Essays in Applied Economics: New Techniques in Aggregate Data Analysis

Thesis by

Michael Alan Udell

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1995

(Submitted May 18, 1995)

ACKNOWLEDGEMENTS

Financial support for the research and writing of this dissertation was provided by the John Randolph Haynes and Dora Haynes Foundation, the Anna and James McDonnell Memorial Fellowship, and the Graduate Research Assistantship awarded by the California Institute of Technology.

Jeff Dubin, Michael Graetz, and Louis Wilde helped immeasureably with the first chapter of this dissertation. In particular, none of the models presented in this chapter, or the companion research in chapter 2 of this dissertation, could have been derived or estimated without Jeff's help. A version of this chapter appeared in *The Review of Economics and Statistics*, 1992.

Rod Kiewiet was instrumental in encouraging the research on the political business cycle reported in the third chapter. It was his class in applied political science that stimulated the idea to use new historical estimates of GNP and unemployment to test the relation between general economic conditions and congressional election outcomes. Much of the text has been improved by his comments and direction.

Many people at Caltech have also contributed to this effort. In particular, Jeff Dubin, Rod Kiewiet, and John Ledyard exhibited great patience as this dissertation lurched, stopped, and twisted its way to completion in the three years since I left Caltech. Several students at Caltech also helped by providing a supportive and stimulating environment in which to learn and do research. In particular Guofu Tan, Bob Cull, Langche Zeng, Charles Noussair, Jeff Prisbrey and Gretchen Kalsow deserve mention. Along with my peers at Caltech, many persons affiliated with the Institute helped by providing a sense of community for me and my family during the four years that we were in residence. I would also like to thank Bernie Schmitt, my supervisor at the Joint Committee on Taxation, for providing a supportive environment in which to complete this research.

Both of my families endured this process with me. My parents deserve mention because they worried sufficiently so about my completion of this task that I felt relieved of the need to worry that I might not complete it. Most of all, I would like to thank my wife, Linda. Without her love, tireless support and encouragement this undertaking would never have happened. And finally, my two daughters, Erin and Lillian, who make it all worth while.

Dedication

To Linda, Erin, and Lillian

ABSTRACT

This dissertation consists of three essays in applied economics. Common to each essay is the use of aggregate data. The first two essays address the demand for tax return preparation services and the effect of that demand on tax evasion. The third essay is an analysis of the effect of general economic conditions on congressional House elections.

In the first chapter, I analyze taxpayer choices of return preparation services. In particular, I distinguish between two types of nonpaid preparation, six types of paid third parties, and self-preparation. Among other things, I find significant differences in the factors which explain the demand for paid third parties who are and are not able to represent clients before the IRS. Among these factors are increases in IRS audit rates, the frequency of IRS penalties, and the complexity of the tax return.

The second chapter builds upon the results of the first chapter and analyzes the effect that different modes of tax return preparation have on tax evasion. Specifically, I allow for three assisted modes of return preparation and self-preparation to effect the level of tax evasion detected on returns they prepared. Chief among my findings is that while returns prepared by third parties who are able to represent their clients before the IRS are characterized by the greatest amounts of income and the most complex tax situations, after controlling for these factors, these returns exhibit less non-compliance than those prepared by the other modes of return preparation.

The third chapter addresses a long standing controversy as to whether general

economic conditions effect the outcome of congressional House elections. I test this general hypothesis by applying new estimates of historical gross national product and unemployment, as well as robust regression techniques useful for the identification of influential observations, to a variety of models in the political science literature that address the outcomes of congressional House elections. The use of new data strongly supports the proposition that changes in gross national product and unemployment have a significant impact on the congressional House election outcomes.

Contents

List of Figures

List of Tables

Chapter 1

The Demand for Tax Return Preparation Services

1.1 Introduction

The IRS has estimated that 44.2 percent of the individual returns filed in 1979 were self prepared, and that these returns accounted for 22.8 percent of detected noncompliance.[1] Returns prepared with third party assistance accounted for 55.8 percent of filings and 77.2 percent of the detected noncompliance. Among returns prepared with third party assistance, however, only 16.7 percent used what we call here a "tax practitioner" (public accountant or attorney), yet these returns accounted for 42.7 percent of underreported tax, 28.5 percent used other paid preparers accounting

[1]This chapter was published in *The Review of Economics and Statistics*, pages 75 to 82, 1992, with coauthors Jeffrey Dubin, Michael Graetz, and Louis Wilde.

for 30.8 percent of underreported tax, and the remaining 10.6 percent used nonpaid assistance accounting for 3.7 percent of underreported tax.[2] A detailed breakdown of the distribution of preparation modes for 1979 and the associated proportions of tax noncompliance are reported below in Table 1.

Economic theory has only recently focused on the demand for third party assistance and its role in tax compliance. Scotchmer (1989), emphasizing the *informational* aspects of third party assistance, suggests that if risk neutral taxpayers are not sure of their true taxable income, they will tend to overreport their tax liability because the cost of underreporting exceeds the cost of overreporting due to IRS penalties. In her models, taxpayers hire the services of informed third parties who reduce or eliminate uncertainty, and thereby lower the amount of reported tax liability.

Reinganum and Wilde (1991) analyze a game theoretic model of taxpayer, tax practitioner, and IRS behavior that focuses on the *service* aspects of third party assistance. They characterize four types of equilibria, depending on whether taxpayers prefer to use tax practitioners and whether the IRS prefers them to use tax practitioners. When tax practitioner penalties for noncompliance are sufficiently low and the efficiency gains from using practitioners are sufficiently high, the IRS prefers taxpayers to prepare their own returns, but taxpayers prefer to use a tax practitioner. In this case, the use of a tax practitioner is associated with lower compliance and higher

[2] CPAs and attorneys are automatically able to represent clients before the IRS. Other paid preparers may provide these services only after passing a written exam sponsored by the IRS, and fulfilling certain continuing education requirements. Because most public accountants meet these requirements, we include them in the tax practitioner category.

Table 1.1: Returns and Proportion of Non-Compliance

Mode of Preparation	Total Returns Filed Number	Frequency	Proportion of Noncompliance
Self	39959000	.442	.228
Non-Paid			
IRS assisted[a]	1265686	.014	.004
Other Non-Paid[b]	8317320	.092	.033
Paid Preparers			
National Tax Service	8588600	.095	.054
Local Tax Service	10488000	.116	.132
Other Paid	6690000	.074	.122
Practitioners			
Public Accountant[c]	5605200	.062	.102
CPA[d]	6057200	.067	.258
Attorney and CPA	3435450	.038	.067
Total	90406000	1.000	1.000

[a]The IRS Assistance category consists of three services. They are IRS Advice, with a population frequency of .011; IRS Prepared, with a population frequency of .001; and IRS Reviewed, with a population frequency of .004.

[b]The other nonpaid category usually consists of a family member who helps prepare the return. We have combined in this category all VITA prepared returns. VITA is an acronym for Volunteer Income Tax Assistance, which consists of unpaid volunteers who prepare returns after receiving limited instruction, typically from the IRS.

[c]Public Accountants are licensed at the state level with requirements varying by state. Only four states, North Carolina, Virginia, Kansas, and Wyoming, do not regulate Public Accountants. National Tax Services, which in 1979 consisted entirely of H & R Block and Beneficial Financial Services Co., and Local Tax Services often provide their own training, but do not require employees to be Public Accountants.

[d]CPA is an abbreviation for Certified Public Accountant.

audit rates.

Finally, Mazur and Nagin (1987), emphasizing the *strategic* aspects of third party assistance, observe that many paid preparers are "not just passive scribes whose function is limited to relieving their clients of the mechanical requirements of return preparation. To varying degrees they provide information on legal requirements and the penalties for their breach, develop strategies for reducing tax liability, provide counsel on the risks of executing such strategies and inform clients of topical enforcement priorities."

Most empirical analyses of the demand for third party assistance have estimated the effects of types of income, other return characteristics, and various socio-economic factors on the choice between self preparation and all forms of third party assistance. Using a variety of data sources including surveys (Slemrod and Sorum, 1984; Collins, Milliron and Toy, 1988; Hite, 1987), TCMP data (Dubin, Graetz, and Wilde, 1989; Erard, 1993), and the IRS Master File (Long and Caudill, 1987), researchers have found that the demand for third party assistance increases with amounts of total or complex income, age, return complexity, tax rates, the number of dependents, and self employment, and decreases with the level of education.

A complete characterization of the impact of third parties on tax compliance would include analyses of taxpayer choices of the kind of third party assistance, the effect of third parties on tax compliance, and the IRS posture towards returns completed by third parties. We begin this expanded characterization by analyzing taxpayers' choices among 12 modes of third party assistance, using nested logit techniques applied to aggregate data from the Special Academic Research File of the 1979 Individual Return Taxpayer Compliance Measurement Program (TCMP). Although the best data on third party assistance released to date by the IRS, this data file limits the analysis to 696 aggregated observations from 12 IRS audit classes and 58 IRS districts.[3]

The paper is organized as follows. Section 2 presents an econometric specification of the demand for tax preparation services. Section 3 describes the data we use and models we estimate. Section 4 presents our results and conclusions.

1.2 A Model of Paid Preparer Selection

We group types of return preparers according to natural patterns of substitutability. One group consists of CPA's, attorneys, and public accountants, which we label "practitioners," most of whom are licensed to represent taxpayers before the IRS. The second group consists of national tax services, local tax services, and other paid preparers who generally are not able to represent their clients before the IRS. The third group consists of third party assistors who are free of charge. A final group consists of self-preparers. We account for this pattern of substitutability in our estimation by using a two-stage nested multinomial logit model.[4]

[3] For precise definitions of IRS audits classes as well as recent estimates of overall noncompliance see IRS (1990).

[4] For a detailed discussion of the nested logit model, see McFadden (1978).

First we estimate taxpayer choices among the four broad categories: practitioners, paid preparers, nonpaid preparers, and self preparation. Then we estimate taxpayers' choices between the modes of assistance within each category. Let $i = 1, \ldots, I$ index the preparation category, and let $j = 1, 2, \ldots, J_i$ index the specific assistance modes within category i. We assume that the probability that an individual chooses alternative ij, P_{ij}, can be written as the product of the conditional probability $P_{j|i}$ and marginal probability P_i with each probability in the multinomial logit form

$$P_{j|i} = \frac{e^{\beta c'_{j|i} X}}{\sum_{j=1}^{J_i} e^{\beta'_{j|i} X}} \tag{1.1}$$

$$P_i = \frac{e^{\alpha'_i Y + \theta_i I_i}}{\sum_{i=1}^{J_i} e^{\alpha'_i Y + \theta_i I_i}}. \tag{1.2}$$

Here X and Y are vectors of observed attributes which vary by audit class and IRS district and I_i is the expected maximum utility (or inclusive value) a taxpayer derives from alternatives in the i^{th} category with

$$I_i = \log \sum_{j=1}^{J_i} e^{\beta'_{j|i} X}. \tag{1.3}$$

The unknown parameters in this model are $\beta_{j|i}$, α_i, and θ_i with the latter being a measure of the dissimilarity of alternatives in the i^{th} category.[5]

[5] McFadden (1978) proves that if θ_i lies in the closed interval $[0, 1]$, the resultant nested logit is consistent with random utility maximization.

Since our data is aggregated to the IRS district level, we estimate $P_{j|i}$ and P_i using an aggregate form of the nested logit model.[6] Specifically, we form the log odds for equations (1) and (2) as

$$\log\left[\frac{P_{j|i}}{P_{1|i}}\right] = \beta'_{j|i}X \qquad \forall j \neq 1, \tag{1.4}$$

and

$$\log\left[\frac{P_i}{P_1}\right] = \alpha'_i Y + \theta_i(I_i - I_1) \qquad \forall i \neq 1. \tag{1.5}$$

Estimation of equations (4) and (5) requires estimates of the choice probabilities for which we use the observed selection frequencies and measures of X and Y for which we use group weighted averages. Denote the probability that an individual in the k^{th} IRS district chooses alternative ij as P_{ij}^k. Let the number of cases in the k^{th} IRS district be N^k and denote the frequency of occurrences of alternative ij in district k as F_{ij}^k. Suppressing the superscript k we can write

$$\log\left[\frac{F_{j|i}}{F_{1|i}}\right] = \beta'_{j|i}X + \mu_{j|i}, \tag{1.6}$$

and

$$\log\left[\frac{F_i}{F_1}\right] = \alpha'_i Y + \theta_i(I_i - I_1) + \mu_i, \tag{1.7}$$

[6]The estimation of nested logit models has to date relied exclusively on individual level data. We extend the nested logit method for use with aggregate data following Berkson (1944) and Theil (1969). Implicit in this development is the assumption of homogeneity of individuals within the aggregation classes. For a discussion of the robustness properties of logit techniques when applied to aggregate data, see also Allenby and Rossi (1991).

with $\mu_{j|i} = \log\left[F_{j|i}/F_{1|i}\right] - \log\left[P_{j|i}/P_{1|i}\right]$ and $\mu_i = \log\left[F_i/F_1\right] - \log\left[P_i/P_1\right]$. [7]

Straightforward calculations show that

$$
\begin{aligned}
E(\mu_{j|i}) &= 0 \\
V(\mu_{j|i}) &= \frac{1}{N}\left(\frac{1}{P_{j|i}} + \frac{1}{P_{1|i}}\right) \quad \text{for } j \neq 1, && (1.8) \\
\mathrm{cov}(\mu_{j|i}, \mu_{1|i}) &= \frac{1}{NP_{1|i}} \quad \text{for } j \neq 1, && (1.9)
\end{aligned}
$$

and that

$$
\begin{aligned}
E(\mu_i) &= 0, \\
V(\mu_i) &= \frac{1}{N}\left(\frac{1}{P_j} + \frac{1}{P_1}\right) \quad \text{for } j \neq 1, && (1.10) \\
\mathrm{cov}(\mu_j, \mu_1) &= \frac{1}{NP_1} \quad \text{for } j \neq 1. && (1.11)
\end{aligned}
$$

Since P_{ij} varies with the alternative and N varies with the size of each audit class and IRS district, equations (8) and (10) present a classic form of heteroscedasticity. Furthermore, a nonzero covariance is also present (equations (9) and (11)) since the log odds equations (6) and (7) use common comparison groups. Minimum chi-square estimation of the nested logit model corrects simultaneously for this covariance structure.[8]

[7] Our data set has an average of 70 individuals per cell. Monte Carlo evidence reported in Domencich and McFadden (1975) shows that when cell sizes are large (over 30 observations) the Berkson-Theil estimator of equations (6) and (7) has very small sample bias (see Table 5.1, page 113).

[8] The appendix discusses the derivation and estimation of the model presented above as well as

1.3 Model Specification

The TCMP program involves line-by-line audits of approximately 50,000 randomly chosen individual tax returns. Our data file aggregates the results of the 1979 TCMP audits by IRS district and audit class. Six audit classes comprise returns where income was not derived principally from farm or sole-proprietorship activities, three others where income was derived principally from sole-proprietorships, and a final three where income was derived principally from farm activities. For every line for each audit class in each of the districts, taxpayers' reported amounts are recorded as well as the adjusted amounts recommended by the TCMP audit. For the 1979 TCMP, additional information was recorded on the 12 return preparation modes used by taxpayers. To avoid districts with too few observations, we combine all attorneys into a single category and assign the following mnemonics to the resulting practitioner alternatives:

PA	=	Public Accountant
ATT	=	Attorney, or Attorney and CPA
CPA	=	Certified Public Accountant

The three paid-preparer alternatives are:

---

present estimates of equations (6) and (7).

LTS = Local Tax Service

OPP = Other Paid Preparer

NTS = National Tax Service

Finally, we combine several nonpaid assistance alternatives and assign the following mnemonics:

IRS = IRS prepared, IRS reviewed, or IRS assisted

ONP = Volunteer Income Tax Assistance, or Other Nonpaid Preparers

SELF = Self–prepared

As discussed in the Introduction, three basic motives for taxpayers to use third party assistance have been identified: information (Scotchmer); service (Reinganum and Wilde); and strategic (Mazur and Nagin). Our choice of explanatory variables parallels these motives. In particular, we assume that the probability of choosing any mode of third party assistance or self-preparation is a function of the IRS audit rate (AUDIT), the frequency of penalties (PENALTY), the percent of the adult population with a high school education (EDUCATE), the percent of the adult population over 65 years old (AGE65), state, local, and real-estate tax deductions (TAXDED), other deductions (OTHDED), the number of exemptions (EXEMPT), the sum of wage, interest, and dividend income (SIMPLE), the sum of schedule C gross receipts and partnership income (BUSINESS), other income (OTHINC), and the number of forms included with the taxpayer's return (FORMS).

Although we make no attempt to test one against another, our predictions regarding the effects of the explanatory variables on the demand for third party assistance are based on the service, information, and strategic motives. We next summarize our general expectations regarding the effects of the explanatory variables on the demand for third party assistance with occasional reference to specific motives.

Since neither nonpaid nor paid preparers generally can represent clients before the IRS, increases in the IRS enforcement activity variables RATE or PENALTY should increase the demand for practitioners.[9] Because more educated taxpayers are better able to prepare their own returns, represent themselves before the IRS, determine their own true tax liability, and engage in appropriately strategic behavior, increases in EDUCATE should reduce the demand for all modes of third party assistance and increase self-preparation. An increase in AGE65 may possibly increase the demand for preparers and practitioners based on the service and information motives.[10]

Our tax return variables reflect the TCMP adjusted levels for three basic components of the tax calculation: income, deductions, and exemptions.[11] We use three measures of income: the sum of wages, interest, and dividends, the sum of schedule C gross receipts and partnership income, and all other income. Increases in SIMPLE

[9]The audit rate is based on audits accomplished data provided by the Examination Division of the IRS. The frequency of penalties is obtained directly from the aggregated TCMP data file.

[10]The percent of the adult population with a high school education and the percent of the adult population over age 65 are from the Statistical Abstract of the United States for the year 1979.

[11]The TCMP provides not only what the tax return reported amounts were, but what the IRS believed the true amount to be. We use the corrected amounts of deductions and exemptions as meaningful measures of the true amounts of these items.

should decrease the demand for third party assistance and increase self-preparation, since for taxpayers with predominately wage, interest, or dividend income, there is less need for service or information, less opportunity for strategic behavior and little likelihood of IRS enforcement attention. Increases in BUSINESS should increase the demand for practitioners since owners of small businesses do face substantial IRS enforcement attention, considerable complexity, and may as a normal business practice require the services of attorneys or highly qualified accountants. We have no *a priori* expectations with respect to the effect of increases in OTHINC on the demand for third party assistance and self-preparation, but include it in our specification in order to account for all income.

We divide deductions into state, local, and real estate taxes and other deductions. Since taxpayers who pay state and local taxes may be subject to state enforcement attention as well as IRS enforcement attention, increases in TAXDED may increase the demand for practitioners.[12] Unlike state, local, and real estate taxes which are not likely to be contested, other deductions often are contested. This suggests that increases in OTHDED should also increase the demand for practitioners. By our definition, EXEMPT primarily measures family size. To the extent that larger families have greater demand for tax services, an increase in EXEMPT should increase the demand for both preparers and practitioners.[13] Finally, the number of forms is a

[12] See Dubin, Graetz, and Wilde (1989) for a discussion of the relationship between state and federal enforcement activity.

[13] We subtract exemptions for over 65 years of age and blindness from total exemptions. The remainder is a measure of the number of persons supported by the taxpayer.

straightforward measure of return complexity and as such increases in FORMS should increase the demand for both preparers and practitioners.

1.4 Results and Conclusions

The choice of third party assistance is estimated in two stages, as discussed in Section 2 and illustrated in Figure 1 below.[14]

[14]The branch structure of the choice tree in figure 1 is suggested by key features of the choice of return preparation assistance. From the top of the tree downward, there are four general categories of tax return preparation assistance specified. These are Self-prepared returns and the three modes of assisted return preparation. The three modes are Non-paid asistance, comprised of family members or friends and of IRS sponsored assistance; Paid preparers who cannot represent their clients in the case of an IRS audit, comprised of the National Tax Services, Local Tax Services, and Other Paid Preparers; and Practitioners, comnprised of Public Accountants, CPA's, and Attorneys, who can represent their clients in the event of an IRS audit, as well as provide expert opinions for tax positions held on the tax return.

One restrictive property of the nested mulitnomial logit model is known as the independence of irrelevant alternatives (IIA). This property states that the ratio of the probabilities of choosing any two alternatives is independent of the attributes of any other alternative in the choice set. When applied to the nested multinomial logit, this property implies that the IIA property must hold at each multinomial logit along each branch of the choice tree. A violation of this property at any stage of the nested logit model would invalidate the nested multinomial model.

Hausman and McFadden (1984) derive tests for violations of the IIA property in the multinomial logit and the nested multinomial logit models. The first of these, a Hausman (1980) type specification test can be estimated for each multinomial logit by comparing the parameter estimates of two multinomial logits that differ only in the exclusion of a subset of the alternatives. If the parameter estimates are similar, then the test statistic will be small and the IIA property cannot be rejected by the inclusion of the subset of alternatives.

More specifically, let the choice set for the Non-paid$'$ branch be composed of the {IRS, Other} alternatives; for the Paid$'$ branch as {National Tax Service, Local Tax Service, Other Paid Tax Service}, and for the Practitioner branch as {Public Accountants, CPA's, Attorneys} where the prime designates a subset of the choice set. Denote alternative choice sets as Non-Paid {IRS, Other, Self}; Paid {National Tax Service, Local Tax Service, Other Paid, Public Accountants}, and Practitioner$'$ {CPA's, Attorneys}. These alternative choice sets generate a similar two stage model as depicted in figure 1 except that the top stage has three alternatives rather than four because the Self prepared returns are grouped with all other Non-paid alternatives and in the second stage Public Accountants have been moved into Paid preparers who are not Practitioners. This is but one of a large number of alternative configurations for a nine alternative model, but it represents the most plausible alternative construction because it tests for two empirically relevant issues. The first of these is whether there is any difference between returns that are Self-prepared and those that use Non-Paid assistance. The second is whether Public Accountants, many of whom are enrolled agents able to represent clients before the IRS, are more like Practitioners than Paid preparers who cannot

Figure 1.1: Choice of Tax Return Preparation Services

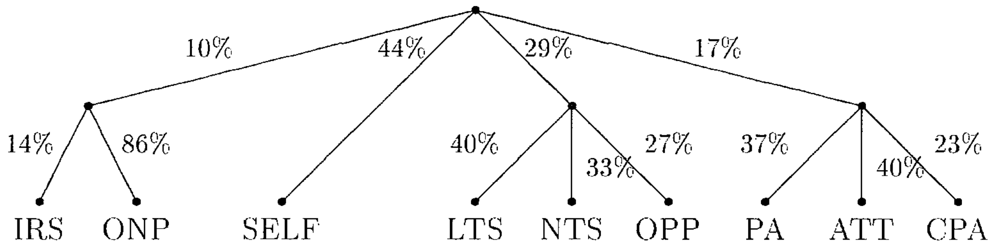

Table 2 presents the elasticities of the unconditional choice probabilities P_{ij} with respect to the underlying explanatory variables.[15]

Our results generally are consistent with our expectations. Increases in the IRS audit rate or frequency of penalties increase significantly the demand for practitioners. For the audit rate this effect is limited to CPA's, but for the frequency of penalties the effect applies to all three types of practitioners. For the audit rate, this increase comes at the expense of national tax services while for the frequency of penalties it comes largely at the expense of self-preparation.

represent their clients before the IRS.

Denote the $k \times 1$ vector of coefficients estimated by equation (1.6) as β and the $k \times k$ estimated covariance matrix in equations (1.8) and (1.9) as COV. Hausman and McFadden (1984) show that the statistic

$$(\beta_{\text{Paid}'} - \beta_{\text{Paid}})'(COV_{\text{Paid}'} - COV_{\text{Paid}})^{-1}(\beta_{\text{Paid}'} - \beta_{\text{Paid}})$$

is asymptotically distributed chi-square with k degrees of freedom under the null hypothesis that the IIA property is valid for the larger choice set. We estimated equation (1.6) for the choice sets Paid and Paid$'$; Non-Paid and Non-Paid$'$; and Practitioner and Practitioner$'$ and constructed the test statistic with 11 degrees of freedom for each pair of choice sets. The values of the test statistic were 38, 26, and 15 respectively, indicating that the choice sets Non-Paid$'$ and Paid$'$ violated the IIA property while the Practitioner$'$ set did not. As a result, we excluded the Self-prepared alternative from the Non-Paid choice set and included the Public Accountant alternative with the Practitioner choice set, as shown in figure 1.

[15]These elasticities are based on sequential estimation of the nested logit model, which includes the estimation of the dissimilarity parameters as discussed in Section 2. These parameters are 0.28, 0.85, and 1.4 for nonpaid preparers, paid preparers, and practitioners, respectively. The last of these is not significantly different from 1. The calculation of the elasticities is complicated by the fact that each explanatory variable has a distinct affect on all alternatives at each level of the nested logit tree. The mean values of all variables and a discussion of their units is available on request.

As predicted, an increase in the percent of the adult population with a high school education increases self-preparation. However, it does not decrease the demand for all forms of third party assistance. In particular, other nonpaid preparers show a statistically significant fall in demand, and the demand for national tax services actually shows a significant increase. An increase in the percent of the adult population over age 65 increases the demand for practitioners and decreases the demand for paid preparers.

Some of our strongest results are obtained with respect to tax return variables. Increases in wages, interest, and dividends significantly decrease the demand for all forms of third party assistance except IRS assistance, and very significantly increase self-preparation. Increases in schedule C gross receipts and partnership income, as expected, increase the demand for practitioners and have no significant effect on any other particular mode of preparation including self-preparation. Finally, an increase in other income, perhaps surprisingly, has no statistically significant effect on any form of third party assistance or self-preparations.

With respect to deductions, an increase in state, local, or real estate taxes significantly increases the demand for practitioners. However, it actually decreases the demand for local and national tax services, an effect which we did not expect. Among all the explanatory variables, the number of exemptions has the strongest effects; increases in it significantly decrease self-preparation and increase the demand for all forms of paid third party assistance. It also significantly decreases the demand for nonpaid preparers. Finally, the number of forms also has a pattern of effects consistent with our expectations; increases in the number of forms increase the demand for practitioners, decrease self-preparation, and have little effect on the demand for other modes of third party assistance.

These results generally confirm the results obtained by other investigators. However, in many cases we find significant differences between the effect of explanatory variables on paid-preparers as compared to practitioners. In particular, many of the results obtained in previous studies which differentiate only between self-preparation and all forms of third party assistance seem to be driven by the demand for practitioners. We also have obtained new results regarding the effect of audit rates and the frequency of penalties on the demand for third party assistance. These results will be particularly important when we turn to the effects of third party assistance on noncompliance. Nevertheless, some of our results must be viewed as preliminary. In particular, while our results with respect to the percent of the adult population over 65 and the number of exemptions conform to our expectations, better data and further investigations are needed fully to understand the relationship between these variables and the demand for third party assistance.

Even at this early stage, we can, however, speculate on the change in noncompliance induced by a shift in return preparation. For example, consider a change in information reporting which has the effect of increasing the amount of simple income relative to other forms of income. In this case, we would predict an increase in self

Table 1.2: Response Elasticities by Mode of Preparation

	Nonpaid Preparers		Paid Preparers		
Mode	IRS	ONP	LTS	OP	NTS
CONSTANT	.577	4.609	-1.664	-.701	-1.614
	(.715)	(8.591)	(-2.189)	(-.890)	(-2.066)
RATE	-.586	.247	.291	-.171	-.734
	(-1.942)	(1.239)	(.849)	(-.493)	(-2.011)
PENALTY	-.176	.033	-.016	.126	.047
	(-.785)	(.277)	(-.274)	(1.723)	(.667)
EDUCATE	-.852	-3.085	.461	-.080	1.088
	(-1.458)	(-8.519)	(1.145)	(-.185)	(2.570)
AGE65	.099	-.530	-.192	-.575	-.250
	(.356)	(-2.967)	(-1.402)	(-3.293)	(-1.601)
TAXDED	.100	.052	-.161	.087	-.180
	(.516)	(.534)	(-2.543)	(1.294)	(-2.547)
OTHDED	-.338	-.039	.134	-.038	-.140
	(-1.511)	(-.332)	(1.732)	(-.445)	(-1.655)
EXEMPT	-.973	-1.534	1.100	1.264	1.372
	(-2.239)	(-6.371)	(6.299)	(6.426)	(6.986)
SIMPLE	-.107	-.392	-.249	-.330	-.294
	(-.539)	(-3.320)	(-4.042)	(-4.458)	(-4.045)
BUSINESS	.856	-.694	.017	.021	.031
	(1.235)	(-1.625)	(.359)	(.406)	(.571)
OTHINC	.235	-.269	-.101	-.112	-.074
	(.462)	(-1.041)	(-1.450)	(-1.347)	(-.822)
FORMS	.077	.208	.492	.266	.152
	(.124)	(.637)	(2.895)	(1.388)	(.780)

Note: t-statistics in parenthesis.

Table 1.3: Response Elasticities by Mode of Preparation, cont'd.

	Practitioners			Self-prepared
Mode	PA	ATT	CPA	Self
CONSTANT	-3.774	-5.568	-3.849	.715
	(-6.232)	(-8.485)	(-6.621)	(4.808)
RATE	.266	.203	.487	-.050
	(.292)	(.947)	(2.407)	(-.417)
PENALTY	400	.281	.529	-.190
	4.728)	(2.887)	(6.529)	(-5.791)
EDUCATE	-.252	.558	-.485	.323
	(-.645)	(1.265)	(-1.308)	(1.934)
AGE65	.623	.787	.232	.106
	(2.641)	(3.022)	(1.026)	(1.605)
TAXDED	.139	.155	.098	-.001
	(3.609)	(3.889)	(2.755)	(-.062)
OTHDED	.041	.265	.466	-.084
	(.356)	(2.227)	(4.058)	(-2.495)
EXEMPT	1.368	1.179	.936	-.931
	(7.415)	(5.432)	(5.487)	(-12.537)
SIMPLE	-.242	-.348	-.157	.348
	(-3.444)	(-4.583)	(-2.420)	(12.026)
BUSINESS	.099	.085	.114	.009
	(2.100)	(1.725)	(2.439)	(.207)
OTHINC	-.058	-.060	-.062	.118
	(-1.043)	(-.987)	(-1.176)	(2.480)
FORMS	.692	1.440	1.199	-.673
	(4.472)	(8.550)	(8.246)	(-7.423)

Note: t-statistics in parenthesis.

preparation. If the level of noncompliance attributable to self prepared returns does not change, then a net reduction in noncompliance would result. However, since the choice of a type of tax return preparation may itself be conditioned on the level of tax evasion, consistent determination of the levels of noncompliance attributable to specific preparer types must take account of the potential for self-selection bias.

References

Allenby, Greg and Rossi, Peter. "There Is No Aggregation Bias: Why Macro Logit Models Work." *Journal of Business and Economic Statistics* 9 (January 1991):1-14.

Berkson, Joseph. "Application of the Logistic Function to Bio-Assay." *Journal of the American Statistical Association* 39 (September 1944):357-365.

Collins, Julie; Milliron, Valerie; and Toy, Daniel. "Tax Preparer Usage and its Implications." Mimeographed. University of North Carolina, 1988.

Domencich, Thomas and McFadden, Daniel. *Urban Travel Demand.* New York: North-Holland Publishing Co., 1975.

Dubin, Jeffrey; Graetz, Michael; and Wilde, Louis. "The Report of the United States to the International Fiscal Association on the Administrative and Compliance Costs of Taxation." In *Cahiers of the International Fiscal Association, 43 Congress.* Netherlands: Kluwer, 1989, pp. 311–344.

Erard, Brian. "Taxation with Representation: An Analysis of the Role of Tax Practitioners in Tax Compliance." *Journal of Public Economics* 52 (1993):163-197.

Hausman, Jerry. "Specification Tests in Econometrics." *Econometrica* 46 (1978):1251-1272.

Hausman, Jerry, and McFadden, Daniel. "Specification Tests For the Multinomial Logit Model." *Econometrica* 52 (1984):1219-1240.

Hite, Peggy. "Characteristics of Taxpayers Who Use Preparers: An Exploratory Study." Paper prepared for IRS Research Conference on The Role of Tax Practitioners in the Tax System, Washington D.C., November 16-17, 1987.

Internal Revenue Service. *Net Tax Gap and Remittance Tax Gap Estimates*, Publication 1415. Washington: Research Division, Compliance Estimates Group, April, 1990.

Long, James and Caudill, Steven. "The Usage and Benefits of Paid Tax Return Preparation." *National Tax Journal* 40 (March 1987):35-46.

McFadden, Daniel. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by A. Karlquist et al. Amsterdam: North-Holland Publishing, 1978.

Mazur, Mark and Nagin, Daniel. "Tax Preparers and Tax Compliance: A Theoretical and Empirical Analysis." Mimeographed. Carnegie Mellon University, November, 1987.

Reinganum, Jennifer and Wilde, Louis. "Equilibrium Enforcement and Compliance in the Presence of Tax Practitioners." *Journal of Law, Economics, and Organization* 7 (Spring 1991):163-181.

Scotchmer, Suzanne. "Who Profits from Taxpayer Confusion?" *Economics Letters* 29, No. 1 (1989):49-55.

Slemrod, Joel and Sorum, Nikki. "The Compliance Cost of the U.S. Individual Income Tax System." *National Tax Journal* 37 (December 1984):461-474.

Theil, Henri. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10 (October 1969):251-259.

1.5 Appendix

1.5.1 Minimum Chi Square Estimation

This section concerns the estimation of the multinomial nested logit with aggregate data when all variables are interacted with alternative specific constants. We first derive this model and show the covariance restrictions implied by its construction. We then compare estimation of the aggregate data univariate nested logit using minimum chi square, maximum likelihood methods, and GLS methods.[16]

Let $i = 1, \ldots, I$ index the preparation category, and let $j = 1, 2, \ldots, J_i$ index the mode, conditional upon using category i, to choose from a set of J_i specific assistance modes. We assume that the taxpayer derives utility U_{ij} from alternative ij where

$$U_{ij} = V_{ij} + \epsilon_{ij}, \tag{1.12}$$

and where V_{ij} is a function of non-stochastic observed attributes, and ϵ_{ij} is a random component of unobserved attributes. The probability that the individual chooses alternative ij is

$$\begin{aligned} P_{ij} &= P[U_{ij} \geq U_{kl}, \quad \forall\, kl \neq ij] \\ &= P[V_{ij} + \epsilon_{ij} \geq V_{kl} + \epsilon_{kl}, \quad \forall\, kl \neq ij] \\ &= P[\epsilon_{kl} - \epsilon_{ij} \leq V_{ij} - V_{kl}, \quad \forall\, kl \neq ij]. \end{aligned} \tag{1.13}$$

[16] In the next section I show that minimum chi square and maximum likelihood estimators are asymptotically equivalent.

The strict utility component is assumed to be of the form

$$V_{ij} = \beta'_{ij} X + \alpha'_i Y \tag{1.14}$$

where X and Y are vectors of mean values of observed attributes varying with each of the 12 audit classes in the 58 IRS districts. They also include alternative specific constants.[17]

We also assume that ϵ_{ij} has a generalized extreme value distribution with cumulative distribution function

$$F(\epsilon;\theta) = exp\left(-\sum_{i=1}^{I}\left(\sum_{j=1}^{J_i} e^{-\epsilon_{ij}/\theta_i}\right)^{\theta_i}\right),$$

where $\epsilon = (\epsilon_{ij})$. Under these assumptions, McFadden (1978) demonstrates that P_{ij} can be written as the product of the conditional probability $P_{j|i}$ and marginal probability P_i where each probability is in the multinomial logit form:

$$P_{j|i} = \frac{e^{\beta'_{j|i}X}}{\sum_{j=1}^{J_i} e^{\beta'_{j|i}X}} \tag{1.15}$$

$$P_i = \frac{e^{\alpha'_i Y + \theta'_i I_i}}{\sum_{i=1}^{J_i} e^{\alpha'_i Y + \theta'_i I_i}}, \tag{1.16}$$

[17]Specification of the strict utility without interactions with alternative specific constants would lead to the universal logit model (McFadden (1975)) which is not consistent with the GEV specification and does not satisfy stochastic utility maximazation. However, it also is not restricted by the IIA assumption.

where I_i denotes the expected maximum utility (or inclusive value) taxpayers derive from alternatives in the i^{th} category,

$$I_i = \log \sum_{j=1}^{J_i} e^{\beta'_{j|i} X}, \tag{1.17}$$

and θ_i is a measure of the dissimilarity of alternatives in the i^{th} category.[18] We estimate a version of this model in which the $j's$ are unique to specific $i's$.

To extend the sequential multinomial logit model to the aggregate multinomial logit model, we first introduce the superscripts k to denote an IRS audit class, and t to denote an IRS district. Let N^{kt} denote the number of cases in the k^{th} IRS audit class in the t^{th}IRS district. The frequency of the ij^{th} case in audit class k in district t is then

$$F_{ij}^{kt} \equiv \frac{1}{N^{kt}} \sum_{h=1}^{N^{kt}} \delta_{hij}^{kt} \tag{1.18}$$

where δ_{hij}^{kt} is an indicator equal to one if the h^{th} person in audit class k in district t chooses alternative ij. Assuming constant probability for alternative ij in audit

[18] McFadden (1978) proves that if the random component of the utility function is specified with the GEV distribution, and if θ is between 0 and 1, the resultant nested logit is consistent with random utility maximization. If $\theta = 1$, then the choice of an alternative is unaffected by the presence or absence of other alternatives, i.e., all alternatives are independent of each other. In this restrictive case, the joint probability P_{ij} is of the multinomial logit form. A value of θ in the interval $(0, 1)$ constitutes rejection of this independence, and the joint probability P_{ij} is of the nested multinomial logit form. In this case, alternatives are independent only within a category. If $\theta = 0$, the nested logit model assesses the choices among alternatives in category i as if there were a single maximal alternative.

class/district kt, it follows that for large enough N^{kt}

$$
\begin{aligned}
E(\delta_{hij}^{kt}) &= P_{ij}^{kt} \\
E(F_{ij}^{kt}) &= \frac{1}{N^{kt}} \sum_{h=1}^{N^{kt}} P_{ij}^{kt} = P_{ij}^{kt}.
\end{aligned} \tag{1.19}
$$

We estimate the log odds of alternative ij relative to alternative $i1$ in each IRS audit class/district as[19]

$$
\log\left[\frac{F_{j|i}^{kt}}{F_{1|i}^{kt}}\right] = \beta_{ij}' X^{kt} + \mu_{j|i}^{kt} \tag{1.20}
$$

$$
\text{for} \quad i = 1, 2, ..., I \quad \text{and} \quad j = 2, 3, ..., J_i
$$

$$
\text{where} \quad \mu_{j|i}^{kt} = \log\left[\frac{F_{j|i}^{kt}}{F_{1|i}^{kt}}\right] - \log\left[\frac{P_{j|i}^{kt}}{P_{1|i}^{kt}}\right]. \tag{1.21}
$$

and for the second stage we estimate the log odds of category i relative to category 1 as

$$
\log\left[\frac{F_{i}^{kt}}{F_{1}^{kt}}\right] = \alpha_i' Y^{kt} + \theta_i'(I_i^{kt} - I_1^{kt}) + \mu_i^{kt} \tag{1.22}
$$

$$
\text{for} \quad i = 1, 2, ..., I
$$

$$
\text{where} \quad \mu_{i}^{kt} = \log\left[\frac{F_{i}^{kt}}{F_{1}^{kt}}\right] - \log\left[\frac{P_{i}^{kt}}{P_{1}^{kt}}\right]. \tag{1.23}
$$

[19] Where $x^{kt} = \frac{1}{N^{kt}} \sum_{h=1}^{N^{kt}} x_h^{kt}$

Rewrite equation (21) as

$$
\begin{aligned}
\mu_{j|i}^{kt} &= \log F_{j|i}^{kt} - \log F_{1|i}^{kt} - \log P_{1|i}^{kt} + \log P_{j|i}^{kt} \\
&= \log F_{j|i}^{kt} - \log P_{j|i}^{kt} - \left(\log F_{1|i}^{kt} - \log P_{1|i}^{kt}\right)
\end{aligned}
$$

Now do first order Taylor's series approximations of $F_{j|i}$ around $P_{j|i}$ and of $F_{1|i}$ around $P_{1|i}$.

$$
\begin{aligned}
\log F_{j|i}^{kt} &\cong \log P_{j|i}^{kt} + (F_{j|i}^{kt} - P_{j|i}^{kt})\left(\frac{1}{P_{j|i}^{kt}}\right) + o(F_{j|i}^{kt} - P_{j|i}^{kt}) \\
\log F_{1|i}^{kt} &\cong \log P_{1|i}^{kt} + (F_{1|i}^{kt} - P_{1|i}^{kt})\left(\frac{1}{P_{1|i}^{kt}}\right) + o(F_{1|i}^{kt} - P_{1|i}^{kt})
\end{aligned}
$$

substituting the above expansions into (21) yields

$$
\mu_{kj} \cong \frac{(F_{j|i}^{kt} - P_{j|i}^{kt})}{P_{j|i}^{kt}} - \frac{(F_{1|i}^{kt} - P_{1|i}^{kt})}{P_{1|i}^{kt}}. \tag{1.24}
$$

Therefore,

$$
E(\mu_{j|i}^{kt}) \cong 0 \tag{1.25}
$$

and

$$
V\left(\mu_{j|i}^{kt}\right) = \frac{V(F_{j|i}^{kt})}{(P_{j|i}^{kt})^2} + \frac{V(F_{1|i}^{kt})}{(P_{1|i}^{kt})^2} - \frac{2\,cov(F_{j|i}^{kt}, F_{1|i}^{kt})}{P_{j|i}^{kt} P_{1|i}^{kt}}
$$

$$
\begin{aligned}
&= \frac{P^{kt}_{j|i}(1-P^{kt}_{j|i})}{(P^{kt}_{j|i})^2 N^{kt}} + \frac{P^{kt}_{1|i}(1-P^{kt}_{1|i})}{(P^{kt}_{1|i})^2 N^{kt}} + \frac{2\,P^{kt}_{j|i}P^{kt}_{1|i}}{P^{kt}_{j|i}P^{kt}_{1|i}N^{kt}} \\
&= \left[\frac{1-P^{kt}_{j|i}}{P^{kt}_{j|i}} + \frac{1-P^{kt}_{1|i}}{P^{kt}_{1|i}} + 2\right]\tfrac{1}{N^{kt}} \\
&= \tfrac{1}{N^{kt}}\left(\frac{1}{P^{kt}_{j|i}} + \frac{1}{P^{kt}_{1|i}}\right) \quad for j \neq 1 \qquad (1.26)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{cov}\,(\mu^{kt}_{j|i}, \mu^{kt}_{l|i}) &= E(\mu^{kt}_{j|i}\mu^{kt}_{l|i}) \quad \text{for} j \neq l \quad j, l \neq 1 \\
&= E\left[\left(\frac{(F^{kt}_{j|i}-P^{kt}_{j|i})}{P^{kt}_{j|i}} - \frac{(F^{kt}_{1|i}-P^{kt}_{1|i})}{P^{kt}_{1|i}}\right)\left(\frac{(F^{kt}_{l|i}-P^{kt}_{l|i})}{P^{kt}_{l|i}} - \frac{(F^{kt}_{1|i}-P^{kt}_{1|i})}{P^{kt}_{1|i}}\right)\right] \\
&= \frac{cov\,(F^{kt}_{j|i}F^{kt}_{l|i})}{P^{kt}_{j|i}P^{kt}_{l|i}} - \frac{cov\,(F^{kt}_{j|i}F^{kt}_{1|i})}{P^{kt}_{j|i}P^{kt}_{1|i}} - \frac{cov\,(F^{kt}_{l|i}F^{kt}_{1|i})}{P^{kt}_{l|i}P^{kt}_{1|i}} + \frac{V\,(F^{kt}_{1|i})}{(P^{kt}_{1|i})^2} \\
&= -\tfrac{1}{N^{kt}} + \tfrac{1}{N^{kt}} + \tfrac{1}{N^{kt}} + \frac{P^{kt}_{1|i}(1-P^{kt}_{1|i})}{(P^{kt}_{1|i})^2 N^{kt}} = \tfrac{1}{N^{kt}}\frac{1}{P^{kt}_{1|i}} \qquad (1.27)
\end{aligned}
$$

The same calculations show that

$$
\begin{aligned}
E(\mu^{kt}_i) &= 0 \qquad (1.28) \\
V(\mu^{kt}_i) &= \frac{1}{N^{kt}}\left(\frac{1}{P^{kt}_i} + \frac{1}{P^{kt}_1}\right) \qquad \text{for } j \neq 1 \qquad (1.29) \\
\mathrm{cov}(\mu^{kt}_j, \mu^{kt}_1) &= \frac{1}{N^{kt}P^{kt}_1} \qquad \text{for } j \neq 1. \qquad (1.30)
\end{aligned}
$$

Note the resulting structure of the system covariance matrix illustrated in figure 2.

Table 1.4: Minimum Chi Square Covariance Matrix

$$
\left[\begin{array}{cc|cc}
\frac{1}{N^{kt}}\left(\frac{1}{P^{kt}_{j|i}}+\frac{1}{P^{kt}_{1|i}}\right) & 0 & \frac{1}{N^{kt}}\left(\frac{1}{P^{kt}_{1|I}}\right) & 0 \\
 & \ddots & & \ddots \\
0 & \frac{1}{N^{KT}}\left(\frac{1}{P^{KT}_{j|i}}+\frac{1}{P^{KT}_{1|i}}\right) & 0 & \frac{1}{N^{KT}}\left(\frac{1}{P^{KT}_{1|I}}\right) \\
\hline
\frac{1}{N^{kt}}\left(\frac{1}{P^{kt}_{1|I}}\right) & 0 & \frac{1}{N^{kt}}\left(\frac{1}{P^{kt}_{J|I}}+\frac{1}{P^{kt}_{1|I}}\right) & 0 \\
 & \ddots & & \ddots \\
0 & \frac{1}{N^{KT}}\left(\frac{1}{P^{KT}_{1|I}}\right) & 0 & \frac{1}{N^{KT}}\left(\frac{1}{P^{KT}_{J|I}}+\frac{1}{P^{KT}_{1|I}}\right)
\end{array}\right]
$$

Where $j = 2, 3, \ldots, l, \ldots, J$ and $i = 1, 2, \ldots, I$. For each equation in the system, the variance is a diagonal matrix whose elements are given by equation (26). The off diagonal elements of the variance matrices are zero by the assumed independence of the observations. The system covariance terms are the off-diagonal matrices and are given by equation (27). The off-diagonal terms for the covariance matrices are zero as a result of the assumed independence of alternatives. [20] This convenient structure results from the assumption of independence between observations and alternatives. This system covariance structure allows for minimum chi square estimation of the multinomial logit, correcting simultaneously for two sources of heteroskedasticity. From (19), since $P^{kt}_{j|i}$ changes with each alternative, the variance is heteroskedastic. Further, the term N^{kt} in the denominator is not constant, but varies with the size of the IRS district and the audit class adding an additional source of heteroskedasticity.

[20]This feature of the the covariance matrices is due to the IIA property of the MNL. The covariance terms result from the construction of the reduced form equations and represent the correlation of the alternatives with the comparison group.

Additionally, correlation with the comparison group is controlled by the covariance terms.

The system covariance matrix may easily be calculated in the following manner. Let i be a vector of ones whose length equals $J-1$. For a binary logit this is 1, for a trinary logit it is 2, etc... Denote the identity matrix whose rank equals $J-1$ as I. Let the comparison group be denoted by the subscript 1, and the remaining alternatives as $j = 2, 3, ..., J$. Let N denote a vector whose length is KT and whose elements are the number of cases in each aggregation. Here, an aggregation is an IRS district/audit class pair. The number of cases is the number of returns in that district and audit class pair. Let P_1 be a vector of length KT of the probability of choosing the comparison group, and finally, create a vector P_j of length KT of probabilities of choosing the j^{th} alternative. This last vector is to be "stacked" so that for each alternative to be compared, the vector has the choice probabilities arranged $1, 2, \ldots, KT$, for $J-1$ alternatives; e.g., For the binary logit $\frac{1}{P_{2|i}}$ is a length KT vector; for the trinary logit

$$\begin{bmatrix} \frac{1}{P_{2|i}^{11}} \\ \vdots \\ \frac{1}{P_{2|i}^{KT}} \\ \frac{1}{P_{3|i}^{11}} \\ \vdots \\ \frac{1}{P_{3|i}^{KT}} \end{bmatrix}$$

is of length $2KT$. Then the system covariance matrix is

$$V(\mu) = (\frac{1}{N} \otimes I) \left[(\text{diag}\frac{1}{P_1} \otimes ii') + \text{diag}(\frac{1}{P_j}\text{for j=2,3,...J}) \right] = \Sigma. \tag{1.31}$$

The Aitken's estimator for this model is

$$\hat{\beta} = (X'(\Sigma^{-1})X)^{-1}(X'(\Sigma^{-1})Y) \tag{1.32}$$

where

$$X = \begin{bmatrix} X_2 - X_1 \\ X_3 - X_1 \\ \vdots \\ X_J - X_1 \end{bmatrix}$$

$$Y = \begin{bmatrix} \log\left(\frac{F_2}{F_1}\right) \\ \log\left(\frac{F_3}{F_1}\right) \\ \vdots \\ \log\left(\frac{F_J}{F_1}\right) \end{bmatrix}.$$

1.5.2 Asymptotic Equivalence of Minimum Chi Square and Maximum Likelihood Estimators

To show the asymptotic equivalence between minimum chi square estimators and maximum likelihood estimators for models with multinomial distributions let

$$\begin{aligned} N &= \text{total number of responses} \\ x_i &= \text{number of responses in the } i^{th} \text{ cell} \\ P_i &= \text{probability of response in } i^{th} \text{ cell } \; i = 1, ..., K \end{aligned}$$

We specify the cell probabilities as a function of the set of parameters $\beta = (\beta_1, \beta_2, ..., \beta_s)$ where $s < K$, and choose values of these parameters to minimize Pearson's chi square

$$\chi^2 = \sum_i \frac{(x_i - NP_i(\beta))^2}{NP_i(\beta)} \tag{1.33}$$

Differentiating with respect to β yields

$$\begin{aligned} \frac{\partial \chi^2}{\partial \beta} &= \left[-2N \sum_i \frac{x_i - NP_i(\beta)}{NP_i(\beta)} - \sum_i \frac{(x_i - NP_i(\beta))^2}{NP_i^2(\beta)} \right] \frac{\partial P_i(\beta)}{\partial \beta} \\ -\frac{1}{2} \frac{\partial \chi^2}{\partial \beta} &= \sum_i \left[\frac{x_i - NP_i(\beta)}{P_i(\beta)} + \frac{(x_i - NP_i(\beta))^2}{2NP_i(\beta)} \right] \frac{\partial P_i(\beta)}{\partial \beta} = 0 \end{aligned}$$

Now for large N, the second term within the brackets becomes negligible, and the above expression reduces to

$$\begin{aligned}
-\frac{1}{2}\frac{\partial \chi^2}{\partial \beta} &= \sum_i \left[\frac{x_i - NP_i(\beta)}{P_i(\beta)}\right] \frac{\partial P_i(\beta)}{\partial \beta} \\
&= \sum_i \frac{x_i}{P_i(\beta)} \frac{\partial P_i(\beta)}{\partial \beta} - \sum_i \frac{NP_i(\beta)}{P_i(\beta)} \frac{\partial P_i(\beta)}{\partial \beta}
\end{aligned}$$

Observe that since N is constant, and $\sum_i P_i = 1$, the second term on the right sums to 0 and the expression reduces further to

$$\sum_i \frac{x_i}{P_i(\beta)} \frac{\partial P_i(\beta)}{\partial \beta} = 0 \tag{1.34}$$

A solution to these i equations approximately minimizes the chi square.

For maximum likelihood estimation, the log of the multinomial p.d.f. is

$$\begin{aligned}
\log f(x_i) &= \log \left[\frac{N!}{\prod_i x_i!} \prod_i x_i(\beta)^{x_i}\right] \\
&= \log \left[\frac{N!}{\prod_i x_i!}\right] + \sum_i x_i \log P_i(\beta) - N \log N
\end{aligned}$$

Differentiating with respect to β yields

$$\frac{\partial \log f(x_i)}{\partial \beta} = \sum_i \frac{x_i}{P_i(\beta)} \frac{\partial P_i(\beta)}{\partial \beta} = 0 \tag{1.35}$$

Solving these i equations yields maximum likelihood estimates for P_i, which are asymptotically equivalent to the minimum chi square estimates.

1.5.3 Future Research

This chapter and its companion that follows extend the use of multinomial logit models for individual level response data to situations where the observations are aggregations of responses. This chapter extends the nested multinomial logit model of McFadden (1978) to the case where the observations are replicates. The following chapter extends the discrete-continuous framework of Dubin and McFadden (1984) to the case where the observations are replicates. Common to both examples is the use of a sequential estimation strategy. The following is a brief discussion of some remaining issues in extending these models to situations where aggregate level responses are used with a sequential estimation procedure.

Sequential estimation procedures are often used when the computational burden of the likelihood function of a fully specified model can be reduced to a series of easily estimated steps. The nested mulitnomial logit model with aggregate data specified by equations (15) to (17) and (25) to (30) easily lends itself to sequential estimation using minimum chi square estimation. Equations (15) to (17) along with (25) to (30) constitue the extension of the minimum chi square estimation procedure of Berkson (1955) and Theil (1971) to the nested mulitnomial logit with aggregate data. A full information procedure to estimate the nested multinomial logit would entail simultaneous, non-linear estimation of the lower stage (equation (15)) and upper stage (equation (16)) probabilities as well as the inclusive values (equation (17)). The approach adopted in this research was to estimate each stage separately and include an estimate of the inclusive value in estimation of the top stage. Amemiya (1978) has shown that this two-step procedure results in incorrect standard errors for the top stage estimates because sequential estimation treats the inclusive values as non-stochastic regressors, when in fact they are only estimates (although consistent) of the inclusive value. He developes a procedure by which correct standard errors for sequential estimation procedures can be obtained. This procedure relies on a modification to the top stage estimation (equation (16)) by using the hessian of the first stage estimates (equation (15)). Brownstone and Small (1989) propose a similar procedure. To date, these correction procedures have not been applied to sequential estimation with aggregate data.

To extend their work to the aggregate data nested multinomial logit the proof from the previous section that the minimum chi square estimator is asymptotically equivalent to the maximum likelihood estimator must be extended to the case with replicated observations for the nested logit likelihood function . When this is done, it should be possible to extend the gradient methods developed by Amemiya (1978) and Brownstone and Small (1989) to the case of the nested multinomial logit with replicated observations.

In the next chapter, an aggregate data version of the discrete-continuous system derived by Dubin and McFadden (1984) is estimated. A sequential estimation procedure is used in which first an aggregate data nested multinomial logit model is estimated and second, an aggregate data continuous variable regression model is estimated incorporating a correction for selection bias. This procedure uses estimated choice probabilities from the nested logit in the construction of selection bias correction terms for consistent estimation of the continuous variable (in this case, the amount of tax evasion found on tax returns). Unfortunately, as Dubin (1985) has shown, even for the case of individual level data the formulae for correcting the standard errors from this two stage procedure are quite involved. These procedures would need to be extended first to nested logit models with replicated data, and second to continuous variable regression models with replicated data.

References For Appendix

Amemiya, Takashi. *Advanced Econometrics.* Cambridge: Harvard University Press, 1985.

Berkson, Joseph. "Maximum Likelihood and Minimum Chi-Squared Estimates of the Logistic Function." *Journal of the American Statistical Association* 50 (1955):130-162.

Brownstone, David and Small, Kenneth. "Efficient Estimation of Nested Logit Models." *Journal of Business and Economic Statistics* 7 (1989):67-74.

Dubin, Jeffrey. *Consumer Durable Choice and the Demand for Electricity.* Amsterdam: North Holland Press, 1985.

McFadden, Daniel. "On Independence, Structuer, and Simultaneity in Transportation Demand Analysis." Working Paper 7511. Institute of Transportation Studies, University of California, Berkeley, 1975.

McFadden, Daniel. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location* edited by A. Karlgvist et al. Amsterdam: North Holland Press, 1978.

McFadden, Daniel, "Econometric Models of Probabilistic Choice." In *Structural analysis of Discrete Data with Econometric Applications* edited by Charles F. Manski and Daniel McFadden. Cambridge: MIT Press, 1981.

Theil, Henri, "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10 (1969):251-259.

Chapter 2

Tax Return Preparers and Tax Evasion

2.1 Introduction

The IRS estimates that for tax year 1992, 66 billion dollars of tax evasion will occur on individual income tax returns prepared by third party preparers (IRS 1990). The largest amount of this tax evasion will be associated with a relatively small percentage of all returns that are prepared by CPA's, attorneys, and Public Accountants, many of whom are tax practitioners.[1] Two other types of third party return preparation

[1] Certified Public Accountants (CPA's), attorneys, and Public Accountants who are in good standing within their professional organizations, and who meet certain continuing education requirements established by the IRS's Director of Practice are granted tax practitioner status. Tax practitioners not only prepare tax returns for a fee, they also may represent the taxpayer in matters before the IRS, including an audit, and provide expert opinions on positions maintained on a tax return that effectively shield the taxpayer from large penalties. Public Accountants are licensed at the state level with requirements varying by state. Only four states, North Carolina, Virginia, Kansas, and Wyoming do not regulate Public Accountants. Tax Practitioner behavior is governed by the Trea-

assistance account for the remainder of the 66 billion dollars of tax evasion. They are the non-paid preparers, comprised mostly of volunteers trained by the IRS to prepare tax returns, or a family member of the taxpayer,[2] and paid preparers who are not tax practitioners.[3]

In this paper I analyze the role that third party preparers of individual tax returns have on tax evasion. In particular, this research sheds light on how much of the tax evasion on returns can be attributable to the type of return preparation used. I estimate a four alternative simultaneous switching regression model and allow the amount of tax evasion found in each alternative to be endogenous with the choice of return preparation. The four types of return preparation used are non-paid assistance (Non-Paid), paid assistance who are not tax practitioners (Paid), tax practitioners (Practitioners), and self-prepared returns (Self). The number, relative frequency, and proportion of tax evasion attributable to these modes for tax year 1979 is summarized in table 1. I show that even after controlling for taxpayer characteristics, the mode of return preparation used affects tax compliance. An important finding is that the use of a tax practitioner lowers the amount of tax evasion while the use of Non-

sury Department's Circular No. 230, which describes both grounds for, and penalties applicable for violations of, acceptable conduct by tax practitioners. Table 1 shows that almost 17 percent of the returns filed in 1979 used a tax practitioner.

[2] Non-paid assistance includes returns that were prepared, advised, or reviewed by the IRS; returns prepared by unpaid volunteers under the VITA (Volunteer Income Tax Assitance) program sponsered by the IRS; or more generally by a family member. Preparers in this category face no legal burdens associated with providing tax return preparation assistance.

[3] Paid preparers include national tax services such as H & R Block, and local tax services that are not tax practitioners. These firms set their own standards of conduct, unlike CPA's, attorneys, and Public Accountants, and provide their own training. Moreover, paid preparers are not empowered to represent the taxpayer before the IRS in the case of an audit, and have no authority to provide an expert opinion to justify a position maintained by the taxpayer. Table 1 shows that nearly 29 percent of the returns filed in 1979 used a paid preparer.

Paid assistance or Paid assistance has no effect on tax evasion. Another important conclusion from this research is that complexity of the tax return per se does not increase the amount of non-compliance with the tax code if Practitioners prepare the return. In fact, increased complexity may increase compliance with the tax code if there is a concomitant increase in the use of Practitioners. I offer some possible explanations for this result in the Conclusion.

Table 2.1: Returns and Proportion of Non-Compliance by Mode of Preparation

	Returns		Proportion of
Mode of Preparation	Number	Frequency	Noncompliance
SELF	39959000	.442	.228
NON-PAID	9583006	.106	.037
PAID PREPARERS	25766600	.285	.308
PRACTITIONERS	15097850	.167	.427
TOTAL	90406000	1.000	1.000

This paper is organized as follows. Section two summarizes the literature relevant to tax return perparation and tax evasion. Section three presents a model of the demand for tax evasion conditional upon the four modes of tax return preparation and describes the models estimated and the data used. Section four presents results of the estimation and contains conclusions.

2.2 Literature Review

The diversity of services provided, skill levels, and business intentions of third party tax return preparers has made it difficult for economists to develop a unified theory of tax return preparer behavior. Although no single result applies to all of the modes of third party return preparation, there are situations when the use of a tax preparer results in less taxes paid.[4] For example, Scotchmer (1989a) found that third party preparers reduce uncertainty about taxable income, which reduces reported tax liability. A similar result was found by Reinganum and Wilde (1989) who focus on the efficiency of third party preparers in lowering the cost of gathering information and preparing the tax return. Subsequent work by Scotchmer (1989b), Graetz, Reinganum and Wilde (1989), and Klepper, Mazur and Nagin (1991) focused on the effect of tax practitioners per se on non-compliance. Scotchmer (1989b) showed that lax enforcement of professional standards for tax practitioners could result in lower tax compliance. Graetz, Reinganum and Wilde (1989) and Klepper, Mazur and Nagin (1991) emphasized the unique abilities of tax practitioners to issue expert opinions and to exploit the vagueness of particular sections of the tax code in reducing the reported tax liability. Both underscore the ability of tax practitioners to influence

[4]Tax evasion is defined as the difference between a taxpayer's reported liability and the true tax liability. In certain situations, determining the true tax liability is difficult. This complicates the analysis of tax evasion because in these situations, taxpayers, tax return preparers, and the IRS disagree on what the correct tax should be. Resolving these differences often requires a compromise on the part of the parties involved. In order to progress towards an understanding of the myraid of factors that may effect the reported tax liability of a tax return, many theorists have abstracted this problem away. In this research, I assume that the IRS determined tax liability is the correct liability.

the determination of tax liability. These analyses suggest that Practitioners, and to a lesser extent Paid preparers, should lower the amount of non-compliance found on tax returns.

Nitzan and Tzur (1989) and Melamud, Wolfson, and Ziv (1991) modelled the relationship between the IRS and tax practitioners. In particular, they focused on the ability of the IRS to grant monopoly profits to tax practitoners by controlling the number of tax practitioners in exchange for an attestation function by the tax practitioners when preparing tax returns. In this capacity, tax practitioners, through training, professional codes of conduct, and through special penalties applicable to tax practitioners verify amounts reported on tax returns which improves tax compliance.

The theoretical literature illustrates the diversity of perspectives from which to analyze the role of third party tax preparers in general, and tax practitioners in particular on tax compliance. One hypothesis tested in this paper is that as the complexity of the tax return increases, non-compliance falls on Paid preparer and Practitioner prepared returns. A second hypothesis is that the probability of receiving a penalty will be lowered as a result of using a tax practitioner. One consequence of this "insurance" is a reduction in penalties concomittant an increase in taxes as a result of an IRS audit. This would occur because tax practitioners can provide expert opinions, which insulate the taxpayer from the imposition of large penalties. A third hypothesis is that tax practitioners lower tax non-compliance by fulfilling an attestation function for the IRS.

The empirical literature is more consistent in its findings on the demand for tax return preparation services or the effect of tax return preparers on compliance. In a series of papers, Slemrod and Sorum (1984), Slemrod (1989), Collins, Milliron and Toy (1987), Hite (1987), and Dubin, Graetz, Udell and Wilde (1992) found in general that greater amounts of income, capital gains, probability of self-employment, sole-proprietor income, itemized deductions, return complexity, age of taxpayer, and marginal tax, penalty, and audit rates all increased the use of paid third party preparers, while greater educational levels attained and/or greater knowledge of the tax code reduced the use of paid third party preparers. With the exception of Dubin, Graetz, Udell and Wilde (1992) these researches combined (or could not separate) Practitioners from Paid preparers. They show that in addition to the preceeding results, greater amounts of wage, interest, and dividend income reduce the demand for Practitioners.

Long and Caudill (1987) modelled both the demand for tax return assistance and reported tax liability. Their findings for the demand for paid assistance are similar to those above, while they find that in general the reported tax liability from returns prepared by paid preparers is less than unassisted modes of return preparation. Erard (1993) modelled the demand for tax return preparation and for tax evasion among self prepared, paid preparers who were not practitioners, and practitioners. He found that the demand for tax practitioners and paid preparers who were not tax practitioners increases with capital gains, small business or farm activity, rents and royalties, the number of tax forms attendant the return, being over 65 years of age, previous audit history, the marginal tax rate, and the IRS audit rate. Unlike tax practitioners, he found that the demand for paid preparers decreased with the marginal tax rate. One striking result of this is that the use of tax practitioners lowers tax compliance.

The empirical literature shows that, in general, greater amounts of income, and more complex returns, increase the demand for third party return preparation. The effect of this increased demand for paid third party assistance on the amount of tax evaded depends upon which type of return preparation assistance is demanded. If increases in income and complexity increase the demand for tax practitioners, then this may result in less tax evasion because of their tax expertise, ability to exploit ambiguity in the tax code, and attestation function. Alternatively, if third party return preparation assistance is largely a matter of convenience, then there may be little if any effect on tax evasion from their use. In the model presented below, I test these competing hypotheses.

2.3 Model

2.3.1 Specification

With the exceptions of Long and Caudill (1987) and Erard (1993), the empirical literature on tax evasion has not controlled for the endogeneity of third party tax return preparers on tax evasion.[5] These analyses control for the endogeneity of the mode of preparer, but because of statistical procedures, restrict the choice set for mode of return preparation. I extend the research on the demand for tax return preparation services presented in Dubin, Graetz, Udell, and Wilde (1992) and use audited tax return information to model the effect of self prepared (SELF), non-paid prepared (NON-PAID), paid prepared who are not practitioners (PAID), and practitioners (PRACTITIONERS) on tax evasion with a simultaneous switching regression model.

Dubin, Graetz, Udell, and Wilde (1992) specify the taxpayer's choice of mode of return preparation with a random utility model. Briefly, let $i = 1, ..., I$ denote the mode of return preparation. Let the taxpayer's utility U_i from the i^{th} mode of return preparation be given by

$$U_i = V_i + \epsilon_i \tag{2.1}$$

where the strict utility component V_i is a function of non-stochastic observed attributes, and ϵ_i is a random component of unobserved attributes. For the taxpayer

[5] Long and Caudill (1987) use unaudited 1983 tax return information to model the difference between professional tax return preparation (combining tax practitioners and non-practitioner paid modes) and non-paid modes of tax return preparation (combining self and non-paid assisted modes) on reported tax liability. Erard (1993) uses audited 1979 tax return information with a trinary distinction between non-paid prepared (combining self and non-paid assisted modes), paid-prepared that was not a practitioner, and practitioner-prepared returns to model their effect on tax evasion. As Dubin, Graetz, Udell and Wilde (1992) show, restricting the choice of mode of return preparation to two or three alternatives can produce misleading inferences about the motives for the demand for tax return preparation assistance. We build on their research and analyze the effect of a larger choice set on the demand for tax evasion. While Long and Caudill (1987) and Erard (1993) use micro data and this research uses aggregate data, we believe that the costs associated with the aggregation of alternatives is greater than the costs associated with the aggregation of the data. In the following section, we describe the method of data aggregation used.

we observe the choice of the i^{th} mode of return preparation when

$$\begin{aligned}\delta_i &= 1 \text{ iff } U_i > U_m \ \forall_m \ \ m \neq i \\ &= 0 \text{ otherwise.}\end{aligned} \tag{2.2}$$

Denote the probability of choosing the i^{th} mode of return preparation be

$$\begin{aligned}P_i &= P\left[U_i \geq U_m \quad \forall_m \ \ m \neq i\right] \\ &= P\left[V_i + \epsilon_i \geq V_m + \epsilon_m \ \ \forall_m \ \ m \neq i\right] \\ &= P\left[\epsilon_m - \epsilon_i \leq V_i - V_m \quad \forall_m \ \ m \neq i\right].\end{aligned} \tag{2.3}$$

Assuming that the random component of utility ϵ is distributed with the independent extreme value distribution and that non-random component of utility is of the form

$$V_i = \Gamma_i X_i$$

where X_i is a column vector of exogenous characteristics of a taxpayer that selected the i^{th} mode of return preparation and Γ_i is a row vector of unknown coefficients, McFadden (1978) demonstrates that the probability P_i is of the multinomial logit form:

$$P_i = \frac{e^{\Gamma_i X_i}}{\sum_{i=1}^{I} e^{\Gamma_i X_i}}. \tag{2.4}$$

The equation for the amount of tax evasion, Y, on a return prepared by the i^{th}

mode of return preparation is

$$Y_i = \alpha_i + \beta_i X_i + \eta \quad \text{if } \delta_i = 1 \tag{2.5}$$

where Y_i is a scalar dependent variable, α_i is a preparer specific parameter, β_i is a row vector of parameters, X_i is a column vector of exogenous characteristics of taxpayers and their returns associated with the i^{th} mode of return preparation, and η is a scalar random component with mean 0 and constant variance σ^2.

In the presence of correlation between η and δ_i, ordinary least squares estimation of equation (5) yields inconsistent estimates of α_i and β_i. Such correlation might arise because an unobservable feature of the taxpayer that increases the probability of selecting a tax practitioner, such as a propensity to aggresively minimize reported tax liability, may be correlated with the amount of unreported tax discovered by the IRS upon audit. In situations where this correlation may arise, Dubin and McFadden (1984) and Dubin (1985) have derived several estimators that allow consistent estimates of α_i and β_i.

To proceed, define $\nu = \eta - E(\eta \mid \delta_i = 1)$. Then $E(\nu \mid \delta_i = 1) = 0$. Dubin and McFadden (1984) further assume that the conditional distribution of η given $\epsilon_1 \ldots \epsilon_I$ has mean $\left(\frac{\sqrt{2}\sigma}{\lambda}\right) \sum_{i=1}^{I} \rho_i \epsilon_i$ and variance $\sigma^2(1 - \sum_{i=1}^{I} \rho_i^2)$ where $\sum_{i=1}^{I} \rho_i = 0$ and $\sum_{i=1}^{I} \rho_i^2 < 1$. Then ρ_i is the correlation of η and ϵ_i where ϵ_i has unconditional mean 0 and unconditional variance $\frac{\lambda^2}{2}$, and η has unconditional mean 0 and unconditional variance σ^2. Under these assumptions, Dubin and McFadden (1984) show that

$$\begin{aligned} E(\eta \mid \delta_i = 1) &= \sum_{m=1}^{I} \left(\frac{-\sqrt{6\sigma^2}\rho_m}{\pi} \right) \left[\frac{\log P_m}{(1-P_m)} \right] [P_m - \delta_{im}] \\ &= \sum_{m \neq i}^{I} \left(\frac{-\sqrt{6\sigma^2}\rho_m}{\pi} \right) \left[\frac{P_m \log P_m}{1-P_m} + \log P_i \right] \end{aligned} \quad (2.6)$$

where $\delta_{im} = 1$ when $i = m$ and 0 otherwise, σ^2 is the unconditional variance of η and where ρ_m is the correlation of the m^{th} mode of return preparation with η. Equation (7) specifies a correction for selection bias when the choice model is of the multinomial logit form and the continuous variable outcome associated with that choice is a regression model. For the I alternative model, equation (7) specifies $I-1$ selection bias correction terms. Each of these terms can be decomposed into a correction variable

$$C(P_m, P_i) = \left[\frac{P_m \log P_m}{1-P_m} + \log P_i \right] \quad (2.7)$$

and a correction parameter

$$\gamma_m = \frac{-\sqrt{6\sigma^2}\rho_m}{\pi}. \quad (2.8)$$

Including these terms in respecifying equation (5) with a correction for selection bias, yields for each mode of return preparation

$$Y_i = \alpha_i + \beta_i X_i + \sum_{m \neq i}^{I} \gamma_m C(P_m, P_i) + \nu \quad \text{if } \delta_i = 1. \tag{2.9}$$

Dubin (1985) shows that when the true value of P_m in equation (7) is not known, a consistent estimate of P_m may be substituted for P_m.[6]

Finally, I impose no cross equation restrictions on the coefficients, β_i. This later assumption seems reasonable because there is no a priori reason to believe that exogenous variables such as total income, age, and family size should have the same effect on tax evasion given the mode of return preparation. Different values of the coefficients would indicate that these variables affect the amount of tax evasion to different degrees depending upon what mode of return preparation is used. This reflects the different levels of expertise and familiarity with the tax code and IRS procedures that the preparer types represent.

Equations (2), (3) and (5) can be interpreted as a switching regression model with I switches corresponding to the I modes of tax return preparation and the effect of that demand on tax evasion. Equation (6) specifies the distribution of the errors associated with equations (3) and (5) and equation (7) specifies a correction

[6] Alternatively, consistent estimates of α_i and β_i can be made with instrumental variables estimation using a consistent estimate of δ_i. One such instrument would be P_i (or a consistent estimate of P_i), the probability of choosing the i^{th} mode of return preparation as specified by equation (4). I attempted instrumental variables estimation, using the predicted probabilities of mode demand as instruments for the mode specific dummy variables in equation (5). However, instrumental variables estimates were very sensitive to the limited number of degrees of freedom of this model, and did not yield stable estimates.

for selection bias. The unknown parameters, α_i, β_i, and $\sum_{m\neq i}^{I}\gamma_i$ are estimated for each of the I equations on the subset of observations that selected the i^{th} mode of return preparation.

2.3.2 Aggregation

In a previous paper, Dubin, Graetz, Udell and Wilde (1992) estimate equation (4) using aggregate data from the 1979 Taxpayer Compliance Measurement Program (TCMP).[7] To estimate equation (10) with aggregate data, let $k = 1, ..., K$ be the number of taxpayers in an IRS district; $i = 1, .., I$ be the number of modes of tax return preparation available in an IRS district; and $J = 1, .., J$ be the number of IRS districts. Let $\delta_{ijk} = 1$ if the k^{th} taxpayer in the j^{th} IRS district selects the i^{th} mode of return preparation, and zero otherwise. Define the number of taxpayers in the j^{th} IRS district who select the i^{th} mode of return preparation as

$$N_{ij} = \sum_{k=1}^{K} \delta_{ijk}. \tag{2.10}$$

[7] This data set is the result of efforts by the Internal Revenue Service (IRS) to assess the size and extent of non-compliance with the filing of individual Federal income tax returns. For public use, the IRS prepared aggregate data extracts of the 1979 individual tax return micro-data set. For the extracts used in this research, the aggregation takes place over all taxpayers in the 58 IRS districts, which are geographically exclusive and exhaustive of the United States. One minor exception to this is the exclusion of tax returns filed by U.S. citizens employed by the Federal government who are resident overseas. Fourty-four of the districts are states. Of the remaining 14 districts, four are in New York and there are two each in California, Texas, Pennsylvania, Illinois, and Ohio.

The mean value of the amount of tax evasion found on tax returns prepared by the i^{th} mode of return preparation in the j^{th} IRS district is

$$\overline{Y}_{ij} = \sum_{k=1}^{K} \frac{\delta_{ijk}}{N_{ij}} Y_{ijk}. \tag{2.11}$$

For the k^{th} taxpayer in the j^{th} IRS district who selects the i^{th} mode of return preparation, equation (9) is

$$Y_{ijk} = \alpha_i + \beta_i X_{ijk} + \sum_{m \neq i}^{I} \gamma_m C(P_{mjk}, P_{ijk}) + \nu_{jk} \quad \text{if } \delta_{ijk} = 1, \tag{2.12}$$

where the probabilities P_{mjk}, P_{ijk} in $C(P_{mjk}, P_{ijk})$ are the probabilities that the k^{th} taxpayer selects the m^{th} and i^{th} mode of return preparation within the j^{th} IRS district office. Summing across the K individuals within the i^{th} preparer class and the j^{th} IRS district equation (12) becomes

$$\begin{aligned}\sum_{k=1}^{K} \delta_{ijk} Y_{ijk} = {} & \alpha_i \sum_{k=1}^{K} \delta_{ijk} + \beta_i \sum_{k=1}^{K} \delta_{ijk} X_{ijk} + \sum_{m \neq i}^{I} \gamma_m \sum_{k=1}^{K} \delta_{ijk} C(P_{mjk}, P_{ijk}) + \\ & \sum_{k=1}^{K} \delta_{ijk} \nu_{jk} \quad \text{if } \delta_{ijk} = 1. \end{aligned} \tag{2.13}$$

Equation (13) is the aggregate amount of evasion associated with the i^{th} preparer class in the j^{th} IRS district. By combining equations (10) and (11), the aggregate amount of evasion in equation (13) can be rewritten as

$$\sum_{k=1}^{K} \delta_{ijk} Y_{ijk} = N_{ij} \overline{Y}_{ij}. \tag{2.14}$$

The same procedure used in equation (14) can be used to rewrite each variable and the constant term as the product of the number of taxpayers in the i^{th} preparer class in the j^{th} IRS district, N_{ij}, and the mean value of the variable for taxpayers in the i^{th} preparer class in the j^{th} IRS district so that equation (13) becomes

$$\begin{aligned} N_{ij}\overline{Y}_{ij} &= N_{ij}\alpha_i + N_{ij}\beta_i\overline{X}_{ij} + N_{ij}\sum_{m\neq i}^{I} \gamma_m \overline{C}_{ij} + \\ &\quad \psi_{ij} \quad \text{if } \delta_{ij} = 1 \end{aligned} \tag{2.15}$$

where

$$\overline{C}_{ij} = \frac{1}{N_{ij}} \sum_{k=1}^{K} \delta_{ijk} C(P_{mjk}, P_{ijk}) \tag{2.16}$$

is an average of the correction bias terms defined in equation (7) over the K individuals in the j^{th} IRS district. Rewrite equation (15) replacing $\overline{C}_{ij}$ which is the average of the correction bias terms, by the correction bias term evaluated at the average probability of taxpayers who choose the m^{th} and i^{th} modes of return preparation respectively so that

$$\begin{aligned} N_{ij}\overline{Y}_{ij} &= N_{ij}\alpha_i + N_{ij}\beta_i\overline{X}_{ij} + N_{ij}\sum_{m\neq i}^{I} \gamma_m C(\overline{P}_{mj}, \overline{P}_{ij}) + \\ &\quad \psi_{ij}^{1} \quad \text{if } \delta_{ij} = 1 \end{aligned} \tag{2.17}$$

where

$$\overline{P}_{ij} = \sum_{k=1}^{K} \frac{\delta_{ijk}}{N_{ij}} P_{ijk}. \tag{2.18}$$

and where

$$\psi_{ij}^{1} = \psi_{ij} + \left[N_{ij} \sum_{m \neq i}^{I} \overline{C}_{ij} - N_{ij} \sum_{m \neq i}^{I} C(\overline{P}_{ij}, \overline{P}_{mj}) \right]. \tag{2.19}$$

Still, this definition of the correction bias term requires knowledge of the individual choice of return preparer probabilities which were not available for this research. Instead, I use an estimate of $\overline{P}_{ij}$ from Dubin, Graetz, Udell and Wilde (1992). With their estimates of the choice of mode of return preparation, I replace the average of the individual choice probabilities for taxpayers selecting the i^{th} mode of return preparation in the j^{th} IRS district, $\overline{P}_{ij}$, by the choice probability of an average individual in the j^{th} IRS district selecting the i^{th} mode of return preparation, P_{ij}.[8] Using probabil-

[8] These probabilities were estimated by a multinomial nested logit model using aggregate data and minimum chi square techniques described in Dubin, Graetz, Udell, and Wilde (1992). They are defined as

$$P_{mj} = \frac{e^{\Gamma_m \overline{X}_j}}{\sum_{m=1}^{I} e^{\Gamma_m \overline{X}_j}}$$

where

$$\begin{aligned} \overline{X}_j &= \sum_{i=1}^{I} N_{ij} \frac{\overline{X}_{ij}}{N_j} \\ &= \sum_{i=1}^{I} \frac{N_{ij}}{N_j} \left(\frac{\sum_{k=1}^{K} \delta_{ijk} X_{ijk}}{N_{ij}} \right) \end{aligned}$$

and where

$$N_j = \sum_{i=1}^{I} \sum_{k=1}^{K} \delta_{ijk}$$

ities of taxpayer averages rather than the average of taxpayer probabilities introduces two possible sources of error. The first problem is that often P_{mj} is a biased estimate of $\overline{P}_{mj}$.[9] A second potential source of error is that we do not observe directly the true value of P_{mj}, but use an estimate in its place. This introduces an approximation error with using P_{mj}. Both sources of error can be mitigated by estimating P_{mj} over sufficiently homogeneous classes of taxpayers using a consistent estimator for aggregate choice shares. This approach was followed by Dubin, Graetz, Udell and Wilde (1992), who used data grouped over 696 mutually exclusive and exhaustive categories that placed taxpayers into nearly homogeneous aggregation classes, and estimated aggregate choice shares using a minimum chi square estimation procedure.[10] Substituting P_{mj} for $\overline{P}_{mj}$, equation (17) can be rewritten as

$$\begin{aligned} N_{ij}\overline{Y}_{ij} &= N_{ij}\alpha_i + \beta_i N_{ij}\overline{X}_{ij} + N_{ij}\sum_{m\neq i}^{I}\gamma_m C(P_{mj}, P_{ij}) + \\ & \quad \psi_{ij}^2 \quad \text{if } \delta_{ij} = 1 \end{aligned} \tag{2.20}$$

$$= \sum_{i=1}^{I} N_{ij}$$

is the total number of taxpayers in the j^{th} IRS district.

[9]The bias occurs because the logit probabilities defined by equation (4) are non-linear in X. By application of Jensen's inequality, it can be shown that for high probability events, the probability of the average response is greater than the average of the probabilities of the individual response. The converse holds for low probability events. Allenby and Rossi (1991) show that under certain conditions logit models fitted to aggregate data do not produce biased estimates. They argue that when heterogenous consumers have the same choice set and the same information about that choice set, the resulting aggregate choice shares of individual consumers closely resemble the aggregate choice shares of a representative consumer and therefore, very little information is lost from using aggregate market share data.

[10]See Dubin, Graetz, Udell and Wilde (1992) for details.

where

$$\psi_{ij}^2 = \psi_{ij} + \left[N_{ij} \sum_{m \neq i}^{I} \overline{C}_{ij} - N_{ij} \sum_{m \neq i}^{I} C(\overline{P}_{mj}, \overline{P}_{ij})\right] + \left[N_{ij} \sum_{m \neq i}^{I} C(\overline{P}_{mj}, \overline{P}_{ij}) - C(P_{mj}, P_{ij})\right].$$

Equation (20) relates the total amount of evasion associated with the i^{th} mode of return preparation in the j^{th} IRS district to aggregate taxpayer characteristics for that IRS district and preparer class. To apply least squares estimation to equation (20), note that the error term, ψ_{ij}^2, has an expected value of 0 if $\delta_{ij} = 1$ since it is the sum of K terms, each which has conditional expectation equal to 0. Furthermore, the variance of ψ_{ij}^2 is of order $N_{ij} = \sum_{k=1}^{K} \delta_{ijk}$. Therefore, a correction for heteroscedasticity can be made to equation (20) by dividing each member through by $N_{ij}^{\frac{1}{2}}$.[11]

2.3.3 Data

The 1979 TCMP file for individual returns involves line-by-line audits of approximately 50,000 randomly selected tax returns. Our file aggregates the results of the 1979 TCMP audits by the 58 IRS district and four modes of return preparation. The four modes of return preparation used for estimation are self-prepared (SELF), non-paid prepared (NON-PAID, including IRS assisted and other non-paid assistance), paid preparers (PAID) who were not enrolled agents (mostly H & R Block or Beneficial Finance Corporation), and practitioners (PRACTITIONERS, including enrolled

[11] The number of returns that were self prepared averaged 69,248 per IRS district office, while the number of returns for non-paid averged 15,860; for paid perparers 44,521; and for practitioners was 26,241.

agents that were Public Accountants, CPA's, or attorneys). Both the taxpayer's reported amounts and the adjusted amounts recommended by the TCMP audit are recorded.[12] Our dependent variable (EVASION) is the difference between the taxpayer reported liability and the IRS examiner's corrected liability net of overreported liability.

To test the hypothesis that tax evasion decreases with the complexity of the tax situation, I include the number of forms filed with the tax return (FORM).[13] I also include two variables that are generally believed to be positively correlated with tax evasion. They are the sum of income from schedules C, D, E, F, and Form 4797 (COMPLEX) and state, local and real estate tax deductions (ASSET). The later acts as a measure of state and local tax burdens, while the former is associated with federal tax burdens.

I include the frequency with which penalties were assesed in the TCMP audit (PENALTY) to test whether the frequency of Penalty situations act as a deterrent to tax evasion. A central focus of this research is to test whether the choice of mode return preparation affects the amount of tax evasion. To test for this dependence, I include a preparer a specific variable, $C(P_{mj}, P_{ij})$, defined in the previous section to control for the potential correlation between the choice of mode of return preparation and the amount of evasion.

[12] We use the corrected amounts of deductions and exemptions as meaningful measures of the true amounts of these items.

[13] These forms include Schedule C for Profit or Loss from a Business; schedule D for Capital Gains and Losses; schedule E the Supplemental Income Schedule to report income from rents, royalties, and trusts; schedule F for Farm Income and Expenses; and Form 4797 for Sales of Business Property.

To complete our specification, I include three additional variables. They are the sum of wage, salary, interest, and dividend income (SIMPLE), the number of dependents eligible to be claimed by the taxpayer (EXEMPTION), and the number of taxpayers over 65 years of age (OVER65). Together, SIMPLE and COMPLEX account for nearly all of a taxpayer's income. Their mnemonics reflect the complexity of the rules in the tax code governing the reporting of income from these sources. EXEMPTION and OVER65 capture two important demographic features. By our definition, EXEMPTION primarily measures family size. An increase in EXEMPTION, all else held constant, should increase the demand for tax evasion as the additional cost of a family member typically exceeds the value of the exemption. I include OVER65 in our specification to test whether the mode of return assistance used by the elderly is related to the demand for tax evasion, or to the service and information aspects of tax return preparation. The mean values of these variables for each mode of return preparation are reported in table 2.[14]

[14] Note that the aggregation scheme described in the previous section places restrictions on the use of variables that are auxilary to the 1979 TCMP data set. In particular, under this aggregation scheme, the audit rate data that was available for this research would have resulted in constant audit rates across preparer classes within an IRS district. Because there would be no within district audit rate variation, it is not included in our estimation. However, omitting the IRS audit rate from estimation of equation (5) is not the same as omitting the IRS audit rate from the analysis since the IRS district level audit rate is employed in the estimation of the demand for tax return preparation services as described in Dubin, Graetz, Udell, and Wilde (1992). In their analysis, the probability of selecting the i^{th} mode of return preparation was estimated using audit class specific data within an IRS distirict.

Table 2.2: Mean Values of Variables by Mode of Preparation

Variable	Self	Non-paid	Paid	Practitioner
SIMPLE	13,900	10,151	18,900	38,100
COMPLEX	963	740	2,255	7,252
ASSET	777	358	958	2,046
EXEMPTION	183	71	341	802
OVER 65	0.072	0.112	0.151	0.217
FORM	1.210	1.175	1.510	2.210
PENALTY	0.034	0.055	0.070	0.096
EVASION	112	130	225	655

Notes: Amounts in Dollars
Frequencies in Proportion of Returns

For each of the four modes of return preparation, I estimate equation (20) using weighted least squares with the following specification[15](where the subscript j has been suppresed from all variables):

$$\begin{aligned}
\text{EVASION}_i = \; & \alpha_i \; + \beta_{1i}\,\text{SIMPLE}_i + \beta_{2i}\,\text{COMPLEX}_i \\
& + \beta_{3i}\,\text{ASSET}_i + \beta_{4i}\,\text{EXEMPTION}_i + \beta_{5i}\,\text{OVER65}_i \\
& + \beta_{6i}\,\text{QUESTIONABLE}_i + \beta_{7i}\text{FORM}_i + \beta_{8i}\,PENAL_i \\
& + \sum_{m \neq i}^{K} \gamma_m C(P_m, P_i) + \text{ERROR}.
\end{aligned}$$

[15]The correction terms have been normalized.

2.4 Results

2.4.1 The Demand for Tax Evasion

Table 3 presents ordinary least squares estimates of equation (20). An increase in either SIMPLE or COMPLEX income increases the amount of tax evasion on the Practitioner mode of return preparation while only increases in COMPLEX income increase the amount of evasion found on Paid prepared returns. I find no significant effect from either income variable on evasion for Self prepared or Non-paid prepared returns. Our results for state and local taxes (ASSET), family size (EXEMPTIONS), and taxpayers over the age of 65 years (OVER 65) show no effect on tax evasion, with the lone exception that greater state and local tax burdens increase the amount of evasion found on Practitioner prepared returns.

Increases in the penalty rate (PENALTY) increase the amount of evasion detected among Self prepared and Paid prepared modes of return preparation.[16] Although we find no support for a deterrent effect from penalties, note that relative to other modes of return preparation, Practitioners reduce the exposure to penalties, even though the returns that they prepare certainly entertain situations where penalties would apply for non-compliance. Increases in the complexity of the tax return, as measured by the average number of forms and schedules (FORM) attendant the return, decreases

[16]The current penalty regime, with substantially higher penalty rates, was created largely during the penalty reforms placed into law with the 1989 Omnibus Budget Reconcilliation Act. For example, the penalty for intentional disregard of rules with respect to the paying of income tax was 5 percent of the underpayment of tax in 1979 (per the Internal Revenue Code of 1954 section 6653(a)) but is currently 20 percent (per the Internal Revenue Code of 1986 section 6662(b) as amended in 1989).

the amount of evasion found on returns prepared by Practitoners, but not for any other mode of return preparation. Finally, the coefficient of the selectivity correction parameter, $C(P_m, P_i)$, is significant, and positive, for the Practitioner mode. From equation (6), this means that there is a negative correlation between the unobservable characteristics of the choice of the Practitioner mode to prepare a tax return and the amount of evasion detected on that return and supports the hypothesis that Practitioners reduce non-compliance. Note that the insignificance of the selectivity correction parameter for the Paid preparer mode supports the view that paid tax return preparers who are not Practitioners are providing a convenience role more than a tax minimizing role for their clients.

Table 2.3: Estimates of Tax Evasion

Weighted Variable	Self	Non-paid	Paid	Practitioner
ONE	-8.617	-0.427	-87.698	61.159
	(-0.283)	(-0.007)	(-2.361)	(1.172)
SIMPLE	0.001	0.005	0.010	0.007
	(0.091)	(0.378)	(1.150)	(1.830)
COMPLEX	0.084	0.090	0.074	0.094
	(1.541)	(1.116)	(2.693)	(6.842)
ASSET	0.035	0.067	-0.035	0.076
	(0.707)	(0.498)	(-0.918)	(3.626)
EXEMPTION	-0.046	0.113	-0.038	-0.034
	(-0.199)	(0.280)	(-0.268)	(-0.026)
OVER 65	-415.762	-141.249	-69.943	-427.087
	(-1.059)	(-0.311)	(-0.212)	(-1.384)
FORM	77.789	63.246	-89.682	-345.340
	(0.557)	(0.295)	(-0.723)	(-3.902)
PENAL	1855.020	513.482	1846.570	36.152
	(2.908)	(0.781)	(4.267)	(0.085)
CORRECTION		-57.010	15.041	135.024
		(-0.759)	(0.217)	(2.539)
R-squared	.79			
Number of Observations	232			

Note: t-statistics in parenthesis.

To derive the elasticity concepts used in table 4 below, start with a definition of total evasion, Y, as

$$Y = \sum_{i=1}^{I}\sum_{j=1}^{J} N_{ij}\overline{Y}_{ij}. \tag{2.21}$$

Taking expectations yields

$$\begin{aligned} E(Y) &= E\left(\sum_{i=1}^{I}\sum_{j=1}^{J} N_{ij}\overline{Y}_{ij} \mid \delta_{ij} = 1\right) P_{ij} + \\ & \quad E\left(\sum_{i=1}^{I}\sum_{j=1}^{J} N_{ij}\overline{Y}_{ij} \mid \delta_{ij} = 0\right) (1 - P_{ij}) \end{aligned} \tag{2.22}$$

where

$$P_{ij} = \text{Prob}(\delta_{ij} = 1).$$

Note that the second term in the equation (22) is 0 because when $\delta_{ij} = 0$ it follows that $\overline{Y}_{ij} = 0$. Therefore, expected total evasion is

$$E(Y) = \sum_{i=1}^{I}\sum_{j=1}^{J} E(N_{ij}\overline{Y}_{ij} \mid \delta_{ij} = 1)P_{ij} \tag{2.23}$$

The first term in equation (23), $N_{ij}\overline{Y}_{ij}$ is given by equation (20), and the second term, P_{ij}, is given by equation (4). P_{ij} was estimated by Dubin, Graetz, Udell, and Wilde (1992).[17]

Two elasticity concepts are presented in table 4. The first four columns show the

[17] Note that in equation (4), the subscript for the j^{th} IRS district has been suppresed.

mode specific short run elasticities of tax evasion. They are estimated using equation (23) by holding constant the demand for tax return preparation services. This treats the P_{ij} in equation (23) as constant and is given by

$$\begin{aligned} \epsilon_i^{short} &= \sum_{j=1}^{J} \frac{\partial(N_{ij}\overline{Y}_{ij})}{\partial \overline{X}_{ij}} \frac{\overline{X}_{ij}}{N_{ij}\overline{Y}_{ij}} P_{ij} \\ & \sum_{j=1}^{J} \beta_i \frac{\overline{X}_{ij}}{N_{ij}\overline{Y}_{ij}} P_{ij}. \end{aligned} \tag{2.24}$$

The fifth column presents the sum of the short run elasticities over all modes of return preparation. The sixth column presents the long run elasticities of tax evasion. For this calculation, the demand for each mode of tax return preparation is allowed to adjust, and is given by

$$\epsilon^{long} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\frac{\partial(N_{ij}\overline{Y}_{ij})}{\partial \overline{X}_{ij}} P_{ij} + N_{ij}\overline{Y}_{ij} \frac{\partial P_{ij}}{\partial \overline{X}_{ij}} \right] \frac{\overline{X}_{ij}}{N_{ij}\overline{Y}_{ij}} \tag{2.25}$$

where $\frac{\partial P_{ij}}{\partial \overline{X}_{ij}}$ was calculated in Dubin, Graetz, Udell, and Wilde (1992).

The unconditional effect of the two income variables, SIMPLE and COMPLEX, is to increase the amount of tax evasion among all modes of return preparation. The first four columns of table 4 decompose the overall short run elasticity of 0.402 for SIMPLE income into each return preparation mode's weighted contribution. The three assisted modes of return preparation have contributions to the short run total elasticity of 0.038 for Non-paid, 0.238 for Paid, and 0.071 for Practitioner prepared returns. Increases in COMPLEX income have the greatest effect on tax evasion, with

Table 2.4: Short Run and Long Run Elasticities of Tax Evasion

	Short Run Elasticities					Long Run Elasticities
Mode	Self	Non-paid	Paid	Practitioner	Total	Total
Variable	(1)	(2)	(3)	(4)	(5)	(6)
SIMPLE	0.055	0.038	0.238	0.071	0.402	0.325
COMPLEX	0.321	0.049	0.210	0.182	0.762	0.771
ASSET	0.108	0.015	-0.042	0.042	0.123	0.135
EXEMPTION	-0.033	0.006	-0.016	-0.007	-0.050	-0.057
OVER 65	-0.118	-0.011	0.013	-0.025	-0.141	-0.165
FORM	0.374	0.055	-0.171	-0.204	0.054	-0.412
PENAL	0.250	0.021	0.163	0.001	0.435	0.398

a total short run elasticity of tax evasion of 0.762. Interestingly, the largest component of the short run effect is from self prepared returns with a short run elasticity of 0.320 followed next by paid preparer's at 0.210 followed by Practitioner prepared returns at 0.182. The overall effect of state and local income taxes, ASSET, on tax evasion is quite small, with a conditional elasticity of 0.123 and an unconditional elasticity of 0.135. However, consistent with our expectations greater amounts of ASSET increase evasion on Practitioner prepared returns with an elasticity, albeit small, of 0.042. Interestingly, family size, as measured with EXEMPTION, has very little overall effect on tax evasion. The overall elasticity of evasion with respect to AGE65 is also quite small at -0.141, with the bulk of this coming from self prepared returns, with a mode elasticity of evasion of -0.118.

Our strongest results are for the variable (FORMS). The short run total elasticity is 0.054. However, the small size of this effect belies its distributional character because the short run elasticities for Practitioner and Paid prepared returns are -0.203 and -0.171 respectively. These are more than offset by the short run elasticities for the Self and Non-paid modes of return preparation, at 0.374 and 0.055 respectively. The effect of complexity of the tax situation on tax evasion is clearly demonstrated by the negative elasticities for tax return preparer professionals versus the positive elasticities for non-professional modes of return preparation. For those able to purchase tax expertise through a Paid preparer or a Practitioner, increased complexity results in lower tax changes upon audit, while the opposite holds for the Self and Non-Paid modes of return preparation.[18] These results are even more striking when increased complexity is allowed to affect the demand for tax return preparation assistance. Referring to table 5, which shows the elasticities of demand for each of the four modes of return preparation with respect to increases in the number of forms and schedules attendant the tax return, increases in complexity of the tax situation dramatically increase the demand for Practitioners and significantly reduce the demand for Self preparation. When these effects on the level of demand for each of the four modes of return preparation are accounted for, the elasticity of tax evasion with respect to complexity (FORMS) is negative. That is, increased complexity *decreases* the amount of tax evasion, as the final column of table 4 shows.[19]

Finally, increases in the number of situations eligible for penalty assessments (PENALTY)

[18]This result also supports the hypothesis that Practitioners provide an attestation function for the IRS in return for monopoly rents from barriers to entry.

[19]This result is in contrast to that of Erard (1993) who found that the use of a Practitioner increases tax evasion.

increase the amount of evasion detected, with an overall conditional elasticity of 0.435, with most of this coming from the Self and Paid modes of return preparation. Note that the conditional elasticity of tax evasion for Practitioners is insignificant. While these results do not support the hypothesis that increases in penalties provide a deterrent effect to tax evasion, they do suggest that relative to other modes of return preparation, use of a Practitioner does reduce the taxpayer's chance of receiving a penalty. This result supports the hypothesis advanced by Graetz, Reinganum and Wilde (1989) that the ability of Practitioners to issue expert opinions insulates taxpayers from receiving penalties.

2.4.2 Conclusion

I find strong evidence to support the view that tax evasion increases with income (COMPLEX) with an overall elasticity of evasion with respect to amounts of COMPLEX income of 0.771. I find only marginal support for a deterrent effect from penalties. With the incidence of penalty assessments very low in the TCMP data, this may reflect more about the unwillingness of TCMP examiners to impose penalties than the deterrent effect of the penalty regime per se. I also find support for the hypothesis that increased complexity of the tax return decreases tax evasion. This is the result of both the increased demand for Practitioners and the reduction in tax evasion associated with returns prepared by Practitioners. This later feature of Practitioner prepared returns supports the perspective modelled by Nitzan and Tzur (1989) and

Melamud, Wolfson and Ziv (1991) that Practitioners provide an attestation role for the IRS.[20] This result also suggests that policies that would increase the demand for Practitioners, such as a deduction for Practitioner costs from gross income, may be cost effective because of their ability to increase compliance.

[20]Interestingly, this result supports the result of Long and Caudill (1987) that paid tax return preparers lower the reported liability. Our own estimation of the affect of tax return preparation on reported liability confirms their findings. The use of a Practitioner reduces reported liability. (These results are available upon request.) When combined, the picture of Practitioners that emerges is that they 1) decrease reported tax liability, and 2) decrease the amount of tax evasion detected upon audit. When considering the effect of Practitioners on taxes paid, which is the total effect from taxes reported and from a tax audit, Practitioners reduce the variance associated with paying Federal income tax.

Table 2.5: Elasticities of Demand for Mode of Return Preparation

Variable	Self	Non-paid	Paid	Practitioner
SIMPLE	0.345	-0.328	-0.273	-0.255
	(12.397)	(-3.195)	(-6.982)	(-7.033)
COMPLEX	-0.001	-0.226	0.018	0.088
	(-0.003)	(-0.746)	(0.630)	(3.275)
ASSET	-0.005	0.048	-0.117	0.166
	(-0.209)	(0.544)	(-2.945)	(7.867)
EXEMPTION	-0.937	-1.505	1.212	1.224
	(-12.680)	(-7.009)	(11.238)	(10.901)
OVER 65	0.081	-0.459	-0.324	0.602
	(1.416)	(-2.943)	(-3.761)	(6.081)
FORM	-0.624	0.254	0.284	0.972
	(-6.526)	(0.846)	(2.471)	(10.771)
PENAL	-0.182	0.010	0.051	0.386
	(-5.869)	(0.092)	(1.413)	(9.886)
AUDIT	-0.015	0.155	-0.214	0.298
	(-0.111)	(0.871)	(-1.033)	(2.605)

Note: t-statistics in parenthesis.

References

Allenby, Greg and Rossi, Peter. "There is No Aggergation Bias: Why Macro Logit Models Work," *Journal of Business and Economic Statistics* 9 (1991):1-14.

Collins, Julie; Milliron, Valerie and Toy, Daniel. "Tax Preparer Usage and its Implications." Unpublished Manuscript. University of North Carolina, 1988.

Dubin, Jeffrey. *Consumer Durable Choice And The Demand For Electricity*, North-Holland, Amsterdam, 1985.

Dubin, Jeffrey; Graetz, Michael; Udell, Michael and Wilde, Louis."The Demand for Tax Preparation Services." *Review of Economics and Statistics* 74 (1992):75-82.

Dubin, Jeffrey; and McFadden, Daniel. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52 (1984):345-362.

Erard, Brian. "Taxation with Representation: An Analysis of the Role of Tax Practitioners in Tax Compliance." *Journal of Public Economics* 52 (1993):163-197.

Graetz, Michael, Reinganum, Jennifer, and Wilde, Louis. "Expert Opinions and Taxpayer Compliance: A Strategic Analysis." Social Science Working Paper No. 710. Pasadena: California Instituute of Technology, 1989.

Hausman, James; and Taylor, William."Panel Data and Unobservable Individual Effects" *Econometrica* 49 (1981):1377-1398.

Heckman, James J. "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica* 46 (1978):475-492.

Hite, Peggy. "Characteristics of Taxpayers Who Use Preparers: An Exploratory Study." Paper prepared for IRS Research Conference on The Role of Tax Practitioners in the Tax System, Washington D.C., November 16-17, 1987.

Internal Revenue Service. 1990. *Income Tax Compliance Research: Net Tax Gap and Remittance Gap Estimates.* Publication 1415. Washington D.C.

Klepper, Steven and Nagin, Daniel. "The Role of Tax Preparers in Tax Compliance." *Policy Sciences* 22 (1989):167-194.

Klepper, Steven; Mazur, Mark, and Nagin, Daniel. "Expert Intermediaries and Legal Compliance: The Case of Tax Preparers." *The Journal of Law & Economics* 54 (1991):205-229.

McFadden, Daniel. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by A. Karlquist et al. Amsterdam: North-Holland Publishing, 1978.

Melamud, Nahum; Wolfson, Mark and Ziv, Amir."Should Tax Preparation Fees be Subsidized? A Game Theoretic Analysis." mimeo, 1991.

Nitzan, Shmuel and Tzur, Joseph. "Restricted Competition as a Reward for Public Service." *European Journal of Political Economy* 5 (1989):519-532.

Reinganum, Jennifer and Wilde, Louis. "Equilibrium Enforcement and Compliance in the Presence of Tax Practitioners." *Journal of Law, Economics, and Organization* 7 (1991):163-181.

Scotchmer, Suzanne. "Who Profits from Taxpayer Confusion?" *Economics Letters* 29 (1989):49-55.

Scotchmer, Suzanne. "The Effect of Tax Advisors on Tax Compliance." In *Taxpayer Compliance: Social Science Perspectives* pp. 182-199. Philadelphia: University of Pennsylvania Press. 1989.

Slemrod, Joel. "The Return to Tax Simplification: An Econometric Analysis." *Public Finance Quarterly* 17 (1989):3-27.

Slemrod, Joel, and Sorum, Nikki. "The Compliance Cost of the U.S. Individual Income Tax System." *National Tax Journal* 37 (1984):461-474.

Chapter 3

Congressional Elections and the Economy

3.1 Introduction

One of the most influential studies in the history of political science is that of Kramer (1971). It has inspired a large and steady flow of research on economic conditions and political outcomes that shows no sign of slackening. The results of his time series analysis indicate that short-term fluctuations in the economy account for much of the variance in the share of the national vote won by congressional candidates of the incumbent president's party. Most of the explanatory power is associated with a single measure—percentage change in real per capita personal income in the year of the election. Initial reaction was decidedly skeptical, with critics arguing either that economic fluctuations had no detectable influence upon congressional election outcomes (Stigler 1973; Arcelus and Meltzer 1975), or that such effects were confined to the small subset of elections in which economic performance was particularly poor (Bloom and Price 1975). For the most part, however, political scientists discounted these refutations as readily as the electorate would appear to function of presidential approval scores and annual change in real per capita disposable income, the matter seemed settled. Since then, other influential studies have reported strong links between economic fluctuations and the incumbent party's performance in presidential elections (Fair 1978), senatorial and gubernatorial elections (Stein 1990), and in elections held in several advanced industrial democracies (Lewis-Beck 1987). No proposition in political science enjoys a consensus (except, perhaps, for the one just asserted). But the idea that the electoral fortunes of the incumbent party hinge on the vigor of the economy was, as Erikson (1990) puts it, "as widely accepted ... as any hypothesis about elections and voting behavior" (p. 373).

Recently, however, widespread agreement has given way to renewed controversy. Echoing Stigler, Erikson (1990) concludes that economic conditions do not generally affect the choices voters make in congressional elections, and that any observed impact of economic conditions upon election outcomes is more apparent than real. Rosenthal and Alesina (1989) make a similar claim that alternative, non-economic models of congressional elections fit the data as well as Kramer-type models that do. Others, echoing Bloom and Price (1975), argue that economic conditions play an important role in only some types of congressional elections, or in only certain periods of time

(Gough 1984; Radcliff 1988). In short, new questions have been raised about the ubiquity of "economic voting." No longer does it seem wise to simply dismiss Kramer's early critics as engaging in specification search for the worst-fitting equation.

In this paper we hope to shed some light on this controversy by systematically assessing the robustness of both Kramer's findings and those of major critics. There are three major components to this undertaking. First, we seek to guage the robustness of these findings and counter-findings to variations in specification—in the dependent variable in particular. Secondly, we assess the extent to which their findings are produced by a few disproportionately influential sample observations. During the past several years econometricians have become increasingly troubled by the properties of conventional regression techniques. In particular, parameter estimates generated by least-squares regression techniques can be extremely sensitive to the values taken on by a small number of observations on a small number of variables. We apply the technique of least median squares (LMS) as developed by Rousseeuw and Leroy (1987) to test for outliers in the data sets used. This problem is of special concern here, in that the time series studies in this area are based upon a small number of observations.

Thirdly, we investigate the extent to which Kramer's findings and others' counter-findings depend upon the income and unemployment estimates that are used. Kramer (1983) and researchers who followed him have regarded these aggregate economic statistics as particularly hard kinds of data, not like the squishier stuff generated in opinion surveys. Several recent studies (most notably by Romer, 1986a,b, 1988, and 1989; Balke and Gordon, 1989), however, cast doubt on the quality of much of the economic data—specifically, official government estimates of national income, product, and unemployment for the period prior to World War II—upon which Kramer's and subsequent studies are based. On the basis of new data, improvements in the analysis of old data, or different assumptions about the structure of the prewar economy, they derive revised estimates of prewar economic time series that differ significantly from the official statistics. These studies raise the possibility that the effects registered upon election outcomes by economic conditions depend upon which estimates of income and unemployment are used.

I show that new estimates of national unemployment by Romer (1986b) and Darby (1976) and gross national product (put into per capita terms) by Romer (1986a, 1988, 1989) and Balke and Gordon (1989) substantially improve the explanatory power of Kramer (1971) type models of the effects of general economic conditions on congressional elections and provide strong evidence to support the proposition that general economic conditions influence congressional House election outcomes. I also find that the strength of the proposition that general economic conditions effect congressional House election outcomes is significant regardless of whether the dependent variable is specified as the precentage of the popular vote cast for republican candidates as in Kramer (1971) or Erikson (1990) or as the deviation of this percentage from the average of the total votes cast for republican candidates over the preceeding eight elections, as in Tufte (1975). Moreover, regression diagnostics described below consistently identify three elections as being "different" from the remaining elections in the sense that their inclusion in least squares estimation would significantly alter estimates of the size of the effect that general economic conditions have on congressional House elections. Those observations are 1) the 1894 election; 2) the 1912 election, and 3) the election of 1932. Of less significant note, in several instances the elections of 1942, 1944 and 1980 are also identified as observations that exert excessive influence on least squares estimates.

3.2 Robustness to Changes in Specification

The assault on Kramer's results came largely in the form of models with alternative specifications that were maintained to be more sensible, while at the same time showing little if any support for the hypothesis that general economic conditions affect voting behavior. Broadly, these criticisms focused on 1) changes in the independent variables (Stigler, 1973 and Erikson, 1990); 2) changes in the specification of the dependent variable (Tufte, 1975); 3) the inclusion of particular observations (Stigler, 1973) or entire sets of observations (Gough, 1984) the stability of parameter estimates over time (Radcliff, 1988), and across mid-term versus presidential elections (Bloom and Price, 1975 and Campbell, 1985), and 5) the level of data aggregation from national statistics (Kramer, 1971) to state level data (Radcliff, 1988), to county level data (Owens and Olson, 1980) to individual level data (Fiorina, 1978). This research focuses on the first four of these issues.

Stigler (1973) tested the sensitivity of Kramer's variables to alternative constructions and finds that changing the one-year time frame of the economic variables to two years, and many other arbitrary changes in specification, affected the magnitude of the coefficients associated with the economic indicators.[1] In some respects, Stigler's sensitivity testing was more supportive of Kramer's basic findings than is generally realized. in general, these changes in specification affected the magnitude of the coefficients associated with the economic indicators, but they usually continued to exceed conventional levels of statistical significance.

Erikson (1990) identifies a change in specification that makes a dramatic difference in the observed impact of the economy upon congressional elections. Looking first at Tufte's (1975) analysis of postwar midterm elections, he notes that the dependent variable in this study is specified as the difference between the incumbent (president) party's vote share and the average of its vote shares won in the previous eight elections—what Tufte calls the "standardized vote." He finds that the goodness-of-fit of Tufte's model, impressive as it is, can nonetheless be matched by a much simpler regression of the incumbent party's vote share on its vote share in the previous election.[2] On the basis of this and other comparisons, he concludes that "the appro-

[1] In particular, he showed that a two year difference in the unemployment rate, rather than Kramer's one year difference, produced a significant negative coefficient on unemployment changes in models with an incumbency and nominal per capita income measures. Redefining nominal per capita income in a like fashion produced these findings as well. Moreover, shuffling certain years in and out of the specification, most notably the years prior to 1900, for which unemployment data was particularly poor, and 1912, 1918, 1942, and 1944, also produced significant and insignificant coefficients on the unemployment rate.

[2] If one is interested solely in forecasting election outcomes, Erikson's specification is unobjec-

priate control for past voting" is simply the lagged dependent variable. In regressions of this type, fluctuations in real per capita income appear to exert no influence upon the incumbent party's performance in midterm elections. Furthermore, Erikson argues that what at first appear to be effects associated with the income variable in midterm elections are due to spurious correlation with presidential coattails. Once the vote share won by the president is entered into the equation, estimates of the effects of the economy in on-year congressional elections are also indistinguishable from zero.

The next major study to report negative findings concerning economic conditions and elections was that of Owens and Olson (1980) who looked at the 1972, 1974, and 1976 congressional elections. Their research on congressional elections shows them to be primarily local affairs, candidate-oriented rather than party-oriented, and not, as Kramer (1971) had supposed them, close approximations of "the Downsian case of relatively anonymous candidates competing as members of a common party team" (p. 135). In their analysis, they regress changes in the vote for individual congressional candidates upon district-level changes in real income and inflation. In equation after

tionable. This specification does present problems, however, for parameter estimation. The autoregressive model

$$Y_t = \alpha + \beta Y_{t-1} + \epsilon_t, \quad t = 1, 2, \ldots,$$

is stationary only if $|\beta| < 1$ (Dickey and Fuller 1979). If $|\beta| = 1$, the variance of Y_t is $t\sigma^2$ and the time series is a "random walk." If $|\beta| > 1$, the variance of Y_t grows exponentially as t increases. Under these circumstances it is not possible to reliably estimate coefficients for other variables of interest, and that the time series should be transformed, e.g., by differencing. We are indebted to Doug Rivers for alerting us about the properties of estimators when time series have unit roots. The problem of exponentially increasing variance when $|\beta| > 1$ will not necessarily occur in a regression with intermittent observations, such as Tufte's or Erikson's analysis of midterms only. On the other hand, failure of the model to specify that the incumbent party's vote generally "bounces back" during the favorable on-year elections that are ommitted implies misspecification bias.

equation, the regression coefficents of these two terms are either statistically indistinguishable from zero, or, if significant, in the wrong direction.

One obvious limitation of Owens and Olson's study is that it looks only at three elections (1972, 1974, and 1976), and then only at House districts in which an incumbent was running for reelection. The later point may introduce selectivity bias. Jacobson and Kernell (1981) find that a significant share of the impact of economic conditions upon the two parties' national vote shares is due to the strategic decision of incumbents whether or not to seek reelection. Incumbents of the president's party are more likely to run again when they expect a strong economy going into election day, while potentially strong candidates of the opposition party are correspondingly more likely to remain on the sidelines.[3] The situation is reversed when it looks like the economy will be faltering. Owens and Olson's inability to compute district-level economic data for most districts further reduces their sample to less than a third of the districts in each election. Even if we accept their findings, however, it must be kept in mind that they analyze the electoral effects of cross-sectional, inter-district variation in local economic conditions rather than temporal variation in national economic conditions. Rather than implying that the economy "doesn't matter" in congressional elections, their findings can readily be interpreted as corroboration for Kinder and Kiewict's (1979) hypothesis that it is in fact changes in national economic conditions to which voters respond.

[3] Of course the most recent mid term election, in which members of the President's party were defeated wholesale despite the healthy state of the economy, throws cold water on this explanation.

3.3 Disproportionately Influential Observations

Stigler (1973) made several adjustments to Kramer's data. First, he limited anlaysis to years after 1900 because there was no reliable data on unemployment prior to that. Second, he showed in an analysis of post-1934 elections, that including a year Kramer omitted, 1942, while excluding another that he had used, 1946, dramatically reduces the size of the real personal income coefficient. To be sure, changing one observation out of a total of 14 represents a fairly large (over 7 percent) modification in the sample. But this pattern is indicative of the presence of one or more disproportionately influential observations that exert a high degree of "leverage" on estimated regression coefficients. Stigler also discovered that the 11 values in the pre-1920 income series Kramer used were miscoded. When his key equation (reproduced below) was reestimated after correcting these errors, the coefficient of the income term was somewhat smaller, but that of inflation became much larger and statistically significant.

Following a line of argument first advanced by Bloom and Price (1975), Campbell (1985) maintained that incumbents are hurt only by a substantial downturn in economic fortunes during years prior to the mid-term election. Gough (1984) also examined this hypothesis by rerunning Kramer's equation after omitting data from the Depression era, and found that in the remaining "normal" years, the impact of change in real income was indistinguishable from zero. Tufte (1978) argued that Presidential popularity is responsive to prior year economic conditions, and Presidential popularity at the time of the mid-term election does affect the outcome.

Radcliff (1988) maintained that in the years following World War II, Congressional House elections became increasingly uncoupled from general macroeconomic effects. Building on Owens and Olson's (1980) null results about macroeconomic effects of House elections in 1972, 1974, and 1976, he argues that the increased importance of candidate incumbency had made House elections more responsive to local issues rather than national economic conditions.

With the limited number of observations available to analyze the effects of general economic conditions, and with the great variation in political and social contexts surrounding each election, it is tempting to identify subsets of observations that one may believe should be treated differently than the majority of observations. The preceeding literature, while in no manner exhaustive, is indicative of the problem that empirical researchers find themselves in when confronted with a reasonably small number of observations from which to test a hypothesis. That is, small modifications to a general hypothesis invariably produce substantial improvements in estimated models when tested on a few number of observations.[4] In this research I use a relatively new approach towards the identification of outliers and excessively influential observations. Rather than rely on perceived changes in the political or social environment, I use a purely statistical procedure developed by Rousseeuw and LeRoy (1987) known as Least Median Squares (LMS) regression. The strength of this procedure is

[4]Two other approaches are the elimination of particular observations, as Kramer (1971) did with the elections of 1912, 1918, 1942 and 1944, and the use of time varying parameter estimation techniques (Radcliff, 1988). The gains in model efficiency, as measured by regression standard errors and R^2, can be dramatic even from the differential treatment of a single observation.

that it fits the data to the model rather than searching for the model that fits the data.[5]

3.4 Estimates of Income and Unemployment

Macroeconomists have long been aware that observed variance in all major conventional indicators of economic activity, including industrial production, unemployment, and GNP, is only about half as great in the years following World War II than in the four or five decades prior to the war. Most have taken this as gratifying evidence of the success of postwar stabilization policies in tamping down the amplitude of the business cycle in the U.S. economy. Romer (1986a, 1986b, 1988, 1989), however, argues that the greater variance in pre-war indicators of economic activity is instead an artifact of the way in which pre-war estimates of economic activity were constructed.

3.4.1 Unemployment Statistics

The standard unemployment series Kramer used in his analyses comes from two sources. Beginning in 1940, the officially reported unemployment rate has been based upon household survey data collected by the Current Population Survey. Estimates for prior years are those reported by Lebergott (1964). Seeking to make his estimates

[5]Their approach is but one of an emerging set of tools for the applied researcher known as robust estimation procedures. The notion of robustness used here is defined explicitely in the following sections and follows Rousseeuw and LeRoy (1987). While their estimation procedure provides an important diagnostic for ordinary least squares estimation on small data sets, it is not without its weaknesses. For a criticism of least median squares regression, see Hettmansperger and Sheather (1992).

as consistent as possible with the post-1940 figures, Lebergott adopted the Bureau of Labor Statistics' definition of unemployment as the difference between the total number of workers in the labor force and the total number of employed. Estimating an annual series for the former by first calculating labor force participation rates for various demographic groups on the basis of decennial census data, he then aggregated these figures and multiplied the resulting figures by annual estimates of total population. Employment totals were estimated on the basis of a variety of industry-level sources, including the Census of Manufacturing, which was administered every five years prior to World War II, and biennially thereafter. For the years in which these figures were not available, he interpolated figures on the basis of fluctuations in the Shaw (1947) series on commodity output, Kuznets' (1961) GNP estimates for 1919 to 1929, and various industry- and state-level reports.

In making these interpolations, Lebergott assumed that the labor force participation rates he observed in census years remained fixed in intercensal years. According to Romer (1986b), this procedure ignores changes in labor productivity and in the "stickiness" of labor markets across the business cycle.[6]

[6] Romer's estimates are based on the assumption that labor force dynamics in the prewar era were similar to those in the postwar era. In light of the many major changes that have occurred in the American labor force, this may be questionable. One reason why the variance in unemployment rates may have been greater in the prewar era is that the labor force contained large numbers of people who today are considered too young or too old to work. Eighty years ago, employers may have been far less reluctant to shed child and adolescent laborers as they were adult labor in the 1970s. Unionization, which peaked in the 1950s and has fallen ever since, probably also imparted some degree of stickiness in the retention of workers over the course of the business cycle. In addition, the demographics of the post war labor force has changed considerably. In particular, there has been a large increase in the proportion of women 35 to 44 years of age in the labor force, from 39 percent in 1950 to 77 percent in 1992 along with a decline in the participation rates of males from 87 percent in 1950 to 67 percent in 1992 (Kutscher, 1993). The increased participation of women in the labor

More specifically, during periods of recession, the labor force (or at least its rate of growth) contracts, as "discouraged" workers cease job search and so drop out, and new workers refrain from joining in the first place. In a recovery, in contrast, the labor force expands as the number of entrants grows and the number of discouraged workers falls. Incorporating these sources of cyclic behavior in the size of the labor force, Romer (1986b) produces a series of annual unemployment estimates that is a good deal smoother than the Lebergott series.[7] It is, of course, hard to say how these many changes net out.

Another major adjustment to Lebergott's figures is that of Darby (1976), who adjusts Lebergott's series from 1930 to 1943 by shifting workers employed in government programs such as the WPA out of the unemployed category and into the employed. For 1938, the year of the largest adjustment, this increases employment by 3.5 million workers and reduces the estimated unemployment rate from 19.1 to 12.5 percent. Table 1 compares annual changes in the pre-1945 unemployment rate based upon BLS data with the revised estimates made by Romer and Darby.

force would tend to reduce volatility in the official unemployment rate if they are disproportionately more likely to exit the labor force after losing a job. The decline of farm employment, on the other hand, probably increased the volatility of unemployment figures, as farm labor tends to be subject to more underemployment than unemployment.

[7] Emphasizing this point, Balke and Gordon compare the volatility of the prewar to postwar Gottlieb-Commerce construction index with the construction components of the Shaw series and find that the former are significantly more volatile than the Shaw series. Finally, they compare the prewar/postwar volatility of the transportation and communication series with the volatility of the Shaw series and find that this series also is more volatile in the prewar period than the Shaw series. The volatility of the transportation and communication sectors is notable because these are the only annual data that predate 1909.

3.4.2 Income and Product Statistics

Most of his data are the official inflation and unemployment series reported by the Bureau of Labor Statistics, and the annual series on personal income reported by the Department of Commerce, the 1895 to 1929 segment of his personal income series is Kendrick's (1961) GNP estimates for that period, multiplied by a scale factor to approximate personal income (Kramer 1971, p. 143).

The U.S. Department of Commerce began tabulating National Income and Product Accounts in 1929. This work is currently done by a subunit of the Department, the Bureau of Economic Analysis. The BEA estimates gross national product by aggregate personal consumption expenditures, private investment (including net change in inventories), net exports, and government purchases of goods and services. Alternative methods of calculation by the BEA include aggregating by product (final sales and net change in inventories for durable goods, nondurable goods, services, and structures) and by sector (households, institutions, government, farm and nonfarm business, and housing). Personal income is derived by subtracting from GNP capital consumption, indirect business taxes and business transfer payments, corporate profits, contributions for social insurance, and adding personal dividend income, interest income, and government transfer payments to individuals. Alternatively, personal income can be calculated as the sum of wage and salary income, interest and dividend income, transfer payments, and, after capital adjustment, rental and proprietor's income. In principle, each of these methods can be related to each other through accounting identities and should yield identical figures. This never actually occurs, but the extent to which totals arrived at through alternative methods approximate each other is a good gauge of the accuracy and reliability of the BEA's estimates. The amount of information that the BEA processes in order to make these calculations is impressive. They use data generated by the Census Bureau's annual surveys of the wholesale, retail, manufacturing, construction, and service and government sectors, but also receives data on farm income and expenditures from the Department of Agriculture, wage and salary data from the Bureau of Labor Statistics, and federal budget data from the Office of Management and Budget. The BEA collects its own data on international trade and on capital stocks.

Over the years there has been substantial improvement in both the quantity and quality of the data available to the BEA. In previous decades many of the Census Bureau's current annual surveys were conducted biannually or not at all, and the BEA depended to a much greater extent upon data provided by other federal agencies and by state and local governments. Describing the calculation of income and product accounts shortly after World War II, Ruggles (1949) candidly observed that "the Department of Commerce is forced to use ingenious and sometimes roundabout methods of estimating the various components in order to avoid the large gaps that exist in the statistical material. Some of the sources are very much more reliable than others" (p. 104).[8]

[8] One consequence of these ongoing improvements is that every few years the BEA revises many of its product, income, and price series, often all the way back to 1929. In 1985, for example, the entire 1929 to 1982 annual series of National Income and Product Accounts was revised in light of

Although current methods for estimating GNP have often been the target of criticism, the current controversy in the macroeconomics literature centers on BEA estimates of income and GNP prior to the initiation of the National Income and Product Accounts. The official estimates for these years, reported in the *Historical Statistics of the United States, Colonial Times to 1970* and in other sources, are those derived by John Kendrick, an economist at the BEA during the 1950s. These estimates were made on the basis of much sparser data. Kendrick's estimates of GNP for the 1889 to 1908 period are based primarily upon Kuznets' earlier unpublished estimates, which in turn are based primarily upon William Shaw's (1947) annual series on commodity output from the agriculture, mining, and manufacturing sectors. Shaw derived his annual estimates of commodity output by using data gathered in (decadal) census years as benchmarks, and then interpolating values for intervening years on the basis of a variety of industry- and state-level reports. The major improvement Kendrick made was to incorporate much better estimates of government expenditures.[9] For the 1909 to 28 period, the official Commerce Department statistics compiled by Kendrick are based upon a series on consumer expenditures put together by Dewhurst (1947) and upon other secondary data sources. Unfortunately, the Commerce Department

improved estimates of imputed interest paid by life insurance companies, consumer expenditures for goods, consumer purchases of energy, and, most notably, new estimates, derived from the Internal Revenue Service's Taxpayer Compliance Measurement Program, of the extent to which taxpayers under-report income (Bureau of Economic Analysis 1985). Redefining and reclassifying old data also produces changes in the accounts. For example, in the 1985 revisions, medical vendor payments, which had previously been included in the government purchases category, were placed instead under personal consumption expenditures (thus raising estimates of personal income as well).

[9]See Romer (1988), pp. 97 to 98, for details. The sources of the unpublished Kuznets data are not reported.

failed to document precisely how these estimates were derived; according to Romer (1988), "all that researchers know about the creation of the Commerce series is what Kendrick remembers" (p. 94).

Romer (1988) makes a *prima facie* case against the 1908 to 1929 Commerce GNP series by arguing that it is substantially different than Kendrick's (1961) subsequent (and thus presumably superior) estimates. She also offers persuasive evidence that the official Commerce series is flawed by comparing the behavior of disaggregated components of the two series during the severe downturn of 1919 to 1921. The Kuznets (1961) GNP series, upon which the later Kendrick estimates are based, shows a large drop of 22 percent in expenditures on durable goods but a small increase in expenditures on nondurables and on services—precisely the pattern present in postwar recessions. The official Commerce series, in contrast, also shows a large cyclical drop in expenditures on services as well. This is not terribly plausible in her view, and imparts substantially more volatity to the Commerce GNP estimates.

As for pre-1908 estimates of GNP, Romer's (1989) major criticism is of Kuznets' (and thus Kendrick's) assumption that non-commodity components of GNP fluctuate as much as the commodity components, or, more precisely, that deviations from the trend in unobserved non-commodity components of GNP move in a one-to-one relationship with deviations from the trend in the observed, commodity-based component of GNP. Commodity output from the agricultural, mining, and manufacturing sectors constituted about one-third of GNP during this period (Balke and Gordon 1989).

Romer argues that most of the time GNP actually moves much less over the course of a business cycle than does commodity output: "for the postwar era it is widely accepted that the noncommodity components of GNP such as services, trade, and transportation, are much less cyclically sensitive than the commodity component" (p. 94). Kuznets (1961) supports his assumption with a "freehand regression curve" fit to the deviations from trend in GNP on the deviations from trend in commodity output form 1909 to 1938. His eye is actually pretty good, as Romer's computer-assisted regression of GNP onto commodity output during this period came up with an estimate of 0.9. She argues, however, that this result is dominated by the unusual experience of the Great Depression: " ...in very severe depressions services and distribution may collapse as much as commodity output ...This could be due to revisions in estimates of permanent income or to increasingly binding liquidity constraints ...If these kinds of responses do indeed occur ...this could cause GNP and commodity output to genuinely move together one-for-one in such periods." To emphasize how unique the Great Depression was, Romer notes that, "the percentage deviation of production from trend at the trough of the Great Depression was approximately two and a half times as large as the percentage deviation of production from trend in the worst depression of the 1869 to 1908 period." She then estimates the relationship between deviations from the trend in GNP and deviations from the trend in commodity output over two sample periods, 1909 to 1928 and 1947 to 1985. Using a varying coefficients regression model, she estimates the coefficient on deviations in commodity output as 0.58, rather than 0.9 when the Great Depression years are included.

In light of these problems, Romer makes several modifications to the pre-1929 GNP series. First, she replaces the 1909 to 1928 Commerce estimates, based primarily on the Shaw series, with a slightly augmented version of the Kuznets-Kendrick GNP estimates (Romer 1989). These income-payments-based estimates of GNP rely on data from a variety of sources, including the IRS, Census Bureau, and state and industry reports. The advantage of these data is primarily the increased coverage of economic activity. For the pre-1908 years, Romer uses her significantly lower estimate of the covariance between change in commodity output and change in GNP to recompute GNP estimates.

Balke and Gordon's (1989) efforts to improve existing GNP estimates for the pre-1929 era focus on incorporating components of GNP that are not present in the Shaw series, most notably from the construction, transportation, and communication sectors. For years prior to 1915 they use a construction output series assembled by Gottlieb (1965), and splice it to the Commerce Department series on construction output covering 1915 to the present. For the transportation sector, they use Frickey's (1947) index of rail, street, and canal output from 1869 to 1889 and Kendrick's (1961) annual transportation index from 1889 to 1929. These sectors combine to account for nearly 25 percent of GNP prior to 1929.

To show that estimates of GNP based exclusively on the Shaw commodity output series may be misleading, they construct an alternative commodity output series based on the Frickey data, Fabricant's (1942) 1900 to 1932 index of manufacturing output, and the Federal Reserve Board index for manufacturing from 1940 to the present. Interestingly, they report that their new hybrid series, as well as the individual components, displayed *more* volatility than the Shaw commodity output series. In contrast to Romer, Balke and Gordon conclude that reliance on the Shaw commodity output series may impart too little volatility to the resultant GNP series. Further, they show that once the construction, communication, and transportation sectors are added to the estimation of GNP, inclusion of the Great Depression years imparts no extra volatility (they estimate the coefficient of deviations from trend of commodity output on deviations from trend of GNP as less than 0.4, compared to Romer's 0.58). To form new estimates of pre-1929 GNP, Balke and Gordon combine the Gottlieb-Commerce series on construction with the Frickey-Kendrick series on transportation and communication and the Shaw commodity output series, using the same benchmark years to derive trend GNP as Romer (1989).

Balke and Gordon also estimate an entirely new GNP deflator series from the measures of consumer prices compiled by Hoover (1960) and Rees (1961). Unlike the Shaw price series, these series consist of final prices for goods paid by consumers. These data exhibit smoother movements in prices than the conventional GNP deflator series in the years prior to 1908, but more volatile movements during the 1909 to 1928 period, primarily because of the inflation associated with World War I. They then use a combination of the conventional GNP deflator and the Hoover-Rees consumer price index deflator to arrive at the final estimates of real GNP. In contrast to Romer's revised estimates, estimates of annual percentage changes in real per capita GNP based upon the Balke-Gordon series generate a series that is more volatile than the Commerce series as is shown in Figure 2.

3.5 Model

3.5.1 Specification

Kramer (1971) specifies a linear probability model as[10]:

$$\begin{aligned} y_t &= \beta_0 + \beta_1\delta_t + \beta_2\delta_t\left(\frac{R_t - R_{t-1}}{R_t}\right) + \\ &\quad \beta_3\delta_t\left(\frac{P_t - P_{t-1}}{P_t}\right) + \beta_4\left(U_t - U_{t-1}\right) + \mu_t \end{aligned} \tag{3.1}$$

where

y = share of total vote for Republican party in House Congressional elections

δ = +1 if President is Republican

[10]It is well know that a significant shortcoming of the linear probability model is that it can produce probability estimates outside of the interval (0,1). (See, for example, Maddala (1983)). This is especially true when observations for the dependent variable are either 0 or 1 (as in the case with individual data observations) or when the values of the dependent variable approach either 0 or 1 (as in the case with aggregate data observations), because any significant deviation in the independent variables in the forecast period from the values in the period observed could easily result in predicted probabilities outside of the unit interval.

In chapter one of this dissertaion, a log-odds transformation of the dependent variable was specified to map the probabilities from the interval (0,1) to the real line. This mapping implicitely bounds the predicted probabilities onto the interval (0,1) regardless of the values the independent variables may take. There are two reasons why the log odds transformation is not used in this chapter. First, the popular vote for republican party candidates in U.S. House of Representative elections has been well contained within the interval (0.35,0.65) throughout the period from 1894 to 1992. For pracitical purposes, truly extreme variations in the independent variables, much more so than occurred during the Great Depression which is included in the historical record, would be needed to generate predictions outside of this interval. Since the probability of this happening is remote, we need not concern ourselves with it here. Moreover, if such draconian changes were to occur, the overall validation of predictions based on this model would be suspect since such an observation would lie far outside of the range of historical data. Second, to emphasize how 1) alternative estimates of historical data; 2) a longer historical record from which to estimate the model and; 3) new statistical techniques for the detection of outlying observations can verify or dispel Kramer's original formulation I use the same general specification that he used.

$$\begin{aligned}
& -1 \;\; \text{otherwise} \\
R &= \text{real per capita income} \\
P &= \text{cost of living index} \\
U &= \text{Unemployment Rate} \\
t &= t^{th} \;\; \text{year of Congressional election}
\end{aligned}$$

Unlike Kramer, I assume that the residual μ_t follows a first-order autoregressive process such that

$$\mu_t = \rho\mu_{t-1} + \epsilon_t$$

where $|\rho| < 1$ and ϵ_t has mean 0 and variance σ_ϵ^2.[11] This is a more general specification than that used by Kramer (1971) and allows for potential serial correlation in the time series.[12]

Before investigating the sensitivity of new estimates of income and unemployment,

[11] The first order autoregressive process assumes that

$$\begin{aligned}
E(\mu_t) &= 0 \\
E(\mu_t\mu_t) &= \rho\sigma_\mu^2 \\
E(\mu_t\mu_{t-1}) &= \sigma_\mu^2 = \frac{\sigma_\epsilon^2}{1-\rho^2}
\end{aligned}$$

where ρ is the correlation between observations in the year t and $t-1$.

[12] Technically, the correction is for a second order autoregressive process and the rho reported in Tables 1 through 5 represent the second period correlation in the data. However, because we eliminate non-election year observations of the dependent variable, we estimate the model using an AR1 correction.

we incorporate the following alterations into Kramer's original analysis. First, we extended the time series forward from 1964 to 1992 and backward a bit from 1896 to 1892, and include the years that Kramer excluded from his analysis—1912, 1918, 1942, and 1944. These additions add significantly to the empirical record, boosting the number of observations from 31 to 51. In light of the very high correlation that exists between changes in unemployment and changes in per capita real income, we follow Stigler's (1973) and Fair's (1978) practice of including an income term or an unemployment term only.[13]

I also tested two specifications of the dependent variable. The first is defined as the share of total vote for the Repbulican party in House congressional elections, and was used by Kramer (1971), and by Erikson (1990) among others. The second, developed by Tufte (1975), is the deviation of the share of the total vote for the republican party in House congressional elections from the average of the eight preceeding elections for the same. This is intended to measure election outcomes as changes from a party's long term, average voter support.

3.5.2 Regression Diagnostics

In order to check the robustness of the regression results to one or a few highly influential cases, I reestimated the various equations using the regression diagnostic

[13]It may come as a surprise to researchers that historical GNP and unemployment series often use the same underlying sources of data. Most notable among these is the use of the Shaw commodity output series (1944) in Lebergott's (1957) construction of unemployment and all constructions of GNP used herein.

procedures of Rousseeuw and Leroy (1987). Their robust estimation procedure detects outlying observations by an algorithim that minimizes the square of the median residual (LMS or least median square) of the fitted model. A weighted least squares model is estimated by applying a weight of 0 to observations detected as outliers by the LMS algorithim and a weight of 1 to all other observations. Rousseeuw and Leroy denote this special weighted least squares model as least trimmed squares (LTS).[14]

Briefly their procedure is as follows. Let $t = 1894, \ldots, 1992$ denote the years of congressional elections, and let

$$y_t = \beta x_t + \mu_t \tag{3.2}$$

denote the classical linear model with β a $1 \times p$ vector of coefficients and x_t a $p \times 1$ vector of variables defined in equation (1). Define

$$r_t = y_t - \hat{\beta} x_t$$

as the residual for the t^{th} observation where $\hat{\beta}$ is an estimate of β. A least median

[14] LTS should not be confused with trimmed least squares, in which the largest and smallest values of a variable are eliminated (trimmed) from least squares estimation. The choice of weights as 0 and 1 is mainly for convenience in estimation. Other weight schemes could be considered. Rousseeuw and Leroy show that LTS has the same breakdown point as LMS.

square (LMS) solution to the above model is[15]

$$\text{Minimize}_{\hat{\beta}} \left\{ \text{med}_t (y_t - \hat{\beta} x_t)^2 \right\}. \tag{3.3}$$

A measure of the dispersion of the residuals based on the minimal median is the following scale factor

$$s^o = 1.4826 \left(1 + \frac{5}{n-p}\right) \sqrt{\text{med}_t r_t^2(\hat{\beta})} \tag{3.4}$$

where the first terms are a correction for small sample sizes.[16] With this scale estimate, define the standardized residuals

$$\frac{r_t}{s^o}$$

and apply the weight function w_i to the entire data set where

$$w_t = \begin{cases} 1 & \text{if } \left|\frac{r_t}{s^o}\right| \leq 2.5 \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

The weighted least squares procedure that uses these weights is called least trimmed

[15]There is no explicit formula to compute the LMS solution. Rather, a solution to this problem requires examination of a very large number of subsets of the data.

[16]The scale estimate is used in much the same manner as the regression standard deviation is to generate standardized residuals. The difference is that unlike the least sum of squares standard error, this scale estimate provides a robust measure of the dispersion of the residuals. The LMS estimates use this scale estimate to identify outlier observations.

squares (LTS).[17] Rousseeuw and Leroy (1987) show that the LTS estimator has an asymptotically normal distribution.[18]

Statistical inference from the LTS procedure is only approximate because the underlying distribution theory depends in a complicated way on the choice of observations that receive weights of 0 or 1. However, monte-carlo evidence demonstrates that the t-statistics generated in the LTS model are close to their true values.[19] As a final estimation step, I rerun the AR1 correction on the subset of observations selected by the LMS algorithim. This procedure is labelled GTS, for generalized trimmed squares.

A second statistical problem that can seriously affect the stability of least squares estimates is the degree to which combinations of the independent variables can result in the near singularity of the data matrix. To detect ill-conditioned data(The term ill-conditioned is synonymous with near singularity of the $X'X$ data matrix.), I calculate the singular value decomposition for the data matrix X following Belsley, Kuh and Welsch (1980). First, each column of X is scaled by the euclidean norm of the

---

[17] With these weights, a final scale estimate for the LMS model is $\sigma^* = \sqrt{\frac{\sum_{t=1892}^{1992} w_t r_t}{\sum_{t=1892}^{1992} w_t - p}}$ where p is the number of parameters estimated. This final scale estimate is useful as a means of comparison of OLS and LMS residuals.

[18] See Theorem 4 in Rousseeuw and Leroy (1987).

[19] See tables 4 and 5 on pages 212-213 of Rousseuw and Leroy (1987). Moreover, inferences generated from a model with outliers are clearly incorrect because the underlying distribution is no longer gaussian. There are two significant advantages to this procedure. The first is that while the sum of squared residuals is affected by a single outlier, the median residual is unaffected by as many as $\frac{n}{2} - 1$ of the observations being outliers. The second advantage to their procedure is the ability to detect multiple outliers. Other procedures to detect highly influential observations such as the use of diagonal elements of the matrix $X(X'X)^{-1}X'$ (the so called hat matrix described in Belseley, Kuh, Welsch 1980) are appropriate only for the detection of a single outlier. The value of this technique for present purposes is that it fits the model to the data according to a criterion that is far less sensitive than ordinary least squares regression to outlier values that are highly influential.

elements. Then the singular value decomposition of X is defined as

$$X = UDV'$$

where $U'U = V'V = I$, and D is a diagonal matrix with nonnegative diagonal elements called the singular values. The singular values are equal to the square root of the eigenvectors of the matrix $X'X$. The ratio of the largest of these diagonal values to the smallest diagonal value is the condition number of X. Belsley, Kuh and Welch (1980) show that the condition number provides a measure of the sensitivity of the change in the solution of the linear system of equations $Y = \beta' X$ when elements of X and Y change. In particular, the condition number is bounded from below at 1, in which case the columns of the data matrix X are orthonormal. As the condition number becomes large, the data reflect a greater degree of ill-conditioning, implying greater sensitivity of the coefficients β to small changes in either X and Y.[20] The condition number for each of the regressions that use the Tufte and Kramer specifications is 2 while for each of the models that use the Erikson specification which includes a lagged value of the dependent variable are 22. Therefore, ill-conditioning does not appear to be an issue for the models tested.

[20] What constitutes a large condition number is an empirical question. Belsley, Kuh and Welsch present evidence that indexes less than 10 indicate weak dependencies in the data, while those between 30 and 100 indicate strong relations amongst the data.

3.5.3 Data

Several data series are used to estimate the model specified in the previous section. The following five tables report estimates of the equations (1) and (3) using two different dependent variables, three measures of per capita income, two measures of national unemployoyemnt, and one measure of the rate of inflation.

For each of the five tables, the dependent variable in the first three columns is defined as the deviation of the vote share for republican candidates in congressional House elections from the average of the same in the previous eight elections, following Tufte (1975).[21] The first column presents generalized least squares (GLS) estimates of equation (1) on the full data set, which for the Tufte dependent variable contains 44 observations, from 1906 to 1992. The second column presents least median squares (LMS) estimates of equation (3) and the third column presents generalized trimmed squares (GTS) estimates of equation (1) with observations deleted that were identified by the LMS procedure as having excessive influence on the least squares estimates using the cirterion of equation (5). The dependent variable for the next six columns is defined as the share of the total vote for the republican party in House congressional elections, following Kramer (1971). The fourth column presents GLS estimates of equation (1) on the 51 congressional elections from 1892 to 1992. The fifth column presents LMS estimates of equation (3) and the sixth column presents GTS estimates of equation (1) after deleting observations identified by LMS estimation as having

[21] The republican party share of the popular vote in congressional House elections is taken from Ornstein (1993).

excessive influence on the least squares estimates using the criterion of equation (5). The final three columns use the same dependent variable as do the models in columns (4), (5) and (6) but also include as an additional regressor the share of the total vote for the republican party in House congressional elections from the previous election, following Erikson (1990). The seventh column, like the fourth, uses the 51 congressional elections from 1892 to 1992 and presents GLS estimates of equation (1). The eighth column, like the fifth, presents LMS estimates of equation (3) while the ninth column presents GTS estimates of equation (1) after deleting observations identified by the criterion of equation (5) as having excessive influence on the least squares estimates.

The first table presents estimates of equation (1) using per capita income prepared by the Bureau of Economic Analysis as reported in *Historical Statistics of the United States, Colonial Times to 1970* (HSUS) up to 1970. From 1970 to 1992 we use per capita income reported in *Statistical Abstract of the United States: 1993* (denoted SUAS93) and in SUAS82-83. Inflation statistics are derived from the consumer price index as reported in HSUS through 1970 and from the SUAS93 and SUAS82-83 through 1992.[22] Table 2 replaces the pre 1930 BEA per capita income with the

[22]The BLS CPI is series E135-166 in HSUS. The estimates prior to 1913 are from unidentified non-BLS sources; from 1912 to 1919 from surveys of shipbuilding and industrial centers for World War I wage negotiations; and from 1919 thru the present from the BLS Monthly Labor Review. The BEA per capita income series is from series F1 and F3 in HSUS for 1930 onward. Prior to 1930 the GNP series constructed by Kendrick for the Commerce Department was adjusted to match the personal income estimates using the ratio between personal income in 1919 to GNP in 1919, following Kramer (1971). This same adjustment was made to the Romer (1988, 1989) and Balke and Gordon (1989) GNP series as well. These series were then converted to real dollars using the Commerce Deptartment's 1982 base GNP deflator and a new 1982 base GNP deflator for years prior to 1929 constructed by Balke and Gordon (1989).

series derived by Romer (1988, 1989), while table 3 replaces it with the series derived by Balke and Gordon (1989). Table 4 substitutes the Bureau of Labor Statistics historical unemployment series as reported in HSUS for the per capita income series of tables 1 through 3.[23] Table 5 replaces this series with that of Romer (1986b) through 1930; and with Darby's (1976) modifications for 1931 through 1943. The share of the republican vote for congressional House elections, inflation rate, three per capita income measures and two unemployment rate measures are shown in table 6 in the appendix.

Each table also includes a test of the hypothesis, following Radcliff (1988), that congressional House elections have become unresponsive to national economic conditions since the end of World War II. A Chow test of the difference between congressional House elections up to 1946 and congressional House elections after 1946 is represented for each of the generalized least squares estimates (columns one, four, and seven) and for each of the generalized trimmed least squares estimates (columns three, six, and nine).[24]

[23] The BLS unemployment data are from series D85-86 in the HSUS. For years after 1970 unemployment data are taken from SUAS93 and SUAS82-83.

[24] The Chow test is an F-statistic with $(k+1)$ and $n_1 + n_2 - k_1 - k_2$ degrees of freedom formed as

$$\frac{(RSS - URSS)\,/\,(k+1)}{URSS/\,(n_1 + n_2 - k_1 - k_2)}$$

where RSS denotes the residual sum of squares of estimation of equation (1) on the full set of observations; where $URSS$ denotes the unrestricted sum of squares of estimation of equation (1) on the full set of observations where each coefficient is allowed to take on different values for the pre 1948 and the post 1946 sets of observations; where n_1 denotes the number of observations prior to 1948; n_2 denotes the number of observations after 1946; k_1 denotes the number of coefficients estimated on the n_1 set of observations; k_2 is the number of coefficients estimated on the n_2 set of observations and k is the number of coefficients on the full $(n_1 + n_2)$ set of observations.

3.6 Results

The first three tables present estimates of equation (1) by GLS, LMS (using equation (3)), and GTS procedures using per capita income measures derived from BEA, Romer (1988, 1989), and Balke and Gordon (1989) respectively.

Table 1 presents estimates using BEA estimates of GNP. Columns (1), (4) and (7) present GLS estimates on the 51 congressional House elections from 1892 to 1992. These estimates show a small but significant positive effect from changes in per capita income on the republican share of the total vote. The party of the President incumbency coefficient is negative, and significant, indicating (as in Kramer (1971)) a slight anti-incumbency effect. The coefficients on inflation are not significant, but in column (9) following Erikson (1990), the share of the vote for republican party candidates in the prior congressional election is positive and significant. Finally, when the dependent variable is defined as a difference between the current vote for republican candidates in congressional House elections and the average of the preceeding eight congressional House elections, following Tufte (1975) in the first column, the hypothesis that general economic conditions have the same effect on congressional House elections prior to 1946 as after 1946 cannot be rejected. When the dependent variable is defined simply as the share of the vote for the republican party in the current congressional House elections, following Kramer (1971) and Erikson (1990), then this hypothesis can be rejected at the 5 percent level of significance. As with Kramer's (1971) original models, these three models explain slightly more than one third of the variation in the dependent variables.

Least median squares estimates for the preceeding three specifications are presented in columns (2), (5), and (8). Rather than focusing on the estimates, attention should be drawn to the number of observations that the LMS procedure identifies as having excessive influence on least squares estimates. For example, in column (2) LMS estimation identified 1906 and 1918 as outliers when the dependent variable is defined following Tufte (1975), and in column (5), 1894 and 1912, when the dependent variable was defined following Kramer (1971). Including the lagged value of the vote share for republican candidates in congressional House elections as an additional regressor in column (8) increases the number of observations identified as outliers. For the model that incorporates BEA estimates of GNP, the elections of 1894, 1898, 1912, 1914, 1932, and 1948 were identified as outlying observations using the criterion established with equation (5).

Generalized trimmed squares estimates are presented in columns (3), (6), and (9). These three columns show GTS estimates on the subset of observations not identified as outliers with LMS estimation. GTS estimates in columns (3) and (6) show a modest improvement in the overall fit with the coefficients on inflation now becoming significant, and negative, as expected. That is an increase in the price level hurts the incumbent party at the polls. Note that in the last column, GTS estimation shows a marked improvement in the overall fit, with an R^2 of 0.65 and large and highly significant effects from the lagged dependent variable and from inflation, with a smaller, and barely significant coefficient on per capita income. Finally, as we shall see in subsequent tables, one of the casualties of using alternative estimates of GNP and unemployment and of using GTS estimation is the demise of signficance of incumbency of the President's party.

Table 3.1: Estimates With BEA GNP 1892 to 1992

	Tufte			Kramer			Erikson		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	GLS	LMS	GTS	GLS	LMS	GTS	GLS	LMS	GTS
constant	-0.004	-0.003	-0.005	0.490	0.470	0.489	0.330	0.202	0.241
	(-0.009)		(-0.010)	(0.009)		(0.009)	(0.062)		(0.047)
y_{t-2}							0.329	0.592	0.504
							(0.123)		(0.098)
δ	-0.019	-0.005	-0.019	-0.022	0.001	-0.016	-0.018	0.014	-0.001
	(-0.009)		(-0.010)	(-0.009)		(-0.008)	(-0.008)		(-0.001)
inflation	-0.273	-0.152	-0.258	-0.259	-0.625	-0.287	-0.255	-0.701	-0.409
	(0.153)		(-0.163)	(0.156)		(-0.138)	(-0.158)		(-0.114)
income	0.310	0.145	0.315	0.225	0.190	0.227	0.276	0.237	0.146
	(0.100)		(0.100)	(0.104)		(0.087)	(0.105)		(0.081)
R^2	.41		.39	.35		.40	.38		.65
rho	.28		.30	.26		.34			
N	44	42	42	51	49	49	51	46	46
F-statistic Pre-1946/ Post-1946	1.78		1.72	2.38		3.46**	2.54**		0.04

Notes: Standard errors in parenthesis.
** denotes significance at 95 percent confidence level.

In table 2, estimates of per capita income from BEA GNP are replaced with estimates of per capita income using the GNP series created by Romer (1988,1989). Generally, the coefficients of these models are slightly larger than those in table 1, although the levels of significance are similar. Observations identified as outliers using LMS estimation are 1918 for the model using a dependent variable defined as in Tufte (1975) in column (3); 1894, 1912, and 1932 for the basic Kramer (1971) model in column (6); and 1894, 1912, 1914, and 1922 for the model that includes a lagged dependent variable in column (9) following Erikson (1990). The Romer (1988, 1989) GNP based per capita income produces the same conclusions with respect to the hypothesis that the pre 1946 estimates are the same as the post 1946 estimates as with the BEA GNP based per capita income used in table 1. That is, no difference is detected when the dependent variable is defined as a difference from the standardized vote as in columns (1) and (3) but the hypothesis is rejected when the dependent variable is defined as the vote share of the republican party in the current election, as in columns (4), (6) and (7). Finally, the overall fits of these models are indistinguishable from those of table 1.

Table 3.2: Estimates With Romer GNP 1892 to 1992

	Tufte			Kramer			Erikson		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	GLS	LMS	GTS	GLS	LMS	GTS	GLS	LMS	GTS
constant	-0.004	0.015	-0.005	0.490	0.481	0.487	0.337	0.324	0.232
	(-0.009)		(-0.009)	(0.009)		(0.008)	(0.062)		(0.052)
y_{t-2}							0.315	0.337	0.519
							(0.128)		(0.107)
δ	-0.019	0.012	-0.019	-0.024	-0.008	-0.008	-0.019	-0.007	-0.007
	(-0.009)		(-0.009)	(-0.009)		(-0.008)	(-0.008)		(-0.006)
inflation	-0.269	-0.806	-0.259	-0.246	-0.605	-0.418	-0.238	-0.596	-0.315
	(-0.149)		(-0.159)	(-0.152)		(-0.141)	(-0.156)		(-0.117)
income	0.369	0.237	0.368	0.282	0.109	0.226	0.322	0.521	0.285
	(0.111)		(0.111)	(0.117)		(0.103)	(0.118)		(0.090)
R^2	.42		.39	.36		.42	.39		.58
rho	.25		.27	.27		.27			
N	44	43	43	51	48	48	51	47	47
F-statistic Pre-1946/ Post-1946	1.01		1.10	3.51**		6.00**	2.69**		1.98

Notes: Standard errors in parenthesis.
** denotes significance at 95 percent confidence level.

The third table presents estimates using per capita income derived from a GNP series created by Balke and Gordon (1989). In general, these estimates are similar to those of table 2 using the per capita income series constructed with Romer's GNP series. The major difference is in the number of observations identified as outliers using LMS estimation. For the models in the first three columns that use the difference between the standardized vote and the current vote for republican candidates in congressional House elections, LMS estimation identified 1912, 1932, and 1934 as outliers, while for the models in columns (4) through (6) that used the vote share of the republican party in the current election as a dependent variable, LMS estimation identified 1894, 1912, and 1932 as outliers. Interestingly, when the lagged value of the republican vote share is included as an additional regressor in the columns (7) through (9) models, LMS estimation identifies 1894, 1912, 1914, 1922, 1932, 1942, 1946, 1948, and 1980 as outliers. As compared with column (9) of table 2 which is the same model but with per capita income defined from Romer's GNP series, the first four years are the same, but the later years of 1932, 1942, 1946, and 1980 are unique to the Balke and Gordon GNP construction. Since the 1980 value does not change at all in either construction of GNP, this may be cause for some concern about what the LMS estimation is identifying. Keeping in mind that the Balke and Gordon (1989) series is more volatile than the Romer series, what LMS estimation tells us in this instance is that when the additional years of 1932, 1942, and 1946 are deleted, 1980 becomes an outlier as well.[25]

[25] Data sets that contain multiple outliers often exhibit this property, known as masking, in which certain observations are not identified as outliers until other outlying observations are removed. See Atkinson (1986) for a discussion.

Table 3.3: Estimates With Balke and Gordon GNP 1892 to 1992

	Tufte			Kramer			Erikson		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	GLS	LMS	GTS	GLS	LMS	GTS	GLS	LMS	GTS
constant	-0.003	-0.001	0.001	0.491	0.479	0.489	0.313	0.223	0.272
	(-0.009)		(0.131)	(0.009)		(0.008)	(0.061)		(0.050)
y_{t-2}							0.364	0.547	0.440
							(0.126)		(0.102)
δ	-0.019	-0.002	-0.005	-0.023	-0.004	-0.008	-0.018	-0.002	-0.004
	(-0.009)		(-0.006)	(-0.008)		(-0.008)	(-0.008)		(-0.006)
inflation	-0.278	-0.488	-0.441	-0.254	-0.633	-0.427	-0.256	-0.522	-0.459
	(-0.151)		(-0.113)	(-0.154)		(-0.141)	(-0.155)		(-0.130)
income	0.345	0.309	0.314	0.276	0.037	0.237	0.325	0.379	0.290
	(0.103)		(0.078)	(0.111)		(0.094)	(0.112)		(0.118)
R^2	.42		.55	.37		.43	.39		.65
rho	.29		.21	.29		.30			
N	44	41	41	51	48	48	51	42	42
F-statistic Pre-1946/ Post-1946	0.98		3.56**	3.52**		6.62**	3.23**		0.01

Notes: Standard errors in parenthesis.
** denotes significance at 95 percent confidence level.

The fourth and fifth tables exchange the per capita income variable used in tables one, two, and three with two different measures of the national unemployment rate. Table four uses an unemployment rate series constructed by the BLS and reported in HSUS and table five replaces this series for the years prior to 1943 with that of Romer (1986b) through 1930 and Darby (1976) up to 1943. Unlike Kramer's (1971) results, these models show a large and significant negative relation between changes in the unemployment rate and the vote share of republican candidates in House congressional elections. In neither table is the coefficient on the incumbency of the President's party significant.

Turning to table four, columns (1), (4), and (7) show estimates of the three specifications on the full data set. Perhaps the most striking result of this table is the very large negative and significant effect from changes in the unemployment rate on the republican vote share in congressional House elections. The coefficient on inflation is negative and significant in the Tufte (column (1)) and Kramer (column (4)) specifications of the dependent variable, but not in the Erikson (column (7)) model. Each model explains a little more than one third of the total variation in the republican vote share for congressional House elections. LMS estimation identified the election years of 1908 and 1918 as outliers for the Tufte dependent variable in column (2); 1894, 1912, and 1932 for the Kramer specification in column (5); and 1892, 1894, 1898, 1912, 1914, 1922, 1924, 1932, 1948, and 1980 for the Erikson specification in column (8). The GTS estimates that delete these observations show little difference from the GLS estimates on the full data set for the Tufte specification in column (3) and the Kramer specification in column (6). However, for the Erikson specification in column (9), the elimination of ten observations, eight of which are from the pre World War II record, results in a substantial improvement in the model fit, with an R^2 of 0.82, and also leads to the rejection of the hypothesis that the pre 1946 model coefficients are the same as the post 1946 coefficients. As with the first three tables, that used per capita income rather than changes in unemployment rates, the Tufte models of columns (1) and (3) do not reject this hypothesis, while for the most part, the Kramer models of columns (4) and (6) do reject this hypothesis.

Table 3.4: Estimates With BLS Unemployment 1892 to 1992

	Tufte			Kramer			Erikson		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	GLS	LMS	GTS	GLS	LMS	GTS	GLS	LMS	GTS
constant	-0.005	0.011	-0.005	0.489	0.478	0.487	0.334	0.329	0.230
	(-0.009)		(-0.009)	(0.009)		(0.008)	(0.062)		(0.039)
y_{t-2}							0.318	0.337	0.534
							(0.129)		(0.079)
δ	-0.010	0.017	-0.011	-0.016	-0.004	-0.003	-0.013	-0.001	0.007
	(-0.009)		(-0.010)	(-0.009)		(-0.008)	(-0.008)		(0.005)
inflation	-0.307	-0.771	-0.246	-0.870	-0.633	-0.445	-0.263	-0.617	-0.583
	(-0.154)		(-0.166)	(0.157)		(-0.144)	(-0.159)		(-0.092)
Unemp.	-0.782	-0.629	-0.869	-0.550	-0.040	-0.321	-0.656	-0.974	-0.741
	(-0.265)		(-0.277)	(-0.260)		(-0.241)	(-0.251)		(-0.165)
R^2	.38		.37	.34		.38	.38		.82
rho	.19		.19	.22		.22			
N	44	42	42	51	48	48	51	41	41
F-statistic Pre-1946/ Post-1946	0.85		0.67	2.64**		4.85**	2.11		3.15**

Notes: Standard errors in parenthesis.
** denotes significance at 95 percent confidence level.

In table five, replacing the BLS unemployment rate with that of Romer (1986b) and Darby (1976) produces coefficients on unemployment nearly twice as large (and negative) as those of table four. Moreover, the overall fits of the models estimated on the full data set in columns (1), (4), and (7) are larger than any other set of models presented in this paper, and in the case of the Tufte type model in column (1), explains nearly one half of the variance in the difference between the republican share of the vote in the current congressional House election and that of the previous eight elections. LMS estimation identified the election years of 1906 and 1918 as outliers for the Tufte dependent variable, column (2); 1892, 1894, 1896, 1912, 1924, 1926, 1932, and 1980 for the Kramer specification in column (5); and 1894, 1912, 1914, and 1934 for the Erikson specification in column (8). GTS estimation on these modified data sets results in noticeably improved overall model fits, with the Kramer type model of columns (4) and (6) increasing from an R^2 of 0.39 to 0.61, and the models in columns (7) and (9) that use an Erikson type specification increasing from 0.43 to 0.58. Interestingly, with the Romer and Darby unemployment data, only the Kramer type models of columns (4) and (6) reject the hypothesis that the pre 1946 model coefficients are signficantly different from the post 1946 model coefficients.

Table 3.5: Estimates With Romer and Darby Unemployment 1892 to 1992

	Tufte			Kramer			Erikson		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable	GLS	LMS	GTS	GLS	LMS	GTS	GLS	LMS	GTS
constant	-0.001	0.001	-0.001	0.491	0.481	0.485	0.335	0.096	0.258
	(-0.006)		(-0.007)	(0.009)		(0.006)	(0.059)		(0.051)
y_{t-2}							0.321	0.776	0.466
							(0.123)		(0.104)
δ	-0.004	-0.019	-0.005	-0.016	0.004	0.004	-0.013	0.025	-0.004
	(-0.008)		(-0.008)	(0.009)		(0.006)	(-0.081)		(-0.006)
inflation	-0.388	-0.073	-0.328	-0.286	-0.726	-0.621	-0.268	-0.609	-0.295
	(-0.131)		(-0.146)	(-0.149)		(-0.116)	(-0.151)		(-0.117)
Unemp.	-1.555	-0.567	-1.543	-0.959	-0.892	-0.806	-1.280	-1.422	-0.887
	(-0.314)		(-0.326)	(-0.328)		(-0.256)	(-0.308)		(-0.261)
R^2	.49		.47	.39		.61	.43		.58
rho	.06		.04	.23		.17			
N	44	42	42	51	43	43	51	47	47
F-statistic Pre-1946/ Post-1946	1.52		1.13	3.08**		3.69**	2.58**		1.41

Notes: Standard errors in parenthesis.
** denotes significance at 95 percent confidence level.

3.7 Conclusion

In this paper, a variety of modifications to a model of the effects of general macroeconomic conditions on congressional House election outcomes as specified by equation

(1) were tested. First and foremost among our conclusions is that regardless of the construction of alternative data series for GNP and unemployment, increasing the size of the data set on which the various specifications of equation (1) are tested results in strong support for the belief that general economic conditions effect the popular vote in congressional House elections. Briefly, increases in the rate of inflation or in the unemployment rate reduce the proportion of the popular vote cast for republican candidates in congressional House elections while increases in per capita income or in the vote share in the prior congressional election increase the vote share. Moreover, I find no effect from the incumbency of the President's party. These models generally explain between one third and one half of the variation of the republican vote share in congressional House elections. Among the three specifications tested, the most stable in the sense that the fewest outliers are detected using the least median squares estimation is when the dependent variable is defined as the difference between the current vote share for republican candidates in House congressional elections with those of the prior eight congressional elections following Tufte (1975). This is not too surprising since the use of a standardized vote measure tends to modify the influence of unusual election outcomes.[26] By the same token, models specified with a lagged value of the vote share for republican candidates in congressional House elections, as in

[26] Three elections were consistently identified by LMS estimation as outlier observations. They are 1894, 1912, and 1932. The later two have obvious interpretations as unusual elections because of the strong third party challenge in 1912 and the depression in 1932. Interestingly, other researchers have also identified these later years. Kramer (1971) omitted 1912, and Gough (1984) omitted 1932. However, the World War II elections were identified as outliers only once, when the Balke and Gordon (1989) GNP series was used in table 3, and only then for the Erikson (1990) specification in column (9).

Erikson (1990), are the most skittish in the sense that large numbers (as many as ten) of observations are identified by least median squares estimation as outliers. These models also experience the greatest swings in R^2 from a low of 0.58 for models that use the Romer GNP series and Romer and Darby modifications to BLS unemployment data to 0.82 for the model that uses the BLS unemployment series.

The use of revised historical estimates of GNP and unemployment clearly improve the explanatory power of the basic model described in equation (1). The most compelling results for per capita income are contained in table three, which uses a new historical GNP series created by Balke and Gordon (1989) to construct per capita income. The generalized trimmed squares estimates presented in columns (3), (6) and (9) of table three are remarkably stable across each of the three alternative specifications, for all of the variables with the coefficient on the change in inflation at about -0.45 and the coefficient on the change in per capita income at about 0.30. Moreover, models using the Balke and Gordon (1989) GNP data have R^2's of 0.55, 0.43 and 0.65 for the Tufte, Kramer, and Erikson specifications respectively. When the models are fitted with new historical unemployment data from Romer (1986b) and Darby (1976), again there is a sizeable and consistent improvement over the standard historical unemployment series, as reported in table five. The generalized trimmed squares estimates presented in table 5 report coefficients on the change in inflation of between -0.3 and -0.6 while the coefficients on the change in the rate of unemployment vary between -1.5 and -0.8.

I have mixed results on the test of the hypothesis that general economic conditions prior to 1946 had a different effect on the proportion of the vote for republican candidates in congressional House elections than for the years after 1946. The only model that uses a Tufte type dependent variable, in column (3) of each table, that rejects this hypothesis is when the Balke and Gordon (1989) GNP series is used (in table 3) to construct per capita income, and the only model that uses the Erikson (1990) type construction, in column (9) of each table, which rejects this hypothesis is when the BLS unemployment data are used, in table 4. On the other hand, generalized trimmed squares estimation of Kramer type models, in column (6) of each table, reject this hypothesis every time.

References

Alesina, Alberto and Rosenthal, Howard. 1989. "Partisan Cycles in Congressional Elections and the Macroeconomy." *American Political Science Review* 83:373–98.

Arcelus, Francisco and Meltzer, Allan. 1975. "The Effect of Aggregate Economic Variables on Congressional Elections." *American Political Science Review* 69:1232–39.

Atkinson, A.C.. 1986. "Masking Unmasked." *Biometrika* 73:533—41.

Balke, Nathan and Gordon, Robert. 1989. "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy* 97:38–92.

Belseley, David; Kuh, Edwin and Welch, Robert. 1980. *Regression Diagnostics.* New York: John Wiley and Sons.

Bloom, Howard and Price, Douglas. 1975. "Voter Response to Short-run Economic Conditions: The Asymmetric Effect of Prosperity and Recession." *American Political Science Review* 69:1240–54.

Born, Richard. 1986. "Strategic Politicians and Unresponsive Voters." *American Political Science Review* 80:599–612.

Bureau of Economic Analysis. 1985. *Survey of Current Business*, July.

Campbell, James. 1985. "Explaining Presidential Losses in Midterm Congressional Elections." *Journal of Politics* 47:1140–57.

Darby, Michael. 1976."Three-and-a-Half Million U.S. Employees Have Been Mislaid: Or, an Explanation of Unemployment, 1934–1941." *Journal of Political Economy* 84:1–16.

Dewhurst, Frederic and Associates. 1947. *America's Needs and Resources* New York, Twentieth Century Fund.

Dickey, David and Fuller, Wayne. 1979. "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." *Journal of the American Statistical Association* 74:427–31.

Erikson, Robert. 1990. "The Puzzle of Midterm Loss." *Journal of Politics* 50:1012-29.

———. 1990. "Economic Conditions and the Congressional Vote: A Review of the Macrolevel Evidence." *American Journal of Political Science* 34:373–99.

Fair, Ray. 1978. "The Effect of Economic Events on Votes for President." *Review of Economics and Statistics* 60:159–73.

Fiorina, Morris. 1978. "Economic Retrospective Voting in American Elections." *American Journal of Political Science* 22:426—43.

Frickey, Edwin. 1947. *Production in the United States, 1860–1914.* Cambridge, MA: Harvard University Press.

Gallman, Robert. 1966."Gross National Product in the United States, 1834–1909." In *Output, Employment, and Productivity in the United States after 1800.* (Studies in Income and Wealth, v. 30., Conference on Research in Income and Wealth.) New York: Columbia University Press.

Gottlieb, Manuel. 1965. "New Measures of Value of Nonfarm Building for the United States, Annually 1850–1939." *Review of Economics and Statistics* 47:412–19.

Gough, Paul. 1984. "Economic Conditions and Congressional Elections: An Attempt to Control for the Depression." *American Politics Quarterly* 12:71–88.

Hettmansperger, Thomas and Sheather, Simon. 1992. "A Cautionary Note on the Method of Least Median Squares." *The American Statistician* 46:79–83.

Hoover, Ethel. 1960. "Retail Prices after 1850." In *Trends in the American Economy in the Nineteenth Century.* (Studies in Income and Wealth, v. 24, Conference on Research in Income and Wealth.) Princeton, NJ: Princeton University Press.

Jacobson Gary. 1990. "Does the Economy Matter in Midterm Elections?" *American Journal of Political Science* 34:400–5.

———, and Kernell, Samuel. 1981. *Strategy and Choice in Congressional Elections.* New Haven, CT: Yale University Press.

Kendrick, John. 1961. *Productivity Trends in the United States.* Princeton, NJ: Princeton University Press.

Kiewiet, D. Roderick. 1983. *Macroeconomics and Micropolitics.* Chicago: University of Chicago Press.

Kinder, Donald and Kiewiet, D. Roderick. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting." *American Journal of Political Science* 23:495–527.

Kramer, Gerald. 1971. "Short-Term Fluctuations in U. S. Voting Behavior, 1896–1964." *American Political Science Review* 65:131–43.

———. 1983. "The Ecological Fallacy Revisited: Aggregate- Versus Individual-Level Findings on Economics and Elections and Sociotropic Voting." *American Political Science Review* 77):92–111.

Kuznets, Simon. 1961. *Capital in the American Economy.* Princeton, NJ: Princeton University Press.

Lebergott, Stanley. 1964. *Manpower in Economic Growth: The American Record Since 1800.* New York: McGraw-Hill.

Lebergott, Stanley. 1986. "Discussion". *Journal of Economic History* 46:367–71.

Lewis-Beck, Michael. 1988. *Economics and Elections: The Major Western Democracies.* Ann Arbor, MI: The University of Michigan Press.

Maddala, G.S. 1977. *Econometrics.* New York: McGraw-Hill.

Maddala, G.S. 1983.*Limited Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Markus, Gregory. 1988. "The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-sectional Analysis." *American Journal of Political Science* 32:137–54.

Owens, John and Olson, Edward. 1980. "Economic Fluctuations and Congressional Elections." *American Journal of Political Science* 24:469–94.

Ornstein, Norman, Malbin, Michael and Mann, Thomas. 1994. *Vital Statistics of Congress 1993 to 1994.* Washington D.C.: American Enterprise Institute.

Pearson, F. A. and Myers, William . 1948. "Prices and Presidents." *Farm Economics* 163:4210–18.

Radcliff, Benjamin. 1988. "Solving a Puzzle: Aggregate Analysis and Economic Voting Revisited." *Journal of Politics* 50:440-58.

Rees, Albert. 1961. *Real Wages in Manufacturing, 1890–1914.* General Series, no. 70. Princeton, NJ: Princeton University Press.

Rivers, R. Douglas. 1986. "Microeconomics and Macropolitics." Social Science Working Paper 486, California Institute of Technology.

Romer, Christina. 1986a. "Is the Stabilization of the Postwar Economy a Figment of the Data?" *American Economic Review* 76:314–34.

———. 1986b. "Spurious Volatility in Historical Unemployment Data." *Journal of Political Economy* 94:1–37.

———. 1988. "World War I and the Postwar Depression: A Reinterpretation Based on Alternative Estimates of GNP" *Journal of Monetary Economics* 22:91–115.

———. 1989. "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869–1908." *Journal of Political Economy* 97:1–37.

Rousseeuw, Peter and Leroy, Annick. 1987. *Robust Regression and Outlier Detection.* New York: John Wiley and Sons.

Ruggles, Richard. 1949. *National Income and Income Analysis.* New York: McGraw-Hill.

Shaw, William. 1947. *Value of Commodity Output Since 1869.* New York: National Bureau of Economic Research.

Stein, Robert. 1990. "Economic Voting for Governor and U.S. Senator: The Electoral Consequences of Federalism." *Journal of Politics* 52:29–53.

Stigler, George. 1973."General Economic Conditions and National Elections." *American Economic Review* 63:160–7.

Tufte, Edward. 1975. "Determinants of the Outcomes of Midterm Congressional Elections." *American Political Science Review* 69:812–26.

Tufte, Edward. 1978. *Political Control of the Economy.* Princeton: Princeton University Press.

Tufte, Edward. 1978. "Evidence Selection in Political-Economic Studies: The Unusual Distribution of Published R^2."

U.S. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970* Government Printing Office, Washington, D.C.

U.S. Bureau of the Census. 1982*Statistical Abstract of the United States: 1982-83* Government Printing Office, Washington, D.C.

U.S. Bureau of the Census. 1993*Statistical Abstract of the United States: 1993* Government Printing Office, Washington, D.C.

3.8 Appendix

Table 3.6: Data: 1892 to 1929

Year	Share of Total Vote for Republican Party	BLS CPI 1967=100	Real Per Capita Income Constant $ 1987 BEA	Romer	Balke & Gordon	Unemployment Rate BLS	Romer & Darby
1892	0.414	27.0	2510	2490	2520	3.0	3.7
1893		27.0	2380	2450	2470	11.7	8.1
1894	0.618	26.0	2250	2360	2350	18.4	12.3
1895		25.0	2470	2490	2580	13.7	11.1
1896	0.567	25.0	2310	2440	2480	14.4	11.9
1897		25.0	2530	2580	2630	14.5	12.4
1898	0.456	25.0	2580	2660	2640	12.4	11.6
1899		25.0	2820	2810	2890	6.5	8.7
1900	0.511	25.0	2920	2950	2910	5.0	5.0
1901		25.0	3150	3080	3200	4.0	4.6
1902	0.493	26.0	3180	3150	3190	3.7	4.3
1903		27.0	3220	3200	3220	3.9	4.4
1904	0.538	27.0	3130	3160	3280	5.4	5.1
1905		27.0	3370	3330	3510	4.3	4.6
1906	0.507	27.0	3660	3530	3590	1.7	3.3
1907		28.0	3650	3570	3460	2.8	3.6
1908	0.497	27.0	3330	3390	3210	8.0	6.2
1909		27.0	3780	3660	3520	5.1	5.1
1910	0.465	28.0	3690	3670	3460	5.9	5.9
1911		28.0	3740	3680	3510	6.7	6.3
1912	0.340	29.0	3910	3820	3660	4.6	5.3
1913		29.7	3980	3900	3730	4.3	4.9
1914	0.569	30.1	3600	3730	3380	7.9	6.6
1915		30.4	3680	3960	3460	8.5	7.2
1916	0.537	32.7	4280	4340	3960	5.1	5.6
1917		38.4	4250	4390	3910	4.6	5.2
1918	0.569	45.1	4540	4520	4160	1.4	3.4
1919		51.8	4110	4090	4020	1.4	2.9
1920	0.642	60.0	3950	3930	3880	5.2	5.2
1921		53.6	3710	3690	3680	11.7	8.7
1922	0.517	50.2	3940	3920	3880	6.7	6.9
1923		51.1	4380	4360	4350	2.4	4.8
1924	0.555	51.2	4370	4340	4380	5.0	5.8
1925		52.5	4410	4380	4420	3.2	4.9
1926	0.569	53.0	4630	4600	4620	1.8	4.0
1927		52.0	4580	4550	4580	3.3	4.6
1928	0.565	51.3	4650	4620	4610	4.2	5.0
1929		51.3	4740	4740	4740	3.2	4.6

Table 3.7: Data: 1930 to 1966

Year	Share of Total Vote for Republican Party	BLS CPI 1967=100	Real Per Capita Income Constant $ 1987 BEA	Romer	Balke & Gordon	Unemployment Rate BLS	Romer & Darby
1930	0.526	50.0	4320			8.7	8.9
1931		45.6	4010			15.9	15.3
1932	0.414	40.9	3440			23.6	22.5
1933		38.8	3290			24.9	20.6
1934	0.461	40.1	3440			21.7	16.0
1935		41.1	3760			20.1	14.2
1936	0.441	41.5	4240			16.9	9.9
1937		43.0	4340			14.3	9.1
1938	0.514	42.2	4030			19.0	12.5
1939		41.6	4330			17.2	11.3
1940	0.487	42.0	4520			14.6	9.5
1941		44.1	5170			9.9	6.0
1942	0.539	48.8	6170			4.7	3.1
1943		51.8	7300			1.9	1.8
1944	0.494	52.7	7770			1.2	
1945		53.9	7740			1.9	
1946	0.558	58.5	6480			3.9	
1947		66.9	5970			3.9	
1948	0.481	72.1	6040			3.8	
1949		71.4	5890			5.9	
1950	0.510	72.1	6290			5.3	
1951		77.8	6620			3.3	
1952	0.503	79.5	6840			3.0	
1953		80.1	7020			2.9	
1954	0.471	80.5	6860			5.5	
1955		80.2	6990			4.4	
1956	0.494	81.4	7130			4.1	
1957		84.3	7150			4.3	
1958	0.434	86.6	7190			6.8	
1959		87.3	7230			5.5	
1960	0.448	88.7	7330			5.5	
1961		89.6	7430			6.7	
1962	0.475	90.6	7610			5.5	
1963		91.7	7770			5.7	
1964	0.428	92.9	8080			5.2	
1965		94.5	8405			4.5	
1966	0.480	97.2	8733			3.8	

Table 3.8: Data: 1967 to 1992

Year	Share of Total Vote for Republican Party	BLS CPI 1967=100	Real Per Capita Income Constant $ 1987 BEA	Real Per Capita Income Constant $ 1987 Romer	Real Per Capita Income Constant $ 1987 Balke & Gordon	Unemployment Rate BLS	Unemployment Rate Romer & Darby
1967		100.0	9035			3.8	
1968	0.482	104.2	9346			3.6	
1969		109.8	9581			3.5	
1970	0.445	116.3	9659			4.9	
1971		121.3	9696			5.9	
1972	0.464	125.3	10006			5.6	
1973		133.1	10500			4.9	
1974	0.405	147.7	10480			5.6	
1975		161.2	10260			8.5	
1976	0.421	170.5	10550			7.7	
1977		181.5	10850			7.1	
1978	0.447	195.4	11280			6.1	
1979		217.4	11500			5.8	
1980	0.480	246.8	11570			7.1	
1981		272.4	11650			7.6	
1982	0.433	289.1	11490			9.7	
1983		298.4	11640			9.6	
1984	0.470	311.1	12180			7.5	
1985		322.5	12530			7.2	
1986	0.446	328.1	12820			7.0	
1987		340.1	13190			6.2	
1988	0.455	355.4	13600			5.9	
1989		372.5	13890			5.2	
1990	0.450	392.7	14030			5.4	
1991		409.2	13820			6.6	
1992	0.456	421.5	14040			7.3	