

Construction of Aggregate and Regional Bank Data using the Call Reports

Data Manual

Wouter J. den Haan, Steven W. Sumner, and Guy Yamashiro

December 5, 2002

A.1. Overview

The purpose of this manual is twofold. First, we discuss the construction of the bank data set. This data set is made available at <http://weber.ucsd.edu/~wdenhaan> and contains total loans, commercial and industrial (C&I) loans, real estate loans, consumer loans, deposits, bank equity, and total assets for the U.S. economy, the nine census regions, and the 50 U.S. states and the District of Columbia. The time series are constructed with bank data from the Consolidated Reports of Condition and Income (the Call Reports) available from the Federal Reserve Bank of Chicago.¹ The Call Reports are quarterly income statement and balance sheet data that all federally insured² banks are required to submit to the Federal Deposit Insurance Corporation (FDIC). At the aggregate level, we construct both a level data set that runs from the first quarter of 1976 to the last quarter of 2000 and an index series, that corrects for mergers, that runs from the first quarter of 1977 to the last quarter of 2000. At the regional level we only construct the index series that correct for mergers. When banks from different regions merge, the reported loan volume for the regional level series change since banks' headquarters report data. Even if the merger has caused a substantial change in financial conditions, its importance is clearly overstated by the reported reassignment of loan volume from one region to another. We prefer, therefore, to use the regional index series because they exclude merging banks in the periods around a merger.³ The data set also includes data for small banks for which merger activity has been much less important.

The second purpose of this appendix is to compare the (total) loan series constructed with the Call Reports to the H.8 data that are issued by the Board of Governors of the Federal Reserve System.⁴ The H.8 is the data set researchers have traditionally used to examine the behavior of aggregate bank balance sheet variables in response to monetary shocks.⁵ The main difference between the H.8 and the Call

¹ The data were downloaded from

<http://www.chicagofed.org/economicresearchanddata/data/bhcdatabase/bhcdatabase.cfm>.

² The data cover all banks regulated by the Federal Reserve System, the Federal Deposit and Insurance Corporation and the Comptroller of the Currency.

³ When the data are corrected for mergers then the regional series do not exactly add up to the corresponding aggregate.

⁴ The data were downloaded from <http://www.federalreserve.gov>.

⁵ King (1986), Romer and Romer (1990), Bernanke and Blinder (1992), Ramey (1993) and Bacchetta and Ballabriga (2000) rely on predecessors of the H.8 published by the Board of Governors of the Federal Reserve System in, *Banking and Monetary Statistics, 1941-1970* and, *Annual Statistical Digests*. Although this time-series is available back to 1941, the current form of the H.8, used by Lown,

Reports is that the Call Reports are mandatory for all federally insured banks while the H.8 is based on a voluntary bank survey. The main advantage of the Call Reports, therefore, is that it provides a nearly comprehensive survey of banks operating in the United States. Due to the complexity of the panel, however, there are difficulties in working with the data. The major problem is that the data are designed primarily for regulatory purposes, and as a result small banks are not required to report in as much detail as larger banks.⁶ More importantly, as the banking sector evolves and regulations change, the variables reported and how they are measured also change—posing a major challenge in the effort to construct consistent time-series.

This appendix is organized as follows. In Section A.2 we discuss the construction of consistent end-of-period level series. Since the literature typically uses the averages from the H.8 data set we also describe how to construct quarterly averages⁷ from the Call Reports. In Section A.3 we describe the construction of the index series that corrects for mergers and in Section A.4 we describe how to construct end-of-period series for small banks. In Section A.5 we discuss data revisions that at times occur in the Call Reports. In Section A.6, we discuss the H.8 loan series and in Section A.7 we compare the H.8 series to the Call Reports series.

A.2. Forming Consistent Level Series

The sheer amount of data available in the Call Reports coupled with the fact that the data are designed primarily for regulatory purposes make working with the data difficult. The task of defining variables correctly through time is made more difficult by the major overhaul to the Call Report format in the first quarter of 1984. Since 1984 banks were in general required to provide more detail concerning assets and liabilities, resulting in discontinuities in many series. There are, however, some exceptions. For example, prior to 1984 holdings of various types of federal securities and state and local securities were reported separately. Following 1984, these securities were combined into a single series along with holdings of all other bonds, stocks and securities. Consequently, tracking disaggregated holdings of these securities is not possible.

Thus, the complexity of the panel necessitates careful work to yield time-series suitable for empirical analysis. Ensuring consistency of time-series depends on two factors. First, the variables of interest must be correctly defined over time. The *splices* used to do this are discussed in Section A.2.1. Second, one must identify the types of banks one wants to consider and problematic observations, such as observations with negative loans, must be eliminated. The *screens* that define the

Morgan and Rohatgi (2000), is only available (online) back to the beginning of 1973. Bernanke and Blinder (1992) find similar results using either form of data.

In the past four different forms, varying in level of detail, were filed. Banks with foreign offices, regardless of size, must file Report of Condition and Income FFIEC 031. Banks without foreign offices and assets of 300 million dollars or more file Report of Condition and Income FFIEC 032. Banks without foreign offices and assets less than 300 million dollars, but more than 100 million dollars file Report of Condition and Income FFIEC 033. Finally, banks without foreign offices and assets of less than 100 million dollars file Report of Condition and Income FFIEC 034.

⁷ We will argue, however, that the quarterly averages from the Call Reports are less reliable than the corresponding end-of-quarter series.

universe of banks we consider are discussed in Section A.2.2. To correctly define variables through time one should refer to the available dictionary files that provide brief descriptions—including reporting frequency and any changes in definition, which may have occurred—of every variable available on the Call Reports. However, for many of the key balance sheet variables Kashyap and Stein have done much of the necessary work and their findings are available in the appendices of Kashyap and Stein (1995, 2000) and on the website of the Federal Reserve Bank of Chicago.

A.2.1. Definition of Variables (Splices)

In defining the bank loan variables used in this paper we closely follow Kashyap and Stein (2000). In Table A.1 we document how each variable is constructed using the code names from the call reports. The second column gives the name under which the series can be found in our data set. We will now give a more precise description of the balance sheet items and discuss the choices we made to construct consistent time series.

Loans

Although we follow the methods of Kashyap and Stein (2000) quite closely, the reported variable definitions differ from theirs. The main difference is that Kashyap and Stein focus on loans reported on a consolidated basis, that is loans issued by both domestic and foreign branches (rcfd data), while we focus on loans issued by domestic branches only (rcon data). If one wants to examine the effect of monetary policy shocks on regional and aggregate real activity, then it makes more sense to focus on loans issued by domestic branches (rcon data).⁸ Moreover, the rcfd loan series display a discontinuity. We will elaborate on this issue in Section A.2.3.

From the dictionary files we get the following descriptions.

- Total loans (rcon1400): The aggregate gross book value of total loans (before deduction of valuation reserves) including: (a) acceptances of other banks and commercial paper purchased in open market; (b) acceptances executed by or for account of reporting bank and subsequently acquired by it through purchase or discount; (c) customers' liability to reporting bank on drafts paid under letter of credit for which bank has not been reimbursed; and (d) "cotton overdrafts" or "advances", and commodity or bill of lading drafts payable upon arrival of goods against which drawn for which reporting bank has given deposit credit to customers. Also includes: (a) paper rediscounted with Federal Reserve or other banks; and (b) paper pledged as collateral to secure bills payable, as marginal collateral to secure bills rediscounted, or for any other purpose. Beginning in the first quarter of 1984 this item includes rcon 2165, lease financing receivables. The inclusion of this series results in an inconsistency

⁸ A foreign loan could still be included in the data if it is a loan extended by a domestic bank office to a foreign agent. Moreover, a domestic bank office can extend a loan to a multinational company who uses the funds for its foreign operations. See Houston, James, and Marcus (1997) for more on this issue.

in the total loan series. We correct by adding lease financing receivables to total loans prior to 1984.

- C&I loans (rcon1600): Includes loans for commercial and industrial purposes to business enterprises (proprietorships, partnerships, and corporations), whether secured or unsecured, single payment or installment. These loans may take the form of direct or purchased loans. Includes the reporting bank's own acceptances that it holds in its portfolio when the account party is a commercial or industrial enterprise. Also includes loans to individuals for commercial, industrial, and professional purposes but not for investment or personal expenditure purposes. Excludes all commercial and industrial loans held in trading accounts.
- Real Estate Loans (rcon1410): Includes all loans, whatever the purpose, secured primarily by real estate as evidenced by mortgages, deeds of trust, land contracts, or other instruments, whether first or junior liens (e.g., equity loans, second mortgages) on real estate. Includes all such loans whether originated by the bank or purchased from another party.
- Consumer Loans (rcon1975): Includes all loans, not secured primarily by real estate, to individuals for medical expenses, personal taxes, vacations, consolidation of personal (nonbusiness) debts; for the purchase of private passenger automobiles, household appliances, furniture, mobile homes, trailers, boats, etc.; and loans not secured primarily by real estate for the purpose of purchasing real estate that will be used as a residence of the borrower's family.

Average Loans

Relative to quarterly averages, the end-of-quarter series constructed above may either contain data variations that are important for the empirical analysis conducted or may contain uninformative noise that will reduce the efficiency of the econometric procedure. In general, it is difficult to determine whether working with the end-of-quarter series or with the quarterly averages is better.⁹ But, since empirical studies on the monetary transmission mechanism often use quarterly averages we also tried to construct a quarterly average using the Call Reports.

It is important to note that not all balance sheet variables on the Call Reports are available as quarterly averages.¹⁰ For the items that are available as quarterly averages, banks have the option of reporting either:

⁹ It is important to realize that while averaged data may be less noisy than end-of-period data, averaging does introduce an MA component into the series, which implies that the typical model used to study the monetary transmission mechanism, the VAR with a finite number of lags, is misspecified.

¹⁰ The averaged items are reported on Schedule RC-K. This particular schedule is used to calculate yields and costs of funds on selected assets and liabilities. Since this schedule consists of quarterly averages it would be unusual for an amount reported on this schedule to equal its outstanding balance (end-of-quarter data). In this schedule, banks are permitted to report loan detail in terms of general loan categories that are based upon each bank's own internal loan categorization system.

1. an average of daily figures for the quarter,
2. an average of weekly figures (e.g., the Wednesday of each week), or
3. an average of the three month-end figures of the current quarter and the last month-end figure of the previous quarter (for banks reporting on FFIEC 034).

The fact that banks have flexibility in reporting averages is clearly an undesirable feature of the series. Furthermore, there are two less important issues that we must discuss. First, there is a difficulty in terms of directly comparing the average loan series to the end-of-quarter loan series. The available quarterly average loan series rcon3360 does not quite match up with the end-of-quarter loan series rcon1400. End-of-quarter total loans and leases gross, rcon1400, is the sum of rcon2122 and unearned income on loans, rcon2123. While the average loan series, rcon3360, does match up with rcon2122 there is no average measure for rcon2123. However, the unearned income on loans is negligible and so does not affect comparability with the end-of-quarter, rcon1400 very much.

Second, the quarterly average total loan series rcon3360 excludes lease-financing receivables, net of unearned income prior to 1984, but includes it afterwards. The average series for this variable, rcon3484, is available only from 1987 on. To construct a consistent time-series, however, we would need rcon3484 either for the entire period prior to 1984 or for the entire period after 1984. The magnitude of the series, however, is small enough such that an average loan series including it and a series excluding it for the period it is available are visually indistinguishable when both series are plotted in the same graph.

Figure A.1 plots the percentage changes of the rcon end-of-quarter and the rcon quarterly average series. The quarterly average series is a smoother version of the end-of-quarter series. This property is documented by the relative standard deviations. The standard deviation for the percentage change of the quarterly average series is equal to 0.0131 while the standard deviation for the percentage change of the end-of-quarter series is equal to 0.0159. The correlation coefficient for the two series is equal to 0.83.

Finally, the observation for rcon3360 in the fourth quarter of 1983 is over 50% lower than the values in the two surrounding quarters. This is obviously a mistake. In the empirical work we have replaced this observation by the average of the two surrounding quarters.

Bank Assets

The bank asset series in the Call reports are rcfd2170 and rcon2170. The total assets series is defined in the dictionary files as simply the sum of all asset items and must equal the sum of (i) total liabilities and (ii) equity capital.¹¹

¹¹ The sum of these two items is reported in rcfd3300 (total liabilities, limited-life preferred stock and equity capital).

Total Deposits

The bank deposit series in the Call Reports are rcf2200 and rcon2200 and they are defined as the sum of (i) total transaction accounts, (ii) total nontransaction accounts¹² and (iii) the sum of noninterest-bearing deposits and interest-bearing deposits.

Bank Equity

The code for the bank capital series in the Call Reports is rcf3210. The definition of bank capital given in the dictionary file that accompanies the call report is the sum of the following components: (i) the value of common and preferred stock at par value and related surpluses,¹³ (ii) undivided profits and capital reserves, (iii) cumulative foreign currency translation adjustment, and (iv) less the net unrealized loss on marketable equity securities. In examining these series, one can see that there have been changes to the definitions over time. In particular, some of the series do not exist prior to 1989.

The series for preferred stock and its surplus and the cumulative foreign currency translation adjustment series are not consistent. The preferred stock series is defined as the sum of different types of stock (perpetual preferred and limited-life) and is not consistent across the sample. Likewise the series for cumulative foreign currency translation adjustment only exists for the period since 1984. As an alternative definition of bank equity capital, we construct a consistent time series for the sum of common stock and its surplus, undivided profits, and capital reserves less the net unrealized loss on marketable equity securities. In Table A.1 we give the splices used to construct our equity series. In Figure A.2 we compare the two series and the graph documents that the two series are virtually identical so at least at the aggregate level the magnitude of the omitted terms is negligible.

Figure A.2 also documents that even at the aggregate level there are some obvious, large changes in bank capital that could overshadow the time-series properties of the bank capital series. The most dramatic change in bank capital occurs in the second quarter of 1987. The primary cause of this jump was the increase in loan loss reserves for expected losses on loans to developing countries (set off by Brazil's announcement that it would cease paying on its debts) made by the largest of banks (Federal Reserve Bulletin, 1988). Because these drops in bank capital do not necessarily reflect actual losses of bank capital, we choose to incorporate the loan loss reserve into the bank capital variable. Additionally, loan loss reserves are included in the measure of Tier 2 regulatory capital, which is used in calculating a bank's required capital-to-asset ratio. This has the advantage of smoothing the bank capital series and preventing the large changes in loan loss reserves from overshadowing the time-series behavior of the bank capital variable.¹⁴

¹² This item includes money market deposit accounts.

¹³ The surplus represents amounts paid in excess of the par (or stated) value from the issuance of common or preferred stock.

¹⁴ Another possible method that could be used to deal with this jump in the equity capital series would be to drop the largest ten banks from the sample, since the increase in the loan loss reserves was attributable to these banks.

A.2.2. Definition of Bank Universe (Screens)

The following four screens define our banking universe:

1. Banks located within the fifty states and DC ($0 < \text{rssd9210} < 57$).¹⁵
2. Insured banks ($\text{rssd9424} = 1, 2 \text{ or } 6$).¹⁶
3. Institutions chartered as commercial banks ($\text{rssd9048} = 200$).¹⁷
4. When loans or assets have a non-positive value then we eliminate that bank from the sample in that period.¹⁸ For assets we use rcfd2170 since no consistent counterpart is available for the rcon series.

Note that we restrict the bank universe to banks that are chartered as commercial banks. In particular, in contrast to Kashyap and Stein (2000), we exclude non-deposit trust companies ($\text{rssd9048}=250$) and savings banks ($\text{rssd9048}=300$) from the sample. First, the analysis of the monetary transmission mechanism as well as the interpretation of the results is more straightforward if we restrict ourselves to a more homogeneous group of banks. Also, the excluded banks report on a semi-annual and, in the case of savings banks, on an irregular basis at the beginning of the sample. Inclusion of these institutions would lead to an oscillating series.^{19,20} Throughout the first few years of the sample, the number of institutions licensed as commercial banks make up about 94% of the entire sample. Although the commercial bank share declined over time it was still approximately 88% in 1999. In addition, we do not include branches ($\text{rssd9048}=0$). Branches report irregularly and often do not report loans. Furthermore, if branch data were to be included one would have to subtract this information from the data provided by the bank headquarters of that branch to avoid double counting. The final screen is a screen to eliminate any banks reporting incorrectly.

¹⁵ There are some observations (usually empty sets) that have $\text{rssd9210} = 0$.

¹⁶ In particular, $\text{FDIC/BIF} = 1$, $\text{FDIC/SAIF} = 2$, and $\text{FDIC/BIF} \& \text{FDIC/SAIF} = 6$. This screen, thus, includes in the sample all banks insured by the Savings Association Insurance Fund (SAIF), Banker's Insurance Fund (BIF) and the FDIC.

¹⁷ Institutions chartered as commercial banks include trust companies other than non-deposit trust companies and credit card companies with commercial bank charters.

¹⁸ A bank displaying negative assets is a clear case of reporting error. Only four entries in the Call Reports are allowed to be negative (the most prominent of which is the entry for equity capital).

¹⁹ The loan portfolio of non-deposit trust companies is so small that including them in the bank universe does not lead to a noticeable change in the graph for total loans (also see Section A.7). Including savings banks, however, does lead to a visible oscillation of the series in the beginning of the sample.

²⁰ Since Kashyap and Stein worry only about getting sensible growth rates for each period, these banks would not completely drop out of their sample. Instead they would drop out only around those periods of missing observations.

A.2.3. Consolidated or Domestic Loan Data?

A major point to be aware of when working with the Call Reports is whether data are reported for all branches or for domestic branches only. Kashyap and Stein (2000) report:

“In general the largest banks only provide data on a consolidated foreign and domestic basis. This requires using the rcfd data series (rather than the rcon series). As a rule for the banks that do not have foreign operations one can assume that the rcfd data will be identical to the rcon data. Be advised that detailed data on the foreign [operations] is not available.”

For this reason we do not recommend using the rcon asset and equity series. However, Kashyap and Stein’s statement is not true for all balance sheet items and in particular is not true for deposits and loans. In the instructions to the Call Reports there is a section on reporting by office that states:

“Some information in the Reports of Condition and Income is to be reported by type of office (e.g., for domestic offices, for foreign offices, or for International Banking Facilities (IBF’s) as well as for the consolidated bank. Where information is called for by type of office, the information reported shall be the office component of the consolidated item”²¹

In particular, form FFIEC 031 contains two columns for many balance sheet items. For those balance sheet items, the first column requires the consolidated (rcfd) information and the second column the information from domestic offices (rcon).²² One advantage of the rcon loan series is that it is possible to construct a cleaner time-series for loans from the rcon data, in particular, the regulatory change in 1978 that creates an inconsistency in the rcfd loan series²³ does not affect the rcon series. Also, if one wants to analyze the effect of monetary policy shocks on domestic real activity and domestic financial variables, it makes more sense to use the rcon data, which does not include loans issued by foreign branches.

Figure A.3 plots the percentage change in the rcon and the rcfd loan series for the sample period from 1979 to the second quarter in 2000. As documented by the graph, the correlation between the two series is quite high; the correlation for the sample period from 1979 to the second quarter of 1999 is equal to 0.88. The changes in the rcon series are slightly more volatile than the changes in the rcfd series. Over the same sample the standard deviations are equal to 0.0154 and 0.0138 for the rcon and the rcfd series, respectively.

²¹ See p. 10 of “Instructions for Preparation of Consolidated Reports of Condition and Income” at <http://www.fdic.gov/regulations/resources/call/crinst/callinst2001.html>.

²² Balance sheet items that have to be reported on a consolidated and a domestic basis are loans and its various breakdowns and deposits and its breakdowns.

²³ Prior to 1978 the rcfd loan data is actually on a domestic basis and data for foreign branches is not even available.

A.3. Correcting for mergers

To construct consistent loan growth rate series we use the following screens:²⁴

1. Observation t of series i is dropped if the (log) growth rate is more than five standard deviations away from the cross-sectional mean growth of series i in period t .
2. Observation t of loan series i is only included when four preceding quarterly growth-rate observations of each of the four loan series considered are available.
3. Positive assets, $rcfd2170 > 0$.
4. Insured banks only ($rssd9424 = 1, 2$ or 6).
5. Banks involved in a merger were dropped for the quarter of the merger and the following.
6. Banks located in the fifty states and DC ($0 < rssd9210 < 57$).

From these screens it is possible to construct consistent time-series for the growth rates of individual bank variables. Using the data for the growth rates one can then construct indices for the regions as well as for the national aggregate. We will refer to the percentage changes of the end-of-period call series constructed in Section A.2 as the “CALL” loan series and to the percentage change constructed in this section as the “CALL(corrected)” loan series. As documented in Figure A.4, the aggregate total loan series are extremely similar. The correlation coefficient is equal to 0.97 for the sample from 1977:2 to 2000:2. At the regional level, however, mergers clearly affect the uncorrected loan series. To document this we report in Table A.2, the correlation coefficients between the two total loan series for the nine census regions. The coefficients in the first column are for the period from 1977:2 to 2000:2 and those in the second column are for the period from 1977:2 to 1999:2. Occasional outliers that are observed in periods when a large merger occurs heavily influence the correlation coefficients. This becomes clear when we compare the two samples. Just excluding the last four quarters raises the correlation coefficient substantially in the South Atlantic, East South Central and Pacific census regions because large mergers affected the uncorrected loan series during the last four quarters in these regions. In particular, in the Pacific region the correlation coefficient increases from 0.41 to 0.89 when the last four quarters, that include a 44% drop in the third quarter of 1999, are excluded. These findings are caused by the merger between Bank of America (Pacific) and Nations Bank (South Atlantic) and the choice to locate the headquarters of the new corporation in Charlotte NC, which brought about an enormous reassignment of loans across regions.

²⁴ Here we follow Kashyap and Stein (2000), but we exclude non-deposits trusts and savings banks from the sample even though this set of screens could in principle correct for the irregular reporting of these institutions in the beginning of the sample.

Using the index series for aggregate analysis?

Each index series starts at the observed level in the first quarter of 1977. If a bank is excluded, for example, because of a merger, then our procedure implicitly assumes that this bank's balance sheet variables are growing at the same rate of the other banks in the same region. In constructing the state-level index series the average growth rate of other banks in the same state is used, in constructing the census-level index series the average growth rate of other banks in the same census region is used, and in constructing the national index series the average growth rate of banks across the nation is used. If one is interested in time-series analysis using aggregate data one has three choices. First, one can use the unadjusted level series. Mergers and takeovers are not a problem in this data set and there are virtually no other reasons why banks disappear out of the sample.²⁵ Second, one can use the aggregate index series. Third, one can use the sum of the state-level index series. If a bank is more likely to grow at the rate of other banks in the same region than at the national average, then the third choice may be a better choice than the second.

A.4. Construction of Data Sets for Small Banks

To construct data sets for small banks we first apply the screens discussed above to determine which banks are included in the sample in a specific quarter. After this we rank the banks by the value of their total assets (consolidated basis, i.e., rcf2170) and then define small banks as those banks whose asset value is less than the asset value corresponding to the asset value of a certain percentile.

Studying the behavior of loan variables of small banks is interesting in itself, since especially small banks are thought to have difficulty financing their loan portfolio during periods of monetary tightening, which could very well worsen the economic downturn.²⁶ In addition, small banks have not been involved in mergers as much as large banks. For that reason, the data set for small banks might be cleaner than the data that include all banks. Finally, domestic branches of small banks are less likely to issue loans to foreign firms or issue loans to multinational companies that are used for investments abroad which is an advantage if one wants to study the interaction between bank variables and domestic real activity.

For deposits and total assets we report both data for domestic branches (rcon series) and for the consolidated data for both domestic and foreign branches (rcfd series). When banks have no foreign offices then these series would be the same. This is exactly what we observe for the 90 and 95 percentile small bank data sets.

A.5. Data Revisions

The data that is made available on the Chicago website is sometimes revised. This happens mainly for the most recent data but we have encountered cases where even data ten years back was revised. There are several reasons why the data are revised. In the past the merger file has been substantially updated and now includes a much larger set of financial institutions. Data for individual banks' balance sheet

²⁵ Bankruptcies typically also lead to a merger or takeover by a bank in the sample.

²⁶ See, for example, Kashyap and Stein (1995,2000).

variables can change, for example, because the FDIC or the Federal Reserve Board discovers that a bank misclassified a particular bank transaction. In this case the auditing institution may ask the bank to correct this mistake in past reports as well.

A.6. The Federal Reserve's H.8 Statistical Release

The Board of Governors of the Federal Reserve publishes a weekly aggregate balance sheet, the H.8 Statistical Release, for all banks operating in the U.S. and for domestically chartered banks.²⁷ Data are available on a weekly and a monthly frequency, and on a seasonally adjusted or non-seasonally adjusted basis.²⁸ The monthly series is simply the average of the weekly releases. While the H.8 series is based on the Call Reports, the data actually come from voluntary weekly (Wednesdays) bank credit reports—forms FR2416, FR2644 and FR2069—submitted to the Federal Reserve.²⁹ Since participation is voluntary the data collected only represent a sample of U.S. banks, which are subsequently *blown-up* to be representative of the entire bank universe. The Board reports series both for *domestic* banks that only include data for domestic banks and it reports series for *all banks* that include data for foreign banks operating in the U.S.. The instructions to the H.8 survey explicitly state: “*All individual asset and liability items should exclude, to the extent possible, the asset and liability relationships with foreign offices.*” The domestic H.8 data should, thus, be comparable to the rcon data from the Call Reports, i.e., the unconsolidated data.

The bank universe of the H.8 (in terms of Call Report codes) for the domestic banks is defined as follows:³⁰

1. Banks located within the fifty states and DC ($0 < \text{rssd9210} < 57$).
2. Insured banks only ($\text{rssd9424} = 1, 2$ or 6).
3. Domestic Banks:
 - a. $\text{rssd9331} = 1$ (Commercial bank) and $\text{call8786} = 1$,³¹ or
 - b. $\text{rssd9331} = 1$ and $\text{call8786} = 2$, or
 - c. $\text{rssd9048} = 250$.

²⁷ Data are also reported for the following subgroups: large domestic banks, small domestic banks, and foreign related institutions.

²⁸ Monthly data are seasonally adjusted using the X12 procedure. Since monthly data are averages of the weekly data the weekly data are seasonally adjusted to conform to the monthly adjustment.

²⁹ FR2416 is the most complete and is filed by participating large commercial banks. The size cutoff for filing form FR2416 is 6 billion dollars or more in total assets (as of the first quarter of 1996). However, some banks with less than 6 billion dollars in assets are added to the sample to ensure adequate regional participation. Form FR2644 is less complete, collecting no deposit data, and is a random stratified sample of about 1,100 U.S. commercial banks that do not report on form FR2416. Finally, U.S. branches and agencies of foreign banks with 1.1 billion dollars or more in assets (including IBF's) file form FR2069. Other large branches and agencies are added to the sample to ensure adequate world coverage.

³⁰ William Watkins of the Federal Reserve Board provided this information to us.

³¹ The screen $\text{call8786} = 1$ selects consolidated entities with foreign office(s) and the screen $\text{call8786} = 2$ selects entities with domestic office consolidation.

The bank universe of the domestic H.8 is, thus, almost identical to the universe defined in Section A.2. One difference is that the H.8 in addition to commercial banks, also includes non-deposit trust companies (rssd9048=250).³² As discussed in Section A.2, some of these institutions report irregularly in the beginning of the Call Report sample, but as documented in Section A.7 below their loan portfolio is very small relative to the total loan portfolio of commercial banks.

Although not used here, we report for completeness the screens that define the H.8 bank universe for all “all banks”. Besides the domestic institutions defined by the third screen above, it also includes the following institutions:

4. rssd9331 = 95 (American Express International Banking Corporation),
5. rssd9331 = 12 (New York state chartered subsidiary of a foreign bank),
6. rssd9331 = 9 (U.S. branch of a foreign bank),
7. rssd9331 = 11 (Agencies of foreign banks),
8. rssd9331 = 13 (State licensed agency, acting as a bank under IBA),
9. rssd9331 = 21 (Banking Edge Corporation) and call8786 = 1, and
10. rssd9331 = 23 (Banking Agreement Corporation) and call8786 = 1.

A.7. The H.8 versus the Call Reports

As discussed in the first section of this manual much of the work that has been done on aggregate monetary transmission and bank balance sheet variables has relied on the H.8 Statistical Release. The advantage of the H.8 over the Call Reports is that it is available for a longer period.³³ The key drawback of the H.8 is, of course, that it is based on a voluntary survey of banks that is adjusted to be representative of the entire banking sector. We do not know how the data are adjusted and whether the adjustment procedure has remained constant over time.³⁴ Also, the data are available only on the aggregate level. In contrast, all insured banks are required to submit Call Reports and so the Call Report data provides a nearly complete survey of U.S. banks and, since the data are on the individual bank level, it allows for constructing data by region.

It is important to realize that the Call Reports are a quarterly snapshot of the banking industry. So in the case of loans, banks must report loans outstanding as of the close of business on the last calendar day of the quarter. The H.8 weekly releases are also snapshots, albeit weekly (Wednesday) snapshots of the sample of reporting banks. We match the end-of-quarter Call Report observation with the Wednesday observation that is as close as possible to the end of the quarter.

³² Both the screens in 3a and 3b, that are used by the Federal Reserve Board, and the screen rssd9048 = 200 used by us and by Kashyap and Stein select commercial and industrial banks.

³³ The data are available online back to the beginning of 1973. However, if one refers to the Board of Governors, Federal Reserve System *Banking and Monetary Statistics, 1941-1970* and *Annual Statistical Digests* a form of the data are available back to 1941.

³⁴ For example, for empirical work it would be important to know whether past observations are adjusted using more recent observations.

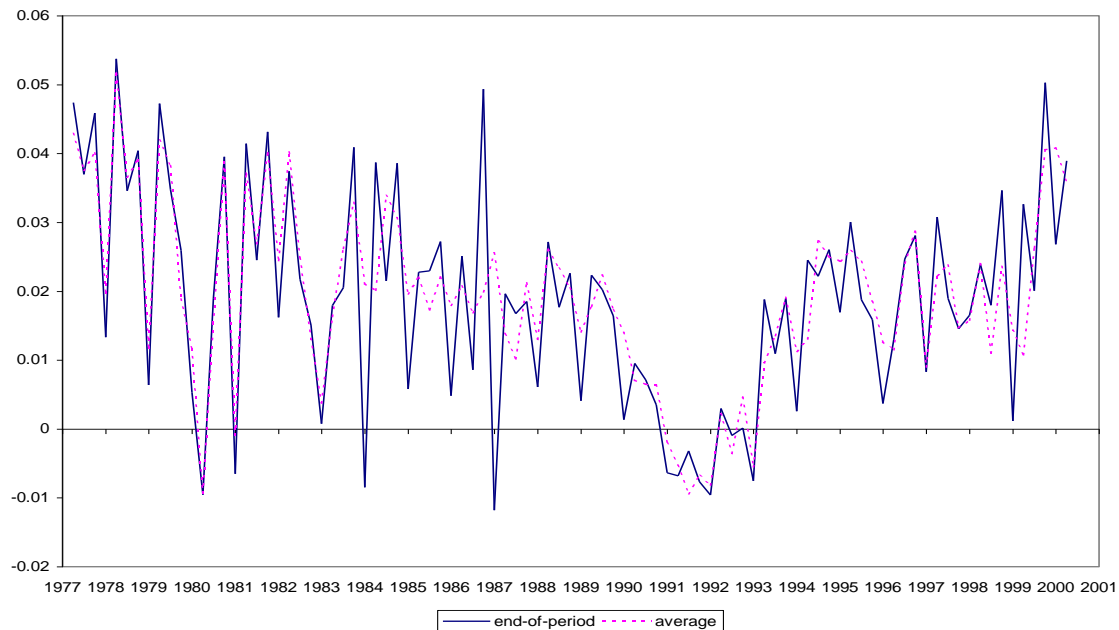
Figure A.5 compares the (log-levels of the) end-of-quarter H.8 series to both end-of-quarter Call Report series. The figure documents that the levels of the H.8 series are close to the levels of the Call Report series even though the Call Report series do not include non-deposit trust companies. Figure A.6 plots the percentage changes in these three series and Table A.3 reports the correlation coefficients between these growth rate series. The graph and the table document the high correlation between the end-of-quarter data from the Call Reports and the corresponding end-of-quarter data published by the Board of Governors.

Figure A.7 plots the quarterly average domestic loan growth rate series published by the Board of Governors and compares it with the corresponding average series from the Call Reports. The graph as well as the correlation coefficient reported in Table A.3 documents that there are bigger differences between the series when we consider quarterly averages. Given the high correlation found for the end-of-quarter series, we suspect that the freedom in choosing how to report averages on the Call Reports makes these series less reliable.

To see whether the differences between the reported loan series are important for typical empirical exercises, we report in Figure A.8 the impulse response function of total loans in response to a federal funds rate shock. The impulse response function is calculated using a VAR that contains the federal funds rate, the price level, aggregate income and one of the loan series.³⁵ Since we consider a total of five different loan series, we have five different estimated VAR's and five different sets of impulse response functions for each identification scheme. Each VAR has two lags and is estimated over the sample period from 1977 to the second quarter of 2000. We see in Figure A.8 that the loan impulses display substantial differences. The impulses for loans when the Call Report data are used are substantially smaller than when the H.8 series are used. For example, the value of the impulse response after eight periods for the end-of-period H.8 loan series is 33% higher than the impulse for the Call loan series and 55% higher than the impulse response for the Call series that corrects for mergers. Not in all cases are the results different. Figure A.9 plots the impulse response function of total loans in response to a one standard deviation funds rate shock when income is held constant. This impulse response function is equal to the difference between the impulse response function in response to a federal funds rate shock (monetary downturn) and the response caused by a set of income shocks (non-monetary downturn) that generate the same income response as the one observed in response to a shock in the federal funds rate. Now, the VAR's with the different loan series generate responses of fairly similar magnitude.

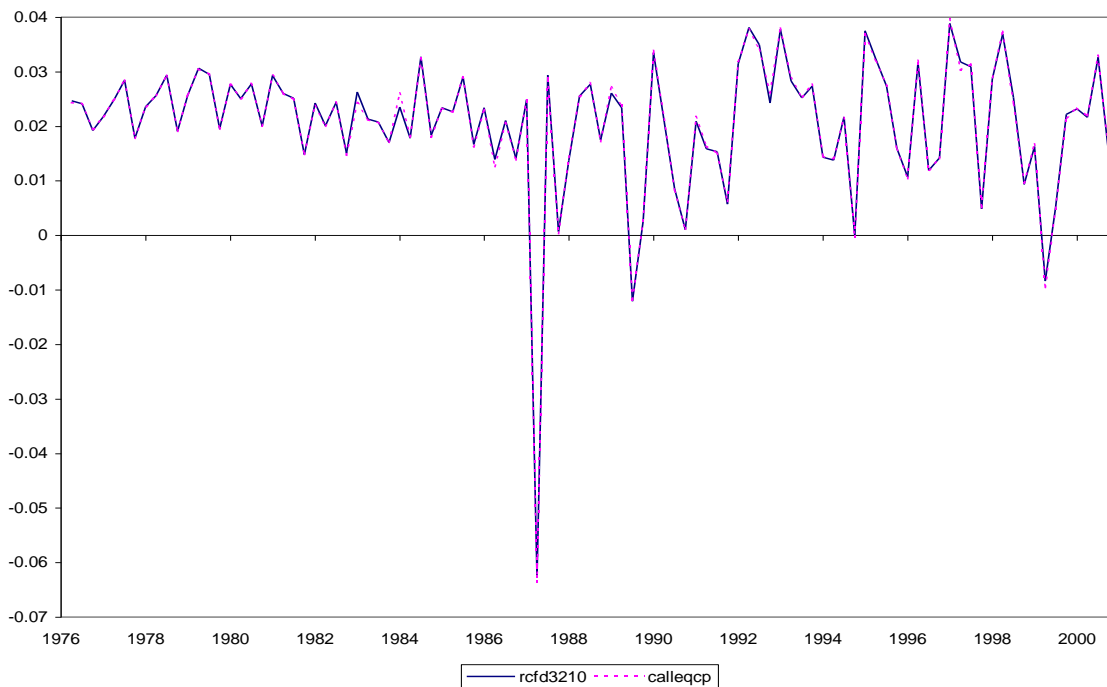
³⁵ This ordering is used in the Cholesky decomposition of the estimated variance-covariance matrix. That is, we assume that innovations in a variable do not have contemporaneous effects on variables that appear earlier in the ordering.

Figure A.1: Comparison Between End-of-period and Average Loan Series
(percentage changes)



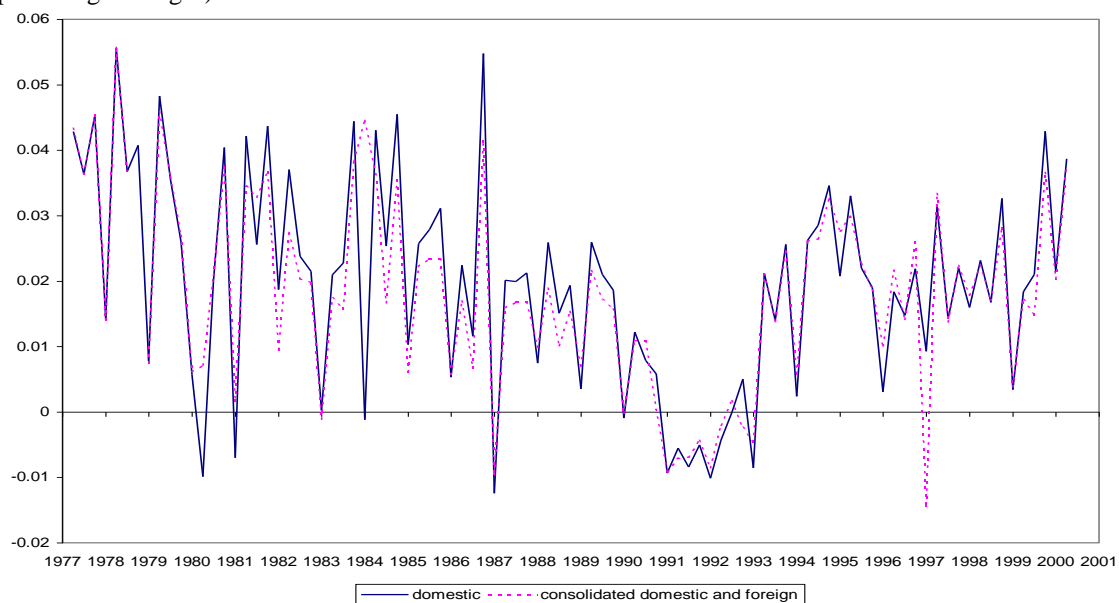
Note: This graph plots the end-of-quarter domestic loan series and the corresponding quarterly averages constructed from the Call Reports.

Figure A.2: Comparison between Call Report bank equity series and consistent definition
(percentage changes)



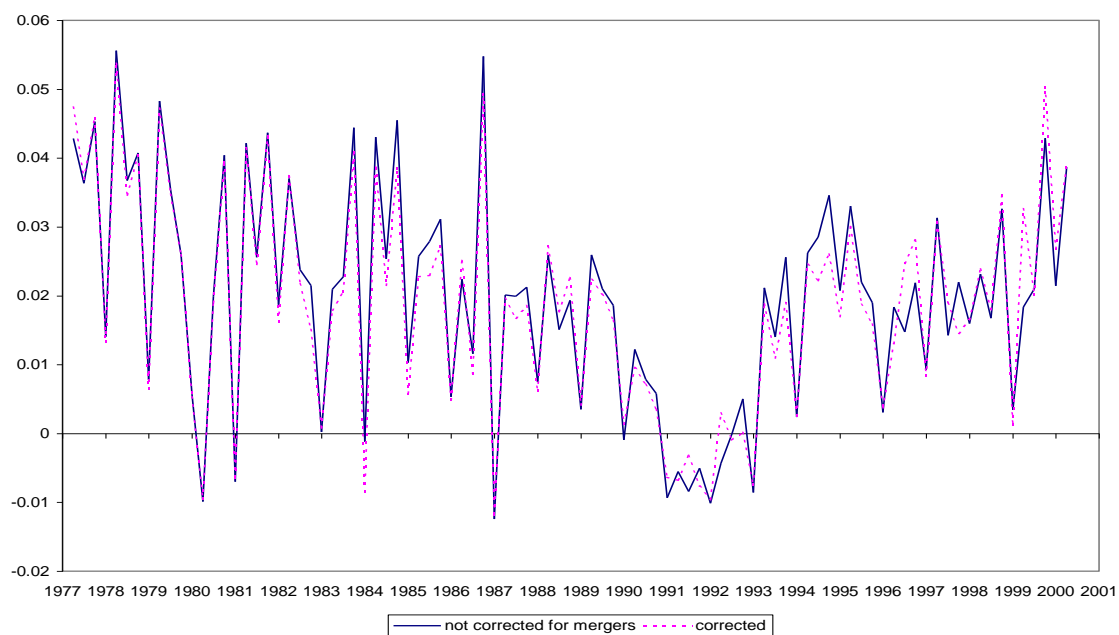
Note: This graph plots the (inconsistent) equity series from the Call Reports and the consistent series we constructed.

Figure A.3: Comparison Between Domestic and Consolidated Loans
(percentage changes)



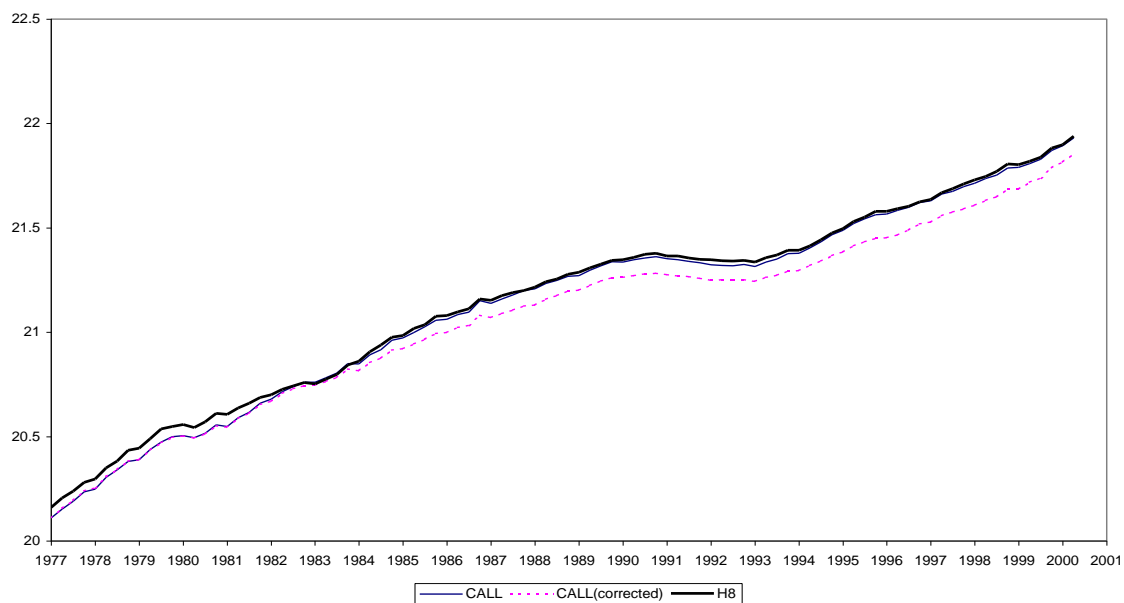
Note: This graph plots the domestic (rcon) and the consolidated (rcfd) loan series constructed from the Call Reports that are not corrected for mergers. For the consolidated data the observation for the fourth quarter of 1978, the period of the break, is excluded.

Figure A.4: Comparison Between (for Mergers) Corrected and Uncorrected Series
(percentage changes)



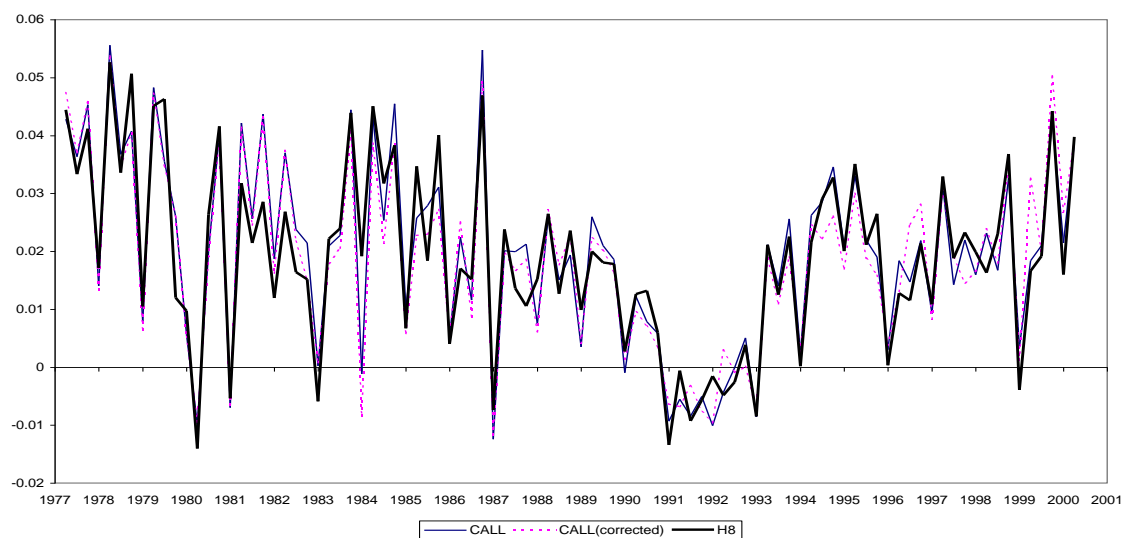
Note: This graph plots the end-of-quarter domestic loan series and the corresponding series constructed using the screens from Kashyap and Stein (2000) that corrects for mergers.

Figure A.5: Comparison Between Call Report Data and H.8 Data
(end-of-quarter, log-levels)



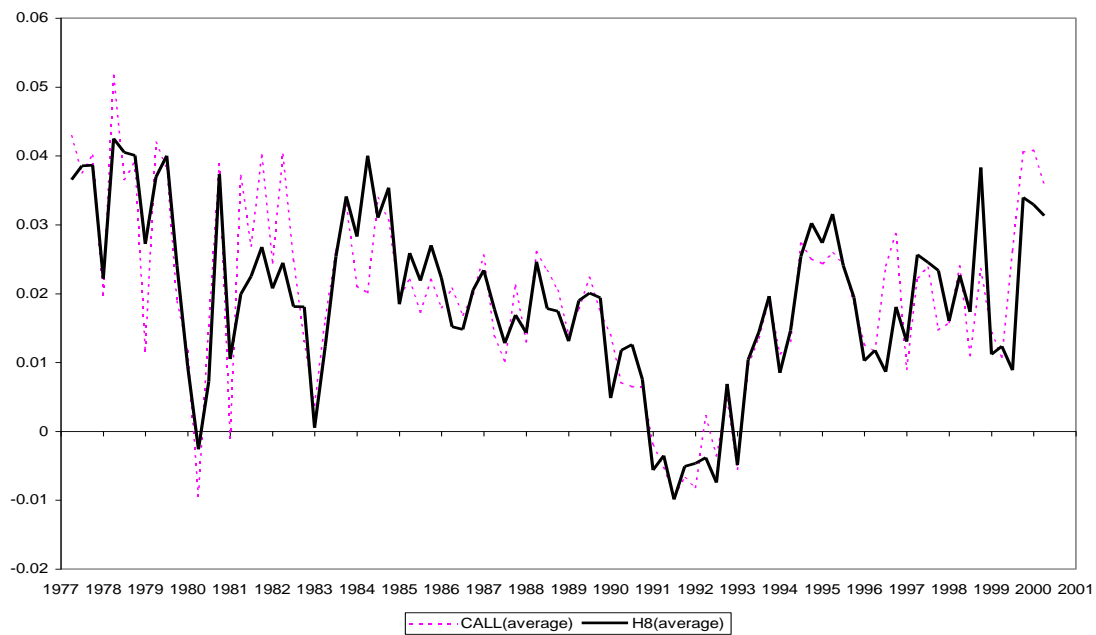
Note: This graph plots the end-of-quarter domestic loan series (CALL), the corresponding series constructed using the screens from Kashyap and Stein (2000), that corrects for mergers (CALL corrected) and the corresponding series published by the Board of Governors (H8).

Figure A.6: Comparison Between Call Report Data and H.8 Data
(end-of-quarter, percentage changes)



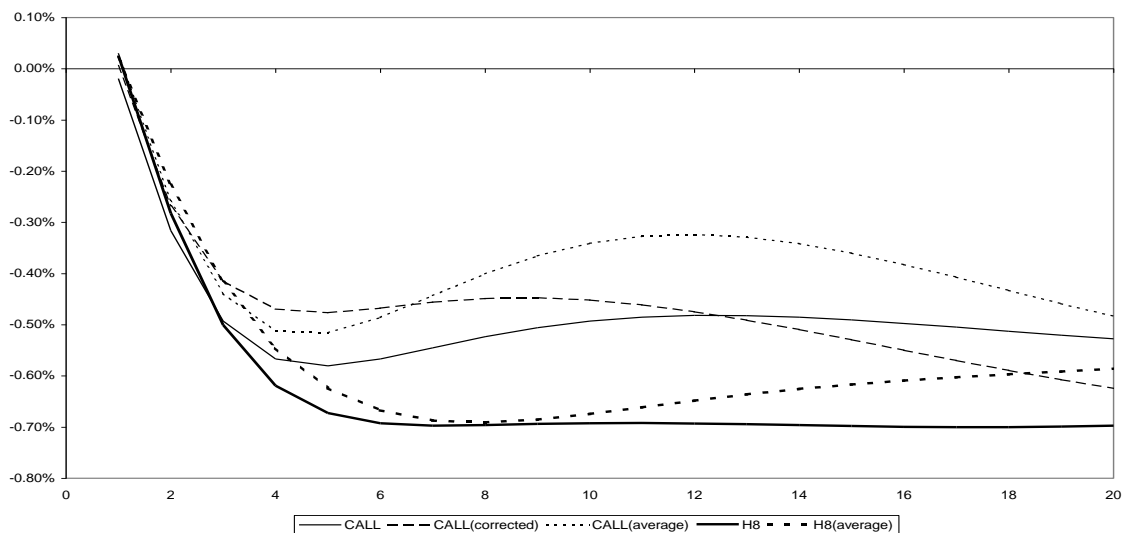
Note: This graph plots the end-of-quarter domestic loan series (CALL), the corresponding series constructed using the screens from Kashyap and Stein (2000) that corrects for mergers (CALL corrected), and the corresponding series published by the Board of Governors (H8).

Figure A.7: Comparison Between Call Report Data and H.8 Data
(quarterly averages, percentage changes)



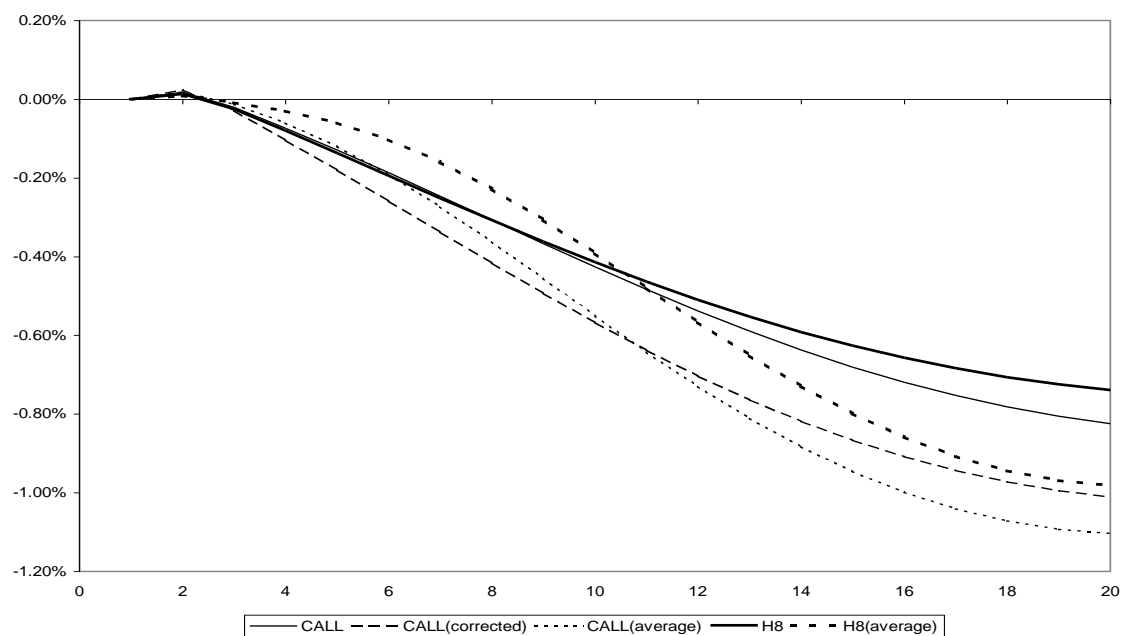
Note: This graph plots the quarterly-averages domestic loan series (CALL average) and the corresponding series published by the Board of Governors (H8 average).

Figure A.8: Response of Loans to Funds Rate Shock



Note: This graph plots the impulse response functions of total loans in response to a one standard deviation funds rate shock for five separate VARs. Each VAR contains the federal funds rate, the consumer price index, the indicated total aggregate loan series and real aggregate income.

Figure A.9: Loan Responses relative to those during similar non-monetary downturn



Note: This graph plots the differences between the impulse response function of total loans in response to a one standard deviation funds rate shock and the impulse response function of total loans during a similar non-monetary downturn for five separate VARs. Each VAR contains the federal funds rate, the consumer price index, the indicated total aggregate loan series and real aggregate income.

Table A.1: Bank variables

Variable	code	Code	Period
Total loans	calltln	rcon1400 + rcon2165 rcon1400	'76:1-'83:4 '84:1-'00:2
C&I loans	callciln	rcon1600	'76:1-'00:2
Real estate loans	callreln	rcon1410	'76:1-'00:2
Consumer loans	callcoln	rcon1975	'76:1-'00:2
Total loans (average)	callavln	rcon3360	'76:1-'00:2
Bank Equity	calleqcp	rcfd3230+rcfd3240+rcfd3247 rcfd3230+rcfd3839+rcfd3632-rcfd0297 rcfd3230+rcfd3839+rcfd3632+rcfd8434	'76:1-'89:4 '90:1-'93:4 '94:1-'00:4
Loan Losses	calllnls	rcfd3123	'76:1-'00:2
Bank Equity + Loan Losses	callecll	rcfd3230+rcfd3240+rcfd3247+rcfd3123 rcfd3230+rcfd3839+rcfd3632-rcfd0297+rcfd3123 rcfd3230+rcfd3839+rcfd3632+rcfd8434+rcfd3123	'76:1-'89:4 '90:1-'93:4 '94:1-'00:4
Bank Assets (not recommended)	rcon2170	rcon2170	'76:1-'00:2
Bank Assets (consolidated)	callasst	rcfd2170	'76:1-'00:2
Deposits	calldeps	rcon2200	'76:1-'00:2
Deposits (consolidated)	rcfd2200	rcfd2200	'76:1-'00:2
Alternative bank equity	rcfd3210	rcfd3210	'76:1-'00:2

Table A.2: Comparison of Loan Series before and after Merger Correction

Region	Correlation Coefficient (1977:2-2000:2)	Correlation Coefficient (1977:2-1999:2)
New England	0.78	0.80
Mid Atlantic	0.87	0.87
East North Central	0.79	0.78
West North Central	0.69	0.69
South Atlantic	0.55	0.76
East South Central	0.53	0.79
West South Central	0.73	0.73
Mountain	0.43	0.44
Pacific	0.41	0.89

Note: This table reports the correlation between the loan growth rate series that does and the loan growth rate series that does not correct for mergers for the indicated sample period.

Table A.3: Correlation Coefficients of Loan Growth Rates (1977:2-2000:2)

	CALL	CALL (corrected)	H.8	CALL (average)	H.8 (average)
CALL	1.00				
CALL (corrected)	0.97	1.00			
H.8	0.94	0.90	1.00		
CALL (average)				1.00	
H.8 (average)				0.82	1.00

Note: This table displays the correlation coefficients between the various aggregate loan series.