\tau]P(X\in A)$$

$$-E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]P(X\in A)\Big].$$

The truncated expectation of $(S_{2,n}-\omega_2)^3$ is as follows:

- $\bar{E}[(S_{2,n}-\omega_2)^3, A] = \frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^6, A] + \sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^4X_j^2, A]$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_i^2 X_j^2 X_k^2, A]\Big]$$

$$- 3E[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}\left[X_i^2 X_j^2, A\right]\Big]$$

$$+ 3E^2[X_{t-1}^2]\bar{E}[X_{t-1}^2, A] - E^3[X_{t-1}^2]P(X \in A).$$

The truncated expectation of $(S_{2,n} - \omega_2)^4$ is as follows:

- $$\bar{E}[(S_{2,n} - \omega_2)^4, A] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^8, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_i^6 X_j^2, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_i^4 X_j^4, A]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_i^4 X_j^2 X_k^2, A] + \sum_{i\neq j\neq k\neq l,t-n}^{t-1} \bar{E}[X_i^2 X_j^2 X_k^2 X_l^2, A]\Big]$$

$$- \frac{4}{n^3}E[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^6, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_i^4 X_j^2, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_i^2 X_j^2 X_k^2, A]\Big]\Big]$$

$$+ \frac{6}{n^2}E^2[X_{t-1}^2]\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}\left[X_i^2 X_j^2, A\right]\Big]$$

$$- 4E^3[X_{t-1}^2]\bar{E}[X_{t-1}^2, A] - E^4[X_{t-1}^2]P(X \in A).$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^3$ is as follows:

- $$\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^3, A] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^7, A]$$

$$+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^6, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_i^5X_j^2, A]$$

$$+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^4, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4X_k^2, A]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2X_k^2, A]$$

$$+ \sum_{i\neq j\neq k\neq l,t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2X_l^2, A]\Big]$$

$$- \frac{1}{n}\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_\tau^6, A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^4X_j^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_i^2X_j^2X_k^2,A]\Big]\Big]$$

$$-3E[X_{t-1}^2]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^5,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2,A]\Big]$$

$$+\frac{3}{n}\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^2X_j^2,A]\Big]$$

$$+3E^2[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^3,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2,A]\Big]$$

$$-3\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E^2[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]$$

$$-E^3[X_{t-1}^2]\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau,A]$$

$$+\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E^3[X_{t-1}^2]P(X\in A)$$

$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^7,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^6,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^5X_j^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^4,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4X_k^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2X_k^2,A]$$

$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2X_l^2,A]$$

$$-\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^6,A]-\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^6,A]$$

$$-\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_i^4X_j^2,A]-\sum_{i\neq j,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^4X_i^2,A]$$

$$-\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^4X_k^2,A]$$

$$-\sum_{i\neq j\neq k,t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2X_k^2,A]$$

$$
\begin{aligned}
&- \sum_{i\neq j\neq k\neq l, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2X_l^2, A]\Big] \\
&+\frac{1}{n^3}\Big[-3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^5, A] \\
&- 3E[X_{t-1}^2]\sum_{i\neq j, t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4, A] \\
&- 3E[X_{t-1}^2]\sum_{i\neq j, t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2, A] \\
&- 3E[X_{t-1}^2]\sum_{i\neq j\neq k, t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2, A] \\
&+ 3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4, A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j, t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^4, A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j, t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2, A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j\neq k, t-n}^{t-1}E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2, A]\Big] \\
&+\frac{1}{n^2}\Big[3E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^3, A] \\
&+ 3E^2[X_{t-1}^2]\sum_{i\neq j, t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2, A]\Big] \\
&+\frac{1}{n}\Big[-3E^2[X_{t-1}^2]\bar{E}[X_{t-1}^2, A]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau] \\
&- E^3[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau, A] \\
&+ E^3[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]P(X\in A)\Big].
\end{aligned}
$$

The truncated expectation of $(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3$ is as follows:

- $\bar{E}[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3, A] = \dfrac{1}{n^4}\Big[\displaystyle\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^7, A]$

$$
\begin{aligned}
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^6,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^5X_j^2,A]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^4,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4X_k^2,A]\\
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2X_k^2,A]+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2X_l^2,A]\Big]\\
&-\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^6,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^4X_j^2,A]\\
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_i^2X_j^2X_k^2,A]\Big]\Big]\\
&-3E[X_{t-1}^2]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^5,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4,A]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2,A]\Big]\\
&+\frac{3}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_i^2X_j^2,A]\Big]\\
&+3E^2[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^3,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2,A]\Big]\\
&-3\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E^2[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]\\
&-E^3[X_{t-1}^2]\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau,A]+\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E^3[X_{t-1}^2]P(X\in A)\\
=&\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^7,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^6,A]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^5X_j^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^4,A]\\
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4X_k^2,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2X_k^2,A]\\
&+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2X_l^2,A]-\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^6,A]\\
&-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_j^6,A]-\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_i^4X_j^2,A]
\end{aligned}
$$

$$
\begin{aligned}
&- \sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_j^4X_i^2,A] - \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_j^4X_k^2,A] \\
&- \sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2X_k^2,A] \\
&- \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2X_l^2,A]\Big] \\
+&\frac{1}{n^3}\Big[-3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^5,A] - 3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4,A] \\
&- 3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2,A] \\
&- 3E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2,A] \\
&+ 3E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4,A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_j^4,A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2,A] \\
&+ 3E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1}E[X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2,A]\Big] \\
+&\frac{1}{n^2}\Big[3E^2[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^3,A] + 3E^2[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2,A]\Big] \\
+&\frac{1}{n}\Big[-3E^2[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau] \\
&- E^3[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau,A] \\
&+ E^3[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]P(X\in A)\Big].
\end{aligned}
$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)^2$ is as follows:

- $\bar{E}[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)^2,A] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^6,A]$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^4X_j^2,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^5Y_{j+1}X_j,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2,A]$$

$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2,A]\Big]$$

$$-2\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^5,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2,A]\Big]$$

$$+\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\Big]^2\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}\left[X_i^2X_j^2,A\right]\Big]$$

$$-2E[X_{t-1}^2]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2,A]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j^3,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2,A]\Big]$$

$$+4\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]E[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}\left[Y_{t+1}X_tY_{\tau+1}X_\tau^3,A\right]$$

$$+\sum_{i\neq j,t-n}^{t-1}\bar{E}\left[Y_{t+1}X_tY_{i+1}X_iX_j^2,A\right]\Big]$$

$$-2\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[Y_{t+1}X_tY_{\tau+1}X_\tau]\Big]^2E[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]$$

$$+E^2[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j,A]\Big]$$

$$
\begin{aligned}
&- 2\frac{1}{n}\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]E^2[X_{t-1}^2]\frac{1}{n}\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau, A] \\
&+ \frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\Big]^2 E^2[X_{t-1}^2]P(X\in A) \\
=&\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^6, A] \\
&+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^4X_j^2, A] \\
&+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^5Y_{j+1}X_j, A] \\
&+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4, A] + \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3, A] \\
&+ \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2, A] \\
&+ \sum_{i\neq j\neq k\neq l,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2, A] \\
&- 2\sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^5, A] \\
&- 2\sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j^5, A] \\
&- 2\sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^4, A] \\
&- 2\sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_i^4, A] \\
&- 2\sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_k^4, A] \\
&- 2\sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{i+1}X_i^3X_j^2, A] \\
&- 2\sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j^3X_i^2, A] \\
&- 2\sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j^3X_k^2, A]
\end{aligned}
$$

$$- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2X_k^2, A]$$

$$- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_i^2X_k^2, A]$$

$$- 2 \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_k^2X_l^2, A]$$

$$+ \sum_{\tau=t-n}^{t-1} E^2[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4, A] + \sum_{i\neq j,t-n}^{t-1} E^2[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^4, A]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]\bar{E}[X_j^4, A]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]\bar{E}[X_k^4, A]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E^2[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2, A]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} E^2[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2, A]$$

$$+ \sum_{i\neq j,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]\bar{E}[X_i^2X_j^2, A]$$

$$+ \sum_{i\neq j\neq k,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]\bar{E}[X_j^2X_k^2, A]$$

$$+ \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]\bar{E}[X_k^2X_l^2, A]\Big]$$

$$+\frac{1}{n^3}\Big[- 2E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^4, A]$$

$$- 2E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}^2X_i^2X_j^2, A]$$

$$- 2E[X_{t-1}^2] \sum_{i\neq j,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j^3, A]$$

$$- 2E[X_{t-1}^2] \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2, A]$$

$$+ 4E[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau^3, A]$$

$$
\begin{aligned}
&+ 4E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j^3, A] \\
&+ 4E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{i+1}X_iX_j^2, A] \\
&+ 4E[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_i^2, A] \\
&+ 4E[X_{t-1}^2] \sum_{i \neq j \neq k, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_jX_k^2, A]\Big] \\
+&\frac{1}{n^2}\Big[- 2E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A] \sum_{\tau=t-n}^{t-1} E^2[Y_{t+1}X_tY_{\tau+1}X_\tau] \\
&- 2E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j] \\
&+ E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{\tau+1}^2X_\tau^2, A] \\
&+ E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} \bar{E}[Y_{t+1}^2X_t^2Y_{i+1}X_iY_{j+1}X_j, A] \\
&- 2E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[Y_{t+1}X_tY_{\tau+1}X_\tau]\bar{E}[Y_{t+1}X_tY_{\tau+1}X_\tau, A] \\
&- 2E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]\bar{E}[Y_{t+1}X_tY_{j+1}X_j, A] \\
&+ E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E^2[Y_{t+1}X_tY_{\tau+1}X_\tau]P(X \in A) \\
&+ E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[Y_{t+1}X_tY_{i+1}X_i]E[Y_{t+1}X_tY_{j+1}X_j]P(X \in A)\Big].
\end{aligned}
$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2)^2$ is as follows:

- $\bar{E}[(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2)^2, A] = \frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2Y_{\tau+1}^2X_\tau^6, A]$

$$
\begin{aligned}
&+ \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^4, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}^2X_i^4X_j^2, A] \\
&+ \sum_{i \neq j \neq k, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2X_k^2, A] + \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2Y_{i+1}X_i^5Y_{j+1}X_j, A]
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_i^3Y_{j+1}X_j^3,A]\\
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_i^3Y_{j+1}X_jX_k^2,A]\\
&+\sum_{i\neq j\neq k\neq l,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2X_l^2,A]\Big]\\
&-2\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau^5,A]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_i^3X_j^2,A]\\
&+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_tY_{i+1}X_iX_j^2X_k^2,A]\Big]\\
&+\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\Big]^2\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}\left[X_i^2X_j^2,A\right]\Big]\\
&-2E[X_{t-1}^2]\frac{1}{n^3}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^4,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2,A]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j^3,A]+\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2,A]\Big]\\
&+4\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}\left[X_tY_{\tau+1}X_\tau^3,A\right]\\
&+\sum_{i\neq j,t-n}^{t-1}\bar{E}\left[X_tY_{i+1}X_iX_j^2,A\right]\Big]\\
&-2\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\Big]^2E[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]\\
&+E^2[X_{t-1}^2]\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^2,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j,A]\Big]\\
&-2\frac{1}{n}\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]E^2[X_{t-1}^2]\frac{1}{n}\sum_{\tau=t-n}^{t-1}\bar{E}[X_tY_{\tau+1}X_\tau,A]\\
&+\frac{1}{n^2}\Big[\sum_{\tau=t-n}^{t-1}E[X_tY_{\tau+1}X_\tau]\Big]^2E^2[X_{t-1}^2]P(X\in A)\\
=&\frac{1}{n^4}\Big[\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^6,A]+\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^4,A]
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1}^2 X_i^4 X_j^2, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1}^2 X_i^2 X_j^2 X_k^2, A] \\
&+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i^5 Y_{j+1} X_j, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i Y_{j+1} X_j X_k^4, A] \\
&+ \sum_{i\neq j,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i^3 Y_{j+1} X_j^3, A] + \sum_{i\neq j\neq k,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i^3 Y_{j+1} X_j X_k^2, A] \\
&+ \sum_{i\neq j\neq k\neq l,t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i Y_{j+1} X_j X_k^2 X_l^2, A] \\
&- 2 \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] \bar{E}[X_t Y_{\tau+1} X_\tau^5, A] \\
&- 2 \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j^5, A] \\
&- 2 \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{i+1} X_i X_j^4, A] \\
&- 2 \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j X_i^4, A] \\
&- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j X_k^4, A] \\
&- 2 \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{i+1} X_i^3 X_j^2, A] \\
&- 2 \sum_{i\neq j,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j^3 X_i^2, A] \\
&- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j^3 X_k^2, A] \\
&- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{i+1} X_i X_j^2 X_k^2, A] \\
&- 2 \sum_{i\neq j\neq k,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j X_i^2 X_k^2, A] \\
&- 2 \sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j X_k^2 X_l^2, A]
\end{aligned}
$$

$$+\sum_{\tau=t-n}^{t-1} E^2[X_tY_{\tau+1}X_\tau]\bar{E}[X_\tau^4,A] + \sum_{i\neq j,t-n}^{t-1} E^2[X_tY_{i+1}X_i]\bar{E}[X_j^4,A]$$

$$+\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\bar{E}[X_j^4,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\bar{E}[X_k^4,A]$$

$$+\sum_{i\neq j,t-n}^{t-1} E^2[X_tY_{i+1}X_i]\bar{E}[X_i^2X_j^2,A] + \sum_{i\neq j\neq k,t-n}^{t-1} E^2[X_tY_{i+1}X_i]\bar{E}[X_j^2X_k^2,A]$$

$$+\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\bar{E}[X_i^2X_j^2,A]$$

$$+\sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\bar{E}[X_j^2X_k^2,A]$$

$$+\sum_{i\neq j\neq k\neq l,t-n}^{t-1} E[X_tY_{i+1}X_i]E[X_tY_{j+1}X_j]\bar{E}[X_k^2X_l^2,A]\Big]$$

$$+\frac{1}{n^3}\Big[-2E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1}\bar{E}[X_t^2Y_{\tau+1}^2X_\tau^4,A] - 2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}^2X_i^2X_j^2,A]$$

$$-2E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_j^3,A]$$

$$-2E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1}\bar{E}[X_t^2Y_{i+1}X_iY_{j+1}X_jX_k^2,A]$$

$$+4E[X_{t-1}^2]\sum_{\tau=t-n}^{t-1} E[X_tY_{\tau+1}X_\tau]\bar{E}[X_tY_{\tau+1}X_\tau^3,A]$$

$$+4E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_j^3,A]$$

$$+4E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_tY_{i+1}X_iX_j^2,A]$$

$$+4E[X_{t-1}^2]\sum_{i\neq j,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_jX_i^2,A]$$

$$+4E[X_{t-1}^2]\sum_{i\neq j\neq k,t-n}^{t-1} E[X_tY_{i+1}X_i]\bar{E}[X_tY_{j+1}X_jX_k^2,A]\Big]$$

$$+\frac{1}{n^2}\Big[-2E[X_{t-1}^2]\bar{E}[X_{t-1}^2,A]\sum_{\tau=t-n}^{t-1} E^2[X_tY_{\tau+1}X_\tau]$$

$$
\begin{aligned}
&- 2E[X_{t-1}^2]\bar{E}[X_{t-1}^2, A] \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] \\
&+ E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} \bar{E}[X_t^2 Y_{\tau+1}^2 X_\tau^2, A] + E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} \bar{E}[X_t^2 Y_{i+1} X_i Y_{j+1} X_j, A] \\
&- 2E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E[X_t Y_{\tau+1} X_\tau] \bar{E}[X_t Y_{\tau+1} X_\tau, A] \\
&- 2E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] \bar{E}[X_t Y_{j+1} X_j, A] \\
&+ E^2[X_{t-1}^2] \sum_{\tau=t-n}^{t-1} E^2[X_t Y_{\tau+1} X_\tau] P(X \in A) \\
&+ E^2[X_{t-1}^2] \sum_{i \neq j, t-n}^{t-1} E[X_t Y_{i+1} X_i] E[X_t Y_{j+1} X_j] P(X \in A)\Big].
\end{aligned}
$$

Appendix E

Appendix for Chapter 7

E.1 Expansion of truncated central moments

We begin by expanding powers and products of the statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$, and the corresponding truncated expectations:

- $\bar{E}[S_{1,n}, A] = \frac{1}{n}\bar{E}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau, A\right] = \left(1 - \frac{n_b}{n}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]$
- $\bar{E}[S_{2,n}, A] = \frac{1}{n}\bar{E}\left[\sum_{\tau=t-n}^{t-1} X_\tau^2, A\right] = \bar{E}[X_{t-1}^2, A]$
- $\bar{E}[S_{3,n}, A] = \frac{1}{n}\bar{E}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau, A\right] = \frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A]$
- $\bar{E}[S_{4,n}, A] = \frac{1}{n}\bar{E}\left[\sum_{\tau=t-n}^{t-1} Y_{2,\tau+1}X_\tau, A\right] = \bar{E}[Y_{2,t}X_{t-1}, A].$

Next, we expand $\bar{E}[S_{1,n}S_{2,n}, A]$:

$$
\begin{aligned}
S_{1,n}S_{2,n} =& \frac{1}{n^2}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n}^{t-1} X_\tau^2 \\
=& \frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right] \\
=& \frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 + \sum_{i\neq j} Y_{1,i+1}X_iX_j^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right].
\end{aligned}
$$

The truncated expecation of the three terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3, A\Big] = (n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A],\\
&\bar{E}\Big[\sum_{i\neq j} Y_{1,i+1}X_iX_j^2, A\Big] = ((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A],\\
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A].
\end{aligned}
$$

The truncated expectation of $S_{1,n}S_{2,n}$ is as follows:

- $$
\begin{aligned}
\bar{E}\left[S_{1,n}S_{2,n}, A\right] = &\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&+\left(1-\frac{1}{n}(2n_b+1)+\frac{n_b}{n^2}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&+n_b\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A].
\end{aligned}
$$

Next we expand $\bar{E}[S_{2,n}S_{3,n}, A]$:

$$
\begin{aligned}
S_{2,n}S_{3,n} =&\frac{1}{n^2}\sum_{\tau=t-n}^{t-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\\
=&\frac{1}{n^2}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^2 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]\\
=&\frac{1}{n^2}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3 + \sum_{i\neq j} Y_{2,i+1}X_iX_j^2 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right].
\end{aligned}
$$

The truncated expecation of the three terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3, A\Big] = n_b\bar{E}[Y_{2,t}X_{t-1}^3, A],\\
&\bar{E}\Big[\sum_{i\neq j} Y_{2,i+1}X_iX_j^2, A\Big] = (n_b^2-n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A],\\
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A].
\end{aligned}
$$

The truncated expectation of $S_{2,n}S_{3,n}$ is as follows:

- $$\begin{aligned}\bar{E}\left[S_{2,n}S_{3,n},A\right] =&\frac{n_b}{n^2}\bar{E}[Y_{2,t}X_{t-1}^3,A] + \frac{n_b}{n^2}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2,A]\\ &+\frac{n_b}{n^2}(n-n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2,A].\end{aligned}$$

The truncated expectation of $S_{2,n}S_{4,n}$ is as follows:

- $$\bar{E}\left[S_{2,n}S_{4,n},A\right] =\frac{1}{n}\bar{E}[Y_{2,t}X_{t-1}^3,A] + \left(1-\frac{1}{n}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2,A].$$

Next we expand $\bar{E}[S_{1,n}S_{3,n},A]$:

$$S_{1,n}S_{3,n} = \frac{1}{n^2}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau.$$

The truncated expectation of $S_{1,n}S_{3,n}$ is as follows:

- $$\bar{E}\left[S_{1,n}S_{3,n},A\right] =n_b\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1},A]$$

Next we expand $\bar{E}[S_{1,n}^2,A]$:

$$S_{1,n}^2 = \frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]^2 = \frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^2 + \sum_{i\neq j} Y_{1,i+1}X_iY_{1,j+1}X_j\right].$$

The truncated expectation of the two terms are as follows:

$$\begin{aligned}&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^2,A\Big] = (n-n_b)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2,A],\\ &\bar{E}\Big[\sum_{i\neq j} Y_{1,i+1}X_iY_{1,j+1}X_j,A\Big]\\ &\qquad\qquad= ((n-n_b)^2-(n-n_b))\bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2},A\right].\end{aligned}$$

The truncated expectation of $S_{1,n}^2$ is as follows:

- $$\bar{E}[S_{1,n}^2,A] = \left(\frac{1}{n}-\frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2,A]$$

$$+\left(1-\frac{1}{n}(2n_b+1)+\frac{n_b}{n^2}(n_b+1)\right)\bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2},A\right].$$

Next we expand $\bar{E}[S_{2,n}^2, A]$:

$$S_{2,n}^2=\frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-1}X_\tau^2\right]^2=\frac{1}{n^2}\left[\sum_{\tau=t-n}^{t-1}X_\tau^4+\sum_{i\neq j}X_i^2X_j^2\right].$$

The truncated expectation of $S_{2,n}^2$ is as follows:

- $\bar{E}[S_{2,n}^2,A]=\frac{1}{n}\bar{E}[X_{t-1}^4,A]+\left(1-\frac{1}{n}\right)\bar{E}[X_{t-1}^2X_{t-2}^2,A].$

Next we expand $\bar{E}[S_{3,n}^2, A]$:

$$S_{3,n}^2=\frac{1}{n^2}\left[\sum_{\tau=t-n_b}^{t-1}Y_{2,\tau+1}X_\tau\right]^2=\frac{1}{n^2}\left[\sum_{\tau=t-n_b}^{t-1}Y_{2,\tau+1}^2X_\tau^2+\sum_{i\neq j}Y_{2,i+1}X_iY_{2,j+1}X_j\right].$$

The truncated expectation of the two terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1}Y_{2,\tau+1}^2X_\tau^2,A\Big]=n_b\bar{E}\left[Y_{2,t}^2X_{t-1}^2,A\right],$$
$$\bar{E}\Big[\sum_{i\neq j}Y_{2,i+1}X_iY_{2,j+1}X_j,A\Big]=\Big(n_b^2-n_b\Big)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2},A\right].$$

The truncated expectation of $S_{3,n}^2$ is as follows:

- $\bar{E}[S_{3,n}^2,A]=\frac{n_b}{n^2}\bar{E}\left[Y_{2,t}^2X_{t-1}^2,A\right]+\frac{n_b}{n^2}(n_b-1)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2},A\right].$

The truncated expectation of $S_{4,n}^2$ is as follows:

- $\bar{E}[S_{4,n}^2,A]=\frac{1}{n}\bar{E}\left[Y_{2,t}^2X_{t-1}^2,A\right]+\left(1-\frac{1}{n}\right)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2},A\right].$

Next we expand $\bar{E}[S_{1,n}S_{2,n}^2, A]$:

$$S_{1,n}S_{2,n}^2=\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-1}X_\tau^2\right]^2$$

$$
\begin{aligned}
=&\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2+\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]^2\\
=&\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^2+\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^2+2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]\\
=&\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4+\sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2+\sum_{\tau=t-n_b}^{t-1} X_\tau^4+\sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2\right.\\
&\left.+\,2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]\\
=&\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5+\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4+\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2\right.\\
&+\sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2+\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\sum_{\tau=t-n_b}^{t-1} X_\tau^4\\
&+\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2+2\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3\sum_{\tau=t-n_b}^{t-1} X_\tau^2\\
&\left.+2\sum_{i\neq j\,,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right].
\end{aligned}
$$

The truncated expectation of the eight terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5, A\Big]=(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A],\\
&\bar{E}\Big[\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4, A\Big]=((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A],\\
&\bar{E}\Big[\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2, A\Big]=2((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A],\\
&\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2, A\Big]\\
&\qquad\qquad=((n-n_b)^3-3(n-n_b)^2+2(n-n_b))\\
&\qquad\qquad\qquad\cdot\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A],\\
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\sum_{\tau=t-n_b}^{t-1} X_\tau^4, A\Big]=n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A],
\end{aligned}
$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2, A\Big]$$
$$= (n-n_b)(n_b^2-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A],$$
$$\bar{E}\Big[2\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = 2n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A],$$
$$\bar{E}\Big[2\sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big]$$
$$= 2n_b((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A].$$

The truncated expectation of $S_{1,n}S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{1,n}S_{2,n}^2, A] &= \left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A]\\
&+\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A]\\
&+2\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]\\
&+\left(1-\frac{3}{n}(n_b+1)+\frac{1}{n^2}(2+6n_b+3n_b^2)-\frac{n_b}{n^3}(2+3n_b+n_b^2)\right)\\
&\qquad\cdot\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]\\
&+n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A]\\
&+n_b(n_b-1)\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A]\\
&+2n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A]\\
&+2n_b\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A].\end{aligned}$$

Next we expand $\bar{E}[S_{1,n}^2S_{2,n}, A]$:

$$\begin{aligned}S_{1,n}^2S_{2,n} &= \frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]^2\left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]\\
&=\frac{1}{n^3}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^2+\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iY_{1,j+1}X_j\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2+\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n^3}\Bigg[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^4 + \sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^2 X_j^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \\
&+ \sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i^3 Y_{1,j+1} X_j + \sum_{i\neq j\neq k\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j X_k^2 \\
&+ \sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \Bigg].
\end{aligned}
$$

The truncated expectation of the six terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^4, A\Big] = (n-n_b)\bar{E}[Y_{1,t-n_b}^2 X_{t-n+b-1}^4, A], \\
&\bar{E}\Big[\sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^2 X_j^2, A\Big] = ((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A], \\
&\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A], \\
&\bar{E}\Big[\sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i^3 Y_{1,j+1} X_j, A\Big] \\
&\qquad\qquad = 2((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A], \\
&\bar{E}\Big[\sum_{i\neq j\neq k\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j X_k^2, A\Big] \\
&\qquad\qquad = ((n-n_b)^3-3(n-n_b)^2+2(n-n_b)) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A], \\
&\bar{E}\Big[\sum_{i\neq j\ ,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = \\
&\qquad\qquad n_b((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A].
\end{aligned}
$$

The truncated expectation of $S_{1,n}^2 S_{2,n}$ is as follows:

- $$
\begin{aligned}
\bar{E}[S_{1,n}^2 S_{2,n}, A] = &\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n+b-1}^4, A] \\
&+ \left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&+ n_b \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A] \\
&+ 2\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3} n_b(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&+ \left(1 - \frac{3}{n}(n_b+1) + \frac{1}{n^2}(2+6n_b+3n_b^2) - \frac{n_b}{n^3}(2+3n_b+n_b^2)\right) \\
&\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&+ n_b \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3} n_b(n_b+1)\right) \\
&\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A].
\end{aligned}
$$

Next we expand $\bar{E}[S_{3,n} S_{2,n}^2, A]$:

$$
\begin{aligned}
S_{3,n} S_{2,n}^2 &= \frac{1}{n^3} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau\right] \left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]^2 \\
=&\frac{1}{n^3} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau\right] \left[\sum_{\tau=t-n_b}^{t-1} X_\tau^2 + \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]^2 \\
=&\frac{1}{n^3} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau\right] \left[\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^2 + \Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^2 + 2\sum_{\tau=t-n_b}^{t-1} X_\tau^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right] \\
=&\frac{1}{n^3} \Bigg[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^5 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^4 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 X_j^2 \\
&+ \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2 X_k^2 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 \\
&+ \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 + 2\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \\
&+2\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Bigg].
\end{aligned}
$$

The truncated expectation of the eight terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^5, A\Big] = n_b \bar{E}[Y_{2,t} X_{t-1}^5, A], \\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^4, A\Big] = n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^4, A],
\end{aligned}
$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2, A\Big] = 2n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2, A\Big] = (n_b^3-3n_b^2+2n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A],$$

$$\begin{aligned}\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2, A\Big]&\\ = n_b((n-n_b)^2-(n-n_b))&\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A],\end{aligned}$$

$$\bar{E}\Big[2\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = 2n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A],$$

$$\begin{aligned}\bar{E}\Big[2\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big]&\\ = 2n_b(n_b-1)(n-n_b)&\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A].\end{aligned}$$

The truncated expectation of $S_{3,n}S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{3,n}S_{2,n}^2, A] &= \frac{n_b}{n^3}\bar{E}[Y_{2,t}X_{t-1}^5, A] + \frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A]\\ &+ 2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] + \frac{n_b}{n^3}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]\\ &+ n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]\\ &+ n_b\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]\\ &+ 2n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]\\ &+ 2n_b(n_b-1)\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A].\end{aligned}$$

The truncated expectation of $S_{4,n}S_{2,n}^2$ is as follows:

- $$\begin{aligned}\bar{E}[S_{4,n}S_{2,n}^2, A] &= \frac{1}{n^2}\bar{E}[Y_{2,t}X_{t-1}^5, A] + \left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A]\\ &+ 2\left(\frac{1}{n}-\frac{1}{n^2}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] + \left(1-\frac{3}{n}+\frac{2}{n^2}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A].\end{aligned}$$

Next we expand $\bar{E}[S_{3,n}^2 S_{2,n}, A]$:

$$
\begin{aligned}
S_{3,n}^2 S_{2,n} &= \frac{1}{n^3}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\right]^2 \left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]\\
&=\frac{1}{n^3}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i Y_{2,j+1}X_j\right]\left[\sum_{\tau=t-n_b}^{t-1} X_\tau^2 + \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]\\
&=\frac{1}{n^3}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^4 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^2 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right.\\
&\quad + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3 Y_{2,j+1}X_j + \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_i Y_{2,j+1}X_j X_k^2\\
&\quad \left. + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i Y_{2,j+1}X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right].
\end{aligned}
$$

The truncated expectation of the six terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^4, A\Big] = n_b \bar{E}[Y_{2,t}^2 X_{t-1}^4, A],\\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^2, A\Big] = n_b(n_b-1)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2, A],\\
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-n_b-1}^2, A],\\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3 Y_{2,j+1}X_j, A\Big] = 2n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3 Y_{2,t-1}X_{t-2}, A],\\
&\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_i Y_{2,j+1}X_j X_k^2, A\Big]\\
&\qquad\qquad = n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A],\\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i Y_{2,j+1}X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big]\\
&\qquad\qquad = n_b(n_b-1)(n-n_b)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A].
\end{aligned}
$$

The truncated expectation of $S_{3,n}^2 S_{2,n}$ is as follows:

- $$\begin{aligned}\bar{E}[S_{3,n}^2 S_{2,n}, A] &= \frac{n_b}{n^3}\bar{E}[Y_{2,t}^2 X_{t-1}^4, A] + \frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2, A] \\ &+ n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-n_b-1}^2, A] + 2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3 Y_{2,t-1}X_{t-2}, A] \\ &+ \frac{n_b}{n^3}(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A] \\ &+ n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A].\end{aligned}$$

The truncated expectation of $S_{4,n}^2 S_{2,n}$ is as follows:

- $$\begin{aligned}\bar{E}[S_{4,n}^2 S_{2,n}, A] &= \frac{1}{n^2}\bar{E}[Y_{2,t}^2 X_{t-1}^4, A] + \left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2, A] \\ &+ 2\left(\frac{1}{n} - \frac{1}{n^2}\right)\bar{E}[Y_{2,t}X_{t-1}^3 Y_{2,t-1}X_{t-2}, A] \\ &+ \left(1 - \frac{3}{n} + \frac{2}{n^2}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A].\end{aligned}$$

Next we expand $\bar{E}[S_{1,n}S_{2,n}S_{3,n}, A]$:

$$\begin{aligned}S_{1,n}S_{2,n}S_{3,n} &= \frac{1}{n^3}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n}^{t-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \\ =&\frac{1}{n^3}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right] \\ =&\frac{1}{n^3}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \\ &+ \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big].\end{aligned}$$

The truncated expectation of the four terms are as follows:

$$\begin{aligned}E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3\Big] &= n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A], \\ E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2\Big] &= n_b((n-n_b)^2 - (n-n_b)) \\ &\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A],\end{aligned}$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big] = (n-n_b)(n_b^2-n_b)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A].$$

The truncated expectation of $S_{1,n}S_{2,n}S_{3,n}$ is as follows:

- $$\bar{E}[S_{1,n}S_{2,n}S_{3,n}, A] = n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A]$$
$$+ n_b\left(\frac{1}{n}-(2n_b+1)\frac{1}{n^2}+n_b(n_b+1)\frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]$$
$$+ n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A]$$
$$+ n_b(n_b-1)\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A].$$

Next we expand $\bar{E}[S1, nS_{2,n}^3, A]$:

$$S_{1,n}S_{2,n}^3 = \frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]^3$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]^3$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^3 + \Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^3 + 3\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right.$$
$$\left. +3\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^2\right]$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^6 + \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^4X_j^2 + \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} X_i^2X_j^2X_k^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^6\right.$$
$$+ \sum_{i\neq j,\, t-n_b}^{t-1} X_i^4X_j^2 + \sum_{i\neq j\neq k,\, t-n_b}^{t-1} X_i^2X_j^2X_k^2 + 3\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 + \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2\Big)\sum_{\tau=t-n_b}^{t-1} X_\tau^2$$
$$\left. +3\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^4 + \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2\Big)\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]$$

$$
\begin{aligned}
&= \frac{1}{n^4}\Bigg[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^7 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^6 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^5X_j^2 \\
&+ \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^4 + \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4X_k^2 + \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2X_k^2 \\
&+ \sum_{i\neq j\neq k\neq l,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2X_l^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^6 \\
&+ \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,\, t-n_b}^{t-1} X_i^4X_j^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j\neq k,\, t-n_b}^{t-1} X_i^2X_j^2X_k^2 \\
&+ 3\Big(\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2 \\
&+ \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2\Big) \sum_{\tau=t-n_b}^{t-1} X_\tau^2 + 3\Big(\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2\Big) \\
&\cdot\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^4 + \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2\Big)\Bigg] \\
=&\frac{1}{n^4}\Bigg[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^7 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^6 + \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^5X_j^2 \\
&+ \sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^4 + \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4X_k^2 + \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2X_k^2 \\
&+ \sum_{i\neq j\neq k\neq l,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2X_l^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^6 \\
&+ \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,\, t-n_b}^{t-1} X_i^4X_j^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j\neq k,\, t-n_b}^{t-1} X_i^2X_j^2X_k^2 \\
&+ 3\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 + 3\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \\
&+ 3\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 + 3\sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \\
&+ 3\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{\tau=t-n_b}^{t-1} X_\tau^4 + 3\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2 \\
&+3\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^4 + 3\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2X_j^2\Bigg].
\end{aligned}
$$

The truncated expectation of the eighteen terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^7, A\Big] = (n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^7, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^6, A\Big] = ((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^6, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^5X_j^2, A\Big] = 3((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-n_b-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^4, A\Big] = 3((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4X_k^2, A\Big]$$

$$= 3((n-n_b)^3-3(n-n_b)^2+2(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-n_b-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2X_k^2, A\Big]$$

$$= 3((n-n_b)^3-3(n-n_b)^2+2(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-n_b-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k\neq l,\, t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2X_l^2, A\Big]$$

$$= ((n-n_b)^4-6(n-n_b)^3+11(n-n_b)^2-6(n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-n_b-4}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} X_\tau^6, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^6, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,\, t-n_b}^{t-1} X_i^4X_j^2, A\Big]$$

$$= 3n_b(n-n_b)(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j\neq k,\, t-n_b}^{t-1} X_i^2X_j^2X_k^2, A\Big]$$

$$= n_b(n-n_b)(n_b^2-3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big]$$
$$= n_b((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-1}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big]$$
$$= 2n_b((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-1}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big]$$
$$= n_b((n-n_b)^3-3(n-n_b)^2+2(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-1}^2, A],$$
$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{\tau=t-n_b}^{t-1} X_\tau^4, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^4, A],$$
$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{i\neq j,\,t-n_b}^{t-1} X_i^2X_j^2, A\Big]$$
$$= n_b(n-n_b)(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2X_{t-2}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^4, A\Big]$$
$$= n_b((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^4, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{i\neq j,\,t-n_b}^{t-1} X_i^2X_j^2, A\Big]$$
$$= n_b(n_b-1)((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2X_{t-2}^2, A].$$

The truncated expectation of $S_{1,n}S_{2,n}^3$ is as follows:

- $$\bar{E}[S_{1,n}S_{2,n}^3, A] = \left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^7, A]$$
$$+\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^6, A]$$
$$+3\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-n_b-2}^2, A]$$
$$+3\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^4, A]$$

$$+ 3\left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-n_b-3}^2, A]$$
$$+ 3\left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$
$$+ \left(1 - \frac{1}{n}(4n_b+6) + \frac{1}{n^2}(6n_b^2+18n_b+11) - \frac{1}{n^3}(4n_b^3+18n_b^2+22n_b+6)\right.$$
$$\left.+\frac{n_b}{n^4}(n_b^3+6n_b^2+11n_b+6)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-n_b-4}^2, A]$$
$$+ n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^6, A]$$
$$+ 3n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4X_{t-2}^2, A]$$
$$+ n_b(n_b^2-3n_b+2)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$
$$+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-1}^2, A]$$
$$+ 3n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-1}^2, A]$$
$$+ 6n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-1}^2, A]$$
$$+ 3n_b\left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-1}^2, A]$$
$$+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^4, A]$$
$$+ 3n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2X_{t-2}^2, A]$$
$$+ 3n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^4, A]$$
$$+ 3n_b(n_b-1)\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2X_{t-2}^2, A].$$

Next we expand $\bar{E}[S_{3,n}S_{2,n}^3, A]$:

$$S_{3,n}S_{2,n}^3 = \frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]^3$$

$$
\begin{aligned}
=&\frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2+\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]^3\\
=&\frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\right]\left[\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^3+\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^3+3\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^2\sum_{\tau=t-n_b}^{t-1} X_\tau^2\right.\\
&\left.+3\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^2\right]\\
=&\frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\right]\left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^6+\sum_{i\neq j,\,t-n}^{t-n_b-1} X_i^4X_j^2+\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} X_i^2X_j^2X_k^2+\sum_{\tau=t-n_b}^{t-1} X_\tau^6\right.\\
&+\sum_{i\neq j,\,t-n_b}^{t-1} X_i^4X_j^2+\sum_{i\neq j\neq k,\,t-n_b}^{t-1} X_i^2X_j^2X_k^2+3\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4+\sum_{i\neq j,\,t-n}^{t-n_b-1} X_i^2X_j^2\Big)\sum_{\tau=t-n_b}^{t-1} X_\tau^2\\
&\left.+3\Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^4+\sum_{i\neq j,\,t-n_b}^{t-1} X_i^2X_j^2\Big)\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]\\
=&\frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{\tau=t-n}^{t-n_b-1} X_\tau^6+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{i\neq j,\,t-n}^{t-n_b-1} X_i^4X_j^2\right.\\
&+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} X_i^2X_j^2X_k^2+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^7+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^6\\
&+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_i^5X_j^2+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^4+\sum_{i\neq j\neq k,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4X_k^2\\
&+\sum_{i\neq j\neq k,\,t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2X_k^2+\sum_{i\neq j\neq k\neq l,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2X_l^2\\
&+3\Big(\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big)\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4+\sum_{i\neq j,\,t-n}^{t-n_b-1} X_i^2X_j^2\Big)\\
&+3\Big(\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^5+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2\\
&\left.+\sum_{i\neq j\neq k,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2\Big)\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\right]\\
=&\frac{1}{n^4}\left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{\tau=t-n}^{t-n_b-1} X_\tau^6+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{i\neq j,\,t-n}^{t-n_b-1} X_i^4X_j^2\right.\\
&+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} X_i^2X_j^2X_k^2+\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^7+\sum_{i\neq j,\,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^6
\end{aligned}
$$

$$+ \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^5 X_j^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 X_j^4 + \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^4 X_k^2$$

$$+ \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 X_j^2 X_k^2 + \sum_{i\neq j\neq k\neq l,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2 X_k^2 X_l^2$$

$$+ 3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^3 + 3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2$$

$$+ 3 \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^3 + 3 \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2$$

$$+ 3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^5 + 3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^4$$

$$+3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 X_j^2 + 3 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^2 X_k^2 \Bigg].$$

The truncated expectation of the eighteen terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \sum_{\tau=t-n}^{t-n_b-1} X_\tau^6, A \Big] = n_b(n-n_b) \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^6, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^4 X_j^2, A \Big] = 3n_b((n-n_b)^2 - (n-n_b))$$

$$\cdot \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^4 X_{t-n_b-2}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \sum_{i\neq j\neq k,\, t-n}^{t-n_b-1} X_i^2 X_j^2 X_k^2, A \Big] = n_b((n-n_b)^3 - 3(n-n_b)^2 + 2(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau^7, A \Big] = n_b \bar{E}[Y_{2,t} X_{t-1}^7, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i X_j^6, A \Big] = n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^6, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^5 X_j^2, A \Big] = 3n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1}^5 X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 X_j^4, A \Big] = 3n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1}^3 X_{t-2}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4X_k^2, A\Big] = 3n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2X_k^2, A\Big] = 3n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k\neq l,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2X_l^2, A\Big] = (n_b^4-6n_b^3+11n_b^2-6n_b)$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4, A\Big] = n_b(n-n_b)(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^4, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-n_b-1} Y_{2,\tau+1}X_\tau^3\sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2, A\Big] = n_b((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2X_{t-n_b-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2, A\Big] = n_b(n_b-1)((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^5\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n-n_b)(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = 2n_b(n-n_b)(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n_b^2-3n_b+2)(n-n_b)$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-n_b-1}^2, A].$$

The truncated expectation of $S_{3,n}S_{2,n}^3$ is as follows:

- $\bar{E}[S_{3,n}S_{2,n}^3, A] = n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^6, A]$

$$+ 3n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4X_{t-n_b-2}^2, A]$$

$$+ n_b\left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$

$$+ \frac{n_b}{n^4}\bar{E}[Y_{2,t}X_{t-1}^7, A] + \frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A]$$

$$+ 3\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] + 3\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A]$$

$$+ 3\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A]$$

$$+ 3\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A]$$

$$+ \frac{n_b}{n^4}(n_b^3-6n_b^2+11n_b-6)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]$$

$$+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^4, A]$$

$$+ 3n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^4, A]$$

$$+ 3n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$

$$+ 3n_b(n_b-1)\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right)$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$

$$+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-n_b-1}^2, A]$$

$$+ 3n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-n_b-1}^2, A]$$

$$+ 6n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-n_b-1}^2, A]$$

$$+ 3n_b(n_b^2-3n_b+2)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-n_b-1}^2, A].$$

The truncated expectation of $S_{4,n}S_{2,n}^3$ is as follows:

- $\bar{E}[S_{4,n}S_{2,n}^3, A] = \frac{1}{n^3}\bar{E}[Y_{2,t}X_{t-1}^7, A] + \left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A]$

$$+ 3\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] + 3\left(\frac{1}{n^2} - \frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A]$$

$$+ 3\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A]$$

$$+3\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2,A]$$
$$+\left(1-\frac{6}{n}+\frac{11}{n^2}-\frac{6}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2,A].$$

Next we expand $\bar{E}[S_{1,n}^2S_{2,n}^2,A]$:

$$S_{1,n}^2S_{2,n}^2=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}X_\tau\right]^2\left[\sum_{\tau=t-n}^{t-1}X_\tau^2\right]^2$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^2+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j\right]\left[\sum_{\tau=t-n}^{t-n_b-1}X_\tau^2+\sum_{\tau=t-n_b}^{t-1}X_\tau^2\right]^2$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^2+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j\right]\left[\Big(\sum_{\tau=t-n}^{t-n_b-1}X_\tau^2\Big)^2+\Big(\sum_{\tau=t-n_b}^{t-1}X_\tau^2\Big)^2\right.$$
$$\left.+2\sum_{\tau=t-n}^{t-n_b-1}X_\tau^2\sum_{\tau=t-n_b}^{t-1}X_\tau^2\right]$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^2+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j\right]\left[\sum_{\tau=t-n}^{t-n_b-1}X_\tau^4+\sum_{i\neq j,\,t-n}^{t-n_b-1}X_i^2X_j^2\right.$$
$$\left.+\sum_{\tau=t-n_b}^{t-1}X_\tau^4+\sum_{i\neq j,\,t-n_b}^{t-1}X_i^2X_j^2+2\sum_{\tau=t-n}^{t-n_b-1}X_\tau^2\sum_{\tau=t-n_b}^{t-1}X_\tau^2\right]$$
$$=\frac{1}{n^4}\left[\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^6+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}^2X_i^2X_j^4+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}^2X_i^4X_j^2\right.$$
$$+\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1}Y_{1,i+1}^2X_i^2X_j^2X_k^2+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_i^5Y_{1,j+1}X_j$$
$$+\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_jX_k^4+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_i^3Y_{1,j+1}X_j^3$$
$$+\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j^3X_k^2+\sum_{i\neq j\neq k\neq l,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_jX_k^2X_l^2$$
$$+\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^2\sum_{\tau=t-n_b}^{t-1}X_\tau^4+\sum_{\tau=t-n}^{t-n_b-1}Y_{1,\tau+1}^2X_\tau^2\sum_{i\neq j,\,t-n_b}^{t-1}X_i^2X_j^2$$
$$+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j\sum_{\tau=t-n_b}^{t-1}X_\tau^4+\sum_{i\neq j,\,t-n}^{t-n_b-1}Y_{1,i+1}X_iY_{1,j+1}X_j\sum_{i\neq j,\,t-n_b}^{t-1}X_i^2X_j^2$$

$$+\,2\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^4 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 + 2\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^2 X_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2$$
$$+\,2\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i^3 Y_{1,j+1} X_j \sum_{\tau=t-n_b}^{t-1} X_\tau^2$$
$$\left. +2\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j X_k^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \right].$$

The truncated expectation of the seventeen terms are as follows:

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2 X_\tau^6, A\Big] = (n-n_b)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^6, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^2 X_j^2, A\Big] = ((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^4, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^4 X_j^2, A\Big] = 2((n-n_b)^2-(n-n_b))\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1}^2 X_i^2 X_j^2 X_k^2, A\Big] = ((n-n_b)^3-3(n-n_b)^2+2(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i^5 Y_{1,j+1} X_j, A\Big] = 2((n-n_b)^2-(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5 Y_{1,t-n_b-1} X_{t-n_b-2}, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j X_k^4, A\Big] = ((n-n_b)^3-3(n-n_b)^2+2(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^4, A],$$
$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i^3 Y_{1,j+1} X_j^3, A\Big] = 2((n-n_b)^2-(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}^3, A],$$
$$\bar{E}\Big[\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1} X_i Y_{1,j+1} X_j^3 X_k^2, A\Big] = 4((n-n_b)^3-3(n-n_b)^2+2(n-n_b))$$
$$\cdot\,\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}^3 X_{t-n_b-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k\neq l,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iY_{1,j+1}X_jX_k^2X_l^2, A\Big]$$

$$= ((n-n_b)^4 - 6(n-n_b)^3 + 11(n-n_b)^2 - 6(n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-n_b-4}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^4, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^4, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^2 \sum_{i\neq j,\,t-n_b}^{t-1} X_i^2X_j^2, A\Big] = n_b(n_b-1)(n-n_b)$$

$$\cdot \bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^2X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iY_{1,j+1}X_j \sum_{\tau=t-n_b}^{t-1} X_\tau^4, A\Big] = n_b((n-n_b)^2 - (n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iY_{1,j+1}X_j \sum_{i\neq j,\,t-n_b}^{t-1} X_i^2X_j^2, A\Big] = n_b(n_b-1)((n-n_b)^2 - (n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2X_{t-2}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}^2X_\tau^4 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^4X_{t-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}^2X_i^2X_j^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = n_b((n-n_b)^2 - (n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\,t-n}^{t-n_b-1} Y_{1,i+1}X_i^3Y_{1,j+1}X_j \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big] = 2n_b((n-n_b)^2 - (n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\,t-n}^{t-n_b-1} Y_{1,i+1}X_iY_{1,j+1}X_jX_k^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2, A\Big]$$

$$= n_b((n-n_b)^3 - 3(n-n_b)^2 + 2(n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-1}^2, A].$$

The truncated expectation of $S_{1,n}^2 S_{2,n}^2$ is as follows:

- $$\begin{aligned}
\bar{E}[S_{1,n}^2 S_{2,n}^2, A] &= \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^6, A] \\
&+ \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^4, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2, A] \\
&+ \left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&+ \left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^4, A] \\
&+ 2\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}^3, A] \\
&+ 4\left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}^3 X_{t-n_b-3}^2, A] \\
&+ \Bigg(1 - \frac{1}{n}(4n_b+6) + \frac{1}{n^2}(6n_b^2+18n_b+11) - \frac{1}{n^3}(4n_b^3+18n_b^2+22n_b+6) \\
&+ \frac{n_b}{n^4}(n_b^3+6n_b^2+11n_b+6)\Bigg) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-n_b-4}^2, A] \\
&+ n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^4, A] \\
&+ n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2 X_{t-2}^2, A] \\
&+ n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^4, A] \\
&+ n_b(n_b-1)\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2 X_{t-2}^2, A] \\
&+ 2n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-1}^2, A]
\end{aligned}$$

$$
\begin{aligned}
&+ 2n_b \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1) \right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-1}^2, A] \\
&+ 4n_b \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1) \right) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&+ 2n_b \left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2) \right) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-1}^2, A].
\end{aligned}
$$

Next we expand $\bar{E}[S_{3,n}^2 S_{2,n}^2, A]$:

$$
\begin{aligned}
S_{3,n}^2 S_{2,n}^2 = &\frac{1}{n^4} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1} X_\tau \right]^2 \left[\sum_{\tau=t-n}^{t-1} X_\tau^2 \right]^2 \\
= &\frac{1}{n^4} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j \right] \left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \right]^2 \\
= &\frac{1}{n^4} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j \right] \left[\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \Big)^2 + \Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2 \Big)^2 \right. \\
&\left. + 2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \right] \\
= &\frac{1}{n^4} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j \right] \left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 + \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 \right. \\
&\left. + \sum_{\tau=t-n_b}^{t-1} X_\tau^4 + \sum_{i\neq j,\, t-n_b}^{t-1} X_i^2 X_j^2 + 2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2 \right] \\
= &\frac{1}{n^4} \left[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^2 \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 \right. \\
&+ \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2 X_j^2 \\
&+ \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^6 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^4 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^4 X_j^2 \\
&+ \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^2 X_k^2 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^5 Y_{2,j+1} X_j
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_jX_k^4 + \sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3Y_{2,j+1}X_j^3 \\
&+ \sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_j^3X_k^2 + \sum_{i\neq j\neq k\neq l,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_jX_k^2X_l^2 \\
&+ 2\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2X_\tau^4 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + 2\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2X_i^2X_j^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \\
&+ 2\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_i^3Y_{2,j+1}X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \\
&+2\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_jX_k^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \Bigg].
\end{aligned}
$$

The truncated expectation of the seventeen terms are as follows:

$$
\begin{aligned}
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2X_\tau^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^4, A], \\
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2X_\tau^2 \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2, A\Big] = n_b((n-n_b)^2-(n-n_b)) \\
&\qquad\qquad\qquad\qquad \cdot \bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A], \\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^4, A\Big] = n_b(n_b-1)(n-n_b) \\
&\qquad\qquad\qquad\qquad \cdot \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^4, A], \\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}X_iY_{2,j+1}X_j \sum_{i\neq j,\, t-n}^{t-n_b-1} X_i^2X_j^2, A\Big] = n_b(n_b-1)((n-n_b)^2-(n-n_b)) \\
&\qquad\qquad\qquad\qquad \cdot \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2, A], \\
&\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2X_\tau^6, A\Big] = n_b\bar{E}[Y_{2,t}^2X_{t-1}^6, A], \\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2X_i^2X_j^2, A\Big] = n_b(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^4, A], \\
&\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2X_i^4X_j^2, A\Big] = 2n_b(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-2}^2, A],
\end{aligned}
$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^2 X_k^2, A\Big] = (n_b^3 - 3n_b^2 + 2n_b)\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^5 Y_{2,j+1} X_j, A\Big] = 2n_b(n_b-1)\bar{E}[Y_{2,t} X_{t-1}^5 Y_{2,t-1} X_{t-2}, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j X_k^4, A\Big] = (n_b^3 - 3n_b^2 + 2n_b) \cdot \bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^4, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 Y_{2,j+1} X_j^3, A\Big] = 2n_b(n_b-1)\bar{E}[Y_{2,t} X_{t-1}^3 Y_{2,t-1} X_{t-2}^3, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j^3 X_k^2, A\Big] = 4(n_b^3 - 3n_b^2 + 2n_b) \cdot \bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^3 X_{t-3}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k\neq l,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j X_k^2 X_l^2, A\Big] = (n_b^4 - 6 - n_b^3 + 11n_b^2 - 6n_b) \cdot \bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A],$$

$$\bar{E}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}^2 X_\tau^4 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}^2 X_{t-1}^4 X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1}^2 X_i^2 X_j^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = n_b(n_b-1)(n-n_b) \cdot \bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2 X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j,\, t-n_b}^{t-1} Y_{2,i+1} X_i^3 Y_{2,j+1} X_j \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = 2n_b(n_b-1)(n-n_b) \cdot \bar{E}[Y_{2,t} X_{t-1}^3 Y_{2,t-1} X_{t-2} X_{t-n_b-1}^2, A],$$

$$\bar{E}\Big[\sum_{i\neq j\neq k,\, t-n_b}^{t-1} Y_{2,i+1} X_i Y_{2,j+1} X_j X_k^2 \sum_{\tau=t-n}^{t-n_b-1} X_\tau^2, A\Big] = (n-n_b)(n_b^3 - 3n_b^2 + 2n_b) \cdot \bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^2 X_{t-n_b-1}^2, A].$$

The truncated expectation of $S_{3,n}^2 S_{2,n}^2$ is as follows:

- $\bar{E}[S_{3,n}^2 S_{2,n}^2, A] = n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^4, A]$

$$+ n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^4, A]$$
$$+ n_b(n_b-1)\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)$$
$$\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ \frac{n_b}{n^4}\bar{E}[Y_{2,t}^2X_{t-1}^6, A]+\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^4, A]$$
$$+ 2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-2}^2, A]+\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$
$$+ 2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^5Y_{2,t-1}X_{t-2}, A]$$
$$+ \frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^4, A]$$
$$+ 2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}^3, A]$$
$$+ 4\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-3}^2, A]$$
$$+ \frac{n_b}{n^4}(n_b^3-6n_b^2+11n_b-6)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2, A]$$
$$+ 2n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-n_b-1}^2, A]$$
$$+ 2n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-n_b-1}^2, A]$$
$$+ 4n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A]$$
$$+ 2n_b(n_b^2-3n_b+2)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-n_b-1}^2, A].$$

The truncated expectation of $S_{4,n}^2S_{2,n}^2$ is as follows:

- $$\bar{E}[S_{4,n}^2S_{2,n}^2, A]=\frac{1}{n^3}\bar{E}[Y_{2,t}^2X_{t-1}^6, A]+\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^4, A]$$
$$+ 2\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-2}^2, A]+\left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$
$$+ 2\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^5Y_{2,t-1}X_{t-2}, A]$$
$$+ \left(\frac{1}{n}-\frac{3}{n^2}+\frac{2}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^4, A]$$
$$+ 2\left(\frac{1}{n^2}-\frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}^3, A]$$

$$+ 4\left(\frac{1}{n} - \frac{3}{n^2} + \frac{2}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-3}^2, A]$$
$$+ \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2, A].$$

Next we expand $\bar{E}[S_{1,n}S_{2,n}^2S_{3,n}, A]$:

$$S_{1,n}S_{2,n}^2S_{3,n} = \frac{1}{n^4}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \left[\sum_{\tau=t-n}^{t-1} X_\tau^2\right]^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau$$
$$=\frac{1}{n^4}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \left[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^2\right]^2$$
$$=\frac{1}{n^4}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \Bigg[\Big(\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2\Big)^2 + \Big(\sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big)^2$$
$$+2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2\Bigg]$$
$$=\frac{1}{n^4}\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \Big[\sum_{\tau=t-n}^{t-n_b-1} X_\tau^4 + \sum_{i\neq j,t-n}^{t-n_b-1} X_i^2X_j^2 + \sum_{\tau=t-n_b}^{t-1} X_\tau^4$$
$$+ \sum_{i\neq j,t-n_b}^{t-1} X_i^2X_j^2 + 2\sum_{\tau=t-n}^{t-n_b-1} X_\tau^2 \sum_{\tau=t-n_b}^{t-1} X_\tau^2\Big]$$
$$=\frac{1}{n^4}\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4$$
$$+ \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2 + \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2$$
$$+ \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^5 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4$$
$$+ \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2 + \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j\neq k,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2$$
$$+ 2\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3 + 2\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2$$
$$+ 2\sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3 + 2\sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big]$$

The truncated expectation of the twelve terms are as follows:

$$E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^5\Big] = n_b(n-n_b)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X^5_{t-n_b-1}, A],$$

$$E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^4\Big] = n_b((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}, A],$$

$$E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_i^3X_j^2\Big] = 2n_b((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}, A],$$

$$E\Big[\sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau \sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2X_k^2\Big] = n_b((n-n_b)^3-3(n-n_b)^2+2(n-n_b))$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^5\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X^5_{t-1}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^4\Big] = n_b(n_b-1)(n-n_b)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X^4_{t-2}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_i^3X_j^2\Big] = 2n_b(n_b-1)(n-n_b)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X^3_{t-1}X^2_{t-2}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau \sum_{i\neq j\neq k,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2X_k^2\Big] = n_b(n_b^2-3n_b+2)(n-n_b)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3\Big] = n_b(n-n_b)\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{2,t}X^3_{t-1}, A],$$

$$E\Big[\sum_{\tau=t-n}^{t-n_b-1} Y_{1,\tau+1}X_\tau^3 \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big] = n_b(n_b-1)(n-n_b)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{2,t}X_{t-1}X^2_{t-2}, A],$$

$$E\Big[\sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{\tau=t-n_b}^{t-1} Y_{2,\tau+1}X_\tau^3\Big] = n_b((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}^3, A],$$

$$E\Big[\sum_{i\neq j,t-n}^{t-n_b-1} Y_{1,i+1}X_iX_j^2 \sum_{i\neq j,t-n_b}^{t-1} Y_{2,i+1}X_iX_j^2\Big] = n_b(n_b-1)((n-n_b)^2-(n-n_b))$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}X_{t-2}^2, A].$$

The truncated expectation of $S_{1,n}S_{2,n}^2S_{3,n}$ is as follows:

- $\bar{E}[S_{1,n}S_{2,n}^2S_{3,n}, A] = n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^5, A]$

$$+ n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A]$$

$$+ 2n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]$$

$$+ n_b\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$

$$\cdot \bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$

$$+ n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^5, A]$$

$$+ n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^4, A]$$

$$+ 2n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3X_{t-2}^2, A]$$

$$+ n_b(n_b^2-3n_b+2)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]$$

$$+ 2n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}^3, A]$$

$$+ 2n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}X_{t-2}^2, A]$$

$$+ 2n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}^3, A]$$

$$+ 2n_b(n_b-1)\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}X_{t-2}^2, A].$$

The expanded expressions for powers and products of the statistics $S_{1,n}$, $S_{2,n}$, and $S_{3,n}$ given above are used to expand truncated central moments of first and second order. We expand $\bar{E}[(S_{1,n} - \omega_{1,n}), A]$:

$$\begin{aligned}\bar{E}[(S_{1,n}-\omega_{1,n}), A] &= \bar{E}[S_{1,n}, A] - \bar{E}[\omega_{1,n}, A] \\ &= \left(1 - \frac{n_b}{n}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] - \omega_{1,n} P(X \in A).\end{aligned}$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{1,n} - \omega_{1,n}), A] = &\Big[\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] - E[Y_{1,t-n_b} X_{t-n_b-1}] P(X_1 \in A)\Big] \\ &+ \frac{n_b}{n}\Big[E[Y_{1,t-n_b} X_{t-n_b-1}] P(X_1 \in A) - \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A]\Big].\end{aligned}$$

Next, we expand $\bar{E}[(S_{2,n} - \omega_2), A]$:

- $$\bar{E}[(S_{2,n} - \omega_2), A] = \bar{E}[X_{t-1}^2, A] - \omega_2 P(X_1 \in A).$$

Next, we expand $\bar{E}[(S_{3,n} - \omega_{3,n}), A]$:

$$\bar{E}[(S_{3,n} - \omega_{3,n}), A] = \bar{E}\,[S_{3,n}, A] - \bar{E}[\omega_{3,n}, A] = \frac{n_b}{n} \bar{E}[Y_{2,t} X_{t-1}, A] - \omega_{3,n} P(X_1 \in A).$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})$ is as follows:

- $$\bar{E}[(S_{3,n} - \omega_{3,n}), A] = \frac{n_b}{n}\Big[\bar{E}[Y_{2,t} X_{t-1}, A] - E[Y_{2,t} X_{t-1}] P(X_1 \in A)\Big].$$

The truncated expectation of $(S_{4,n} - \omega_4)$ is as follows:

- $$\bar{E}[(S_{4,n} - \omega_4), A] = \bar{E}[Y_{2,t} X_{t-1}, A] - E[Y_{2,t} X_{t-1}] P(X_1 \in A).$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2), A]$:

$$\begin{aligned}\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n} - \omega_2), A] = &\bar{E}[S_{1,n} S_{2,n}, A] - \omega_{1,n} \bar{E}[S_{2,n}, A] - \omega_2 \bar{E}[S_{1,n}, A] \\ &+ \omega_{1,n} \omega_2 P(X_1 \in A)\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1}, A] \\
&\quad + \left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2}, A] \\
&\quad + n_b \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-1}, A] - \omega_{1,n} \bar{E}[X^2_{t-1}, A] \\
&\quad - \omega_2 \left(1 - \frac{n_b}{n}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] + \omega_{1,n}\omega_2 P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)$ is as follows:

- $$
\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n} - \omega_2), A] &= \Big[\bar{E}[Y_{1,t-n_b} X_{t-n_b} X^2_{t-n_b-2}, A] \\
&\quad - E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X^2_{t-1}, A] - \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&\quad + \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] P(X \in A)\Big] \\
&+ \frac{1}{n}\Big[\bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1}, A] - (2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2}, A] \\
&\quad + n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-1}, A] \\
&\quad + n_b E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X^2_{t-1}, A] + \omega_2 n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&\quad - n_b \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] P(X \in A)\Big] \\
&+ \frac{n_b}{n^2}\Big[(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2}, A] - n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-1}, A] \\
&\quad - \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1}, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{2,n} - \omega_2)(S_{3,n} - \omega_{3,n}), A]$:

$$
\begin{aligned}
\bar{E}[(S_{2,n}-\omega_2)(S_{3,n} - \omega_{3,n}), A] &= \bar{E}[S_{2,n} S_{3,n}, A] - \omega_2 \bar{E}[S_{3,n}, A] - \omega_{3,n} \bar{E}[S_{2,n}, A] \\
&\quad + \omega_2 \omega_{3,n} P(X \in A) \\
&= \frac{n_b}{n^2} \bar{E}[Y_{2,t} X^3_{t-1}, A] + \frac{n_b}{n^2}(n_b - 1)\bar{E}[Y_{2,t} X_{t-1} X^2_{t-2}, A] \\
&\quad + \frac{n_b}{n^2}(n - n_b)\bar{E}[Y_{2,t} X_{t-1} X^2_{t-n_b-1}, A] - \omega_2 \frac{n_b}{n} \bar{E}[Y_{2,t} X_{t-1}, A] \\
&\quad - \omega_{3,n} \bar{E}[X^2_{t-1}, A] + \omega_2 \omega_{3,n} P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{2,n} - \omega_2)(S_{3,n} - \omega_{3,n})$ is as follows:

- $\bar{E}[(S_{2,n} - \omega_2)(S_{3,n} - \omega_{3,n}), A] = \dfrac{n_b}{n}\Big[\bar{E}[Y_{2,t} X_{t-1} X^2_{t-2}, A] - \omega_2 \bar{E}[Y_{2,t} X_{t-1}, A]$

$$- E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] + \omega_2 E[Y_{2,t}X_{t-1}]P(X \in A)\Big]$$
$$+ \frac{n_b}{n^2}\Big[\bar{E}[Y_{2,t}X_{t-1}^3, A] + (n_b - 1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] - n_b\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\Big].$$

The truncated expectation of $(S_{2,n} - \omega_2)(S_{4,n} - \omega_4)$ is as follows:

- $\bar{E}[(S_{2,n} - \omega_2)(S_{4,n} - \omega_4), A] = \Big[\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] - \omega_2\bar{E}[Y_{2,t}X_{t-1}, A]$
$- E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] + \omega_2 E[Y_{2,t}X_{t-1}]P(X \in A)$
$- \bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A] + \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\Big]$
$+ \frac{1}{n}\Big[\bar{E}[Y_{2,t}X_{t-1}^3, A] - \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\Big].$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{3,n} - \omega_{3,n}), A]$:

$$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})(S_{3,n} - \omega_{3,n}), A] &= \bar{E}[S_{1,n}S_{3,n}, A] - \omega_{1,n}\bar{E}[S_{3,n}, A] - \omega_{3,n}\bar{E}[S_{1,n}, A] \\
&\quad + \omega_{1,n}\omega_{3,n}P(X \in A) \\
&= n_b\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A] - \frac{n_b}{n}\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}, A] \\
&\quad - \left(1 - \frac{n_b}{n}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] + \omega_{1,n}\omega_{3,n}P(X \in A).
\end{aligned}$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{3,n} - \omega_{3,n})$ is as follows:

- $\bar{E}[(S_{1,n}-\omega_{1,n})(S_{3,n} - \omega_{3,n}), A] = \frac{n_b}{n}\Big[\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A]$
$- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}, A] - E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]$
$+ E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]P(X \in A)\Big]$
$+\frac{n_b^2}{n^2}\Big[- \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A] + E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}, A]$
$+ E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]$
$- E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]P(X \in A)\Big].$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})^2, A]$:

$$\bar{E}[(S_{1,n}-\omega_{1,n})^2, A] = \bar{E}[S_{1,n}^2, A] - 2\omega_{1,n}\bar{E}[S_{1,n}, A] + \omega_{1,n}^2 P(X \in A)$$

$$
\begin{aligned}
&= \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&\quad + \left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right) \bar{E}\left[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A\right] \\
&\quad - 2\omega_{1,n}\left(1 - \frac{n_b}{n}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] + \omega_{1,n}^2 P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})^2$ is as follows:

- $$
\begin{aligned}
\bar{E}[(S_{1,n} - \omega_{1,n})^2, A] = &\Big[\bar{E}\left[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A\right] \\
&- 2E[Y_{1,t-n_b} X_{t-n_b}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] + E^2[Y_{1,t-n_b} X_{t-n_b}] P(X \in A)\Big] \\
+ \frac{1}{n}&\Big[\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] - (2n_b+1)\bar{E}\left[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A\right] \\
&+ 4n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&- 2n_b E^2[Y_{1,t-n_b} X_{t-n_b-1}] P(X \in A)\Big] \\
+ \frac{n_b}{n^2}&\Big[(n_b+1)\bar{E}\left[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A\right] - \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&- 2n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&+ n_b E^2[Y_{1,t-n_b} X_{t-n_b-1}] P(X \in A)\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{2,n} - \omega_2)^2, A]$:

$$
\begin{aligned}
\bar{E}[(S_{2,n} - \omega_2)^2, A] &= \bar{E}[S_{2,n}^2, A] - 2\omega_2 \bar{E}[S_{2,n}, A] + \omega_2^2 P(X \in A) \\
&= \frac{1}{n}\bar{E}[X_{t-1}^4, A] + \left(1 - \frac{1}{n}\right)\bar{E}[X_{t-1}^2 X_{t-2}^2, A] - 2\omega_2 \bar{E}[X_{t-1}^2, A] + \omega_2^2 P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{2,n} - \omega_2)^2$ is as follows:

- $$
\begin{aligned}
\bar{E}[(S_{2,n} - \omega_2)^2, A] = &\Big[\bar{E}[X_{t-1}^2 X_{t-2}^2, A] - 2\omega_2 \bar{E}[X_{t-1}^2, A] + \omega_2^2 P(X \in A)\Big] \\
&+ \frac{1}{n}\Big[\bar{E}[X_{t-1}^4, A] - \bar{E}[X_{t-1}^2 X_{t-2}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{3,n} - \omega_{3,n})^2, A]$:

$$
\begin{aligned}
\bar{E}[(S_{3,n} - \omega_{3,n})^2, A] = &\bar{E}[S_{3,n}^2, A] - 2\omega_{3,n}\bar{E}[S_{3,n}, A] + \omega_{3,n}^2 P(X \in A) \\
= &\frac{n_b}{n^2}\bar{E}\left[Y_{2,t}^2 X_{t-1}^2, A\right] + \left(\frac{n_b^2}{n^2} - \frac{n_b}{n^2}\right)\bar{E}\left[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}, A\right]
\end{aligned}
$$

$$- 2\omega_{3,n}\frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A] + \omega_{3,n}^2 P(X \in A).$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})^2$ is as follows:

- $$\bar{E}[(S_{3,n} - \omega_{3,n})^2, A] = \frac{n_b}{n^2}\Big[\bar{E}\left[Y_{2,t}^2X_{t-1}^2, A\right] + (n_b - 1)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right] - 2n_bE[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A] + n_bE^2[Y_{2,t}X_{t-1}]P(X \in A)\Big].$$

The truncated expectation of $(S_{4,n} - \omega_4)^2$ is as follows:

- $$\bar{E}[(S_{4,n} - \omega_4)^2, A] = \Big[\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right] - 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A] + E^2[Y_{2,t}X_{t-1}]P(X \in A)\Big] + \frac{1}{n}\Big[\bar{E}[Y_{2,t}^2X_{t-1}^2, A] - \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A]\Big].$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2, A]$:

$$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n} - \omega_2)^2, A] &= \bar{E}[S_{1,n}S_{2,n}^2, A] - \omega_{1,n}\bar{E}[S_{2,n}^2, A] - 2\omega_2\bar{E}[S_{1,n}S_{2,n}, A] \\
&\quad + 2\omega_{1,n}\omega_2\bar{E}[S_{2,n}, A] + \omega_2^2\bar{E}[S_{1,n}, A] - \omega_{1,n}\omega_2^2P(X \in A) \\
&= \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A] \\
&\quad + \left(\frac{1}{n} - \frac{1}{n^2}(2n_b + 1) + \frac{1}{n^3}n_b(n_b + 1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&\quad + 2\left(\frac{1}{n} - \frac{1}{n^2}(2n_b + 1) + \frac{1}{n^3}n_b(n_b + 1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&\quad + \left(1 - \frac{3}{n}(n_b + 1) + \frac{1}{n^2}(2 + 6n_b + 3n_b^2) - \frac{n_b}{n^3}(2 + 3n_b + n_b^2)\right) \\
&\qquad \cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&\quad + n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A] \\
&\quad + n_b(n_b - 1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A] \\
&\quad + 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A]
\end{aligned}$$

$$+ 2n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A]$$
$$- \frac{\omega_{1,n}}{n}\bar{E}[X_{t-1}^4, A] - \omega_{1,n}\left(1-\frac{1}{n}\right)\bar{E}[X_{t-1}^2X_{t-2}^2, A]$$
$$- 2\omega_2\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]$$
$$- 2\omega_2\left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right)\bar{E}[Y_{t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]$$
$$- 2\omega_2 n_b\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A] + 2\omega_{1,n}\omega_2\bar{E}[X_{t-1}^2, A]$$
$$+ \omega_2^2\left(1-\frac{n_b}{n}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] - \omega_{1,n}\omega_2^2P(X\in A).$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2$ is as follows:

- $$\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^2, A] = \Big[&\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] - 2\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&+ 2\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] + \omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] \\
&- \omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]P(X\in A)\Big] \\
+\frac{1}{n}\Big[&\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] + 2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&- 3(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&+ 2n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A] - E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^4, A] \\
&+ (n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] - 2\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A] \\
&+ 2(2n_b+1)\omega_2\bar{E}[Y_{t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&- 2\omega_2n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A] - 2n_b\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] \\
&- n_b\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] + n_b\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]P(X\in A)\Big] \\
+\frac{1}{n^2}\Big[&\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A] - (2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&- 2(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&+ (2+6n_b+3n_b^2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&+ n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A] + n_b(n_b-1)\bar{E}[Y_{t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A] \\
&+ 2n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A] \\
&- 2n_b(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A]
\end{aligned}$$

$$
\begin{aligned}
&+ n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[X_{t-1}^4, A] - n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2, A] \\
&+ 2n_b \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] - 2n_b(n_b+1)\omega_2 \bar{E}[Y_{t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&+ 2n_b^2 \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A]\Big] \\
+&\frac{1}{n^3}\Big[- n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5, A] + n_b(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] \\
&+ 2n_b(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&- n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&- n_b^2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^4, A] + n_b^2(n_b-1)\bar{E}[Y_{t-n_b} X_{t-n_b-1} X_{t-1}^2 X_{t-2}^2, A] \\
&- 2n_b^2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-1}^2, A] \\
&+ 2n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2)^2, A]$:

$$
\begin{aligned}
\bar{E}[(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2)^2, A] &= \bar{E}[S_{3,n} S_{2,n}^2, A] - \omega_{3,n}\bar{E}[S_{2,n}^2, A] - 2\omega_2 \bar{E}[S_{2,n} S_{3,n}, A] \\
&+ 2\omega_2\omega_{3,n}\bar{E}[S_{2,n}, A] + \omega_2^2 \bar{E}[S_{3,n}, A] - \omega_2^2 \omega_{3,n} P(X \in A) \\
=&\frac{n_b}{n^3}\bar{E}[Y_{2,t} X_{t-1}^5, A] + \frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^4, A] \\
&+ 2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t} X_{t-1}^3 X_{t-2}^2, A] + \frac{n_b}{n^3}(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-3}^2, A] \\
&+ n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^4, A] \\
&+ n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] \\
&+ 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t} X_{t-1}^3 X_{t-n_b-1}^2, A] \\
&+ 2n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-n_b-1}^2, A] \\
&- \omega_{3,n}\frac{1}{n}\bar{E}[X_{t-1}^4, A] - \left(1 - \frac{1}{n}\right)\omega_{3,n}\bar{E}[X_{t-1}^2 X_{t-2}^2, A] \\
&- 2\omega_2\frac{n_b}{n^2}\bar{E}[Y_{2,t} X_{t-1}^3, A] - 2\frac{n_b}{n^2}(n_b-1)\omega_2\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2, A] \\
&- 2\frac{n_b}{n^2}(n-n_b)\omega_2\bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2, A] + 2\omega_2\omega_{3,n}\bar{E}[X_{t-1}^2, A] \\
&+ \frac{n_b}{n}\omega_2^2\bar{E}[Y_{2,t} X_{t-1}, A] - \omega_2^2\omega_{3,n} P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^2,A] &= \frac{n_b}{n}\Big[\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2},A]\\
&\quad - E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}X^2_{t-2},A] - 2\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1},A]\\
&\quad + 2\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1},A] + \omega_2^2\bar{E}[Y_{2,t}X_{t-1},A] - \omega_2^2 E[Y_{2,t}X_{t-1}]P(X\in A)\Big]\\
&+\frac{1}{n^2}\Big[n_b\bar{E}[Y_{2,t}X_{t-1}X^4_{t-n_b-1},A] - n_b(2n_b+1)\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2},A]\\
&\quad + 2n_b\bar{E}[Y_{2,t}X^3_{t-1}X^2_{t-n_b-1},A] + 2n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-n_b-1},A]\\
&\quad - n_bE[Y_{2,t}X_{t-1}]\bar{E}[X^4_{t-1},A] + n_bE[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}X^2_{t-2},A]\\
&\quad - 2n_b\omega_2\bar{E}[Y_{2,t}X^3_{t-1},A] - 2n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2},A]\\
&\quad + 2n_b^2\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1},A]\Big]\\
&+\frac{1}{n^3}\Big[n_b\bar{E}[Y_{2,t}X^5_{t-1},A] + n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X^4_{t-2},A]\\
&\quad + 2n_b(n_b-1)\bar{E}[Y_{2,t}X^3_{t-1}X^2_{t-2},A] + n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3},A]\\
&\quad - n_b^2\bar{E}[Y_{2,t}X_{t-1}X^4_{t-n_b-1},A] + n_b^2(n_b+1)\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2},A]\\
&\quad - 2n_b^2\bar{E}[Y_{2,t}X^3_{t-1}X^2_{t-n_b-1},A] - 2n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-n_b-1},A]\Big].\end{aligned}$$

The truncated expectation of $(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^2$ is as follows:

- $$\begin{aligned}\bar{E}[(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^2,A] &= \Big[-E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}X^2_{t-2},A]\\
&\quad + 2\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1},A] + \omega_2^2\bar{E}[Y_{2,t}X_{t-1},A] - \omega_2^2E[Y_{2,t}X_{t-1}]P(X\in A)\\
&\quad - 2\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2},A] + \bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3},A]\Big]\\
&+\frac{1}{n}\Big[\bar{E}[Y_{2,t}X_{t-1}X^4_{t-n_b-1},A] - E[Y_{2,t}X_{t-1}]\bar{E}[X^4_{t-1},A]\\
&\quad + E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}X^2_{t-2},A] - 2\omega_2\bar{E}[Y_{2,t}X^3_{t-1},A]\\
&\quad + 2\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2},A] + \bar{E}[Y_{2,t}X_{t-1}X^4_{t-2},A] + 2\bar{E}[Y_{2,t}X^3_{t-1}X^2_{t-2},A]\\
&\quad - 3\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3},A]\Big]\\
&+\frac{1}{n^2}\Big[\bar{E}[Y_{2,t}X^5_{t-1},A] - \bar{E}[Y_{2,t}X_{t-1}X^4_{t-2},A] - 2\bar{E}[Y_{2,t}X^3_{t-1}X^2_{t-2},A]\\
&\quad + 2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}X^2_{t-3},A]\Big].\end{aligned}$$

Next, we expand $\bar{E}[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2), A]$:

$$
\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2), A] &= \bar{E}[S_{1,n}^2 S_{2,n}, A] - 2\omega_{1,n}\bar{E}[S_{1,n}S_{2,n}, A] - \omega_2\bar{E}[S_{1,n}^2, A] \\
&\quad + \omega_{1,n}^2\bar{E}[S_{2,n}, A] + 2\omega_{1,n}\omega_2\bar{E}[S_{1,n}, A] - \omega_{1,n}^2\omega_2 P(X\in A) \\
&= \left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n+b-1}^4, A] \\
&\quad + \left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] \\
&\quad + n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A] \\
&\quad + 2\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3 Y_{1,t-n_b-1}X_{t-n_b-2}, A] \\
&\quad + \left(1-\frac{3}{n}(n_b+1)+\frac{1}{n^2}(2+6n_b+3n_b^2)-\frac{n_b}{n^3}(2+3n_b+n_b^2)\right) \\
&\qquad\qquad \cdot\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2, A] \\
&\quad + n_b\left(\frac{1}{n}-\frac{1}{n^2}(2n_b+1)+\frac{1}{n^3}n_b(n_b+1)\right) \\
&\qquad\qquad \cdot\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2, A] \\
&\quad - 2\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\omega_{1,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A] \\
&\quad - 2\left(1-\frac{1}{n}(2n_b+1)+\frac{n_b}{n^2}(n_b+1)\right)\omega_{1,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&\quad - 2n_b\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\omega_{1,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A] \\
&\quad - \omega_2\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&\quad - \left(1-\frac{1}{n}(2n_b+1)+\frac{n_b}{n^2}(n_b+1)\right)\omega_2\bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A\right] \\
&\quad + \omega_{1,n}^2\bar{E}[X_{t-1}^2, A] + 2\omega_{1,n}\omega_2\left(1-\frac{n_b}{n}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] - \omega_{1,n}^2\omega_2 P(X\in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2)$ is as follows:

- $$
\begin{aligned}
\bar{E}[(S_{1,n}-\omega_{1,n})^2(S_{2,n}-\omega_2), A] &= \Big[\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2, A] \\
&\quad - 2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&\quad - \omega_2\bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A\right] + E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] \\
&\quad + 2\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}, A\right] - \omega_2 E^2[Y_{1,t-n_b}X_{t-n_b-1}]P(X\in A)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{n}\Big[\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] + 2\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&\quad - 3(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&\quad + n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&\quad - 2E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&\quad + 2(3n_b+1)E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&\quad - 2n_b E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] - \omega_2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&\quad + (2n_b+1)\omega_2 \bar{E}\,[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&\quad - 2n_b E^2[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] \\
&\quad - 4n_b \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&\quad + 2n_b \omega_2 E^2[Y_{1,t-n_b} X_{t-n_b-1}]P(X \in A)\Big] \\
&+\frac{1}{n^2}\Big[\bar{E}[Y_{1,t-n_b}^2 X_{t-n+b-1}^4, A] - (2n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] \\
&\quad + n_b \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A] - 2(2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&\quad + (2+6n_b+3n_b^2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&\quad - n_b(2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&\quad + 4n_b E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&\quad - 2n_b(3n_b+2)E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&\quad + 4n_b^2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] + n_b \omega_2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&\quad - n_b(n_b+1)\omega_2 \bar{E}\,[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&\quad + n_b^2 E^2[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] \\
&\quad + 2n_b^2 \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&\quad - n_b^2 \omega_2 E^2[Y_{1,t-n_b} X_{t-n_b-1}]P(X \in A)\Big] \\
&+\frac{1}{n^3}\Big[-n_b \bar{E}[Y_{1,t-n_b}^2 X_{t-n+b-1}^4, A] + n_b(n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] \\
&\quad - n_b^2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A] \\
&\quad + 2n_b(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&\quad - n_b(2+3n_b+n_b^2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&\quad + n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A]
\end{aligned}
$$

$$- 2n_b^2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}, A]$$
$$+ 2n_b^2(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A]$$
$$- 2n_b^3 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}, A]\Big].$$

Next, we expand $\bar{E}[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2), A]$:

$$\begin{aligned}
\bar{E}[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2), A] &= \bar{E}[S^2_{3,n}S_{2,n}, A] - 2\omega_{3,n}\bar{E}[S_{3,n}S_{2,n}, A] - \omega_2\bar{E}[S^2_{3,n}, A] \\
&\quad + \omega^2_{3,n}\bar{E}[S_{2,n}, A] + 2\omega_{3,n}\omega_2\bar{E}[S_{3,n}, A] - \omega^2_{3,n}\omega_2 P(X\in A) \\
&= \frac{n_b}{n^3}\bar{E}[Y^2_{2,t}X^4_{t-1}, A] + \frac{n_b}{n^3}(n_b-1)\bar{E}[Y^2_{2,t}X^2_{t-1}X^2_{t-2}, A] \\
&\quad + n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y^2_{2,t}X^2_{t-1}X^2_{t-n_b-1}, A] + 2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X^3_{t-1}Y_{2,t-1}X_{t-2}, A] \\
&\quad + \frac{n_b}{n^3}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^2_{t-3}, A] \\
&\quad + n_b(n_b-1)\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^2_{t-n_b-1}, A] \\
&\quad - 2\omega_{3,n}\frac{n_b}{n^2}\bar{E}[Y_{2,t}X^3_{t-1}, A] - 2\frac{n_b}{n^2}(n_b-1)\omega_{3,n}\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}, A] \\
&\quad - 2\frac{n_b}{n^2}(n-n_b)\omega_{3,n}\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}, A] \\
&\quad - \omega_2\frac{n_b}{n^2}\bar{E}\left[Y^2_{2,t}X^2_{t-1}, A\right] - \omega_2\frac{n_b}{n^2}(n_b-1)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right] \\
&\quad + \omega^2_{3,n}\bar{E}[X^2_{t-1}, A] + 2\omega_2\omega_{3,n}\frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A] - \omega^2_{3,n}\omega_2 P(X\in A).
\end{aligned}$$

The truncated expectation of $(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)$ is as follows:

- $$\begin{aligned}
\bar{E}[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2), A] &= \frac{1}{n^2}\Big[n_b\bar{E}[Y^2_{2,t}X^2_{t-1}X^2_{t-n_b-1}, A] \\
&\quad + n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X^2_{t-n_b-1}, A] \\
&\quad - 2n_b^2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}, A] \\
&\quad - n_b\omega_2\bar{E}\left[Y^2_{2,t}X^2_{t-1}, A\right] - \omega_2 n_b(n_b-1)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right] \\
&\quad + n_b^2E^2[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}, A] + 2n_b^2\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A] \\
&\quad - n_b^2\omega_2 E^2[Y_{2,t}X_{t-1}]P(X\in A)\Big] \\
&\quad + \frac{1}{n^3}\Big[n_b\bar{E}[Y^2_{2,t}X^4_{t-1}, A] + n_b(n_b-1)\bar{E}[Y^2_{2,t}X^2_{t-1}X^2_{t-2}, A] \\
&\quad - n_b^2\bar{E}[Y^2_{2,t}X^2_{t-1}X^2_{t-n_b-1}, A] + 2n_b(n_b-1)\bar{E}[Y_{2,t}X^3_{t-1}Y_{2,t-1}X_{t-2}, A]
\end{aligned}$$

$$+ n_b(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]$$
$$- n_b^2(n_b - 1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A]$$
$$- 2n_b^2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A] - 2n_b^2(n_b - 1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]$$
$$+ 2n_b^3 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\Big].$$

The truncated expectation of $(S_{4,n} - \omega_4)^2(S_{2,n} - \omega_2)$ is as follows:

- $\bar{E}[(S_{4,n} - \omega_4)^2(S_{2,n} - \omega_2), A] = \Big[- \omega_2\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right]$

$$+ E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] + 2\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A]$$
$$- \omega_2 E^2[Y_{2,t}X_{t-1}]P(X \in A) + \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]$$
$$- 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\Big]$$
$$+\frac{1}{n}\Big[- \omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2, A] + \omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A]$$
$$+ \bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2, A] + 2\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}, A]$$
$$- 3\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A] - 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A]$$
$$+ 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] + 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\Big]$$
$$+\frac{1}{n^2}\Big[\bar{E}[Y_{2,t}^2X_{t-1}^4, A] - \bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2, A]$$
$$- 2\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}, A] + 2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]\Big].$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)(S_{3,n} - \omega_{3,n}), A]$:

$$\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)(S_{3,n} - \omega_{3,n}), A] = \bar{E}[S_{1,n}S_{2,n}S_{3,n}, A] - \omega_{1,n}\bar{E}[S_{2,n}S_{3,n}, A]$$
$$- \omega_2\bar{E}[S_{1,n}S_{3,n}, A] - \omega_{3,n}\bar{E}[S_{1,n}S_{2,n}, A] + \omega_{1,n}\omega_2\bar{E}[S_{3,n}, A] + \omega_{1,n}\omega_{3,n}\bar{E}[S_{2,n}, A]$$
$$+ \omega_2\omega_{3,n}\bar{E}[S_{1,n}, A] - \omega_{1,n}\omega_2\omega_{3,n}P(X \in A)$$
$$= n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A]$$
$$+ n_b\left(\frac{1}{n} - (2n_b + 1)\frac{1}{n^2} + n_b(n_b + 1)\frac{1}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]$$
$$+ n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A]$$
$$+ n_b(n_b - 1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A]$$

$$
\begin{aligned}
&- \frac{n_b}{n^2}\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}^3, A] - \frac{n_b}{n^2}(n_b-1)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&- \frac{n_b}{n^2}(n-n_b)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\\
&- n_b\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A]\\
&- \left(\frac{1}{n}-\frac{n_b}{n^2}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&- \left(1-\frac{1}{n}(2n_b+1)+\frac{n_b}{n^2}(n_b+1)\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&- n_b\left(\frac{1}{n}-\frac{n_b}{n^2}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A]\\
&+ \omega_{1,n}\omega_2\frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A] + \omega_{1,n}\omega_{3,n}\bar{E}[X_{t-1}^2, A]\\
&+ \omega_2\omega_{3,n}\left(1-\frac{n_b}{n}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] - \omega_{1,n}\omega_2\omega_{3,n}P(X\in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n})$ is as follows:

- $\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)(S_{3,n}-\omega_{3,n}), A]$

$$
\begin{aligned}
=&\frac{n_b}{n}\Big[\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A] - \omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A]\\
&- E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&+ \omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}, A] + E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A]\\
&+ \omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]\\
&- \omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]P(X\in A)\Big]\\
+&\frac{n_b}{n^2}\Big[\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&- (2n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&+ \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A] + (n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A]\\
&- (n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&+ 2n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\\
&+ n_b\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A] - E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&+ (2n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&- n_b E[Y_{2,t} X_{t-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] \\
&- n_b \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{2,t} X_{t-1}, A] \\
&- n_b E[Y_{1,t-n_b} X_{t-n_b-1}] E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2, A] \\
&- n_b \omega_2 E[Y_{2,t} X_{t-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&+ n_b \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] E[Y_{2,t} X_{t-1}] P(X \in A) \Big] \\
+&\frac{n_b^2}{n^3} \Big[- \bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&+ (n_b + 1) \bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] - \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{2,t} X_{t-1}^3, A] \\
&- (n_b - 1) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{2,t} X_{t-1} X_{t-2}^2, A] + E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{2,t} X_{t-1}^3, A] \\
&+ (n_b - 1) E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2, A] \\
&- n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2, A] + E[Y_{2,t} X_{t-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&- (n_b + 1) E[Y_{2,t} X_{t-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&+ n_b E[Y_{2,t} X_{t-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] \Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^3, A]$:

$$
\begin{aligned}
\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^3, A] &= \bar{E}[S_{1,n} S_{2,n}^3, A] - \omega_{1,n} \bar{E}[S_{2,n}^3, A] - 3\omega_2 \bar{E}[S_{1,n} S_{2,n}^2, A] \\
&\quad + 3\omega_{1,n}\omega_2 \bar{E}[S_{2,n}^2, A] + 3\omega_2^2 \bar{E}[S_{1,n} S_{2,n}, A] - 3\omega_{1,n}\omega_2^2 \bar{E}[S_{2,n}, A] \\
&\quad - \omega_2^3 \bar{E}[S_{1,n}, A] + \omega_{1,n}\omega_2^3 P(X \in A) \\
&= \left(\frac{1}{n^3} - \frac{n_b}{n^4} \right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^7, A] \\
&\quad + \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1) \right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^6, A] \\
&\quad + 3 \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1) \right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5 X_{t-n_b-2}^2, A] \\
&\quad + 3 \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1) \right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^4, A] \\
&\quad + 3 \left(\frac{1}{n} - \frac{3}{n^2}(n_b + 1) + \frac{1}{n^3}(3n_b^2 + 6n_b + 2) - \frac{n_b}{n^4}(n_b^2 + 3n_b + 2) \right) \\
&\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4 X_{t-n_b-3}^2, A] \\
&\quad + 3 \left(\frac{1}{n} - \frac{3}{n^2}(n_b + 1) + \frac{1}{n^3}(3n_b^2 + 6n_b + 2) - \frac{n_b}{n^4}(n_b^2 + 3n_b + 2) \right)
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1} X^2_{t-n_b-2} X^2_{t-n_b-3}, A] \\
&+ \left(1 - \frac{1}{n}(4n_b+6) + \frac{1}{n^2}(6n_b^2+18n_b+11) - \frac{1}{n^3}(4n_b^3+18n_b^2+22n_b+6)\right. \\
&\left. + \frac{n_b}{n^4}(n_b^3+6n_b^2+11n_b+6)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2} X^2_{t-n_b-3} X^2_{t-n_b-4}, A] \\
&+ n_b \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^6_{t-1}, A] \\
&+ 3n_b(n_b-1) \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^4_{t-1} X^2_{t-2}, A] \\
&+ n_b(n_b^2-3n_b+2) \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-1} X^2_{t-2} X^2_{t-3}, A] \\
&+ 3n_b \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X^5_{t-n_b-1} X^2_{t-1}, A] \\
&+ 3n_b \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^4_{t-n_b-2} X^2_{t-1}, A] \\
&+ 6n_b \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1} X^2_{t-n_b-2} X^2_{t-1}, A] \\
&+ 3n_b \left(\frac{1}{n} - \frac{3}{n^2}(n_b+1) + \frac{1}{n^3}(3n_b^2+6n_b+2) - \frac{n_b}{n^4}(n_b^2+3n_b+2)\right) \\
&\cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2} X^2_{t-n_b-3} X^2_{t-1}, A] \\
&+ 3n_b \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1} X^4_{t-1}, A] \\
&+ 3n_b(n_b-1) \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right) \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1} X^2_{t-1} X^2_{t-2}, A] \\
&+ 3n_b \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2} X^4_{t-1}, A] \\
&+ 3n_b(n_b-1) \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \\
&\cdot \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^2_{t-n_b-2} X^2_{t-1} X^2_{t-2}, A] \\
&- \frac{\omega_{1,n}}{n^2} \bar{E}\left[X^6_{t-1}, A\right] - 3\left(\frac{1}{n} - \frac{1}{n^2}\right) \omega_{1,n} \bar{E}\left[X^4_{t-1} X^2_{t-2}, A\right] \\
&- \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right) \omega_{1,n} \bar{E}\left[X^2_{t-1} X^2_{t-2} X^2_{t-3}, A\right] \\
&- 3\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \omega_2 \bar{E}[Y_{1,t-n_b} X^5_{t-n_b-1}, A] \\
&- 3\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3} n_b(n_b+1)\right) \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X^4_{t-n_b-2}, A] \\
&- 6\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3} n_b(n_b+1)\right) \omega_2 \bar{E}[Y_{1,t-n_b} X^3_{t-n_b-1} X^2_{t-n_b-2}, A]
\end{aligned}
$$

$$- 3\left(1 - \frac{3}{n}(n_b+1) + \frac{1}{n^2}(2+6n_b+3n_b^2) - \frac{n_b}{n^3}(2+3n_b+n_b^2)\right)\omega_2$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$

$$- 3n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A]$$

$$- 3n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A]$$

$$- 6n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A]$$

$$- 6n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A]$$

$$+ 3\omega_{1,n}\omega_2\frac{1}{n}\bar{E}[X_{t-1}^4, A] + 3\omega_{1,n}\omega_2\left(1 - \frac{1}{n}\right)\bar{E}[X_{t-1}^2X_{t-2}^2, A]$$

$$+ 3\omega_2^2\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]$$

$$+ 3\omega_2^2\left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]$$

$$+ 3\omega_2^2 n_b\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A] - 3\omega_{1,n}\omega_2^2\bar{E}[X_{t-1}^2, A]$$

$$- \omega_2^3\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] + \omega_{1,n}\omega_2^3 P(X \in A).$$

The truncated expectation of $(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3$ is as follows:

- $$\bar{E}[(S_{1,n}-\omega_{1,n})(S_{2,n}-\omega_2)^3, A] = \Big[\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-n_b-4}^2, A]$$

$$- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$

$$- 3\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$

$$+ 3\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] + 3\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]$$

$$- 3\omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] - \omega_2^3\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]$$

$$+ \omega_2^3 E[Y_{1,t-n_b}X_{t-n_b-1}]P(X \in A)\Big]$$

$$+\frac{1}{n}\Big[3\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-n_b-3}^2, A]$$

$$+ 3\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$

$$- (4n_b+6)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-n_b-4}^2, A]$$

$$+ 3n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-1}^2, A]$$

$$- 3E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^4X_{t-2}^2, A]$$

$$
\begin{aligned}
&+ (n_b + 3)E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\
&- 3\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] - 6\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&+ 9(n_b + 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&- 6n_b \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2, A] + 3\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^4, A] \\
&- 3(n_b + 1)\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2 X_{t-2}^2, A] + 3\omega_2^2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&- 3\omega_2^2 (2n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] + 3\omega_2^2 n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] \\
&+ 3\omega_2^2 n_b E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] + \omega_2^3 n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&- \omega_2^3 n_b E[Y_{1,t-n_b} X_{t-n_b-1}]P(X \in A)\Big] \\
+\frac{1}{n^2}\Big[&\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^6, A] + 3\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5 X_{t-n_b-2}^2, A] \\
&+ 3\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^4, A] \\
&- 9(n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4 X_{t-n_b-3}^2, A] \\
&- 9(n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&+ (6n_b^2 + 18n_b + 11)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2 X_{t-n_b-4}^2, A] \\
&+ 3n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4 X_{t-1}^2, A] + 6n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2 X_{t-1}^2, A] \\
&- 9n_b(n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2 X_{t-1}^2, A] \\
&+ 3n_b \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^4, A] \\
&+ 3n_b(n_b - 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2 X_{t-2}^2, A] \\
&- E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^6, A] \\
&+ 3(n_b + 1)E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^4 X_{t-2}^2, A] \\
&- (3n_b + 2)E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] \\
&- 3\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5, A] + 3(2n_b + 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] \\
&+ 6(2n_b + 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&- 3(3n_b^2 + 6n_b + 2)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&- 3n_b \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^4, A] \\
&- 3n_b(n_b - 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2 X_{t-2}^2, A] \\
&- 6n_b \omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-1}^2, A] \\
&+ 6n_b(2n_b + 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
& - 3n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^4_{t-1}, A] \\
& + 3n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^2_{t-1}X^2_{t-2}, A] - 3n_b\omega_2^2\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}, A] \\
& + 3n_b(n_b+1)\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A] \\
& - 3n_b^2\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}, A]\Big] \\
+ & \frac{1}{n^3}\Big[\bar{E}[Y_{1,t-n_b}X^7_{t-n_b-1}, A] - (2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^6_{t-n_b-2}, A] \\
& - 3(2n_b+1)\bar{E}[Y_{1,t-n_b}X^5_{t-n_b-1}X^2_{t-n_b-2}, A] \\
& - 3(2n_b+1)\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^4_{t-n_b-2}, A] \\
& + 3(3n_b^2+6n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}X^2_{t-n_b-3}, A] \\
& + 3(3n_b^2+6n_b+2)\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A] \\
& - (4n_b^3+18n_b^2+22n_b+6)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-n_b-4}, A] \\
& + n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^6_{t-1}, A] + 3n_b(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-1}X^2_{t-2}, A] \\
& + n_b(n_b^2-3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}X^2_{t-2}X^2_{t-3}, A] \\
& + 3n_b\bar{E}[Y_{1,t-n_b}X^5_{t-n_b-1}X^2_{t-1}, A] - 3n_b(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}X^2_{t-1}, A] \\
& - 6n_b(2n_b+1)\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A] \\
& + 3n_b(3n_b^2+6n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-1}, A] \\
& + 3n_b\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^4_{t-1}, A] + 3n_b(n_b-1)\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-1}X^2_{t-2}, A] \\
& - 3n_b(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^4_{t-1}, A] \\
& - 3n_b(n_b-1)(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}X^2_{t-2}, A] \\
& + n_b E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^6_{t-1}, A] - 3n_b E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^4_{t-1}X^2_{t-2}, A] \\
& + 2n_b E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^2_{t-1}X^2_{t-2}X^2_{t-3}, A] + 3n_b\omega_2\bar{E}[Y_{1,t-n_b}X^5_{t-n_b-1}, A] \\
& - 3n_b(n_b+1)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}, A] \\
& - 6n_b(n_b+1)\omega_2\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}, A] \\
& + 3n_b(n_b^2+3n_b+2)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A] \\
& + 3n_b^2\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-1}, A] \\
& + 3n_b^2(n_b-1)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}X^2_{t-2}, A] + 6n_b^2\omega_2\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-1}, A] \\
& - 6n_b^2(n_b+1)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A]\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{n^4}\Big[-n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^7, A] + n_b(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^6, A] \\
&\quad + 3n_b(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-n_b-2}^2, A] \\
&\quad + 3n_b(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^4, A] \\
&\quad - 3n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-n_b-3}^2, A] \\
&\quad - 3n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&\quad + n_b(n_b^3+6n_b^2+11n_b+6)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-n_b-4}^2, A] \\
&\quad - n_b^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^6, A] - 3n_b^2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4X_{t-2}^2, A] \\
&\quad - n_b^2(n_b^2-3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2X_{t-3}^2, A] \\
&\quad - 3n_b^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5X_{t-1}^2, A] + 3n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4X_{t-1}^2, A] \\
&\quad + 6n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2X_{t-1}^2, A] \\
&\quad - 3n_b^2(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2X_{t-1}^2, A] \\
&\quad - 3n_b^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^4, A] - 3n_b^2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2X_{t-2}^2, A] \\
&\quad + 3n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^4, A] \\
&\quad + 3n_b^2(n_b-1)(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2X_{t-2}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3, A]$:

$$
\begin{aligned}
\bar{E}[(S_{3,n}-\omega_{3,n})(S_{2,n}-\omega_2)^3, A] &= \bar{E}[S_{3,n}S_{2,n}^3, A] - \omega_{3,n}\bar{E}[S_{2,n}^3, A] - 3\omega_2\bar{E}[S_{3,n}S_{2,n}^2, A] \\
&\quad + 3\omega_{3,n}\omega_2\bar{E}[S_{2,n}^2, A] + 3\omega_2^2\bar{E}[S_{3,n}S_{2,n}, A] - 3\omega_{3,n}\omega_2^2\bar{E}[S_{2,n}, A] \\
&\quad - \omega_2^3\bar{E}[S_{3,n}, A] + \omega_{3,n}\omega_2^3P(X\in A) \\
&= n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^6, A] \\
&\quad + 3n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4X_{t-n_b-2}^2, A] \\
&\quad + n_b\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right) \\
&\qquad\qquad \cdot\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&\quad + \frac{n_b}{n^4}\bar{E}[Y_{2,t}X_{t-1}^7, A] + \frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A] \\
&\quad + 3\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] + 3\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A]
\end{aligned}
$$

$$
\begin{aligned}
&+ 3\frac{n_b}{n^4}(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A] \\
&+ 3\frac{n_b}{n^4}(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A] \\
&+ \frac{n_b}{n^4}(n_b^3 - 6n_b^2 + 11n_b - 6)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A] \\
&+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^4, A] \\
&+ 3n_b(n_b - 1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^4, A] \\
&+ 3n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&+ 3n_b(n_b - 1)\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&+ 3n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-n_b-1}^2, A] \\
&+ 3n_b(n_b - 1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-n_b-1}^2, A] \\
&+ 6n_b(n_b - 1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-n_b-1}^2, A] \\
&+ 3n_b(n_b^2 - 3n_b + 2)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-n_b-1}^2, A] \\
&- \frac{\omega_{3,n}}{n^2}\bar{E}\left[X_{t-1}^6, A\right] - 3\left(\frac{1}{n} - \frac{1}{n^2}\right)\omega_{3,n}\bar{E}\left[X_{t-1}^4X_{t-2}^2, A\right] \\
&- \left(1 - 3\frac{1}{n} + 2\frac{1}{n^2}\right)\omega_{3,n}\bar{E}\left[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A\right] \\
&- 3\frac{n_b}{n^3}\omega_2\bar{E}[Y_{2,t}X_{t-1}^5, A] - 3\frac{n_b}{n^3}(n_b - 1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A] \\
&- 6\frac{n_b}{n^3}(n_b - 1)\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] \\
&- 3\frac{n_b}{n^3}(n_b^2 - 3n_b + 2)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] \\
&- 3n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A] \\
&- 3n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b + 1) + \frac{1}{n^3}n_b(n_b + 1)\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&- 6n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A] \\
&- 6n_b(n_b - 1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A] \\
&+ 3\omega_2\omega_{3,n}\frac{1}{n}\bar{E}[X_{t-1}^4, A] + 3\omega_2\omega_{3,n}\left(1 - \frac{1}{n}\right)\bar{E}[X_{t-1}^2X_{t-2}^2, A]
\end{aligned}
$$

$$+ 3\omega_2^2 \frac{n_b}{n^2} \bar{E}[Y_{2,t} X_{t-1}^3, A] + 3\omega_2^2 \frac{n_b}{n^2}(n_b - 1)\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2, A]$$
$$+ 3\omega_2^2 n_b \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2, A] - 3\omega_2^2 \omega_{3,n} \bar{E}[X_{t-1}^2, A]$$
$$- \omega_2^3 \frac{n_b}{n} \bar{E}[Y_{2,t} X_{t-1}, A] + \omega_2^3 \omega_{3,n} P(X \in A).$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2)^3$ is as follows:

- $$\bar{E}[(S_{3,n} - \omega_{3,n})(S_{2,n} - \omega_2)^3, A] = \frac{1}{n}\Big[n_b \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A]$$
$$- 3n_b\omega_2 \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] - n_b E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A]$$
$$+ 3n_b\omega_2^2 \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2, A] + 3n_b\omega_2 E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2, A]$$
$$- n_b\omega_2^3 \bar{E}[Y_{2,t} X_{t-1}, A] - 3n_b\omega_2^2 E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2, A]$$
$$+ n_b\omega_2^3 E[Y_{2,t} X_{t-1}] P(X \in A)\Big]$$
$$+ \frac{1}{n^2}\Big[3n_b \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^4 X_{t-n_b-2}^2, A]$$
$$- 3n_b(n_b+1) \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A]$$
$$+ 3n_b \bar{E}[Y_{2,t} X_{t-1}^3 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A]$$
$$+ 3n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] - 3n_b\omega_2 \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^4, A]$$
$$+ 3n_b(2n_b+1)\omega_2 \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] - 6n_b\omega_2 \bar{E}[Y_{2,t} X_{t-1}^3 X_{t-n_b-1}^2, A]$$
$$- 6n_b(n_b-1)\omega_2 \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-n_b-1}^2, A] - 3n_b E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^4 X_{t-2}^2, A]$$
$$+ 3n_b E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + 3n_b\omega_2^2 \bar{E}[Y_{2,t} X_{t-1}^3, A]$$
$$+ 3n_b(n_b-1)\omega_2^2 \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2, A] - 3n_b^2\omega_2^2 \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2, A]$$
$$+ 3n_b\omega_2 E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^4, A] - 3n_b\omega_2 E[Y_{2,t} X_{t-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2, A]\Big]$$
$$+ \frac{1}{n^3}\Big[n_b \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^6, A] - 3n_b(2n_b+1) \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^4 X_{t-n_b-2}^2, A]$$
$$+ n_b(3n_b^2 + 6n_b + 2) \bar{E}[Y_{2,t} X_{t-1} X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A]$$
$$+ 3n_b \bar{E}[Y_{2,t} X_{t-1}^3 X_{t-n_b-1}^4, A] + 3n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-n_b-1}^4, A]$$
$$- 3n_b(2n_b+1) \bar{E}[Y_{2,t} X_{t-1}^3 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A]$$
$$- 3n_b(n_b-1)(2n_b+1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A]$$
$$+ 3n_b \bar{E}[Y_{2,t} X_{t-1}^5 X_{t-n_b-1}^2, A]$$
$$+ 3n_b(n_b-1) \bar{E}[Y_{2,t} X_{t-1} X_{t-2}^4 X_{t-n_b-1}^2, A]$$

$$
\begin{aligned}
&+ 6n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-n_b-1}^2, A] \\
&+ 3n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-n_b-1}^2, A] - 3n_b\omega_2\bar{E}[Y_{2,t}X_{t-1}^5, A] \\
&- 3n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A] - 6n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] \\
&- 3n_b(n_b^2-3n_b+2)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] + 3n_b^2\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A] \\
&- 3n_b^2(n_b+1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A] + 6n_b^2\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A] \\
&+ 6n_b^2(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A] - n_bE[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^6, A] \\
&+ 3n_bE[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4X_{t-2}^2, A] - 2n_bE[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]\Big] \\
+&\frac{1}{n^4}\Big[- n_b^2\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^6, A] + 3n_b^2(n_b+1)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4X_{t-n_b-2}^2, A] \\
&- n_b^2(n_b^2+3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&+ n_b\bar{E}[Y_{2,t}X_{t-1}^7, A] + n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A] \\
&+ 3n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] + 3n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A] \\
&+ 3n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A] \\
&+ 3n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A] \\
&+ n_b(n_b^3-6n_b^2+11n_b-6)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A] \\
&- 3n_b^2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^4, A] - 3n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^4, A] \\
&+ 3n_b^2(n_b+1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&+ 3n_b^2(n_b-1)(n_b+1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&- 3n_b^2\bar{E}[Y_{2,t}X_{t-1}^5X_{t-n_b-1}^2, A] - 3n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-n_b-1}^2, A] \\
&- 6n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-n_b-1}^2, A] \\
&- 3n_b^2(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-n_b-1}^2, A]\Big].
\end{aligned}
$$

The truncated expectation of $(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^3$ is as follows:

- $$
\begin{aligned}
\bar{E}[(S_{4,n}-\omega_4)(S_{2,n}-\omega_2)^3, A] = &\Big[- E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A] \\
&+ 3\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] - \omega_2^3\bar{E}[Y_{2,t}X_{t-1}, A] \\
&- 3\omega_2^2E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] + \omega_2^3E[Y_{2,t}X_{t-1}]P(X\in A) \\
&+ 3\omega_2^2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] - 3\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&+ \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]\Big] \\
&+\frac{1}{n}\Big[- 3E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4X_{t-2}^2, A] + 3E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A] \\
&+ 3\omega_2^2\bar{E}[Y_{2,t}X_{t-1}^3, A] - 3\omega_2^2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] \\
&+ 3\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4, A] - 3\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] \\
&- 3\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A] - 6\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] \\
&+ 9\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] + 3\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A] \\
&+ 3\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A] - 6\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]\Big] \\
&+\frac{1}{n^2}\Big[- 3\omega_2\bar{E}[Y_{2,t}X_{t-1}^5, A] + 3\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A] \\
&+ 6\omega_2\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] - 6\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] \\
&- E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^6, A] + 3E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4X_{t-2}^2, A] \\
&- 2E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2X_{t-3}^2, A] + \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A] \\
&+ 3\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] + 3\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A] \\
&- 9\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A] - 9\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A] \\
&+ 11\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]\Big] \\
&+\frac{1}{n^3}\Big[\bar{E}[Y_{2,t}X_{t-1}^7, A] - \bar{E}[Y_{2,t}X_{t-1}X_{t-2}^6, A] - 3\bar{E}[Y_{2,t}X_{t-1}^5X_{t-2}^2, A] \\
&- 3\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^4, A] + 6\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4X_{t-3}^2, A] \\
&+ 6\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2X_{t-3}^2, A] - 6\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2X_{t-4}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2)^2, A]$:

$$
\begin{aligned}
&\bar{E}[(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2)^2, A] = \bar{E}[S_{1,n}^2S_{2,n}^2, A] - 2\omega_2\bar{E}[S_{1,n}^2S_{2,n}, A] \\
&\quad - 2\omega_{1,n}\bar{E}[S_{1,n}S_{2,n}^2, A] + \omega_{1,n}^2\bar{E}[S_{2,n}^2, A] + 4\omega_{1,n}\omega_2\bar{E}[S_{1,n}S_{2,n}, A] \\
&\quad + \omega_2^2\bar{E}[S_{1,n}^2, A] - 2\omega_{1,n}^2\omega_2\bar{E}[S_{2,n}, A] - 2\omega_{1,n}\omega_2^2\bar{E}[S_{1,n}, A] + \omega_{1,n}^2\omega_2^2P(X \in A) \\
&= \left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^6, A] \\
&\quad + \left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^4, A] \\
&\quad + 2\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^4X_{t-n_b-2}^2, A]
\end{aligned}
$$

$$+\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A]$$
$$+2\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5Y_{1,t-n_b-1}X_{t-n_b-2}, A]$$
$$+\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^4, A]$$
$$+2\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}^3, A]$$
$$+4\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}^3X_{t-n_b-3}^2, A]$$
$$+\left(1-\frac{1}{n}(4n_b+6)+\frac{1}{n^2}(6n_b^2+18n_b+11)-\frac{1}{n^3}(4n_b^3+18n_b^2+22n_b+6)\right.$$
$$\left.+\frac{n_b}{n^4}(n_b^3+6n_b^2+11n_b+6)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-n_b-4}^2, A]$$
$$+n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^4, A]$$
$$+n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^2X_{t-2}^2, A]$$
$$+n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^4, A]$$
$$+n_b(n_b-1)\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2X_{t-2}^2, A]$$
$$+2n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^4X_{t-1}^2, A]$$
$$+2n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-1}^2, A]$$
$$+4n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)$$
$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2, A]$$
$$+2n_b\left(\frac{1}{n}-\frac{3}{n^2}(n_b+1)+\frac{1}{n^3}(3n_b^2+6n_b+2)-\frac{n_b}{n^4}(n_b^2+3n_b+2)\right)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-1}, A]$$

$$- 2\omega_2 \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y^2_{1,t-n_b}X^4_{t-n+b-1}, A]$$

$$- 2\omega_2 \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right) \bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-n_b-2}, A]$$

$$- 2\omega_2 n_b \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-1}, A]$$

$$- 4\omega_2 \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A]$$

$$- 2\omega_2 \left(1 - \frac{3}{n}(n_b+1) + \frac{1}{n^2}(2+6n_b+3n_b^2) - \frac{n_b}{n^3}(2+3n_b+n_b^2)\right)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-n_b-3}, A]$$

$$- 2\omega_2 n_b \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-1}, A]$$

$$- 2\omega_{1,n} \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y_{1,t-n_b}X^5_{t-n_b-1}, A]$$

$$- 2\omega_{1,n} \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-n_b-2}, A]$$

$$- 4\omega_{1,n} \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-n_b-2}, A]$$

$$- 2\omega_{1,n} \left(1 - \frac{3}{n}(n_b+1) + \frac{1}{n^2}(2+6n_b+3n_b^2) - \frac{n_b}{n^3}(2+3n_b+n_b^2)\right)$$

$$\cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A]$$

$$- 2\omega_{1,n} n_b \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-1}, A]$$

$$- 2\omega_{1,n} n_b(n_b-1) \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}X^2_{t-2}, A]$$

$$- 4\omega_{1,n} n_b \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right) \bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-1}, A]$$

$$- 4\omega_{1,n} n_b \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A]$$

$$+ \omega^2_{1,n}\frac{1}{n}\bar{E}[X^4_{t-1}, A] + \omega^2_{1,n}\left(1 - \frac{1}{n}\right) \bar{E}[X^2_{t-1}X^2_{t-2}, A]$$

$$+ 4\omega_{1,n}\omega_2 \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}, A]$$

$$+ 4\omega_{1,n}\omega_2 \left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A]$$

$$+ 4\omega_{1,n}\omega_2 n_b \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A]$$
$$+ \omega_2^2 \left(\frac{1}{n} - \frac{n_b}{n^2}\right) \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A]$$
$$+ \omega_2^2 \left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right) \bar{E}\left[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A\right]$$
$$- 2\omega_{1,n}^2\omega_2 \bar{E}[X_{t-1}^2, A] - 2\omega_{1,n}\omega_2^2\left(1 - \frac{n_b}{n}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] + \omega_{1,n}^2\omega_2^2 P(X \in A).$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2)^2$ is as follows:

- $\bar{E}[(S_{1,n} - \omega_{1,n})^2(S_{2,n} - \omega_2)^2, A]$

$$\begin{aligned}
&= \Big[\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-n_b-4}^2, A] \\
&\quad - 2\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2, A] \\
&\quad - 2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&\quad + E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] \\
&\quad + 4\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&\quad + \omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A] - 2\omega_2E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X_{t-1}^2, A] \\
&\quad - 2\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] + \omega_2^2E^2[Y_{1,t-n_b}X_{t-n_b-1}]P(X \in A)\Big] \\
&+ \frac{1}{n}\Big[\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&\quad + \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^4, A] \\
&\quad + 4\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}^3X_{t-n_b-3}^2, A] \\
&\quad - (4n_b+6)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-n_b-4}^2, A] \\
&\quad + 2n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-1}^2, A] \\
&\quad - 2\omega_2\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&\quad - 4\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}, A] \\
&\quad + 6(n_b+1)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2, A] \\
&\quad - 2n_b\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2, A] \\
&\quad - 2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&\quad - 4E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]
\end{aligned}$$

$$+ (8n_b + 6)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A]$$
$$- 4n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A]$$
$$+ E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^4_{t-1}, A]$$
$$- (2n_b + 1)E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^2_{t-1}X^2_{t-2}, A]$$
$$+ 4\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}, A]$$
$$- 4(3n_b + 1)\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A]$$
$$+ 4n_b\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}, A]$$
$$+ \omega_2^2\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}, A]$$
$$- (2n_b + 1)\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A]$$
$$+ 4n_b\omega_2E^2[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[X^2_{t-1}, A]$$
$$+ 4n_b\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A]$$
$$- 2n_b\omega_2^2E^2[Y_{1,t-n_b}X_{t-n_b-1}]P(X \in A)\Big]$$
$$+\frac{1}{n^2}\Big[\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^4_{t-n_b-2}, A] + 2\bar{E}[Y^2_{1,t-n_b}X^4_{t-n_b-1}X^2_{t-n_b-2}, A]$$
$$- 3(n_b + 1)\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A]$$
$$+ 2\bar{E}[Y_{1,t-n_b}X^5_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A]$$
$$- 3(n_b + 1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^4_{t-n_b-3}, A]$$
$$+ 2\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{1,t-n_b-1}X^3_{t-n_b-2}, A]$$
$$- 12(n_b + 1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X^3_{t-n_b-2}X^2_{t-n_b-3}, A]$$
$$+ (6n_b^2 + 18n_b + 11)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-n_b-4}, A]$$
$$+ n_b\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^4_{t-1}, A]$$
$$+ n_b(n_b - 1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-1}X^2_{t-2}, A]$$
$$+ 2n_b\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A]$$
$$+ 4n_b\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-1}, A]$$
$$- 6n_b(n_b + 1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X^2_{t-n_b-3}X^2_{t-1}, A]$$
$$- 2\omega_2\bar{E}[Y^2_{1,t-n_b}X^4_{t-n_b-1}, A] + 2(2n_b + 1)\omega_2\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-n_b-2}, A]$$
$$- 2n_b\omega_2\bar{E}[Y^2_{1,t-n_b}X^2_{t-n_b-1}X^2_{t-1}, A]$$
$$+ 4(2n_b + 1)\omega_2\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}, A]$$

$$
\begin{aligned}
&- 2(3n_b^2 + 6n_b + 2)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&+ 2n_b(2n_b + 1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&- 2E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5, A] \\
&+ 2(3n_b + 1)E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] \\
&+ 4(3n_b + 1)E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&- 2(6n_b^2 + 9n_b + 2)E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&- 2n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^4, A] \\
&- 2n_b(n_b - 1)E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2 X_{t-2}^2, A] \\
&- 4n_b E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-1}^2, A] \\
&+ 4n_b(3n_b + 1)E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2, A] \\
&- 2n_b E^2[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[X_{t-1}^4, A] \\
&+ n_b(n_b + 2)E^2[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[X_{t-1}^2 X_{t-2}^2, A] \\
&- 8n_b\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&+ 4(3n_b + 2)\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&- 8n_b^2\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A] \\
&- n_b\omega_2^2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2, A] \\
&+ n_b(n_b + 1)\omega_2^2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&- 2n_b^2\omega_2 E^2[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[X_{t-1}^2, A] \\
&- 2n_b^2\omega_2^2 E[Y_{1,t-n_b} X_{t-n_b-1}] \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}, A] \\
&+ n_b^2\omega_2^2 E^2[Y_{1,t-n_b} X_{t-n_b-1}] P(X \in A)\Big] \\
+&\frac{1}{n^3}\Big[\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^6, A] - (2n_b + 1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^4, A] \\
&- 2(2n_b + 1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2, A] \\
&+ (3n_b^2 + 6n_b + 2)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&- 2(2n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&+ (3n_b^2 + 6n_b + 2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^4, A] \\
&- 2(2n_b + 1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}^3, A] \\
&+ 4(3n_b^2 + 6n_b + 2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2}^3 X_{t-n_b-3}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&- (4n_b^3 + 18n_b^2 + 22n_b + 6)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-n_b-4}^2, A] \\
&+ n_b \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^4, A] + n_b(n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2 X_{t-2}^2, A] \\
&- n_b(2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^4, A] \\
&- n_b(n_b-1)(2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2 X_{t-2}^2, A] \\
&+ 2n_b \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-1}^2, A] \\
&- 2n_b(2n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2 X_{t-1}^2, A] \\
&- 4n_b(2n_b+1)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&+ 2n_b(3n_b^2 + 6n_b + 2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2 X_{t-1}^2, A] \\
&+ 2n_b \omega_2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4, A] - 2n_b(2n_b+1)\omega_2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^2, A] \\
&+ 2n_b^2 \omega_2 \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-1}^2, A] \\
&- 4n_b(n_b+1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 Y_{1,t-n_b-1} X_{t-n_b-2}, A] \\
&+ 2n_b(n_b^2 + 3n_b + 2)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-n_b-3}^2, A] \\
&- 2n_b^2(n_b+1)\omega_2 \bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{1,t-n_b-1} X_{t-n_b-2} X_{t-1}^2, A] \\
&+ 4n_b E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^5, A] \\
&- n_b(6n_b+4) E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] \\
&- 2n_b(6n_b+4) E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&+ 2n_b(4n_b^2 + 9n_b + 4) E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&+ 4n_b^2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^4, A] \\
&+ 4n_b^2(n_b-1) E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2 X_{t-2}^2, A] \\
&+ 8n_b^2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-1}^2, A] \\
&- 4n_b^2(3n_b+2) E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-1}^2, A] \\
&+ n_b^2 E^2[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^4, A] - n_b^2 E^2[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[X_{t-1}^2 X_{t-2}^2, A] \\
&+ 4n_b^2 \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1}^3, A] \\
&- 4n_b^2(n_b+1)\omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2, A] \\
&+ 4n_b^3 \omega_2 E[Y_{1,t-n_b} X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} X_{t-1}^2, A]\Big] \\
+&\frac{1}{n^4}\Big[- n_b \bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^6, A] + n_b(n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^2 X_{t-n_b-2}^4, A] \\
&+ 2n_b(n_b+1)\bar{E}[Y_{1,t-n_b}^2 X_{t-n_b-1}^4 X_{t-n_b-2}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&- n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&+ 2n_b(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5Y_{1,t-n_b-1}X_{t-n_b-2}, A] \\
&- n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^4, A] \\
&+ 2n_b(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}^3, A] \\
&- 4n_b(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}^3X_{t-n_b-3}^2, A] \\
&+ n_b(n_b^3+6n_b^2+11n_b+6)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-n_b-4}^2, A] \\
&- n_b^2\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^4, A] \\
&- n_b^2(n_b-1)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^2X_{t-2}^2, A] \\
&+ n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^4, A] \\
&+ n_b^2(n_b+1)(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2X_{t-2}^2, A] \\
&- 2n_b^2\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^4X_{t-1}^2, A] \\
&+ 2n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}^2X_{t-n_b-1}^2X_{t-n_b-2}^2X_{t-1}^2, A] \\
&+ 4n_b^2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-1}^2, A] \\
&- 2n_b^2(n_b^2+3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{1,t-n_b-1}X_{t-n_b-2}X_{t-n_b-3}^2X_{t-1}^2, A] \\
&- 2n_b^2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A] \\
&+ 2n_b^2(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&+ 4n_b^2(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&- 2n_b^2(n_b^2+3n_b+2)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&- 2n_b^3E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A] \\
&- 2n_b^3(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A] \\
&- 4n_b^3E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A] \\
&+ 4n_b^3(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2, A]$:

$$
\begin{aligned}
\bar{E}[(S_{3,n}-\omega_{3,n})^2(S_{2,n}-\omega_2)^2, A] &= \bar{E}[S_{3,n}^2S_{2,n}^2, A] - 2\omega_2\bar{E}[S_{3,n}^2S_{2,n}, A] \\
&- 2\omega_{3,n}\bar{E}[S_{3,n}S_{2,n}^2, A] + \omega_{3,n}^2\bar{E}[S_{2,n}^2, A] + 4\omega_{3,n}\omega_2\bar{E}[S_{3,n}S_{2,n}, A] \\
&+ \omega_2^2\bar{E}[S_{3,n}^2, A] - 2\omega_{3,n}^2\omega_2\bar{E}[S_{2,n}, A] - 2\omega_{3,n}\omega_2^2\bar{E}[S_{3,n}, A] + \omega_{3,n}^2\omega_2^2P(X\in A)
\end{aligned}
$$

$$
\begin{aligned}
=&n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t-n_b}^2X_{t-n_b-1}^2X_{t-1}^4,A]\\
&+n_b\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2,A]\\
&+n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^4,A]\\
&+n_b(n_b-1)\left(\frac{1}{n^2}-\frac{1}{n^3}(2n_b+1)+\frac{n_b}{n^4}(n_b+1)\right)\\
&\qquad\qquad\qquad\cdot\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2,A]\\
&+\frac{n_b}{n^4}\bar{E}[Y_{2,t}^2X_{t-1}^6,A]+\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^4,A]\\
&+2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-2}^2,A]+\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-3}^2,A]\\
&+2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^5Y_{2,t-1}X_{t-2},A]\\
&+\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^4,A]\\
&+2\frac{n_b}{n^4}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}^3,A]\\
&+4\frac{n_b}{n^4}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-3}^2,A]\\
&+\frac{n_b}{n^4}(n_b^3-6n_b^2+11n_b-6)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2,A]\\
&+2n_b\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-n_b-1}^2,A]\\
&+2n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-n_b-1}^2,A]\\
&+4n_b(n_b-1)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2,A]\\
&+2n_b(n_b^2-3n_b+2)\left(\frac{1}{n^3}-\frac{n_b}{n^4}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-n_b-1}^2,A]\\
&-2\omega_2\frac{n_b}{n^3}\bar{E}[Y_{2,t}^2X_{t-1}^4,A]-2\omega_2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2,A]\\
&-2\omega_2n_b\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2,A]\\
&-4\omega_2\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2},A]\\
&-2\omega_2\frac{n_b}{n^3}(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2,A]\\
&-2\omega_2n_b(n_b-1)\left(\frac{1}{n^2}-\frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2,A]\\
&-2\omega_{3,n}\frac{n_b}{n^3}\bar{E}[Y_{2,t}X_{t-1}^5,A]-2\omega_{3,n}\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4,A]\\
&-4\omega_{3,n}\frac{n_b}{n^3}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2,A]
\end{aligned}
$$

$$- 2\omega_{3,n}\frac{n_b}{n^3}(n_b^2 - 3n_b + 2)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]$$
$$- 2\omega_{3,n}n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]$$
$$- 2\omega_{3,n}n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$- 4\omega_{3,n}n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]$$
$$- 4\omega_{3,n}n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]$$
$$+ \omega_{3,n}^2\frac{1}{n}\bar{E}[X_{t-1}^4, A] + \omega_{3,n}^2\left(1 - \frac{1}{n}\right)\bar{E}[X_{t-1}^2X_{t-2}^2, A]$$
$$+ 4\omega_2\omega_{3,n}\frac{n_b}{n^2}\bar{E}[Y_{2,t}X_{t-1}^3, A] + 4\omega_2\omega_{3,n}\frac{n_b}{n^2}(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]$$
$$+ 4\omega_2\omega_{3,n}\frac{n_b}{n^2}(n-n_b)\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]$$
$$+ \omega_2^2\frac{n_b}{n^2}\bar{E}\left[Y_{2,t}^2X_{t-1}^2, A\right] + \omega_2^2\frac{n_b}{n^2}(n_b-1)\bar{E}\left[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A\right]$$
$$- 2\omega_2\omega_{3,n}^2\bar{E}[X_{t-1}^2, A] - 2\omega_2^2\omega_{3,n}\frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A] + \omega_2^2\omega_{3,n}^2P(X \in A).$$

The truncated expectation of $(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2)^2$ is as follows:

- $$\bar{E}[(S_{3,n} - \omega_{3,n})^2(S_{2,n} - \omega_2)^2, A] = \frac{1}{n^2}\Big[n_b\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$- 2n_b\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2, A]$$
$$- 2n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A]$$
$$- 2n_b^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ n_b^2E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A] + 4n_b^2\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]$$
$$+ n_b\omega_2^2\bar{E}[Y_{2,t}^2X_{t-1}^2, A] + n_b(n_b-1)\omega_2^2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A]$$
$$- 2n_b^2\omega_2E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] - 2n_b^2\omega_2^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A]$$
$$+ n_b^2\omega_2^2E^2[Y_{2,t}X_{t-1}]P(X \in A)\Big]$$
$$+\frac{1}{n^3}\Big[n_b\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^4, A] - n_b(2n_b+1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^4, A]$$
$$- n_b(n_b-1)(2n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$

$$+ 2n_b\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-n_b-1}^2, A] + 2n_b(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-n_b-1}^2, A]$$
$$+ 4n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-n_b-1}^2, A]$$
$$+ 2n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-n_b-1}^2, A]$$
$$- 2n_b\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^4, A] - 2n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2, A]$$
$$+ 2n_b^2\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2, A] - 4n_b(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}, A]$$
$$- 2n_b(n_b^2-3n_b+2)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]$$
$$+ 2n_b^2(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2, A]$$
$$- 2n_b^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]$$
$$+ 2n_b^2(2n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$- 4n_b^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]$$
$$- 4n_b^2(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]$$
$$+ n_b^2E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4, A] - n_b^2E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A]$$
$$+ 4n_b^2\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A]$$
$$+ 4n_b^2(n_b-1)\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]$$
$$- 4n_b^3\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\Big]$$
$$+\frac{1}{n^4}\Big[- n_b^2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^4, A] + n_b^2(n_b+1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$- n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^4, A]$$
$$+ n_b^2(n_b-1)(n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ n_b\bar{E}[Y_{2,t}^2X_{t-1}^6, A] + n_b(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^4, A]$$
$$+ 2n_b(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-2}^2, A]$$
$$+ n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$
$$+ 2n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^5, A]$$
$$+ n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^4, A]$$
$$+ 2n_b(n_b-1)\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}^3, A]$$
$$+ 4n_b(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-3}^2, A]$$
$$+ n_b(n_b^3-6n_b^2+11n_b-6)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2, A]$$
$$- 2n_b^2\bar{E}[Y_{2,t}^2X_{t-1}^4X_{t-n_b-1}^2, A] - 2n_b^2(n_b-1)\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-n_b-1}^2, A]$$

$$- 4n_b^2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-n_b-1}^2, A]$$
$$- 2n_b^2(n_b^2-3n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-n_b-1}^2, A]$$
$$- 2n_b^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^5, A] - 2n_b^2(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A]$$
$$- 4n_b^2(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A]$$
$$- 2n_b^2(n_b^2-3n_b+2)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]$$
$$+ 2n_b^3E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]$$
$$- 2n_b^3(n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]$$
$$+ 4n_b^3E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]$$
$$+ 4n_b^3(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]\Big].$$

The truncated expectation of $(S_{4,n}-\omega_4)^2(S_{2,n}-\omega_2)^2$ is as follows:

- $\bar{E}[(S_{4,n}-\omega_4)^2(S_{2,n}-\omega_2)^2, A] = \Big[\omega_2^2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A]$

$$- 2\omega_2E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] - 2\omega_2^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}, A]$$
$$+ \omega_2^2E^2[Y_{2,t}X_{t-1}]P(X\in A) - 2\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]$$
$$+ 4\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] + \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2, A]$$
$$- 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]\Big]$$
$$+\frac{1}{n}\Big[\omega_2^2\bar{E}[Y_{2,t}^2X_{t-1}^2, A] - \omega_2^2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}, A]$$
$$- 2\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2, A] - 4\omega_2\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}, A]$$
$$+ 6\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A] - 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]$$
$$+ E^2[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4, A] + 4\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A]$$
$$- 4\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] + \bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2X_{t-3}^2, A]$$
$$+ \bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^4, A] + 4\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}^3X_{t-3}^2, A]$$
$$- 6\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2X_{t-4}^2, A] - 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A]$$
$$- 4E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] + 6E[Y_{2,t}X_{t-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]\Big]$$
$$+\frac{1}{n^2}\Big[- 2\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^4, A] + 2\omega_2\bar{E}[Y_{2,t}^2X_{t-1}^2X_{t-2}^2, A]$$
$$+ 4\omega_2\bar{E}[Y_{2,t}X_{t-1}^3Y_{2,t-1}X_{t-2}, A] - 4\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{2,t-1}X_{t-2}X_{t-3}^2, A]$$

$$
\begin{aligned}
&+ \bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^4, A] + 2\bar{E}[Y_{2,t}^2 X_{t-1}^4 X_{t-2}^2, A] \\
&- 3\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] + 2\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^5, A] \\
&- 3\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^4, A] + 2\bar{E}[Y_{2,t} X_{t-1}^3 Y_{2,t-1} X_{t-2}^3, A] \\
&- 12\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^3 X_{t-3}^2, A] + 11\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A] \\
&- 2E[Y_{2,t} X_{t-1}]\bar{E}[Y_{2,t} X_{t-1}^5, A] + 2E[Y_{2,t} X_{t-1}]\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^4, A] \\
&+ 4E[Y_{2,t} X_{t-1}]\bar{E}[Y_{2,t} X_{t-1}^3 X_{t-2}^2, A] - 4E[Y_{2,t} X_{t-1}]\bar{E}[Y_{2,t} X_{t-1} X_{t-2}^2 X_{t-3}^2, A]\Big] \\
+&\frac{1}{n^3}\Big[\bar{E}[Y_{2,t}^2 X_{t-1}^6, A] - \bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^4, A] - 2\bar{E}[Y_{2,t}^2 X_{t-1}^4 X_{t-2}^2, A] \\
&+ 2\bar{E}[Y_{2,t}^2 X_{t-1}^2 X_{t-2}^2 X_{t-3}^2, A] - 2\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^5, A] \\
&+ 2\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^4, A] - 2\bar{E}[Y_{2,t} X_{t-1}^3 Y_{2,t-1} X_{t-2}^3, A] \\
&+ 8\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2}^3 X_{t-3}^2, A] - 6\bar{E}[Y_{2,t} X_{t-1} Y_{2,t-1} X_{t-2} X_{t-3}^2 X_{t-4}^2, A]\Big].
\end{aligned}
$$

Next, we expand $\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2(S_{3,n} - \omega_{3,n}), A]$:

$$
\begin{aligned}
&\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2(S_{3,n} - \omega_{3,n}), A] = \bar{E}[S_{1,n} S_{2,n}^2 S_{3,n}, A] - \omega_{1,n}\bar{E}[S_{2,n}^2 S_{3,n}, A] \\
&\quad - 2\omega_2\bar{E}[S_{1,n} S_{2,n} S_{3,n}, A] - \omega_{3,n}\bar{E}[S_{1,n} S_{2,n}^2, A] + 2\omega_{1,n}\omega_2\bar{E}[S_{2,n} S_{3,n}, A] \\
&\quad + \omega_2^2\bar{E}[S_{1,n} S_{3,n}, A] + \omega_{1,n}\omega_{3,n}\bar{E}[S_{2,n}^2, A] + 2\omega_2\omega_{3,n}\bar{E}[S_{1,n} S_{2,n}, A] - \omega_{1,n}\omega_2^2\bar{E}[S_{3,n}, A] \\
&\quad - 2\omega_{1,n}\omega_2\omega_{3,n}\bar{E}[S_{2,n}, A] - \omega_2^2\omega_{3,n}\bar{E}[S_{1,n}, A] + \omega_{1,n}\omega_2^2\omega_{3,n}P(X \in A) \\
&= n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1}^5, A] \\
&\quad + n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right)\bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^4, A] \\
&\quad + 2n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b + 1) + \frac{n_b}{n^4}(n_b + 1)\right)\bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1}^3 X_{t-n_b-2}^2, A] \\
&\quad + n_b\left(\frac{1}{n} - \frac{3}{n^2}(n_b + 1) + \frac{1}{n^3}(3n_b^2 + 6n_b + 2) - \frac{n_b}{n^4}(n_b^2 + 3n_b + 2)\right) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{2,t} X_{t-1} Y_{1,t-n_b} X_{t-n_b-1} X_{t-n_b-2}^2 X_{t-n_b-3}^2, A] \\
&\quad + n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{2,t} X_{t-1}^5, A] \\
&\quad + n_b(n_b - 1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{2,t} X_{t-1} X_{t-2}^4, A] \\
&\quad + 2n_b(n_b - 1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b} X_{t-n_b-1} Y_{2,t} X_{t-1}^3 X_{t-2}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&+ n_b(n_b^2 - 3n_b + 2)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] \\
&+ 2n_b\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}^3, A] \\
&+ 2n_b(n_b-1)\left(\frac{1}{n^3} - \frac{n_b}{n^4}\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}X_{t-2}^2, A] \\
&+ 2n_b\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}^3, A] \\
&+ 2n_b(n_b-1)\left(\frac{1}{n^2} - \frac{1}{n^3}(2n_b+1) + \frac{n_b}{n^4}(n_b+1)\right) \\
&\qquad\qquad\qquad \cdot \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}X_{t-2}^2, A] \\
&- \frac{n_b}{n^3}\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}^5, A] - \frac{n_b}{n^3}(n_b-1)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A] \\
&- 2\frac{n_b}{n^3}(n_b-1)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A] \\
&- \frac{n_b}{n^3}(n_b^2 - 3n_b + 2)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A] \\
&- n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A] \\
&- n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&- 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A] \\
&- 2n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{1,n}\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A] \\
&- 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A] \\
&- 2n_b\left(\frac{1}{n} - (2n_b+1)\frac{1}{n^2} + n_b(n_b+1)\frac{1}{n^3}\right)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&- 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A] \\
&- 2n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A] \\
&- \left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A] \\
&- \left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&- 2\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&- \left(1 - \frac{3}{n}(n_b+1) + \frac{1}{n^2}(2 + 6n_b + 3n_b^2) - \frac{n_b}{n^3}(2 + 3n_b + n_b^2)\right)
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A] \\
&- n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^4_{t-1}, A] \\
&- n_b(n_b-1)\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}X^2_{t-2}, A] \\
&- 2n_b\left(\frac{1}{n^2} - \frac{n_b}{n^3}\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}X^2_{t-1}, A] \\
&- 2n_b\left(\frac{1}{n} - \frac{1}{n^2}(2n_b+1) + \frac{1}{n^3}n_b(n_b+1)\right)\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-1}, A] \\
&+ 2\frac{n_b}{n^2}\omega_{1,n}\omega_2\bar{E}[Y_{2,t}X^3_{t-1}, A] + 2\frac{n_b}{n^2}(n_b-1)\omega_{1,n}\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-2}, A] \\
&+ 2\frac{n_b}{n^2}(n-n_b)\omega_{1,n}\omega_2\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}, A] \\
&+ n_b\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A] \\
&+ \frac{1}{n}\omega_{1,n}\omega_{3,n}\bar{E}[X^4_{t-1}, A] + \left(1 - \frac{1}{n}\right)\omega_{1,n}\omega_{3,n}\bar{E}[X^2_{t-1}X^2_{t-2}, A] \\
&+ 2\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\omega_2\omega_{3,n}\bar{E}[Y_{1,t-n_b}X^3_{t-n_b-1}, A] \\
&+ 2\left(1 - \frac{1}{n}(2n_b+1) + \frac{n_b}{n^2}(n_b+1)\right)\omega_2\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A] \\
&+ 2n_b\left(\frac{1}{n} - \frac{n_b}{n^2}\right)\omega_2\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-1}, A] - \omega_{1,n}\omega_2^2\frac{n_b}{n}\bar{E}[Y_{2,t}X_{t-1}, A] \\
&- 2\omega_{1,n}\omega_2\omega_{3,n}\bar{E}[X^2_{t-1}, A] - \left(1 - \frac{n_b}{n}\right)\omega_2^2\omega_{3,n}\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] \\
&+ \omega_{1,n}\omega_2^2\omega_{3,n}P(X \in A).
\end{aligned}
$$

The truncated expectation of $(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2(S_{3,n} - \omega_{3,n})$ is as follows:

- $$
\begin{aligned}
&\bar{E}[(S_{1,n} - \omega_{1,n})(S_{2,n} - \omega_2)^2(S_{3,n} - \omega_{3,n}), A] \\
&= \frac{n_b}{n}\Big[\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A] \\
&\quad - E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}X^2_{t-n_b-2}, A] \\
&\quad - 2\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}, A] \\
&\quad - E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X^2_{t-n_b-2}X^2_{t-n_b-3}, A] \\
&\quad + 2\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X^2_{t-n_b-1}, A] + \omega_2^2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}, A] \\
&\quad + E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X^2_{t-1}X^2_{t-2}, A]
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\omega_2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&- \omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}, A] \\
&- 2\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2, A] \\
&- \omega_2^2 E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}, A] \\
&+ \omega_2^2 E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]P(X \in A)\Big] \\
+&\frac{n_b}{n^2}\Big[\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&+ 2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&- 3(n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&+ 2\bar{E}[Y_{2,t}X_{t-1}^3Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&+ 2(n_b-1)\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A] \\
&+ n_b E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&+ (2n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A] \\
&- 2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A] \\
&- 2(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A] \\
&- 2\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A] \\
&+ 2(2n_b+1)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A] \\
&- 2\omega_2\bar{E}[Y_{2,t}X_{t-1}^3Y_{1,t-n_b}X_{t-n_b-1}, A] \\
&- 2(n_b-1)\omega_2\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2Y_{1,t-n_b}X_{t-n_b-1}, A] \\
&- E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A] \\
&- 2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A] \\
&+ 3(n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A] \\
&- 2n_b E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A] \\
&+ 2\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A] \\
&+ 2(n_b-1)\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A] \\
&- 2n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A] \\
&- 2n_b\omega_2 E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]
\end{aligned}
$$

$$
\begin{aligned}
&- n_b\omega_2^2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1},A] + E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4,A] \\
&- n_bE[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2,A] \\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2,A] \\
&+ 2\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3,A] \\
&- 2(2n_b+1)\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2,A] \\
&+ 2n_b\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2,A] \\
&+ n_b\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1},A] \\
&+ 2n_b\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2,A] \\
&+ n_b\omega_2^2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1},A] \\
&- n_b\omega_2^2E[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]P(X\in A)\Big] \\
+&\frac{n_b}{n^3}\Big[\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^5,A] \\
&- (2n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4,A] \\
&- 2(2n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2,A] \\
&+ (3n_b^2+6n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2,A] \\
&+ \bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^5,A] + (n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^4,A] \\
&+ 2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3X_{t-2}^2,A] \\
&+ (n_b^2-3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2,A] \\
&+ 2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}^3,A] \\
&+ 2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}X_{t-2}^2,A] \\
&- 2(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}^3,A] \\
&- 2(n_b-1)(2n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}X_{t-2}^2,A] \\
&- E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^5,A] \\
&- (n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4,A] \\
&- 2(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2,A] \\
&- (n_b^2-3n_b+2)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2,A] \\
&+ n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4,A] \\
&+ n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4,A]
\end{aligned}
$$

$$
\begin{aligned}
&- n_b(2n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]\\
&- n_b(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]\\
&+ 2n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]\\
&+ 2n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]\\
&+ 2n_b(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]\\
&+ 2n_b(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]\\
&+ 2n_b\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&- 2n_b(n_b+1)\omega_2\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&+ 2n_b\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3, A]\\
&+ 2n_b(n_b-1)\omega_2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&- E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A]\\
&+ (2n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A]\\
&+ 2(2n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]\\
&- (3n_b^2+6n_b+2)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]\\
&- n_bE[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A]\\
&- n_b(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A]\\
&- 2n_bE[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A]\\
&+ 2n_b(2n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A]\\
&- 2n_b\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3, A]\\
&- 2n_b(n_b-1)\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&+ 2n_b^2\omega_2E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2, A]\\
&- n_bE[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^4, A]\\
&+ n_bE[Y_{1,t-n_b}X_{t-n_b-1}]E[Y_{2,t}X_{t-1}]\bar{E}[X_{t-1}^2X_{t-2}^2, A]\\
&- 2n_b\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3, A]\\
&+ 2n_b(n_b+1)\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2, A]\\
&- 2n_b^2\omega_2E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2, A]\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{n_b^2}{n^4}\Big[-\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^5, A]\\
&+(n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A]\\
&+2(n_b+1)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]\\
&-(3n_b^2+6n_b+2)\bar{E}[Y_{2,t}X_{t-1}Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]\\
&-\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^5, A]-(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^4, A]\\
&-2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}^3X_{t-2}^2, A]\\
&-(n_b^2-3n_b+2)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]\\
&-2\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}^3, A]\\
&-2(n_b-1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&+2(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}^3, A]\\
&+2(n_b-1)(n_b+1)\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2Y_{2,t}X_{t-1}X_{t-2}^2, A]\\
&+E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^5, A]\\
&+(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^4, A]\\
&+2(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-2}^2, A]\\
&+(n_b^2-3n_b+2)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-3}^2, A]\\
&-n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^4, A]\\
&+n_b(n_b+1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-n_b-1}^2X_{t-n_b-2}^2, A]\\
&-2n_bE[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}^3X_{t-n_b-1}^2, A]\\
&-2n_b(n_b-1)E[Y_{1,t-n_b}X_{t-n_b-1}]\bar{E}[Y_{2,t}X_{t-1}X_{t-2}^2X_{t-n_b-1}^2, A]\\
&+E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^5, A]\\
&-(n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^4, A]\\
&-2(n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-n_b-2}^2, A]\\
&+(n_b^2+3n_b+2)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-n_b-3}^2, A]\\
&+n_bE[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^4, A]\\
&+n_b(n_b-1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-1}^2X_{t-2}^2, A]\\
&+2n_bE[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}^3X_{t-1}^2, A]\\
&-2n_b(n_b+1)E[Y_{2,t}X_{t-1}]\bar{E}[Y_{1,t-n_b}X_{t-n_b-1}X_{t-n_b-2}^2X_{t-1}^2, A].
\end{aligned}
$$

Appendix F

Appendix for Chapter 8

The following proposition was obtained from [18].

Proposition F.1 *Let $X_1, \cdots, X_n$ be i.i.d. real valued random variables. If $E|X_1|^j < \infty$, $j \geq 2$, then there are constants $C_j > 0$ and $D_j > 0$ such that*

$$E|\bar{X} - \mu|^j \leq C_j E|X_1|^j n^{-j/2},$$

$$|E(\bar{X} - \mu)^j| \leq D_j E|X_1|^j n^{-(j+1)/2}, j \quad \textit{odd}.$$

The following lemma is from [76].

Lemma F.2 (Hurt) *Let $i_1, \cdots, i_r$ be nonnegative real numbers, $\sum_{k=1}^{r} i_k = j$, $j > 0$, and $X_1, \cdots, X_r$ be random variables. Then*

$$E[\,|X_1|^{i_1} \ldots |X_r|^{i_r}] \leq \left\{[E|X_1|^j]^{i_1} \ldots [E|X_r|^j]^{i_r}\right\}^{1/j},$$

assuming only that the moments exist.

The following theorem was obtained from [94], p.154.

Theorem F.3 *If $X_n \xrightarrow{P} X$ and $|X_n| \leq Y$ with Y integrable, then X is integrable, and $E[X_n] \to E[X]$.*

Bibliography

[1] H. Akaike. Information theory and an extension of the likelihood principle. In B.N. Petrov and F. Csáki, editors, *Proceedings of the Second International Symposium on Information Theory*, pages 267–281, Budapest, 1973. Akadémiai Kiado.

[2] P.G. Allen and R. Fildes. Econometric forecasting. In J.S. Armstrong, editor, *Principles of Forecasting*. Springer, New York, 2001.

[3] G.S. Alogoskoufis and R. Smith. The phillips curve, the persistence of inflation, and the lucas critique: Evidence from exchange rate regimes. *American Economic Review*, 81:1254–1275, 1991.

[4] Donald W.K. Andrews. Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory*, 4:458–67, 1988.

[5] D.W.K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61:821–856, 1993.

[6] D.W.K Andrews and W. Ploberger. Optimal changepoint tests for normal linear regression. *Journal of Econometrics*, 70:9–38, 1996.

[7] T.M. Apostol. *Mathematical Analysis*. Addison-Wesley, Reading, Massachusetts, 1974.

[8] Tom M. Apostol. *Calculus*, volume 1. John Wiley and Sons, New York, second edition, 1967.

[9] J. Scott Armstrong, editor. *Principles of Forecasting*. Springer, New York, 2001.

[10] J.S. Armstrong. *Long-Range Forecasting from Crystal Ball to Computer*. John Wiley and Sons, New York, 2nd edition, 1985.

[11] J.S. Armstrong. Extrapolation for time-series and cross-sectional data. In J.S. Armstrong, editor, *Principles of Forecasting*, pages 217–243. Springer, New York, 2001.

[12] L. Arnold. Convergence in probability of random power series and a related problem in linear topological spaces. *Israel Journal of Mathematics*, 5(2):127–134, 1967.

[13] J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78, 1998.

[14] R.H. Berk. Limiting behavior of posterior distributions when the model is incorrect. *Annals of Mathematical Statistics*, 37:51–58, 1966.

[15] R.H. Berk. Consistency a posteriori. *Annals of Mathematical Statistics*, 41:894–906, 1970.

[16] R.J. Bhansali. Effects of not knowing the order of an autoregressive process on the mean squared error of prediction 1. *Journal of the American Statistical Association*, 76(375):588–597, 1981.

[17] R.J. Bhansali and F. Papangelou. Convergence of moments of least squares estimators for the coefficients of an autoregressive process of unknown order. *The Annals of Statistics*, 19(3):1155–1162, 1991.

[18] Peter J. Bickel and Kjell A. Doksum. *Mathematical Statistics, Vol 1.* Prentice Hall, New Jersey, 2001.

[19] Patrick Billingsley. *Probability and Measure.* Wiley, New York, 1995.

[20] D. Binkley and J.B. Birch. Estimation of regression coefficients under model misspecification and non-normality of errors. *J. Stat. Comput. and Simul*, 19:155–170, 1984.

[21] T. Birkel. Moment bounds for associated sequences. *Ann. Probab*, 16:1184–1193, 1988.

[22] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[23] T Bollerslev, R.Y. Chou, and K.F. Kroner. Arch modeling in finance : A selective review of the theory and empirical evidence. *Journal of Econometrics*, 52:5–59, 1992.

[24] T. Bollerslev, R.F. Engle, and D.B. Nelson. Arch models. In R. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4. North-Holland, Amsterdam, 1994.

[25] G.E.P. Box and G.M. Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, 1976.

[26] David R. Brillinger. A note on the rate of convergence of a mean. *Biometrika*, 49:574–576, Dec 1962.

[27] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 2nd edition, 1991.

[28] R.L. Brown, J. Durbin, and J.M. Evans. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, Series B, 37:149–192, 1975.

[29] G. Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28:591–605, 1960.

[30] C.S.J. Chu, M. Stinchcombe, and H. White. Monitoring structural change. *Econometrica*, 64:1045–1065, 1996.

[31] K. L. Chung. *A Course in Probability Theory*. Harcourt, New York, 1974.

[32] M.P. Clements and D. Henry. *Forecasting Economic Time Series*. Cambridge University Press, 1998.

[33] M.P. Clements and D. Henry. *Forecasting Non-stationary Economic Time Series*. Cambridge University Press, 1999.

[34] D.R. Cox. Tests of separate families of hypotheses. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 105–123, Berkeley, 1961. University of California Press.

[35] D.R. Cox. Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society*, B:406–424, 1962.

[36] H. Cramer. *Mathematical Methods of Statistics.* Princeton University Press, Princeton,NJ, 1946.

[37] James Davidson. *Stochastic Limit Theory.* Oxford University Press, New York, 1994.

[38] Yu. A. Davydov. Convergence of distributions generated by stationary stochastic processes. *Theor. Probability Appl*, 13:691–696, 1968.

[39] A.W. Van der Vaart. *Asymptotic Statistics.* Cambridge University Press, New York, 1998.

[40] Harvey Diamond and Gregory Gelles. Gaussian approximation of expected utility. *Economics Letters*, 64:301–307, 1999.

[41] Francis X. Diebold. The past, present, and future of macroeconomic forecasting. *Journal of Economic Perspectives*, 12(2):175–192, 1998.

[42] T.E. Dielman and E.L. Rose. Forecasting in least absolute value regression with autocorrelated errors: A small-sample study. *International Journal of Forecasting*, 10:539–547, 1994.

[43] Ian Domowitz and Halbert White. Misspecified models with dependent observations. *Journal of Econometrics*, 20:35–58, 1982.

[44] J. L. Doob. *Stochastic Processes.* John Wiley and Sons, New York, 1953.

[45] P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, 84:313–342, 1999.

[46] P. Doukhan and F. Portal. Moments de variables aléatoires mélangeantes. *C.R. Acad. Sc. Paris*, 297:129–132, 1983.

[47] R.F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007, 1982.

[48] R.F. Engle and C.W.J Granger. Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55:251–76, 1987.

[49] R.C. Fair and R.J. Shiller. Comparing information in forecasts from econometric models. *American Economic Review*, 80:375–389, 1990.

[50] M. Forni and L. Reichlin. Let's get real: A dynamic factor analytical approach to disaggregated business cycles. Technical report, University of Modena and University of Bruxelles, 1997.

[51] Philip Franklin. *A Treatise on Advanced Calculus*. John Wiley and Sons, New York, 1940.

[52] R. Frisch. Propagation problems and impulse problems in dynamic economics. In *Economic Essays in Honour of Gustav Cassel*. Allen and Unwin, London, 1933.

[53] R. Garcia and P. Perron. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics*, 78:111–125, 1996.

[54] J. Geweke. The dynamic factor analysis of economic time-series models. In D.J. Aigner and A.S. Goldberger, editors, *Latent Variables in Socioeconomic Models*, pages 365–83. North-Holland, Amsterdam, 1977.

[55] L.G. Godfrey. *Misspecification Tests in Econometrics*, volume 16 of *Econometric Society Monographs*. Cambridge University Press, 1988.

[56] Clive W.J. Granger and Paul Newbold. *Forecasting Economic Time Series*. Academic Press, New York, 1977.

[57] C.W.J. Granger. Investigating casual relations by econometric models and cross-spectral methods. *Econometrica*, 37:424–38, 1969.

[58] C.W.J. Granger. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16:121–30, 1981.

[59] C.W.J Granger and P. Newbold. Some comments on the evaluation of economic forecasts. *Applied Economics*, 5:35–47, 1973.

[60] C.W.J. Granger and Y. Teräsvirta. *Modeling Nonlinear Economic Relationships*. Oxford University Press, Oxford, 1993.

[61] William H. Greene. *Econometric Analysis*. Prentice-Hall, New Jersey, 4th edition, 2000.

[62] T. Haavelmo. The probability approach in econometrics. *Econometrica*, 12:1–118, 1944. Supplement.

[63] G. Haberler. Prosperity and depression. Technical report, League of Nations, Geneva, 1937.

[64] James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994.

[65] J.D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–84, 1989.

[66] B.E. Hansen. Tests for parameter instability in regressions with i(1) processes. *Journal of Business and Economic Statistics*, 10:321–335, 1992.

[67] A.C. Harvey. *The Econometric Analysis of Time Series*. Philip Allan, Deddington, 1981.

[68] A.C. Harvey. *Time Series Models*. Philip Allan, London, 1981.

[69] D.F. Hendry, A.R. Pagan, and J.D. Sargan. Dynamic specification. In Z. Griliches and M.D. Intriligator, editors, *Handbook of Econometrics*, volume 2, chapter 18. North-Holland, Amsterdam, 1984.

[70] D.F. Hendry and J.F. Richard. On the formulation of empirical models in dynamic econometrics. *Journal of Econometrics*, 20:3–33, 1982.

[71] Walter Hlawitschka. The empirical nature of taylor-series approximations to expected utility. *The American Economic Review*, 84(3):713–719, 1994.

[72] R.R. Hocking. Misspecification in regression. *The American Statistician*, 28(1):39–40, 1974.

[73] R.R. Hocking. The analysis and selection of variables in linear regression. *Biometrics*, 32:1–49, 1976.

[74] P. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, 1967. University of California Press.

[75] Ken Hung and Frank B. Alt. The effect of misspecificaion in vector autoregressive-moving average models on parameter estimation and forecasting. *Commun. Statist. Simula*, 18(2):467–479, 1989.

[76] J. Hurt. Asymptotic expansions of functions of statistics. *Aplikace Matematiky*, 21:444–456, 1976.

[77] I.A. Ibragimov and Yu. A. Linnik. *Independent and stationary sequences of random variables*. Groningen Wolters-Noordhoff, 1971.

[78] C. Inclan and G.C. Tioa. Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89:913–923, 1994.

[79] W. S. Jevons. *Investigations in Currecy and Finance*. Macmillan, London, 1884.

[80] K. Kadiyala. Mixed regression estimator under misspecification. *Economics Letters*, 21:27–30, 1986.

[81] Jean-Pierre Kahane. *Some random series of functions*. Cambridge Univeristy Press, New York, 1985.

[82] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82:35–45, 1960.

[83] P. Kennedy. *A Guide to Econometrics*. The MIT Press, Cambridge, MA, 1992.

[84] Rasula A. Khan. Approximation for the expectation of a function of the sample mean. *Statistics*, 38(2):117–122, April 2004.

[85] Lawrence R. Klein. *Economic Fluctuations in the United States 1921-1941*. Number 11 in Cowles Commission Monograph. John Wiley, New York, 1950.

[86] Konrad Knopp. *Infinite Sequences and Series*. Dover, New York, 1956.

[87] Konrad Knopp. *Theory and application of infinite series*. Dover, New York, 1990.

[88] A. Kolmogorov. Sur l'interpolation et l'extrapolation des suites stationnaires. *C.R. Acad. Sci. Paris*, 208:2043–2045, 1939.

[89] A. Kolmogorov. Interpolation und extrapolation von stationären zufälligen folgen. *Bull. Acad. Sci. (Nauk) U.R.S.S., Ser. Math.*, 5:3–14, 1941.

[90] L. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.

[91] Naoto Kunitomo and Taku Yamamoto. Properties of predictors in misspecified autoregressive time series models. *Journal of the American Statistical Association*, 80(392):941–950, 1985.

[92] E.L. Lehmann. *Theory of Point Estimation*. John Wiley and Sons, New York, 1983.

[93] H. Levy and H. M. Markowitz. Approximating expected utility by a function of mean and variance. *The American Economic Review*, 69(3):308–317, 1979.

[94] M. Loève. *Probability Theory I*, volume I. Springer Verlag, New York, 4 edition, 1977.

[95] Otto Loistl. The erroneous approximation of expected utility by means of a taylor's series approximation: Analytic and computational results. *The American Economic Review*, 66(5):904–910, 1976.

[96] S. Louhichi. Moment inequalities for sums of certain dependent random variables. *Theory Probab Appl*, 47:649–664, 2003.

[97] R.E. Lucas. Econometric policy evaluation: A critique. In K. Brunner and A. Meltzer, editors, *The Phillips Curve and Labor Markets*, volume 1 of *Carnegie-Rochester Conferences on Public Policy*, pages 19–46, Amsterdam, 1976. North-Holland.

[98] Esfandiar Maasoumi. How to live with misspecification if you must. *Journal of Econometrics*, 44:67–86, 1990.

[99] Arthur W. Marget. Morgenstern on the methodology of economic forecasting. *The Journal of Political Economy*, 37(3):312–339, June 1929.

[100] D.L. McLeish. A maximal inequality and dependent strong laws. *Anals of Probability*, 3:828–836, 1975.

[101] H.L. Moore. *Economic Cycles - Their Law and Cause*. Macmillan, New York, 1914.

[102] H.L. Moore. *Generating Economic Cycles*. Macmillan, New York, 1923.

[103] Mary S. Morgan. *The history of econometric ideas*. Cambridge University Press, New York, 1990.

[104] O. Morgenstern. *Wirtschaftsprognose: eine Untersuchung ihrer Voraussetzungen und Moglichkeiten*. Julius Springer, Vienna, 1928.

[105] T. Murai. The value-distribution of random taylor series in the unit disk. *J. London Math. Soc.*, 24:480–494, 1981.

[106] C.R. Nelson. The prediction performance of the frb-mit-penn model of the u.s. economy. *American Economic Review*, 62:902–917, 1972.

[107] Gary W. Oehlert. A note of the delta method. *The American Statistician*, 46(1):27–29, 1992.

[108] K. Ohtani and Y. Honda. On small sample properties of the mixed regression predictor under misspecification. *Commun. Statist. Theor. Meth.*, 13(22):2817–2825, 1984.

[109] M. Peligrad. Invariance principles for mixing sequences of random variables. *Ann. Probab.*, 10:968–981, 1982.

[110] M. Peligrad. On the central limit theorem for ρ-mixing sequences of random variables. *Ann. Probab.*, 15:1387–1394, 1987.

[111] W.M. Persons. The problem of business forecasting. In *The Problem of Business Forecasting*. Pollak Foundation for Economic Research Publications, London, pitman edition, 1924.

[112] W.M. Persons. Some fundamental concepts of statistics. *Journal of the American Statistical Association*, 19:1–8, 1924.

[113] M. Hashem Pesaran and Allan Timmermann. Model instability and choice of observation window. Technical Report 99-19, University of California, San Diego, 1999.

[114] M. Hashem Pesaran and Allan Timmermann. Small sample properties of forecasts from autoregressive models under structural breaks. *Journal of Econometrics*, 129:183–217, 2005.

[115] V.V. Petrov. *Limit Theorems of Probability Theory.* Oxford University Press, New York, 1995.

[116] W. Ploberger, W. Kramer, and K. Kontrus. A new test for structural stability in the linear regression model. *Journal of Econometrics*, 40:307–318, 1989.

[117] Ser-Huang Poon and Clive W.J. Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, XLI:478–539, 2003.

[118] D. Quah and T.J. Sargent. A dynamic index model for large cross sections. In J.H. Stock and M.W. Watson, editors, *Business Cycles, Indicators and Forecasting*, pages 285–310. University of Chicago Press for NBER, Chicago, 1993.

[119] C.R. Rao. *Linear Statistical Inference and Its Applications.* Wiley, New York, 1973.

[120] P. Rao. Some notes on misspecification in multiple regressions. *The American Statistician*, 25(5):37–39, 1971.

[121] P. Rao. Some notes on the errors-in-variables model. *The American Statistician*, 27(5):217–218, 1973.

[122] E. Rio. In'egalit'es de moments pour les suites stationnaries et fortement m'elangeantes. *C.R. Acad. Sci. Paris*, 318:355–360, 1994.

[123] D.H. Roberson. *A Study of Industrial Fluctuations*. P.S. King and Son, London, 1915.

[124] T.J. Robinson and J.B. Birch. Model misspecification in parametric dua modeling. *J. Statist. Comput. Simul.*, 66:113–126, 2000.

[125] P. A. Samuelson. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Review of Economic Studies*, 37(112):537–542, 1970.

[126] I. Sánchez and D. Pena. Properties of predictors in overdifferenced nearly nonstationary autoregression. *Journal of Time Series Analysis*, 22(1):45–66, 2001.

[127] T.J. Sargent and C. Sims. Business cycle modeling without pretending to have too much prior theory. In C. Sims, editor, *New Methods of Business Cycle Research*. Federal Reserve Bank of Minneapolis, Minneapolis, 1977.

[128] Q.M. Shao. *Limit theorems for sums of dependent and independent random variables*. PhD thesis, Univ. Science and Thechnology of China, 1989.

[129] Q.M. Shao. On the invariance principle for ρ-mixing sequences of random variables. *Chinese Ann. Math.*, 10B:427–433, 1989.

[130] C.A. Sims. Money, income and casuality. *American Economic Review*, 62:540–52, 1972.

[131] C.A. Sims. Macroeconomics and reality. *Econometrica*, 48:1–48, 1980.

[132] E. Slutsky. The summation of random causes as the source of cyclic processes. *Econometrica*, 5:105–46, 1927.

[133] V. I. Smirnov. *A Course of Higher Mathematics*, volume 1. Pergamon, 1964.

[134] J.H. Stock and M.W. Watson. New indexes of coincident and leading economic indicators. In O. Blanchard and S. Fischer, editors, *NBER Macroeconomics Annual*, pages 351–94. MIT Press, Cambridge Mass., 1989.

[135] J.H. Stock and M.W. Watson. Adaptive diffusion indexes. Technical report, Kennedy School, Harvard University, and Woodrow Wilson School, Princeton University, 1997.

[136] W.F. Stout. *Almost Sure Convergence*. Academic Press, New York, 1974.

[137] H. Theil. *Economic Forecasts and Policy*. North-Holland, Amsterdam, 2nd edition, 1961.

[138] H. Theil. *Applied Economic Forecasting*. North-Holland, Amsterdam, 1966.

[139] A. Timmermann. Structural breaks, incomplete information and stock prices. *Journal of Business and Economic Statistics*, 19:299–314, 2001.

[140] J. Tinbergen. An econometric approach to business cycle problems. Hermann and Cie, Paris, 1937.

[141] J. Tinbergen. Statistical testing of business-cycle theories, vol. i a method and its application to investment activity; vol ii: Business cycles in the united states of america, 1919-1932. Technical report, League of Nations, Geneva, 1939.

[142] Howell Tong. *Non-linear Time Series: A Dynamical Systems Approach*. Clarendon Press Oxford, New York, 1990.

[143] S. C. Tsiang. The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *American Economic Review*, 62(3):354–371, 72.

[144] William J. Vetter. Matrix calculus operations and taylor expansions. *SIAM Review*, 15(2):352–369, 1973.

[145] E. Walker and R.G. O'Brien. Using restricted least squares to delineate the effects of misspecification in linear models. *The Statistician*, 41:467–476, 1992.

[146] T.D. Wallace. Efficiencies for stepwise regressions. *Journal of the American Statistical Association*, 59(308):1179–1182, 1964.

[147] Kenneth F. Wallis. Macroeconomic forecasting: A survey. *The Economic Journal*, 99:28–61, 1989.

[148] Alexander Weinmann. *Uncertain Models and Robust Control*. Springer-Verlag, New York, 1991.

[149] H. White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50(1):1–26, 1982.

[150] H. White. Using least squares to approximate unknown regression functions. *International Economic Review*, 21(1):149–170, 1982.

[151] H. White and I. Domowitz. Nonlinear regression with dependent observations. *Econometrica*, 52:143–162, 1984.

[152] Halbert White. *Estimation, Inference and Specification Analysis*. Economic Society Monographs. Cambridge University Press, New York, 1996.

[153] Halbert White. *Asymptotic Theory for Econometricians*. Academic Press, San Diego, revised edition, 2001.

[154] P. Whittle. *Prediction and Regulation by Linear Least Square Methods*. D. Van Nostrand, Princeton, 1963.

[155] Peter Whittle. *Prediction and Regulation by Linear Least-Squares Methods*. University of Minnesota Press, Minnesota, 2nd edition, 1983.

[156] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. Wiley, New York, 1949.

[157] P. Wijekoon and G. Trenkler. Mean square error matrix superiority of estimators under linear restrictions and misspecification. *Economics Letters*, 30:141–149, 1989.

[158] H. Wold. *The Analysis of Stationary Time Series*. Almquist and Wicksell, Uppsala, 1938.

[159] Ryozo Yokoyama. Moment bounds for stationary mixing sequences. *Z. Wahrsch. verw. Gebiete*, 52:45–57, 1980.

[160] K. Yoshihara. Moment inequalities for mixing sequences. *Kodai Math J.*, 1:316–328, 1978.

[161] G.U. Yule. On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions*, 226A, 1927.

[162] Lin Zhengyan and Lu Chuanrong. *Limit Theory for Mixing Dependent Random Variables*. Science Press/Kluwer Academic Publishers, 1996.