REVIEW

SEPTEMBER/OCTOBER 1999

Michael Dueker is a research officer at the Federal Reserve Bank of St. Louis. Nick H. Meggos and Thomas A. Pollmann provided research assistance.

Measuring Monetary Policy Inertia in Target Fed Funds Rate Changes

Michael Dueker

Inflation in the United States has been relatively low and stable following the difficult disinflation during the early 1980s. Since then, the Federal Reserve has been perceived as following policies designed to preempt rising inflation. At the same time, empirical studies of Federal Reserve policy actions find that policy responses are generally stodgy and that a great deal of interest rate smoothing takes place, relative to rule-based policy prescriptions, such as the Taylor rule (Taylor, 1993). Several explanations have arisen as to why a central bank that is focused primarily on inflation control would exhibit policy inertia when changes seem indicated. Rudebusch (1999) and Orphanides (1998) study data uncertainty as a possible justification for stodgy policy responses, because revisions take place before estimates of output and price indices are final and policymakers want to avoid acting on false signals. This justification falls under the umbrella of recognition lags in that it takes time for monetary policymakers to observe that inflation has moved decisively upward or that an output gap has developed. Sack (1998) suggests that uncertainty regarding the structure of the economy may be a reason for the Federal Reserve to respond cautiously. Goodfriend (1997) notes that if aggressive Federal Reserve actions subsequently proved to be mistimed, the public's trust in the Fed would diminish. Hence, policymakers might tend to act less aggressively, again due to uncertainty regarding the timeliness of currently contemplated policy actions. Woodford (1999) offers a different explanation of policy inertia that is not based on model or instrument uncertainty. In Woodford's view, every policy action also is a hint of how policy will be conducted in the future. In this case, part of the central bank's credibility rests on making current policy roughly consistent with the path to which it had hinted through earlier policies. In this case, some policy inertia results as policymakers balance their impetus to respond to current news against their implicit prior commitment to a particular path for interest rates.

This article presents two distinct measures of inertia in the target fed funds rate, relative to the prescriptions of Taylor's (1993) policy rule, including the inertia inherent in the discreteness of the target fed funds rate. Previous empirical estimates of policy inertia have not addressed the discrete nature of target funds rate changes; instead, they use quarterly or monthly averages of the funds rate, which can mask an important feature of the policymaking process. The Federal Open Market Committee (FOMC) sets its objective, or target, for the federal funds rate, the interest rate that banks charge each other for overnight loans of Federal Reserve deposits. In practice, the FOMC usually changes the target fed funds rate in discrete amounts by multiples of 25 basis points. Thus, an important facet of policy inertia takes the form of the following question: How far does the FOMC let the prevailing target funds rate get out of line, relative to a shadow desired level that changes continuously? Another facet, which has been the subject of previous study, is the degree to which the fed funds rate obeys a partial-adjustment mechanism and the sluggishness implied by such a mechanism. A joint view of these two facets of policy inertia ought to provide a more complete picture, from which subsequent research can investigate whether policy appears to display either too much or the right amount of inertia.

In particular, this article presents an econometric model of discrete changes in the target fed funds rate in order to estimate thresholds at which the FOMC decides to change the target and by how much. The model defines a latent desired target level, so the threshold coefficients represent the sizes of gaps between the desired and actual target funds rates that are necessary to induce target changes of various sizes. These estimated thresholds are compared with the sizes of the actual changes in the target fed funds rate—usually 25 or 50 basis points—to arrive at a quantitative measure of the Fed's readiness or reticence to initiate changes in the target funds rate. The complete set of threshold coefficients also provides estimates of cutoff levels at which the Fed chooses to make either a small or large change in the target funds rate. The payoff to estimating threshold coefficients is that we can then separate any slugghishness in changes to the target fed funds rate introduced by its discrete nature from sluggishness due to data or model uncertainty. Also, tests for asymmetry in the thresholds between increases and decreases in the target funds rate are possible.

DISCRETE NATURE OF TARGET CHANGES

The discrete nature of changes in the target fed funds rate poses special challenges to empirical analysis. Almost all changes in the target funds rate are in multiples of 25 basis points. In this article, I use data on the target funds rate that start in 1985 (Rudebush, 1995). A plot of the target funds rate is shown in Figure 1. To match the frequency of some of the explanatory variables, I calculate the change in the target funds rate from the last business day of the month to the last business day of the previous month. Prior to 1990, some changes were as small as six basis points, with some other small changes of 12.5 basis points. I classified monthly changes of less than 18 basis points as "no change" with two exceptions: In September 1985 and November 1988 the target funds rate increased by 12.5 basis points for the second month in a row; I counted the combined change as a 25 basis point increase during the second month.

Table 1 summarizes the five discrete categories I use for the changes in the target funds rate, where the target funds rate is denoted FF^T. The last column shows that, despite a number of odd-sized changes in the target through 1989, the means within the five categories correspond very closely to multiples of 25 basis points: (–.50, –.25, 0, +.25, +.50). The target funds rate has decreased on net since 1985, resulting in more 25 basis point decreases than increases.

TAYLOR'S RULE IGNORING DISCRETENESS

Following Judd and Rudebusch (1998), I use Taylor's rule to motivate an empirical specification for modeling changes in the federal funds rate. Taylor (1993) suggested that the FOMC's behavior from 1987 to 1992, with respect to setting a target for the federal funds rate, appeared to be well-summarized by a simple monetary policy rule. This interest-rate rule, which became known as Taylor's rule, is one in which the Federal Reserve changes its objective for the federal funds rate in response to the gap between the actual and desired levels of inflation, and to the percentage gap between actual and potential output. In its original specification, Taylor's rule takes the following form under the assumptions that the FOMC's long-run desired rate of inflation is 2 percent and that the equilibrium real short-term interest rate also is 2 percent (Orphanides, 1998):

$$(1) \quad FF_t = 1 + 1.5\pi_t + .5\left(y - y^p\right)_t,$$

where FF stands for the federal funds rate, π for the inflation rate, y for the log of actual output and y^P for the log of potential output. In long-run equilibrium, the output gap will be zero and the equilibrium real rate of interest $(FF - \pi)$ will equal two by assumption. Thus, equation 1 implies a long-run inflation target of 2 percent.

The most important empirical lesson from Taylor's rule has not been to check whether actual policy is consistent with 2 percent inflation, however. Instead, Taylor (1998) gives considerable importance to the coefficient on inflation having an absolute value greater than one, because if the Fed were to raise the fed funds rate by more than any increase in inflation, then the real interest rate would increase, thereby dampening inflationary pressures in the economy. Any policy equation in which the response of the nominal interest rate instrument to a change in inflation is greater than one is said to have the Taylor-rule property, regardless of the implied long-run inflation target. In empirical research, Taylor's original equation often is modified to include a lagged dependent variable. Rules with gradual adjustment have the Taylor property if the long-run response of the interest-rate instrument to a change in inflation is greater than one:

$$(2) \quad FF_t = \rho FF_{t-1} + \lambda_o + \lambda_1 \pi_t + \lambda_2 \left(y - y^P\right)_t + \delta \Delta FF_{t-1}$$

$$\lambda_1 / (1 - \rho) > 1$$

For positive values of ρ, equation 2 allows for more gradual adjustment of the federal funds rate to changes in inflation. Sack (1998) and Judd and Rudebusch (1998) have used such a partial adjustment approach to model FOMC policy. The caveat is that ordinary least squares estimation of the partial-adjustment mechanism from equation 2 ignores the discreteness in the actual changes to the target fed funds rate. Estimates of equation 2 with monthly data starting in 1985 appear in Table 2, where π is the year-over-year inflation rate calculated from the consumer price index without the food and energy components. The funds rate data, *FF*, are the monthly averages of the effective funds rate in the upper panel of Table 2 and are the end-of-month values of the target fed funds rate in the lower panel of Table 2. The log of real GDP (deflated by the chain-weighted price index) is y, and y^P is the log of the potential output series from the Congressional Budget Office. Note that y and y^P only are available at the quarterly frequency. The output gap tends to evolve slowly, however, so I used the same value of the output gap for three months running in order to estimate equation 2 at a monthly frequency. Figure 2 plots the output gap since 1985 in logs, so that 100 times the log gap is the output gap as a percentage.

Table 2 shows that for either

Figure 1

Fed Funds Target

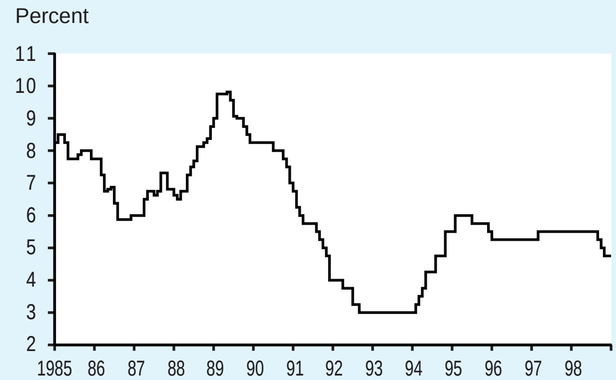


Table 1

Observation Categories Based on Size of Target Rate Change. Monthly Sample: 1/85 — 12/98

Category	Criterion	Frequency	Mean D *FFT* within category
1	$\Delta FF^T < -.25$	11	–.523
2	$-.25 \le \Delta FF^T \le -.18$	25	–.248
3	$-.18 < \Delta FF^T < +.18$	111	+.0056
4	$+.18 \le \Delta FF^T \le +.25$	10	+.244
5	$\Delta FF^T > +.25$	10	+.538

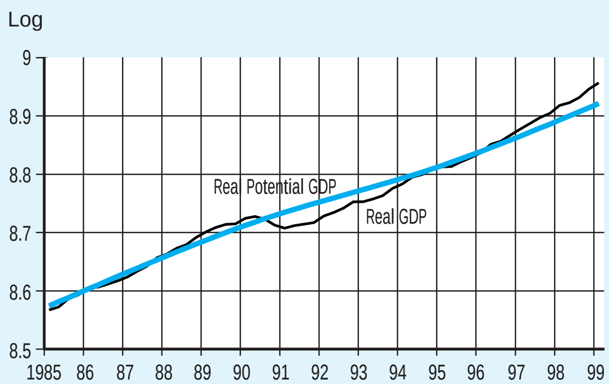
Table 2

Estimated Taylor Rules from Equation 2
Monthly Data: 1/85 — 12/98

Coef cient	Description	Value	Std. Error
	Results for Effective Funds Rate		
λ_1	Inflation	0.086	0.044
λ_2	Output Gap	0.063	0.022
ρ	Autoregression	0.937	0.022
δ	Lagged change	0.379	0.074
λ_0	Constant	0.061	0.072
	Results for Target Funds Rate		
λ_1	Inflation	0.080	0.044
λ_2	Output Gap	0.065	0.023
ρ	Autoregression	0.939	0.023
δ	Lagged change	0.342	0.075
λ_0	Constant	0.068	0.072

Figure 2

Real GDP and Real Potenial GDP
(in Logs)



dependent variable the long-run response of the funds rate to a percentage point increase in inflation is greater than one $\lambda_1 / (1 - \rho) = 1.37$ for the effective fed funds rate and 1.31 for the target fed funds rate. In Table 2, the estimated long-run response of the funds rate to a one percentage point increase in the output gap, $\lambda_2 / (1 - \rho)$, is about one, which is larger than the 0.5 coefficient Taylor wrote in his original rule. Judd and Rudebusch (1998) also find that estimated feedback coefficients on the output gap are larger than originally specified by Taylor in equation 1.

A DISCRETE EMPIRICAL MODEL WITH THRESHOLDS

In this section, I examine a Taylor-rule specification of the target federal funds rate in an ordered probit model that has been enhanced with time-series features. An ordered probit model, unlike the ordinary least-squares regression from Table 2, takes into account the discrete nature of target changes. As mentioned in the introduction, the discrete-variable model includes threshold coefficients that provide information on another possible manifestation of policy inertia: the degree of interest-rate misalignment required to induce a discrete change in the target level of the funds rate.

The five categories defined in Table 1 fit naturally into an ordered probit framework. As in probit models in general, one assumes that there is a continuous latent or desired level of the target funds rate, which is denoted as FF^{T*}. A standard assumption in probit models is that the latent variable is a linear function of some lagged explanatory variables, X, plus a normally distributed, mean-zero error term, ε:

$$(3) \qquad FF_t^{T*} = X_{t-1}'\beta + \varepsilon_t .$$

In keeping with the Taylor-rule specification of equation 2, the X variables are inflation, the output gap, and an intercept. Given the evidence of a partial-adjustment mechanism from the estimates of equation 2, I also add a lagged dependent variable and the lagged change in the dependent variable to equation 3, following Judd and Rudebusch (1998):

$$(4) \qquad FF_t^{T*} = \rho FF_{t-1}^{T*} + \lambda_0 + \lambda_1 \pi_t + \lambda_2 \left(y - y^p\right)_t + \delta \Delta FF_{t-1}^{T*} + \varepsilon_t .$$

The assumed mapping between the latent variable and the observable discrete changes in the target funds rate is

$$(5)\quad \Delta FF_t^T \textit{ is in category } j \textit{ if}$$

$$FF_t^{T*} - FF_{t-1}^T \textit{ is in } \left(c_{j-1}, c_j\right),$$

where the categories are defined in Table 1 and c is a vector of threshold coefficients. The difference $FF_t^{T*} - FF_{t-1}^T$ represents the latent "desired" change in the target funds rate, as defined by the econometric model. The estimated threshold coefficients indicate the degree of underlying pressure that is needed to induce a discrete change in practice. For the threshold coefficients to have this meaning, the latent variable must be restricted to economically relevant values. I do this by constraining the intercept such that the mean of the latent variable is the same as the mean of the observed target rate:

$$\overline{FF^{T*}} = \overline{FF^T} = 5.91.$$

This restriction implies that the discrete variable has the same sample-wide mean as the continuous latent variable.

MODEL ESTIMATION

The log-likelihood function for the observed changes in the target fed funds rate is

$$(6)\quad lnf\left(\Delta FF_t^T\right) = \sum_{j=1}^{5} D\left(cat.j\right)$$
$$ln\{[\Phi((FF_{t-1}^T - \rho EFF_{t-1}^{T*}$$
$$- \lambda_1 \pi_{t-1} - \lambda_2 (y - y^p)_{t-1}$$
$$- \delta E \Delta FF_{t-1}^{T*} + c_j)/\sigma)$$
$$-\Phi((FF_{t-1}^T - \rho EFF_{t-1}^{T*}$$
$$- \lambda_1 \pi_{t-1} - \lambda_2 \left(y - y^p\right)_{t-1}$$
$$\delta E \Delta FF_{t-1}^{T*} + c_{j-1})/\sigma)]\},$$

where $\Phi(.)$ is the cumulative normal density function and D(cat. j) is a dummy variable for category j at time t. For the lagged dependent variable, FF_t^{T*}, we need to take an expected value, because we do not observe the realized residual, ε. Therefore, we use

$$EFF_t^{T*} = \rho EFF_{t-1}^{T*} + \lambda_0$$
$$+ \lambda_1 \pi_{t-1} + \lambda_2 \left(y - y^p\right)_{t-1}$$
$$-\delta E \Delta FF_{t-1}^{T*} + E\Big[\varepsilon_t \,\Big|\, \Delta FF_t$$
$$\textit{is in cat.} j\Big]$$

$$E\left[\varepsilon_t \middle| \Delta FF_t \textit{ is in cat.} j\right] =$$
$$-\sigma / \left(2\pi\right)^{.5}\Big[exp\left\{-c_j^2 / \left(2\sigma^2\right)\right\}$$
$$- exp\left\{-c_{j-1}^2 / \left(2\sigma^2\right)\right\}\Big]$$
$$/ f\left(\Delta FF_t^T\right)$$

where λ_0 is always restricted to make

$$\overline{FF^{T*}} = \overline{FF^T} =$$
$$5.91,\ c_j > c_{j-1},$$
$$j = 1, \ldots, 4, c_0 = -\infty,$$

and $c_5 = \infty$.

The variance parameter, σ^2, was assigned a value and was not estimated, since we cannot identify jointly the threshold constants c and the variance parameter. The value of σ^2 was set at 0.12, which implies a standard deviation of the disturbance to the desired fed funds rate of about 35 basis points per month. If a much larger variance were chosen, the threshold constants would increase in absolute value, but some experimentation showed that estimates of the threshold coefficients are fairly stable across a range of plausible values of the variance.

Table 3

Autoregressive Ordered Probit Model of Target Fed Funds Rate Monthly Sample: 1/85 — 12/98

Coef cient (by variable)	Value	Category mean (from Table 1)	Std. Error
Lagged EFF^{T*}	0.906		0.032
Inflation	0.152		0.063
Output gap	0.102		0.029
Lagged $E\Delta FF^{T*}$	0.118		0.067
Threshold c_1	–0.799	–0.523	0.186
Threshold c_2	–0.439	–0.248	0.191
Threshold c_3	0.459	0.244	0.191
Threshold c_4	0.658	0.538	0.188

Log - Likelihood Value = –150.4

ORDERED PROBIT ESTIMATION RESULTS

Monthly data also were used in the ordered probit estimation, with the same caveat mentioned earlier about the measures used for the output gap. Quarterly changes in the target funds rate would not have fit cleanly into the 25-basis-point categories illustrated for the monthly data in Table 1. Table 3 presents the parameter estimates for equation 4. Interestingly, the discrete-variable model also shows evidence of Taylor's (1998) criterion for an inflation-fighting monetary policy; that is, the long-run response of the latent target funds rate (FF^{T*}) to a unit increase in inflation is greater than one: $\lambda_1 / (1 - \rho) = 1.62$. In particular, the estimates suggest that it takes about nine months for the latent target funds rate to increase by more than one percentage point in response to a one-percentage-point increase in the inflation rate. Thus, from a discrete-variable model of the target funds rate, we have recovered parameter estimates that have the Taylor-rule property—a greater than one-to-one response of the interest rate instrument to a unit increase in inflation. This finding adds corroborative evidence to the simple regression results from Table 2. The long-run response of the latent target funds rate to a percentage point increase in the output gap also closely matches the estimates from the simple regresssions: $\lambda_2 / (1 - \rho) = 1.09$.

In addition, the discrete-variable model provides information on the thresholds at which the FOMC tends to move the target funds rate. To induce either an increase or a decrease in the target of 25 basis points, the latent target funds rate, FF^{T*}, must be about 45 basis points above or below the actual target, according to the point estimates. In other words, the gap between the latent desired level of the target fed funds rate and the actual target level must be considerably larger than 25 basis points to induce a 25 basis-point change.

The symmetry of the threshold levels across increases and decreases suggests that the FOMC does not require more pressure to raise the target funds rate than to lower it. That is, we cannot come close to rejecting the hypothesis that $c_2 + c_3 = 0$. Less obvious is the failure to reject symmetry across the thresholds for large target changes: $c_1 + c_4 = 0$. The standard error on the sum is 0.36, so we cannot reject symmetry for large changes either. The thresholds are significantly greater in absolute value than the 25 and 50 basis-point levels corresponding with the category means. We can easily reject the joint hypothesis that $c_3 = -c_2 = 0.25$, and $c_4 = -c_1 = 0.50$, with a Wald test statistic that has a probability value well below 0.001. Similarly, the probability value of the Wald test that $c_3 = -c_2 = 0.25$ is almost equally low. Thus, the threshold coefficients reveal significant evidence that the discrete nature of the target federal funds rate adds a dimension of sluggishness to monetary policy responses to inflation and output gaps that goes beyond the partial adjustment specification of equations 2 and 4.

SUMMARY AND CONCLUSIONS

Previous studies have shown that movements in the federal funds rate can be described by a Taylor-rule equation with interest-rate smoothing via a partial adjustment mechanism. The contribution of this article is to examine the smoothing or policy inertia within a model of discrete target funds rate changes. To do so, I estimate the thresholds that govern the relationship between an underlying partial-adjustment model of the latent or desired target level and the discrete changes observed in practice. The estimates show a substantial degree of sluggishness in the discrete responses of the target level of the funds rate to gaps between the underlying desired target level and the actual target level. Nevertheless, the estimated threshold coefficients are less than twice the size of their corresponding discrete changes, so that there is a limit to the size of any misalignment of the target funds rate brought by its discrete nature. Furthermore, the estimates of the parameters that govern the latent desired level of the target funds rate display the long-run responsiveness property that Taylor (1998) associates with a sound inflation-fighting policy rule.

In addition, one area where the estimates of threshold coefficients would be useful is the federal funds futures market. Market participants would not want to forecast the fed funds target rate changes using ordinary least-squares estimates, given that discrete target changes face substantial threshold hurdles. With estimates of the thresholds, participants in the fed funds futures market can gauge more accurately the likelihood of a change in the target level in the near term, and thereby more accurately forecast the future monthly average for the federal funds rate.

REFERENCES

Goodfriend, Marvin. "Interest Rate Policy and the Inflation Scare Problem: 1979-1992" Federal Reserve Bank of Richmond *Economic Quarterly* (Winter 1993), pp. 1-24.

______. "Monetary Policy Comes of Age: A 20th Century Odyssey," Federal Reserve Bank of Richmond *Economic Quarterly* (Winter 1997), pp. 1-22.

Huh, Chan G., and Kevin J. Lansing. "Federal Reserve Credibility and Inflation Scares," Federal Reserve Bank of San Francisco *Economic Review* (1998), pp. 3-16.

Judd, John P., and Glenn D. Rudebusch. "Taylor's Rule and the Fed: 1970-1997," Federal Reserve Bank of San Francisco *Economic Review* (1998), pp. 3-16.

Orphanides, Athanasios. "Monetary Policy Evaluation with Noisy Information," Federal Reserve Board of Governors FEDS Working Paper 98-50 (November 1998).

Rudebusch, Glenn D. "Federal Reserve Interest Rate Targeting, Rational Expectations and the Term Structure." *Journal of Monetary Economics,* (April 1995) pp. 245-74.

______. "Is the Fed Too Timid? Monetary Policy in an Uncertain World," Federal Reserve Bank of San Francisco Working Paper (March 1999).

Sack, Brian. "Does the Fed Act Gradually? A VAR Analysis," Federal Reserve Board of Governors FEDS Working Paper 98-17 (March 1998).

Taylor, John B. "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (December 1993), pp. 195-214.

______. "Monetary Policy and the Long Boom," this *Review* (November/December 1998), pp. 3-11.

Woodford, Michael. "Optimal Monetary Policy Inertia," 1999 manuscript, Princeton University.

Michael R. Pakko is a senior economist at the Federal Reserve Bank of St. Louis. Gilberto Espinoza, Daniel Steiner, and Stephen Majesky provided research assistance.

The U.S. Trade De cit and the "New Economy"

Michael R. Pakko

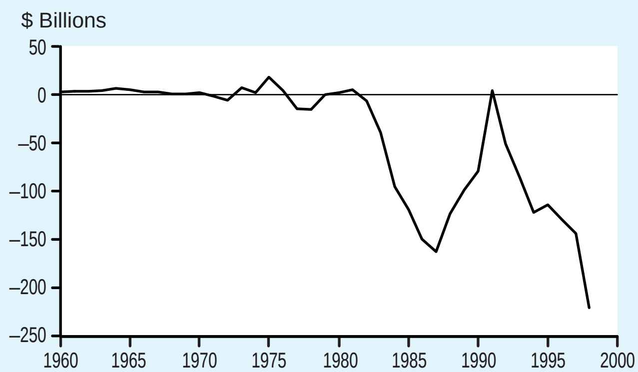
The performance of the U.S. economy during the 1990s has been universally hailed as stellar. Economic growth has been strong, unemployment has reached its lowest rate in over a generation, and inflation has remained relatively low. Consumer confidence has been high, helping to maintain strong growth in consumption expenditures, and investment spending has experienced a sustained growth rate that is unparalleled during the second half of the twentieth century. Many have gone so far as to declare that current conditions and prospects for the future represent a "new economy" or "new paradigm" in which these favorable trends might continue indefinitely.

One economic indicator that often is viewed with alarm, however, is the nation's growing trade deficit. In 1998, the U.S. trade deficit reached a record level, and when final data for 1999 is available, it is projected to be even higher. Each new release of trade data prompts the financial press to trumpet headlines announcing new record deficits. In both the media and popular opinion, trade deficits often are portrayed negatively, being blamed on the unfair trading practices of our trading partners or on a lack of U.S. competitiveness in world markets. Trade deficits often are attributed with reducing economic growth or resulting in lost jobs, and they almost always are discussed using terms with negative connotation. (For example, a widening deficit is frequently described as a "deterioration.")

Figure 1 illustrates recent movements in the most comprehensive measure of the

Figure 1

Current Account Balance
Annual Data



U.S. international trade position, the current account. Simple logic suggests that the downward trend established during the 1990s cannot be maintained indefinitely—if it were to do so, the United States would ultimately exceed its ability to pay for the rising tide of imports.

Nevertheless, few economists consider such a disastrous scenario likely. Long before the trade deficit could overwhelm the economy, interest rates, exchange rates, and relative national incomes would adjust to re-establish more balanced trade patterns. A key question that remains after acknowledging such market forces, however, is *how* such an adjustment ultimately will take place. If it is a smooth, gradual process, the favorable trends in productivity and incomes in the United States need not be interrupted significantly. If the adjustment were to be sharp and disruptive, however, the claims of proponents about the new economy would begin to ring hollow.

Understanding the underlying causes of the present U.S. trade deficit is an important part of evaluating their future impact on the economy. This article discusses the factors to consider in such an evaluation, focusing on a broad measure of

Figure 2

Components of the Current Account
Annual Data

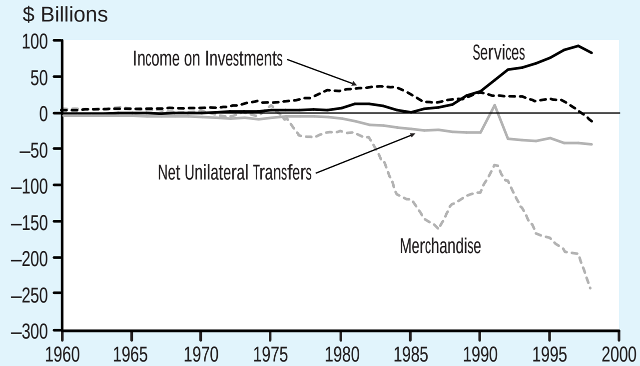


Table 1

Composition of the Current Account in 1998

(Billions of dollars)

Merchandise:		
Exports	670.2	
Imports	–917.2	
Balance		–246.9
Services:		
Exports	263.7	
Imports	–181.0	
Balance		82.7
Income on Investment:		
Inflows	258.3	
Outflows	–270.5	
Balance		–12.2
Net Unilateral Transfers:		–44.1
CURRENT ACCOUNT		–220.6

SOURCE: Department of Commerce, Bureau of Economic Analysis.
NOTE: Components may not sum to total due to rounding.

the U.S. trade position: its current account. The analysis suggests that recent trade deficits are driven by the same market forces that are otherwise manifested in the booming economy of the new paradigm theories.

If the present trade deficit is a temporary (albeit protracted) outcome of the adjustment to a new, higher long-run growth path for the economy, then we should not consider it to be a pressing concern. If it is driven by unsustainable, perhaps speculative imbalances, however, the deficit might forebode the ultimate demise of the longest economic expansion in U.S. history. The conditions under which the present trade deficit ultimately will be reversed might therefore be considered an important indicator of whether the new economy has entered a new, more mature phase, or whether its promises were illusory.

UNDERSTANDING THE CURRENT ACCOUNT

Components of the Current Account

The most comprehensive measure of the U.S. trade position is the *current account*, which is comprised of four categories. A listing of these categories and their magnitudes in 1998 is shown in Table 1, and the current account's recent behavior over time is illustrated in Figure 2.

The largest component, and the one that accounts for nearly the entire deficit, is merchandise. This component also is the most variable, accounting for most of the fluctuations in the current account over time.

In contrast to the deficit in merchandise trade, the United States has a stable surplus in services trade. This surplus has been growing consistently for more than two decades, but trade in services remains quite a bit smaller than merchandise trade. A commonly used measure of the trade deficit—often used loosely as interchangeable with the current account concept—is the combined merchandise and services accounts.[1]

The third category of the current account is income on investments. As foreign residents have accumulated U.S. assets over time, rising debt-service payments reflected in this category have

[1] This measure has the advantage of being available monthly, while other components of the current account are calculated only on a quarterly basis.

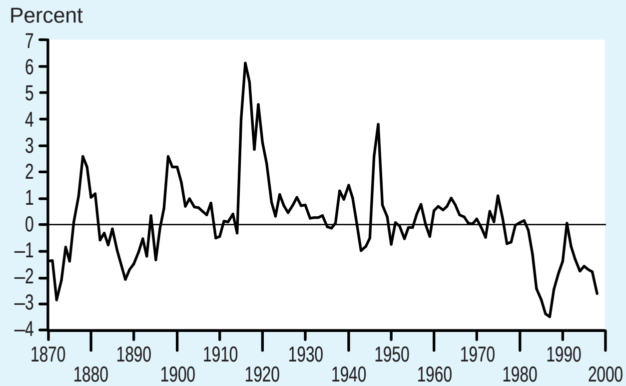
grown. Before 1998, the income-on-investments component represented a net inflow of payments. Since then, this component has reflected a net outflow as interest payments on foreign investment in the United States have risen above payments of interest on U.S. investments overseas. Nevertheless, the magnitude and variability of this component contribute little to the behavior of the overall current account.

The final category is net unilateral transfers. Because the United States is a major donor country in economic aid, this category is consistently in a deficit position.[2] As in the case of the investment-income component, net unilateral transfers are fairly small, and contribute little to the magnitude or variability of the overall current account.

Figure 3

Current Account as a Percent of GNP, 1870-1998

Annual Data



Historical series is from *International Historical Statistics, The Americas 1750-1993,* 4th Edition by B.R. Mitchell.

Putting the Deficit in Perspective

As shown in Figure 1, the current account appears quite volatile over the past decade or so, and recently has approached unprecedented levels. It is more meaningful, however, to gauge the magnitude of the current account deficit against the size of the total economy. In a growing economy, it is perfectly natural for the absolute magnitude of trade flows to be increasing over time. Hence, when we look at the current account deficit relative to the total production or income in the U.S. economy, the recent decline in our net export position —while still large—is not entirely unprecedented. To illustrate this point, Figure 3 shows the U.S. current account as a fraction of gross national product (GNP)—a broad measure of total economic activity. Although the current account deficit was a record $221 billion in 1998, this figure represented only 2.6 percent of GNP. In relative terms, the peak deficit of the 1980s was larger, reaching 3.5 percent of GNP in 1987.

Figure 3 also adds a longer historical perspective to the analysis.[3] It shows that even though the relative magnitude of current account fluctuations in the 1980s and 1990s is greater than during the 1960s and 1970s, swings in the U.S. current account balance in recent years are not quite as exceptional in the context of the past century or more.

The Determinants of Deficits

The recent steep decline in the U.S. trade position is significant, however, and perhaps not something we should simply dismiss. Deficits often are cited as either a cause or a symptom of economic weakness. The underlying implication of such a position is that selling is good, while buying is bad. When stated this starkly, the assumption loses much of its intuitive appeal.

In truth, deficits are neither causes nor symptoms of weakness, but are among the many macroeconomic quantities that are determined jointly by the decisions and interactions of households, firms, and governments in the United States and abroad. In the short-run, the current account can be affected by exchange rate fluctuations—which alter relative prices of imports and exports—and by differences in income growth at home and abroad. In fact, one of the fundamental forces behind the recent widening of our trade deficit has been the strength of the recent U.S. expansion relative to the growth rates of our major trading partners. As U.S. income growth outpaces growth abroad, our demand for both

2 A notable exception was in 1991, when payments to the United States from allies in the Persian Gulf war resulted in a net inflow in the unilateral transfers account.

3 Data for years prior to 1960 are from Mitchell (1998). GNP (rather than GDP) is used as the measure of aggregate economic activity to maintain consistency with the historical data.

THE SIMPLE ALGEBRA OF SAVING, INVESTMENT, AND THE CURRENT ACCOUNT

To derive the fundamental relationship among saving, investment, and the current account, one must begin with the national income accounting identity, which states that the total quantity of goods produced domestically (Y_{prod}) and imported (M) are used for consumption (C), investment (I), purchased by the government (G) or exported (X):

(1) $Y_{prod} + M = C + I + G + X,$

or, in a more conventional form,

(1') $Y_{prod} = C + I + G + NX,$

where $NX = X – M$ represents net exports.

Household income (Y_{inc}) is used to purchase consumption goods, save for the future (S), and pay taxes (T):

(2) $Y_{inc} = C + S^P + T,$

where the superscript P designates private saving. The government also saves (or dissaves) to the extent that tax revenues exceed (or fall short of) government spending:

(3) $S^G = T – G.$

Noting that every transaction involves a matched sale and purchase, aggregate equilibrium requires that the total value of goods produced is equal to the total value of income $Y_{prod} = Y_{inc}$. Using this equilibrium condition, and substituting the definitional relationships (2) and (3) into (1) yields the savings/investment/current account nexus:

(4) $NX = (S^P + S^G) – I.$

Consequently, a trade surplus is associated with an excess of saving over investment, while a trade deficit occurs when saving falls short of investment.

domestic goods *and* imports rise, while foreign demand for our exports languishes. This is one sense in which a current account deficit *reflects* underlying strength in the U.S. economy.

Even more fundamentally, the current account or net export balance reflects the outcome of the collective saving and investment decisions in an economy. (See shaded insert.) The relationship can be summarized as

(1) Net Exports = National Saving – Investment,

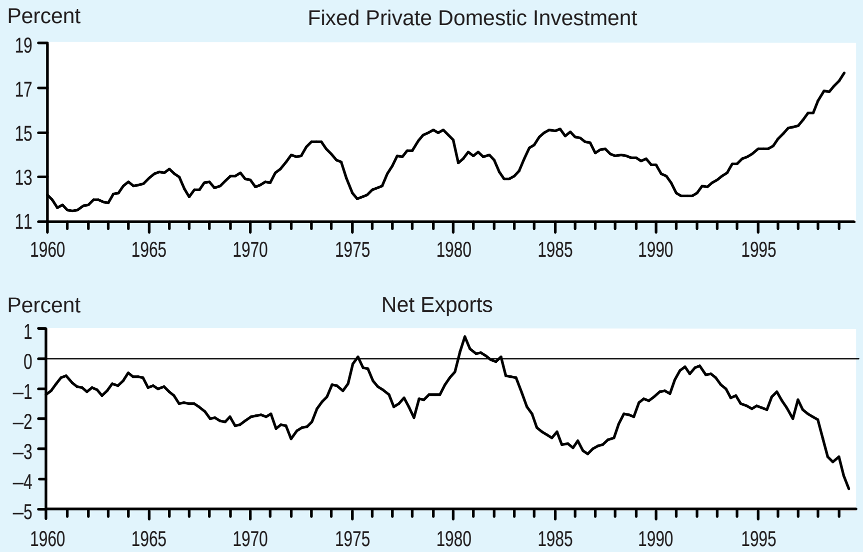
where the relevant measure of national saving includes both private sector saving and government saving (which is positive for government surpluses, negative for deficits), and

(2) National Savings = Private Saving – Government Deficits.

To understand this relationship more intuitively, note that a trade deficit reflects an excess of purchases over sales. Just as is the case for a household or a business firm that has current expenses exceeding current income, the difference must be financed through borrowing. Whether or not this borrowing is wise depends on what is being purchased. For example, a household that is continually running up credit card debt to finance current consumption, or a firm that is accumulating debt to cover operating losses, might well be following an unwise and unsustainable practice. On the other hand, when borrowing is undertaken to finance investments that will yield a flow of profits or services over

Figure 4

Fixed Investment and Net Exports as a Percent of GDP
Quarterly Data



time, it is a perfectly sound policy. The question of whether our national current account deficit is good or bad similarly hinges on the questions of why we are borrowing from the rest of the world, and what we are doing with the resources we are borrowing.

The relationship expressed in equation 1 represents a complex interaction of households, firms, and governments both at home and abroad. As such, it can be misleading to think of a clear, consistent causal relationship among the various components of the equation. Rather, it should be interpreted as summarizing an accounting identity that must hold in the context of all the components of the equation being affected simultaneously by overall economic conditions.

Sometimes it is useful to identify possible paths of causality within the overall relationship, however. For instance, much was made during the 1980s of the "twin deficits" of the United States—a combination of government deficits and current account deficits. Although it is beyond the scope of this article to discuss the issues involved in evaluating the claim that government budget deficits caused or contributed to the trade deficits of the 1980s, equations 1 and 2 demonstrate the plausibility of such a relationship. The basic current account/savings/investment relationship in equation 1 also demonstrates the conditions under which such an hypothesis would hold—namely that private savings and investment do not adjust to offset the effect of the government deficit.

Regardless of whether or not the relatively large current account deficits of the 1980s were related to government budget deficits, that surely cannot be the explanation for the 1990s experience. The U.S. government budget has been in a surplus position since 1997.

The low savings rate of U.S. households often is pointed to as one factor contributing to our negative current account and net export position. Generally speaking, the savings rate does not fluctuate markedly enough to be a key determinant of fluctuations in the current account. The decline of the personal savings rate throughout the 1990s—falling from over 5 percent at the beginning of the decade to nearly zero in 1998—has been a factor contributing to the widening current account deficit. I will suggest below, however, that this

Figure 5

Savings and Investment in a Large Open Economy

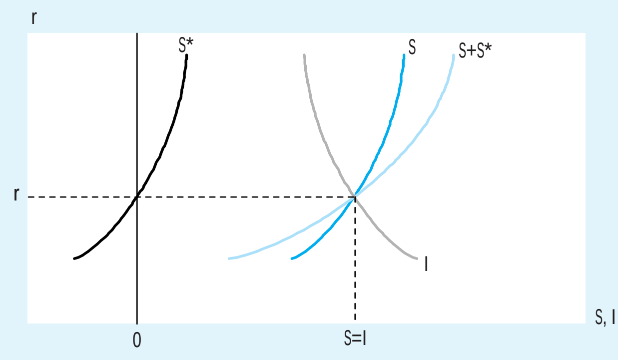
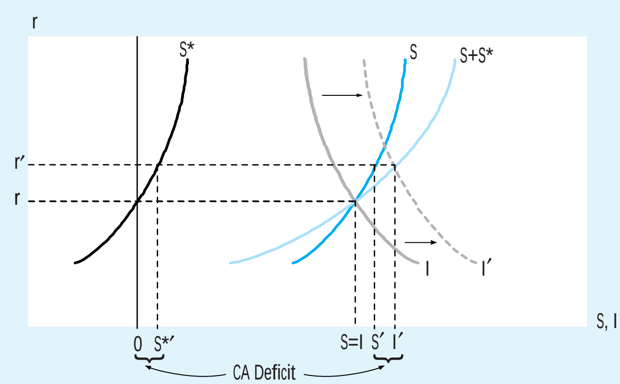


Figure 6

Investment Demand and the Current Account



decline reflects the very same underlying forces driving the deficit, rather than being a root cause.

Investment has been remarkably strong during the current expansion. In 1996, fixed private domestic investment as a percent of (gross domestic product) GDP matched its previous peak of 15.0 percent. In 1997 and 1998, investment spending rose to new record highs of 15.7 percent and 16.8 percent of GDP. Generally speaking, large swings in investment correspond to commensurate movements in the trade deficit, as illustrated in Figure 4.

Although the relationship is not exact, there is a clear tendency for large upswings in investment to be associated with widening trade deficits, and for troughs to be associated with surpluses, or at least smaller deficits. This is particularly true for the 1990s: As investment spending as a percent of GDP has surged through most of 1996-98, the U.S. trade deficit has expanded in tandem.

The Analytics of Investment, Savings, and Deficits

Figures 5 and 6 illustrate a diagrammatic representation of the savings/investment/current account relationship. Figure 5 shows an economy in the situation of balanced trade. Domestic savings, *S*, and the supply of potentially available foreign savings, *S**, both are represented as increasing in response to a higher domestic interest rate, *r* (relative to interest rates abroad). Investment demand, *I*, declines when domestic interest rates rise because the interest rate is the key cost of financing investment spending. When the supply of domestic savings matches investment demand, as is the case in Figure 5, no net inflow or outflow of foreign savings is necessary ($S^* = 0$) and the current account is balanced.

In Figure 6, the demand for investment in new capital goods has increased. Assuming for simplicity that the position of the supply curves for domestic and foreign savings are unaffected, the quantities of both domestic and foreign savings rise (to S' and $S^{*\prime}$) in response to the upward pressure on the interest rate. Investment spending rises and the country experiences a current account deficit equal to the shortfall of domestic savings relative to investment, $I' - S'$.

Note that if the opportunity to draw on the pool of foreign savings was not available in this instance, the increase in investment demand could only be financed by domestic savings (the intersection between the *S* and *I* curves in Figure 6). If this were the case, the interest rate would have to rise further in order to establish the balance between savings and investment, limiting the amount of investment spending that would actually take place.

DEFICITS AND THE NEW ECONOMY

With this analysis in mind, what can we say about the relationship between the widening trade deficit of the 1990s and the new economy?

What is the New Economy?

Two, or sometimes three, factors generally are cited as the underlying positive performance of the economy during the 1990s.[4] First is the adoption of new technologies in information processing and telecommunications. A second factor is the world-wide commodity glut that has caused sharp declines in the prices of some key U.S. imports—particularly oil. An additional factor often cited is the competitive effects of globalization.

The globalization argument suggests that competition from abroad has forced U.S. firms to keep costs down and prompted workers to scale back expected wage increases. While it is undoubtedly true that foreign competition is an important consideration for many firms, there is little empirical evidence that foreign competition significantly affects wages or aggregate income in the United States.[5] Even though trade has taken on increased prominence in the composition of economic activity in the United States, it still comprises only a small share of total GDP. In 1998, exports and imports represented only 13 percent and 16 percent of GDP, respectively. For all practical purposes, trends in the United States depend on domestic factors, with influences from abroad taking on a distinctively secondary role.[6]

It is the main argument of this article that the U.S. trade position, in relation to the rest of the world *reflects* the underlying determinants of recent economic performance, rather than being a fundamental cause in and of itself.

The other two factors underlying the new economy—low oil prices and technological advances—fall in the general category of supply shocks.[7] Both would be expected to raise economic output and to increase the demand for domestic labor and capital. To the extent that investment spending rises to meet higher demand for physical capital, this type of supply shock also tends to give rise to a current account deficit—as described in the previous section.

The two types of supply shocks differ in key respects, however. A decline in the price of an imported factor of production, like oil, often is reversed—witness the increases in oil prices in early 1999. Investment demand will increase only to the extent that productive capacity can be brought online and maintained to exploit the favorable cost environment. In and of itself, a temporary oil price decline is more likely to increase capital utilization, rather than capital investment. Moreover, a temporary increase in economic activity would be expected to raise aggregate savings, as households set aside some of the windfall income gains for the future. Hence, domestic savings and investment demand both rise, and there is little, if any, pressure for widening trade and current account deficits.

The adaptation of new technologies to the production process, on the other hand, is more likely to be associated with a sustained increase in investment as new equipment replaces "vintage" capital.[8] To the extent that capital productivity is expected to rise permanently (or at least for an extended period of time), the rise in investment demand will be larger than it would be for a temporary supply shock. At the same time, domestic savings might rise very little—and might even fall, which has happened during the 1990s. This is because there is less incentive to set aside a portion of the increase in income for the future, with the future looking so bright. Consequently, longer-lasting supply shocks would be expected to induce larger current account movements than more short-term shocks.

Assessing Recent Deficits

The decline in U.S. net exports as a fraction of national income eventually will be reversed, either as the investment boom runs its course or as foreigners become increasingly unwilling to finance mounting

4 In his public speeches, Federal Reserve Board Chairman, Alan Greenspan (1997, 1999) has described the factors underlying current strong economic conditions.

5 Burtless (1995) provides a survey of economic research on the effect of trade on U.S. wages and incomes.

6 Krugman (1996a,b) forcefully and persuasively argues this contention about the quantitative irrelevance of globalization, which he calls "globaloney."

7 These types of supply shocks figure prominently in so-called real business cycle theories.

8 The notion that technological progress is embedded in new capital replacing vintage capital is explored by Jorgenson (1966) and more recently by Greenwood, et al (1997). Greenwood et al find that as much as 60 percent of postwar productivity growth can be attributed to this type of investment-specific technological change.

Figure 7

Capital Goods Imports as a Percent of Total Merchandise Imports

Quarterly Data

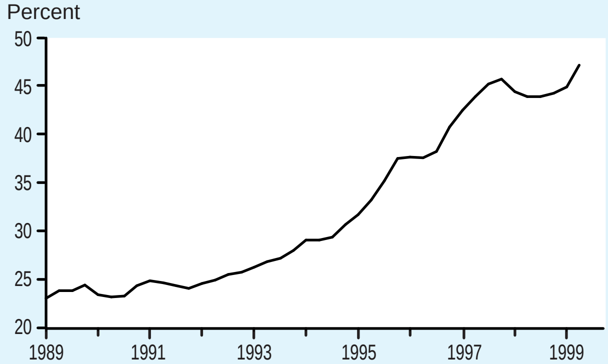
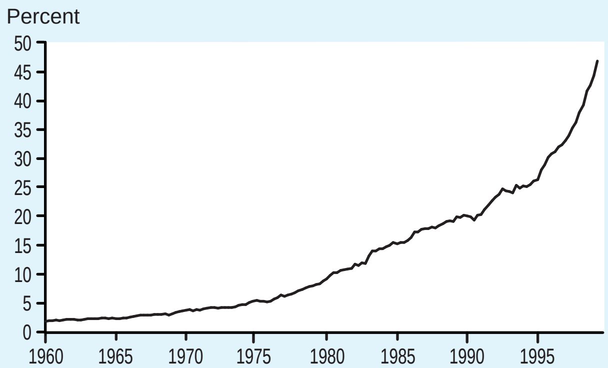


Figure 8

Information Processing Equipment as a Percent of Total Fixed Investment

Quarterly Data



deficits and the debt. The question of whether the reversal is likely to take place as an orderly adjustment, or as a "crash-and-burn" scenario, is crucial for evaluating the prospects for continued economic strength suggested by advocates of the new economy view. Which of these scenarios is more likely depends, in turn, on the factors underlying the burgeoning trade deficit. The key question to ask is: "What are we doing with the resources that we're borrowing from the rest of the world?"

Recalling the analogy to individual households or businesses, we maintain that borrowing to finance frivolous consumption is a recipe for disaster, while borrowing to invest in assets that will pay off in future flows of goods and services is more likely to be a prudent course.

Figure 7 considers one dimension of the question: "What are we doing with the borrowing?" Looking literally at the composition of imports, we see that the rise in the capital goods imports as a share of total imports during the 1990s has been remarkable, rising from around 25 percent at the beginning of the decade to more than 44 percent in 1997 and 1998. Figure 8 illustrates an important feature of the composition of total investment during the 1990s. The share of investment in information processing equipment and technologies as a fraction of total investment has been rising steadily since the mid-1970s, but has increased dramatically during the latter half of the 1990s.

Moreover, the United States is taking the world lead in investing in cutting-edge technologies. In 1997, for example, spending on information technology accounted for a full 4.5 percent of U.S. GDP, compared to only 2.6 percent in Japan and 2.3 percent in Western Europe (Koretz, 1999).

These measures suggest that unprecedented rates of investment in the latter half of the 1990s are associated with the widespread and rapid adoption of new technologies. To the extent that these investments do, in fact, pay off in future higher productivity and output growth, undoubtedly we will look back on this period as setting the stage for what truly will be a new economy.

Until very recently, there has been little indication that the adoption of new technologies has resulted in any significant gains in productivity. In fact, the early stages of the 1990s' economic expansion were characterized by very slow productivity and employment growth by historical standards. This is consistent with economic models of technological advancement, however.[9] During the early stages of technological breakthroughs, like those we are witnessing in information processing and telecommunications, a period of slow growth is predicted as new technologies are integrated and adapted to production

[9] Theoretical analyses of the effects of breakthroughs in such general purpose technologies are explored in depth in Aghion and Howitt (1998).

processes. Only after this transition phase does productivity rise. The relatively high and rising rates of productivity for the United States during the latter part of 1998 and early 1999 might forebode the beginning of long-awaited productivity gains.

Allen (1997) assesses various theories explaining the lack of obvious productivity gains over the course of the 1990s. He cites as an historical precedent the work of David (1990), who compared the modern information technology revolution to the invention of the electric dynamo in the nineteenth century. David suggested that fully exploiting the new technology represented by the dynamo took decades. In the meantime, its effect on productivity lagged its ultimate potential.

If the analogy holds true, it is not surprising that productivity growth has not yet accelerated as much as one might think with the adoption of new technology. The ultimate benefits of adopting new technologies only will become apparent over the course of years to come. Once the initial surge in investment demand subsides, we would expect the deterioration in the U.S. current account to show signs of reversal, suggesting that the economic expansion associated with this transition has reached a more mature stage.

CONCLUSION

This article has described the basic determinants of the current account, challenging the common, but simple notion that trade deficits are inherently bad. In fact, deficits are neither good nor bad: Rather, they are reflections of the more fundamental underlying forces affecting the economy.

In the context of the U.S. economy during the 1990s, rising trade and current account deficits are consistent with the notion that strong investment spending is associated with the adoption of new technologies, with the anticipation of rapid economic growth in the future suppressing domestic saving. The resulting weakness of the U.S. current account balance is, therefore, a reflection of an economy that is strong, but in transition. A turnaround of the deficit is likely to be an important indicator of when that transition is complete. Only after we reach this more mature phase of the current economic expansion will we be able to fully evaluate the claims of those who suggest that we are on the threshold of a new economy in which rising rates of productivity and economic growth will last far into the future.

REFERENCES

Aghion, Philippe, and Peter Howitt. *Endogenous Growth Theory*, MIT Press, 1998.

Allen, Donald S. "Where's the Productivity Growth (from the Information Technology Revolution)?" this *Review* (March/April 1997), pp. 15-25.

Burtless, Gary. "International Trade and the Rise in Earnings Inequality," *Journal of Economic Literature* (June 1995), pp. 800-16.

David, Paul. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *The American Economic Review* (May 1990), pp. 355-61.

Greenspan, Alan. Remarks at the Annual Convention of the American Bankers Association, Boston, Massachusetts, October 5, 1997.

________. "The American Economy in a World Context," Remarks at the 35th Annual Conference on Bank Structure and Competition of the Federal Reserve Bank of Chicago, Chicago, Illinois, May 6, 1999.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krussell. "Long-Run Implications of Investment-Specific Technological Change," *The American Economic Review* (June 1997), pp. 342-62.

Jorgenson, Dale W. "The Embodiment Hypothesis," *Journal of Political Economy* (February 1966), pp. 1-17.

Koretz, Gene. "Info Tech: Who's Spending What," *Business Week* (May 3, 1999), p. 27.

Krugman, Paul R. *Pop Internationalism*, MIT Press, 1996a.

________. "The Adam Smith Address: What Difference Does Globalization Make?" *Business Economics* (January 1996b), pp. 7-10.

Mitchell, B.R. *International Historical Statistics: The Americas 1750-1993*, Stockton Press, 1998.

Donald S. Allen is an economist for the Federal Reserve Bank of St. Louis. Thomas A. Pollmann provided research assistance.

Seasonal Production Smoothing

Donald S. Allen

Inventory investment dynamics appear to dominate the economy's historical movement around its long-run path. Blinder and Maccini (1991) show that the average movement in inventory investment during post-war recessions account for 87 percent of the peak-to-trough movement in Gross National Product (GNP).[1] Because inventory fluctuations have historically played such a major role in business cycles (and possibly seasonal fluctuations), it is important to understand the theoretical motivation for inventory holdings and the implied dynamics. The received view, established by Holt, Modigliani, Muth and Simon (1960), is that inventories are used to smooth production in the presence of increasing marginal cost (convex costs). An empirical fact of inventory investment, however, is that production is more volatile than sales. The failure to confirm production smoothing empirically has been explained by inadequacies of the data or exceptions to the abstraction of convex costs.

Intuition suggests that even with convex costs, firms may not be likely to smooth production over periods longer than a year. Production horizons are likely to be shorter than a year and inventory holding costs may make it uneconomical to hold inventory for as long as a year. Many industries have well documented seasonal patterns in demand allowing them to plan production in concert with available capacity, required lead times, and labor market flexibility. In addition, evidence has been uncovered suggesting that seasonal fluctuations in output also can be affected by inventory/production decisions. For example, Carpenter and Levy (1998) use frequency domain analysis and find a large and statistically significant average squared coherence between inventory investment and the change in output in the manufacturing sector at both seasonal and business-cycle frequencies. It seems appropriate, therefore, to focus some attention on inventory decisions at seasonal frequencies.

In this paper we look for evidence of seasonal production smoothing in seasonally unadjusted monthly data on manufacturing and retail inventories and sales. Using detrended, seasonally unadjusted data we find that the variance of production is less than the variance of sales for 23 out of 35 industries. The equivalent test using seasonally adjusted data found none with production varying less than sales. We interpret this as stronger evidence of production smoothing than found in previous literature.

The Fourier series of the inventory-to-sales (I/S) ratio of the industries with the lowest variance of production relative to sales revealed strong seasonal components (annual, six months, and three months). A strong seasonal component in the I/S ratio suggests a possible negative seasonal correlation between sales and inventory and is an intuitive indication that smoothing occurs at higher frequencies.[2] The results confirm Ghali's (1987) finding of seasonal smoothing using detrended, seasonally unadjusted data for the cement industry. The results also suggest that a model other than production smoothing may be more appropriate for explaining trend movements in production relative to sales.

Background and Literature Survey

Holt, Modigliani, Muth, and Simon (1960) established the analytical framework demonstrating that optimizing firms facing convex production costs and uncertain demand are motivated to smooth production and

[1] There is some debate whether more recent practices still induce this effect. Allen (1995) suggests that recent improvements in inventory management may be reducing the "boom/bust" effects of inventory swings, but empirical confirmation is not strong. Filardo (1995) finds no evidence of a muted inventory cycle.

[2] If inventories are high when sales are low, and vice versa, then the I/S ratio will fluctuate accordingly. Obviously, if inventory remained constant and sales varied seasonally, the I/S ratio also would fluctuate seasonally, so this is not an exact metric of smoothing. The seasonality of the ratio does suggest, however, the frequency over which smoothing is taking place.

use inventories to buffer demand shocks. If the marginal cost of production is increasing, then storing output during periods of low demand is prudent as long as storage costs are sufficiently low. (See the shaded insert on page 5.) Much of the research in inventory since Holt et al., has focused on the efficacy of using the production-smoothing paradigm at the macroeconomic level. If firms use inventories to smooth production, then production should vary less than sales. Empirical testing of this hypothesis has yielded mixed results. Using a simple test of the ratio of the variance of production to the variance of sales, a majority of researchers have found a ratio greater than 1.0, contradicting the theory.

As a rule, data on production are not available. Production can be readily estimated, however, by adding current period sales to the change in inventory from last period. If production exceeds (is less than) sales in a given period, then the difference must go to increasing (decreasing) inventories. This can be represented by the following equation:

$$P_t = S_t + \Delta I_t ,$$

where P is production, S is sales and ΔI is the change in inventory. This fundamental equation implies an important relationship among the variances and covariance of P, S, and ΔI:

$$Var(P) = Var(S) + Var(\Delta I) + 2Cov(S, \Delta I)$$

For the variance of production to be less than the variance of sales, the covariance of sales and the change in inventories, $Cov(S, \Delta I)$ must be negative and greater in absolute value than half the variance of inventories.

Testing this covariance relationship, Miron and Zeldes (1988) find no support for production smoothing using both seasonally adjusted and unadjusted data after removing an estimated linear trend from monthly data. Blinder (1986) also finds little empirical support for the basic production smoothing model. He identifies, however, conditions under which the facts could be compatible with production smoothing, to wit, if:

- Cost shocks are present,
- Firms see demand shocks before they make their production decisions,
- Demand shocks build before they decay,
- Or, technological parameters dictate a rapid speed of adjustment.

If firms do not face convex production costs, production smoothing is not optimal. Ramey (1991) finds indication of nonconvex costs in some industries. Blinder and Maccini (1991) observe that wholesale and retail trade, and the materials and supplies portion of manufacturing inventory, make up a large portion of total inventories and are likely to face nontrivial "quasi-fixed" costs of ordering. This type of cost structure makes an (*S,s*) inventory rule more economical. That is, firms will wait until inventory falls below a trigger point (*s*) then order sufficient stocks to raise inventory to an upper bound (*S*). This way the quasi-fixed costs are spread over a larger quantity. This behavior, sometimes called "bunching" will result in a higher volatility of production than sales. This leads Blinder and Maccini to conclude that the (*S,s*) paradigm is more consistent with the empirical evidence.

Another source of empirical failure may be the data. Ghali (1987) demonstrated that seasonal adjustment and aggregation will remove evidence of seasonal smoothing, and Lai (1991) shows that aggregation can distort the data sufficiently to negate production-smoothing tests. Some researchers, using disaggregated physical product data, find some support for production smoothing. Fair (1989), and Krane and Braun (1991) use disaggregated physical product data for the United States and confirm smoothing in several industries, while Beason (1993) has similar success with Japanese data. Dimelis and Ghali (1994) detect statistically significant evidence of smoothing in disaggregated physical product data for three out of five industries, using the variance bounds test introduced by West (1986) and generalized by Kollintzas (1995).

Physical unit information is more appropriate for testing the implications of inventory management. It makes more sense, when discussing the motivation for holding inventory, to talk about the number of cars in stock than the value. Unfortunately, the most readily accessible data, particularly at an aggregate level, are the nominal values of inventory. One way of getting closer to physical quantities is to remove the effects of price changes. Finding the appropriate price deflator to convert the nominal values to real values is a difficult task. And, even when data are converted to remove price increases, trend growth in the real level of sales also can disguise smoothing. If sales are trending up (down), then production will also trend up (down). If firms smooth production annually and adjust the target level of smoothed production each year, then the variance induced by the trend growth also will distort the smoothing measure.

Miron (1996) page 18 finds noticeably less seasonal variation in price variables than quantity variables.

> Seasonal movements in both real and nominal price variables are noticeably smaller than those in quantity variables. For example, the standard deviation of the seasonals in the growth rates of prices is 0.2 percent, and seasonal dummies explain only 3.1 percent of the total variation. The same conclusions hold qualitatively for nominal interest rates, real interest rates, nominal wages, and real wages. Miron (1996).

This observation means that if we remove the trend from seasonally unadjusted data, the high frequency movements are more likely to reflect changes in quantities. This provides justification for the data transformation that we discuss in the next section.

DATA SOURCE AND TRANSFORMATION

The data used are from the Census Bureau's monthly data on manufacturing

Table 1

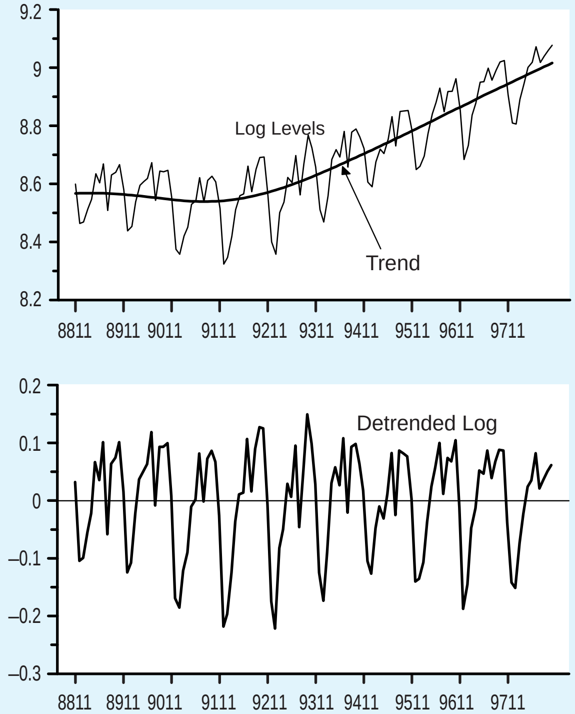
Industries Analyzed	Years of Data
All Manufacturing Industries	1/58 to 12/98
Stone, Clay and Glass (SIC 32)	1/58 to 12/98
Primary Metals (SIC 33)	1/58 to 12/98
Nonferrous and other Primary Metals	1/58 to 12/98
Fabricated Metal Products (SIC 34)	1/58 to 12/98
Industrial Machinery and Equipment (SIC 35)	1/58 to 12/98
Electrical Machinery (SIC 36)	1/58 to 12/98
Transportation Equipment (SIC 37)	1/58 to 12/98
Instruments/Related Products (SIC 38)	1/58 to 12/98
All Other Durable Goods	1/58 to 12/98
Nondurable Goods Manufacturing Industries	1/58 to 12/98
Tobacco Products (SIC 21)	1/58 to 12/98
Textile Mill Products (SIC 22)	1/58 to 12/98
Paper and Allied Products (SIC 26)	1/58 to 12/98
Chemical and Allied Products (SIC 28)	1/58 to 12/98
Petroleum and Coal Products (SIC 29)	1/58 to 12/98
Automotive Equipment	1/58 to 12/98
Home Goods and Apparel	1/58 to 12/98
Consumer Staples	1/58 to 12/98
Machinery and Equipment *	1/68 to 12/98
Business Supplies	1/58 to 12/98
Construction Materials/Supplies /Intermediate	1/58 to 12/98
Capital Goods Industries	1/58 to 12/98
Producers' Durable Equipment *	1/68 to 12/98
Household Durable Goods	1/58 to 12/98
All Retail	1/87 to 12/98
Retail: Durable Goods Stores	1/87 to 12/98
Retail: Bldg Matls/Hdwre/Garden Supply/ Mobile Home Dealers (SIC 52)	1/87 to 12/98
Retail: Automotive Dealers (SIC 55)	1/87 to 12/98
Retail: Furniture, Home Furnishings & Eqpt Stores (SIC 57)	1/87 to 12/98
Retail: Nondurable Goods Stores	1/87 to 12/98
Retail: General Merchandise Group Stores (SIC 53)	1/87 to 12/98
Retail: Department Stores ex Leased Departments (SIC 531)	1/87 to 12/98
Retail: Food Stores (SIC 54)	1/87 to 12/98
Retail: Apparel and Accessory Stores (SIC 56)	1/87 to 12/98

* Data available starting in 1968.

Figure 1

Stone, Clay and Glass Products

(Monthly Sales, Seasonally Unadjusted, Millions of Current Dollars)



and retail inventories and sales, seasonally unadjusted and adjusted. Production is computed by adding the change in inventories to sales each period. A total of 35 series, 25 manufacturing and 10 retail, were analyzed. Table 1 lists the series and the years of data used. Most manufacturing data covered the period between 1958 and 1998. Retail data covered the period between 1987 and 1998.

HP Filtering

After taking logs of the data, a nonlinear trend was removed using a Hodrick-Prescott (HP) filter with a penalty set to 14,400, the default value for monthly data in *Eviews*. This method removes the "low frequency" components from the data, whether due to price increase or trend growth. Figure 1 illustrates the transformation of the data for the most recent 10 years for the Stone, Clay and Glass manufacturing industry. The smooth line shows the trend that is extracted to get the filtered data.

Frequency Domain

Looking at the data in the frequency domain highlights the effect of the seasonal adjustment and the HP filter.[3] The Fourier series representation decomposes the data into the contribution of individual frequencies to the variance. The vertical axis indicates the value of the coefficient applied to that frequency. If there is a trend present, there will be a large contribution from the low frequencies. If there is a strong contribution at a particular frequency compared to others, there will be a noticeable spike at that frequency. The Fourier series displayed here are truncated at 1.0 to focus attention on the higher frequency components. Figure 2 compares the Fourier series of the logarithm of sales for Stone, Clay and Glass Industries for the raw seasonally unadjusted series, the seasonally adjusted series, and the series after an HP filtered trend is removed from the raw (unadjusted) data. The spikes in the unadjusted data occur at cycles of 12 months, six months, four months and 3 months, reflecting harmonics of the seasonal cycle. The appearance of harmonics in the data may reflect the aggregation of individual firms with seasonal cycles that are offset, (i.e. some may have peak sales in winter while others peak during the summer). The seasonally adjusted data have no spikes. The HP filtered data show the absence of the low frequency components (zero near the origin), while the high frequency contributions appear to be intact. Figure 3 shows the I/S ratio of selected industries, and Figure 4 shows the Fourier series of the ratios. In the next section, we report the results of the variance ratio test, then compare this to the frequency spectra of the I/S ratios of the sectors.

Results

The typical measure of production smoothing is the ratio of the variance of

[3] Harvey (1994), chapter 6, provides an excellent exposition of frequency domain analysis.

THE PRODUCTION SMOOTHING WITH TREND

Figure A

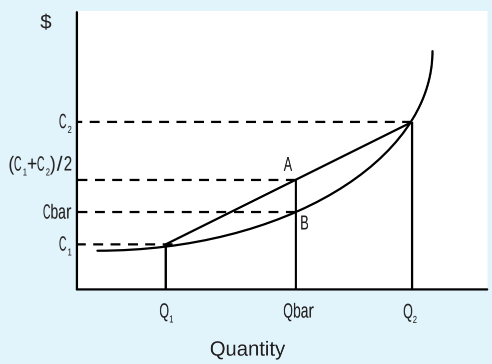


Figure B

Periodic Production Smoothing

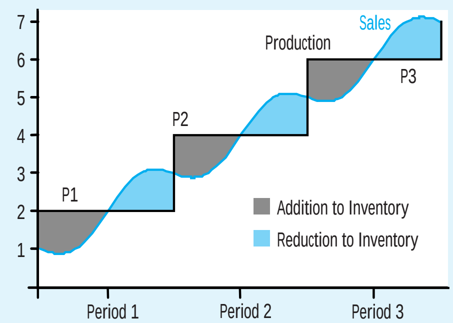


Figure A illustrates the production smoothing motivation when increasing marginal costs exist. If Q_1 and Q_2 represent the demand in periods 1 and 2 respectively, then the point A represents the average cost, $(C_1 + C_2)/2$, if Q_1 is produced during period 1 and Q_2 is produced during period 2. Point B represents the average cost, *Cbar*, if $(Q_1+Q_2)/2$ is produced during both periods, with the excess produced in period 1 carried over to period 2. The trade off is between the cost of storage for one period versus the saving from smoothing. The difference between A and B must be greater than the cost of holding inventory to justify smoothing. Note also that if mean demand is expected to decrease below current production for an extended period (i.e., Q_2 is current demand and Q_1 is next period's expected demand), then it becomes optimal to reduce production and serve part of the current demand from inventory. Thus production-smoothing motivation can lead to level changes if forecast sales change direction.

Figure B illustrates how periodic adjustments to production to match trend growth can result in lumpy movements in production even with production smoothing. The blue line indicates trend growth in sales with a seasonal component. If sales are forecast and production planned at the beginning of each period, then P1 represents the production level for the first period, P2 the second period, and P3 the third period. During the first half of each period, production exceeds sales and the difference goes into inventory. During the second half of the period, production is less than sales and the difference is made up out of inventory. Each period, production is smoothed. Because of the trend growth in sales, however, production jumps at the start of each new period. If data over all three periods are used, the variance of production may exceed the variance of sales.

production to the variance of sales.[4] A ratio more than 1.0 implies that production is more volatile than sales and, therefore, contradicts the smoothing hypothesis. A negative correlation between sales and the change in inventory may be insufficient to produce a lower variance in production than in sales. Tables 2, 3, and 4 summarize

[4] A more appropriate test of production smoothing would be a comparison of the variance of production to the variance of *forecasted* sales.

Table 2

Manufacturing

		Variance	Variance	Covariance	Variance	Variance Ratio
		Sales	Inventory Investment	Sales and Inven. Invest.	Production	
All Manufacturing Industries	NSA	0.00272	0.00029	–0.00015	0.00270	0.99554
	SA	0.00061	0.00014	0.00009	0.00092	1.50964
Stone, Clay and Glass (SIC 32)	NSA	0.00883	0.00076	–0.00176	0.00606	0.68635
	SA	0.00117	0.00023	0.00000	0.00141	1.20013
Primary Metals (SIC 33)	NSA	0.01106	0.00072	–0.00015	0.01148	1.03841
	SA	0.00731	0.00054	0.00035	0.00855	1.16922
Nonferrous and other Primary Metals	NSA	0.00729	0.00059	0.00010	0.00808	1.10774
	SA	0.00462	0.00037	0.00021	0.00541	1.17196
Fabricated Metal Products (SIC 34)	NSA	0.00399	0.00112	–0.00027	0.00456	1.14233
	SA	0.00104	0.00065	0.00011	0.00190	1.83141
Industrial Machinery and Equipment (SIC 35)	NSA	0.00721	0.00123	–0.00089	0.00667	0.92451
	SA	0.00133	0.00050	0.00029	0.00240	1.80271
Electrical Machinery (SIC 36)	NSA	0.00496	0.00086	–0.00020	0.00542	1.09311
	SA	0.00104	0.00034	0.00022	0.00182	1.75383
Transportation Equipment (SIC 37)	NSA	0.01422	0.00118	–0.00148	0.01244	0.87488
	SA	0.00388	0.00039	–0.00014	0.00399	1.02722
Instruments/Related Products (SIC 38)	NSA	0.00418	0.00121	–0.00074	0.00391	0.93591
	SA	0.00072	0.00058	0.00003	0.00137	1.89900
All Other Durable Goods	NSA	0.00634	0.00048	–0.00046	0.00590	0.92985
	SA	0.00155	0.00024	0.00012	0.00204	1.31248
Nondurable Goods Manufacturing Industries	NSA	0.00164	0.00019	–0.00006	0.00171	1.03902
	SA	0.00038	0.00011	0.00005	0.00058	1.52003
Tobacco Products (SIC 21)	NSA	0.01631	0.02009	–0.00220	0.03199	1.96182
	SA	0.00396	0.00469	–0.00005	0.00855	2.15891
Textile Mill Products (SIC 22)	NSA	0.00668	0.00099	–0.00073	0.00620	0.92822
	SA	0.00139	0.00025	0.00012	0.00188	1.35705
Paper and Allied Products (SIC 26)	NSA	0.00227	0.00023	0.00008	0.00265	1.16834
	SA	0.00109	0.00011	0.00014	0.00148	1.35517
Chemical and Allied Products (SIC 28)	NSA	0.00338	0.00038	–0.00036	0.00304	0.89854
	SA	0.00096	0.00019	0.00010	0.00134	1.40595
Petroleum and Coal Products (SIC 29)	NSA	0.00534	0.00080	0.00032	0.00678	1.27019
	SA	0.00423	0.00058	0.00036	0.00552	1.30557

Table 3

"Other" Manufacturing

		Variance	Variance	Covariance	Variance	Variance Ratio
		Sales	Inventory Investment	Sales and Inven. Invest.	Production	
Automotive Equipment	NSA	0.02848	0.00097	–0.00213	0.02518	0.88435
	SA	0.00898	0.00027	–0.00012	0.00901	1.00339
Home Goods and Apparel	NSA	0.00898	0.00138	–0.00195	0.00647	0.72049
	SA	0.00088	0.00024	0.00014	0.00140	1.58834
Consumer Staples	NSA	0.00133	0.00035	0.00004	0.00177	1.32663
	SA	0.00032	0.00014	0.00002	0.00049	1.52705
Machinery and Equipment	NSA	0.00722	0.00130	–0.00142	0.00567	0.78549
	SA	0.00096	0.00041	0.00012	0.00161	1.68163
Business Supplies	NSA	0.00183	0.00033	–0.00015	0.00186	1.01557
	SA	0.00051	0.00019	0.00004	0.00078	1.54868
Construction Materials/Supplies /Intermediate	NSA	0.00563	0.00041	–0.00052	0.00499	0.88617
	SA	0.00124	0.00020	0.00012	0.00167	1.35123
Capital Goods Industries	NSA	0.00694	0.00144	–0.00163	0.00513	0.73934
	SA	0.00080	0.00048	0.00010	0.00148	1.85593
Producers' Durable Equipment	NSA	0.00667	0.00083	–0.00066	0.00619	0.92711
	SA	0.00120	0.00033	0.00010	0.00173	1.43817
Household Durable Goods	NSA	0.00840	0.00149	–0.00133	0.00723	0.86042
	SA	0.00114	0.00047	0.00018	0.00196	1.72187

the results, showing the variance of sales, inventories and the covariance of sales and the change in inventories. Of the 35 seasonally unadjusted series, there are only three manufacturing industries with positive covariances between sales and the change in inventories: Nonferrous and Other Primary Metals (a sub-category of Primary Metals); Paper and Allied Products; and, the Petroleum and Coal Products. By contrast, the covariances of all but three manufacturing series with seasonally adjusted data are positive. The seasonally adjusted retail data indicate some with negative covariance of inventory investment and sales, but none sufficiently negative to result in a variance ratio less than 1.0.

Manufacturing

The variance ratio of the detrended seasonally unadjusted data for all manufacturing is less than 1.0, but only barely; leaving unanswered the question of whether the production-smoothing model is adequate at this level of aggregation. At the two-digit SIC level of aggregation, SIC codes 33, 34, and 36 have variance ratios greater than 1.0 for the detrended log data, while SIC codes 32, 35, 37, and 38, as well as the "all other durable goods" category have variance ratios less than 1.0. The implication is that most durable goods industries smooth production over high frequency periods.[5] The seasonally adjusted data do not show smoothing, indicating that

[5] This implication holds at least if we interpret production smoothing as meaning that the growth rate of production varies less than the growth rate of sales.

Table 4

Retail

		Variance	Variance	Covariance	Variance	Variance Ratio
		Sales	Inventory Investment	Sales and Inven. Invest.	Production	
All Retail	NSA	0.00630	0.00222	–0.00206	0.00439	0.69691
	SA	0.00011	0.00009	–0.00001	0.00018	1.68785
Retail: Durable Goods Stores	NSA	0.00618	0.00264	–0.00116	0.00650	1.05145
	SA	0.00046	0.00038	–0.00004	0.00077	1.67262
Retail: Bldg Matls/Hdwre/Garden Supply/ Mobile Home Dealers (SIC 52)	NSA	0.02100	0.00199	–0.00361	0.01576	0.75081
	SA	0.00065	0.00034	0.00003	0.00104	1.60598
Retail: Automotive Dealers (SIC 55)	NSA	0.00805	0.00402	–0.00241	0.00727	0.90217
	SA	0.00075	0.00087	–0.00016	0.00130	1.72889
Retail: Furniture, Home Furnishings & Eqpt Stores (SIC 57)	NSA	0.01237	0.00749	–0.00483	0.01021	0.82485
	SA	0.00041	0.00058	0.00004	0.00107	2.63839
Retail: Nondurable Goods Stores	NSA	0.00809	0.00270	–0.00325	0.00430	0.53181
	SA	0.00006	0.00004	0.00000	0.00009	1.45016
Retail: General Merchandise Group Stores (SIC 53)	NSA	0.04588	0.01966	–0.02139	0.02276	0.49608
	SA	0.00011	0.00037	–0.00004	0.00041	3.60050
Retail: Department Stores ex Leased Departments (SIC 31)	NSA	0.04906	0.02083	–0.02225	0.02539	0.51750
	SA	0.00013	0.00043	–0.00005	0.00046	3.70666
Retail: Food Stores (SIC 54)	NSA	0.00190	0.00019	–0.00006	0.00198	1.04026
	SA	0.00008	0.00002	0.00000	0.00010	1.28350
Retail: Apparel and Accessory Stores (SIC 56)	NSA	0.03814	0.01931	–0.01861	0.02023	0.53035
	SA	0.00030	0.00048	0.00002	0.00082	2.73757

removing the higher frequencies from the data masks evidence of smoothing.

In the nondurable goods category in Table 2, only Textile Mill Products (SIC 22) and Chemical and Allied Products (SIC 28) have variance ratios less than 1.0 for the detrended log seasonally unadjusted data. The aggregate nondurable goods industries has a variance ratio greater than 1.0. Intuitively, we would expect production of some nondurables to be less amenable to storage. For instance, Tobacco Products may be largely influenced by crop size rather than by demand, while demand may be less elastic seasonally. Of the nine other manufacturing sectors, which reflect a lower level of aggregation, evidence of smoothing is revealed in seven when seasonally unadjusted data are used. (See Table 3). This suggests that aggregation may be playing a role as well.

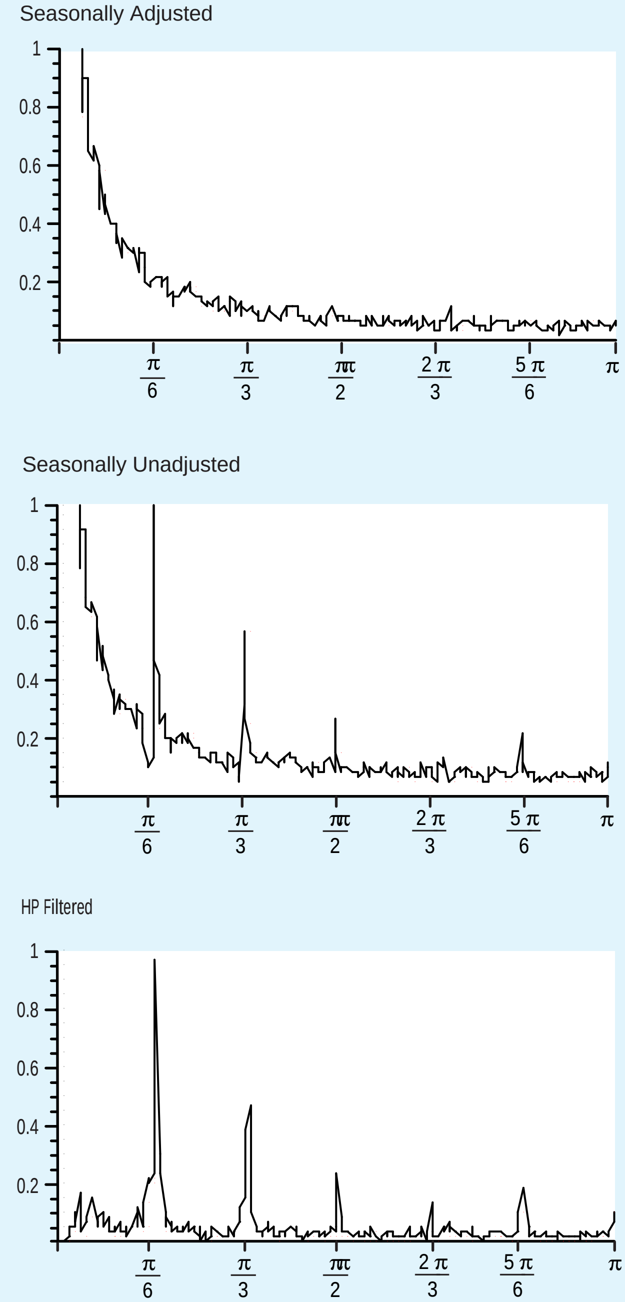
Retail Sector

The seasonally unadjusted data for the retail sector reveal smoothing by most industries, suggesting that some retail

firms may accumulate inventory in anticipation of seasonal increases in sales. Retail Food Stores (SIC 54) and aggregate of Retail Durable Goods stores are the only two of the 10 series that have a variance ratio higher than 1.0. Given that fixed costs associated with transportation are likely to induce (*S,s*) behavior at the retail level, detecting smoothing may appear to be contradictory. Here again, however, the frequency of observation influences the detection of the underlying decision rule.
It is likely that adjustments to inventory based on the (*S,s*) rule takes place at frequencies less than one month. So, on average, inventory moves between upper and lower bounds within a month. Consequently, monthly data reveals seasonal movements in the bandwidth, while obscuring higher frequency (*S,s*) movements. Seasonal smoothing at the retail level does not preclude (*S,s*) behavior at higher frequency. In addition, aggregation over a large number of establishments is likely to dampen high frequency movements.

Figure 2

Fourier Spectrum of Stone, Clay and Glass Industries Sales
(Seasonally Adjusted and Unadjusted)



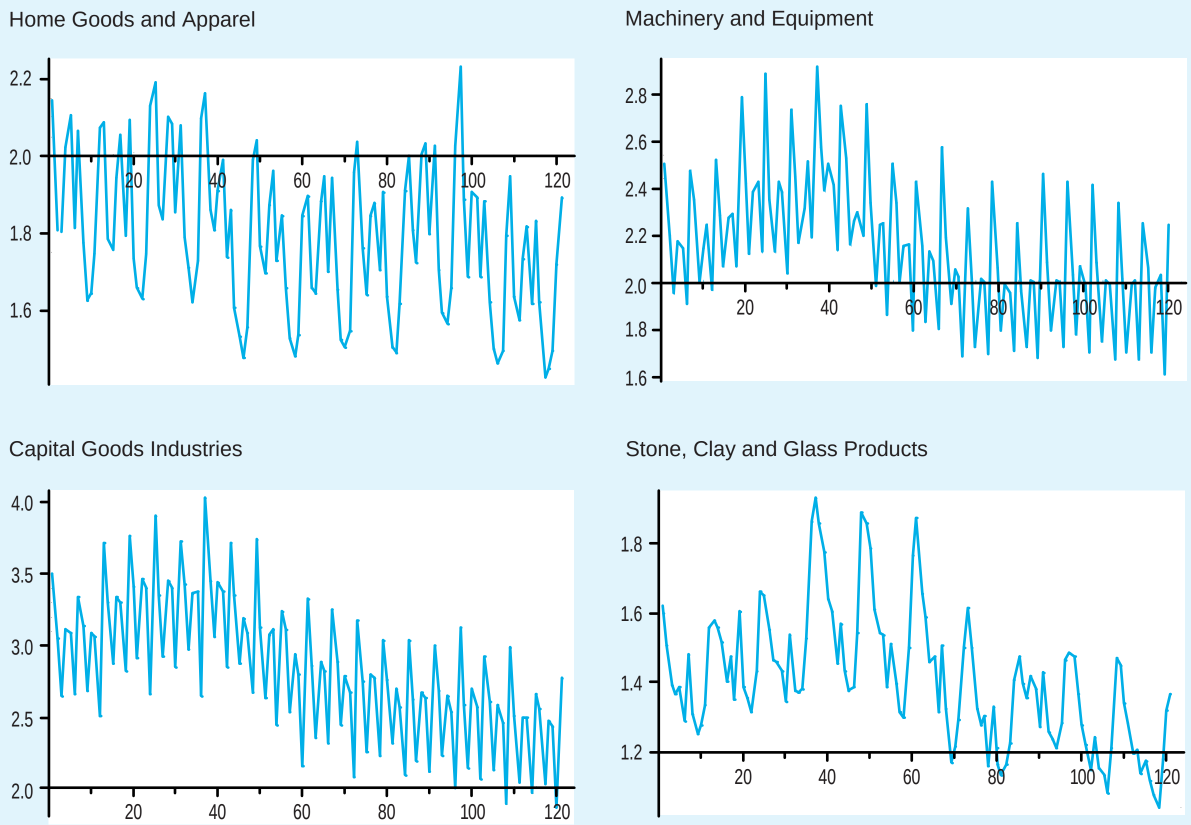
CONCLUSIONS

Empirical evidence of production smoothing has been relatively elusive. Part of the problem appears to be the tendency to use seasonally adjusted data. This paper finds that smoothing takes place in a large proportion of manufacturing industries at seasonal frequencies or higher. Seasonal adjustment of the data removes this evidence. Removing the trend from the data allows us to exclude changes in production associated with trend growth in sales. This confirms empirical results of Allen (1997B), which suggest that inventory management at the firm level reflects planned and unplanned changes. The trend component of production reflects planned additions to inventory levels based on trend movement in sales, while the higher frequency component of production reflects smoothing over shorter horizons. Evidence of seasonal smoothing in the retail sectors suggests that retail firms also manage inventory to smooth seasonal fluctuations in sales. Although smoothing is not generally associated with retail inventory management, it is not inconsistent with (*S,s*) behavior at frequencies higher than the observed data.

In summary, we find evidence of production smoothing at relatively high frequencies when the trend is removed from seasonally unadjusted data. We interpret this to mean that using data that has been seasonally adjusted and includes trend growth, limits the ability to extract the

Figure 3

Inventory-to-Sales Ratios Selected Industries



underlying motivation for holding inventories. To the extent that seasonal cycles mimic business cycles, analysis of production/inventory behavior at seasonal frequencies may provide insights into business cycle dynamics.

REFERENCES

Allen, Donald S. "A Multi-Sector Inventory Model," *Journal of Economic Behavior and Organization*, (January 1997A), pp. 55-87.

________. "Do Inventories Moderate Fluctuations in Output?" this *Review*, (July/August 1997B), pp. 39-50.

________. "Changes in Inventory Management and the Business Cycle," this *Review*, (July/August 1995), pp. 17-26.

Beason, Richard. "Tests of Production Smoothing in Selected Japanese Industries," *Journal of Monetary Economics*, (June 1993), pp. 381-94.

Blinder, Alan S. "Retail Inventory Behavior and Business Fluctuations," *Brookings Papers on Economic Activity*, 2 (1981), pp. 443-505.

________. "Can the Production Smoothing Model of Inventory Behavior be Saved?" *Quarterly Journal of Economics*, (August, 1986), pp. 431-53.

________., and Louis J. Maccini. "The Resurgence of Inventory Research: What Have We Learned?" *Journal of Economic Surveys* (1991), pp. 291-328.

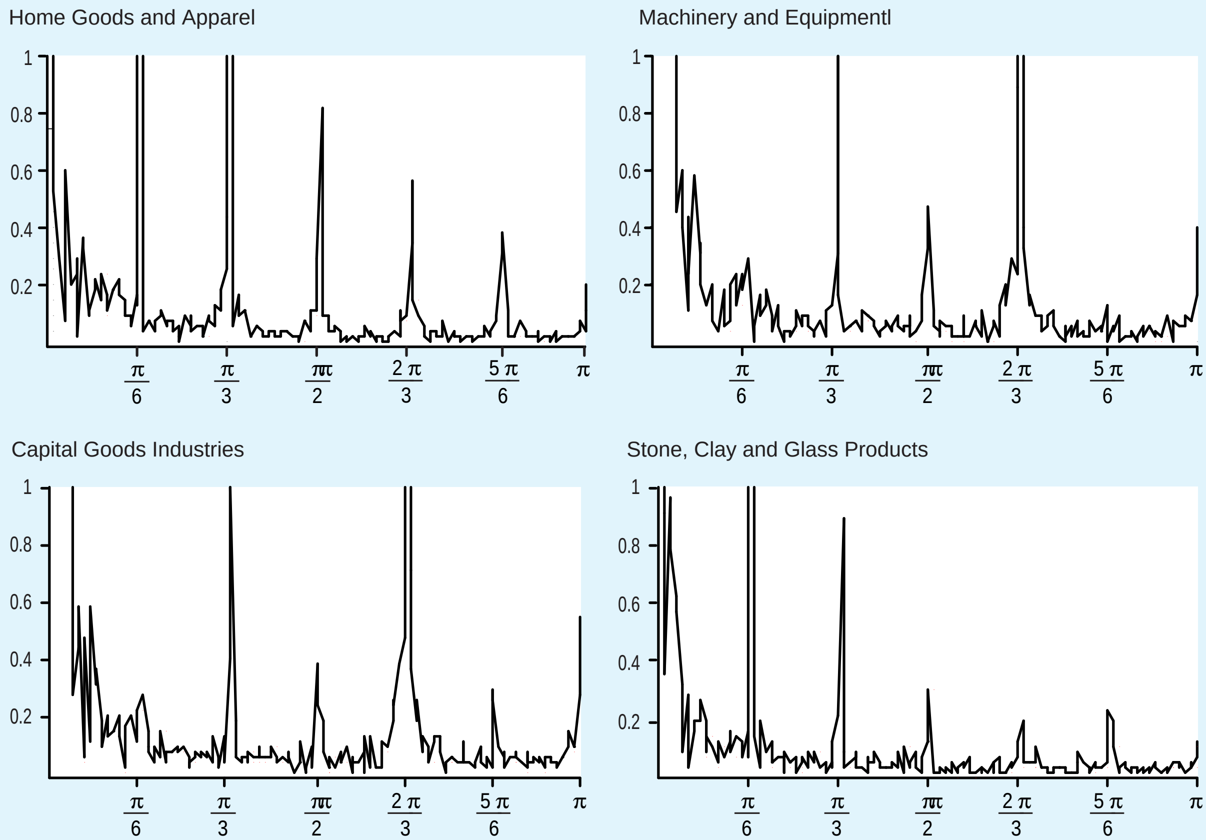
Carpenter, Robert E., and Daniel Levy. "Seasonal Cycles, Business Cycles, and the Comovement of Inventory Investment and Output," *Journal of Money, Credit, and Banking*, (August 1998), pp. 331-46.

Dimelis, Sophia P., and Moheb A. Ghali. "Classical and Variance Bounds Tests of the Production Smoothing Hypothesis," *International Journal of Production Economics*, (June 1994), pp. 15-22.

Fair, Ray C. "The Production-Smoothing Model is Alive and Well," *Journal of Monetary Economics*, (November 1989), pp. 353-70.

Figure 4

Fourier Spectrum of Inventory-to-Sales Ratios, Selected Industries



Filardo, Andrew J. "Recent Evidence on the Muted Inventory Cycle," Federal Reserve Bank of Kansas City *Economic Review*, (second quarter 1995), pp. 27-43.

Ghali, Moheb A. "Seasonality, Aggregation and the Testing of the Production Smoothing Hypothesis," *The American Economic Review* (June 1987), pp. 464-69.

Harvey, Andrew C. *Time Series Models*, Second Edition, (1994), The MIT Press, Cambridge, Massachusetts.

Holt, Charles C., Franco Modigliani, John F. Muth, and Herbert A. Simon. *Planning Production, Inventories, and Work Force* (1960), Prentice-Hall Inc., Englewood Cliffs, NJ.

Kollintzas, Tryphon. "A Generalized Variance Bounds Test with an Application to the Holt et al. Inventory Model," *Journal of Economic Dynamics and Control,* (Jan/Feb 1995), pp. 59-89.

Krane, Spencer D., and Stephen N. Braun. "Production Smoothing Evidence from Physical-Product Data," *Journal of Political Economy,* (June 1991), pp. 558-81.

Lai, Kon S. "Aggregation and Testing of the Production Smoothing Hypothesis," *International Economic Review*, (May 1991), pp. 391-403.

Miron, Jeffrey A., and Stephen P. Zeldes. "Seasonality, Cost Shocks, and the Production Smoothing Model of Inventories," *Econometrica*, (July 1988), pp. 877-908.

Miron, Jeffrey A. *The Economics of Seasonal Cycles*, (1996), Massachusetts Institute of Technology, Cambridge, Massachussets.

Ramey, Valerie A. "Nonconvex Costs and the Behavior of Inventories," *Journal of Political Economy*, (April 1991), pp. 306-34.

West, Kenneth D. "A Variance Bounds Test of the Linear Quadratic Inventory Model," *Journal of Political Economy*, (April 1986), pp. 374-401.

Appendix

FOURIER SERIES

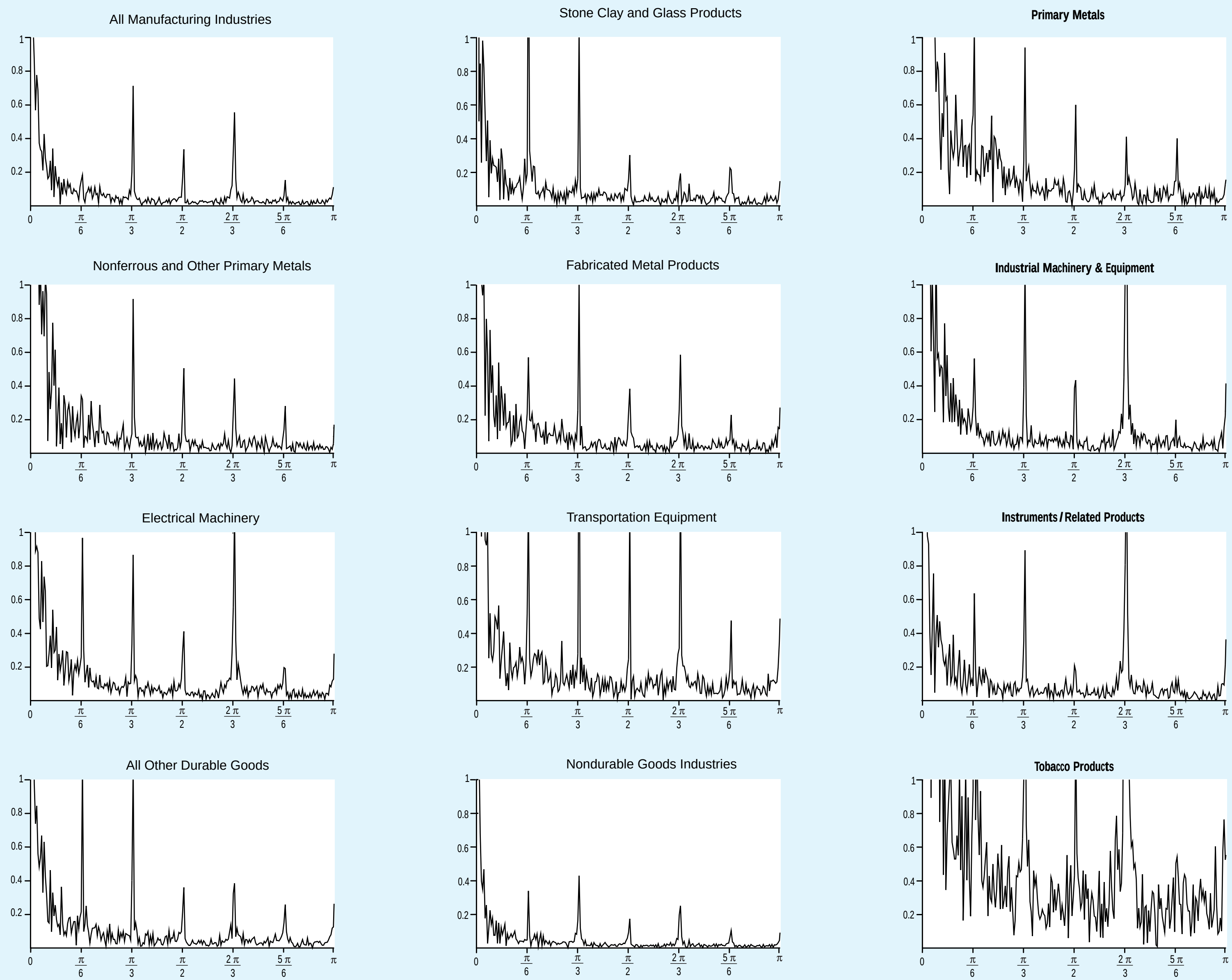
Figures 1A-3A in the appendix show plots of the Fourier spectra of the I/S ratios, and Figures 1B-3B show detrended sales and inventory movements of all 35 series. The horizontal axes of the plot of the spectra are labeled in multiples of π. Annual cycles are at $\pi/6$, cycles of 6 months are at $\pi/3$, and so on. The magnitude of the spike at each frequency gives an indication of the dominant cycles. In Figure 1A, the Fourier spectrum of the I/S ratio of the Stone, Clay and Glass Products sector has a high annual component (compared to 6-month). Figure 1B shows for that industry a negative correlation between detrended inventory and sales. By comparison, the Instruments and Related sector shows a high 3-month (quarterly) component (compared to annual) in Figure 1A. The corresponding chart in Figure 1B shows the high frequency composition of sales in this sector while inventory shows more of an annual cycle. Whereas the seasonally unadjusted variance ratio of Stone, Clay and Glass Products was 0.6864, the variance ratio for Instruments and Related Products was 0.9359.

For the three industries with positive covariance between sales and the change in inventory, Nonferrous and Other Metals (Figure 1A) and Paper and Allied Products (Figure 2A) show seasonal spikes in the I/S spectra, while no significant seasonality is depicted for Petroleum and Coal Products (Figure 2A). The positive co-movement between sales and inventory for all three is observable in Figures 1B and 2B. For the Petroleum and Coal industry the positive co-movement between sales and inventory eliminates all seasonal components from the I/S ratio while the other two industries show higher swings in sales than inventory, leaving some seasonality in the ratio.

The charts for this appendix follow on pages 34-39.

Figure 1A

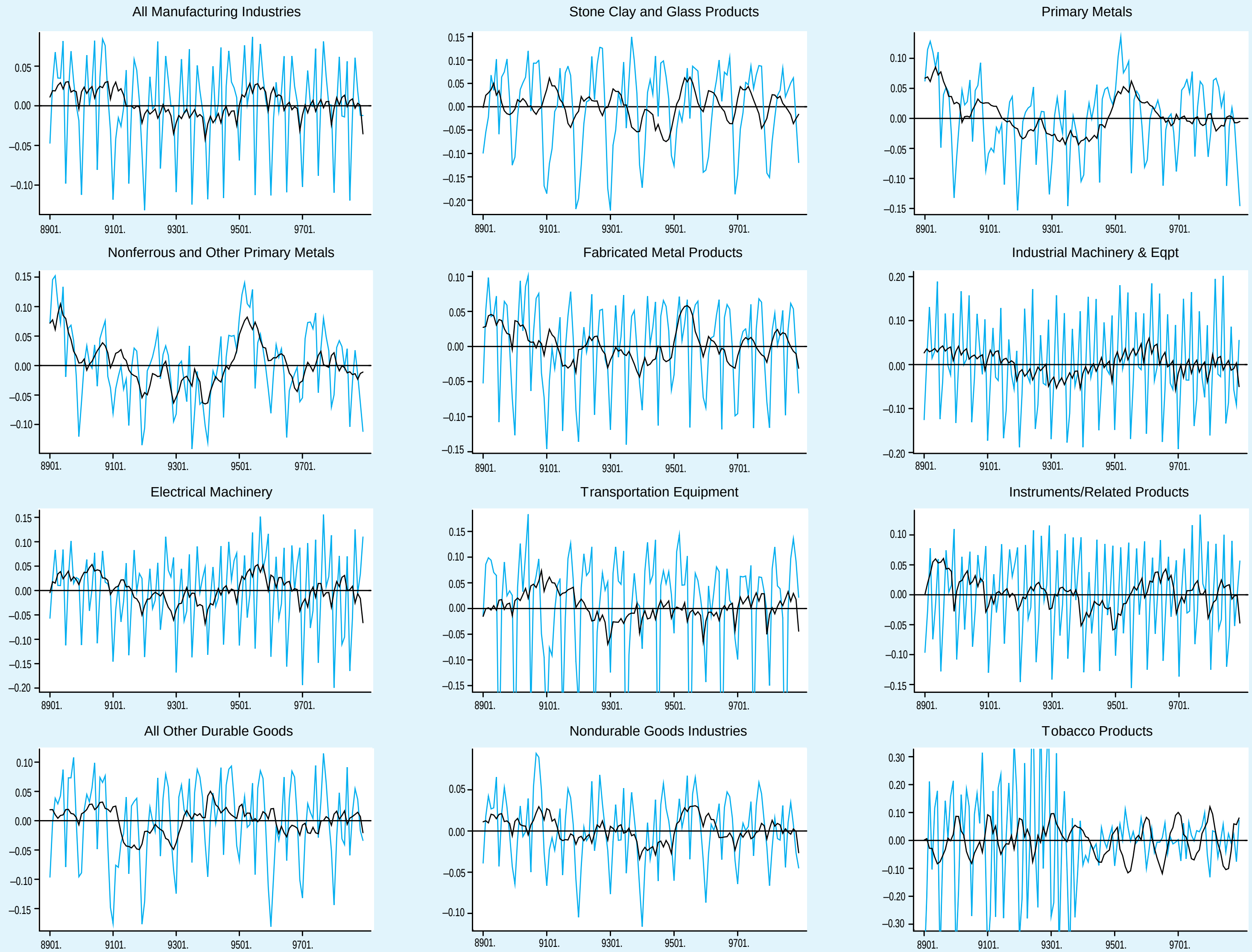
Fourier Series of Inventory-Sales Ratio



The Fourier Series shown in these charts are made up of complex numbers. Ther vertical scale measures the absolute values of the Fourier Series. The horizontal scale shows the frequency measured on a scale from zero to π.

Figure 1B

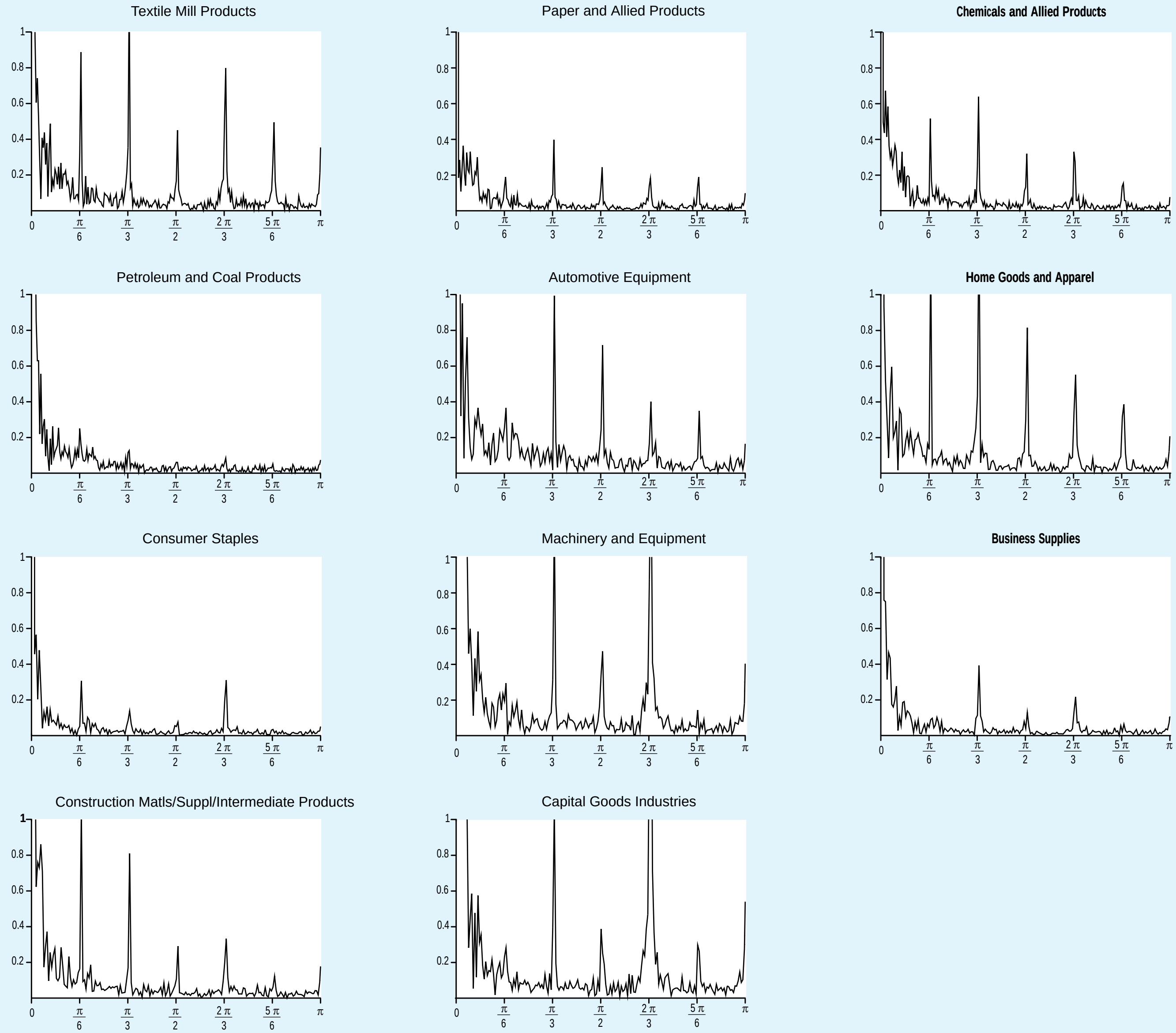
Detrended Log of Inventory (black) and Sales (blue)



The vertical scale depicts log deviations from trend of sales (shown in blue) and of inventories (shown in black). The horizontal scale is time. The data are monthly from January 1989 to December 1998.

Figure 2A

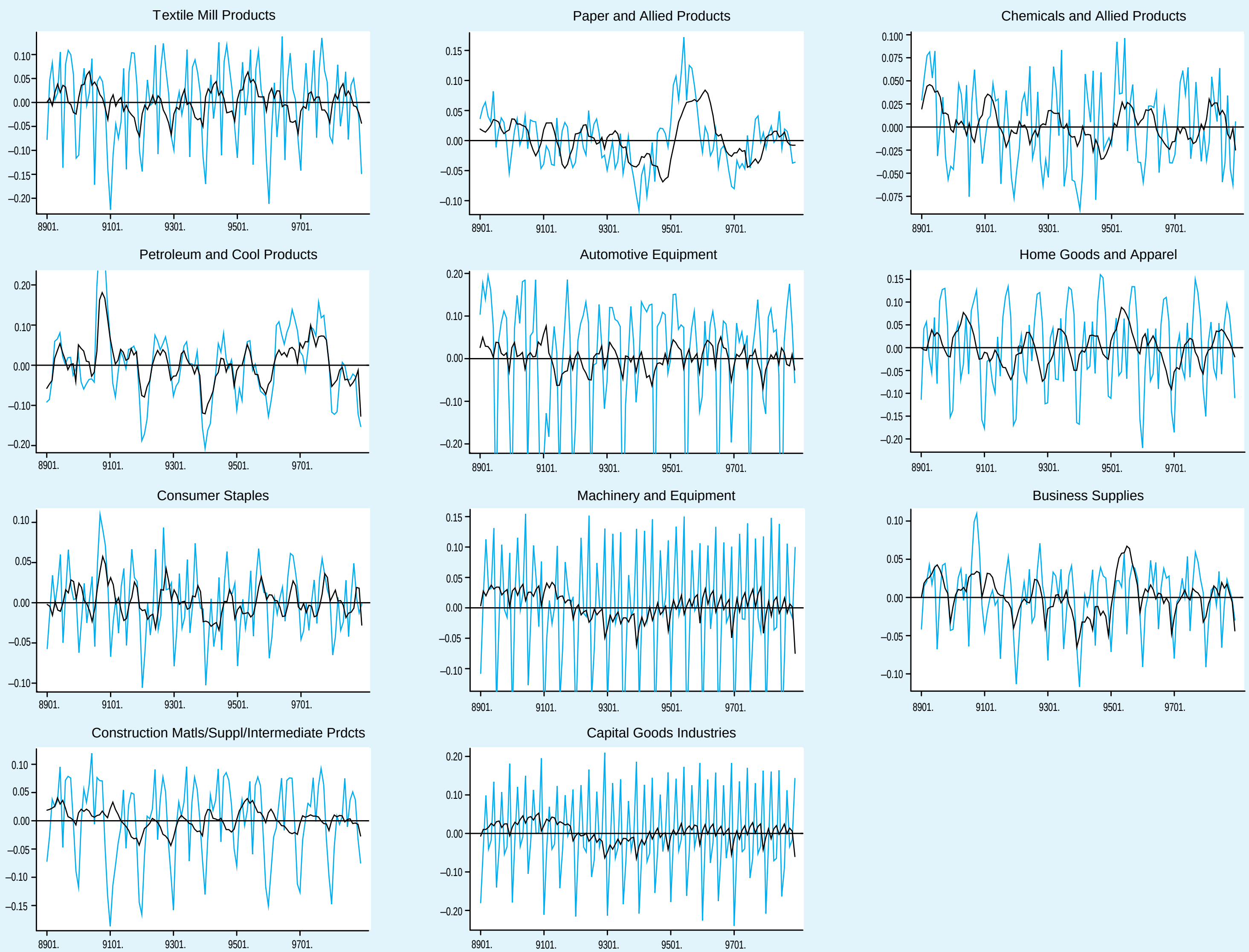
Fourier Series of Inventory-to-Sales Ratio



The Fourier Series shown in these charts are made up of complex numbers. The vertical scale measures the absolute values of the Fourier Series. The horizontal scale shows the frequency measured on a scale from zero to π.

Figure 2B

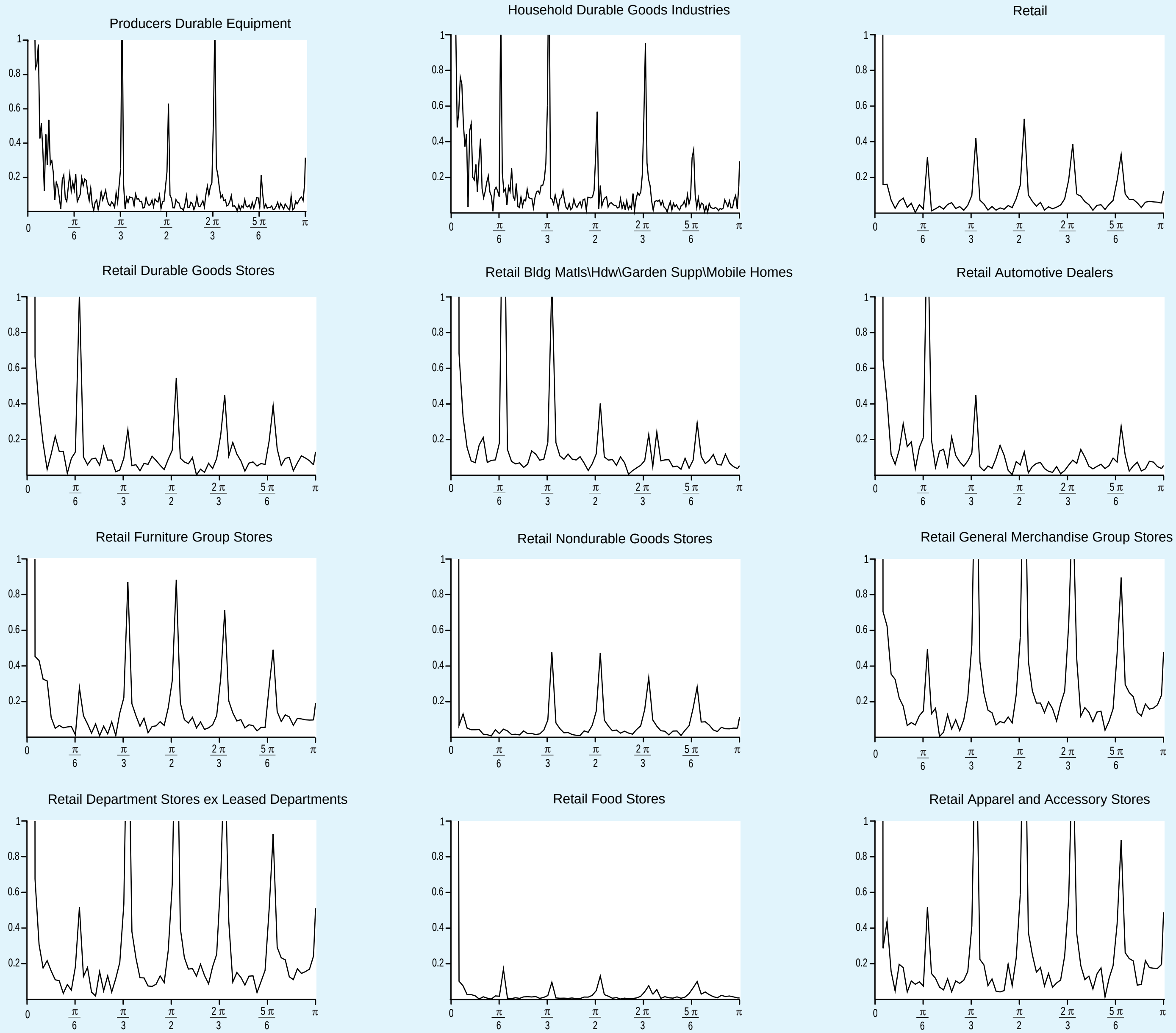
Detrended Log of Inventory (black) and Sales (blue)



The vertical scale depicts log deviations from trend of sales (shown in blue) and of inventories (shown in black). The horizontal scale is time. The data are monthly from January 1989 to December 1998.

Figure 3A

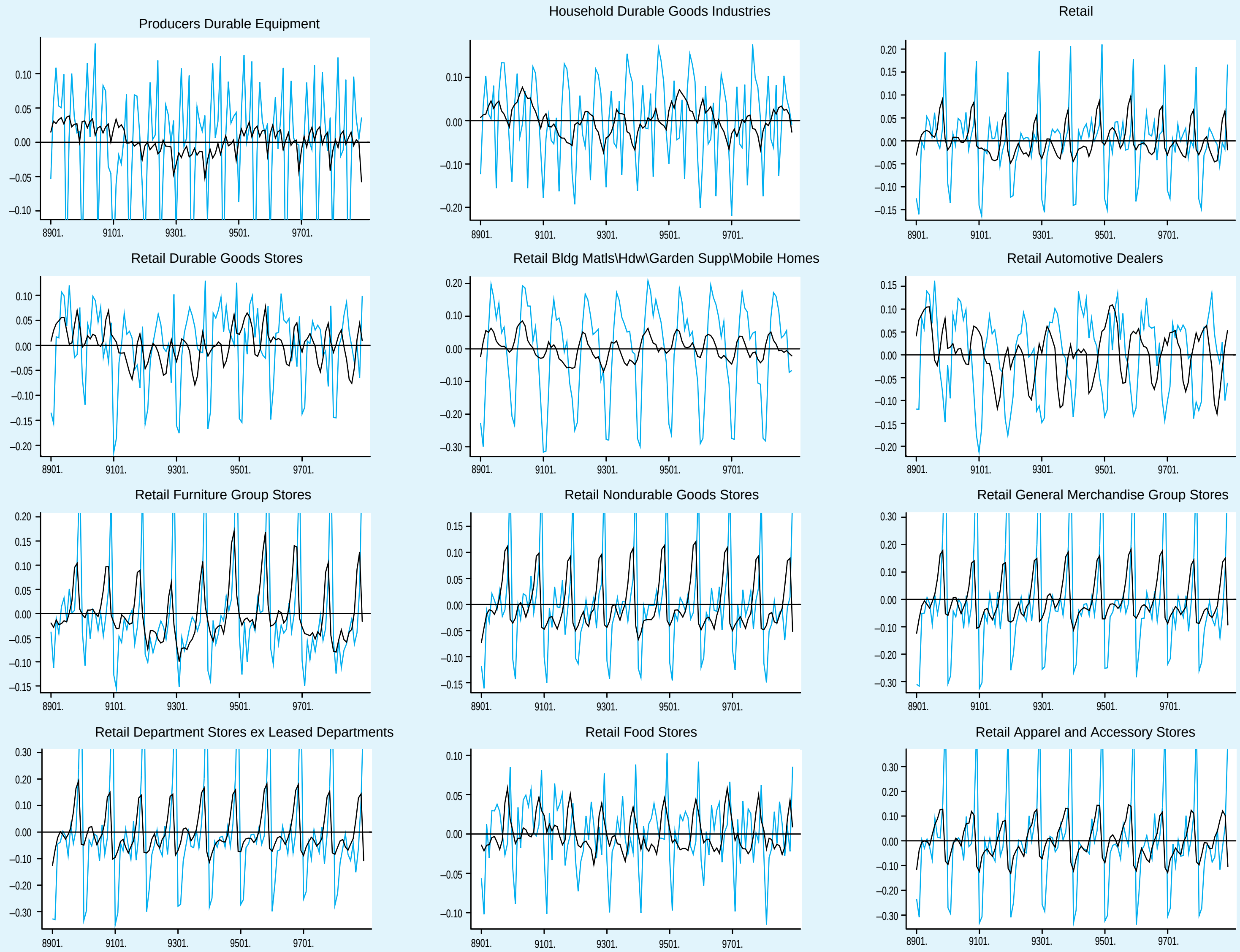
Fourier Series of Inventory-to-Sales Ratio



The Fourier Series shown in these charts are made up of complex numbers. Ther vertical scale measures the absolute values of the Fourier Series. The horizontal scale shows the frequency measured on a scale from zero to π.

Figure 3B

Detrended Log of Inventory (black) and Sales (blue)



The vertical scale depicts log deviations from trend of sales (shown in blue) and of inventories (shown in black). The horizontal scale is time. The data are monthly from January 1989 to December 1998.

William R. Emmons is a research economist and Frank A. Schmid is a senior research economist at the Federal Reserve Bank of St. Louis.
Robert Webb and Marcela Williams provided research assistance.

Credit Unions and the Common Bond

William R. Emmons and Frank A. Schmid

Cooperative financial institutions have their roots in 19th century Europe, appearing first in the United States during the early 20th century. Cooperative financial institutions are ubiquitous in both developed and developing countries today, posing something of a puzzle in the former group of countries where one might have expected corporate financial institutions with professional management and sophisticated capital-market oversight to have displaced them. This has not occurred, however, as some groups of cooperative financial institutions in developed countries are holding steady or even increasing their market shares. In the United States, the most prominent types of cooperative financial institutions today are mutual savings and loans, mutual savings banks, mutual insurance companies, and credit unions.

Credit unions are regulated and insured financial institutions dedicated to the saving, credit, and other basic financial needs of selected groups of consumers. By law, credit unions are cooperative enterprises controlled by their members—under the principle of "one-person one-vote." In addition, credit union members must be united by a "common bond of occupation or association, or (belong) to groups within a well-defined neighborhood, community, or rural district" (Supreme Court, 1998, p. 2, quoting from the Federal Credit Union Act of 1934).

Despite the rather low profile and mundane operations of the vast majority of credit unions, these institutions have long been a source of controversy in the United States. Public awareness of this long-simmering debate was piqued recently by a Supreme Court case pitting commercial banks against credit unions and their federal regulator (Supreme Court, 1998). The Court found in favor of banks in this case, ruling that the federal credit-union regulator, the National Credit Union Administration, must cease granting federally chartered credit unions the right to combine multiple common bonds (fields of membership) within a single institution. Less than six months later, however, President Clinton signed into law new legislation that essentially reversed the Supreme Court's ruling.

This paper provides background on credit unions and the debate they have spurred in the United States. In addition, we present new evidence relevant to the credit-union debate concerning fields of membership (common bonds). Our analysis is based on a theoretical model of credit-union formation and consolidation. Using an extensive dataset and a nonlinear empirical approach, we find that credit-union participation rates generally decline as the group of potential members becomes larger, holding all else equal. That is, the larger the pool from which a single-group credit union can draw, the less effective it is in attracting members.

We also provide new evidence on two more general banking policy issues. First, we find evidence to support the structure-conduct-performance paradigm of local banking competition. This is the prediction, derived from theoretical considerations, that more concentrated markets ultimately lead to higher prices and lower quantities. Policymakers have used this paradigm extensively when justifying intervention in the market for corporate control in financial services. Using the Herfindahl index calculated for local bank deposit market shares as a measure of local market structure, we find that higher levels of market concentration are associated with higher participation rates at credit unions. This is consistent with the notion that banking competition is weaker in more concentrated markets, which increases the attractiveness of credit unions.

The second banking policy issue we address is that of possible scale economies among financial institutions. Our empirical results indicate that credit unions generally encounter significant scale economies, whether scale is measured by the log of total assets or by the log of the number of credit-union members. The latter finding, however, applies only to relatively large credit unions.

It is important to point out several limitations of this study. As in all empirical investigations, we can describe relationships in the existing data but we cannot predict exactly how these relationships would appear under a different set of operating conditions. For example, an extended period of growth by many credit unions could alter the extent of scale economies that exist. Similarly, significant changes in credit-union regulation might result in different empirical regularities than those identified here. It also is important to keep in mind that we abstract from managerial agency problems in credit unions in this article (see Emmons and Schmid, 1999, for an extensive discussion of this issue). Finally, it is hazardous to draw conclusions about public policy toward credit unions on the basis of this rather narrowly focused investigation. We hope to provide insights into the effects of common-bond requirements, not to provide a comprehensive framework for evaluating competition in the financial-services sector as a whole.

The paper is organized as follows: The first section provides some institutional and historical background on credit unions, while the second section outlines the current credit-union debate in the United States. The third section develops a theoretical model of credit-union formation and consolidation. The model stresses the countervailing influences on participation rates of (1) scale economies in production, and (2) decreasing within-group membership affinity as a credit union grows. The model provides intuition for why the number of common bonds within a credit union might be important for their formation and growth. The third section also describes a simulation of the theoretical model that can be used to generate some comparative-static results. The fourth section briefly describes the dataset and the econometric methods we employ in analyzing federally chartered occupational credit unions. The fifth section presents our empirical results, and the sixth section draws conclusions. An appendix describes the data we use.

BACKGROUND ON CREDIT UNIONS

This section provides some institutional background to help motivate the theoretical and empirical analyses later in the article. The key points this section seeks to illuminate are the restrictions on credit-union expansion and the arguments that have been made to support or oppose these restrictions. The sections that follow investigate the extent to which the common-bond requirement acts as a binding constraint on credit-union operations.

Overview of Credit Unions in the United States

Credit unions numbered 11,392 at year-end 1996, serving some 70 million individual members (U.S. Treasury, 1997, p. 15). At the same time, there were 11,452 commercial banks and thrift institutions (savings and loan associations and mutual savings banks). Credit-union assets were only $327 billion, compared to $5,606 billion held by commercial banks and thrifts (U.S. Treasury, 1997, p. 21). A more direct standard of comparison might be community banks and thrifts, however. At year-end 1996, there were 7,049 community banks and thrifts (defined as all federally insured banks and thrifts with less than $100 million in assets) holding combined assets of $324 billion (U.S. Treasury, 1997, p. 21). A comparison of credit unions and community banks and thrifts is particularly meaningful because institutions of both types are relatively focused institutions, and hence, are unable to grow beyond certain limits. For example, a single-employer occupational credit union is authorized to serve only the employees of the sponsoring firm and their immediate relatives, who may total no more than a few hundred people. A community bank or thrift may operate in only one geographical area. In addition, credit unions are restricted in the types of financial services they may provide, with traditional consumer financial services at the core of virtually all credit unions' activities. Community banks and thrifts may offer a similar array of services.

Both federal and state agencies grant credit-union charters. Regardless of the type of charter they hold, the deposits (or technically, "shares") of virtually all credit unions are now federally insured by the National Credit Union Administration (NCUA). Federal credit unions are regulated by the NCUA while state-chartered credit unions are regulated by an agency of the chartering state.

Of the 7,068 federally chartered institutions at year-end 1996, about three quarters were occupational credit unions (U.S. Treasury, 1997, p. 19).[1] In an occupational credit union, one or more firms sponsor a credit union, sometimes providing office space, paid time off for volunteer workers, and perhaps other forms of support. The remaining federal credit unions were either single-group associational or community credit unions, or multiple-group credit unions with predominantly associational, community, or more than one type of membership (i.e., several groups that span the usual classifications).

By size, most credit unions (65 percent of federally insured institutions) had less than $10 million in assets (U.S. Treasury, 1997, p. 19). Large credit unions exist, however, and they are an important part of the sector. For example, the 11 percent of credit unions with more than $50 million in assets (1,284 institutions) accounted for 74 percent of total credit-union assets.

Credit unions play a limited role in the U.S. financial system, catering to the basic saving, credit, and other financial needs of well-defined consumer groups. More than 95 percent of all federal credit unions offer automobile and unsecured personal loans, while a similar proportion of large credit unions (more than $50 million in assets) also offer mortgages; credit cards; loans to purchase planes, boats, or recreational vehicles; ATM access; certificates of deposit; and personal checking accounts (U.S. Treasury, 1997, p. 23). Very small credit unions typically offer a limited range of services, are staffed by member-volunteers, and are likely to receive free or subsidized office space. Larger credit unions offer a broader array of services. They may employ some full-time workers, including the manager, and are more likely to pay a market-based rent for office space.

Historically, members of credit unions were drawn from groups that were underserved by traditional private financial institutions; these consumers tended to have below-average incomes or were otherwise not sought out by banks. While credit-union members today still must share a common bond to be eligible for membership, the demographic characteristics of credit-union members have become more like the median American. While only 1 percent of the U.S. adult population aged 18 or over belonged to a credit union in 1935, some 33 percent of the adult population had joined by 1989 (American Bankers Association, 1989, p. 29). Subsequent strong growth of new credit-union charters has increased that proportion.[2]

According to a credit-union survey in 1987, 79 percent of all Americans who were eligible to join a credit union had done so (American Bankers Association, 1989, p. 29). Given the prominent role of occupational credit unions, a majority of members are in the prime working ages of 25-44 (American Bankers Association, 1989, p. 30). Perhaps surprisingly, given the origins of credit unions, current members are overrepresented in upper-middle

[1] We concentrate on federally chartered credit unions because the NCUA does not vouch for the accuracy of data provided by state-chartered credit unions, which report directly to their state's regulatory authorities.

[2] The estimated 70 million current credit-union members represent a bit more than 34 percent of the 1996 U.S. population over 16 years of age numbering 204 million (U.S. Census Bureau, <http://www.census.gov>).

income strata, defined as household incomes between $30,000 and $80,000 in 1987. Overall, it appears that credit unions, banks, and thrifts are more direct competitors today than when credit unions first appeared.

A Brief Legislative History of Credit Unions in the United States

The predecessors of American credit unions were cooperative banking institutions of various sorts in Canada and Europe during the 19th century. The first credit union in the United States was formed in Manchester, New Hampshire, in 1909 (U.S. Treasury, 1997, p. 15). Soon thereafter, Massachusetts created a charter for credit unions. The credit-union movement swept across the United States from there, meeting with particular success in the New England and upper Midwestern states.

These early cooperative financial institutions often had a social, political, or religious character in addition to their explicit economic function. While the social and political aspects of the cooperative movement were acknowledged and accepted by the United States Congress, the Federal Credit Union Act (FCUA) of 1934 was focused more narrowly on the economic potential of credit unions.

The legislation itself was modeled closely on state credit-union statutes that had appeared during the early decades of the 20th century in the Northeast and upper Midwestern states. The FCUA clearly reflected Congressional intent to create a class of federally chartered financial institutions that would operate in a safe and sound manner:

> … the ability of credit unions to "come through the depression without failures, when banks have failed so notably, is a tribute to the worth of cooperative credit and indicates clearly the great potential value of rapid national credit union extension." (Supreme Court, 1998, p. 17, citing the FCUA, S.Rep. No. 555.)

The likelihood that federal credit unions would serve consumers not served by banks was an additional element in Congressional deliberations:

> Credit unions were believed to enable the general public, which had been largely ignored by banks, to obtain credit at reasonable rates. (Supreme Court, 1998, p. 17.)

Partly because credit unions are mutual associations, they were not subjected to federal taxation as were shareholder-owned commercial banks and thrift institutions. Mutuality cannot be the only reason why credit unions are not taxed, however. Other mutually owned enterprises are subject to taxation. As for the benefits of tax exemption, credit unions (or any other firm) could avoid paying taxes by paying out all "profits" to members in the form of lower borrowing rates or higher deposit rates. The real importance of the tax exemption is that credit unions can retain earnings tax free. Advocates argue that this is justified because credit unions cannot raise equity in a public offering, so they must be able to build capital internally.

It is clear from the legislative history surrounding the passage of the FCUA in 1934 that Congress saw the common-bond requirement as critical to the success of credit unions:

> The common bond requirement "was seen as the cement that united credit union members in a cooperative venture, and was, therefore, thought important to credit unions' continued success. ..."
>
> "Congress assumed implicitly that a common bond amongst members would ensure both that those making lending decisions would know more about applicants and that borrowers would be more reluctant to default." (Supreme Court, 1998, pp. 17-18, citing 988 F.2d, at 1276.)

The subsequent history of credit unions in the United States largely has fulfilled the promise envisioned by

Congress in 1934. Credit unions have grown and spread across the country. Although hundreds of individual credit unions failed during the 1980s and early 1990s, the National Credit Union Insurance Fund (NCUSIF, formed in 1970) avoided accounting insolvency—in marked contrast to the Federal Savings and Loan Insurance Corporation and the Bank Insurance Fund of the Federal Deposit Insurance Corporation (Kane and Hendershott, 1996). Credit unions control a small but growing share of household deposits, and some of our empirical results indicate that they may play a role in maintaining a high level of retail banking competition in some local markets.

THE CURRENT CREDIT-UNION DEBATE

The special status and comparative success of credit unions in recent decades, coinciding as it has with a period of stress on thrift and commercial-banking institutions, has led to political conflicts between advocates of credit unions and banks. This conflict reached its high point in a series of court decisions culminating at the U.S. Supreme Court in October 1997. The particular case at issue involved the AT&T Family Credit Union and the NCUA's interpretation of the 1934 FCUA allowing multiple common bonds of membership. Brought by several banks and the American Bankers Association, the case was ultimately decided in February 1998 (on a 5-4 decision) in favor of the banks who sued to stop the NCUA from granting more multiple-group credit-union charters. The bankers' victory was short-lived, however, as Congress almost immediately drafted new legislation that enables credit unions to continue growing much as before—including multiple common bonds within a single credit union. The shaded insert summarizes the key provisions of the Act.

Attacks on credit unions have come from a wide range of viewpoints, the proponents of which have wielded sometimes contradictory arguments. Some of the arguments used in the recent Supreme Court decision concerning the role of the common-bond requirement in credit unions reflect the unsettled nature of the debate. We focus on two strands of the credit-union debate here, namely the arguments stressing inefficient governance structures on the one hand and unfair competition on the other.

Some have argued that credit unions are inherently inefficient due to their one-member one-vote governance structure. One might expect decision-making in a credit union to be of poor quality due to a lack of professionalism (i.e., volunteer managers and workers), free-riding of members in monitoring the management, and weak incentives for members to intervene when action is needed to correct specific problems or deficiencies.[3] According to this argument, credit unions may waste scarce resources and they may eventually impose significant costs on individual sponsoring firms or the economy as a whole.

The second prominent line of argument aimed at credit unions takes a nearly opposite view of their organizational effectiveness. This view presumes that credit unions operate efficiently enough to offer consistently better terms on savings and credit services than those offered by commercial banks and thrifts. Bank and thrift managers and owners often present this point of view in public discourse. To be sure, those arguing that credit unions represent unfair competition ascribe some or all of their competitive advantages to subsidies such as their tax-exempt status or sponsor subsidies rather than inherent efficiency.

Proponents of the first view—that credit unions are inherently inefficient—have a difficult time explaining why the number of credit unions and credit-union members continues to grow, and why members express high levels of satisfaction with the services they receive. If most credit unions were very inefficient, one might expect their members to become disaffected and their role in the financial system to diminish over time.

[3] Free-riding is when members choose not to exert monitoring effort because they assume someone else will do it for them.

THE CREDIT UNION MEMBERSHIP ACCESS ACT

President Clinton signed the Credit Union Membership Access Act on August 7, 1998, following approval in the Senate on July 28 and in the House of Representatives on August 4. The act substantially reverses a Supreme Court ruling handed down on February 25, 1998, that would have barred federally chartered credit unions from accepting multiple membership groups, each with its own common bond.

This landmark credit-union legislation represents a major defeat for the top lobbying group representing commercial banks, which had argued successfully at the Supreme Court that credit unions with multiple common bonds violated both the letter and the spirit of federal legislation dating from 1934. The subsequent legislative response in support of multiple common bonds at credit unions was swift and overwhelming, passing both chambers with large majorities.

The act contains three provisions upholding the rights of federal credit unions to serve membership groups encompassing multiple common bonds. First, all federal credit unions that already included multiple common bonds before February 25, 1998, were allowed to continue operating without interruption. Second, all federal credit unions were given the right to accept additional membership groups with multiple common bonds so long as the relevant groups have fewer than 3,000 members. Third, the act gives the National Credit Union Administration the right to grant exemptions to the 3,000-member limit under certain circumstances, such as when the group in question could not reasonably support its own credit union.

The act also:

- Requires annual independent audits for insured credit unions with total assets of $500 million or more.
- Authorizes and clarifies a federally insured credit union's right to convert to a mutual savings bank or savings association without prior NCUA approval.
- Limits business loans to members to 12.25 percent of total assets.
- Establishes new capital standards for insured credit unions similar to those enacted for banks and thrifts in 1991.
- Gives the NCUA authority to base deposit-insurance premiums on the reserve ratio of the insurance fund.
- Directs the Treasury to report to Congress on differences between credit unions and other federally insured financial institutions, including the potential effects of applying federal laws—including tax laws—to credit unions.

Hailing the new legislation, President Clinton said, "This bill ensures that consumers continue to have a broad array of choices in financial services….and [makes] it easier for credit unions to expand where appropriate." Meanwhile, a spokeswoman for the American Bankers Association termed it "ironic" that the bill was presented as a measure to protect credit unions because in the long run, she said, it will dilute them, turning them into larger and larger institutions.

Source: *BNA Banking Report,* "House Passes Credit Union Bill; Clinton Wastes No Time Signing It," August 10, 1998, Vol. 71, No. 6.

On the other hand, proponents of the second view—that credit unions are unfair competitors due in part to subsidies—cannot explain easily why credit-union sponsors and governments are such strong supporters of credit unions. It is hard to understand how large net subsidies could be delivered to credit-union members over time without more opposition arising from constituencies that might be paying the subsidies, such as shareholders or employees who do not belong to their firm's occupational credit union, or taxpayers who belong to no credit union. In fact, the most vocal complaints about alleged subsidies for credit unions are heard from banks and thrifts, whose resentment of credit-union competition could be expected even if there were no subsidies flowing to credit unions.

Ironically, the juxtaposition of these two lines of attack against credit unions appeared in the argumentation of the Supreme Court majority that decided the AT&T Family Credit Union case in favor of commercial banks. At one point in its opinion, the majority cited the legislative history surrounding the 1934 Federal Credit Union Act as support for the view that credit unions are a fragile—even flawed—type of institution, reasoning that:

> Because, by its very nature, a cooperative institution must serve a limited market, the legislative history of Section 109 demonstrates that one of the interests "arguably…to be protected" by Section 109 is an interest in limiting the markets that federal credit unions can serve. (Supreme Court, 1998, footnote 6, pp. 8-9.)

Thus, a credit union would become inefficient if it grew beyond its "limited market," as defined by its common bond.

At a different point in its opinion, however, the majority accepted the argument that credit unions with multiple groups of members would be *more* formidable competitors to banks and thrifts than single-group institutions. The majority argued that an expansive interpretation of the 1934 Act "would allow the chartering of a conglomerate credit union whose members included the employees of every company in the United States (1998, p. 4)." In other words, credit unions would overwhelm banks and thrifts unless otherwise constrained.

The irony is, of course, that the argumentation based on the *reductio ad absurdum* of a hypothetical "conglomerate credit union" did not mention the legislative history of the 1934 Act, which had essentially predicted that such a huge credit union would not have been a safe and sound financial institution, nor consequently a viable one in the long run.

THE MODEL AND SIMULATION

How should policymakers think about credit unions? Are they relics of a bygone era, propped up by subsidies and distorting financial-sector competition? Or, are they efficient and focused financial institutions that could, if unleashed, eventually dominate some or all of the retail financial landscape? We do not seek to answer these emotionally charged questions directly. Instead, we focus on the more limited question of what effect the common-bond restriction exerts on credit-union formation and consolidation. In a sense, we are merely attempting to answer the question, "Does the common-bond requirement constrain the existence or growth of credit unions?" We hope that our insights may contribute to a better understanding of the larger policy questions mentioned above.

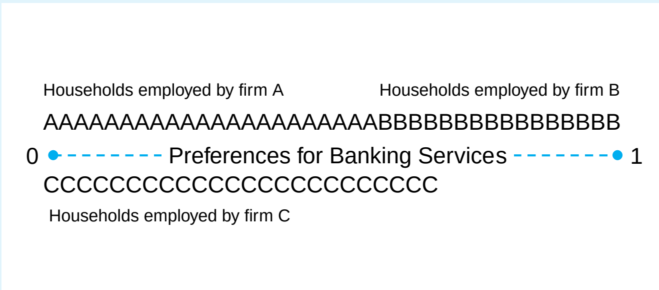
In this section we present a model of credit-union formation and consolidation. We then describe the results of a simulation of the model. Subsequent sections of the paper discuss testable hypotheses emerging from the model, the data we examine, and empirical results.

The Model

We take for granted that credit unions typically are small; that they encounter operating economies of scale as they expand from a very small base of members and assets; and that they face direct competition from banks. The key trade-off we model is between decreasing affinity among members as the potential membership grows (i.e., as a given common bond is

Figure 1

Linear City with Three Common Bonds of Occupation



extended to more people)—making a credit union less effective—versus the increasing scale economies that come with a larger base of members and assets—making a credit union more effective. We show that the ability of credit unions to expand by adding multiple common bonds to their membership affects this trade-off in an important way.

We examine a Hotelling (1929) economy consisting of a "city" that lies on a straight line of unit length. The city's length is covered by a continuum of households. The location of each household corresponds to its preferences for banking services. In particular, each household demands exactly one unit of banking services but the nature of desired services differs among households. Preferences in the real world are, of course, multidimensional, encompassing tastes for different menus of financial services, different levels of service, or different locational preferences. We assume for the sake of simplicity, however, that a household's preferences for banking services can be represented in terms of a single index running from zero to one. Figure 1 depicts the linear-city model.

Because we are interested only in the formation and consolidation of credit unions, we assume that credit unions are scarce (or differentiated) while commercial banks are ubiquitous (or uniform). In other words, consumption of credit-union-provided financial services takes place at the point on the unit interval where a credit union is located, while commercial-bank services are available at a fixed price at any point on the line. This assumption makes household preferences critical for the existence of and participation in credit unions while maintaining the realistic assumption that commercial banks provide an alternative to credit unions (and vice versa).

We assume that the entire city (i.e., every point on the line) is covered by at least one household and at most two households. Without loss of generality, we assume that all points covered by two households are arrayed continuously from zero upward towards, but potentially short of, one on the unit interval. For expositional purposes, we will refer to the households that inhabit the completely covered zero-to-one interval as being above the line and all others as below the line. Thus, two households that possess identical locations (preferences) are said to be "back-to-back" households.

Households are further grouped by affinity, or common bonds. For tractability, we discuss occupational common bonds and limit the number of employers in the economy to three. Each household located above the line contains an employee of either firm A or firm B (but not both). Because all households in employee group A share a common bond, they are located in a contiguous segment of the line that does not overlap the domain of employee group B. All households below the line contain employees of firm C. Each employer may sponsor a credit union, although, as we will see, not all will do so.

We examine two periods (or regimes), differentiated according to the permissibility of forming credit unions with multiple common bonds. All households are born at the start of period 1 and live through the end of period 2. Each household needs to consume one unit of banking services in each period. These services can be provided by an occupational credit union or by a bank in either period.

At the beginning of the first period, households find themselves arrayed along the city's unit interval. The lengths of the firm-A and firm-C segments are distributed as uniform random variables on the [0, 1] interval. The length of the firm-B segment is one minus the length of the firm-A segment.

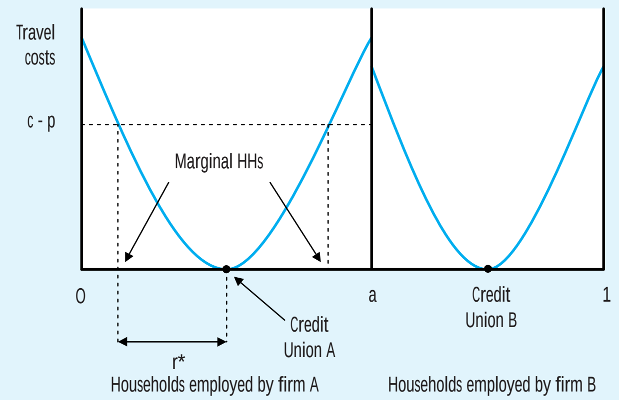
Suppose, first, that each of the three employers sponsors a credit union (in the simulation below, not all firms necessarily sponsor a credit union). All credit unions are restricted to a single employee group during the first period. Each credit union has a life span of one period. The credit unions have idiosyncratic technologies for producing banking services. In particular, each credit union *i* operates with fixed costs $f_i = f_a + f_b e_i$, where f_a and f_b are common to all credit unions and e_i is an i.i.d. uniform random variable. In addition, each credit union faces constant marginal costs of *v* per unit of banking services provided. Thus, the cost function of credit union *i* is $C(m_i) = f_a + f_b e_i + v m_i$, where m_i is the number of actual members in the credit union, and *i* = *A, B,* or *C*. An important feature of this cost function is that average costs, $C(m_i)/m_i$, are declining in the number of members the credit union is able to attract.

At the beginning of period 1, households vote on the credit-union management team for that period. Voting is costless, the one-household one-vote principle applies, and side payments among households are permitted (to allow those with strongly held preferences to "buy" the votes of those with weaker preferences). This implies, by virtue of the "value maximization principle," that the competitive outcome maximizes social welfare (Milgrom and Roberts, 1992, pp. 36-37). The potential members choose a management team that locates the credit union to minimize the sum of member "travel costs" (which are described in the following paragraph). It is clear that the credit union will locate in the center of the preference spectrum of all potential members because we assume that travel costs are quadratically increasing in the distance between member households and the credit union.

Credit-union services are offered at the price p_i to all potential members of a credit union (i.e., employees of the relevant firm). The price equals the credit union's

Figure 2

Travel Costs Facing Households Employed by Firms A and B



average costs (AC_i) because credit unions are not-for-profit institutions. Households face marginal costs of $t \times r$ per unit of distance *r* when travelling to a credit union with *t* being a travel-cost parameter. This is because the credit union's banking services are (in general) not identical to a given household's preferences (i.e., location on the line). The cost of using credit union *i* at a distance r_j from household *j*'s location is $(t/2)r_j^2$. Each household also can access banking services from a commercial bank at a constant price *c*. Together, these assumptions imply that the membership of credit union *i* will comprise all households *j* within the potential membership for which the following inequality holds (see Figure 2):

$$\frac{t}{2} r_j^2 \leq c - p_i .$$

In particular, the marginal—i.e., most distant—households will be the ones (on either side of the credit union) for which the expression holds at equality:

$$\frac{t}{2} r^{*2} = c - p_i .$$

As Figure 2 illustrates, not every potential member joins the credit union. A household relatively far from the credit union (at a distance greater than *r**) buys banking services from a commercial bank instead. The number of members credit

Figure 3

Demand and Supply Curves for Credit-Union Services

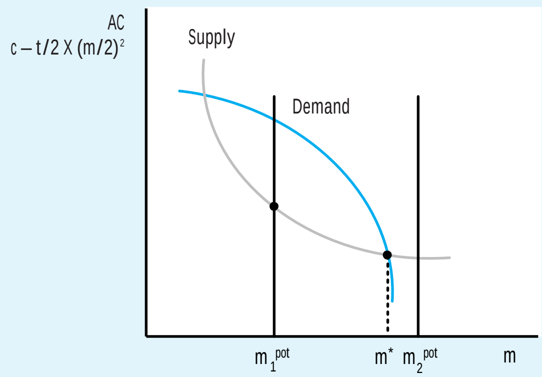
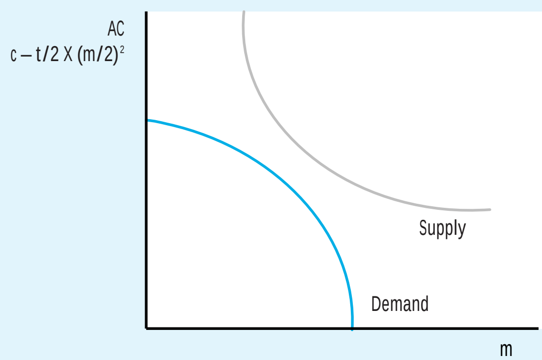


Figure 4

A Case in Which No Credit Union Exists



union *i* attracts is therefore $2r^*$, which we denote m^*.

Because the average cost as a function of the number of credit-union members, *m*, is $(f + vm)/m$, and price must be equal to average cost, we now obtain an expression relating the distance between the marginal member and the credit union, r^*, and the optimal number of members in the credit union, m^*:

$$(1) \qquad c - \frac{t}{2} r^{*2} = \frac{f + vm^*}{m^*}.$$

But we know that $r^* = m^*/2$, so we can substitute in equation 1 for r^* to obtain a cubic expression that determines the optimal number of credit-union members, m^*, expressed in terms of the demand-side parameters *c* and *t* (price of commercial-banking services and the travel-cost parameter, respectively), as well as the supply-side parameters *f* and *v* (fixed and marginal costs of credit-union production of financial services, respectively):

$$(2) \qquad c - \frac{t}{2}\left(\frac{m^*}{2}\right)^2 = \frac{f + vm^*}{m^*}.$$

The economic interpretation of Equation 2, the optimality condition, is straightforward. The left-hand side represents the demand curve for credit-union services, while the right-hand side represents the average-cost curve of a credit union. Credit unions are not-for-profit institutions, so their average-cost curves also are their supply curves. A downward-sloping supply curve indicates that scale economies exist in the range we consider (see Figure 3). For $m^{pot} > m^*$, where m^{pot} is the potential membership of the credit union, we obtain an interior solution. In other words, the participation rate—the fraction of the potential membership that chooses to join the credit union—is lower than one. For $m^{pot} \leq m^*$, on the other hand, the participation rate is equal to one because all potential members choose to join.

Notice that, if the domain of potential members of a credit union is too small, the supply and the demand curves may not intersect. In this case, the credit union cannot operate because there is no positive number for m^* that satisfies equation 2 (see Figure 4).

The second period of the model corresponds to a regime in which the law allows credit unions to serve groups of households united by different common bonds (e.g., employees of both firms A and B). A new management team must be selected at the beginning of period 2 to operate each credit union. New credit unions may be formed in which multiple occupational groups are included. In addition to single-employer credit unions, we now might see four other combinations of common bonds—

a multiple-group credit union encompassing employees of firms:

- A and B (what we term an "A&B" credit union) plus a single-group credit union serving employees of firm C;
- A and C (an "A&C" credit union) plus a single-group credit union serving employees of firm B;
- B and C (a "B&C" credit union) plus a single-group credit union serving employees of firm A; and
- A, B, and C (an "A&B&C" credit union).

As in period 1, none of these credit unions necessarily exists; the particular configuration of parameters will determine the outcome.

We allow for side payments so there is no path dependence as we go from period 1 to period 2 (i.e., period-2 results do not depend on period-1 outcomes). As in period 1, the socially optimal combination of occupational groups is chosen. Also, the new credit unions will again be located in the center of the preference spectrums of their potential members. Before voting, all potential members of the various credit unions observe the (random) technology the new credit unions possess. These are drawn anew at the beginning of period 2. We allow the fixed costs of a multiple-group credit union to deviate systematically from the fixed costs of a single-group credit union. The fixed costs of a multiple-group credit union *i* amount to $f_i = f_a(1+\alpha) + f_b e_i$, with $\alpha > -1 - f_b e_i / f_a$.

After the new credit unions have been established, each household either purchases one unit of banking services from the credit union or it buys them from a commercial bank. The economy ends after period 2.

Finally, we point out several comparative-static features of the model that follow standard intuition despite the existence of a downward-sloping supply curve. The two important demand-side parameters are *t*, the households' travel-cost parameter, and *c*, the cost of alternative banking services as provided by a commercial bank. Recalling Figure 3, which shows the supply curve of the credit union and the demand curve for its services, it is clear that as a household's travel costs rise, the demand for credit-union services declines. We interpret rising travel costs as an increase in the strength of a household's preferences for its ideal bundle of banking services. Such an increase causes the demand curve to shift downward, decreasing m^*. That is, the optimal size of the credit union declines. On the other hand, an increase in the price of commercial-bank provided financial services, *c*, shifts the demand curve up. This has the opposite effect on the optimal size of the credit union, increasing m^*.

The important supply-side parameters of the model are *f*, the credit union's fixed cost, and *v*, the credit union's marginal cost of providing banking services. An increase in *f* pushes up the supply curve of credit-union services, with the sharpest increase at low levels of membership. An increase in the marginal cost of credit-union production also translates into an upward shift of the supply curve. In both cases, the size of the potential membership required to achieve full participation increases.

Simulation of the Model

We simulate the model by drawing repeatedly (10,000 times) a set of five uniformly distributed random numbers from the [0, 1] interval. The first draw determines the length of the segment containing households with an employee of firm A. Recall that the length of the segment containing firm-B households equals one minus the length of the firm-A segment. The second draw determines the length of the segment containing households with an employee of firm C. This determines the length of the line segment that is covered by two households. The last three random numbers enter the three (potential) credit unions' cost functions as stochastic elements of their fixed costs (denoted e_i in the model description above, $i = A, B, C$). These random elements in the credit unions' cost functions ensure that a "conglomerate" credit union consisting of the employees of all three firms is not degenerate—i.e., existing with probabilities of either zero or one.

Recall that in the first period, all credit unions must consist of a single common bond. The first step in the formation of a credit union is a vote by the potential membership on the management team. Since side payments are allowed, the team that minimizes the sum of the travel costs of all potential members—i.e., which picks the most central location—will win. In a second step, all households decide whether to become members or to purchase financial services from a commercial bank.

We calculate the preferred outcome for each group of households in turn (A, B, and C). The equilibrium solution for each employee group must be one of three possibilities: the credit union exists at a corner solution, in which all households participate; the credit union achieves an interior solution with a participation rate less than one; or the credit union does not exist. To compare the various outcomes, we calculate a welfare index for each group of households that is simply the sum of the production costs of the credit union (if it exists), the travel costs incurred by households that use credit unions, and the expenditures of households that obtain financial services from a bank, all multiplied by negative one (to maintain the convention that a higher index value signifies higher social welfare):

$$(3)\quad -(f + mv) - \int_0^{r^*}\int_0^{r^*} tr\, dr\, dr - (m^{pot} - m)c.$$

In the second period, multiple-group credit unions are allowed. We iterate through the possible combinations by first allowing mergers between two given credit unions and forcing the third to operate independently (if it exists). Then we allow all three credit unions to merge. In each regime, households vote on the management team (i.e., choose the credit union's location). In particular, households choose between a team that would operate the credit unions independently and a team that would merge them. Because bribing is allowed, the team that maximizes the welfare index over all potential members will win. It is possible that a stand-alone credit union that could not exist on its own becomes part of a multiple-group credit union. The reason is that the post-merger credit union is able to spread its fixed costs over a larger membership. It also is possible that a credit union that could not exist on its own also is not viable as part of a multiple-group credit union. On the other hand, any employee group that is served by a credit union in period 1 also will be served by a credit union in period 2 because all mergers must be welfare-enhancing. That is, all options for operating credit unions with single common bonds available in period 1 still are possible after permitting multiple common bonds in period 2.

Table 1 displays a summary of the simulation results. The table presents two measures of credit-union activity: the fraction of all employee groups served by a credit union and the fraction of households served by a credit union. When only single-employer credit unions are allowed (period 1), only 6 percent of the 30,000 simulated employers (A, B, and C in each of 10,000 simulations) actually sponsor a credit union and only 4 percent of households actually belong to credit unions. Among households that are eligible to join a credit union, some 50 percent do so. All other households use commercial banks to obtain financial services. We have chosen parameter values to reflect the fact that single-group credit unions are relatively small and may not be viable for many employee groups.

The bottom part of Table 1 presents results when multiple-group credit unions are allowed (period 2). It is clear that the permissibility of multiple common bonds dramatically increases the viability of credit unions. This is a general result in the sense that restricting credit union membership to one employee group is a binding constraint, the relaxation of which may increase the beneficial role of credit unions for employees. When two employee groups may be combined in a single credit union (A and B, A and C, or B and C), the fraction of employee groups in the economy served by a credit union rises to between

Table 1

Simulation Results (1)

Credit Unions in the Economy		Fraction of employee groups served by… (1) Any Credit Union	(2) A Single-Group Credit Union	(3) A Multiple-Group Credit Union	Participation rates as a fraction of…[1] (4) All Households	(5) Those Households Eligible to Join a Single-Group Credit Union	(6) Those Households Eligible to Join a Multiple-Group Credit Union	(7) Those Households Eligible to Join a Credit Union (weighted average of 5 and 6)
Period 1	Welfare index							
Only single-group credit unions	–23,951.89	0.06	0.06	----	0.04	0.50	----	0.50
Period 2	Increase in welfare index							
A&B, C	53.09	0.14	0.04	0.10	0.04	0.50	0.31	0.40
A&C, B	1,048.20	0.50	0.03	0.47	0.37	0.48	0.77	0.74
B&C, A	523.49	0.34	0.04	0.30	0.16	0.45	0.60	0.46
A&B&C	1,123.34	0.49	0.01	0.48	0.30	0.58	0.41	0.42
Optimal Combination	1,590.32	0.94	0.02	0.93	0.43	0.54	0.56	0.56

Parameter values: $f_a = 0.1$; $\alpha = 0$; $f_b = 0.1$; $t = 22$; $c = 1.6$; $v = 1$.

[1] Participation rate is the fraction of eligible households that belongs to a credit union. Rates are weighted by segment lengths.

14 and 50 percent, while the fraction of households served by a credit union rises to between 4 and 37 percent, depending on the combination. When all three employee groups are allowed to combine in a single credit union (A and B and C), the fraction of employee groups served by a credit union jumps to 49 percent, although only 30 percent of households are still served.[4]

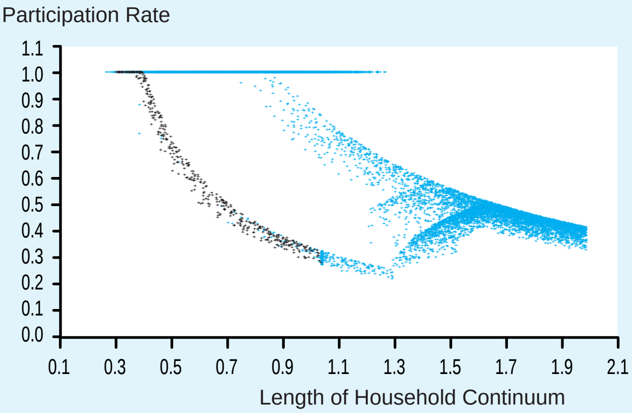
Examination of column 6 indicates that multiple-group credit unions comprising groups A and B or A, B, and C are characterized by relatively low participation rates. This reflects the fact that many members of employee groups A and B are located far from any multiple-group credit union, reducing their incentive to join. The credit union formed by employee groups A and C alone, on the other hand —which are located back-to-back—is characterized by a very high participation rate (77 percent of those eligible actually join). This is because the preferences of these two groups overlap.

In general, how likely households are to join credit unions does not depend primarily on whether multiple-group credit unions are allowed (see column 7, where the exception is the credit union comprising groups A and C, the back-to-back case). In other words, participation rates in multiple-group credit unions are not necessarily higher. Rather, it is the fact that more credit unions are viable when multiple common bonds are allowed that is responsible for the expanded role of credit unions in the economy. A comparison

[4] Household segments not involved in a merger during the second-period simulations face the same economic situation as in the first period. Consequently, they come to the same decision during the second period of whether to operate a credit union as they did during the first period.

Figure 5

Participation Rates as a Function of Potential Membership



of columns 1 and 3 shows that newly viable multiple-group credit unions are indeed the key to greater credit-union access by households, as the lion's share of all credit unions in every possible configuration in period 2 include multiple common bonds.

The final row of Table 1 presents the social optimum, which is the welfare-maximizing combination of single- and multiple-group credit unions that is feasible in the economy. Multiple-group credit unions serve 93 percent of all employee groups in the social optimum, while single-group credit unions serve only 2 percent. Average household participation rates are similar across the two types of credit unions, with the multiple-group average slightly higher.

The averages presented in Table 1 conceal two important features of credit unions in our model, however. Figure 5 is a scatterplot showing the participation rates of all the (optimally formed) credit unions from our 10,000 runs as a function of potential membership. The horizontal scale runs from about 0.1 (the minimum segment length needed to support a credit union under our baseline parameterization) to 2.0 (the sum of two unit-length segments, corresponding to the maximum potential membership of any multiple-group credit union). The two distinct downward-curving sets of points represent the declining participation rates of single-group (the gray points in the lower curve, ending at 1.0) and multiple-group credit unions (the blue points), respectively. If we were to show each type of credit union plotted in Figure 5 in a separate chart, two important features would be obvious from the average participation rates shown in Table 1.

The first feature is that participation rates of multiple-group credit unions tend to lie above those of single-group credit unions for a given number of potential members. This points to the fact that multiple-group credit unions can be closer to the average member's preferences due to the existence of back-to-back households (i.e., households with different employers but identical preferences for banking services). This effect is due entirely to the households in employee group C in our model, whose preferences overlap those of some households in other employee groups, most importantly group A.

Figure 5 also shows the second important feature of the model that is not revealed in the table—the downward slope of both main sets of points. Greater potential membership tends to generate lower participation rates. This always holds for single-group credit unions and for multiple-group credit unions that comprise "horizontally neighboring" membership groups only (i.e., groups A and B). Given the travel costs that represent preference heterogeneity among the potential membership, it is not surprising that credit unions that span a more heterogeneous set of households are able to attract proportionately fewer of them. For credit unions that comprise segments B and C, participation rates initially fall with an increase in the membership base, then rise and later fall again. Credit unions that unite all three employee groups (A&B&C credit unions) exhibit a similar pattern in terms of participation rates. For A&B&C credit unions, a rise in the potential membership is due solely to an increase in the interval spanned by group C, which means an increase in the fraction of households located back-to-back. For back-to-back

Table 2

Simulation Results (2)

Column	Parameter Values α	t	c	Number of times this configuration of credit unions was optimal:[1] (1) A, B, C	(2) A&B, C	(3) A&C, B	(4) B&C, A	(5) A&B&C	(6) Number of employee groups not served by a credit union	Participation rates as a fraction of…[2] (7) All Households	(8) All Households Eligible to Join a Credit Union	(9) Those Households Eligible to Join a Single-Group Credit Union	(10) Those Households Eligible to Join a Multiple-Group Credit Union
(1)	0	22	1.6	746	168	4,452	1,168	3,466	7,959	0.43	0.56	0.54	0.56
(2)	0	22	1.7	302	394	4,580	1,223	3,501	4,782	0.54	0.61	0.64	0.61
(3)	0	24	1.6	926	99	4,436	1,169	3,370	8,793	0.40	0.55	0.50	0.55
(4)	0.05	22	1.6	895	111	4,489	1,148	3,357	8,436	0.43	0.57	0.54	0.57

[1] Based on 10,000 runs.
[2] Weighted by segment lengths.

credit unions of the type A&C, the participation rate increases with a growing membership base if (and only if) the overlap of the intervals spanned by the two segments increases. If (and only if) the overlap shrinks with an increase in the potential membership, the participation rates shrink, too. To sum up, the overall effect of the size of potential membership on the participation rates depends on the relative importance of the various types of multiple-group credit unions.

Table 2 presents comparative-static results for changes in the parameters t, c, and α (the travel cost parameter, the price of banking services of commercial banks, and the parameter in the multiple-group credit unions' fixed costs, respectively). The first row restates the results of the benchmark simulation summarized in the last row of Table 1. Columns 1-5 show the number of times in the 10,000 runs of the simulation that each configuration of credit unions was optimal. The most frequently preferred configuration was a two-group credit union comprising employee groups A and C (column 3), the back-to-back solution. In this configuration, employees of firm B were sometimes served by a credit union and sometimes not; the feasibility of a credit union for employee group B depends on the size of the membership base and the random technology of the potential credit union. The next most frequently preferred configuration involved a three-group credit union. Across all simulations, almost 27 percent of employee groups were left unserved by credit unions even though all mergers were chosen optimally (this figure is calculated from column 6, which is divided by the total number of employee groups in the simulation, 30,000). It is apparent that participation rates of multiple-group credit unions (column 10) are dragged down primarily by the relatively low participation rates in the credit unions with horizontally neighboring groups (i.e., credit unions A&B and A&B&C; recall the result from column 6 of Table 1).

The first comparative-static exercise we performed is summarized in the second row of Table 2. When the price of financial services offered by commercial banks rises, the fraction of employee groups as well as the fraction of households served by credit unions increases, as expected. From column 6 we know that only 16 percent of employee groups have no credit union after the higher cost of bank-provided services is imposed, while only 46 percent of households use a commercial bank (down from 57 percent in the benchmark case; see column 7). Interestingly, all of the multiple-group credit unions increasingly are preferred when banking services become more costly, while only the single-group credit unions become less likely to be optimal. A higher price for bank-provided services is predicted by higher banking concentration in the structure-conduct-performance paradigm, and our comparative-static result demonstrates that credit unions are indeed likely to benefit from more concentrated banking markets.

The second comparative-static result we computed is summarized in the third row of Table 2. When the cost of travelling to a credit union is increased—intuitively, when preferences for banking services become more idiosyncratic or strongly held—both the fraction of employee groups served by credit unions and the participation rate of households decline (columns 6 and 7-10, respectively). Compared to the benchmark case, the number of single-group credit unions in optimal configurations increases (column 1). On the other hand, multiple-group credit unions appear somewhat less attractive (columns 2-5).

The last row of Table 2 displays our third comparative-static result. When the fixed costs of production are systematically higher for multiple-group credit unions than for single-group credit unions, the formation of multiple-group credit unions is less advantageous. Relative to the benchmark case displayed in row 1, fewer groups of employees are served by credit unions (column 6). When they exist, credit unions with multiple-group charters have higher participation rates than in the benchmark case (column 10), which is due mainly to the higher representation of pure back-to-back credit unions (column 3). This leaves the overall household participation rate unchanged at the reported two-digit level of precision (column 7).

Taken together, the comparative-static results in Table 2 indicate that the optimal configuration of credit unions in the economy is sensitive to model parameters such as the market price of bank-provided financial services, the intensity of preferences for specific bundles of banking services, and the potential extra costs associated with multiple-group charters.

Hypotheses

We are now able to state several testable hypotheses that involve the determinants of the participation rate and the average operating costs (i.e., the cost ratio). First, we focus on participation rates at credit unions. Our maintained hypothesis is that a credit union is more successful in providing services to its constituency the less heterogeneous is its membership. This leads to our first testable hypothesis:

- HYPOTHESIS 1. A credit union's participation rate falls with the number of its potential members, all else held constant.

Another hypothesis concerns the effects of local banking-market conditions on credit-union participation rates:

- HYPOTHESIS 2. A credit union's participation rate rises with the level of concentration in the local banking market, all else held constant.

Next we investigate the validity of our maintained assumption that credit unions face scale economies in production:

- HYPOTHESIS 3a. A credit union's cost ratio falls with the number of its potential members, all else held constant.

- HYPOTHESIS 3b. A credit union's cost ratio falls with its level of total assets, all else held constant.

Related questions include the effect of multiple-group charters on the cost ratio and the participation rate. Neither the model nor the simulation address the relationship between multiple common bonds and the cost ratio. As for the impact of multiple common bonds on the participation rate, the model and simulation results are ambiguous. Consequently, it is purely an empirical question how multiple-group credit unions affect operating costs and participation rates, holding all else constant.

DATA AND EMPIRICAL METHODS

We examine a subset of all federally chartered and federally insured occupational credit unions in 1996 (see the appendix for details on construction of the dataset and the variables we use). Table 3 provides a breakdown of our sample according to the type of membership group characterizing each credit union. The table distinguishes between credit unions with a single common bond and those with multiple common bonds. Credit unions sponsored by a single educational institution, for example, numbered 299 in our sample. Credit unions with a membership comprising multiple common bonds, most of which were educationally oriented, numbered 469, and so on for the other membership types. Overall, 1,980 credit unions in our sample had a single common bond (41.8 percent of the sample) while 2,753 credit unions had multiple common bonds among the membership (58.2 percent).

In addition to data on individual credit unions, we collected three types of environmental variables. To control for differences in local economic conditions, we gathered levels and computed growth rates of real gross state product for each state. Measures of economic activity may capture systematic differences in demand for credit union services that we have not modeled explicitly. We also calculated the Herfindahl index of concentration of bank deposit shares in each credit union's local banking market, since concentration measures often are used to control for differences in the competitiveness of local markets. The index is calculated as the sum of the squared market shares of all participants in each local market. Third, we collected data on population density by county, which might be another factor

Table 3

Distribution of Credit Unions by Type Of Membership (TOM)

Number of Credit Unions	TOM Codes [1]	Type of Membership
299	4	Educational
37	5	Military
392	6	Federal, state, local government
744	10-15	Manufacturing
508	20-23	Services
469	34	Multiple group – primarily educational
124	35	Multiple group – primarily military
621	36	Multiple group – primarily federal, state, local government
821	40-49	Multiple group – primarily manufacturing
718	50-53	Multiple group – primarily services

Total: 4,733

[1] National Credit Union Association (NCUA), Instruction No. 6010.2, July 28, 1995.

in credit unions' competition with commercial banks.

We use a semiparametric model to allow the influence of the number of members and total assets on the dependent variable to be nonlinear. The parametric part of the model contains independent variables whose effects may be approximately linear, such as the Herfindahl index. In particular, we use a semiparametric model of a credit union's participation rate of the form:

$$(4) \qquad y_i = x_{pi} \times \beta_p + f(x_i) + \varepsilon_i, \quad i = 1, \ldots, n,$$

where y_i is the i-th observation of the dependent variable; x_{pi} is a row vector consisting of the i-th observation of the explanatory variables of the linear (parametric) part of the model; β_p is a column vector of the parameters of the linear part of the model; x_i is a vector consisting of the i-th observation of the explanatory variables in the nonparametric part of the model; and (ε_i is the i-th realization of the error term. For details on this econometric approach, see the appendix in Emmons and Schmid (1999).

Our hypotheses are framed in terms of two different dependent variables, namely: 1) PARTICIPATION, the participation rate of those eligible to join the credit union,

Table 4

Descriptive Statistics[1]

	Minimum	Median	Mean	Maximum	Standard Deviation
Participation Rate (PARTICIPATION)	3.050×10^{-2}	6.246×10^{-1}	6.142×10^{-1}	1	2.139×10^{-1}
Cost Ratio (COST)	6.268×10^{-3}	3.897×10^{-2}	4.088×10^{-2}	4.169×10^{-1}	1.739×10^{-2}
Total Assets	4.300×10^{4}	6.231×10^{6}	3.300×10^{7}	8.922×10^{9}	1.652×10^{8}
Number of Members	4.500×10^{1}	1.865×10^{3}	6.833×10^{3}	1.601×10^{6}	2.860×10^{4}
Number of Potential Members	7.500×10^{1}	3.193×10^{3}	1.432×10^{4}	2.032×10^{6}	5.540×10^{4}
Herfindahl Index	5.346×10^{-2}	1.966×10^{-1}	2.080×10^{-1}	1	9.469×10^{-2}

[1] 4,733 observations.

defined as the number of actual members divided by the number of potential members as specified in the credit union's charter; and 2) COST, the credit union's total operating expenses divided by total assets. There are four independent variables of interest: the number of members (or the number of potential members when we examine participation rates); total assets (for the COST regression); the Herfindahl index of local bank-deposit concentration (HERF); and the indicator variable MULTGROUP, which is equal to one if the credit union has a multiple-group charter and zero otherwise.

Membership (or potential membership) and total assets are included in the nonparametric part of both regression approaches. They are in logarithmic form and—to avoid simultaneity problems—are lagged by one period. The parametric part of the model includes the other two variables of interest, HERF and MULTGROUP. The parametric part also contains the following control variables: the credit union's home state's real gross state product per capita (REALGSPPC) in the PARTICIPATION regression; the log growth rate of real gross state product (GRREALGSP) in the COST regression; and indicator variables corresponding to the credit union's primary field of membership (in both regression approaches). Fields of membership include educational, military, government, manufacturing, and services. Because there is a constant included in the nonparametric part of the regression equation, we must drop one of the membership indicator variables; we chose the educational indicator variable for exclusion.

The variable REALGSPPC in the PARTICIPATION regression controls for preferences for banking services as they may vary with real income. In the COST regression, GRREALGSP serves as a measure for real growth, which is a main factor in the capacity utilization of credit unions. In the PARTICIPATION regression, the Herfindahl index is lagged to avoid simultaneity problems that may arise from the interaction between credit-union participation rates and concentration in the local banking market.

Table 4 presents descriptive sample statistics for the dependent and some of the independent variables. The participation rate among sample credit unions ranged from 3 percent to 100 percent, with the median at 62 percent. The median cost ratio was 3.90 percent, with a range of 0.63 to 41.70 percent of assets. Although this range may contain some extreme values, we retain all observations because all of them contain information. In addition, our locally weighted regression approach is somewhat robust to outliers. Total assets ranged from $43,000 to $8.92 billion, with the median credit union holding $6.23 million in assets. The number of actual and potential members ranged from 45 to 1.6 million and 75 to 2.03 million, respectively, while median actual and potential membership counts were 1,865 and 3,198, respectively. Finally, Herfindahl indexes in relevant banking markets ranged from 0.0535 to 1.00 with a median value of 0.1966 (recall that the index is defined on the interval ((0,1]).

EMPIRICAL RESULTS

Our results are presented in two sections corresponding to the dependent variable used. The first section discusses results from regressions using credit unions' participation rate while the second section reports results from regressions using credit unions' cost ratio.

Participation Rates

Hypothesis 1 relates the size of a credit union's potential membership to its participation rate. Regressions including PARTICIPATION use (the lagged value of the log of) potential members instead of actual members. Potential members are relevant for evaluating participation rates because all individuals eligible to join constitute the predetermined economic potential that each credit union seeks to exploit.

The series of plots presented in Figures 6-8 are "conditioning plots." The solid lines in Figures 6-8 are point estimates and the dashed lines indicate 90-percent confidence bounds. In each plot, one variable is held at its median value while the other variable (identified on the horizontal axis) is allowed to vary. The graph displays the impact of this independent variable on the level of the dependent variable. In other words, the slope of the graph at a particular point reflects the marginal impact of the independent variable at that point. The intercept is not identified in regressions of this type, so only vertical distances are meaningful (not the level itself). In sum, the key to interpreting these graphs is

Figure 6

Participation Rate – Number of Potential Members

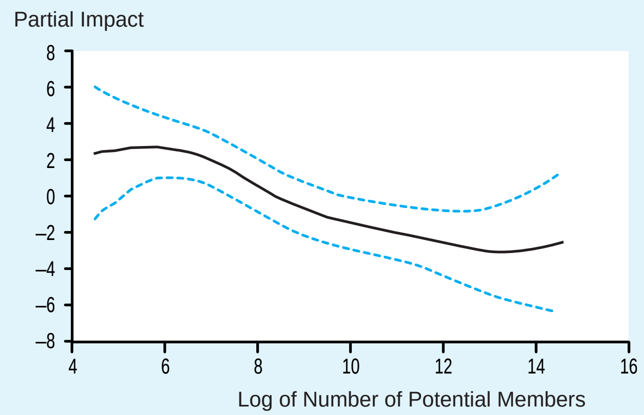


Figure 7

Cost Ratio – Number of Members

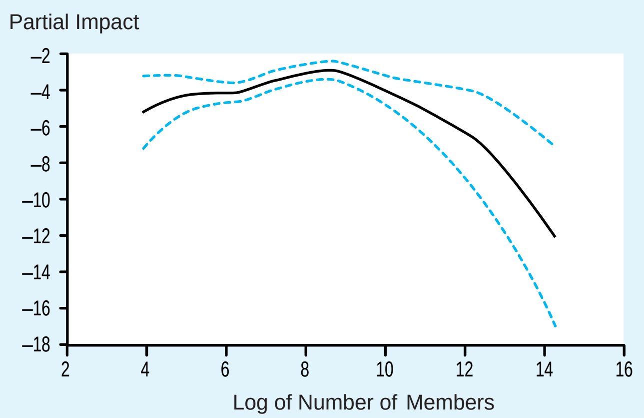


Figure 8

Cost Ratio – Total Assets

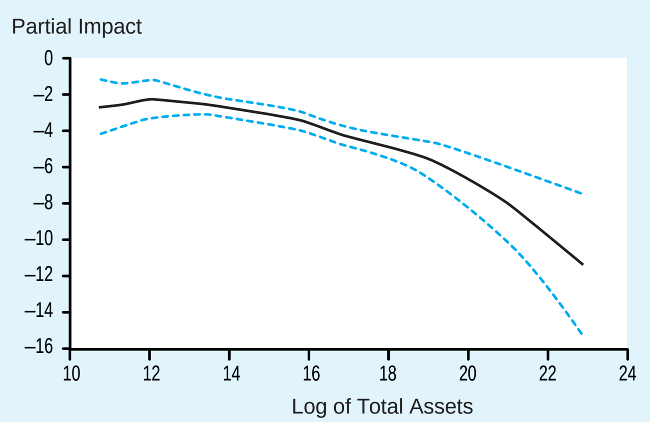


Table 5

Participation Rate

Independent Variable	Coefficient	*t*-statistic
MULTGROUP	8.684×10^{-2}	2.834 ***
HERF (lagged)	3.628×10^{-1}	2.535 **
REALGSPPC	–1.820	–1.090
POPDENS	-3.743×10^{-6}	–1.002
TOM: Military	2.266×10^{-1}	3.004 ***
TOM: Government	6.205×10^{-2}	1.496
TOM: Manufacturing	-2.123×10^{-2}	–0.560
TOM: Services	1.304×10^{-1}	3.284 ***
Number of Observations	4,691	

/* Significant at the 5/1 percent level (two-tailed tests).

to focus on the slope of the curve and on the vertical differences moving along the horizontal axis.[5]

Figure 6 provides evidence supporting Hypothesis 1 (a negative relationship between participation rates and potential membership). The plot supports Hypothesis 1 because the lower confidence bound at small credit unions (small number of potential members) lies above the upper confidence bound for large credit unions, except for the very smallest and very largest credit unions, where the small number of observations widens the confidence intervals. This confirms our findings when we simulated the theoretical model (recall Figure 5) and is consistent with the idea that larger membership pools contain greater heterogeneity of preferences for banking services. This leads to greater differences in the most preferred bundle of banking services between the median member and members in the tails of the preference distribution.

An existing single-group credit union that adds one or more membership groups to its common bond encounters both benefits and costs of expansion. On the one hand, adding membership groups whose preferences are close to those in the existing field of membership (the back-to-back case) may increase the participation rate. This outcome would be predicted as a result of a reduction in average operating costs and hence in the credit union's price of banking services. On the other hand, for a given credit-union charter, a higher number of potential members typically means more heterogeneity and thus a lower participation rate as the travel distance of the members located in the tails of the preference distribution increases.

Table 5 presents results from the parametric part of the model. The results provide evidence on the question of whether credit-union participation rates differ when comparing credit unions with a single common bond to those with multiple common bonds, holding all else equal. Recall from Tables 1 and 2 that credit unions with multiple-group charters will (on average) have higher participation rates than single-group credit unions if back-to-back membership combinations dominate (i.e., if there is a significant overlap of banking preferences among employees of different firms). Table 5 indicates that multiple-group credit unions in our sample indeed have higher participation rates, perhaps reflecting the ability of multiple-group credit unions to capitalize on similar preferences among employees of different firms.

Another interesting result in Table 5 is the positive and significant coefficient on the lagged Herfindahl index of bank deposit concentration in credit unions' local markets (Hypothesis 2). This indicates that, the more concentrated its local banking market is, the higher a credit union's participation rate will be. In other words, credit unions may provide an attractive alternative for consumers who face a relatively uncompetitive local banking market.

Cost Ratio

Hypotheses 3a and 3b refer to tests of an important maintained assumption of our model, namely, that a credit union's operating expenses should decline with an increase in its scale of operation. We also would like to know whether serving multiple-group memberships is costly.

[5] See Cleveland and Devlin (1988).

Figure 7 plots COST against the number of members, while holding the level of total assets constant at its median value in the sample. The plot shows that COST decreases sharply beyond a certain threshold level of membership. For small credit unions, average costs seem to increase slightly with the number of members. While wide confidence intervals indicate that in this range the relationship between average costs and number of members is estimated imprecisely, an initial increase in the average operating costs actually might be supported by the data. For small credit unions, subsidies (such as rent-free office space, volunteer workers, etc.) tend to be relatively more important than for large units. As these subsidies become less important for credit unions with higher numbers of members, measured operating costs might approach shadow operating costs. Overall, the findings support our maintained assumption of declining average costs as the scale of operations increases. Similar evidence is provided in Figure 8, which is generated by the same regression that produced Figure 7. Figure 8 plots the influence of another measure of the size of operations, total assets, on the credit union's average operating costs (Hypothesis 3b) while holding the number of members constant at its median value in the sample.

Table 6 presents results from the parametric part of the model. The results indicate that there is a positive relationship between the existence of multiple membership groups in a credit union and COST. One might think that multiple-group credit unions would have high cost ratios due to agency costs. According to this line of reasoning, as membership groups try to free-ride on each other's monitoring, supervision of management might be inefficiently low. As Emmons and Schmid (1999) show, however, there is no evidence of multiple-group charters causing agency costs.

Finally, the significantly negative coefficient on the Herfindahl index in Table 6 implies that higher levels of bank concentration in a local market lead to lower levels of the cost ratio reported by credit unions.

Table 6

Cost Ratio

Independent Variable	Coefficient	*t*-statistic
MULTGROUP	7.509×10^{-2}	6.847 ***
HERF	-1.813×10^{-1}	–3.427 ***
GRREALGSP	-8.448×10^{-1}	–2.377 **
TOM: Military	1.038×10^{-1}	4.533 ***
TOM: Government	1.149×10^{-1}	8.043 ***
TOM: Manufacturing	1.036×10^{-1}	7.444 ***
TOM: Services	8.915×10^{-2}	6.019 ***
Number of Observations	4,733	

/* Significant at the 5/1 percent level (two-tailed tests).

One possible explanation is that less intense competition from banks allows credit-union managers to enjoy a "quiet life." For example, credit unions may be able to attract or retain members with lower marketing efforts or lower quality services than would be the case in a more competitive market. On the other hand, the quiet life that comes with less competitive markets might allow greater scope for managerial agency costs. If the latter were the case, however, it would generate predictions opposite to our empirical findings of lower average operating expenses.

CONCLUSIONS

Continued expansion of credit unions has been accompanied by public debate and courtroom confrontations. Advocates argue that credit unions provide needed competition to banks and thrifts in local markets for retail financial services. Opponents, including most notably the banks and thrifts themselves, point to various subsidies to credit unions that create an unlevel playing field. Previous research findings do not provide unambiguous conclusions favorable to either camp, while recent federal legislation favorable to credit-union expansion merely has intensified the debate. More research into the fundamental operation of credit unions is needed.

In this article, we investigate the relationships between several features of credit unions, namely the number of members in a credit union, the amount of total assets on its balance sheet, and the existence of single and multiple common bonds among its membership, on the one hand, and two measures of credit-union effectiveness, on the other. We also examine the effect of several environmental variables, including economic conditions and banking concentration in the local market, on credit-union operations.

We find that a larger potential credit-union membership translates into lower credit-union participation rates. Credit unions with multiple common bonds, holding all else constant, have higher participation rates. We also find evidence that credit unions in more concentrated banking markets exhibit higher participation rates.

While greater asset size appears to be associated with lower average operating costs, holding all else equal, we find that a larger number of members is associated with a lower cost ratio only for larger credit unions. Thus, asset size and the size of the membership are distinct aspects of credit-union operations. Multiple-group credit unions have higher costs on average, all else equal. We also find that credit-union cost ratios are lower in more concentrated banking markets, perhaps indicating that credit unions can economize on marketing or service provision when competition from banks is less intense.

Our findings are particularly interesting in light of the recent AT&T Family Credit Union case decided by the Supreme Court in February 1998, and its sequel in the U.S. Congress that culminated in the Credit Union Membership Access Act of August 1998. This new federal legislation upholds the right of federally chartered credit unions to grow under an expansive definition of the common-bond requirement. The new law allows multiple groups of members to belong to a single credit union as long as the members of each group are united by a common bond. This statute therefore upholds regulatory actions taken in recent years and overturns the Supreme Court's narrow reading of the 1934 Federal Credit Union Act restricting a federal credit union to a single common bond.

REFERENCES

American Bankers Association. *The Credit Union Industry: Trends, Structure, and Competitiveness* (Washington, D.C., 1989).

BNA Banking Report. "House Passes Credit Union Bill; Clinton Wastes No Time Signing It," August 10, 1998, Vol. 71, No. 6.

Cleveland, William S., and Susan J. Devlin. "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association* (September 1988), pp. 596-610.

Emmons, William R., and Frank A. Schmid. "Wages and Risk-Taking in Occupational Credit Unions: Theory and Evidence," this *Review* (March/April 1999), pp.13-31.

Hotelling, Harold. "Stability in Competition," *Economic Journal,* (March 1929), pp. 41-57.

Kane, Edward J., and Robert Hendershott. "The Federal Deposit Insurance Fund that Didn't Put a Bite on U.S. Taxpayers," *Journal of Banking and Finance* (September 1996), pp. 1305-27.

Milgrom, Paul, and John Roberts. *Economics, Organization and Management,* Prentice Hall, 1992.

Supreme Court. "National Credit Union Administration, Petitioner, v. First National Bank & Trust Co., et al.; AT&T Family Federal Credit Union, et al., Petitioners, v. First National Bank and Trust Co., et al." Decided Feb. 25, 1998. Nos. 96-843, 96-847. 118 S. Ct. 927.

U. S. Treasury Department. *Credit Unions,* U.S. Government Printing Office, 1997.

Appendix

DATASET AND VARIABLES

The Dataset

We analyze a dataset comprising all federally chartered and federally insured credit unions during the year 1996. The dataset was obtained from the Report of Condition and Income for Credit Unions (NCUA 5300, 5300S), produced by the National Credit Union Administration (NCUA). These reports are issued semiannually in June and December. We used the December data. The flows in the December income statements include the entire year of 1996.

We concentrate on the following Types Of Membership (TOM) groups among occupationally based credit unions: educational; military; federal, state, and local government; manufacturing; and services. This means that we do not include community credit unions, associational credit unions, or corporate credit unions. Lists of TOM classification codes are from the NCUA (Instruction No. 6010.2, July 28, 1995).

We excluded observations for any of the following reasons:

- Missing TOM codes.
- Activity codes other than "active."
- Number of members or of potential members not greater than one; applies to actual and to lagged values.
- Nonpositive values for total assets or lagged total assets.
- Zero number of employees.
- Zero value for "employee compensation and benefits."

Total assets, number of members, potential number of members, and the Herfindahl index were all lagged one year (i.e., 1995 values). All other observations are from year-end 1996.

We calculated county-specific Herfindahl indexes as measures of concentration of the local banking market. A Herfindahl index is defined as the sum of squared market shares. We measured market shares by the fraction of total bank deposits (as of June 30) within a county based on FDIC Summary of Deposits data. These data are available online at <http://www2.fdic.gov/sod/>.

We used either the log level of Real Gross State Product (REALGSP) or its log growth rate to control for cross-sectional differences in macroeconomic conditions facing credit unions. The REALGSP data are in millions of chained 1992 dollars. We obtained the data from the U.S. Department of Commerce, Bureau of Economic Analysis, Regional Economic Analysis Division. The data are available online at <http://www.bea.doc.gov/bea/regional/data.htm>.

Population density at the county level was calculated by dividing the total county population by the total land area of the county (in square miles). Both the county population and land area data were obtained from the U.S. Census Bureau <http://www.census.gov>. The population data are Census Bureau estimates as of July 1, 1996. The land area measurements are from the 1990 census.

Definition of Variables

We transformed the dependent variables in some cases to ensure that they are not bounded. These transformations are necessitated by the assumption of normally distributed error terms. For variables that are restricted to the positive orthant of real numbers, we substitute their natural logarithms. For variables expressed as fractions (i.e., restricted to the interval [0,1]), we applied the logit transformation $log(y/(1-y))$, with *log* being the natural logarithm. In this case, observations equal to one were eliminated from the set of observations; there were no cases in which the transformed variable equaled zero.

Definitions of variables and underlying data sources are listed below. For data taken from the Report of Condition and Income for Credit Unions, produced by the National Credit Union Administration, the relevant item numbers are in brackets.

Dependent Variables. We employed two dependent variables in the regressions:

1) Participation Rate (PARTICIPATION): Number of actual credit-union members [CUSA6091] divided by the number of potential members [CUSA6092]. In the regressions, we use the logit transformation $log(y/(1-y))$. No zero values for the number of members occurred. Forty-two cases of full participation ($y=1$) were eliminated from the dataset for these regressions only.
2) Cost Ratio (COST): Total operating expenses [CUSA4130] divided by total assets [CUSA2170]. In the regression, we use log values.

Independent Variables. When total assets (measured in units of one dollar), the number of members, or the number of potential members served as regressors, they were lagged by one period and transformed into natural logarithms.

1. MULTGROUP: equal to one if the credit union has multiple groups; zero otherwise.
2. HERF: Sum of squared market shares of commercial banks within a county based on total bank deposits. By definition, the Herfindahl index is greater than zero; its maximum value is one.
3. REALGSPPC: Real gross state product per capita (chained 1992 dollars).
4. GRREALGSP: Logarithmic changes in the real gross state product (chained 1992 dollars).
5. POPDENS: Population Density, people per square mile in each local banking market.
6. TOM code variables: equal to one if the credit union is of a specific type (educational, military, government, manufacturing, or services). Because we use an intercept in (the nonparametric part of) the regression, the TOM code variable for the educational credit union was dropped.