Three Episodes in Nineteenth Century United States Banking and Finance

Thesis by

Christopher Hoag

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2003
(Submitted May 2, 2003)

Acknowledgement

I would like to thank my dissertation advisor, Professor Davis, for allowing me to study under his guidance. Thanks to my committee members Professors Bossaerts, Hoffman, and Komunjer for their suggestions and their careful comments. Thanks to Professor Grether for sharing his statistical expertise on many occasions. Thanks also to Professors Fohlin, Jackson, and Sherman for their help.

Thanks to the staff of Millikan Library: Ms. Brown, Ms. Sustaita, Ms. Nollar, and Ms. Peyvan, for their help with obtaining and reading historical documents. Thanks to Ms. Auchampaugh, the graduate secretary. Thanks also to my colleagues, past and present, in the office.

Special thanks to Ms. Mirjana Orovic and Mr. Norm Nelson of the New York Clearing House, for allowing me to collect the data for use in Chapter 3 and in subsequent papers.

"I want to thank all those who made this night necessary."

\- Yogi Berra

Abstract

This dissertation samples three episodes from nineteenth century United States history that conveniently illustrate economic behavior in the arena of banking and finance.

The first chapter considers improvements in cross-market arbitrage due to technological change. The completion of the undersea Atlantic telegraph cable in July 1866 more closely integrated securities markets on two continents. Chapter 1 conducts an event study on one security with a dual listing on the New York and London Stock Exchanges using daily data. The event study provides some evidence that the information lag between the two markets shortened from ten days to zero days. We can recover transatlantic steamship crossing times from securities prices.

The second chapter investigates bank window dressing. Window dressing is a temporary change in portfolio designed to produce a more appealing report to regulators or to the public. Market observers accused national banks of window dressing after the Civil War. Chapter 2 attempts to determine whether or not postbellum Philadelphia banks window dressed their balance sheets. A test finds some evidence for window dressing.

The third chapter conducts an econometric test of Diamond and Dybvig's (1983) theory of bank runs as interpreted by Calomiris and Gorton (1991). Diamond and Dybvig employ an exogenous liquidity shock to depositors in order to develop a theory of bank runs. Calomiris and Gorton interpret the exogenous liquidity shock as a seasonal withdrawal from the nation's agricultural interior. Chapter 3 reexamines the hypothesis that a seasonal interior reserve drain served as the exogenous liquidity shock before the bank panics of 1873 and 1893 in the United States. Using individual bank level data in New York, this paper tests whether the banks that held most of the deposits from the interior, the "interest-paying" banks, experience reserve drains just before the panic. The evidence reveals that a seasonal interior drain could have triggered the 1873 panic, but that Diamond and Dybvig's model cannot be applied to the bank panic of 1893 without a non-seasonal interpretation of the exogenous liquidity shock.

Contents

List of Figures

List of Tables

Introduction

Economic history employs historical data to test economic theory. History provides a useful vantage point for economic analysis. Economic history permits observation of economic agents in different regulatory and institutional environments. As a consequence, historical events produce a rich series of natural experiments. This dissertation samples three such episodes from the nineteenth century United States that conveniently illustrate economic behavior in the arena of banking and finance.

The first episode, Chapter 1, considers the behavior of markets when subject to an exogenous change in the structure of informational flows. In 1866, engineers constructed an undersea telegraph cable across the Atlantic Ocean. The telegraphic connection allowed immediate access to transcontinental information without a long ocean voyage. One immediate application of faster information transmission is to transcontinental arbitrage across securities markets.

The first chapter considers market efficiency. Did market participants take advantage of the new information present in the market? Did the introduction of the telegraph allow for closer pricing of securities on the two markets? From the point of finance theory, we expect investors to arbitrage away price differentials. Markets should become more closely integrated when more precise information is available. This chapter inquires about these issues by comparing daily security pricing of a dual-listed security on the New York and London Stock Exchanges before and after the introduction of the Atlantic telegraph.

The advantage of the historical episode is that we have multiple price transactions while information travels between markets. In modern markets, marginal improvements in the speed of information transmission often exceed the rate price quotation. In this case, the introduction of the telegraph decreased the information lag by vessel travel time across the Atlantic (about ten to fourteen days), while stock markets quote daily prices on securities during this period. So, we can observe information as it travels from one market to another. In fact, markets priced the securities so well that we can recover transatlantic crossing time from securities quotations.

The second episode, Chapter 2, investigates bank window dressing. Window dressing is a temporary change in portfolio designed to produce a more appealing report to regulators or to the public. Window dressing obscures the health of individual banks by systematically altering underlying balance sheets. Bank regulators show concern about window dressing because mismeasured data will bias tests and may lead to policy errors. In order to learn more about window dressing, we turn to an historical example.

Just after the American Civil War, market observers accused national banks of window dressing their balance sheets. The National Bank Act of 1864 required national banks to submit reports of condition on the first Monday of each quarter. Because banks knew about the quarterly reporting day in advance, they had an incentive to acquire extra reserves around the reporting date in order to appear as if they regularly held a more conservative portfolio. Prominent bank regulators argued that window dressing by banks produced systematic quarterly stringencies in national money markets.

The historical example permits the investigation of window dressing, which is normally difficult to detect. A direct test of window dressing requires two distinct simultaneous observations on bank reserves. Often, the only available measurement of bank reserves is provided to a bank regulator. Due to a fortunate historical coincidence, postbellum Philadelphia retains two reported series of reserves on roughly the same set of banks. We can perform a direct test for bank window dressing by comparing these two series and evaluating their consistency with each other.

The third episode, Chapter 3, examines how bank crises begin. Bank panics were a regular occurrence in the nineteenth century in the United States, so history is a convenient opportunity to study bank runs. During this period, banks in the agricultural West and Midwest deposited reserves in money centers such as New York. Sprague (1910) describes how a certain class of banks, called "interest-paying" banks, specialized in accepting these deposits and held most of the balances. Deposits from the interior suffered regular seasonal fluctuations as agrarian bankers withdrew their funds in order to meet seasonal agricultural requirements. During periods of seasonal stress, such as the autumn harvest, large amounts of these deposits were recalled to the agricultural interior. Occasionally, these episodes endangered the solvency of "interest-paying" banks and ended in bank crises. The leadership of the New York Clearing House Association, an important financial institution during the period, recognized the vulnerability of the interest-paying banks and attempted to regulate the payment of interest on demand deposits many times.

The third chapter tests a theory of bank crises using the historical example of seasonal withdrawals. Diamond and Dybvig (1983) employ an exogenous liquidity shock to depositors in order to develop a theory of bank runs. Calomiris and Gorton

(1991) interpret the exogenous liquidity shock as a seasonal withdrawal from the nation's agricultural interior. Calomiris and Gorton evaluate the seasonal hypothesis, but these authors do not provide any statistical evidence for their claims. This chapter performs an econometric evaluation of Calomiris and Gorton's thesis by looking for seasonal withdrawals on interest-paying banks just before panics struck the New York money market. According to the interpretation of the theory, interest-paying banks should suffer a large exogenous shock as a bank run begins. We can test whether or not seasonal withdrawals on interest-paying banks accompanied the onset of financial crises.

With the addition of primary data collection, the method of inquiry is standard to applied economics. Economic historians turn to archives to provide information about past economic events. Once the data have been gathered, econometric tools verify whether or not observed behavior conforms to the predictions of economic theory. Economic historians use the same approach and methods as modern applied economists: we identify a test of the theory, gather the appropriate data, and conduct an econometric evaluation. The only difference is that historical data is a just little dustier than modern data.

We earn two lessons from the exercise. The first lesson reminds us that the theory, however elegant, may or may not accurately represent the real world, and so the theory must be tested. The assumptions about the environment, the description of interaction among agents, and the model of how the agents behave may all conspire against the application of the theory. The second lesson is that viewing theory through history is a useful device. On certain occasions, historical events align to provide potential economic insight. Here we collect three such episodes for the use and the enjoyment of the profession.

"Lombard Street and Wall Street talked with each other as two neighbors across the way."

Henry M. Field, quoted in Friedman and Schwartz ([1963], p. 26).

Chapter 1

The Atlantic Telegraph Company completed an undersea telegraph cable across the Atlantic Ocean in 1866. Prior to the telegraphic connection a minimum eight day voyage by steamship linked the capital markets of the United States and Europe. Thereafter information flowed across the Atlantic at the speed of an electrical impulse.

Chapter 1 considers the change in capital market information flows due to the exogenous technological advance of the telegraph. The introduction of the Atlantic Cable offers an excellent opportunity to investigate a change in the structure of information. Modern marginal increases in the speed of information transmission are small relative to the rate of price quotation. Instead, the telegraph shortened information transmission time by the length of an ocean voyage across the Atlantic, about ten to fourteen days. Since securities markets quote daily prices during this period, we can obtain multiple price observations between information shocks. Thus, we can observe information originating in one market arriving in another.

Did the Atlantic Cable increase market efficiency? Semi-strong efficient markets use all past public information to set current equilibrium prices. One component of public information is the price history on another market. Prices in highly integrated markets should converge to prevent arbitrage opportunities. If financial markets took advantage of new information and transactions costs were sufficiently low, then prices set after the introduction of the telegraph should reflect one-day-old information by telegraph instead of at least eight-day-old information by steamship.

Chapter 1 employs an event study in order to recover a change in the informational structure from stock prices. The event study examines how long price differentials between New York and London persist after large price changes in New York. The event study provides evidence from securities prices that the telegraph decreased the information lag from ten days to zero days. That is, we can recover transatlantic steamship crossing times from securities prices.

Introduction

Did market participants take advantage of technological advances? Previous work demonstrates that the introduction of the telegraph coincided with price convergence between the New York and London Stock Exchanges. Garbade and Silber (1978) attribute an increase in price convergence to the introduction of the Atlantic Cable on July 27, 1866. The authors analyze the performance of US 5-20 1862 federal bonds using daily data from April–July 1866 and October–November of 1867, 1868, and 1871. They find a difference between the mean absolute deviations of the pre- and post-telegraph periods of about 2.7 gold dollars (significant at the 0.01% level) or about 7% of the value of the stock.

The authors' analysis is biased toward finding market integration. Garbade and Silber (1978) base their pre-telegraph estimate of price divergence on three months of data. Unfortunately, these months bracket the May 1866 Overend, Guerney and Co. banking crisis in London. This severe shock to the London market caused high price differentials between the two markets when asset prices in London fell. The high price differential during the crisis overstates the lack of integration before the introduction of the telegraph.

In another study, Michie (1987) claims that the telegraph increased market integration between New York and London. Michie (1987, p. 46–7) compares the performance of two railroad stocks in 1860 to a government bond in 1870. He observes a decrease in the price differential between the two markets and an increase in intraday price spread overlap.

Michie's test is also biased toward finding more integration. Neal (1992a, b) describes 1860 as a "worst case" [Neal (1992b, p. 12)] year for examining market integration. The election of Abraham Lincoln precipitated a small stock market panic when Southerners withdrew funds from Northern banks, again creating large price differentials between the two markets. Further, the comparison of two volatile stocks to a bond may generate price convergence independent of telegraphic information flows.

Both studies use only the timing of the change in regime to make the causal inference. Many other events during the period can explain observed price convergence since the introduction of the Atlantic Cable occurs nearly simultaneously with other major shocks. In addition to the Overend, Guerney crisis, the introduction of the telegraph coincides with a decrease in market uncertainty following the American Civil

War. While previous work is very suggestive that the telegraph caused price convergence, perhaps we can obtain more conclusive evidence.

The direction of cross-market information flows is also of interest. Garbade and Silber (1979) framed the question of whether or not a satellite foreign market exerts an influence on a dominant home market. Not only did information about fundamentals and speculative behavior originating in New York affect the London price, but events on the Continent could affect the New York price. Information about wars between European powers, the impending arrival of gold or securities shipped from the Continent, and the British reaction to American political events signaled foreign demand for U.S. securities. Studies of modern cross-listed firms (such as Hauser, Tanchuma, and Yaari [1998], Lau and Diltz [1994], and Lieberman, Ben-Zion, and Hauser [1999]) conclude that often the home market dominates, but that sometimes price effects can be detected in the other direction. This chapter confirms the result that information flows from the home market to the foreign market. Reversing the event study shows that most information flowed from New York to London only. There does not appear to be any significant influence of the London price series on the New York price.

Price convergence of a liquid dual listed security may be an important precursor of full market integration. Papers by Eun and Shim (1989) and Lin, Engle, and Ito (1994) detect cross-market information flows with market indices. In general, historical financial markets have shown a high degree of integration when compared to their modern counterparts. Using data on the full market, Chabot (2000) argues that the New York and London securities markets may have been highly integrated between 1866-

1885. Neal (1990) presents evidence that cross-listed share prices during the 17th and 18th centuries were remarkably well integrated.

Data

In order to investigate the decrease in the information lag, I consider a single security listed on both the New York and London Stock Exchanges. Securities with a dual listing are susceptible to arbitrage operations. Paper assets are preferred to commodity assets because bulky items suffered sizable shipping costs.

This study examines the $100 common shares of the New York and Erie Railroad. Erie was one of the few dual listed securities that had a consistent series of observations both before and after the introduction of the Atlantic Cable on both sides of the Atlantic. According to an 1865 estimate, 60% of Erie common shares were in foreign, primarily English, hands (*New York Times*, October 28, 1865). Erie also had enough significant price jumps to make an event study feasible. Contemporary observers considered Erie common shares, the "Scarlet Woman of Wall Street," one of the most volatile securities on the New York market. Speculators (including the "Speculative Director" and longtime treasurer of the company, Daniel Drew) subjected Erie to a series of pools, raids, and corners during the period [Adams, Jr. and Adams, (1956)]. The end of the data set witnesses the first chapter of the infamous Erie War, including a hostile takeover attempt by Cornelius Vanderbilt.

Two newspapers, the *New York Times* and the London *Times*, contain observations on Erie common shares from August 1864 to July 1868. The *New York Times* prints daily sales of shares on the New York Stock Exchange between 10 to 12 AM. The London *Times* prints both a list of sales for the day and a range of closing prices. We use the average of the closing price spread because data on trades in London became thin toward the extremes of the data set. Also, London closing prices have the advantage of trading roughly simultaneously with the New York market. During this period, the London market during this period closed at 4 PM, which was 11 AM New York time [Michie, (1987, p. 73)].

Comparing prices on the two markets requires a currency conversion. First, we convert U.S. paper dollars ("greenbacks") to U.S. gold dollars. Gold dollars were then converted to pounds at the rate of a 60-day bill of exchange discounted by the London interest rate. Pounds were then converted back to dollars at a customary par exchange. Finally, we adjust the London price for forward trading as most trades were made for a fortnightly settlement day. Schmidt (1875) employs a simpler version of this procedure in his directive for arbitrageurs. Appendix 1A contains additional information about the data set and the conversion procedure.

Design

As previous research has shown, price convergence occurs near the time when the telegraph was introduced. Figure 1.1 shows the daily price differential between the converted New York price and the London price for Erie shares. A visible decrease in the price differential occurs shortly after the introduction of the telegraph on July 27, 1866, denoted by the center black line. The price differential in London drops from an average of 5% before the telegraph to about 2% afterwards. In fact, the difference in the mean absolute deviation between the two periods was 1.61 (= 2.34 - 0.725) in London, slightly less than 5% of the value of the security. Under an independence assumption, the difference is significant at better than the 1% level (t = 18.68). This result replicates those of Garbade and Silber (1978) and Michie (1987).

However, we expect to find an increase in price convergence because domestic market volatility decreased over the period. Both the Erie price and the gold premium are less volatile after the introduction of the telegraph. F tests of the ratio of the variance of the absolute price differential before the cable to the variance of the absolute price differential after the cable for both Erie shares in New York and the premium on the price of gold in greenback dollars were significant at better than the 1% level (Erie F = 3.57, gold F = 91.23). While Erie has only a slight decrease in the magnitude of the first difference, the gold price becomes much less volatile by the end of 1865.

The change in underlying volatility weakens the statement that the telegraph caused price convergence. Because of the changes in the volatility of Erie and the price of gold before the cable was introduced, we already expect to find convergence. An exogenous decrease in the variance of the security will lead to a smaller mean absolute deviation between the prices of the two markets. A simple before and after test may misattribute causation to the telegraph. Large market shocks that strike before the introduction of the telegraph (such as the Overend-Guerney crisis) will appear to generate price convergence even if the telegraph had never existed. The event study strengthens the inference about the telegraph's causal effect by demonstrating the change in the structure of information.

Historical facts complicate detecting a decrease in the information lag. Information flows between the two markets were distributed stochastically over lags of several different lengths. Before the telegraphic connection information transfer was limited to ships that carried mail over the Atlantic. The speed of the information transmission depended on the speed of the steamship, which could vary due to the vessel's design or the weather. Further, seasonal variation in shipping demand also affects the frequency of information transmission. Vessels carrying new information neither left nor arrived every day. Finally, technological advances in recording and transmitting messages caused telegraphic response times to decline over the period. Delays and outages occasionally occurred when cables broke.

Event studies often rely on a regular series of known events such as dividend announcements. Erie suspended paying dividends in 1866.[1] Instead, we allow a sufficiently large price change in the converted New York price to indicate the presence of new information in the domestic market. Perhaps sufficiently large price changes indicate a change in the underlying fundamentals. We use the converted price because it includes new information based on the share price, the gold price, and the exchange rate.

1. A popular comment declared "The three certainties in life are death, taxes, and no dividends on Erie common."

Define an event day as a change greater than 2.35 percent of par in the converted price of Erie. The threshold 2.35 was chosen to yield approximately 50 events. Small changes in the event threshold do not alter the results.

The event study makes information more detectable in the presence of transaction costs. In a study of modern cross-listing, Neumark, Tinsley, and Tosini (1991) suggest that price reactions must exceed a threshold transaction cost in order to register on another market. Transactions costs by cable were formidable: telegraphing 20 words including name, address, and signature from New York to London cost 21 pounds sterling (about $100) until November 1, 1866 when the price fell by half. Partially offsetting the cost of transmission, the press in both countries reported the prices of prominent securities (including Erie) traded on the other market. Instead of attempting to directly estimate transactions costs, the event study helps to overcome transaction cost barriers by conditioning on a subset of the most volatile events. Large price changes in one market, presumably indicating a change in fundamentals, would be the most visible to another market and the most likely to be arbitraged.

Arbitrageurs employed several strategies. Prior to the introduction of the telegraph, arbitrage mostly consisted of being the first investor to trade based on new information from the other market. Suppose new information increased the New York price. Armed with the new information, an arbitrageur would travel by steamship to London. On arrival, the arbitrageur would quickly buy the security. The arbitrageur would then sell back to the market later at a higher price after the London market had digested the news. Once communication by telegraph was possible, more elaborate schemes were available.[2] Again, suppose new information increased the price in New York. In one common tactic, "selling to arrive," a New York investor sold shares forward ten or fifteen days, usually with the option to deliver at a time of the seller's choice. The investor cabled a confederate in London with orders to purchase shares on the London exchange and send them by steamship to New York. If multiple trades did not offset the partners' accounts, remissions could be made by steamship or by cable. Note that speculating in shares was not exactly a theoretical textbook case of arbitrage. Speculators accepted small risks of garbled messages or large price changes during the brief interval that telegraphic orders traveled.

Test

We conduct an event study to detect a change in the structure of information. The event study measures how long a price differential persists between the two markets as a result of a large price change in New York. Suppose the converted price of Erie in New York reacts to new information. The change in the converted price might occur through alterations in the gold price, the exchange rate, or the stock price. Before the telegraph, the London market’s ignorance of the new information will create a price differential between New York and London. The price differential should persist for the time it takes for new information to travel across the Atlantic. After the introduction of the cable, the

2. Actually, arbitrageurs used the same strategy after the telegraph was in place. Detectives exposed a ring of telegraph operators who smuggled Associated Press financial reports to prominent Wall Street contacts prior to publication (*Philadelphia Inquirer*, July 3, 1867). Wall Street brokers were thus able to trade based on the telegraphic information before it became public through the press.

price differential created by a converted price change in New York should disappear within a few days.

Consider a formalization of the event study. Recall that an event day was defined as a change in the converted New York price of greater than 2.35 percent of par. Number each event day i before the telegraph from 1 to N and each event day j after the telegraph from 1 to M. Each event is a 19-day sequence of price differentials beginning on the event day and ending eighteen days after an event day. Number each day in event time t from 0 to 18 relative to the event day at time t = 0. Denote the New York price converted by the foreign exchange conversion described in Appendix 1A by p_{it}^{NY}. Denote the London price of event i at time t by p_{it}^{L}.

Now construct the series of price differentials. We want to extract price differentials between the two markets resulting from an event in New York. While the absolute price differential $\left|p_{it}^{NY} - p_{it}^{L}\right|$ is a viable candidate, it requires standardization because the two markets do not always have exactly the same prices. Suppose that the two markets already price the security differently before the event began. When an event occurs, we cannot attribute all of the price difference between the two markets to the occurrence of the event. So, we standardize the price differential by subtracting out the price differential of the day before the event day. If no price exists for the day before event day, we use the price difference of one of the three days prior to the event day, in order, and we disallow the event if no price differential can be found within three days.

Hence, denote the standardized absolute price difference between the converted price in New York and the London price during event i at event time t before the telegraph by

$$c_{it} = | p_{it}^{NY} - p_{it}^{L} - (p_{i,-1}^{NY} - p_{i,-1}^{L})| \tag{1.1}$$

and the standardized absolute price difference on event j at event time t after the telegraph by

$$d_{jt} = | p_{it}^{NY} - p_{it}^{L} - (p_{i,-1}^{NY} - p_{i,-1}^{L})| \, . \tag{1.2}$$

We now have N slices of price differentials of length 19 from before the introduction of the telegraph and M slices of price differentials of length 19 from afterward.

Event studies standardize the returns during an event period by the mean of the returns from a pre-estimation period. Deviations from a zero mean will inflate the test statistics (Brown and Warner [1985]). Therefore, price differentials for both the before and after regimes are standardized by their respective means. Let $\bar{c}$ (respectively $\bar{d}$) be the mean of the standardized absolute price differentials before (respectively after) the telegraph that did not fall into an event period. Let

$$x_{it} = | p_{it}^{NY} - p_{it}^{L} - (p_{i,-1}^{NY} - p_{i,-1}^{L})| - \bar{c} \quad \text{and} \tag{1.3a}$$

$$y_{jt} = | p_{it}^{NY} - p_{it}^{L} - (p_{i,-1}^{NY} - p_{i,-1}^{L})| - \bar{d} \tag{1.3b}$$

be the mean-adjusted standardized same day price differentials for the periods before and after the introduction of the telegraph, respectively. Event studies typically estimate the means $\bar{c}$ and $\bar{d}$ from a pre-estimation period. Instead of using a pre-estimation period, we estimate these two quantities with the population means of the entire sample.

Consider the distribution of the standardized price differentials. Suppose that before the Atlantic Cable

$$E(x_{it}) = \mu_t \text{ and } \mathrm{Var}\,(x_{it}) = s_t^2, \tag{1.4}$$

while after the introduction of the telegraph

$$E(y_{jt}) = \nu_t \text{ and } \mathrm{Var}\,(\nu_t) = t_t^2, \tag{1.5}$$

where μ_t and ν_t are the cross-section population means while s_t^2 and t_t^2 are the population variances. Hence, for both regimes and for a given event time t, both the mean-adjusted standardized price differentials and their respective variances are constant across events.

Form cross-sectional averages for each t. Fix a time lag t after an event and define the cross-sectional averages over the events by

$$\bar{x}_t = (1/N_t)\sum_{i=1}^{N_t} x_{it} \qquad \text{and} \qquad \bar{y}_t = (1/M_t)\sum_{j=1}^{M_t} y_{jt}\ , \tag{1.6}$$

where N_t and M_t are the number of observations of x_{it} and y_{jt} respectively. The number of points N_t or M_t in the cross-sectional average can vary due to the number of prices that did not exist when one market or the other was closed. For large enough samples, the central limit theorem yields

$$\bar{x}_t \sim N(\mu_t,\ s_t^2/N_t) \qquad \text{and} \qquad \bar{y}_t \sim N(\nu_t,\ t_t^2/M_t), \tag{1.7}$$

if x_{it} and y_{jt} are independent for a fixed t. Observe that values of x_{it} and y_{jt} may be independent across i and j if events are spaced sufficiently far apart in time even if the daily time series is correlated.[3]

---

3. An alternate method allows the conditional variance to be transmitted across markets. Future research

Now derive a test statistic for each event time t. For each t, let the null hypothesis be that the standardized mean price differential is the same before and after the telegraph: H_o: $\mu_t = \nu_t$. Base a test statistic on

$$z_t = (\bar{x}_t - \bar{y}_t)/\sqrt{((s_t^2/N_t) + (t_t^2/M_t))}\,. \tag{1.8}$$

Under the null hypothesis, z_t is asymptotically normally distributed. The actual distribution solves the Behrens-Fisher problem of drawing from two distributions with unequal variances. In practice, statisticians employ a t distribution as a rough approximation.

Let v be the travel time across the Atlantic Ocean for information-bearing vessels. The minimum travel time from Boston to Liverpool was 8 days 5 hours.[4] Since not every ship made the journey across the Atlantic that quickly, v is probably between eight to fourteen days. Most vessels required between ten to fourteen days.

Let c be the time it takes to transmit information by cable. There was only a one-hour overlap when both the London and New York markets were open together (Michie [1987], p. 73), so we expect c to be between zero to two days. It seems unlikely that c would be larger than one, although Sundays, holidays, and cable outages could delay the receipt of news. After the connection, information ought to be transmitted across the Atlantic with perhaps a lag of c days as opposed to a lag of v days.

The cross-sectional averages $\bar{x}_t$ and $\bar{y}_t$ reflect the change in the structure of information flows. Before the introduction of the telegraph, the cross-sectional averages

---

should investigate this possibility.

4. For more information about transatlantic crossing times, see http://www.blueriband.com.

$\overline{x}_t$ should reflect information transmission time of v, vessel travel time. News originating in New York should take v days to reach London. Suppose that an event occurs in New York. The New York price will react to this information while the London market remains ignorant. Hence, for times $t < v$, we should observe a significant difference between the New York price and the London price, so $\overline{x}_t$ will be significantly positive. For information lags longer than v, the London market finds out about the New York event. So for $t \geq v$ we expect the cross-sectional averages $\overline{x}_t$ to be insignificantly different from zero. After the telegraph, the cross-sectional averages $\overline{y}_t$ reflect information arrived by cable. For times $t < c$ after an event, again a significant price difference should exist as the New York market knows something that the London market does not. Hence, $\overline{y}_t$ is positive. For event times $t \geq c$, the information has traveled and there should be no difference. For $t \geq c$ $\overline{y}_t$ should be insignificant different from zero.

Test 1 subtracts the cross-sectional averages in order to obtain the distribution of z_t. Test 1 uses the t statistics to check whether or not the informational structure changed. For event times $t < c$ we should observe no difference between the mean deviation before the cable and the mean deviation after the cable. No information has been transmitted yet. For event times $t \in [c,v)$ in the post-telegraph period London could have received price information about the event by telegraph, whereas for event times $t \in [c,v)$ in the pre-telegraph period price information has not yet arrived. Thus, for $t \in [c,v)$ a price differential will exist in the pre-telegraph London market while a price differential will not exist in the post-telegraph London market. This price differential will persist until news arrives by ship in London (until $t \geq v$). Hence, under the hypothesis that the telegraph caused the entire price difference, for each event time $t \in [c,v)$ the difference between the mean absolute price difference before the cable and the mean absolute price difference after the cable should be statistically significant.

The event study's estimate of v is probably underestimates the average duration of vessel travel time. Recall that information flows are distributed stochastically around v due to fluctuations in maritime conditions. The t statistics for shorter information lags close to v are more likely to be significant than longer ones. Both fast and slow information-bearing vessels contribute to earlier z_t but only slow vessels contribute significantly to later z_t. Hence, later lengths of lag are less likely to be significant.

Standard event studies estimate an appropriate variance of the test statistics by using data from a pre-estimation period. This procedure requires that the variance does not change during the event window. Instead, we must allow the variance to differ during the event window. Following the example of earlier studies, Brown and Warner ([1985] p. 24) suggest consider alternate variance estimates. This event study estimates the variances of the test statistics by directly calculating the variance of the cross-sectional statistics. That is, we estimate $\mathrm{Var}(\bar{x}_t)$ and $\mathrm{Var}(\bar{y}_t)$ from the existing sample realizations of x_{it} and y_{jt} used to calculate $\bar{x}_t$ and $\bar{y}_t$. These statistics are more conservative than simply using the variances estimated from the entire series.

Clustering impedes event study methodology. In standard event studies described by Brown and Warner, clustering is the simultaneous occurrence of events across securities. In our case, clustering takes the form of one event occurring within the event window of another event. Event clustering is an important possibility in our data because the events are defined endogenously by large price changes. Clustering invalidates independence across test statistics and may also cause the tests to have incorrect size (Brown and Warner, Table 8, p. 21).

A quick estimate of the bias to the coefficients due to clustering examines the z_t statistics for large t. Under both the null and the alternative, for large t the z_t statistics should be zero. For t greater than vessel travel time, information has traveled both by steamship and by telegraph. Mean-adjusted price differentials before and after the introduction of the telegraph should be approximately the same. Hence, we can use the t statistics for large t as the true mean of the z_t under the null hypothesis with event clustering. Bernard (1987) discusses other remedies.

We can also run the test in the reverse direction. We can use the same test to determine whether or not the London market transmitted news to New York. The test statistics of the event study remain the same, and only the definition of the event changes. We would like to define a London event as large price change in the converted price of the London security as it appears in New York. Unfortunately, daily historical dollar-sterling exchange rates are not available in London. The lack of data limits the event study to defining an event in London as a large change in the actual London price of the security. This method omits information revealed by the adjustment of the foreign exchange rate. The omission decreases the power of the test, as we are less likely to identify news in London that could affect the price in New York.

Results

The data generated a distribution of events. Recall that we define an event as a change in the converted New York price of Erie of greater than 2.35 percent of par. This definition yields 49 events. Of these, 29 events fell before and 20 events fell after the introduction of the Atlantic Cable. The theory requires a sufficiently calm market between events. Events must not overlap or else new information may obscure the arrival of previous information. Unfortunately, the events clustered together. Of the 49 events, 34 fell within a week of some other event. For example, five events fell during the seven trading days April 4–11, 1865, when the American Civil War drew to a close. Overlapping events obscure the time lag of the market reaction. Also, clustering may invalidate the assumption of cross-sectional independence.

Table 1.1 displays the results of Test 1, the absolute value event study. Table 1.1 contains a list of t statistics on the variable z_t, one for each time t after an event. The t statistics examine the difference between the mean-adjusted absolute deviations before and after the telegraph for a given length lag t after an event. Each t statistic has approximately between 20 and 44 degrees of freedom.

The event study detects the London market reacting to price changes in the New York market. Table 1.1 presents evidence that vessel travel time, v, was at least nine days and perhaps ten. Table 1.1 shows positively significant differences at the 5% level (two-tailed test) for event times t = 0 through 9, although t = 4, 9, and 12 are marginally significant. Since we observe significant z_t statistics for about eight or nine days, this result supports the hypothesis that the telegraph shortened the information lag by about nine or ten days. Because z_t for event time zero was significant, cable transmission time, c, seems to be only zero days. Information travel on the same day seems fast. Since the prices are roughly simultaneous, price information must travel while both markets are open.

In the reverse direction, there does not seem to be an effect of London price changes on the New York market. Table 1.2 confirms previous research that information generally flows from the home market to the domestic market, but not the other way around. We cannot detect information traveling from London to New York by this method. The only significant t statistics are t = 0, 1, and 2. Again, cross-sectional variances close to the event day were somewhat lower than other variances in the event window, partially explaining the large significant t statistics on these days. However, the poor results could be a consequence of defining an event as a change in the London security price and not the converted price.

Steamship crossing times generally confirm the analysis. The *New York Herald* (January 15, 1865) catalogues the transatlantic passage of several major steamship lines for the year 1864. While certainly other vessels made the voyage, these lines held the postal contracts and were probably among the faster steamers. In fact, the financial press mentioned several of these ships by name during the year as transporting news, specie, and securities. Consider a typical voyage between New York to Liverpool. The average travel time, based on about seventy voyages of two steamship lines, was 11.9 days; westbound was 13.2 days (where five hours was subtracted from eastbound traffic and added to westbound traffic in order to account for the time difference). The fastest westbound steamers proceeded in about nine or ten days. Recovering an information transmission of ten days from the event study seems fast given that average travel time was somewhat longer. Ships did not leave port every day, and most ships were slower than the fastest ship. But since information-bearing vessels probably traveled faster than average, recovering an information lag of about ten days from stock prices roughly accords with observed steamship travel time.

We can test the robustness of the result with a joint test for significance. Simes' (1986) modified Bonferroni test evaluates the joint significance of a set of p-values. Consider T different null hypotheses H_1 to H_T with size α. We want to test the joint null hypothesis $H_0 = \{H_1, \ldots, H_T\}$. For each test compute a p-value p_t. Rank the T p-values from least to greatest, $p_{(1)}$ to $p_{(T)}$. The original Bonferroni test rejects the null hypothesis H_0 if $p_{(1)} \leq \alpha/T$. The original Bonferroni test is quite conservative. Simes' modified Bonferroni test rejects the null hypothesis H_0 if for any i $p_{(i)} \leq i\alpha/T$. Simes proves that the test has correct size for independent uniform distribution and performs Monte Carlo simulations that suggest the modified test may deliver increased power at roughly nominal size for other distributions, even under correlation.

The joint tests confirm the result. In our case, H_t is that the t statistic associated with the test statistic z_t at event time t is statistically significant. The joint hypothesis H_0

is that all T of the t statistics are statistically significant, where $T = 9$ because nine was the last significant t statistic from the event study. Both the original Bonferroni test and Simes' modified Bonferroni test reject the null hypothesis H_0 for $\alpha = 5\%$ for the event study that the London market reacted to information from New York.[5]

We might worry that the test is slightly misspecified. As seen in Table 1, event time statistics of length 15 through 18 are positive. Under both the null and the alternative, both test statistics should be zero. One explanation is that some information bearing vessels did take more than two weeks to bring information across the Atlantic. Suppose that vessels leave port twice a week and that new information originates in New York just after the previous vessel leaves port. In this case, a fast ship that was delayed a few days en route could take more than two weeks to bring information to London. But the effect is present for greater event times $t = 17$ and 18. It is unlikely, although possible, that transatlantic voyages were so long. Another explanation is that positive test statistics could also result from the clustering of events. The price differential that falls at a large event i at time t (relative to event i) might also be the price differential for another event j at a smaller event time $t' < t$ (relative to event j). Thus, test statistics could be slightly inflated.

---

5. Alternatively, we can estimate the change in the structure of information using cointegration analysis. First, fill missing prices by linear interpolation. Regress the London price on leads and lags of the New York price. A residual-based test indicates a cointegrating relation. Estimate the cointegrating vector by Stock and Watson's (1993) DGLS. Hypothesis tests on the cointegrating relation indicate that before the telegraph was in place, information traveled from New York to London in ten to thirteen days, but no significant relationship can be detected from London to New York. After the telegraph was in place, information traveled from New York to London in zero to two days, but there is only limited evidence that information traveled from London to New York.

Even if the test were slightly misspecified, the qualitative results remain similar. Suppose we subtract the mean of the last four t statistics (event times t = 15 to t =18) from the t statistics from event times t = 0 to t = 14. The mean of the t statistics for event time t = 15 to t = 18 was 0.64. After this adjustment, Table 1.1 shows that t statistics for event times 2 and 6 remain significant at the 5% level, while t statistics for event times 1, 5, 7, and 8 are marginally significant at the 10% level. The Bonferroni (and Simes' modified Bonferroni tests) do not reject the null hypothesis that the t statistics for event times t = 0 to t = 8 are jointly significant at the 5% level, but they both reject at the 10% level. So we might infer that cable transmission time was perhaps between zero and one days, and that the fastest information-bearing ships traveled between nine and ten days.

Changing the definition of an event in New York does not radically alter the results. First, we can alter the threshold of the definition of an event. Recall that an event was defined as a price change in the converted price of Erie of greater than 2.35. We can increase or decrease the threshold to generate less or more events. Adjusting the threshold of a price jump to add or subtract 5 events does not change the significance of the t statistics. Further, we can discount events resulting from the Union victories just prior to the end of the Civil War. While a few t statistics become marginally significant (four remain significant at the 5% level, while four more t statistics remain significant at the 10% level), the estimate of vessel travel time does not decrease below nine days even when Civil War events are omitted.

Conclusion

Chapter 1 studies the change in capital market information flows between New York and London due to the introduction of the Atlantic Cable. The evidence from an event study suggests that the information lag shortened from ten days to zero days. The telegraph had an immediate impact on the price convergence of one security. Because London prices quickly incorporated newly available price information from New York, the London market may be semi-strong form efficient.

Appendix 1A

Appendix 1A describes the data set. The data consist of daily prices from the *New York Times* and the London *Times* for the period of August 1864 to July 1868 for the $100 common shares of the New York and Erie Railroad. Both papers quote prices in terms of percentages in the local currency. We select Erie because it was the only dual listed security that had both a relatively consistent series of data and enough volatility to make an event study feasible. Daily data on the full New York and London markets during this period are currently unavailable.

In New York, the data consist of daily sale prices for the First Board of the New York Stock Exchange. These prices quote the actual registered sales transactions concluded by member brokers taking place roughly during the hours of 10–12 AM. While the purpose was to obtain trades exactly synchronized with the London market, the trades could have taken place anywhere within this time frame. The price quoted was the last listed spot price on the first call of the First Board. As the list of trading prices follows a continuous pattern, it seems reasonable to assume that they are listed in chronological order; using the average price of spot price sales on the First Board does not change the results. On the rare occasion when there was no trade on the First Board, we substitute a price from the Second Board (in the afternoon), the closing price of the day, or the 10 AM Open Board of Brokers.

The London data includes two series of prices. The first is the list of actual trades throughout the day. Actual trades are less frequent for the beginning and the end of the data set. The second price series is a closing price spread. Unfortunately, if no trade took place on a particular day then the closing price spread often matched the spread of the previous day. This suggests that if there was no market activity, then the London Times did not update the closing price spread and left the old numbers in print. In this case, no market activity is indistinguishable from identical activity on successive days. Because of the paucity of actual trades at the beginning and end of the data set, we used the average of the closing price spread. While it is common for financial studies to examine only price data, we might be interested in sales volume. Unfortunately, my sources do not record London volume for American securities.

Comparing prices in New York and London requires currency conversion (Michie, [1987], p. 61). Both papers report prices as percents of par in the domestic currency. In order to compare New York and London prices, we convert the New York price into the London price. First, greenback dollars were converted to gold dollars. During this period the US was on the paper greenback standard. Gold dollars traded at a premium to paper dollars. As the gold premium could fluctuate throughout the day, we average the daily high and low from Mitchell (1966). Second, we convert gold dollars to English pounds at an exchange rate of based on the price of a sixty-day bill of exchange in New York adjusted for the interest component. The exchange rate fluctuated around the par exchange rate of 4.87 gold dollars per pound. Finally, London firms quoted US securities at a customary exchange rate of 4.44 (Davis and Hughes, [1960], p. 55). This is also the procedure employed by Schmidt (1875).

Data limitations forced an approximation in the conversion procedure. Examining the London reaction to New York price changes involves converting at the London exchange rate. Unfortunately, we lack daily information on the exchange rate in London. In the absence of such information, we simply converted at the New York rate. Thus, the converted New York price incorporates an exchange rate set in New York that is as old as the stock price information instead of the current London dollar-pound exchange rate.

The exchange rate data derives from issues of the *New York Herald* and the Merchants' Magazine, supplemented by the *New York Times* and *Tribune*. Leading bankers provided buy-sell quotations on sixty-day bills of exchange. Bills of exchange served as the primary instrument of foreign payments, although merchants occasionally substituted Erie shares or US federal bonds. Exchange rates varied across papers due to the heterogeneous quality of issuers and endorsements as well as the lack of a centralized market. Quotations on shorter duration bills were only available during peak exchange movements.

In order to obtain a spot exchange rate I discounted sixty-day bills of exchange. As noted by Perkins (1978) and Officer (1985) and confirmed by Schmidt (1875), the correct discount rate is the London rate. We obtain a monthly series of English interest rates on three-month bankers' bills in Great Britain from the NBER's website (series m13016 at http://www.nber.org/databases/macrohistory/contents/chapter13.html). Linear interpolation created a daily rate. Since the interest component of the bill depends on future expectations of interest rates, we gave arbitrageurs perfect foresight and used the arithmetic average of the discount rate over the duration of the bill. In the pre-telegraph period, we allow arbitrageurs knowledge of the current English interest rate even though this information was not available in the market. American quotations of the English interest rates were too infrequent to permit easy construction of a consistent daily series. Schmidt ([1875], p. 202) states that arbitrageurs allowed fifteen extra days of interest to account for the delay in presentation at London, so I followed this practice as well.

A final adjustment corrected for both markets quoting forward prices. NYSE rules allowed traders to deliver stocks by 2:15 PM the next day. Thus, prices quoted in New York were one-day forward trades. We make no adjustment for the trades in New York. In contrast, by the mid-nineteenth century the London market was a time market, and nearly all trades were settled on a fortnightly settlement day. To account for the forward trades, we discount the London prices by the interpolated market rate of interest. In general, the price change was less than a tick size ($1/8^{th}$).

The data also had to be adjusted for dividend payments. Erie common shares paid three 4% dividends during this period. Often one market would begin quoting the security without the dividend while the other market still included it. Although more complicated methods would yield a more accurate calculation, we simply adjusted the price by the converted magnitude of the dividend. This omission ignores at least two factors. First, the value of the dividend should be discounted at a short-term interest rate when quoted on only one market. Second, dividends payable on the continent were most likely received by ship later than dividends payable in New York. All three dividends occurred before the telegraph connection. The lack of adjustment of interest rates may add to the pre-telegraph variability. Changes in quotations due to the timing of dividend quotations were not counted as events.

Figure 1.1: Erie Common Share Price Difference, 1864-8
Converted NY Price Minus London Price

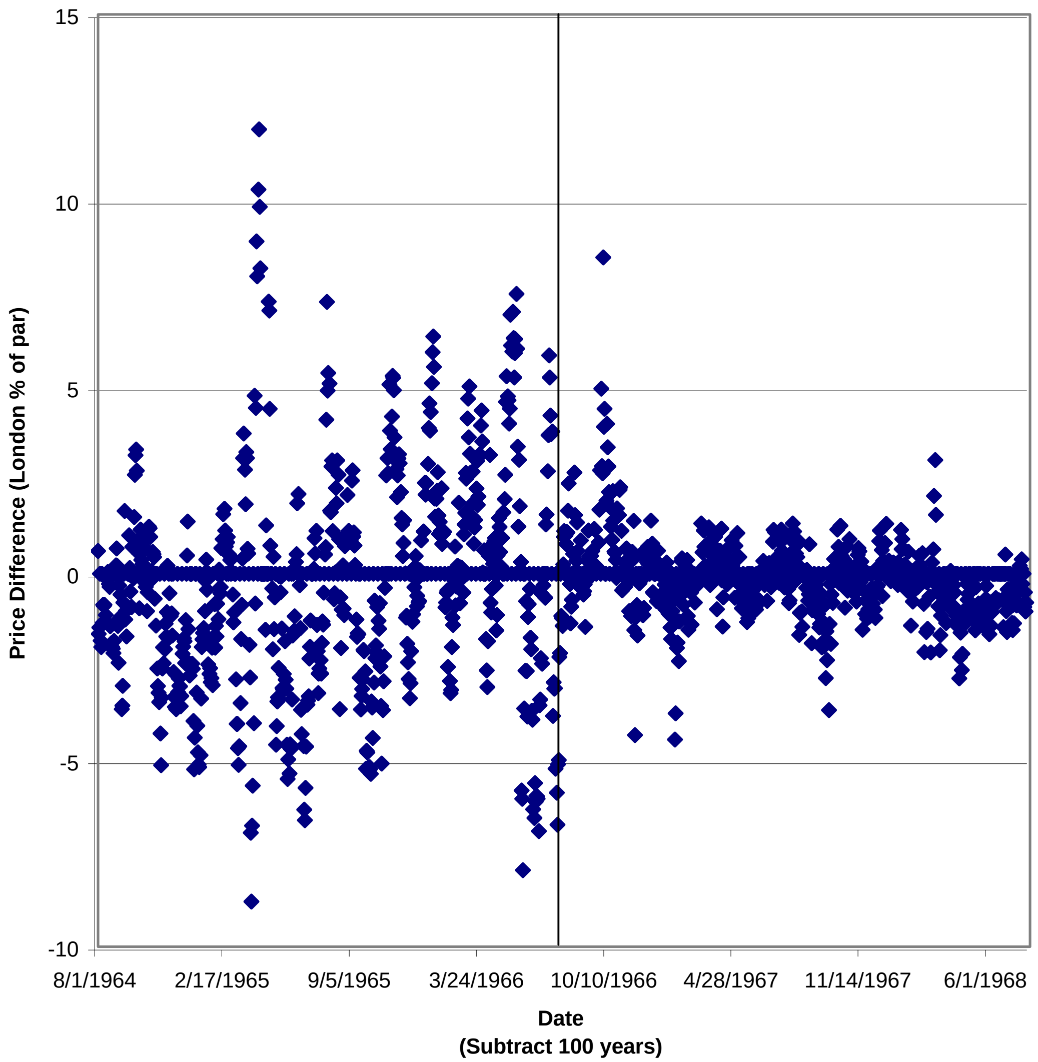


Table 1.1: Event study results, New York events

Event time	Difference in Standardized Means	t statistics	Approximate Degrees of Freedom	Modified t statistics
0	1.125	2.197 **	44	1.553
1	1.387	2.397 **	43	1.753 *
2	2.017	3.356 **	35	2.712 **
3	1.654	2.254 **	27	1.61
4	1.826	2.024 *	28	1.38
5	2.16	2.603 **	24	1.959 *
6	2.631	2.915 **	36	2.271 **
7	2.322	2.399 **	36	1.755 *
8	2.388	2.586 **	32	1.942 *
9	1.536	2.033 *	27	1.389
10	1.948	1.679	21	1.035
11	0.831	0.747	27	0.103
12	2.282	1.882 *	20	1.238
13	0.751	0.697	32	0.053
14	0.509	0.652	33	0.008
15	0.327	0.447	40	-0.197
16	0.343	0.572	34	-0.072
17	0.62	0.962	31	0.318
18	0.444	0.595	34	-0.049

** = significant at 5% level (two-tailed test)
* = significant at 10% level (two-tailed test)

Modified t statistics = subtract mean of last 4 t statistics
(t = 15 to t = 18) = 0.644

Table 1.2: Event study results, London events

Event time	Difference in Standardized Means	t statistics	Approximate Degrees of Freedom
0	1.518	3.447 **	37
1	1.128	2.373 **	36
2	1.356	2.118 **	29
3	0.919	1.324	29
4	0.91	1.395	26
5	-0.272	-0.509	27
6	0.705	0.757	25
7	0.253	0.318	33
8	0.289	0.324	24
9	0.489	0.566	29
10	0.144	0.196	28
11	0.464	0.602	27
12	0.376	0.526	30
13	0.274	0.366	28
14	0.163	0.243	39
15	0.586	0.737	26
16	0.182	0.252	33
17	0.012	0.019	43
18	0.482	0.717	35

** = significant at 5% level (two-tailed test)
* = significant at 10% level (two-tailed test)

"It is known, understood, and anticipated, by all who have dealings with the banks, that they are in the habit of preparing systematically for making creditable exhibits on quarter day."

H. R. Hulburd, Comptroller of the Currency,
Annual Report of the Comptroller of the Currency 1867, p. vii.

Chapter 2

In March 1869, Congress changed the way national banks reported their balance sheets to regulators. Under the National Bank Act of 1864, all national banks submitted balance sheets to the Comptroller of the Currency on the first Monday of each quarter. The 1869 amendment directed the Comptroller of the Currency to call five times a year for reports of previous dates of his choice. Congress intended that the new reporting procedure would eliminate the incentive for banks to window dress their balance sheets.

Window dressing is a temporary change in portfolio designed to produce a more appealing report to regulators or to the public. Window dressing obscures the health of individual banks by systematically altering underlying balance sheets. Mismeasured data can bias tests and may lead to policy errors.

Chapter 2 tests for window dressing of national bank balance sheets during the period 1866–1871. The Philadelphia Clearing House Association provides a weekly series of aggregate balance sheet items. After slight adjustments, we can compare this weekly series to quarterly reports made by the banks to the Comptroller of the Currency.

The comparison test provides some evidence for window dressing. At first, there does not appear to be evidence of window dressing because the two reports coincide before the 1869 change in the reporting law. But a comparison of bank behavior before and after the change in the law indicates that window dressing might have been present. Comparing the two series suggests that if nothing changed other than the law, then the associated Philadelphia banks window dressed their reserves by $900,000 or by about 5%.

Introduction

Congress designed the 1869 amendment to the National Bank Act of 1864 to prevent national banking associations from window dressing their balance sheets. In a statement to the Senate on February 23, 1869, John Sherman of the Senate Finance Committee remarks:

> The present law requires four quarterly statements to be made, giving certain facts. They are required to be made at periodical times, and the banks are generally doctored up or prepared for these reports, so that now a contraction occurs just before the reports are made, and after that an expansion, creating a palpable and visible fluctuation of the currency at these times. This amendment requires five reports during the year, and authorizes the Comptroller to call for them at a day past, so that they will not be prepared to doctor up their reports. (Congressional Globe, 40th Cong., 3rd sess., 1869, pt. 2:1482.)

Senator Sherman attributes variability in the money market to window dressing by national banks during quarterly reporting dates. The purpose of the amendment was to prevent banks and other market participants from anticipating reports by making the bank statements a surprise.

Window dressing decreases the effectiveness of bank regulation. Both academic economists and government regulators use historical data to test hypotheses about the bank industry and to prescribe policy. In the presence of window dressing, bank statistics submitted to regulatory authorities systematically mismeasure benchmark portfolios. Friedman and Schwartz (1970, p. 212–3) state that window dressing may cause academic economists to overestimate historical banking reserves. Also, bank regulators may fear that poor data accuracy could cause policy errors. For example, window dressing may prevent the early detection of bank distress. Regulatory agencies attempt to prevent bank failures (White [1992], p. 14). Problem banks may temporarily dress their balance sheets in order to appear financially sound. In cases where early warning may prevent insolvency, window dressing decreases the usefulness of reported balance sheets.

Modern bank regulators show concern about window dressing. In 1984 the Securities and Exchange Commission censured six banks for window dressing assets by at least 10% (Allen and Saunders [1992], p. 590). In 1962 the Comptroller of the Currency suspected that national banks were window dressing deposits and resources in order to appear larger than competitors. By this time, some call reports lacked surprise value as 24 of the 25 previous December call reports fell on the last business day of year. So in 1962 the Comptroller of the Currency surprised banks by calling for reports on December 28 instead of on the last business day of the year, December 31. When comparing the report to the Comptroller on December 28 and reports to the Federal Reserve on December 31, deposits grew by $5.7 billion or 2.6% over three days. While the increase might represent a year-end effect, bankers admitted window dressing (Comptroller of the Currency [1963], p. 14).

Recent research questions whether or not banks would have an incentive to undertake window dressing. One reason for reporting requirements is that the composition and the performance of bank asset portfolios are usually considered to be private information. For example, theories of bank monitoring such as Calomiris and Kahn (1991) assume that only the banker knows the true underlying bank asset returns while depositors receive an imperfect signal. Recent corporate governance and brokerage analyst scandals suggest that firms hold private information unavailable to investors. But Smirlock and Kaufold (1987) argue that investors could discern the relative exposure of a 1982 sample of US banks to defaults on Mexican debt, even though bank portfolio holdings were not publicly announced. If market participants can perceive true underlying portfolios, there is no incentive for banks to mislead investors by window dressing. However, Smirlock and Kaufold do not discuss how market observers obtain this information. Extracting asymmetric information from share prices becomes more difficult during the nineteenth century because bank shares were highly illiquid. In any event, the presence of window dressing would indicate that bank reports are informative to investors. Presumably, banks would not waste resources on costly window dressing if investors could discern their true asset portfolios.

Bank regulators devised reporting methods to detect window dressing. Window dressing is an information cost paid by bank regulators to obtain private information about bank balance sheets on a fixed date. The solution of Congress in 1869 was to randomize reports over dates in the past. If banks do not know when the reporting date will fall, they will not be able to "doctor up" their reports in advance. A disadvantage of random past date reporting is that variation in the reporting date captures seasonal variation in the data series. In order to eliminate seasonal fluctuation, statisticians at the

Federal Reserve insisted on calling for reports on fixed dates (White [1992], p. 18). Another solution increases the frequency of reports to abbreviated monthly or weekly balance sheets. The cost of compliance and of supervision increases as the frequency and the detail of the reports increases. Alternatively, bankers can submit an average over a substantial length of time. This procedure decreases the usefulness of a point estimate if the variance in underlying variables is large. Regulators often require both time averages and a point estimate. Allen and Saunders (1992) exploit this feature of balance sheet reporting to examine window dressing of U.S. banks.

Allen and Saunders (1992) claim that modern data exhibit upward window dressing in assets of about 2.5%. The authors acquired individual level data on large banks for the period 1978–86. They compare assets as specified on a bank's call report date to a monthly or quarterly average. Allen and Saunders find that assets as measured by a call report point estimate exceeds the previous monthly average by about 2–3% with seasonally detrended data. They present weaker evidence of a decrease in assets after the reporting date. Unfortunately, the authors assume independence in cross section of their measure of window dressing.

Design

Allen and Saunders (1992) discuss several motives for window dressing reserves. Banks have an incentive to upward window dress reserves in order to meet reserve requirements. In addition, high reserves signaled liquidity of the banks' portfolio to depositors in case of a run. Nineteenth century depositors relied on forecasts of bank solvency since deposit insurance did not exist during this period. Instead of holding liquid reserves, banks preferred riskier short-term loans to speculating stockbrokers. The Comptroller notes that the New York banks held almost half ($70 out of $160 million) of their loan portfolio in call or short-term demand loans to stockbrokers (Comptroller of the Currency [1868], p. xxii–xxiii; he classifies an additional $110 million in certified checks as speculative money).

The authors also mention that shareholders may prefer downward window dressing. Downward window dressing indicates higher profits when managers employ otherwise idle reserves. If shareholders were ignorant of their bank's condition, they may approve of extra reserves as currency cushions shareholders against bank runs. Section 12 of the National Bank Act (Robertson, [1968], p.197–8) specified that in case of bank failure shareholders were liable for the par value of the share in addition to their investment. In the end, empirical analysis must resolve the actual direction of the effect.

Ignore other types of window dressing and focus on upward window dressing reserve requirements. While Allen and Saunders empirically test for window dressing assets, nineteenth century financial regulators were more concerned about banks window dressing reserve levels. In order to window dress reserve requirements, banks can either increase reserves or decrease liabilities that required reserve. For reserve cities such as Boston and Philadelphia, section 31 of the National Bank Act of 1864 (Robertson, [1968], p. 203) specifies that reserves must be at least 25% of the sum of deposits and bank note circulation. Reserves included specie (gold coin, silver coin, and gold certificates), U.S. legal tender notes (greenbacks and certain other forms of paper government currency), deposits with a redeeming agent in New York, and clearing house certificates payable in lawful money. At least two-fifths of the required 25% reserve must consist of specie or legal tender notes in the vaults of the bank (see also *Banker's Magazine* [July 1868], p. 62). So increasing the reserve ratio required an increase in specie, legal tenders, or deposits at New York or a decrease in circulation or reserves. Unfortunately, actual reserve ratios are not recoverable from weekly data. The data lack information on deposits due from redeeming agents. However, several of the individual series of the reserve ratio are available: specie, legal tender notes and bank note circulation.

Consider the numerator of the reserve ratio, which includes specie and legal tender notes. Friedman and Schwartz ([1963], p. 28) note that both a greenback dollar and a gold dollar counted as one dollar of reserves even though gold dollars traded at a premium to paper greenback dollars in the market. So for the purpose of window dressing reserves, increasing legal tender notes was more cost-effective.

The temptation to window dress bank liabilities was ambiguous. While a decrease in circulation or deposits would increase the reserve ratio, bankers may have been reluctant to purposefully decrease these balance sheet items. Circulation and deposits may have prestige value as they indicate the size of a bank.

Banks used several methods to window dress reserves. Cornwallis (1879, p. 48–9) notes that banks sought “to borrow deposits for a single day in any way they could, and to reduce loans to directors for a few hours….” The most straightforward method was to convert assets to cash. For example, bankers can augment reserves by suitable manipulations of short-term loans. Bankers can arrange for loans to come due just before the reporting date. One might expect loans to decrease around the reporting date. However, any decrease in loans around the reporting date should appear as an increase in liquid assets, such as legal tender notes. Thus, it suffices to look at a series of legal tender notes.

So one test of window dressing looks for increased holdings of U.S. legal tender notes by banks around the call date. Hence, the effect is exactly what Senator Sherman described: a contraction of the currency around the reporting dates. If banks really window dressed their balance sheets, we can find evidence for the contraction by examining bank holdings of legal tender notes. This study tests whether or not bank window dressing directly caused currency contractions by examining bank portfolios.

Method

One test of window dressing compares reported reserves to an underlying series of holdings. The advantage of using historical data is the availability of two sets of reports of reserves on roughly the same set of banks. In his *Annual Report*, the Comptroller of the Currency published quarterly aggregate balance sheets of national banks by state and reserve city. The test requires an additional series of balance sheet data from between the quarterly reporting dates. Bankers demonstrate reluctance to distribute balance sheets because they reveal private information about bank portfolios. Allen and Saunders (1992) test for window dressing by comparing a monthly or quarterly average to a quarter-end report that were both reported to regulators. Although section

34 of the National Bank Act (Robertson, [1968], p. 204) required banking associations to submit abbreviated monthly reports, these were not available.

Clearing house associations provide a second series of data. The clearing houses of major metropolitan areas published weekly aggregates of member banks' condition. Clearing house associations expedited the daily balance of payments among banks and provided for mutual defense during financial crises. Some major clearing house associations required members to submit weekly average balance sheets, which were then published in local newspapers. Gorton and Mullineaux (1987) describe how self-monitoring by the coalition relieved information asymmetries and permitted clearing at par among member banks. These weekly statements from the clearing house serve as a benchmark of bank holdings.

If the two series measured the same accounting item on an identical set of banks, we could test the two series for consistency and hence for window dressing. However, in most cities, the two sets of banks covered by aggregates statistics do not coincide. The Comptroller's reports of condition include a few national banks that were not members of the local clearing house, while the clearing house aggregate weekly statement would include state-chartered banks that did not fall under the jurisdiction of federal regulators.

The presence of state banks could obscure the presence of window dressing in aggregate data. Aggregate data may not reveal window dressing when only one class of banks submits a report. National banks may borrow reserves from state banks, leaving the aggregate statistics unchanged on the quarterly reporting day. For this reason, the nation's premier money market, New York, is a less desirable choice with only aggregate data. Several New York City banks retained state charters and maintained membership in the New York Clearing House. While individual level data can circumvent this problem, there is another solution.

A fortuitous historical circumstance allows a comparison between the Comptroller's quarterly reports and the clearing house series of Philadelphia. Only national banks belonged to the Philadelphia Clearing House by the end of 1866. Most state banks converted to national banks shortly after Congress announced the taxation of state bank note circulation. State banks did not reenter the Philadelphia market until 1869 and did not join the Clearing House at least until after 1871. Unfortunately, we are left with the problem of national banks remaining outside the clearing house.

The two aggregate series cannot be directly compared because not all of the national banks joined the clearing house. Two Philadelphia national banks remained outside the clearing house as of 1867. These two banks are included in the Comptroller's aggregate reports but are not included in the clearing house series. Once the Comptroller’s aggregate reports are adjusted for the two extra banks, a first test compares the Comptroller’s aggregate reports of legal tender notes to the clearing house weekly average series of legal tender note holdings. If the Comptroller’s reports lie above the benchmark clearing house weekly average, this indicates window dressing.

Comparing the two data series allows for a cleaner test of the hypothesis. Instead of the nonparametric version of the comparison test, we could investigate the time series properties of the clearing house series to test for window dressing. The test looks for increases in the time series of legal tender reserves around the weekly reporting date.

However, the time series analysis suffers from several drawbacks. The series requires seasonal adjustment, but we observe only a small number of seasonal cycles over five years of data. Also, time series analysis requires stationarity and distributional assumptions on the residual process. Finally, the time series test has low power. Suppose that banks window dressed their balance sheets for one day. The weekly average consists of a six-day time average. Thus, a one-day increase in reserves will be divided by six in the weekly average of the clearing house. Even a large amount of window dressing, when divided by six, is difficult to detect against the background variance of the series.[6] The nonparametric test trades these assumptions for the requirement that nothing changed other than the reporting law.

Data

The weekly reports of the aggregate condition of members of the Philadelphia Clearing House approximate the twenty-eight associated banks' underlying portfolios. Clearing houses published aggregates of a few important items including loans, deposits, specie, legal tender notes, and bank note circulation. The *Banker's Magazine* printed at least portions of these series. Supplementing this source were the individual level reports of clearing house member banks contained in the Philadelphia *Public Ledger* and the Philadelphia *Inquirer*. We consider data for the period 12/1866 – 11/1871 (260 weeks) to match the series of reports from the Comptroller.

6. Time series tests of the Philadelphia Clearing House legal tender weekly average series based on a seasonal ARMA model accept the null hypothesis of no window dressing. The time series tests also accept the null hypothesis that banks window dressed their balance sheets by a large amount.

The national banks reported their balance sheets to the Comptroller of the Currency quarterly before the change in the law and five times a year afterwards. In his *Annual Report*, the Comptroller publishes aggregates of full balance sheets of all national banks by state and by reserve city. Abstracts of the aggregate reports from 1867 to 1871 appeared in the *Banker's Magazine*. Since the law changed on March 3, 1869, this yields nine reports before the law changed and twelve reports afterward.

The Comptroller's aggregate reports require adjustments before comparing them to the clearing house series. First, the Comptroller's aggregates include the two national banks that did not belong to the Philadelphia Clearing House. The Comptroller's series can be corrected because the Comptroller's reports and daily newspapers occasionally published quarterly reports of individual banks. Second, we lack precise information about the clearing house definition of legal tender notes. For more information about legal tender data issues, see Appendix 2A.

Test

Chapter 2 tests whether or not banks window dressed their balance sheets. The test compares the sequence of reports of the Comptroller to the underlying weekly averages of the clearing house. If reports to the Comptroller exceed reports to the clearing house before the change in the law, then there is evidence of window dressing. However, this test assumes that the two sets of reports ought to be equal in the absence of window dressing. Consider the counterfactual world without reports to the Comptroller. Perhaps the counterfactual condition of the banks is systematically above or below the weekly average. To account for this possibility, a more general version of the first test compares the excess of the Comptroller's report over the clearing house averages before and after the change in the law. If nothing changed other than the law, a decrease in the excess of the Comptroller's report over the clearing house report from the period before the law changed to the period after the law changed may also suggest window dressing. Ideally, the call reports after the change in the law surprise the banks and reveal their true condition. We can use the reported condition of the banks after the law changed to approximate the counterfactual condition of the banks before the law changed.

For notation, let x_q be the q^{th} aggregate quarterly report of legal tender note holdings by Philadelphia banks (adjusted for the two banks that were not members of the Clearing House). Subdivide the q's into those that fell before and after 1869 amendment. Let $q = 1$ to 9 be the nine quarterly reports (1/67 to 1/69) that fell before the change in the law and let $q = 10$ to 21 be the twelve call reports that fell after the change in the law (4/69 to 10/71). Let y_t represent the series of Philadelphia Clearing House weekly average of U.S. legal tender notes. Let y_q denote the clearing house weekly average of U.S. legal tender notes for the week of the quarterly report q. Now define the differences between the aggregate reports of the Comptroller and the clearing house weekly average. Let $b_q = x_q - y_q$ for q in 1 to 9 and $a_q = x_q - y_q$ for q in 10 to 21.

The test compares the behavior of the banks before and after the 1869 amendment. Suppose that no bank window dresses its balance sheet. If nothing changed other than the law, then the difference between the adjusted Comptroller's reports and the clearing house reports should be the same on average before and after the change in the law. That is, test the null hypothesis that no window dressing took place:

$$E(b_q) = E(a_q). \tag{2.1a}$$

A restricted version of this test assumes that in the absence of window dressing, the Comptroller’s report and the clearing house weekly average should coincide. In this case there should be no difference between the two sets of reports after the change in the law, so restrict $E(a_q) = 0$. Test the null hypothesis of window dressing before the change in the law with:

$$E(b_q) = 0. \tag{2.1b}$$

Replace the expectations by their sample analogues to obtain testable restrictions. The test will appeal to nonparametric statistics due to small sample sizes.

Results

At first glance, Figure 2.1 presents little evidence of window dressing. Figure 2.1 graphs the Philadelphia Clearing House legal tender note series along with the aggregate report to the Comptroller adjusted for the two non-member banks. Figure 2.1 indicates that reports to the Comptroller appear close to the clearing house series before the law changed. But after the 1869 amendment, reports to the Comptroller fall below the bulk of the clearing house. There does appear to be some difference between the two series before and after the change in the law.

The test compares the Comptroller's reports to the weekly averages from the clearing house and suggests that, all else equal, window dressing took place on the order of a million dollars. A t test of (2.1a) estimates the magnitude of the difference between the two samples. An F test does not reject the null hypothesis of equal variances (F = 1.85, 10% critical value = 2.52), so proceed with a t test pooling the variances across the samples of a_q's and b_q's (although an approximate t test with different variances does not change the results). The difference between the mean of the sample of b_q's and mean of the sample of a_q's was $923,668 and was significant (t = 4.11, 5% critical value = 2.09) under an independence assumption. If all else remained equal before and after the change in the law, the t test concludes that window dressing averaged about $900,000 per quarter.

Nonparametric tests verify that the two samples are drawn from different distributions. A non-parametric Mann-Whitney U test examines the null hypothesis that two random samples have the same median (Beyer [1991], p 309; Rice [1991] contains other versions). In this case, the two samples are the excess of the Comptroller's reports over the clearing house average (adjusted for the two missing banks) before and after the change in the law. The Mann-Whitney U statistic checks if the observed b_q's and the a_q's are independent draws from two different distributions. The U statistic was significant at better than the one percent level (U = 11, 1% critical value = 18), suggesting that the two distributions are in fact different.

Most of the apparent difference between the two samples derives from the Comptroller's report falling significantly below the weekly average after the change in the law. Reports to the Comptroller do not appear to exceed the clearing house weekly averages before 1869 amendment. A t test of (2.1b) is not significantly different from zero (t = 0.71, 5% critical value = 2.306). Thus, most of the $900,000 difference must result from the period after the change in the law. The average difference between the adjusted Comptroller's reports and the clearing house weekly averages for the period after the change in the law is significantly different from zero (t = 5.87, 5% critical value = 2.201). The call reports dropped substantially below the weekly average reported to the clearing house after the change in the law. Counterfactually, in the absence of reporting requirements, the reports to the Comptroller would have been about $750,000 under the clearing house weekly averages before the law changed.

Why did reports to the Comptroller fall significantly below the clearing house average after the change in the law? Several possibilities may explain the result. First, reports taken on the first Monday of each quarter may differ systematically from reports taken at other times. If banking practices at the beginning of the quarter could explain window dressing, then surely the Comptroller during this period, an experienced banker, would be aware of it. Also, the timing of call reports after the change in the law varied according to the whim of the Comptroller. Perhaps the Comptroller chose dates specifically after a short-term shock struck the money market. Neither official directive nor informal communication suggests that the Comptroller restricted his attention to the health of banks during adverse monetary conditions. Most often the calls for reports came about a week after the date required for report, so the Comptroller may not have had time to observe a trough in the money market. Further, several calls fell during the seasonal ease of the spring money market.

A second explanation considers intraweek fluctuations in bank holdings. Friedman and Schwartz (1970, p. 213) note that intraweek variation may cause sizable fluctuations in local money markets. If payday was Friday, banks would drain reserves by paying out cash. Call reports taken from the close of business on Saturday or early during the week would therefore lie below a weekly average. All of the reports before the change in the law and most of the call reports after the change in the law fell during the beginning of the banking week. Additional support for this hypothesis notes that the difference between the two reports appears somewhat larger than the average of the other reports for the three call reports that fell on Wednesday or Thursday. However, this story does not explain why the differences on these days are negative. According to the explanation, the banks should have built up cash reserves late in the week. So reports to the Comptroller should exceed the clearing house averages when the report fell late in the week, but they do not. Reports can not fall below the weekly average on every day of the week.

A third possibility notes a decrease in the supply of legal tender notes that occurred mostly before the 1869 amendment. Although this effect goes in the wrong direction to explain why the Comptroller's report fell below the clearing house weekly averages after the law changed, it might explain the average $900,000 discrepancy before and after the change in the law. If the supply of legal tender notes was decreasing, the

Comptroller's Monday morning quarterly report of legal tender notes will appear above a weekly average from the clearing house. Throughout 1867, the Treasury contracted the supply of legal tender notes at a rate of $4 million a month in an attempt to lessen the gold premium. Other forms of paper government currency also declined. From Figure 1, we know that aggregate legal tender notes (including compound interest notes and three percent certificates) in Philadelphia fell from about $20 million to $13 million (excluding the two missing banks) before the change in the law. Hence, the supply of legal tenders in Philadelphia dropped $7 million over about 100 weeks. Distributing the decrease linearly over the two years yields a weekly decrease of $70,000. Comparing reports early in the week to a weekly average halves the trend. So, perhaps $35,000 of the difference between the Comptroller's report and the clearing house weekly average before the change in the law is due to the currency contraction. A similar calculation can be performed after the law. In either case, the decrease in the supply of legal tender notes fails to explain the average $900,000 difference before and after the change and the law.

A final explanation is an unknown mismeasurment of the legal tender series or a change in the definition of the variables. We lack precise information about exactly what the clearing house column "legal tenders" represents. The calculations did not include items such as bills of other national banks, checks and cash items, or exchanges for the clearing house under the heading of legal tenders. None of these items should be considered substitutes for U.S. legal tender notes. Changes in the inclusion of these items fail to explain the large excess of the clearing house reports over reports to the

Comptroller after the change in the law. Appendix 2A provides more information about adjusting the clearing house definition of legal tender notes.

Conclusion

This chapter examined U.S. legal tender note holdings of Philadelphia national banks during the period 1866–1871 in order to detect window dressing. If reporting standards did not change and all else remained equal before and after the change in the law, a t test suggests that Philadelphia banks window dressed on the order of $900,000 out of an average of about $15,000,000 of legal tender note holdings, or about 5% in the aggregate. Again, any evidence for window dressing hinges on the requirement that nothing changed other than the law. A complete solution would explain why reports to the Comptroller of legal tender note holdings of the associated banks fell below their weekly average after the change in the law.

Appendix 2A

Appendix 2A describes adjustments to the legal tender holdings of the Philadelphia banks. The first adjustment facilitates direct comparison between the Comptroller's aggregate reports and the weekly averages of the Philadelphia clearing house. This adjustment corrects the Comptroller’s quarterly aggregates for the two national banks that remained outside the Philadelphia clearing house in order compare the two data series. Unfortunately, individual level data on quarterly reports are not immediately available. One approach notes that Section 34 of the National Bank Act of 1864 (Robertson [1968], p. 204) required national banks to publish abstracts of their individual balance sheets in local newspapers. In compliance with the law, banks placed advertisements of their balance sheets in the paper of their choice. Banks scattered their individual reports among the various papers of the city. One possible solution is to collect these quarterly reports of individual national banks that did not join the clearing house and subtract them from the aggregate. This was easiest for the city of Philadelphia, as there were only two banks outside the Philadelphia Clearing House. The city of Boston had at least four, while New York had about eight national banks outside their respective Clearing Houses during this period.

In the absence of a complete series of quarterly reports, we estimate the legal tender holdings of the two remaining banks outside the clearing house. We collected a handful of individual quarterly reports for the two Philadelphia banks. The collection currently includes all four quarterly reports of 1867 for both banks, but almost no further quarters. The 1867 reports allow estimation of the holdings for the quarters in other years. The Comptroller's *Annual Report* provides individual level data for all national banks for the October report. However, the October report is not an appropriate estimate for the other quarters due to seasonal variation in reserves. So we estimate the remaining reports of a given year by appropriately upweighting the October report for that year. The weight chosen was the ratio of the average of the other three quarterly reports for 1867 to the October report for 1867. As the 1867 quarterly reports for the National Bank of Germantown printed in the newspapers also included specie in its declared amount of legal tender notes, we subtract the specie declared in its 1867 October report (less than $2000).

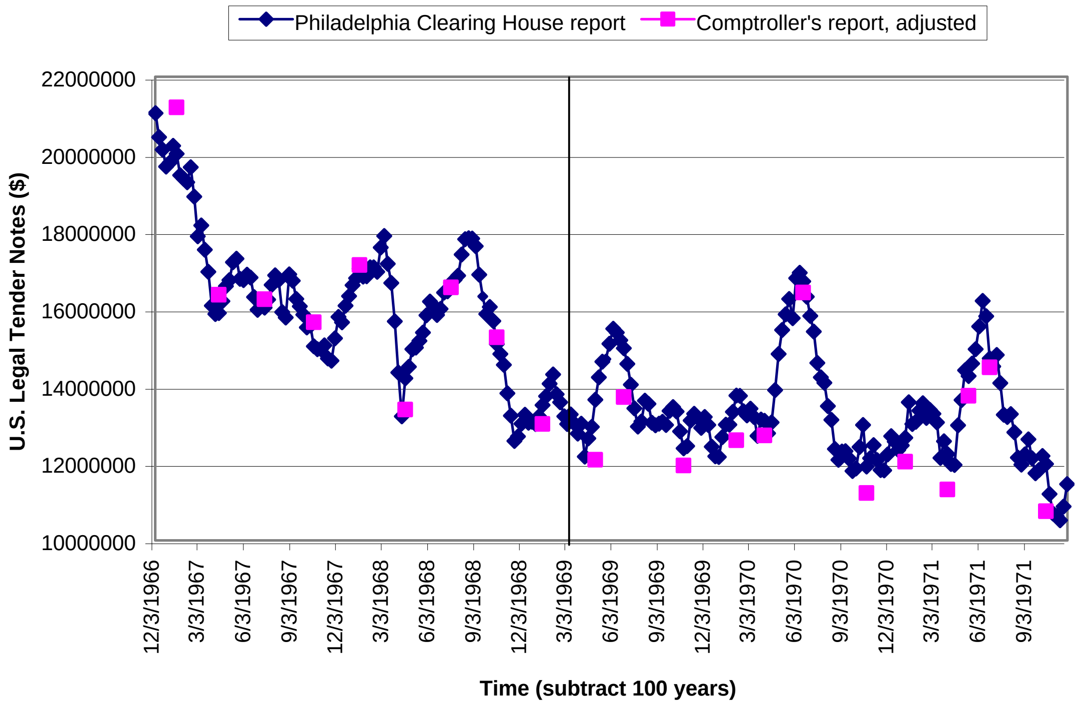
Another difficulty is that the definitions of "legal tender" of the Comptroller and the clearing house may have had different meanings. While the Comptroller annotates his accounting definition in his 1868 report, several items other than legal tender notes may or may not have been included in the clearing house definition. A weekly statement for the New York Clearing House from 1900 (Cannon [1908], p. 184) limits legal tender notes as the only input to the column headed "legal tender." Although they lacked legal tender status, fractional currency, bills of other banks, cash items, and clearing house exchange might plausibly have been placed under the heading "legal tender." However, accounting changes in these items are either far too small or far too large to explain the discrepancy between the two measures of legal tenders that appears when the law instituted call reports.

Further, we lack an official statement about which days of the week the clearing house averages represent. But according to an example in the money column of the

Philadelphia Inquirer (October 1, 1867), clearing house averages are turned in on Saturday and are composed of an average of the closing statements of the previous Saturday through Friday. We used this definition of a weekly average in the absence of more precise information. Before the 1869 amendment, we treat reports from the Comptroller on a Monday before the opening of business as a Saturday report at the close of business.

We obtain information about membership in the Philadelphia Clearing House from local newspapers. Philadephia newspapers printed individual weekly averages of the member banks of the clearing house. This raises a question of why the Philadelphia national banks that were members of the clearing house ought to bother window dressing since their weekly average individual balance sheets were publicly available. Anecdotal evidence suggests that smaller, weaker banks still attempted to increase their reserve holdings over the quarterly report (money column, *Philadelphia Inquirer*, April 8, 1868). However, window dressing by weak banks may have occurred at the expense of reserves of larger and stronger banks, so that no change appears in the aggregate.

Figure 2.1
Philadelphia U.S. Legal Tender Notes, 1866-71



"It follows, therefore, that the seven banks, all of which paid interest upon deposits and which had secured the bulk of the bankers' deposits, were directly responsible for any disturbance in the New York money market, which was due to the use of these funds, and also for any failure to meet demands for their return to banks in the rest of the country."

O. M.W. Sprague (1910), p. 20–1
History of Crises under the National Banking System

Chapter 3

Chapter 3 conducts an econometric test of Diamond and Dybvig's (1983) theory of bank runs as interpreted by Calomiris and Gorton (1991). Diamond and Dybvig employ an exogenous liquidity shock to depositors in order to develop a theory of bank runs. Calomiris and Gorton interpret the exogenous liquidity shock as a seasonal withdrawal from the nation's agricultural interior. This chapter reexamines the hypothesis that a seasonal interior reserve drain served as the exogenous liquidity shock before the bank panics of 1873 and 1893 in the United States. Using individual bank level data in New York, this chapter tests whether the banks that held most of the deposits from the interior, the "interest-paying" banks, experience reserve drains just before the panic. The evidence reveals that the 1873 panic could have been triggered by a seasonal interior drain, as interest-paying banks sustained large withdrawals in the last few weeks before the panic started. But just before the panic of 1893, interest-paying banks did not suffer heavy reserve drains. Diamond and Dybvig's model cannot be applied to the bank panic of 1893 without a non-seasonal interpretation of the exogenous liquidity shock.

Introduction

Although game-theoretic models of bank crises have existed for at least twenty years, there have been few econometric tests of the theory. A long literature beginning with Bryant (1980) and Diamond and Dybvig (1983) models bank runs of depositors on a single bank. Most subsequent empirical papers, such as Demirgüç-Kunt and Detragiache (1998), limit themselves to establishing stylized facts about bank crises rather than explicitly testing the theory. One exception is Madiès (2001), who tests the Diamond and Dybvig (1983) model with a laboratory experiment. Madiès observes bank panics in the laboratory, although a bank run by all subjects is rare. Another exception is Calomiris and Gorton (1991), who use historical data to evaluate an interpretation of Diamond and Dybvig's theory of bank runs.

The model of Diamond and Dybvig (1983) remains one of the most theoretically influential explanations of bank distress. In order to generate bank runs, these authors specify that some depositors, called "impatient" agents, experience a random exogenous liquidity shock that motivates them to withdraw deposits. The agents expect the shock to occur, only the agents do not know the identity of which particular subset of the agents will experience the shock. Two equilibria emerge for the behavior of the agents who are not exogenously motivated to withdraw, the "patient" agents. In the equilibrium with a bank run, each patient depositor expects other depositors to withdraw funds, in which case it is a best reply to withdraw. In the equilibrium without a bank run, each patient depositor expects other depositors to keep their funds in the bank, so no run occurs. The alternation between equilibria explains the presence or absence of bank runs. In their theory, Diamond and Dybvig assume that a certain class of depositors always withdraws without explaining why these exogenous withdrawals take place. More recent theoretical advances, such as Allen and Gale (2001), also rely on an exogenous shock to create bank distress.

In one explanation of the theory, Chari (1989) identifies the exogenous liquidity shocks with seasonal agricultural requirements for currency. In the nineteenth-century United States, banks in the interior often deposited funds with money center banks in New York in order to obtain interest.[7] Periodic agricultural activity, such as planting and harvesting, create a seasonal transactions demand for money. At harvest time, banks on the agricultural interior of the country would withdraw cash from money centers such as New York. The seasonal liquidity requirement fits the theory because the seasonal shock is exogenous and, although its magnitude and timing were somewhat stochastic, market participants anticipated its fluctuations. Further, the interpretation of the exogenous liquidity requirement as a seasonal agricultural withdrawal connects with traditional investigations of the seasonal recurrence of financial crises in central money markets. Jevons (1884) suggests that the monetary needs of country patrons created an annual currency shortage every autumn in the English money market during the mid-nineteenth century. Kemmerer (1910) observes that financial crises in the US often coincided with

---

7. In fact, the pyramid reserve clause of the National Bank Act of 1863 allowed interior banks to count some deposits on money-center banks as cash reserves in the vault for the purpose of fulfilling a reserve requirement. However, the National Bank Act was not the cause of interior banks loaning funds to New York banks. Myers [(1931), Chapter 6] shows that banking practice predated the pyramid reserve clause.

seasonal fluctuations of the money supply. Miron (1986) argues that the Federal Reserve prevented panics between 1914–28 by accommodating seasonal monetary movements.

Following this explanation of the exogenous liquidity shocks, Calomiris and Gorton (1991) equate Diamond and Dybvig's theory of bank panics with seasonal withdrawals on money center banks. Calomiris and Gorton apply the Diamond-Dybvig model to the National Bank Era by identifying country banks in the nation's interior with "depositors" and the large banks of New York with "the bank" in the model. Calomiris and Gorton then evaluate the seasonal withdrawal hypothesis.

Calomiris and Gorton provide many arguments against a seasonal explanation of nineteenth century financial crises. The authors use historical data to refute the seasonal hypothesis, but they never use statistical tests to evaluate their assertions. Although they offer additional evidence against the seasonal hypothesis during and after panics, focus on their analysis of pre-panic periods. First, they describe how currency shipments do not match the seasonal hypothesis. In particular, data on currency shipments after 1899 show no unusual outflows from the New York banks just before the panic of 1907, at least when compared to 1906. Calomiris and Gorton also compare changes in aggregate reserve ratios and aggregate deposits of New York member banks across years during the four calendar weeks preceding each panic. Their premise is that the biggest reserve shocks should correspond to panic years. Instead, they that find that changes in the reserve ratio and deposits in panic years were unexceptional when compared to the same calendar weeks in non-panic years.

But it is not obvious why the perceived lack of correlation between large pre-panic drains and bank panics refutes the theory of Diamond and Dybvig. Unlike Smith (1991), who creates panics with large seasonal shocks, Diamond and Dybvig generate their panics with multiple equilibria. Fluctuations in the size of the shock need not influence equilibrium selection in the version of the model without aggregate risk. Further, Calomiris and Gorton consider only the aggregate reserves of the New York banks. However, interior deposits were not distributed equally among the New York banks. A better test would focus on those New York banks that held interior deposits.

In his classic analysis of the panics, *History of Crises under the National Banking System*, the prominent economist Sprague (1910) highlights the reserve position of a subset of New York banks, called "interest-paying banks" by contemporary observers. A key feature of interest-paying banks was not the payment of interest on demand deposits but rather the solicitation of interior balances. While many banks paid interest on demand deposits, large interest-paying banks specialized in accepting demand deposits from banks located on the interior of the country, so that the interest-paying banks' deposit portfolios consisted largely of interior bank balances. Under the seasonal withdrawal hypothesis we should observe that the interest-paying banks suffer large percentage reserve drains on the eve of a panic. Aggregate statistics could conceal a shock to these individual banks.

Both Sprague and New York Clearing House leadership expressed concern about the stability of interest-paying banks. In the absence of a central bank, the New York Clearing House was the premiere financial institution in the United States during the nineteenth century. All of the largest and most influential New York banks were members of the New York Clearing House, a voluntary association of banks designed to expedite payments among member banks. The leaders of the New York Clearing House understood the vulnerability of the interest-paying banks and attempted to ban the payment of interest on demand deposits several times. In a far-sighted report on banking practices in 1873, George S. Coe and a committee of Clearing House members described how twelve interest-paying banks suffered nearly two-thirds of the total drain reserve drain on all sixty member banks while only consisting of only half of total member deposits (Sprague, p. 93). In the presence of the large reserve drain during the panic, we might expect the drain on the interest-paying banks to begin before the panic starts.

Using individual level data, we can test the assumption of a seasonal drain by examining withdrawals on interest-paying banks. This chapter extends the analysis of Calomiris and Gorton by providing an econometric test of the seasonal interpretation of Diamond and Dybvig. The test employs individual bank level data on New York banks. The seasonal hypothesis implies that interest-paying banks should suffer a reserve drain as the crisis begins. If the existence of seasonal drains should be rejected, then we must either discard the theory of Diamond and Dybvig or provide another explanation for the source of the exogenous liquidity shock.

Several authors consider alternative explanations of bank runs that are not evaluated here. This chapter investigates the hypothesis that an exogenous shock from the interior triggered the panic and takes the position that the drain on interest-paying banks was a cause and not a result of financial instability. Hence, it ignores a causal role for shocks to the value of bank assets as described by Calomiris and Gorton and modeled by Rochet and Vives (2001) and Goldstein and Pauzner (2002). It also ignores international pressure on foreign exchange markets as a source of bank distress. Kaminsky and Reinhart (1999) and Glick and Hutchison (1999) discuss interaction between banking crises and currency crises and consider evidence from international panel data. In the context of the 1893 crisis, Friedman and Schwarz (1963) assert that the panic was the result of international deflation abroad (p. 111). Milller (1996) argues that foreign exchange rate speculation concerning the maintenance of the gold standard caused an internal drain during June–July of 1893.

Design

Instead of testing the theory of Diamond and Dybvig outright, Chapter 3 tests the interpretation as described by Calomiris and Gorton. Several difficulties impede a direct test of Diamond and Dybvig's model with historical data. The theory is static (although Temzelides [1997] considers an evolutionary dynamic version), but most existing bank data are dynamic. Gorton (1988) notes that Diamond and Dybvig do not describe why beliefs change. Consequently, researchers turn to testing the theory's assumptions. If the assumptions of the theory are false, then the theory does not explain observed phenomena. Diamond and Dybvig assume that a certain class of depositors always withdraws. Calomiris and Gorton interpret this assumption as seasonal withdrawals by interior banks on New York banks. Of course, other non-seasonal interpretations of the exogenous liquidity shock could provide evidence for the model of Diamond and Dybvig, but alternate interpretations are not evaluated here.

This chapter tests for seasonal reserve drains on New York banks just before major financial crises. In Diamond and Dybvig's static model, the two observed equilibria are 1) complete liquidation of the bank and 2) withdrawals only by impatient agents (those who were exogenously motivated to withdraw). In both equilibria under the seasonal interpretation, New York banks with a connection to the interior should suffer seasonal reserve drains. However, since in the real world withdrawals are dynamic, we might expect withdrawals to take place over a period of time. So we will look for seasonal reserve drains in the last few weeks before the panic begins.

This chapter tests the theory before the two largest financial crises of the late nineteenth century in the United States, the panics of 1873 and 1893. The first, the panic of 1873, fits the classic pattern of panic chronology. Sprague (1910) and Wicker (2000) describe the seasonal stress on the money supply and the sharp break in asset prices just before the panic. These authorities agree that the beginning of the panic of 1873 coincided with the failure of a prominent banking house in mid-September, 1873. Figure 3.1 graphs aggregate total net deposits (net deposits plus bank note circulation) of New York Clearing House member banks from January 1872 to September 1873. The black line denotes the beginning of the panic. The visible seasonal highs and lows correspond to agricultural requirements of the money market, and it is easy to understand the importance of the seasonal interpretation at this period in the nation's history.

Unfortunately, individual bank level data does not exist during the panics. During crises, the New York Clearing House Association attempted to counteract the panics. One form of relief was the issue of clearing house loan certificates to needy members, a form of temporary emergency loan (Cannon [1908], chapter 10 or Hoag [2002b]). Once protective measures were in place, the associated banks no longer published the balance sheet by individual bank. So, for the purposes of this study, the data end just before the associated banks took concerted action in September 1873.

The year 1893 has both advantages and disadvantages for testing hypotheses about panics. In its favor, most historians note the large internal drain from the interior. Wicker (p. 52) describes how, unlike most other bank panics, interior bank suspensions and failures were unusually numerous during this panic. Since we know interior banks experienced stress, we might expect large withdrawals from New York banks, whether of seasonal origin or not.

The panic of 1893 was somewhat unusual when compared to other banking crises of the nineteenth century for other reasons. Another difference was the long period of unease in financial markets before the actual crisis. Sprague marks the price collapse of a stock market favorite in February 1893 as the first stage of the crisis before the partial suspension of payments in August 1893, when certain banks refused to pay large depositors in cash (Noyes [1894], p. 25). Large declines in asset prices also took place in early May 1893. Defining when the panic actually begins is critical to testing the hypothesis because the hypothesis specifies the behavior of depositors at the beginning of the panic. We might misunderstand panic behavior if we use the wrong starting date for the panic.

The year 1893 limits the available data. In 1893, New York Clearing House leadership recognized the unsettled state of the money market nearly two months before the crisis reached a critical height. The Clearing House issued loan certificates in mid-June 1893, while partial suspension of payments did not occur until August. The early issue of loan certificates means that the data ends in June, so we may not expect the seasonal hypothesis to apply to this case. Mid-summer was not known for its seasonal stress on the money supply. In practice, Calomiris and Gorton also date panics from the issue of New York Clearing House loan certificates. In other years, such as 1873, the Clearing House did not take preemptive action so data up to the actual suspension are available.

Evidence from banking aggregates suggests that dating the panic from June 1893 may be a reasonable choice. Figure 3.2 graphs aggregate total net deposits (net deposits plus bank note circulation) for the New York Clearing House banks for January 1892 to June 1893. The aggregates are available during the panic even though individual level statistics are not. The banks clearly experienced difficulty in early 1893. A serious downturn in aggregate total net deposits begins the week of June 10, 1893 (the black line in the figure), when the Clearing House decided to issue loan certificates. Total net deposits stayed low for several months before returning to its 1892 level. While the money market was unsettled prior to June 1893, it seems reasonable to date the last stage of the crisis of 1893 as beginning in early June.

Data

The data include the weekly statements of the 63 New York Clearing House member banks for the 76-week period January 1892 to June 1893 and for the 60 member banks for the 91-week period January 1872 to September 1873. The weekly bank statement carried weekly averages of key balance sheet items, including net deposits and bank note circulation. Net deposits included such items as deposits due individuals, deposits due to other banks, and unpaid dividends, but subtracted out deposits due from other banks, bank notes of other banks, and other cash items. Define total net deposits as the sum of net deposits and bank note circulation. Appendix 3A contains further information about the data.

Since we lack measures of reserve flows, we turn to measures of deposits. The hypothesis is framed in terms of the flow of reserves: interest-paying banks remitted cash to the interior. However, we only have data on the stock of reserves and must impute the flows as best we can from the stocks. While it might seem most natural to test the hypothesis by examining reserves, changes in reserves do not measure cash outflows by the bank. A bank could make payments without changing the level of reserves by calling in loans or liquidating other assets. Changes in deposits will capture some loan collection and therefore may measure reserve drains better than changes in reserves. However, Myers (p. 409) suggests that empirically changes in reserves might be a better proxy than changes in deposits for seasonal money flows.

One data problem impedes a test for a reserve drain on interest-paying banks. We lack precise information about which banks held interior deposits. Cross-sectional data on the amount of interior deposits accepted or rates of interest paid are not available. Instead, we can use information on the amounts of bankers' balances owed to any bank, whether from the interior or otherwise, as a proxy for holding interior bankers' balances. Four or five times a year state and national banks submitted reports of condition to government regulators. These reports document the holdings of banker's balances by individual bank during the years 1872 and 1892. The year before the panic measures bankers' balances better than the actual year in case there were large drains during the panic. Averaging over all four or five reports avoids seasonal variation in the level of bankers' balances in New York.

The definition of "interest-paying bank" splits the banks into two classes. The classification attempts to detect those banks with a large fraction of their deposit portfolio in interior balances. Banks with the highest proportion of bankers' balances will suffer the largest proportionate reserve drains. Define an interest-paying bank as owing more than one-half of total gross deposits (individual plus bankers') to other banks in the reports to regulators in 1872 or 1892. Banks fall into one class or the other for the entire sample period. This definition selects 9 of 60 banks in 1872 and 13 of 64 banks in 1892 as interest-paying banks. Admittedly, the binary classification is a crude proxy for possessing liabilities due to the interior. Non-interest-paying banks accepted bankers' deposits, so the classification is a difference of degree not of kind. An alternative classification, defining an interest-paying bank as owing 40% or more of its total gross deposits to other banks (selecting 12 of 60 banks in 1872 and 18 of 64 banks in 1892 as interest-paying banks) does not affect the qualitative results. Using the ratio of net bankers' balances to total net deposits does not substantially affect the classification.

By definition, bankers' deposits formed a substantial fraction of the large interest-paying banks' liabilities. The business of large interest-paying banks depended on interior bankers' balances as a source of funds. In both 1872 and 1892, bankers' balances averaged about two-thirds of gross deposits for the interest-paying banks. Non-interest-paying banks held only about 20% of their gross deposit portfolio in bankers' balances.

Bankers' deposits were concentrated in a small number of interest-paying banks. While many national banks paid interest on interior deposits, a few banks commanded most of the balances. The New York interest-paying banks possessed 64% of the bankers' balances held by the New York Clearing House member banks in 1872 and 55% in 1892. Sprague (1910) implicates the seven largest interest-paying banks in 1873, and the definition of an interest-paying bank used here includes Sprague's seven banks. The Coe Report of 1873 issued by the New York Clearing House puts the number at 12. Minor alterations to the classification will not affect the results. Sprague (p. 17–8) notes that the concentration of bankers' deposits was not compensated by size or a more cautious reserve position, as these seven banks held 20% of the capital and 18% of individual deposits of the national banks of New York, as well as the legal minimum reserve ratio of 25%.

Graphical evidence from 1872–3 lends credence to a seasonal explanation of panics. Figures 3.3 and 3.4 record, respectively, the total net deposits for the interest-paying banks and non-interest-paying New York Clearing House member banks in 1872–3. Both interest-paying banks and non-interest-paying banks appear to suffer reserve drains at the end of the panic, although the drop is sharper for the interest-paying banks. However, the seasonality experienced by the non-interest-paying banks is rather surprising. Figures 3.3 and 3.4 look nearly identical. Remember, the non-interest-paying banks hold only a third of the bankers' balances. Yet from seasonal peak to trough in autumn 1872 these banks lose more deposits than the interest-paying banks. This suggests that seasonal effects are not limited to agricultural withdrawals by interior banks or that bankers' balances do not proxy interior deposits.

Graphical evidence from 1892–3 suggests that a seasonal drain did not take place before the panic. Interestingly, aggregate data on member banks in Figure 3.2 does not record the same seasonal pattern as in the earlier years of 1872–3. We know that seasonal fluctuations persisted because Miron (1986) and Champ, Smith, and Williamson (1996) document the seasonality of deposits and nominal interest rates in New York at least until 1910. If anything, non-interest-paying banks rather than interest-paying banks suffered the reserve drain just before the data ends. Figures 3.5 and 3.6 depict, respectively, the total net deposits for interest-paying and for non-interest-paying New York Clearing House member banks in 1892–3. Both sets of banks experience a decline in total net deposits over time. In the last three weeks, the total net deposits of the interest-paying banks continue their mild decline. In contrast, the non interest-paying banks suffer a sharp decrease in total net deposits. This decrease offsets the sharp increase in deposits of late April and early May 1893. So it appears unlikely that econometric analysis will uncover a large shock to the interest-paying banks toward the end of the series. Given the time of the year of each panic, we would probably expect a seasonal drain in autumn 1873 but not in mid-summer 1893, and the tests below agree with this prediction.

Several other authors test the importance of seasonal capital flows. Counter to Sprague and Kemmerer, Goodhart (1969) argues that seasonal westward capital flows should induce higher interest rates in New York and hence create an equilibrating eastward capital flow in the autumn. James ([1978], p. 127–148) critiques Goodhart's theory and evidence. Goodhart's argument characterizes equilibrating capital flows after seasonal flows have taken place, so seasonal flows could still exist. Empirically, 1900–13 was not a representative period of seasonal flows. More importantly for this chapter, Sprague (1910) attempts to link seasonal flows with the panic of 1873. Sprague compares the reserve position of the two classes of banks just before the outbreak of the panic of 1873 using the weekly statements of Clearing House member banks (p. 34–5). He finds a decline in the reserve ratio of both interest-paying and non-interest-paying banks before the panic of 1873. For the purpose of testing the hypothesis that interest-paying banks suffered reserve drains, we want to try to isolate reserve flows rather than examine the reserve ratio. The test reformulates Sprague's method and tests for statistical significance.

Test

This section tests for a reserve drain on interest-paying banks just before the panics begin. Interest-paying banks, by definition, held a large fraction of their deposit portfolio in bankers' balances. If seasonal withdrawals are the exogenous liquidity requirement that triggers crises, we expect interest-paying banks to suffer large percentage reserve drains. We test this hypothesis by examining the aggregate percentage deposit changes of both classes of banks just before the panic begins. Rather than focusing on the reserve losses of individual banks, the test focuses on the aggregate percent changes in deposits of the two classes. Holding bankers' balances is not necessarily the most precise proxy for accepting interior deposits. Slight differentials in the proportion of bankers' balances to total deposits may not correlate with an increase in the vulnerability to seasonal withdrawal. Most likely, the large disparity between the aggregates of the two classes will correlate with seasonal drains if they exist.

In order to test the hypothesis, this chapter employs a time series cross section (TSCS) regression model with separate time-specific variables for each class of banks. A regression model easily accommodates a more robust correlation structure for the residuals. Let IP_{it} be an indicator variable that is 1 if bank i is an interest-paying bank at time t and zero otherwise. Let NIP_{it} be an indicator variable that is 1 if bank i is not an interest-paying bank at time t and zero otherwise. Consider the following TSCS model of bank deposits:

$$y_{it} = \mu + a_i + \gamma y_{it-1} + \theta_t (IP_{it}) + \lambda_t (NIP_{it}) + \varepsilon_{it}, \qquad (3.1)$$

where μ is the constant term, a_i captures individual bank effects, and ε_{it} is an error term for bank i at time t. The individual bank subscript i ranges from 1 to N and the time subscript t ranges from 2 to T because of the lagged dependent variable. The variables θ_t and λ_t represent time shocks common to interest-paying banks and non-interest-paying banks, respectively, at each time t. Additional lagged dependent variables could be included in (3.1), but estimation indicates that the fit was not dramatically improved, so these are omitted. The seasonality apparent in Figure 3.1 should be modeled as well, but capturing annual seasonal fluctuations with at most 20 months of data proves intractable.

The presence of both individual effects and time indicator variables raises problems of perfect multicollinearity. Typical estimation sweeps out the time effects by a matrix operation (see Hsaio [1986], p. 53). In this application, the time indicator variables are the parameters of interest. In order to avoid perfect multicollinearity, at least two indicator variables must be dropped from the model. Interpretation of the time coefficients as time shocks to each class of bank requires the presence of all time indicator variables. Removing a time indicator variable changes the interpretation of the remaining coefficients from a mean shock to a comparison with the omitted period. So, actual estimation omits two individual fixed effect parameters (one of each type of bank) as well as the constant term. The qualitative results turn out to be invariant to the individual bank variables excluded.

Instead of estimating the model with a complete set of T-1 time indicator variables, consider a restricted form of (3.1). Suppose the right-hand side variables consist only of a complete set of T-1 time indicator variables. Then the residual covariance matrix estimated by more general regression techniques such as feasible generalized least squares is singular. In the presence of other regressors the covariance matrix is invertible but becomes ill-conditioned, causing estimated standard errors take implausibly low values. Since we are interested in the behavior of bank reserves just before the panic begins, it is convenient to suppose that the mean drain for both classes of banks is constant over time except for just before the panics. Equation (3.2) restricts (3.1) by equating early time shocks for the first T-3 time periods:

$$y_{it} = \mu + a_i + \gamma y_{it-1} + \theta_0 (IP_{it}) + \lambda_0 (NIP_{it}) + \varepsilon_{it.} \quad t < T\text{-}2 \qquad (3.2)$$
$$y_{it} = \mu + a_i + \gamma y_{it-1} + \theta_t (IP_{it}) + \lambda_t (NIP_{it}) + \varepsilon_{it} \quad t = T\text{-}2, T\text{-}1, T.$$

Equation (3.2) constrains the first T-3 time shocks to be equal for each class of banks. The new parameter θ_0 enforces the restriction $\theta_2 = \theta_3 = ... = \theta_{T-3}$ and λ_0 represents the restriction $\lambda_2 = \lambda_3 = ... = \lambda_{T-3}$. Equation (3.2) then allows separate shocks for the last three periods for each class of bank. Estimating two or four time effects separately instead of three yields similar results.

Equation (3.2) does not provide any causal information about the source of the shock. The purpose of (3.2) is to determine whether interest-paying banks suffered some kind of reserve drain in the last three periods before the panic. Under the seasonal hypothesis, we expect large capital flows leaving New York just before the panic. But even if we observe a drain on interest-paying banks, observed deposit changes need not have a seasonal origin. However, if we do not observe large changes in deposits, then perhaps we would have some evidence that models based on large seasonal outflows might not agree with the data. Further, we observe only the net changes in deposits, so the presumption is that no large offsetting inflows mask the possible existence of seasonal outflows.

Testing reserve drains requires standardization of the dependent variable by the size of the bank. Interest-paying banks were larger than non-interest-paying banks. Naturally, we expect interest-paying banks to experience a greater outflows of reserves simply because they have more liabilities. One option is to consider the percentage change in deposits as the dependent variable. Interest-paying banks held a greater proportion of interior balances in their deposit portfolio. In the event of large withdrawals from the interior, we expect interest-paying banks to sustain a larger percentage drain in deposits than non-interest-paying banks. But the average of the percent changes need not resemble the percent change of the aggregate, especially for small percentages as in the present case.

A slight transformation of the dependent variable improves the interpretability of the coefficients. The transformation allows OLS estimation of the coefficients θ_t and λ_t to be interpreted as aggregate percentage changes for their respective class of banks at time t. Instead of dividing changes in deposits for bank i by the deposits of bank i as in the percent change, divide by the average deposits of the class of bank i (interest-paying or not). Let C_i be the set of banks that are the same type of bank (either interest-paying or not) as bank i. For example, if bank i is an interest-paying bank, then C_i is the set of interest-paying banks. Let TND_{it} be total net deposits (net deposits plus bank note circulation) for bank i at time t. Define y_{it} as

$$y_{it} = \frac{100(TND_{it} - TND_{it-1})}{(1/|C_i|)\sum_{j \in C_i} TND_{jt-1}}, \tag{3.3}$$

where $|C_i|$ is the number of banks in the class of bank i. Multiplying by 100 yields a percentage. If time indicator variables are the only parameters on the right-hand side of (3.2) and the dependent variable y_{it} is defined as in (3.3), then OLS estimation of the parameters θ_t and λ_t measure the aggregate percent change in deposits of the interest-paying banks and the non-interest-paying banks at time t, respectively.

To verify the interpretation of the coefficients, suppose that a complete set of time indicator variables are the only RHS variables in the model. That is, consider estimation of (3.1) with $\mu = \gamma = a_i = 0$ for all i. As is well known, OLS estimation of a panel time indicator variable calculates the cross-sectional mean of the dependent variable at time t. Because the time indicator variables for the two classes of banks are orthogonal, joint estimation including both θ_t and λ_t yields the same estimates as separate OLS estimation for the two classes of banks. The cross-sectional mean of the dependent variable in (3.3) for each class of banks is

$$(1/|C_i|)\sum_{k\in C_i}\left(\frac{100(TND_{kt}-TND_{kt-1})}{(1/|C_i|)\sum_{j\in C_i}TND_{jt-1}}\right)=\frac{\sum_{k\in C_i}100(TND_{kt}-TND_{kt-1})}{\sum_{j\in C_i}TND_{jt-1}}, \qquad (3.4)$$

or the percentage change of total net deposits for the class of bank i at time t. So OLS computes θ_t and λ_t as the aggregate percent change in the total net deposits of interest-paying banks and non-interest-paying banks respectively. Further, the restriction of equal means for the beginning time periods $\theta_2 = ... = \theta_{T-3}$ in (3.2) does not change the interpretation of θ_{T-2}, θ_{T-1}, or θ_T as aggregate percent changes for the interest-paying banks at times T-2, T-1, or T. However, if additional variables such as lagged dependent variables or individual fixed effects are included in the RHS of Equation (3.2), then the coefficients θ_t and λ_t can no longer be interpreted as unconditional mean percent changes.

Chapter 3 tests for a seasonal interpretation of the panics. There are two hypotheses of interest. First, under the seasonal hypothesis the interest-paying banks must experience a decline in deposits. However, this is only a necessary condition as a reserve drain on interest-paying banks need not have a seasonal origin. Second, an alternate possibility is that the exogenous shock fell upon the non-interest-paying banks. Instead of a story about an exogenous seasonal withdrawal, perhaps some other explanation describes why non-interest-paying banks sustained an exogenous shock. In this case we are interested in the withdrawals on non-interest-paying banks, as well as comparing the withdrawals sustained by the two classes of banks. If neither type of bank (nor the aggregate of both types) suffers an exogenous withdrawal just before the panic begins, then Diamond and Dybvig cannot explain panics at all.

A first test determines if interest-paying banks suffered a reserve drain before the panics, whether of seasonal origin or otherwise. Time dummy variables model shocks to the money market. At each time t, non-interest-paying banks experience a common reserve shock θ_t, while interest-paying banks experience a common reserve shock λ_t. The question is whether or not the magnitude of θ_t is large. Even though most versions of Diamond and Dybvig's model can be supported by an arbitrarily small liquidity requirement, most subsequent literature focuses on historically large seasonal shocks, so we will continue in that tradition here. The null hypothesis states that interest-paying banks did not suffer reserve drains in the last few periods just before the panic. One test of the null hypothesis is $\theta_t = 0$ where t = T-2, t = T-1 and t = T. A simple t test evaluates the null hypothesis. We can test for joint significance of the three coefficients by the usual Wald test. If some of the coefficients θ_t from t = T-2 to T should be statistically significant, we would reject the null hypothesis and consider the existence of seasonal drain. We can repeat the same method on the λ_t coefficients to determine whether or not the non-interest-paying banks suffered a drain, in case a story about non-interest-paying banks suffering an exogenous shock fits the data better.

A variation on the first test compares shocks just before the panic to the weekly mean of previous shocks. Instead of testing the hypotheses $\theta_t = 0$ and $\lambda_t = 0$, we can test $\theta_t = \theta_0$ and $\lambda_t = \lambda_0$ for t = T, T-1, and T-2. If time dummy variables are the only RHS variables, observe that θ_0 is just the time mean of θ_t over the periods prior to period T-2. The motivation for this test is to determine whether or not the coefficients at time t are equal to the time mean of shocks during non-pre-panic periods. Since Diamond and Dybvig is a static model, the authors do not specify the rate at which deposits enter the bank. So perhaps the relevant comparison for panic withdrawals is the average rate of deposit change over time. For example, if the data should exhibit a downward time trend, then testing against the mean rather than zero could prevent discovering a false drain when a constant trend was responsible. However, under seasonality we might expect some weeks to have larger shocks than others, so the time mean of shocks may not approximate the appropriate distribution. In this data, estimates of the time means θ_0 and

λ_0 happen to be very close to zero, so the results do not depend on the type of test chosen. Appendix 3B summarizes the results for this test.

A second test compares the two classes of banks directly. If both interest-paying and non-interest-paying banks sustain equal shocks, then the seasonal drain does not explain why non-interest-paying banks suffered reserve drains. The question is whether or not the two classes of banks suffered the same magnitude of shocks just before the panic, when seasonal drains allegedly took place but the panic was not yet underway. Under the null hypothesis that interest-paying banks and non-interest-paying banks suffer the same reserve shocks, the two time specific parameters, θ_t and λ_t, should be the same just before the panics. So for t = T, we can test $\theta_T = \lambda_T$. Under the alternative, λ_T should be larger than θ_T because the interest-paying banks suffered the seasonal reserve drain. We can also test similar hypotheses for times T-1 and T-2. Instead of the usual Wald test of the null hypothesis, consider a direct t test.

A direct t test evaluates the second test comparing the two classes of banks. The null hypothesis equates the time dummy parameters θ_t and λ_t at time t. Because the Wald test of the equality of two coefficients is non-directional, rejection of the null hypothesis provides no information about which coefficient is larger. But a direct t test of the hypothesis that $\theta_t = \lambda_t$ provides a signed test statistic. Following Ramanathan (1995, page 183), define $\delta_t = \theta_t - \lambda_t$ and observe $\theta_t = \lambda_t + \delta_t$. Rewrite the regression in (3.2) above as

$$
\begin{aligned}
y_{it} &= \mu + a_i + \gamma\, y_{it-1} + \lambda_0\,(IP_t + NIP_t) + \delta_0\,(IP_t) + \varepsilon_{it} \quad t < T\text{-}2 \qquad (3.5)\\
y_{it} &= \mu + a_i + \gamma\, y_{it-1} + \lambda_t\,(IP_t + NIP_t) + \delta_t\,(IP_t) + \varepsilon_{it} \quad t = T\text{-}2,\ T\text{-}1,\ T
\end{aligned}
$$

by substituting for θ_t for all relevant t. Note that ($IP_t + NIP_t$) is just a time indicator variable for time t. Then carry out an ordinary t test on d_t to test the original null hypothesis $\theta_t = \lambda_t$, that two classes of banks suffered the same drains at time t. Hence, the test statistics of interest are t tests on δ_T, δ_{T-1}, δ_{T-2}, a Wald test of these three coefficients jointly, and a Wald test of the hypothesis that $\delta_t = 0$ for t = T-2, T-1 and T. For example, if δ_t is positive, then $\lambda_t > \theta_t$, so the interest-paying banks would suffer less of a drain than the non-interest-paying banks at time t. The only real difference between (3.2) and (3.5) is the interpretation of the θ and δ coefficients.

One might question whether or not the data accept the unit root restriction described in (3.3). Equation (3.3) takes first differences of the data, a restriction that is only appropriate when the data contain a unit root. Chang (2002) proposes a panel unit root test that allows for cross-sectional residual correlation. Consider a panel autoregression

$$y_{it} = a_i y_{it-1} + u_{it}, \tag{3.6}$$

where u_{it} is a stationary autoregressive process AR(p_i) that may differ across each cross-sectional unit i. The null hypothesis states that for all i $a_i = 1$, while the alternative allows that for some i $|a_i| < 1$. For each cross-sectional unit i, the test calculates the t score from a test of the hypothesis that a_i equals 1 in a nonlinear instrumental variables regression of (3.6). The function

$$F(y_{it-1}) = y_{it-1} \exp(-c\, y_{it-1}), \tag{3.7}$$

where c is a constant, serves as a nonlinear instrument for the lagged dependent variable. Under the null hypothesis $\Delta y_{it} = u_{it}$, so each instrumental variables regression is augmented by p_i lags of the first difference of the dependent variable. Chang shows that each t test is asymptotically (in the time dimension T) distributed standard normal. The test computes S_N, a cross-sectional average of the t-scores, which is also asymptotically (in T) normally distributed. In contrast to previous panel unit root tests, such as the t-bar test of Im, Pesaran and Shin (1997) and the Fisher test by Maddala and Wu (1999), the Chang S_N test explicitly allows for cross-correlation among the individual panel units. A test for cross-correlation reported below suggests the data require a unit root test with this feature. A small Monte Carlo examination of size and power by Chang shows that the test performs reasonably well for the size of data set studied here. For more information about the application of the Chang S_N test to this data set, see Appendix 3C.

The Chang S_N test rejects the null hypothesis of a panel unit root for the 1873 and the 1893 data. The S_N test on total net deposits yields an S_N statistic equal to –10.23 in 1873 and -3.38 in 1893 with four autoregressive lags ($p_i = 4$ for all i). Since S_N is asymptotically standard normally distributed, the S_N test rejects the null hypothesis of a panel unit root for both panels. The rejection of the null hypothesis implies that at least one of cross-sectional unit series is stationary.

Despite the rejection of the null hypothesis, the unit root was imposed for two reasons. First, testing the static theory requires investigating changes. Diamond and Dybvig's model is a static one-shot game with no prediction about the evolution of the level of deposits over time. However, the seasonal interpretation assumes that the level of deposits changes prior to bank runs. Since the data is dynamic, we need to transform the data in order to test the hypothesis.

The second reason for imposing the unit root is the lack of a practical econometric alternative. Two possibilities under the alternative hypothesis are that the common autoregressive parameter is stationary (for all i $a_i = a < 1$) or that the autoregressive parameter is heterogeneous with some cross-sectional autoregressions possibly nonstationary. In the first scenario, OLS fixed effects regressions yields estimated autoregressive parameters a in excess of 0.99. In this case the data might as well be modeled with a unit root, as the estimated confidence intervals do not change much. The second scenario seems more likely, with some cross-sectional units exhibiting a unit root and others not. Individual unit root tests indicate the presence of a unit root in some cross-sectional units. Unfortunately, there is little guidance about how to proceed when parameters are heterogeneously stationary or nonstationary. Pesaran and Smith (1995) warn about the importance of allowing heterogeneous lagged dependent variables, but most papers consider only stationary autoregressions. Each individual cross-sectional series could be differenced based on its own unit root test. But unmodeled residual correlation across units makes individual tests inefficient and probably biases the tests. A possibility might be to apply Chang's S_N test to subsets of the data, but the selection of subsets may influence the classification of series as I(1) or I(0). Due to the requirements of the theory and in the absence of a practical alternative procedure, for now we impose the unit root on the data despite its statistical rejection. The rejection of a common unit root prevents panel-wide cointegration analysis.

Estimation

Now we turn to estimation of (3.2). Since the time parameters θ_t and λ_t are simply time indicator variables, two standard estimation procedures are fixed effects and random effects. The hypothesis of interest recommends fixed effects estimation over random effects. In the fixed effects procedure, parameters are estimated conditional on the fixed sample. Under random effects, the values of the parameters are drawn from a larger population. Fixed effects models are considered more appropriate than random effects models for inference about a fixed sample (Hsaio [1986], p. 43). This chapter tests a hypothesis about seasonal drains on the specific banks in the sample. Further, as the time dimension T becomes large random effects estimates become fixed effects estimates (Hsaio, p. 49). So for both methodological and practical reasons the method of choice is fixed effects.

Time series cross section (TSCS) methods may allow asymptotic results as T becomes large. Panel data sets often involve only a few time periods T but a large number of cross-sectional observations on individuals N. In contrast, these data sets are nearly square with T = 91 and N = 60 in 1872–3 and T = 76 and N = 63 in 1892–3. In this case, it might be reasonable to employ asymptotic results in the T direction, such as OLS fixed effects (also known as least squares dummy variables, LSDV). The coefficient on lagged dependent variables estimated by LSDV is consistent but biased in small samples (Nickell [1981]). How large is the bias for moderately sized samples? Judson and Owen (1999) perform a Monte Carlo study to determine the magnitude of the bias of LSDV regression coefficients. In the trial most closely paralleling the results below with N = 100, T = 30, and a lagged dependent variable coefficient of 0.2, the average bias was about 10% of the lagged dependent variable coefficient's actual value, although biases of up to 20% fell within two standard deviations. In our case, T = 76 or 91 so the bias is probably about 5% or 3% of the coefficient's actual value since the bias is O(1/T). While it might be possible to employ a bias corrected estimation procedure such as Kiviet (1995), given the small magnitude of the bias and the eventual statistical insignificance of the coefficient it seems reasonable to proceed with LSDV estimation.

Organize the data in long form by grouping the regression equation in (3.2) for each bank i together but in time order. That is, form $y_i = (y_{i1}, ..., y_{iT})'$ and the data vector $x_i = (x_{i1}, ... , x_{iT})'$ where x_{it} is a K x 1 vector of data for the K RHS variables in (3.2) for bank i at time t. Construct the matrix $Y = \text{vec}(y_1, ..., y_N)$ by stacking the N y_i vectors on top of each other, and construct the matrix $X = \text{vec}(x_1, ... ,x_N)$ in the same way. Now the data are organized in long form, with each row of the matrix (Y X) forming an observation on bank i at time t according to (3.2). Let e_{it} be the estimated residual for bank i at time t from the OLS regression of Y on X.

The OLS variance-covariance matrix may require a more robust specification. OLS specifies that the residual covariance matrix is of the form $\sigma^2 I$. This description of the residuals imposes several restrictions. First, estimation by OLS assumes that the residual for bank i at time t is independent from the residual for bank j at time t. Also, OLS assumes that the residual variance σ^2 is constant across banks. Both assumptions are suspect in this application. If banks experience a common monetary shock, there might be cross-sectional correlation at time t among the banks. Further, if some banks experience more variation in withdrawals due to seasonal patterns, their deposits may experience a higher variance. Another possibility is for autocorrelation of the residuals in time. Although OLS regression coefficients (without lagged dependent variables) are still consistent in the presence of these effects, estimates of the standard errors are no longer correct, so tests of hypotheses are invalid.

A standard method of allowing for more robust error structures is feasible generalized least squares (FGLS) as described by Greene (1993). Instead of the usual matrix $\sigma^2 I$, suppose that the residual covariance matrix is given by a more general matrix O. If the residual covariance matrix O was known, we could transform the data by the NT x NT matrix L where L'L = $\hat{\Omega}^{-1}$ and run OLS to obtain fully efficient GLS estimates. In practice, the residual covariance matrix O is unknown, and we perform the GLS transformation with a consistent estimate $\hat{\Omega}$ of the residual covariance matrix O to obtain asymptotic efficiency. In this case, we allow for cross-unit correlation and heteroscedasticity but require that the residuals be uncorrelated over time. Let $\hat{\Sigma}$ estimate the N x N matrix of contemporaneous covariances with typical element:

$$\hat{\Sigma}_{ij} = \left(\sum_{1}^{T} e_{it}\, e_{jt}\right)/T \tag{3.8}$$

The matrix $\hat{\Sigma}_{ij}$ estimates the correlation between banks with the time average of the correlations of the OLS errors. The full estimated panel-corrected variance covariance matrix $\hat{\Omega}$ of the error terms is then

$$\hat{\Omega} = \hat{\Sigma} \otimes I_T \tag{3.9}$$

of dimension NT x NT. The block diagonal form of $\hat{\Omega}$ embodies the assumption that the errors of different banks are uncorrelated in different time periods. The $\hat{b}_{FGLS}$ estimates are

$$\hat{b}_{FGLS} = (X'\hat{\Omega}^{-1}X)^{-1}(X'\hat{\Omega}^{-1}Y) \tag{3.10}$$

with an estimated covariance matrix of

$$\text{Est. Cov}(\hat{b}_{FGLS}) = (X'\hat{\Omega}^{-1}X)^{-1}. \tag{3.11}$$

Note that if the estimated residual covariance matrix $\hat{\Omega}$ were the usual $\hat{s}^2$I, then the estimated coefficients in (3.10) are just OLS estimates and the Est. Cov($\hat{b}$) matrix collapses to the usual OLS covariance matrix $\hat{s}^2(X'X)^{-1}$. FGLS with cross-correlated errors is possible in this context because T ≥ N. If T < N, then the N x N matrix $\hat{\Sigma}$ (and hence $\hat{\Omega}$) would not be invertible (rank($\hat{\Sigma}$) ≤ min{N,T}).

A more robust estimation procedure than FGLS employs panel-corrected standard errors (PCSEs). Beck and Katz (1995) propose PCSEs as a method of correcting for cross-correlation and heteroscedasticity. In the presence of cross-sectional correlation and heteroscedasticity, OLS estimates are consistent but estimates of the covariance matrix are not. The method of PCSEs retains the OLS coefficient estimates and calculates their covariance under the more general estimated residual covariance matrix $\hat{\Omega}$. The estimate of the $\hat{\Omega}$ matrix turns out to be identical to the estimated residual matrix of FGLS, given by (3.8) and (3.9). Write $\hat{b}$ for the vector of estimated OLS coefficients, which the method of PCSEs retains without alteration. Calculate the variance-covariance matrix of $\hat{b}_{PCSE} = \hat{b}$ by the formula

$$\text{Est. Cov}(\hat{b}_{PCSE}) = (X'X)^{-1}(X'\hat{\Omega}X)(X'X)^{-1}. \tag{3.12}$$

Equation (3.12) is the usual covariance matrix for the OLS estimator. Again, if the estimated residual covariance matrix $\hat{\Omega}$ were equal to $\hat{s}^2$I as assumed under OLS, then the Est. Cov($\hat{b}$) matrix collapses to the usual OLS covariance matrix $\hat{s}^2(X'X)^{-1}$. Beck and Katz (1995, 1996) present Monte Carlo evidence that suggests that OLS with PCSEs performs better than other FGLS methods in sample sizes that are observed in practice.

To avoid problems with an unbalanced panel, the estimation drops one small, non-interest-paying bank from each panel. The 1873 data include a bank that closed in May 1873 as the result of a defalcation by a bank officer. The 1893 data include a bank that joined the Clearing House in March 1892. While the least squares dummy variable technique (LSDV) with time specific indicator variables easily handles the unbalanced panel (Greene 1993, p. 623), estimation by PCSEs is less straightforward. One option, casewise deletion, constructs the variance-covariance matrix $\hat{\Sigma}$ using only observations in time where all cross-sectional units are present. This method omits the data when a bank was absent from the calculation of $\hat{\Sigma}$, or about nine weeks of data in the 1893 crisis and about 16 weeks for 1873. An alternative, pairwise deletion, forms the variance-covariance matrix $\hat{\Sigma}$ by using all available pairwise combinations of errors. Pairwise deletion raises questions about the small sample properties of the covariance matrix. Cross-correlation among banks appears more of a concern than the behavior of a new bank or a bank that closed for exogenous reasons. So the results reported in this chapter simply remove the extra banks. Adding the extra banks and estimating $\hat{\Sigma}$ by casewise or pairwise deletion does not affect the qualitative results. None of the other 63 banks failed or otherwise left the Clearing House in 1893. In 1873, one bank failed in the last week of the data set, but a report from the receiver provides data for the last week. For more information about the data, see Appendix 3A.

Results

Before evaluating the results, several robustness tests select an appropriate model. A test detects cross-sectional correlation. Greene ([1993], p. 661) suggests a Breusch-Pagan (1980) Lagrange multiplier test of cross correlation of the residuals. The Breusch-Pagan LM test statistic, a time multiple of the sum of the lower triangular block (not including the diagonal) of the covariance matrix $\hat{\Sigma}$, is

$$\mathrm{BP} = \mathrm{T} \sum_{i=2}^{N} \sum_{j=1}^{i-1} r_{ij}^{2} \ , \tag{3.13}$$

where r^2_{ij} is the estimated Pearson correlation coefficient for the FGLS residual e_{ij}. Greene (1993) notes that the appropriate null hypothesis allows for heteroscedastic errors without cross-sectional correlation since the coefficients of interest are pooled. Under the null hypothesis, BP is asymptotically distributed chi-squared with N(N-1)/2 degrees of freedom. Even though the FGLS estimates are not maximum likelihood, Greene (1993) suggests basing the test on the FGLS residuals. In this case, FGLS estimation of (3.2) for the 1873 data yielded a test statistic BP = 3630 ($p < 0.01$), while for the 1893 data BP =

3553 ($p < 0.01$), easily rejecting the null hypothesis of no cross-sectional correlation in the residuals. Hence, the analysis requires more robust regression techniques than OLS such as FGLS or OLS with PCSEs.

A comparison of OLS, FGLS, and OLS with PCSEs suggests that OLS and FGLS produce overconfident standard errors. Tables 3.1 and 3.2 present initial estimates of (3.2) by OLS with PCSEs for 1873 and 1893 respectively. Beck and Katz (1995) warn that FGLS produces overly confident standard errors. Several of the FGLS estimated standard errors are about one-half or one-third the standard errors of their OLS or PCSE counterparts. However, coefficients estimated by FGLS appear similar to the results of OLS (and PCSEs). Since the data reject the OLS error specification and the FGLS standard errors appear too small, this chapter selects the PCSEs of Beck and Katz (1995) in order to test hypotheses. Note that the time indicator variables in these regressions do not measure aggregate percent changes in deposits because the regressions condition on fixed effects and include the lagged dependent variable.

Estimation by OLS with PCSEs allows several exclusion restrictions. The rightmost column in Tables 3.1 and 3.2 record estimation of (3.2) by OLS with PCSEs. A t test, which is asymptotically valid in this context, shows that the lagged dependent variable was insignificant ($p = 0.54$) and small in magnitude for the 1893 data. Subsequently, a joint F test restricting the individual fixed effects to be equal was accepted ($p = 0.88$), and the resulting indicator variable was insignificant as well ($p = 0.81$). For the 1873 data, the lagged dependent variable was marginally significant ($p = 0.07$), but the magnitude was reasonably small, with $\gamma = -0.07$. Since the results are the same when the lagged dependent variable is present and the significance is marginal, we take the liberty of dropping the variable. The individual effects present accept the restriction of a common intercept using an F test ($p = 1.00$), and the resulting indicator variable was insignificant as well ($p = 0.65$). The insignificance of the individual fixed effects is not entirely surprising. When the dependent variable represents changes, as in (3.3), the fixed effect is a drift term. Observe that bank deposits are not particularly monotonic, at least in the aggregate. For the remainder of the chapter the individual fixed effects and the lagged dependent variable are dropped. Now the only RHS variables remaining in either model are the time indicator variables. The restrictions allow interpretation of the time indicator variables θ_t and λ_t as aggregate percent changes in TND for the interest-paying banks and the non-interest-paying banks respectively.

Tests for autocorrelation reveal that the residuals of some of the estimates suffer from a small amount of autocorrelation. Since the lagged dependent variable was set to zero, serial correlation of the residuals allows consistent estimation of the coefficients but produces biased standard error estimates, invalidating tests of hypotheses. Lagrange multiplier (LM) tests for first-order autocorrelation detect a small amount of autocorrelation in the residuals for both 1873 and 1893. The LM test regresses the OLS residuals on the lagged residuals and the right-hand side variables of the restricted form of (3.2), without individual fixed effects or lagged dependent variables. The 1893 estimated first-order autocorrelation coefficient $\hat{r}$, restricted to be equal across panels, is 0.028. The test statistic $(NT - N)R^2 = 3.73$ is distributed $\chi^2(1)$, marginally accepting the null hypothesis of no autocorrelation ($0.10 > p > 0.05$). As for the 1873 data, the autocorrelation coefficient $\hat{r}$ equals -0.065 and the test statistic $(NT - N)R^2 = 22.96$, rejecting the null hypothesis of no autocorrelation ($p < 0.01$). Although the autocorrelation is small enough to be ignored, we investigate estimation procedures that allow for autocorrelation.

Suppose that the residuals are only first-order serially correlated (AR(1)) within their own cross-sectional unit. One solution, proposed by Beck and Katz (1996), includes a lagged dependent variable to account for autocorrelation. As described above, the coefficient was marginally significant and the results remain the same. Another remedy employs the Prais-Winsten transformation, an instance of FGLS, to eliminate the autocorrelation (Greene [1993], p. 663). Let $\hat{r}_i$ be a consistent estimate of the autocorrelation parameter ρ_i for cross-sectional unit i. Using the above notation for the data, define

$$y_{it}^* = y_{it} - \hat{r}_i y_{it-1} \qquad \text{if } t \geq 2 \qquad (3.14)$$

$$y_{i1}^* = \left(\sqrt{1 - \hat{r}_i^2}\right) y_{i1} \qquad t = 1 \qquad (3.15)$$

for each y_i and apply the same transformation to x_i for each i. Equation (3.14) preserves the first data point. We can then apply OLS with PCSEs to the transformed data in order to account for cross-correlation among the errors. So the estimation procedure first obtains a consistent estimate of the autocorrelation parameter, applies the Prais-Winsten transformation to the data, and then estimates the model by OLS with PCSEs. Beck and Katz (1995) counsel against allowing individual autocorrelation parameters for each cross-sectional regression and suggest restricting the autocorrelation parameter ρ_i to be constant across regressions. These authors see no reason to let the nuisance autocorrelation parameter vary across regressions when the parameters of interest are pooled. Following this advice, the mean of the individual estimated autocorrelation parameters provides a consistent estimate of the common restricted autocorrelation. The initial estimate $\hat{r}_i$ of the autocorrelation parameter ρ_i of each individual cross-sectional unit can be any consistent estimator, so for each cross-sectional unit we use the estimate obtained by an autoregression on the OLS residuals.

A correction for autocorrelation leaves the results unchanged. The estimated autocorrelation coefficient $\hat{r}$ was –0.01 for the 1873 data and –0.076 for the 1893 data, similar to the pooled coefficient from the LM test. Because the autocorrelation was so small, the results were nearly identical to the regression without autocorrelation, and so they are not reported.

With the appropriate model selected, we turn to the estimation results. The results presented here focus on the last three weeks before the panic. For alternate results incorporating the last five weeks, that for the most part preserve the qualitative results discussed here, see Appendix 3A. The magnitudes of the regression coefficients admit a seasonal explanation in 1873 and refute the seasonal hypothesis in 1893. Table 3.3 presents OLS estimation with PCSEs of (3.2), where the lagged dependent variable and the individual fixed effects are dropped. The very low R-squared is expected due to the lack of explanatory variables. First, consider the evidence from the panic of 1873. The interest-paying banks clearly suffered a strong reserve drain in the last three weeks of about 5% per week, with the drain in the last week equal to θ_T = 5.9%. However, the non-interest-paying banks also sustained withdrawals, as their total net deposits declined about 2% per week in the last three weeks, with a drain in the last week equal to λ_T = 3.0%. So it appears plausible that the interest-paying banks suffered larger drains than the non-interest-paying banks in 1873, supporting the seasonal hypothesis. In contrast, the 1893 results show that the interest-paying banks suffered much less of a shock than the non-interest-paying banks. From Table 3.3, the deposits of the non-interest-paying banks declined about λ_T = 3.8% the week before the panic, while the interest-paying banks suffered a much smaller decline of about θ_T = 1.1%. No class of banks reports sizable drains larger than 2% before the last week. So on the eve of the panic of 1893, interest-paying banks did not suffer a sizable shock, and they suffered less of a shock to reserves than non-interest-paying banks. The direction of the effect renders a seasonal explanation for this panic unlikely.

The first test of the seasonal hypothesis test reveals that the interest paying banks suffered a statistically significant reserve drain in 1873 while the non-interest-paying banks did not. One test of the existence of a drain performs an ordinary t test on the θ_t coefficients with the null hypothesis that the coefficient is zero. Table 3.3 reports coefficients estimated from (3.2), the usual indicator variable form of the test. For times t = 89, 90, and 91, the t test shows that the interest-paying banks did suffer significant drains (p = 0.07, 0.06, and 0.01, respectively). A Wald test of the joint null hypothesis that the three θ_t coefficients are insignificant ($\theta_{T-2} = \theta_{T-1} = \theta_T = 0$) is also rejected ($p < 0.01$). In contrast, the non-interest-paying banks do not appear to have suffered a statistically significant drain, despite losing nearly 3% a week in deposits. The t test for the λ_t coefficients for times t = 89, 90, and 91 are insignificant (p = 0.29, 0.66, and 0.20, respectively), and naturally the joint Wald test of the insignificance of the three coefficients is also accepted (p = 0.40). So, it appears that a seasonal drain on interest-paying banks is a plausible explanation for the panic of 1873.

Rather than testing for the existence of reserve drains directly, the second test compares the two classes of banks. Paradoxically, while earlier evidence for 1873 showed that the interest-paying banks suffered a drain while other banks did not, this test shows no significant difference between the reserve drains of the two classes of banks. The test of the hypothesis that the interest-paying banks suffered the same reserve shocks as non-interest-paying banks is given by a test of the significance of the δ_t coefficients in (3.5). Table 3.4 presents direct t test estimates of (3.5) estimated by OLS with PCSEs. For the 1873 panic, the data do not reject the hypothesis of a seasonal drain despite the large magnitude in differences of the reserve drains. The t tests on the coefficients δ_t for t = 89, 90, and 91 accepts the null hypothesis of equal drains (p = 0.51, 0.19, and 0.18, respectively). The joint hypothesis that all three are insignificant is also accepted (p = 0.30). So we cannot reject equality of the two coefficients. Most likely, this is a power problem. The standard errors for the delta parameters were about 2.4%, so a difference of on the order of 3% will probably not be significant unless the parameters are very highly correlated. So according to the first test there is some evidence that the data admit a seasonal explanation, although the second test accepts the hypothesis that the two classes of banks suffer equal drains.

In contrast, the panic of 1893 produces the opposite conclusion, with non-interest-paying banks suffering a reserve drain in the last week while interest-paying banks were unaffected. Table 3.3 also reports the results for estimation of (3.2) from 1893. For times t = 74, 75, and 76, the interest-paying banks clearly did not experience a statistically significant reserve drain (p = 0.72, 0.97, and 0.49, respectively). Of course the joint test of ($\theta_{T-2} = \theta_{T-1} = \theta_T = 0$) is also insignificant (p = 0.90). But the non-interest-paying banks clearly suffered a reserve drain in the last week before the panic. The data reject the null hypothesis that $\lambda_T = 0$ (p = 0.01), but accept the null for t = T-1 (p = 0.20) and T-2 (p = 0.81). A Wald test that the three coefficients are jointly insignificant is rejected due to the drain in the last week (p = 0.04). In this case, the interest-paying banks did not suffer reserve drains before the panic. Not surprisingly given the month the drain started, the panic of 1893 rejects the seasonal interpretation of the theory of Diamond and Dybvig.

The comparison test of the two classes of banks using (3.5) does not reject the hypothesis that the reserve drains experienced by the two classes of banks are approximately the same for 1893. From Table 4, the t tests on the coefficients δ_t for t = 74, 75, and 76 accepts the null of equal drains (p = 0.92, 0.29, and 0.11, respectively), although the coefficient d_T almost marginally rejects the null hypothesis in the wrong direction. Naturally the joint null hypothesis is accepted as well (p = 0.30). So this test upholds the null hypothesis that the two banks suffered identical shocks before the panics. The results of the second test appear to contradict the results of the first test, which stated that non-interest-paying banks did suffer a drain in the last period while interest-paying banks did not. But clearly, the interest-paying banks do not suffer large reserve drains. If anything, interest-paying banks suffered less of a drain then non-interest-paying banks as the sign of both δ_T and δ_{T-1} coefficients are positive. The evidence shows that interest-paying banks did not suffer an unusual drain before the panic of 1893. We reject seasonal drains as a trigger for the panic of 1893. However, because of the significant drain on the non-interest-paying banks in the last week, the data allow that an exogenous drain could have struck the non-interest-paying banks. The drain on the non-interest-paying banks simply reverses a small previous increase in the series. The theory of Diamond and Dybvig (1983) could apply if the theory was reinterpreted to explain a drain on non-interest-paying banks.

Summing up, the first test shows that in 1873 the interest-paying banks suffered a large reserve drain, so the data may admit a seasonal explanation. In 1893, if any banks suffered a reserve drain it was the non-interest-paying banks, not the interest-paying banks. We can reject Calomiris and Gorton's seasonal interpretation of the theory of bank runs by Diamond and Dybvig for the panic of 1893.

A final comment investigates the sequential effect of the panic of 1873 upon the panic of 1893. Suppose that market participants use the experience of 1873 to forecast withdrawals during the panic of 1893. To examine this possibility, fit the 1893 data with

OLS estimates of the 1873 coefficients and generate the resulting error terms. Naturally, the resulting error terms are often positive, because a large drain took place in 1873 but not in 1893. Approximately 12 of the 13 interest-paying banks had positive errors in each of the last three periods. As for the non-interest-paying banks, approximately 36 of the 50 non-interest-paying banks had positive errors in each of the last three periods. This information does not constitute a formal test, as the errors are correlated across banks. But the exercise suggests one of two scenarios. Perhaps investors, based on their experience of 1873, expected a large decline in reserves that did not occur in 1893. In this case the expectation of a bank panic occurring before the data ends was incorrect. Alternatively, the panic had not yet begun and the wise preemptive action of the New York Clearing House that marks the end of the data set was an attempt to prevent the panic before it could start.

Conclusion

Chapter 3 conducts an econometric test of Calomiris and Gorton's (1991) interpretation of the theory of Diamond and Dybvig (1983). Diamond and Dybvig use an exogenous liquidity shock to depositors in order to develop a theory of bank runs. Calomiris and Gorton interpret the liquidity shocks as seasonal withdrawals from money-center banks. Calomiris and Gorton present evidence from the National Bank Era (1863–1912) against the seasonal interpretation, even though they never econometrically test their claims. This chapter employs individual level data on New York Clearing House member banks to develop an econometric test of the seasonal hypothesis. The test focuses on the interest-paying banks, the class of banks that based their business on demand deposits from interior banks, during the two panics of 1873 and 1893.

Although the results do not reject a seasonal explanation of the panic of 1873, they do reject the hypothesis that withdrawals on New York by interior banks sparked the panic of 1893. In 1873, interest-paying banks did suffer a reserve drain before the panic. But the existence of a reserve drain does not prove that it was motivated by exogenous agricultural needs. In contrast, in 1893 the data reject the seasonal interpretation. Interest-paying banks were not responsible for the start of the panic in 1893. Only the non-interest-paying banks suffer a significant reserve drain just before the data ends. Diamond and Dybvig's model cannot be applied to the bank panic of 1893 without a non-seasonal interpretation of the exogenous liquidity shock.

An alternate explanation for the panics allows for endogenous reserve drains based on the strategic calculations of depositors. Calomiris and Gorton and other authors focus on the role of asymmetric information about asset prices. Models of strategic depositor withdrawal such as Goldstein and Pauzner (2002) and Rochet and Vives (2001) formalize the effect of private information about bank assets. Another paper tests this hypothesis by investigating the influence of stock market prices on depositors' decisions before the panic of 1893 (Hoag [2002a]).

Appendix 3A

Appendix 3A details the sources and composition of the data. The New York Clearing House archive contains the weekly statements of the Clearing House banks for 1892–3. Also present are the reports of the Comptroller of the Currency for national and state banks of the city for 1872 and 1892. The *New York Tribune* carries the weekly bank statement for 1872–3, and this report was corrected by the same report printed in the *Commercial and Financial Chronicle*. Unfortunately the *Tribune* is rather hard to read. Neither data set is free from error, although the 1893 data is much more accurate.

Clearing House membership fluctuates during the dataset. The New York Clearing House does not include all of the banks in New York. A few small national banks and about half of the state banks were not members. The larger and older banks, whether state or national, were members of the New York Clearing House. Moen and Tallman (2000) suggest that selecting into a clearing house in 1907 did not have large impact on bank observables. In March 1893, one bank opened for business and joined the Clearing House. Two banks left the Clearing House during the period covered by the 1873 data. The first bank failed as a result of defalcation of a teller in May 1873, which was largely unrelated to the panic. The failure of the second bank is more serious, as it failed during the last week of the data set. This failure appears to be due to the panic conditions. The bank suffered a small run the week of the crisis. The bank was owed money by a broker who suspended, so the bank discontinued payment, and the Comptroller of the Currency initiated bankruptcy proceedings. If the bank had waited to suspend, the bank probably would have obtained aid from the Clearing House and stood with the other members. The bank was eventually liquidated with no loss to depositors. The receiver's report printed in the *New York Tribune* on October 7, 1873 provides the last week of data for this bank. While it might be possible to adjust this point to reflect the weekly average rather than the point estimate of the day the bank failed, small changes in the data point do not affect the results. This last bank, with a full time series of data, was the only bank of the three included in the data set. None of these banks were interest-paying banks. After dropping one bank from each panel, we are left with 60 banks in 1872–3 and 63 banks in 1892–3.

The weekly statements only record net deposits, which could affect the power of the test. The variable net deposits subtracts out a variety of liquid assets from gross deposits. If banks sold certain assets, such as claims on other banks or bank notes of other banks, then net deposits would remain unchanged even though the bank made payments to the interior with the proceeds of these assets. However, interest-paying banks did not hold substantial deposits on other banks. Deposits on other banks formed only about 13% of deposits due to other banks for interest-paying banks, compared to 38% for the non-interest-paying banks in both years. Interior banker's deposits on interest-paying banks were largely unidirectional, with interior banks depositing in New York.

Appendix 3B

Appendix 3B describes the results of two variations of the tests. The first variation includes five weeks of time shocks before the panics instead of three. Adding two additional weeks only changed the result of one hypothesis test. The second variation tests the null hypothesis of the first test against θ_0 or λ_0 instead of against zero. For example, one hypothesis of interest becomes H_0: $\theta_t = \theta_0$. Recall that in the regression in Table 3.3, θ_0 measures the time mean of the shocks θ_t of the periods without a specific time dummy variable. From Table 3.3, note that the coefficient λ_0 is very close to zero in both 1872–3 and 1892–3, while for θ_0 there is a slight upward trend in 1872–3 and a slight downward trend in 1892–3, although all four coefficients are statistically insignificant. As a consequence, the results remain the same as when the hypotheses are tested against zero.

Tables 3.3a and 3.4a are the analogue of Tables 3.3 and 3.4 using five weeks instead of three weeks in the estimation of the restricted version of (3.2), using only time indicator variables on the RHS. The only interesting feature of allowing extra weeks is that in 1893, the interest-paying banks suffered a 3% reserve drain in the fifth week before the panic began. One might associate the withdrawal with the large decline in securities prices on the New York Stock Exchange (Sprague, p. 167) that occurred in early May, 1893. Tables 3.5 and 3.6 present the results of hypothesis tests of the form H_0: $\theta_t = 0$ and H_0: $\theta_t = \theta_0$. Only in two cases does the significance level of the individual coefficients change marginally.

Tables 3.7 and 3.8 display the results of the tests for joint significance. The only case where significance changed was when the two classes of banks were compared over five weeks instead of three weeks in 1893. With five weeks, the Wald test judges that the two classes of banks suffered statistically significantly different drains ($p < 0.01$), while with three weeks the test could not reject identical shocks to the two classes of banks ($p = 0.30$). This reversal resolves the apparent contradiction with the evidence from test 1 that the non-interest-paying banks suffered a drain while the interest-paying banks did not. Even though a shock does strike the interest-paying banks during the fifth week before the panic, the interest-paying banks still did not suffer a large drain in 1893 (H_0: $\theta_{T-4} = \ldots = \theta_T$, $p = 0.43$).

Appendix 3C

Appendix 3C describes two details of the application of the Chang (2002) S_N panel unit root test. The two details treated here are the choice of whether or not to include a time trend and the selection of the lag truncation procedure.

The version of the test employed here corrects for a positive mean but does not adjust for a time trend. Because Chang corrects for a nonzero mean in the Monte Carlo size and power study, it seems wise to do the same. Chang uses adaptive demeaning to remove the mean from the series. The series is transformed by subtracting the mean of the previous t data points. If q_{it} is the original series, then the adaptively demeaned series y_{it} is:

$$y_{it} = q_{it} - \sum_{j=1}^{t-1} q_{jt} \tag{3.16}$$

A time trend in the unit root test seems unwarranted in this context. The data do appear to be downward trending over the sample period. However, the decline observed in total net deposits during the sample period is short-term, as the level of deposits sharply increases back to higher levels after the panic (see Figures 3.1 and 3.2). Since the time trend does not persist outside the sample, it seems spurious to include it in the model. Hence, the method of choice is the adaptive demeaned test.

The S_N test is somewhat sensitive to lag order p_i of the autoregressions. The version of the test in the text fixed the number of lags at four for all cross-sectional units i. Autoregressions of order one ($p_i = 1$ for all i) preserved the results with S_N statistics of –11.38 in 1873 and –4.46 in 1893. While it might be possible to adopt more sophisticated lag truncation rules, in this case the test results are robust to two different lag structures so additional modification seems superfluous.

Figure 3.1: Total Net Deposits of New York Clearing House Member Banks, 1872-3

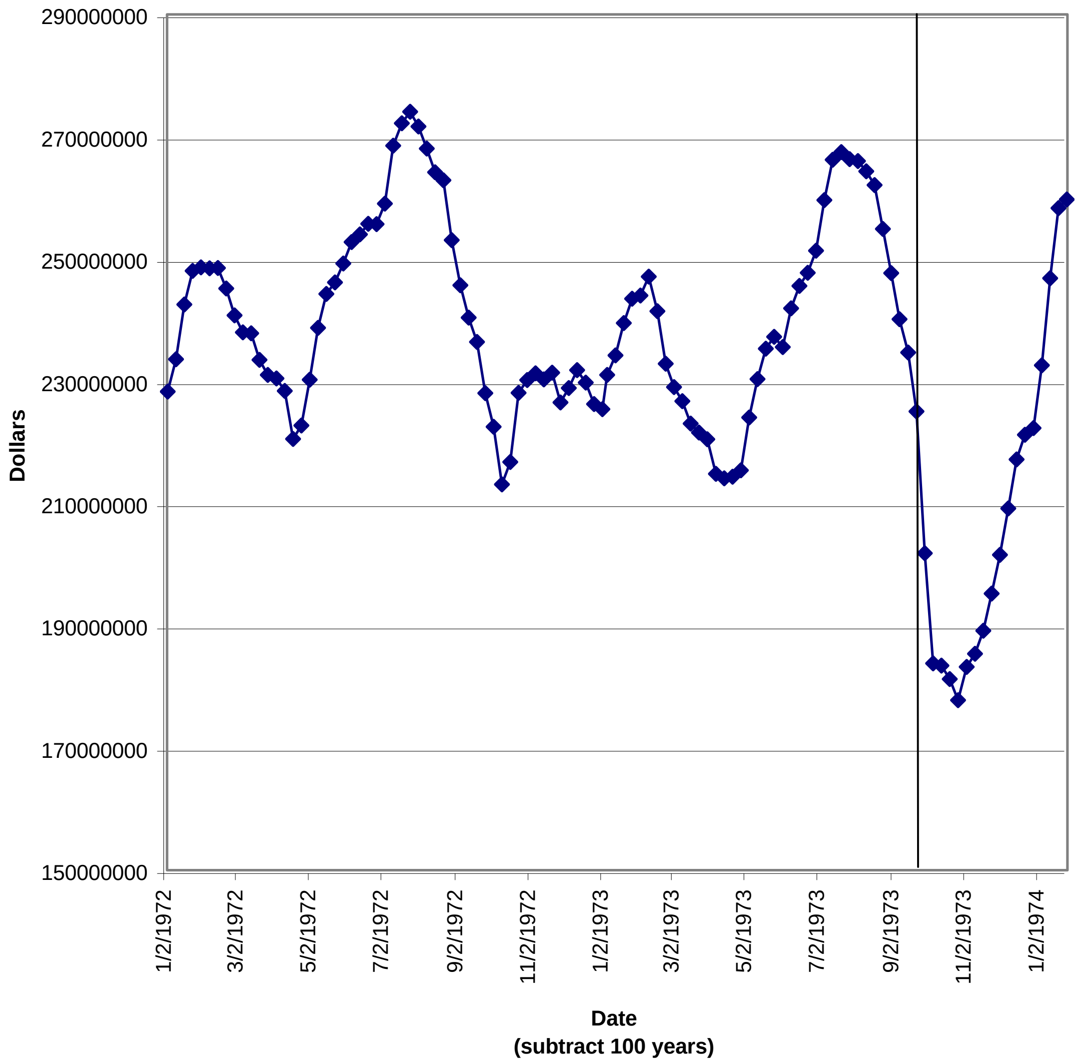


Figure 3.2: Total Net Deposits of New York Clearing House Member Banks, 1892-3

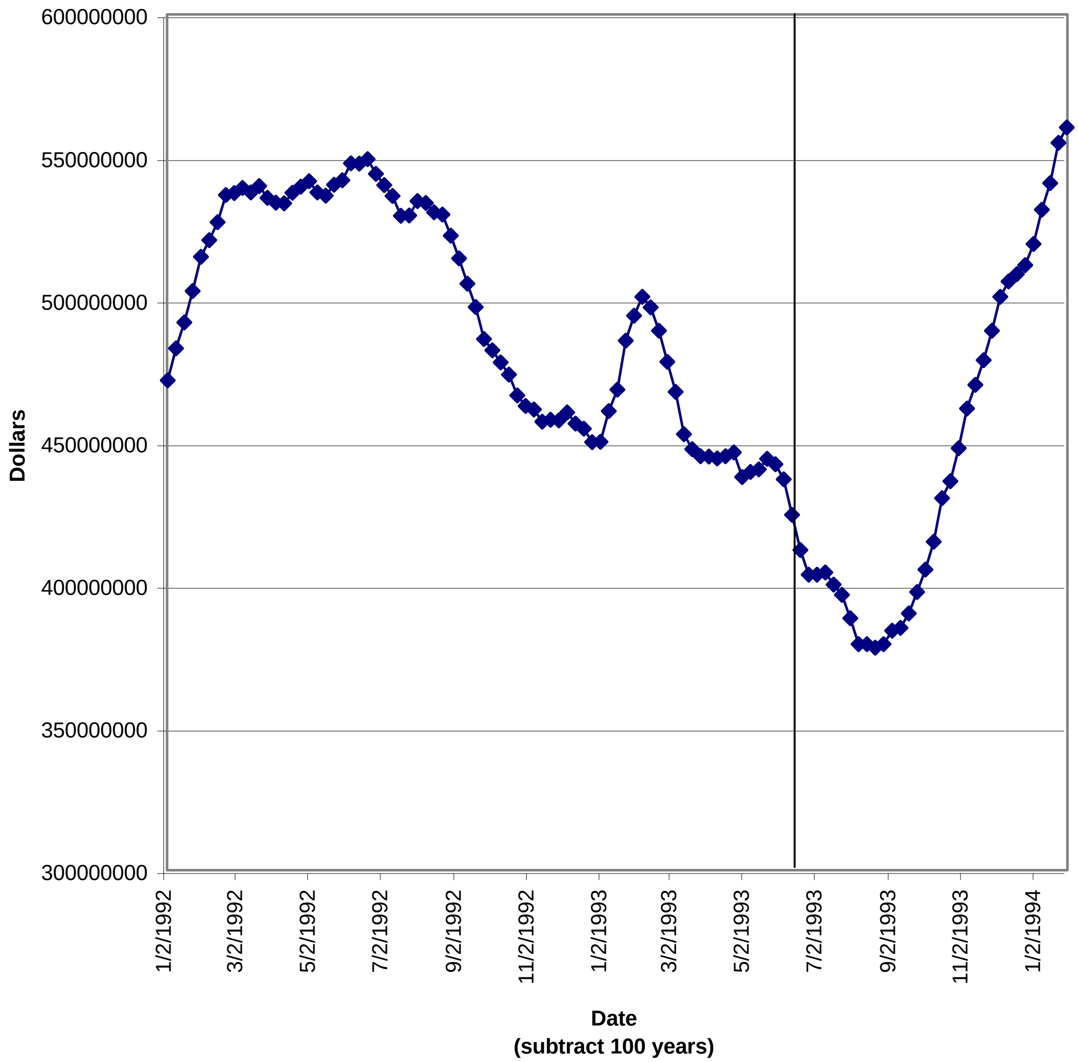


Figure 3.3: Total Net Deposits of Interest-paying Banks, 1872-3

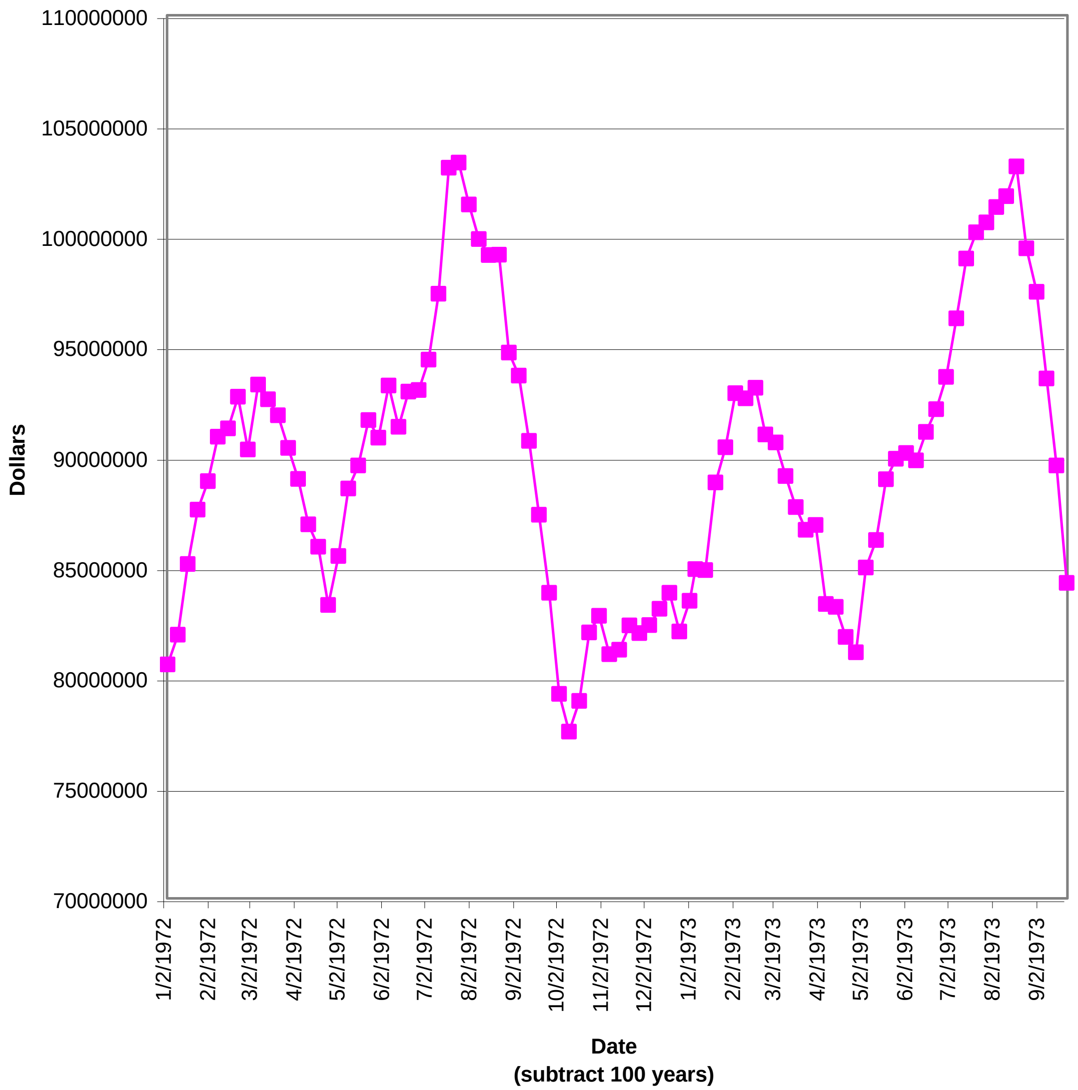


Figure 3.4: Total Net Deposits of Non-interest-paying Banks, 1872-3

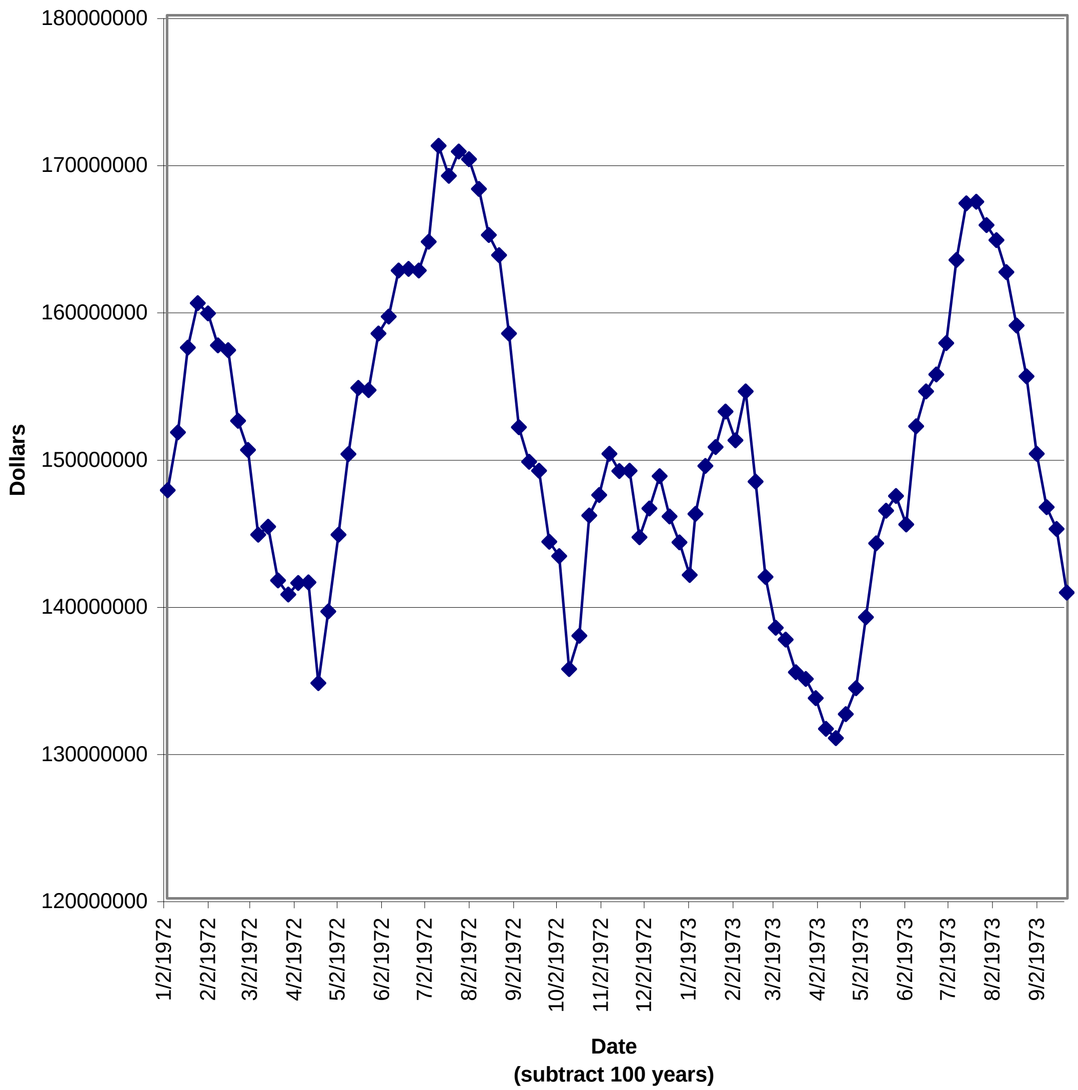


Figure 3.5: Total Net Deposits of Interest-paying Banks, 1892-3

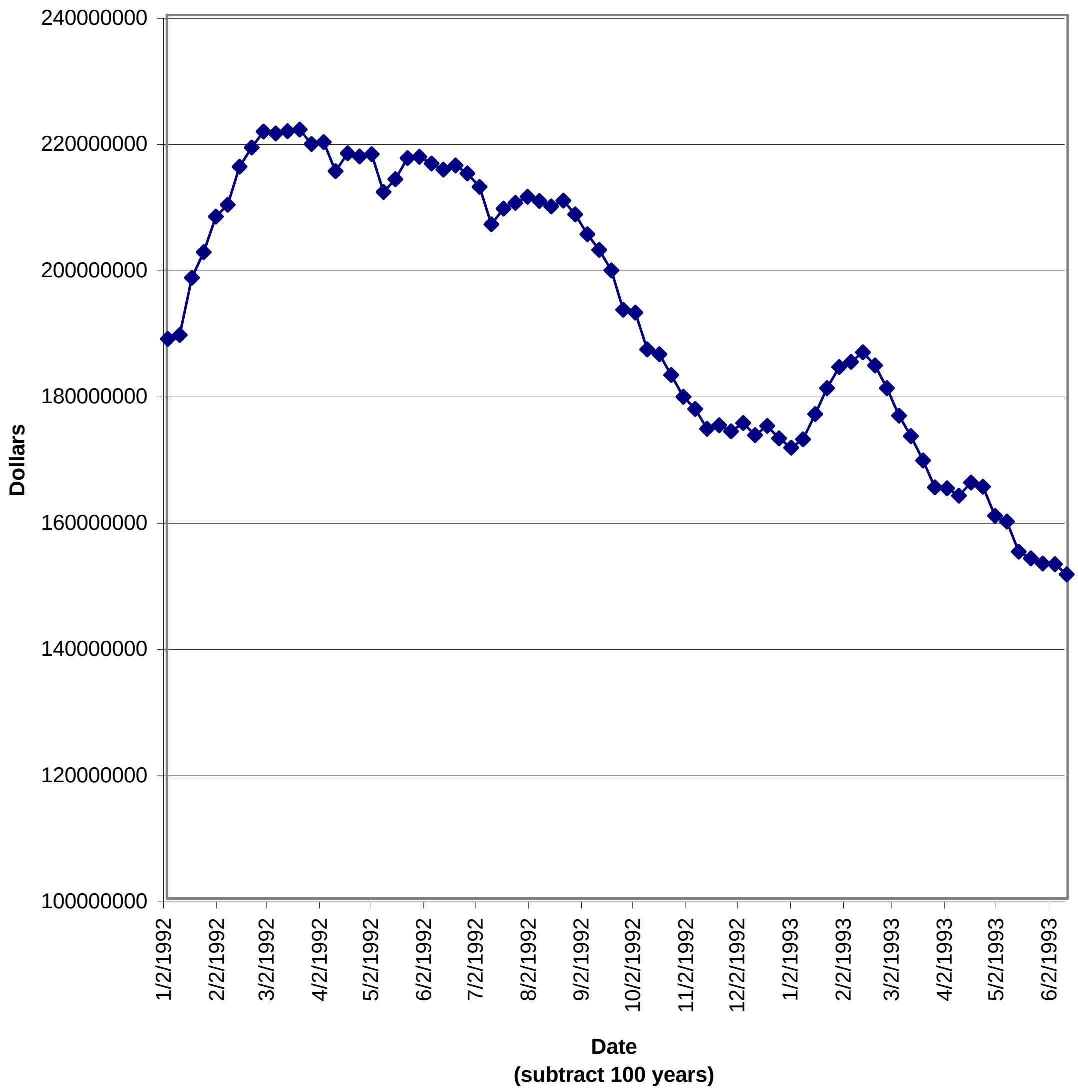


Figure 3.6: Total Net Deposits of Non-interest-paying Banks, 1892-3

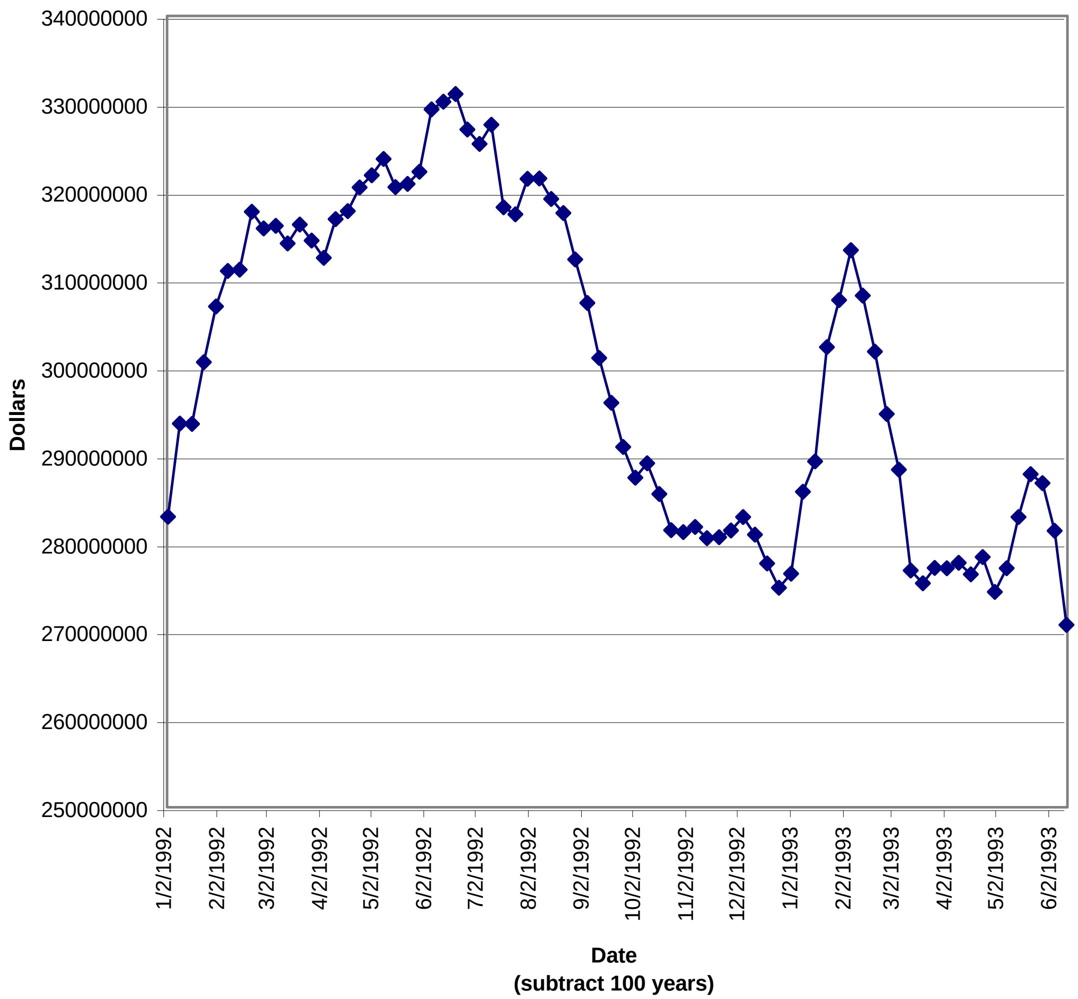


Table 3.1: Comparison of estimation procedures for 1873 data
(approximate t statistics in parentheses, individual fixed effects not reported)

Variable	OLS	FGLS	PCSEs
γ:	-0.067	-0.114	-0.067
	(-4.91)	(-8.28)	(-1.88)
λ_0:	0.219	0.196	0.219
	(0.24)	(0.50)	(0.55)
λ_{T-2}:	-2.443	-1.812	-2.443
	(-1.66)	(-3.14)	(-1.03)
λ_{T-1}:	-0.982	0.019	-0.982
	(-0.67)	(0.03)	(-0.41)
λ_T:	-2.845	-1.705	-2.845
	(-1.93)	(-2.95)	(-1.20)
θ_0:	0.262	0.204	0.262
	(0.29)	(0.88)	(1.12)
θ_{T-2}:	-4.139	-2.565	-4.139
	(-1.41)	(-2.93)	(-1.81)
θ_{T-1}:	-4.462	-2.079	-4.462
	(-1.52)	(-2.37)	(-1.95)
θ_T:	-6.205	-4.509	-6.205
	(-2.12)	(-5.15)	(-2.71)
R^2:	0.010	-	0.010 (from OLS)
Log Likelihood:	-	-13870	-

Number of Observations: 89 weeks * 60 banks = 5340

Dependent variable:

$$y_{it} = \frac{100(TND_{it} - TND_{it-1})}{(1/|C_i|)\sum_{j \in C_i} TND_{jt-1}} \tag{3.3}$$

Model:

$$y_{it} = \mu + a_i + \gamma\, y_{it-1} + \theta_0\,(IP_{it}) + \lambda_0\,(NIP_{it}) + \varepsilon_{it.} \qquad t < T-2 \tag{3.2}$$

$$y_{it} = \mu + a_i + \gamma\, y_{it-1} + \theta_t\,(IP_{it}) + \lambda_t\,(NIP_{it}) + \varepsilon_{it} \qquad t = T-2,\ T-1,\ \text{or}\ T$$

Table 3.2: Comparison of estimation procedures for 1893 data
(approximate t statistics in parentheses, individual fixed effects not reported)

Variable	OLS	FGLS	PCSEs
γ:	0.025	-0.015	0.025
	(1.73)	(-1.11)	(0.62)
λ_0:	0.155	0.150	0.155
	(0.23)	(0.65)	(0.66)
λ_{T-2}:	-0.232	-0.150	-0.232
	(-0.22)	(-0.46)	(-0.16)
λ_{T-1}:	-1.705	-1.435	-1.705
	(-1.59)	(-4.44)	(-1.21)
λ_T:	-3.585	-3.246	-3.585
	(-3.34)	(-10.01)	(-2.55)
θ_0:	-0.682	-0.714	-0.682
	(-0.99)	(-1.22)	(-1.16)
θ_{T-2}:	-0.937	-0.985	-0.937
	(-0.53)	(-1.53)	(-0.58)
θ_{T-1}:	-0.446	-0.091	-0.446
	(-0.25)	(-0.14)	(-0.27)
θ_T:	-1.468	-1.330	-1.468
	(-0.83)	(-2.07)	(-0.90)
R^2:	0.011	-	0.011 (from OLS)
Log Likelihood:	-	-9268	-

Number of Observations: 74 weeks * 63 banks = 4662
Dependent variable:

$$y_{it} = \frac{100(TND_{it} - TND_{it-1})}{(1/|C_i|)\sum_{j\in C_i} TND_{jt-1}} \tag{3.3}$$

Model:

$$y_{it} = \mu + a_i + \gamma\, y_{it-1} + \theta_0\,(IP_{it}) + \lambda_0\,(NIP_{it}) + \varepsilon_{it.} \qquad t < T-2 \tag{3.2}$$

$$y_{it} = \mu + a_i + \gamma\, y_{it-1} + \theta_t\,(IP_{it}) + \lambda_t\,(NIP_{it}) + \varepsilon_{it} \qquad t = T-2, T-1, T$$

Table 3.3: Ordinary Wald test form of regression: restricted PCSE estimation
(approximate t statistics in parentheses)

Variable	1873	1893
λ_0:	0.0465	0.035
	(0.19)	(0.20)
λ_{T-2}:	-2.412	-0.360
	(-1.05)	(-0.24)
λ_{T-1}:	-1.017	-1.886
	(-0.44)	(-1.28)
λ_T:	-2.973	-3.805
	(-1.29)	(-2.58)
θ_0:	0.244	-0.270
	(1.03)	(-1.50)
θ_{T-2}:	-4.022	-0.544
	(-1.81)	(-0.36)
θ_{T-1}:	-4.208	-0.050
	(-1.90)	(-0.03)
θ_T:	-5.938	-1.060
	(-2.68)	(-0.69)
R^2 (from OLS):	0.0039	0.0060
Observations:	5400	4725

Dependent variable:

$$y_{it} = \frac{100(TND_{it} - TND_{it-1})}{(1/|C_i|)\sum_{j\in C_i} TND_{jt-1}} \tag{3.3}$$

Model:

$$y_{it} = \theta_0 (IP_{it}) + \lambda_0 (NIP_{it}) + \varepsilon_{it.} \qquad t < T\text{-}2 \qquad \text{(3.2 restricted)}$$

$$y_{it} = \theta_t (IP_{it}) + \lambda_t (NIP_{it}) + \varepsilon_{it} \qquad t = T\text{-}2, T\text{-}1, T$$

Table 3.3a: Ordinary Wald test form of regression: restricted PCSE estimation with 5 weeks (approximate t statistics in parentheses)

Variable	1873	1893
λ_0:	0.113	-0.019
	(0.46)	(-0.11)
λ_{T-4}:	-2.180	2.109
	(-0.97)	(1.46)
λ_{T-3}:	-3.383	1.721
	(-1.50)	(1.19)
λ_{T-2}:	-2.412	-0.360
	(-1.07)	(-0.25)
λ_{T-1}:	-1.017	-1.886
	(-0.45)	(-1.31)
λ_T:	-2.973	-3.805
	(-1.32)	(-2.64)
θ_0:	0.316	-0.226
	(1.34)	(-1.26)
θ_{T-4}:	-3.597	-3.012
	(-1.66)	(-2.01)
θ_{T-3}:	-1.985	-0.668
	(-0.92)	(-0.45)
θ_{T-2}:	-4.022	-0.544
	(-1.86)	(-0.36)
θ_{T-1}:	-4.208	-0.050
	(-1.94)	(-0.03)
θ_T:	-5.938	-1.060
	(-2.74)	(-0.71)
R^2 (from OLS):	0.0066	0.0089
Observations:	5400	4725

Table 3.4: Direct t test form of regression: restricted PCSE estimation
(approximate t statistics in parentheses)

Variable	1873	1893
λ_0:	0.047	0.035
	(0.19)	(0.20)
λ_{T-2}:	-2.412	-0.360
	(-1.05)	(-0.24)
λ_{T-1}:	-1.017	-1.886
	(-0.44)	(-1.28)
λ_T:	-2.973	-3.805
	(-1.29)	(-2.58)
δ_0:	0.198	-0.305
	(0.76)	(-1.50)
δ_{T-2}:	-1.609	-0.184
	(-0.67)	(-0.11)
δ_{T-1}:	-3.191	1.836
	(-1.32)	(1.06)
δ_T:	-2.965	2.745
	(-1.23)	(1.58)
R^2 (from OLS):	0.0039	0.0060
Observations:	5400	4725

Dependent variable:

$$y_{it} = \frac{100(TND_{it} - TND_{it-1})}{(1/|C_i|)\sum_{j \in C_i} TND_{jt-1}} \qquad (3.3)$$

Model:

$$y_{it} = \lambda_0 (IP_t + NIP_t) + \delta_t (IP_t) + \varepsilon_{it} \quad t < T\text{-}2 \qquad (3.5 \text{ restricted})$$
$$y_{it} = \lambda_t (IP_t + NIP_t) + \delta_t (IP_t) + \varepsilon_{it} \quad t = T\text{-}2, T\text{-}1, T$$

Table 3.4a: Direct t test form of regression: restricted PCSE estimation with 5 weeks (approximate t statistics in parentheses)

Variable	1873	1893
λ_0:	0.113	-0.019
	(0.46)	(-0.11)
λ_{T-4}:	-2.180	2.109
	(-0.97)	(1.46)
λ_{T-3}:	-3.383	1.721
	(-1.50)	(1.19)
λ_{T-2}:	-2.412	-0.360
	(-1.07)	(-0.25)
λ_{T-1}:	-1.017	-1.886
	(-0.45)	(-1.31)
λ_T:	-2.973	-3.805
	(-1.32)	(-2.64)
δ_0:	0.203	-0.207
	(0.78)	(-1.07)
δ_{T-4}:	-1.418	-5.121
	(-0.59)	(-3.16)
δ_{T-3}:	1.398	-2.389
	(0.58)	(-1.47)
δ_{T-2}:	-1.609	-0.184
	(-0.67)	(-0.11)
δ_{T-1}:	-3.191	1.836
	(-1.33)	(1.13)
δ_T:	-2.965	2.745
	(-1.23)	(1.69)
R^2 (from OLS):	0.0066	0.0089
Observations:	5400	4725

Table 3.5: Modified first test in 1873

λ_0 = 0.113 with s.e. 0.245

θ_0 = 0.316 with s.e. 0.235

H_0:	($\lambda_t = 0$)	($\lambda_t = \lambda_0$)
λ_{T-4}	-2.179	
	(0.334)	(0.312)
λ_{T-3}	-3.383	
	(0.134)	(0.123)
λ_{T-2}	-2.413	
	(0.285)	(0.266)
λ_{T-1}	-1.017	
	(0.652)	(0.617)
λ_T	-2.973	
	(0.187)	(0.174)

H0:	($\theta_t = 0$)	($\theta_t = \theta_0$)
θ_{T-4}	-3.597	
	(0.097)	(0.072)
θ_{T-3}	-1.985	
	(0.359)	(0.291)
θ_{T-2}	-4.022	
	(0.063)	(0.047)
θ_{T-1}	-4.208	
	(0.052)	(0.038)
θ_T	-5.938	
	(0.006)	(0.004)

Table 3.6: Modified first test in 1893

λ_0 = -0.019 with s.e. 0.172

θ_0 = -0.226 with s.e. 0.179

H_0:	($\lambda_t = 0$)	($\lambda_t = \lambda_0$)
λ_{T-4}	2.109	
	(0.143)	(0.143)
λ_{T-3}	1.721	
	(0.232)	(0.230)
λ_{T-2}	-0.360	
	(0.803)	(0.814)
λ_{T-1}	-1.886	
	(0.190)	(0.198)
λ_T	-3.805	
	(0.008)	(0.009)

H0:	($\theta_t = 0$)	($\theta_t = \theta_0$)
θ_{T-4}	-3.012	
	(0.044)	(0.065)
θ_{T-3}	-0.668	
	(0.656)	(0.769)
θ_{T-2}	-0.544	
	(0.716)	(0.833)
θ_{T-1}	-0.500	
	(0.973)	(0.907)
θ_T	-1.060	
	(0.479)	(0.580)

Table 3.7: Joint tests and modified joint tests in 1873

Joint Tests:

H_0: (θ_{T-k} = ... = θ_{T-1} = θ_T = 0) = no drain on IP banks

IP banks	3 weeks	5 weeks
F test (p value)	< 0.01	< 0.01

H_0: (λ_{T-k} = ...= λ_{T-1} = λ_T = 0) = no drain on NIP banks

NIP banks	3 weeks	5 weeks
F test (p value)	0.40	0.28

H_0: (δ_{T-k} = ... = δ_{T-1} = δ_T = 0) = no difference between two classes

Difference	3 weeks	5 weeks
F test (p value)	0.30	0.49

Modified joint tests:

H_0: (θ_{T-k} = ... = θ_{T-1} = θ_T = θ_0) = no drain different from mean on IP banks

IP banks	3 weeks	5 weeks
F test (p value)	< 0.01	< 0.01

H_0: (λ_{T-k} = ...= λ_{T-1} = λ_T = λ_0) = no drain different from mean on NIP banks

NIP banks	3 weeks	5 weeks
F test (p value)	0.39	0.26

Table 3.8: Joint tests and modified joint tests in 1893

Joint Tests:

H_0: ($\theta_{T-k} = ... = \theta_{T-1} = \theta_T = 0$) = no drain on IP banks

IP banks	3 weeks	5 weeks
F test (p value)	0.90	0.43

H_0: ($\lambda_{T-k} = ...= \lambda_{T-1} = \lambda_T = 0$) = no drain on NIP banks

NIP banks	3 weeks	5 weeks
F test (p value)	0.04	0.03

H_0: ($\delta_{T-k} = ... = \delta_{T-1} = \delta_T = 0$) = no difference between two classes

Difference	3 weeks	5 weeks
F test (p value)	0.30	< 0.01

Modified joint tests in 1893:

H_0: ($\theta_{T-k} = ... = \theta_{T-1} = \theta_T = \theta_0$) = no drain different from mean on IP banks

IP banks	3 weeks	5 weeks
F test (p value)	0.96	0.58

H_0: ($\lambda_{T-k} = ...= \lambda_{T-1} = \lambda_T = \lambda_0$) = no drain different from mean on NIP banks

NIP banks	3 weeks	5 weeks
F test (p value)	0.04	0.03

References

Adams, Jr., Charles F. and Adams, Henry. 1956. *Chapters of Erie*. Ithaca: Cornell Univ. Press.

Allen, Franklin and Gale, Douglas. 2001. Banking and Markets. Wharton School. Mimeo.

Allen, Linda and Saunders, Anthony M. 1992. Bank window dressing: theory and practice. *J. Bank. and Fin.* 16 (3): 585–623.

Banker's Magazine. 1866–72. Various issues.

Beck, Nathaniel and Katz, Jonathan. 1995. What to do (and not to do) with time-series cross-section data. *Am. Pol. Sci. Rev.* 89 (3): 634–47.

Beck, Nathaniel and Katz, Jonathan. 1996. Nuisance vs. substance: specifying and estimating time-series cross-section models. In *Political Analysis 6*, Freeman, J. R., ed., 1–36. Ann Arbor: Univ. Michigan Press.

Bernard, Victor L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *J. Accounting Res.* 25 (1): 1–48.

Beyer, W., ed. 1991. *CRC Standard Probability and Statistics Tables and Formulae*. CRC Press: Boca Raton, FL.

Box, G. E. P. and Jenkins, G. M. 1976. *Time Series Analysis*. San Francisco: Holden Day.

Brown, Stephen J., and Warner, Jerold B. 1985. Using daily stock returns: the case of event studies. *J. Fin. Econ.* 14 (1): 3–31.

Bryant, John. 1980. A model of reserves, bank runs, and deposit insurance. *J. Bank. and Fin.* 4 (4): 335–344.

Calomiris, Charles. and Gorton, Gary. 1991. Origins of banking panics: models, facts, and bank regulation. In *Financial Markets and Financial Crises*, Hubbard, R. G., ed., 109–73. Chicago: NBER and Univ. Chicago press.

Calomiris, Charles W. and Kahn, Charles M. 1991. The role of demandable debt in structuring optimal banking arrangements. *Am. Ec. Rev.* 81 (3): 497–513.

Cannon, James G. 1908. *Clearing-Houses*. New York: D. Appleton and Co.

Chabot, Benjamin R. 2000. A Single Market? The Stock Exchanges of the United States and London 1866–1885. PhD. Diss., Northwestern University.

Champ, Bruce, Smith, Bruce D., and Williamson, Stephen D. 1996. Currency elasticity and banking panics: theory and evidence. *Canadian Journal of Economics* 29 (4): 828–864.

Chang, Yoosoon. 2002. Nonlinear IV unit root tests in panels with cross-sectional dependency. *J. Econometrics* 110 (2): 261–292.

Chari, V. V. 1989. Banking without deposit insurance or banks panics: lessons from a model of the U.S. national banking system. *Fed. Res. Bank of Minn. Quar. Rev.* 13 (3): 3–19.

Commercial and Financial Chronicle. 1872–3. Various issues.

Congressional Record. 1869. Various issues.

Cornwallis, K. 1879. *The Gold Room and the New York Stock Exchange and Clearing House*. New York: A. S. Barnes & Co.

Davis, Lance, and Hughes, J.R.T. 1960. A Dollar-Sterling Exchange, 1803–1895. *Econ. Hist. Rev*, 13(1): 52–78.

Deketele, Hans. 2000. Masters of the Blue Riband. Online. Internet. May, 2000. <http://www.blueriband.com>.

Diamond Douglas W. and Dybvig, Phillip H. 1983. Bank runs, deposit insurance, and liquidity. *J. Pol. Ec*. 91 (3): 401–20.

Demirgüç-Kunt, Asli. and Detragiache, Enrica. 1998. The determinants of banking crises in developing and developed countries. *IMF Staff Papers* 45 (1): 81–109.

Eun, Cheol S. and Shim, Sangdal. 1989. International transmission of stock market movements. *J. Fin. and Quant. Anal.* 24 (2): 241–256.

Friedman, Milton. and Schwartz, Anna J. 1963. *A Monetary History of the United States*. Princeton, NJ: Princeton Univ Press.

Friedman, Milton and Schwartz, Anna J. 1970. *Monetary Statistics of the United States*. New York: NBER.

Garbade, Kenneth D. and Silber, William L. 1978. Technology, communication and the performance of financial markets: 1840–1975. *J. Fin.*, 33 (3): 819–832.

Garbade, Kenneth D. and Silber, William L. 1979. Dominant and satellite markets: a study of dually-traded securities. *Rev. Ec. Stat.* 61 (3): 455–60.

Glick, Reuven. and Hutchison, Michael. 1999. Banking and currency crises: how common are twins? Federal Reserve Bank of San Francisco. Mimeo.

Goldstein, Itay. and Pauzner, Ady. 2002. Demand deposit contracts and the probability of bank runs. Duke Univ. Mimeo.

Goodhart, C. A. E. 1969. *The New York Money Market and the Finance of Trade, 1900-1913*. Cambridge, MA: Harvard Univ. Press.

Gorton, Gary. 1988. Banking panics and business cycles. *Ox. Ec. Papers* 40 (4): 751–781.

Gorton, Gary and Mullineaux, Donald J. 1987. The joint production of confidence: endogenous regulation and nineteenth century commmerical-bank clearinghouses. *J. Money, Credit, and Banking* 19 (4): 457–68.

Greene, William E. 1993. *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall.

Greene, William. E. 2003. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Hall, Alastair R. 1994. Testing for a unit root in time series with pretest data-based model selection. *J. Bus. Ec. Stat*. 12 (4):461–470.

Hamilton, James. D. 1994. *Time Series Analysis*. Princeton: Princeton Univ. Press.

Harvey, Andrew, Koopman, Siem Jan, and Riani, Marco. 1997. The modeling and seasonal adjustment of weekly observations. *J. Bus. and Ec. Stat*. 15 (3): 354–68.

Hauser, Shmuel., Tanchuma, Yael., and Uzi, Yaari. 1998. International transfer of pricing information between dually listed stocks. *J. Fin. Res.* 21 (2): 139–157.

Hoag, C. 2002a. The influence of asset markets on bank runs. Work in progress.

Hoag, C. 2002b. Interbank lending during financial crises. Work in progress.

Hsaio, Cheng. 1986. *Analysis of Panel Data*. Cambridge: Cambridge Univ. press.

Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B.S. 1992. Seasonal integration and cointegration. In *Modelling Seasonality*, ed. S. Hylleberg, Oxford: Oxford Univ. Press, p. 425–448.

Im, Kyung S., Pesaran M. Hashem, and Shin, Yongcheol. 1997. Testing for unit roots in heterogeneous panels. Dept. Appl. Ec., Univ. Cambridge. Mimeo.

James, John. 1978. *Money and Capital Markets in Postbellum America*. Princeton, NJ: Princeton Univ. Press.

Jevons, W. S. 1884. *Investigations in Currency and Finance*. Reprint. London: Macmillan and Co., 1909.

Judson, Ruth A. and Owen, Ann L. 1999. Estimating dynamic panel data models: a guide for macroeconomists. *Ec. Lett.* 65 (1): 9–15.

Kaminsky, Graciela L. and Reinhart, Carmen M. 1999. The twin crises: the causes of banking and balance-of-payment problems. *Am. Ec. Rev.* 89 (3): 473–500.

Kemmerer, E.W. 1910. *Seasonal Variations in the Relative Demand for Money and Capital in the United States*. Washington: National Monetary Commission, GPO.

Kiviet, Jan F. 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *J. Econometrics* 68 (1): 53–78.

Lau, Sie Ting and Diltz, J. David. 1994. Stock returns and the transfer of information between the New York and Tokyo stock exchanges. *J. Intl. Mon. and Fin* 13 (2): 211–222.

Leybourne, S. J. 1995. Testing for unit roots using forward and reverse Dickey-Fuller regressions. *Ox. Bull. Ec. Stat*. 57 (4): 559–571.

Lieberman, Offer, Ben-Zion, Uri., and Hauser, Shmuel. 1999. A characterization of the price behavior of international dual stocks: an error correction approach. *J. Intl. Mon. and Fin*. 18(2): 289–304.

Lin, Wen-Ling, Engle, Robert F., and Ito, Takatoshi. 1994. Do bulls and bears move across borders? International transmission of stock returns and volatility. *Rev. Fin. Stud*. 7 (3): 507–538.

Maddala, G.S. and Kim, In-Moo. 1998. *Unit Roots, Cointegration, and Structural Change*. Cambridge: Cambridge Univ. Press.

Maddala, G. S. and Wu, Shaowen. 1999. A comparative study of unit root tests with panel data and a new simple test. *Ox. Bull. Ec. St.* 61 (0): 631–52.

Macaulay, F. R. 1938. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the United States since 1856.* New York: NBER.

Madiès, Phillipe. 2001. Self-fulfilling bank panics: how to avoid them? An experimental study. Université Pierre Mendès. Mimeo.

Michie, Ranald C. 1987. *The London and New York Stock Exchanges 1850–1914*. London: Allen and Unwin.

Miller, Victoria. 1996. Exchange rate crises with domestic bank runs: evidence from the 1890s. *J. Intl. Money and Fin.* 15 (4): 637–656.

Miron, Jeffrey A. 1986. Financial panics, the seasonality of the nominal interest rate, and the founding of the Fed. *Am. Ec. Rev.* 76 (1): 125–140.

Mitchell, W. C. 1903. *A History of the Greenbacks*. Chicago: University of Chicago Press.

Mitchell, W.C. [1908] 1966. *Gold, Prices, and Wages under the Greenback Standard*. New York: Augustus M. Kelley Publishers. Reprint.

Moen, Jon R. and Tallman, Ellis W. 2000. Clearinghouse membership and deposit contraction during the panic of 1907. *J. Ec. Hist*. 60 (1): 145–163.

Myers, Margaret G. 1931. *The New York Money Market*. New York: AMS Press.

Neal, Larry. 1990. *The Rise of Financial Capitalism.* Cambridge: Cambridge University Press.

Neal, Larry. 1992a. The disintegration and re-integration of international capital markets in the eighteenth century. In *Business and Economic History*, ed. William J. Hausman, 21: 84–96.

Neal, Larry. 1992b. Technological advance and the progress of capital market integration in the 19th century. Mimeo, University of Illinois.

Neumark, David, Tinsley, P. A., and Tosini, Suzanne. 1991. After-hours stock prices and post-crash hangovers. *J. Fin*. 46(1): 159–178.

New York Clearing House Association archives. New York.

New York Times. 1867–9. Various issues.

New York Tribune. 1872–3. Various issues.

Nickell, Stephen J. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49 (6): 1417–26.

Noyes, Alexander D. 1894. The banks and the panic of 1893. *Pol. Sci. Quar.* 9 (1): 12–30.

Office of the Comptroller of the Currency. U. S. Treasury. 1963. *100th Annual Report of the Comptroller of the Currency, 1962*. Washington, D.C.: U.S. GPO.

Office of the Comptroller of the Currency. *Report of the Comptroller of the Currency*. 1867–71. Washington, D.C.: U.S. GPO.

Officer, Lawrence. H. 1985. Integration in the American Foreign Exchange Markets 1791–1900. *J. Ec. Hist*, 45 (3): 557–585.

Perkins, Edwin J. 1978. Foreign interest rates in American financial markets: a revised series of dollar-sterling exchange rates 1835–1900. *J. Ec. Hist.* 38 (2): 392–409.

Pesaran, M. Hashem. and Smith, Ron. 1995. Estimating long-run relationships from dynamic heterogeneous panels. *J. Econometrics* 68 (1): 79–113.

Philadelphia Inquirer. 1866–71. Various issues.

Philadelphia Public Ledger. 1866–72. Various issues.

Ramanathan, Ramu. 1995. *Introductory Econometrics with Applications*. Fort Worth: Dryden Press.

Rice, John. A. 1995. *Mathematical Statistics and Data Analysis*. Belmont: Wadsworth Publishing Co.

Robertson, Ross. M. 1968. *The Comptroller and Bank Supervision*. Washington, D.C.: Office of the Comptroller of the Currency.

Rochet, Jean-Charles and Vives, Xavier. 2001. Coordination failure and the lender of last resort: was Bagehot right after all? Mimeo.

Schmidt, H. 1875. *Foreign Banking Arbitration: Its Theory and Practice*. London: Effingham Wilson, Royal Exchange.

Smirlock, Michael and Kaufold, Howard. 1987. Bank foreign lending, mandatory disclosure rules and the reaction of bank stock prices to the Mexican debt crisis. *J. Business* 60 (3): 347–364.

Smith, Bruce D. 1991. Bank panics, suspensions, and geography: some notes on the "contagion of fear" in banking. *Ec. Inquiry* 29 (2): 230–248.

Sprague, O. M. W. 1910. *History of Crises under the National Banking System.* Reprint. Fairfield, NJ: Augustus M. Kelley publishers, 1977.

Stock, James H. and Watson, Mark W. 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61 (4): 783-820.

Temzelides, Ted. 1997. Evolution, coordination, and banking panics. *J. Mon. Ec*. 40 (1): 163–183.

White, Eugene. N. 1992. *The Comptroller and the Transformation of American Banking 1960–1990*. Washington, D.C.: Comptroller of the Currency.

Wicker, Elmus. 2000. *Banking Panics of the Guilded Age*. Cambridge: Cambridge Univ. Press.