Outsourcing, firm size, and product complexity: Evidence from credit unions

Is there evidence of the new economy in U.S. GDP data?

The cost of business cycles and the benefits of stabilization

A stable money demand: Looking for the right monetary aggregate
## Contents

First Quarter 2005, Volume XXIX, Issue 1

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Outsourcing, firm size, and product complexity: Evidence from credit unions</td>
<td>Yukako Ono and Victor Stango</td>
<td>Outsourcing business services is a key concern in the modern economy. Focusing on data processing services for credit unions from 1994 to 2003, the authors find that both credit union size and the diversity of their product offerings influence the propensity to outsource. The results suggest that simple scale-economy-based explanations for outsourcing may be inadequate.</td>
</tr>
<tr>
<td>12</td>
<td>Is there evidence of the new economy in U.S. GDP data?</td>
<td>Michael A. Koupapritosas</td>
<td>This article tests whether the trend growth rate of U.S. GDP changed significantly over the “new economy” period from 1996 to 2003. Based on estimates from widely used methods of trend/cycle decomposition, the author finds that the trend growth rate of GDP was not significantly higher over this period. This suggests that the U.S. was the same old economy in the latter half of the 1990s.</td>
</tr>
<tr>
<td>32</td>
<td>The cost of business cycles and the benefits of stabilization</td>
<td>Gadi Barlevy</td>
<td>This article reviews the social cost of U.S. postwar business cycle fluctuations, first calculated by Lucas (1987). Recent work suggests this cost is considerably larger than suggested by Lucas. Despite this, the author argues that it is not obvious that policymakers should have pursued a more aggressive stabilization policy over the postwar period. Still, because volatility is costly, stable growth remains a desirable goal.</td>
</tr>
<tr>
<td>50</td>
<td>A stable money demand: Looking for the right monetary aggregate</td>
<td>Pedro Teles and Ruilin Zhou</td>
<td>A money demand relationship with M1 as the monetary aggregate holds very well until the mid-1980s but not well after that. This could be because the demand for money is not a stable relationship. The authors’ conclusion is that the measure of money is not a stable measure. Technological innovation and changes in regulatory practices in the past two decades have made other monetary aggregates as liquid as M1. Once an appropriately adjusted measure of money is taken into consideration, the stability of money demand is recovered.</td>
</tr>
</tbody>
</table>
Outsourcing, firm size, and product complexity: Evidence from credit unions

Yukako Ono and Victor Stango

Introduction and summary

Outsourcing involves firms’ choosing to procure goods or services from other firms rather than producing them internally. For example, firms can outsource accounting and other business services to service providers or maintain internal departments to meet these needs. An automobile manufacturer can design and produce parts internally or outsource by relying on suppliers for production, design, manufacturing, or some combination of these activities. The choices that firms make regarding outsourcing have increasingly attracted the attention of the media, policymakers, and researchers. This attention stems in part from the fact that outsourcing has become increasingly global in scope, meaning that firms that outsource are often moving production and jobs across international borders. In addition, a growing number of researchers in recent years have identified outsourcing as a key determinant of firm profitability and, therefore, a key component of business strategy. Competitive pressure continually drives firms toward more efficient production. Because outsourcing helps firms to achieve this goal, understanding the drivers of outsourcing improves our understanding of business strategy.

Like any critical business decision, the decision to outsource production or services has benefits and costs. By outsourcing, small firms use more efficient suppliers that can supply goods or services at lower cost. These suppliers are often larger than their clients and have economies of scale that smaller firms could not achieve with in-house production. Lower costs may also result from competition among suppliers in their product markets, providing firms that outsource with multiple options. At the same time, outsourcing imposes transaction costs of writing and enforcing contracts with suppliers. Such benefits and costs of outsourcing would depend on firm characteristics, the suppliers’ industry structure, and the nature of the outsourced function.

In this article, we shed some light on the determinants of outsourcing by studying outsourcing practices of credit unions (CUs). Using data from the National Credit Union Administration (NCUA), we examine the outsourcing practices of CUs in their data processing (DP). Data processing is a critical information management function throughout the financial services industry, as it is in many other industries. Our data are unique in that they contain rich information both on CUs’ DP choices and a number of other firm-level characteristics. This allows us to explore questions that have received relatively little attention from researchers. In particular, we focus on examining how CUs’ decisions to outsource are associated with firm size and the diversity of their product offerings.

Firm size may be important because it affects the scale at which a firm can produce internally if it chooses not to outsource. Scale economies are widely held to influence firms’ outsourcing decisions, particularly for functions that have relatively high fixed costs. Many technology-based functions, such as data processing, fall into this category because they impose significant fixed hardware, software development, and training costs. This suggests that smaller firms should outsource more to take advantage of scale provided by specialized DP vendors. On the other hand, larger firms may have more bargaining power with vendors, rendering them more likely to enter relationships with suppliers. This will be particularly true if large customers make up a significant fraction of a given supplier’s business (Besanko, Dranove, and Shanley, 1996).

Yukako Ono is an economist at the Federal Reserve Bank of Chicago. Victor Stango is an associate professor in the Tuck School of Business at Dartmouth College. The authors wish to thank Thomas Hubbard, Matthew Nixon, Mulyil Shridhar, Craig Furfine, and Tara Rice for helpful comments, and Carrie Jankowksi for excellent research assistance.
We also investigate the relationship between outsourcing and the product offerings of CUs. CUs offer a wide array of financial services, with specific offerings varying across and within firms. Offering a greater number of products may have two effects on the decision to outsource. First, if there are fixed costs associated with offering an individual product, greater product diversity may change the fixed costs of internal production. This may change the scale economies of internal versus external production. A second effect of product diversity may be an increase in the complexity of the firms’ DP requirements. In the literature on transaction cost economics, which we discuss in more detail below, product complexity is considered a primary influence on firms’ decisions to outsource (Masten, 1984). The relationship between complexity and outsourcing is that more diverse product offerings create a greater number of contingencies regarding future vendor-firm interactions. This makes contracting costly and discourages outsourcing.

Using the data from NCUA, we try to estimate the relationship between firm size, product diversity, and outsourcing. Our empirical results show that our two measures of interest—CU size and product diversity—both affect the propensity to outsource. Moreover, they also interact in interesting ways; the relationship between diversity and outsourcing is not simple. Up to a point, greater diversity is associated with more outsourcing, but firms with the greatest product variety are less likely to outsource. This suggests that the countervailing factors affecting outsourcing change in importance with firm size and product diversity.

Beyond these relationships, CU size and product diversity are also linked, with larger CUs offering more products. Holding the number of products constant, we find that for small/medium size CUs, diversity is associated with more outsourcing. For large CUs, diversity is associated with less outsourcing.

**The economics of outsourcing**

In its simplest form, the decision to outsource depends on the relative costs of internal versus external production for a given input. The firm chooses internal production if its net benefits exceed those associated with external production.

Theories of outsourcing attempt to explain firms’ decisions by modeling the factors that affect costs of internal and external production. If production involves significant scale economies, both internal and external production should become cheaper in average cost terms as the size of the producer increases. In general, however, the scale of internal production is limited by other constraints on firm size. This implies that smaller firms are more likely to outsource, because they can rely on scale economies provided by external producers.

Competition in suppliers’ markets also encourages outsourcing. Internal production may not be subject to market discipline because internally produced inputs are not sold in competitive markets. Thus, internal production may be inefficient. A related problem would arise if managers and workers associated with internal production were not compensated in a manner aligned with profit maximization at the firm level or were difficult to monitor and could shirk. In such cases, outsourcing might result in lower input prices. This competitive effect may even make the costs of internal production substantially higher than the costs of using inputs purchased through outsourcing.

**Transaction cost economics and outsourcing**

The theory of the firm literature (Coase, 1937; Williamson, 1975; and others) suggests that while outsourcing is beneficial to many firms because markets have advantages over internal production, it may also be undesirable because market transactions impose costs in some cases. The problem arises when the transaction involves relationship-specific investment: sunk (unrecoverable) costs of entering an outsourcing relationship with a specific vendor. When a transaction involves relationship-specific investment, once the two parties have committed to the relationship, it is possible that one or both parties may try to demand more out of the transaction than was originally agreed upon, taking advantage of the fact that the other party has already made an investment specific to the transactional relationship and so is unlikely to withdraw. This is often referred to as a hold-up problem. As Williamson (1975) notes, each party may use the threat of not trading to appropriate rents from the other; these rents will be directly related to the sunk costs each side has committed to the business relationship. While these sunk costs encourage parties to remain in business relationships once they have begun, the hold-up problem represents a deterrent to outsourcing. If relation-specific costs are large, internal production may be preferable.

In principle, firms in an outsourcing relationship can write contracts to mitigate the risks of such hold-ups. These contracts specify the relationship between market events and payments made from one party to the other. Contracts also specify how contingencies are handled when information about future events is imperfect. Contracts also may define patterns of asset ownership in the business relationship in order to align firms’ incentives in particular ways.
However, entering contracts may prove costly for two reasons. First, contracts carry transaction costs. These are the costs associated with writing, negotiating, and enforcing contracts. Because these costs are often high, contracts in the real world are often incomplete: They do not effectively cover every possible contingency of the transaction. Thus, there will still be incentives for opportunistic behavior even after the contract is written. Given the risks of such opportunistic behavior, firms may forego market transactions (outsourcing) and handle production internally, even if they cannot do so efficiently relative to the market.

Within this transaction cost economics framework, the factors that make contracting more difficult will deter outsourcing. These include the level of sunk costs associated with the transaction; higher sunk costs create greater scope for hold-up. Greater complexity deters outsourcing, because it increases the cost of writing the optimal contract. This could occur because complex products are associated with a wider number of contingencies for future outcomes. These contingencies could pertain to costs for either party to the contract, demand for the final good, or some other aspect of the business environment. Complexity may also make monitoring of the outside production effort difficult.

**Related literature**

There are not many empirical studies on the relationship between firm size and outsourcing. The most recent and very relevant study is Borzekowski (2004), which also uses the data we use in this article. Borzekowski shows the positive association between CU size and its likelihood of outsourcing the DP system. However, as we show later on, CU size and the diversity of its products are also positively correlated. In this article, we examine whether the positive relationship between CU size and the likelihood of outsourcing persists after we control for the diversity of CU products, as well as how size and product diversity interact.

There are only a few empirical papers that examine the relationship between complexity and outsourcing. Masten (1984) studies input procurement in the aerospace industry, showing that more complex inputs are less likely to be outsourced. Among more recent efforts, Baker and Hubbard (2003) study the choice of shippers to use private (in-house) or for-hire (outsourced) drivers as their carriers and find that market segments where drivers perform complex tasks are more likely to be served by in-house drivers and trucks.

**Credit unions and data processing**

With these ideas in mind, we examine firms’ decision to outsource by using the data from call reports that CUs submit to the National Credit Union Administration (NCUA). The data include information on how CUs procure the automated DP systems to manage the records of their share and loan transactions.

CUs are financial institutions that provide banking services to their members. In principle, they are nonprofit organizations, owned by their members. In many cases, the CU is affiliated with an organization from which it draws members, for example, large companies like Boeing, state agencies, the Navy, and the Pentagon all have CUs offering services to their members. Based on the NCUA call reports, the total number of CU members grew from 65.1 million to 83.6 million between June 1994 and December 2003.

CUs earn income from interest on loans and investments, as well as fees charged for their services (such as overdraft fees, ATM fees, and credit card fees). Such income is spent on interest expenses, such as dividends on shares, interest on deposits, as well as non-interest expenses, including employee compensation, benefits, travel and conference expenses, rent, operations, member insurance (that is, borrower’s protection and share insurance), and outside professional services. Often CUs use net proceeds (income minus expense) to maintain or improve the financial services they offer to members or to expand their operations. In many ways, the structure of the CU industry mirrors that of the commercial banking sector, which represents the CUs’ primary competition. Beyond managing checking and saving accounts, CUs offer a wide array of financial services, including more sophisticated saving and investment options, as well as personal loans and mortgages. Because of their status as nonprofit organization, CUs are entitled to preferential tax treatment.

**Data processing**

Like all financial services providers, CUs need to maintain detailed records of their clients’ transactions. The core data for each customer usually include transaction records associated with checking or savings accounts. Managing other financial products, such as credit cards, personal loans, mortgages, as well as share certificates, increases the complexity of DP requirements. Such data may come into the CU through teller transactions, mail, phone, deposit boxes, and ATMs (automated teller machine), or online. While in principle CUs may track customer data manually (on paper), the vast majority of CUs use some form of computer system to handle their DP.

**Internal versus outsourced DP services**

The efficient way to source DP systems is a key concern for CUs; trade publications (for example, the
periodical Credit Union Tech-Talk) and industry conferences reflect this emphasis by devoting considerable attention to information technology (IT) issues and outsourcing in particular. We focus our analysis on CUs that use some form of automated (computerized) data processing system.\(^4\) Our sample comprises approximately 10,000 CUs, with the overall number declining over the sample period 1994–2003 as CUs merge and exit (see table 1).

Among CUs using automated DP systems, some develop their own in-house and others choose various degrees of outsourcing. In the data, CUs are given three options to specify the type of their procurement of the data-processing system. The first is “credit union developed in-house system,” which is the system developed and operated completely internally. The second is “vendor-supplied in-house system,” which refers to a system in which the CU purchases software from a vendor, but operates hardware and software within the CU. And the third is the vendor-supplied online (VOL) system, which is the most complete form of outsourcing. In this article, we focus on the choice between this most complete form of outsourcing and the alternatives, so the term “outsourcing” from here on indicates the use of VOL.

In the VOL arrangement, the hardware and software used for DP are located off-site at the vendor’s service bureau, which handles DP for many or all of its customers. The connection is made through a telecommunications link connected to terminals in the CU and these terminals may be proprietary terminals supplied by the vendor or Windows-based PCs already owned by the CU (or purchased by the CU). As shown in table 1, about 26 percent to 30 percent of the CUs in our sample choose VOL, with the percentage falling slightly over time.

### Credit union size and DP outsourcing

The VOL system is likely to differ from in-house production in its scale requirements. It involves lower fixed costs, both in terms of software development and hardware. For these reasons, we might expect smaller firms to employ VOL more often than larger firms.

CUs vary widely in terms of size. In December 2003, 14.2 percent of the CUs in our sample had less than $2 million in assets, 14.9 percent had between $2 million and $5 million, 16.2 percent had between $5 million and $10 million, 32 percent had between $10 million and $50 million, and 21 percent had $50 million or more. Table 2 shows the size distribution of CUs employing both in-house and outsourced DP services for December 2003. An interesting pattern emerges. The mean firm size for CUs that outsource is smaller than for those that do not outsource, while median firm size for CUs that outsource is larger than for those that do not outsource. Among the CUs that do not use VOL, there are some that are very big, while many others (about 67 percent) are smaller than the median CU that uses VOL. Size distribution is much tighter for those that use VOL, which also suggests that both very small and very large CUs are more likely to retain in-house DP services.\(^5\)

Many small CUs offer less complex products than bigger CUs, thus requiring a lower-tech DP system (such as Microsoft Excel). If lower-tech DP systems have lower fixed costs, it may be worthwhile for smaller firms to handle DP internally. Of course, small firms could outsource these activities as well. However, if search and transaction costs are lumpy or fixed, it may not make sense to outsource such simple activities for which a supplier may not provide large cost advantages. It also might be easier for firms to monitor internal production given the simplicity of their DP requirements.

### TABLE 1

**Credit unions using vendor online DP system**

<table>
<thead>
<tr>
<th>Year</th>
<th>Sample with automated DP system</th>
<th>% with vendor online system</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>10,542</td>
<td>30.7</td>
</tr>
<tr>
<td>1995</td>
<td>10,481</td>
<td>29.2</td>
</tr>
<tr>
<td>1996</td>
<td>10,355</td>
<td>27.2</td>
</tr>
<tr>
<td>1997</td>
<td>10,245</td>
<td>26.4</td>
</tr>
<tr>
<td>1998</td>
<td>10,150</td>
<td>26.4</td>
</tr>
<tr>
<td>1999</td>
<td>9,859</td>
<td>26.4</td>
</tr>
<tr>
<td>2000</td>
<td>9,546</td>
<td>26.6</td>
</tr>
<tr>
<td>2001</td>
<td>9,323</td>
<td>26.4</td>
</tr>
<tr>
<td>2002</td>
<td>9,105</td>
<td>26.0</td>
</tr>
<tr>
<td>2003</td>
<td>8,843</td>
<td>26.2</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on NCUA call reports.

### TABLE 2

**In-house data processing versus outsourcing, year-end 2003**

<table>
<thead>
<tr>
<th></th>
<th>In-house</th>
<th>Outsourcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of credit unions</td>
<td>6,523</td>
<td>2,320</td>
</tr>
<tr>
<td>Assets ($mil.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>78.3</td>
<td>46.1</td>
</tr>
<tr>
<td>Median</td>
<td>9.0</td>
<td>21.6</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>377</td>
<td>154</td>
</tr>
<tr>
<td>10th percentile</td>
<td>10.8</td>
<td>49.3</td>
</tr>
<tr>
<td>90th percentile</td>
<td>164</td>
<td>88.5</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on NCUA call reports for December 2003.
On the other hand, medium and large firms offer more sophisticated products, which require complex DP systems. While performing DP in-house allows flexibility in dealing with complex DP tasks, it could be that only the largest CUs achieve efficient scale for internal DP functions. Transaction costs may also account for this discrepancy. In the context of DP, the relationship-specific investment arises from the necessity to train employees to use the systems specific to a vendor. Vendors may also have specific hardware and data organization requirements that are not easily transferable to competitors' systems. These sunk costs make it difficult for CUs to switch vendors. In such circumstances, CUs would be vulnerable to the opportunistic behavior of vendors. This requires contracting. If the costs of contracting for DP are relatively fixed, smaller firms will be deterred from outsourcing while medium firms will find it worthwhile. Larger firms may find it worthwhile to outsource but may also be able to achieve efficient scale internally.

**Product offerings and outsourcing**

Here, we discuss the relationship between outsourcing and product offerings. Most CUs' data processing requirements involve handling data on share (deposit) information. This includes both share (savings) and share draft (checking) data. Beyond these basic saving and checking accounts, many CUs offer more sophisticated vehicles for saving and investing, as well as various types of loan options. Included among these are saving instruments such as share certificates, IRA accounts, money market accounts, auto loans, credit cards, fixed rate mortgages, variable rate mortgages, and home equity loans. We count how many of these eight products are offered by each CU and use this as a measure for the diversity of product offerings.

Table 3 shows data covering our entire sample period, 1994–2003. On average, for about 21 percent of our sample, the data-processing systems deal with only one additional type of financial transaction beyond the basic savings and checking account data, while for about 16 percent of CUs, the DP systems process seven or eight additional types of transactions.

The degree of product diversity might affect outsourcing decisions in two ways. First, diversity might increase the minimum scale necessary to adequately produce DP services in-house. Diverse products require more sophisticated software, requiring a larger on-site IT staff to maintain the system. This would make multi-product firms more likely to outsource than single-product firms of similar size.

A second factor affecting the relationship between product diversity and outsourcing is transaction costs.

Outsourcing a DP system that manages diverse products would require more detailed contracts and greater contingency coverage. Tellers would also require more training to use a specific vendor system. Such factors increase the sunk costs of entering outsourcing relationships, making hold-up more likely. If CUs with more diverse products have greater scope for data analysis, transaction costs might be higher for those with many product offerings as well. In addition, because software is not located on-site and data are also managed remotely, CUs face the risk of unanticipated downtime for the online system. While disaster recovery is usually covered in the standard vendor contract (Klepper and Jones, 1998), the services may not always be satisfactory. These problems might also be more severe for CUs with complex products. Given that the system is not owned by the CU, the CU would not have full control over how the problems are resolved. All of these factors may encourage CUs with a wide range of products to perform data processing in-house.

Table 3 provides details of the interesting relationship between product offerings and outsourcing. A greater number of products is associated with a greater likelihood of outsourcing, but only to a point. CUs with six additional loan or share data-processing requirements are most likely to outsource their data processing. Those offering either fewer or more than six products are less likely to outsource. Again, this suggests that there are countervailing influences at work. As we mentioned before, as the number of products increases, DP needs become more complex, which might reduce the attractiveness of outsourcing if it increases transaction costs. On the other hand, product diversity might also incur greater fixed costs, which would increase the attractiveness of outsourcing.

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Number of products</th>
<th>% of credit unions</th>
<th>% using vendor online</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.64</td>
<td>12.0</td>
</tr>
<tr>
<td>2</td>
<td>14.73</td>
<td>17.4</td>
</tr>
<tr>
<td>3</td>
<td>12.51</td>
<td>25.2</td>
</tr>
<tr>
<td>4</td>
<td>12.96</td>
<td>36.5</td>
</tr>
<tr>
<td>5</td>
<td>12.23</td>
<td>40.8</td>
</tr>
<tr>
<td>6</td>
<td>11.12</td>
<td>41.3</td>
</tr>
<tr>
<td>7</td>
<td>10.19</td>
<td>33.2</td>
</tr>
<tr>
<td>8</td>
<td>5.62</td>
<td>23.7</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors' calculations based on NCUA call reports.
It is also possible that this relationship between the propensity to outsource and product diversity is simply picking up the relationship between the propensity to outsource and CU size. Figure 1 illustrates the positive relationship between CU size and product variety. CUs that offer more products are typically larger. Because they achieve internal scale economies, larger CUs may prefer not to outsource data processing when their product range is diverse, because the increased complexity in data processing reduces the benefit of outsourcing. However, outsourcing may still be preferred by smaller CUs offering more products, because smaller CUs do not have the same internal scale economies.

To distinguish the effects of size from the effects of complexity, in table 4 we stratify CUs by size and examine how the propensity to outsource changes with the number of products. While some cells in the table use a small number of CUs and are therefore noisy, the table suggests some fairly clear patterns. For smaller CUs, we see that outsourcing becomes more likely as the number of products increases. This is consistent with the notion that smaller CUs do not have scale economies in dealing with complex data processing, forcing them to use vendors if they offer complex products. The pattern also exists for medium-sized CUs, although the effect is not quite so dramatic. For large CUs, however, the relationship is reversed—a greater number of products seems to be associated with in-house DP. This relationship is more consistent with a transaction-cost-based explanation, whereby complex DP creates a difficult contracting environment and, therefore, encourages in-house production.

**Probit analysis**

To further explore these relationships, we perform a probit analysis, specifying cross-sectional variation in outsourcing as a function of size and product diversity. The general empirical framework we employ is a discrete choice model in which a CU outsources when

\[ Y^*_i = \alpha + \beta_1 \text{Size}_i + \beta_2 N_i + \beta_3 (\text{Size}_i \times N_i) + \text{year dummies} + \text{other control variables} + \epsilon_i > 0, \]

where \( Y^*_i \) represents the net benefit of outsourcing for CU \( i \) in year \( t \). CU size is measured by the logarithm of assets that is deflated by GDP deflator (base year is set at 2003) and \( N \) stands for the number of products that a CU offers. As we mentioned, based on table 4, it appears that product mix has different effects based on CU size, so we include the interaction term \( \text{Size}_i \times N_i \). We assume that the error \( \epsilon_i \) is normally distributed and estimate the above equation by performing probit analyses, where a CU outsources when \( Y^*_i > 0 \) based on our whole sample of 98,449 observations. We also include a dummy variable indicating whether the CU is located in an urban area. Existing studies (Hubbard, 2001; Ono, 2001) suggest that local market size affects the propensity to outsource. In our data, while CUs in urban and rural markets offer roughly the same number of products, urban CUs are on average larger. Thus, controlling for location may cloud our interpretation of the size coefficient. We also include dummies indicating the CU’s field of membership, or group that it serves. Such groups are defined by community, association, educational institution, military group, government entity, as well as companies (Borzekowski, 2004). Specific types of membership groups are likely to be associated with other characteristics of CUs as well as their outsourcing propensity; by controlling for the field of membership, we can net out the
TABLE 4
Percentage of CUs outsourcing by size and number of products, 1994–2003

<table>
<thead>
<tr>
<th>Assets (deflated)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $10 million</td>
<td>10.5</td>
<td>15.1</td>
<td>20.8</td>
<td>29.0</td>
<td>33.8</td>
<td>39.5</td>
<td>45.3</td>
<td>34.2</td>
</tr>
<tr>
<td>(19,393)</td>
<td>(12,841)</td>
<td>(9,552)</td>
<td>(6,846)</td>
<td>(2,954)</td>
<td>(948)</td>
<td>(192)</td>
<td>(38)</td>
<td></td>
</tr>
<tr>
<td>$10 million–$50 million</td>
<td>40.3</td>
<td>34.8</td>
<td>40.6</td>
<td>45.9</td>
<td>45.8</td>
<td>47.6</td>
<td>45.6</td>
<td>44.8</td>
</tr>
<tr>
<td>(821)</td>
<td>(1,526)</td>
<td>(2,571)</td>
<td>(5,377)</td>
<td>(7,468)</td>
<td>(6,761)</td>
<td>(4,192)</td>
<td>(1,014)</td>
<td></td>
</tr>
<tr>
<td>$50 million or more</td>
<td>66.3</td>
<td>34.8</td>
<td>38.0</td>
<td>37.5</td>
<td>30.7</td>
<td>28.7</td>
<td>23.6</td>
<td>18.8</td>
</tr>
<tr>
<td>(101)</td>
<td>(135)</td>
<td>(184)</td>
<td>(539)</td>
<td>(1,623)</td>
<td>(3,243)</td>
<td>(5,648)</td>
<td>(4,482)</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>12.00</td>
<td>17.40</td>
<td>25.20</td>
<td>36.50</td>
<td>40.80</td>
<td>41.30</td>
<td>33.20</td>
<td>23.70</td>
</tr>
</tbody>
</table>

Note: Number of observations in parentheses. Source: Authors’ calculations based on NCUA call reports.

The effect of the group that the CU services from that of CU size and the diversity of their services. Table 5 shows the summary statistics of the variables included in the analysis.

Table 6 shows our results. The coefficients for CU size, the number of products (N), and the interaction terms between them are all significant. The coefficient for CU size and N are both positive, and that for the interaction term is negative.11

This suggests that both size and product diversity (N) increase the propensity to outsource, but that at higher levels of N, the relationship reverses. Based on the results in table 6, for CUs with average characteristics, the relationship between the probability of outsourcing and the number of products is:

\[
d (\text{Prob of Outsourcing}) / dN = 0.581 - 0.0340 \text{ Size},
\]

\[
d (\text{Prob of Outsourcing}) / dN \text{ is positive for a CU with log(assets) below 17.09}.12 \text{ For CUs larger than this, offering additional products is associated with a lower probability of outsourcing. Again, this is consistent with the idea that when DP requirements are more complex, larger CUs may prefer to perform the services in-house in order to avoid the costs of specifying many details on the contract. It is also possible that larger CUs experience greater benefits from retaining the flexibility that in-house DP allows, or that DP complexity makes monitoring the outsourced relationship more difficult}.13

For small CUs, on the other hand, the probability of outsourcing is greater for those that offer a wider range of products. It is possible that for small CUs, the benefit of relying on scale economies in a vendor may outweigh the benefits of performing DP in-house.

The outsourcing options might have influenced the number of products that CUs offer. Among CUs that do not use vendor online systems, however, the majority purchase the software from vendors, which usually can accommodate a wide range of products as the VOL. We also ran the probit analysis excluding the CUs that reported they develop the software by themselves, and our results remained qualitatively the same.

Another way of interpreting our empirical results is to focus on size. From table 6,

\[
d (\text{Prob of Outsourcing}) / dN = 0.1679 - 0.340 N.
\]

The effect of Size is zero when N is about five. For CUs offering more than five products, the relationship between the likelihood of outsourcing and CU size is negative. Again, when the product offered is complex, larger CUs may be more likely to perform DP in-house, in order to avoid high transaction costs. For CUs offering fewer than five products, the relationship between size and propensity to outsource is positive. When the degree of product diversity is low, small CUs may find in-house DP less costly, considering

TABLE 5

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log assets (deflated)</td>
<td>16.05</td>
<td>1.60</td>
</tr>
<tr>
<td>Number of products</td>
<td>3.84</td>
<td>2.22</td>
</tr>
<tr>
<td>FOM: Community</td>
<td>0.070</td>
<td>0.26</td>
</tr>
<tr>
<td>FOM: Association</td>
<td>0.058</td>
<td>0.23</td>
</tr>
<tr>
<td>FOM: Education</td>
<td>0.082</td>
<td>0.27</td>
</tr>
<tr>
<td>FOM: Military</td>
<td>0.014</td>
<td>0.12</td>
</tr>
<tr>
<td>FOM: Government</td>
<td>0.118</td>
<td>0.32</td>
</tr>
<tr>
<td>Located in urban areas</td>
<td>0.79</td>
<td>0.41</td>
</tr>
</tbody>
</table>

*Base year is 2003.
*Areas within PMSA under 1994 definition. Source: Authors’ calculations based on NCUA call reports.
the search costs and contract costs associated with outsourcing. In contrast, large firms may have more negotiating power and may receive favorable treatment from vendors; thus for them, the benefits of outsourcing may outweigh the contracting costs as long as the DP requirements are not too complicated. Therefore, when a big CU offers a relatively small range of products—and, consequently, when the contract it has to negotiate if it chooses to outsource is less complicated—the CU might see more benefit from outsourcing compared with performing DP in-house.

**Conclusion**

Outsourcing has become a much examined and debated issue. Researchers are increasingly recognizing that, in addition to the economic issues associated with outsourcing across national borders, outsourcing decisions are a key component of business strategy. Little is known, however, about the factors that affect firms’ outsourcing decisions. We have addressed one aspect of this issue by examining CUs’ outsourcing decisions. We find that both CU size and product diversity are important factors influencing a CU’s decision to outsource DP. While it appears that CU size and product diversity may have independent effects, they also interact; the relationship between outsourcing and CU size depends on the number of products that the CU offers and vice versa.

Our analysis reveals that, in general, larger CUs are more likely to outsource their DP function, although the relationship is reversed for the very largest CUs. This stands in contrast to a simple scale-based explanation for outsourcing. Product diversity in general has an intuitive impact. For smaller CUs without the capacity to handle sophisticated DP functions, having more products increases their propensity to outsource.

Again, for larger CUs the relationship is reversed. Large CUs exhibit a positive relationship between the number of products and in-house data processing. This may reflect larger firms’ desire to make their data processing part of their core competency, a strategy they can pursue because they have sufficient scale.

Our results imply that outsourcing is probably driven by a combination of factors rather than any one simple influence. While scale economies are an important determinant of firms’ outsourcing decisions, the transaction costs associated with using vendors, which vary based on firms’ characteristics, seem to affect their decisions.

| TABLE 6 |

Results of probit analysis

<table>
<thead>
<tr>
<th></th>
<th>dF/dx</th>
<th>Robust standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU size: log assets (deflated)</td>
<td>0.1679***</td>
<td>0.0061</td>
</tr>
<tr>
<td>Number of products</td>
<td>0.5812***</td>
<td>0.0189</td>
</tr>
<tr>
<td>Number of products x CU size</td>
<td>-0.0340***</td>
<td>0.0011</td>
</tr>
<tr>
<td>FOM: Community a</td>
<td>0.0243**</td>
<td>0.0127</td>
</tr>
<tr>
<td>FOM: Association b</td>
<td>-0.0588***</td>
<td>0.0143</td>
</tr>
<tr>
<td>FOM: Education c</td>
<td>-0.0576***</td>
<td>0.0116</td>
</tr>
<tr>
<td>FOM: Military d</td>
<td>-0.0446</td>
<td>0.0266</td>
</tr>
<tr>
<td>FOM: Government e</td>
<td>-0.0179</td>
<td>0.0103</td>
</tr>
<tr>
<td>Dummy: located in urban areas f</td>
<td>0.0094</td>
<td>0.0086</td>
</tr>
<tr>
<td>Dummy: y95 g</td>
<td>-0.0079***</td>
<td>0.0028</td>
</tr>
<tr>
<td>Dummy: y96 h</td>
<td>-0.0291***</td>
<td>0.0041</td>
</tr>
<tr>
<td>Dummy: y97 i</td>
<td>-0.0400***</td>
<td>0.0042</td>
</tr>
<tr>
<td>Dummy: y98 j</td>
<td>-0.0407***</td>
<td>0.0044</td>
</tr>
<tr>
<td>Dummy: y99 k</td>
<td>-0.0442***</td>
<td>0.0045</td>
</tr>
<tr>
<td>Dummy: y00 l</td>
<td>-0.0446***</td>
<td>0.0047</td>
</tr>
<tr>
<td>Dummy: y01 m</td>
<td>-0.0503***</td>
<td>0.0048</td>
</tr>
<tr>
<td>Dummy: y02 n</td>
<td>-0.0598***</td>
<td>0.0043</td>
</tr>
<tr>
<td>Dummy: y03 o</td>
<td>-0.0593***</td>
<td>0.0045</td>
</tr>
<tr>
<td>Predicted probability at mean</td>
<td>0.2730</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>98,271</td>
<td></td>
</tr>
</tbody>
</table>

*dF/dx is for discrete change of dummy variable from 0 to 1; assets are deflated by GDP deflator (base year = 2003).
Notes: Base year is 1994; White-correlated standard errors with clustering over credit unions were calculated; *** indicates significant at 1 percent level;
** indicates significant at 5 percent level; and * indicates significant at 10 percent level.
Source: Authors’ calculations based on NCUA call reports.
NOTES

1 Sunk costs are investment costs that can never be recouped. For example, when an investment made by a firm has no intrinsic value to other firms, cannot be sold in a secondary market, or cannot be allocated to another use within the firm, the investment represents sunk costs.

2 Coase (1937) and Williamson (1975) identified four types of transaction costs. “First, some contingencies which the parties will face may not be foreseeable at the contracting date. Second, even if they could be foreseen, there may be too many contingencies to write into the contract. Third, monitoring the contract may be costly. Fourth, enforcing contracts may involve considerable legal costs.” (Tirole, 1988)

3 Some evidence on the relationship between outsourcing and firm size can be found in Abraham and Taylor (1996) and Ono (2001) although it was not the main focus of these papers. Examining the relationship between manufacturers’ decision to contract out business services and manufacturers’ characteristics, Abraham and Taylor (1996) found support for the scale-based story for outsourcing practices for business services, including janitorial, accounting, and computer services. In contrast, using the Annual Survey of Manufacturers and examining manufacturing establishments’ practices of outsourcing advertising, accounting and bookkeeping, legal services, as well as data processing, Ono (2003) finds evidence inconsistent with the scale-based story.

4 As shown in table 2, 90th percentile assets of CUs with in-house data processing are greater than those of outsourcing CUs.

5 The banking literature draws a distinction between products that appear on the bank’s balance sheet as assets (such as loans) and those that appear as liabilities (such as checking accounts). For our purposes, we assume that consumers view all financial services as “products” defined broadly.

6 Our data are a panel, but for simplicity here we present results from a specification that does not fully exploit this dimension of the data, for example, by including firm fixed effects. We did estimate a fixed-effects model and obtained qualitatively similar results.

7 It is possible that urban locations have a greater supply of IT personnel, allowing CUs to carry an internal IT department at lower cost. At the same time, a dense local IT labor market might be associated with greater turnover of IT personnel. In such a case, CUs may decide to outsource in order to avoid the costs associated with high IT personnel turnover.

8 See Ono (2001) for an analysis of local market effects on outsourcing.

9 We also ran the probit analysis, excluding military and governmental CUs as well as some very small and very large CUs whose log assets (deflated) are below and above 3 standard deviations from the mean. This left us with 85,156 observations. The results of the probit remained qualitatively the same. For this restricted sample, we also ran the probit for each year. The coefficients for size, product complexity, and their interaction terms were qualitatively the same across years.

10 This corresponds to assets of roughly $26 million (deflated by the GDP deflator, base year = 2003).

11 See Baker and Hubbard (2003).
REFERENCES


Is there evidence of the new economy in U.S. GDP data?

Michael A. Kouparitsas

Introduction and summary

Economic theory suggests that temporary cyclical fluctuations in real gross domestic product (GDP) adversely affect the economic well-being of households. For example, when the economy experiences a cyclical downturn, companies lay off workers with resulting negative consequences for the workers and their families. Thus, it is not surprising that cyclical fluctuations in GDP receive a lot of attention from policymakers. Indeed, there is considerable empirical research that shows that cyclical fluctuations in GDP play an important role in the practical conduct of U.S. monetary policy. In general, the U.S. Federal Reserve (Fed) tightens monetary policy (increases interest rates) when the cyclical component of GDP rises and loosens monetary policy (reduces rates) when the cyclical component of GDP falls.

Unfortunately, economists cannot observe the cyclical component of GDP. This is because observed GDP is made up of two unobserved components. The first, called the trend component, refers to the upward sloping part of GDP. For example, figure 1, panel A plots the trend component of GDP under the assumption that it is a constant linear trend (green line). The second, called the cyclical component, refers to the fluctuations around the trend component. Figure 1, panel B plots the cyclical component of GDP that is related to the constant linear trend plotted in panel A.

Economists typically identify the policy-important cyclical component by first making assumptions that allow them to isolate the trend component and then backing out the cyclical component. In general, the biggest challenge in isolating the trend component is estimating its slope. The slope of the trend component is determined by the trend growth rate of GDP (that is, the growth rate of output that would exist if there were no cyclical fluctuations in GDP). Higher trend growth rates imply a steeper trend component.

The debate over the true value of the trend growth rate received a lot of attention in the late 1990s from economic analysts and policymakers. Analysts argued, based on strong observed growth of labor productivity (GDP per worker), that the trend growth rate of GDP had increased significantly. If an increase in the trend growth rate had occurred, this type of structural change would have meant that economists could no longer rely on their longstanding rules of thumb about the relationship between observed GDP and the unobserved cyclical component in formulating policy. This led to speculation by the analysts that the U.S. was a new economy in the late 1990s, in which all the old rules about actual, trend, and cyclical fluctuations of GDP no longer held true.

In this article, I test whether there was in fact significant change in the trend growth rate of U.S. GDP over the new economy era. I do so by applying both long-established and newer techniques of extracting the trend component of U.S. GDP data and then testing to see if the implied trend growth rate of U.S. GDP (that is, its average slope) over the new economy era is significantly higher than the implied trend growth rate of U.S. GDP over the preceding productivity slowdown era.

Irrespective of the method used to extract the trend component, I find that the implied annual trend growth rate of U.S. GDP was about 3 percent over the productivity slowdown period, which is considerably higher than the typical 2.5 percent estimate based on productivity data, and about 3.25 percent over the new economy era. Although I find a positive difference between the new economy and productivity slowdown era estimates, it is not statistically significant. I conclude

Michael A. Kouparitsas is an economist at the Federal Reserve Bank of Chicago.
that, at least in terms of GDP data, the U.S. was the same old economy in the late 1990s.

A simple linear trend model of GDP

An economy is like a biological organism in that it grows exponentially over time. For example, in the simple case of an economy that is growing at the constant rate $\mu$ per time period, the size of this economy measured by GDP at time $t$ is given by

1) $X_t = \Phi e^{\mu t}$

where $X_t$ denotes the level of GDP at time $t$ and $\Phi$ is a constant. Economists generally do not work with the level of GDP, but instead prefer to work with the log of GDP (log GDP). The main reason for this transformation is that growth rate calculations using log GDP are linear, while similar calculations using the level of GDP are non-linear. For example, if we take the log of both sides of equation 1 and denote log GDP at time $t$ by $x_t$, it follows that:

2) $x_t = A + \mu t$

where $A = \log(\Phi)$. In this simple case, log GDP is a linear function of a constant $A$ and a time trend $t$ with coefficient $\mu$. If we were to plot this relationship with the value of log GDP on the vertical axis and time along the horizontal axis, the intercept of log GDP with the vertical axis would be $A$ and the slope of log GDP as we move along the horizontal axis would be $\mu$. An increase in $A$ would shift up log GDP by a constant amount across all periods, so economists call changes in the constant a level shift. Raising the growth rate of GDP $\mu$ increases the slope of log GDP across all periods, so economists call changes in the growth rate a slope change. Models that economists actually use to explain the evolution of GDP over time essentially build on this simple model by allowing for some type of variation in the constant and slope.

Allowing for cyclical variation around the trend

The first significant departure from the simple linear trend model described by equation 2 is that log GDP can be additively decomposed into a trend component $\tau_t$ and a cyclical component $c_t$ as follows:

3) $x_t = \tau_t + c_t$

The trend component captures the upward sloping part of GDP (which was explained in the simple model in equation 2 by a linear trend), while the cyclical component captures fluctuations around the trend component (this component was ignored in equation 2). However, economists do not observe either the trend or cyclical component of GDP. Economists typically proceed along one of three paths in estimating
these unobserved components. First, they estimate the trend component directly and determine the cyclical component as the difference between observed log GDP and the estimated trend component. Second, they estimate the cyclical component directly and determine the trend component as the difference between observed log GDP and the estimated cyclical component. Or, finally, they jointly estimate the trend and cyclical components.

**Estimating a constant growth rate model of GDP**

Early attempts at estimating the trend component took the direct approach by assuming, as in the simple example above, that log GDP had a linear trend with constant slope $\mu$:

$$4) \quad \tau_t = A + \mu,$$

where $t$ denotes the linear trend, $A$ is a constant, and the $t$ subscript denotes the date at which the trend is being measured. Just as in the simple model discussed above, the slope coefficient $\mu$ is the trend growth rate of GDP. The cyclical component is simply the difference between the linear trend and log GDP as follows:

$$5) \quad c_t = x_t - (A + \mu).$$

All the elements of this decomposition are plotted in figure 1. Starting with figure 1, panel A, the black line is log GDP $x_t$, while the green line is the estimated linear trend, $\hat{\tau}_t = \hat{A} + \hat{\mu}$, where $\hat{A}$ and $\hat{\mu}$ are estimated using ordinary least squares. According to my estimate of the linear trend model, the constant annual growth rate of GDP over the entire sample of quarterly GDP data from 1961:Q1 to 2003:Q4 was 3.16 percent. Figure 1, panel B plots the implied cyclical component $c_t$, which is multiplied by 100 so that it can be interpreted as a percentage deviation from the trend component. This figure shows that this relatively simple method of estimating the trend component produces a cyclical measure of GDP that has turning points that line up closely with the National Bureau of Economic Research’s (NBER) peak and trough dates, captured by the gray bars.

**Estimating a time-varying growth rate model of GDP**

Virtually all the recent research aimed at estimating the trend growth rate of GDP has focused on whether it changed significantly over the so-called new economy era from 1995:Q4 to present. Much of this investigation was fueled by the spectacular increase in the observed trend growth rates of labor productivity (GDP per worker) in the mid-1990s. To understand why this is important, we need to review the way that the trend growth rate of GDP is typically estimated. Economists have long recognized that the trend growth rate of GDP is the sum of the trend growth rate of labor productivity and the trend growth rate of employment. This led to the popular bottom-up approach in estimating trend GDP growth, whereby researchers estimate the trend growth rate of productivity and the trend growth rate of employment directly and simply add these components together to get the implied trend growth rate of GDP. Gordon (2003) and many others have used this approach and argued on the basis of their estimates that there was a significant variation in the trend growth rate of GDP over the new economy era.

In particular, Gordon found that the trend growth rate of labor productivity rose significantly from an annual growth rate of 1.5 percent estimated over the productivity slowdown era from 1973:Q1 to 1995:Q3 to 2.5 percent over the new economy era. Under the typical assumption that the trend growth rate of employment is 1 percent, which is based on a constant long-run labor force participation rate and trend growth of the labor force of 1 percent, Gordon concluded that the implied trend growth rate of GDP over the productivity slowdown era was 2.5 percent and that there was a significantly higher implied trend growth rate over the new economy era of 3.5 percent.

In contrast to these researchers, I take a more direct approach to testing whether the trend growth rate of GDP changed over the new economy era. I use the techniques used by Gordon (2003) and others to estimate the trend growth of productivity directly to estimate the trend growth of GDP directly. A possible advantage of this approach is that it does not require auxiliary assumptions about the trend growth of employment, since it uses the same method to estimate the trend growth rate of both productivity and employment.

Following Gordon’s approach to estimating the trend growth rate of productivity, I allow the parameters that govern the slope of the trend component of GDP to vary over the sample. If this exercise shows there has been a statistically significant variation in the parameters that govern the slope of the trend before and after the new economy era, this would imply that we are indeed in a new economy. Alternatively, if I find that variation in the slope over these periods is not statistically significant, this would support the conclusion that the U.S. was the same old economy in the latter half of the 1990s.

**A simple time-varying linear trend model of GDP**

A useful starting point on this path is the time-varying (discrete jump) linear trend model that was
<table>
<thead>
<tr>
<th>Estimation method</th>
<th>Estimation period</th>
<th>Comparison period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity studies (Gordon)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear trend models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear trend (constant)</td>
<td>3.16</td>
<td></td>
</tr>
<tr>
<td>Linear trend (time-varying/discrete jump)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency domain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band-pass filter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unobserved component models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate models (Watson)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate unobserved component (constant)</td>
<td>3.30</td>
<td></td>
</tr>
<tr>
<td>Univariate unobserved component (discrete-jump)</td>
<td>4.20</td>
<td>3.29</td>
</tr>
<tr>
<td>Multivariate models (Gerlach &amp; Smets)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariate unobserved component (constant)</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>Multivariate unobserved component (discrete-jump)</td>
<td>4.12</td>
<td>3.14</td>
</tr>
<tr>
<td>Univariate unobserved component (unit-root)</td>
<td>3.72</td>
<td>3.12</td>
</tr>
<tr>
<td>Multivariate models (Gerlach &amp; Smets)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multivariate unobserved component (constant)</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>Multivariate unobserved component (discrete-jump)</td>
<td>4.12</td>
<td>3.14</td>
</tr>
<tr>
<td>Multivariate unobserved component (unit-root)</td>
<td>3.91</td>
<td>3.02</td>
</tr>
</tbody>
</table>

* Denotes average over period.
* Denotes statistically significant at 5 percent level.
widely used in the 1970s. This model allows the constant and slope of the trend component to vary over discrete intervals:

$$\tau_j = A_{t,j} + \mu_{t,j} t,$$

where $A_{t,j}$ is the constant and $\mu_{t,j}$ is the trend growth rate over the time interval from $i$ to $j$.\(^3\) Following the productivity trend growth literature, I allow for variation in trend component parameters over three periods: the productivity slowdown era from 1973:Q1 to 1995:Q3, the new economy era from 1995:Q4 to 2003:Q4, and the pre-slowdown era from 1961:Q1 to 1972:Q4.

My estimates of the discrete jump linear trend model, reported in table 1 on the previous page, suggest that the trend growth rate of GDP was 3.88 percent in the pre-productivity slowdown era, well above the productivity slowdown estimate of 2.94 and new economy era estimate of 3.39.\(^4\) More importantly, I find that the difference between the new economy and the productivity slowdown trend growth rate estimates is statistically significant, which suggests that the U.S. was a new economy post 1995:Q3.

Figure 2 reveals the impact on the cyclical component of allowing for a time-varying trend growth rate. Differences between trend components are inversely related to differences between cyclical components: As the trend component shifts up, the cyclical component decreases. Although the difference between the constant (dark green line) and discrete jump (light green line) trend components in figure 2 appears to be small, the percentage point difference between the constant (dark green line) and discrete jump (light green line) growth rate cyclical components is quite large. For example, the cycle in 1996:Q1 is $-2.7$ percent for the constant trend growth rate model and $-0.9$ percent for the discrete-jump trend growth rate model. A variation of this size would likely generate a different policy response from the Fed, which highlights the importance to policymakers of estimating the true trend growth rate.

A more important experiment for the current exercise is a comparison of the cyclical component assuming no change in the trend growth rate (dashed green line, which shows what the discrete jump growth rate cyclical component would have been if
the trend growth rate were held constant at its 1995:Q3 level over the post 1995:Q3 new economy period) and the cyclical component when the trend growth rate is allowed to change (light green line). This figure appears to support the new economy theory, because it shows that the policy-important cyclical component that incorporates changes in the trend growth rate of GDP lies everywhere below the same cyclical component assuming no change. This suggests that the Fed would have responded to the change in growth rates over the new economy era by tightening monetary policy less aggressively than if it had maintained the growth rate of the productivity slowdown era.7

**Does GDP have a linear trend?**

Developments in the field of econometrics during the 1980s called into question the usefulness of the simple linear trend model for policy analysis. Armed with new and powerful statistical techniques, economists such as Nelson and Plosser (1982) explored the trend properties of economic time series and discovered that many U.S. time series, including GDP, had stochastic rather than deterministic trends as in the linear trend models.6 Stochastic trends are more general than the deterministic linear trend models described above. The primary difference is that they allow for significant variation in the level of the trend component. In other words, the constant term $A$ in the linear model is a random variable in the stochastic trend model. This development meant that the widely used linear trend models were misspecified.

Economists reacted to this challenge by developing new approaches to modeling economic time series with stochastic trends, known as frequency domain and unobserved component techniques. These methods revealed that the simple linear trend models (including the discrete jump linear trend model estimated above) were poor representations of the data. In particular, they provided misleading results on the nature of the trend and cyclical components of GDP. In light of this finding, economists have largely relied on frequency domain and unobserved component techniques to isolate the trend and cycle components of economic time series.

**Frequency domain estimates of the trend and cyclical components**

Frequency domain techniques were made popular by the modern business cycle literature starting in the 1980s. According to this paradigm, fluctuations in the data at the so-called business cycle frequencies of between 18 months and eight years are considered cyclical movements, $c_t$, while long-run fluctuations occurring at frequencies of greater than eight years make up the trend component $\tau_t$. The main advantage of this approach over unobserved component methods is that it can isolate the noisy short-run movements of economic time series that are a nuisance to
policymakers. Fluctuations in the data occurring at frequencies of less than 18 months are regarded as noise, \( \eta_t \). Using this approach, log GDP is the sum of three components, trend, cycle, and noise:

\[
x_t = \tau_t + c_t + \eta_t.
\]

The most convenient way of extracting these three components from time series data is via an approximate band-pass filter (BPF). Approximate BPFs are essentially centered moving averages. The problem with these approximate filters is that the filtered data ends up being much shorter than the unfiltered time series, because the moving average requires a significant amount of data at the beginning and end of the sample, up to three years in the case of quarterly GDP data. My analysis of the trend growth rate of GDP relies on the approximate BPF method developed by Christiano and Fitzgerald (2003), which is designed to filter the data over the entire sample, thus preserving the sample size.

Figure 3 on page 17 plots the trend and cyclical component of GDP generated by a BPF. The most obvious difference from figure 2 is that the BPF cyclical component is considerably smoother than its linear trend counterparts. This reflects the fact that the BPF cycle does not include the highly irregular noise component. Another advantage of the frequency domain approach over the unobserved component method is that it can endogenously identify changes in the trend growth rate. Figure 3 shows that, in contrast to the discrete-jump linear trend model, the slope changes of the BPF trend are numerous and smooth. The extent of these growth rate changes is revealed in figure 4, which plots the implied annual growth rate of the BPF trend component (green line).

Given the variation in the implied trend growth rate, I test for significant change in the trend growth rate of GDP by testing if the average growth rate of the BPF trend component over the new economy era is significantly higher than the average growth rate of the BPF trend component over the preceding productivity slowdown era (black line). I find that the average BPF trend growth rates are 2.98 for the productivity slowdown period and 3.26 percent for the new economy era. In contrast to the discrete-jump linear trend model, the difference between these estimates is not statistically significant, which suggests that the U.S. was the same old economy in the late 1990s.

**Unobserved component techniques**

Another group of economists led by Watson (1986) took a completely different route to decomposing GDP into its trend and cyclical components by applying unobserved component (UC) techniques. In contrast to the frequency domain approach, UC methods require strong assumptions about the exact form of the trend and cyclical components. Watson’s initial UC model of log GDP responded directly to the work of Nelson and Plosser (1982) by allowing log GDP to have a stochastic trend. In particular, Watson’s model assumed that the trend component of log GDP was a random walk with drift.

The random walk with drift assumption meant that the trend component \( \tau_t \) depended on its most recent past observation \( \tau_{t-1} \), a random component \( \epsilon_{\tau_t} \), and a constant term, typically called drift \( \mu \):

\[
7) \quad \tau_t = \mu + \tau_{t-1} + \epsilon_{\tau_t}.
\]

In the absence of random fluctuations (\( \epsilon_{\tau_t} = 0 \)), the trend component grows at a rate equal to the drift \( \mu \). However, the trend component does not always grow at the trend growth rate because positive random fluctuations lead to trend growth in excess of the drift, while negative random fluctuations cause the trend to grow by less than the drift. It is important to note that while fluctuations in the random component \( \epsilon_{\tau_t} \) have a permanent effect on the level of the trend component, they do not have a permanent effect on the trend.
growth rate. Therefore, the long-run or trend growth rate of GDP is measured by the drift term $\mu$.

Watson’s model assumes the cyclical component is a second order autoregression, which means that the current cyclical component $c_t$ depends on its most recent past two observations $c_{t-1}$, $c_{t-2}$, and a random component $\varepsilon_t$:

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon_t.$$  

The cyclical component is assumed to be a stationary process, which means that random shocks to the cycle $\varepsilon_t$ have no permanent effect on the level of the cycle or log GDP. This requires that $\phi_1 + \phi_2 < 1$. Finally, the noise component cannot be identified, so log GDP $x_t$ is assumed to be the sum of the trend and cyclical components.

Watson’s UC model of log GDP is often referred to as a univariate UC model because although there are many unobserved components, there is only one observed component. The unobserved components are the trend, cyclical, random trend, and random cycle components, while the observed component is log GDP. The unobserved components are identified by assuming that the random trend $e_t$ and random cycle $e_{ct}$ components are uncorrelated.

Table 1 reports the trend growth rate estimates from Watson’s univariate UC model with a constant drift. Despite additional data, the move to chain-weighted real GDP indexes, and recent changes in the measurement of business investment, my estimate of the trend growth rate of GDP is very close to that first reported by Watson. At 3.30 percent, it is slightly higher than the constant growth rate estimate from the linear trend model.

Figure 5 plots the univariate UC constant drift trend and cyclical components. Panel A reveals that the UC trend component is not as smooth as the linear trend component. This highlights level shifts of log GDP caused by fluctuations in the random trend component. Panel B shows that the univariate UC cyclical component also has turning points that closely match the NBER business cycle dates.

**Unobserved component time-varying trend growth rate models**

Watson’s model assumes that variations in the growth rate of the trend are temporary, so it needs to be modified to test for permanent changes in the trend growth rate. I build on Watson’s model by introducing a time-varying drift term $\mu_t$ that allows the growth rate of the trend to change permanently. I consider two cases, a discrete-jump model that allows for lumpy changes in the trend growth rate and a unit-root model that allows for smooth changes in the trend growth rate.
Discrete-jump trend growth rate

The first case assumes that changes in the trend growth rate of GDP take on discrete jumps. In particular, the trend growth rate is assumed to jump to a new level \( \mu_{ij} \) for a fixed period \((t = i \text{ to } j)\). Under this assumption, the trend component is a random walk with drift:

\[
\tau_t = \mu_i + \tau_{i-1} + \epsilon_t,
\]

but now the drift \( \mu_i \) is allowed to vary over fixed periods:

\[
\mu_t = \mu_{ij} \quad \text{for } i \leq t \leq j.
\]

This is analogous to the time-varying discrete-jump linear trend model studied above. Just as in the linear trend case, I allow the drift term to vary over three periods that are widely viewed by empirical researchers to be periods in which the trend growth rate of productivity changed significantly: the productivity slowdown era; the new economy era; and the pre-productivity slowdown era. I test for significant change in the trend growth rate over the new economy era by testing whether the drift term over the new economy period is significantly higher than the drift term over the preceding productivity slowdown period.

Table 1 reports my estimates of trend growth rates for all the models I estimate across the three periods. In the case of the univariate discrete-jump model, the trend growth rate over the new economy era (3.29 percent) is higher than that for the productivity slowdown era (2.92 percent), but the difference between the two rates is not statistically different from zero. Therefore, I cannot reject the null hypothesis that the U.S. was the same old economy in the late 1990s. In contrast, my estimates suggest that the difference between the pre-slowdown and productivity slowdown trend growth rates is a statistically significant 1.28 percent.

Figure 6 plots the univariate UC trend and cyclical components under the discrete-jump assumption. The discrete-jump drift cyclical component (light green line) lies slightly above the constant-drift cyclical component (dark green line) over the early part of the new economy period. This gap diminishes over the latter part of the 1990s, so that the two curves are virtually identical around 2000. Just as in the discrete-jump linear trend model, level differences of this size would likely generate a different policy response from the Fed, which again highlights the importance to policymakers of estimating the true trend component.
However, this comparison of the constant and discrete-jump drift cyclical components is uninformative when it comes to answering the question of whether changes in the trend growth rate had a bearing on monetary policy. As in the linear case, we need to compare the cyclical components under the assumption of no change in the trend growth rate in 1995:Q4 in the discrete-jump linear trend model (the old economy path, not plotted) and the plotted cycle, which allows for changes in the trend growth rate (the new economy path). Irrespective of where the estimation sample ends post 1995:Q3, I find that the path of the cyclical component, assuming no change in the growth rate from 1995:Q3 onwards, is close to the cycle that allows for changes in the trend growth rate in 1995:Q4. This finding suggests that if the Fed used this unobserved component model to estimate the trend and cycle component, but failed to factor in a change in the trend growth rate in 1995:Q4, its monetary policy response would have been indistinguishable from its response with a change in the trend growth rate.

**Unit-root growth rate**

The next time-varying model follows Harvey and Todd (1983) and Clark (1987) in assuming that the trend component is a random walk with a time varying drift:

\[ \eta_t = \eta_{t-1} + \epsilon_{t\mu} \]

but now I allow the drift \( \mu_t \) to vary in a smooth way by allowing it to be also a random walk process:

\[ \eta_t = \mu_{t-1} + \epsilon_{t\mu} \]

where the current trend growth rate \( \mu_t \) depends on its most recent past observation \( \mu_{t-1} \), plus a random component \( \epsilon_{t\mu} \). Under this assumption, fluctuations in trend growth come from two sources: changes in the random trend component \( \epsilon_{t\mu} \), which permanently change the level of the trend component, and changes in the random drift shock \( \epsilon_{t\mu} \), which permanently change the slope of the trend (or long-run trend growth rate).

The unobserved trend, cyclical, and time-varying drift components are identified by assuming that...
the random trend component \( \varepsilon_{\mu} \), random cycle component \( \varepsilon_{\pi} \), and random drift component \( \varepsilon_{\sigma} \) are uncorrelated. In this case, I test for significant change in the underlying trend growth rate by testing whether the average time-varying drift \( \mu \), over the new economy era is significantly higher than the average time-varying drift over the preceding productivity slowdown era.

Figure 7 plots the unit-root trend growth rate for the Watson model (dark grey line) against its discrete-jump (light green line) and constant trend growth rate (dark green line) counterparts. This figure suggests that after a dramatic fall in the trend growth rate over the pre-productivity slowdown era, the trend growth rate held steady at about 3.12 percent over the latter part of the productivity slowdown era and into the new economy era. A comparison of the average growth rates over these periods indicates that there has not been a statistically significant variation in the average trend growth rate over the latter part of the estimation period (see point estimates in table 1). Hence, there is no evidence of the new economy in the GDP data.

The point estimates of the trend and cyclical component parameters of the unit-root model are slightly different to the model with a constant drift. This is echoed in figure 6 by the similarity of the unit-root (dark gray line) and the constant drift (dark green line) trend and cyclical components, especially over the productivity slowdown and new economy eras. The only noticeable difference occurs in the early part of the sample, which reflects the relatively high trend growth rate estimates over the 1960s.

If the underlying trend growth rates did not change, what then explains the high observed GDP growth rates in the late 1990s? Figure 8 suggests that the high GDP growth rates of the late 1990s were the result of a level shift in the trend component, which was driven by a sequence of relatively large positive random fluctuations that had a permanent effect on the level of the trend, but no permanent effect on the trend growth rate.

**Multivariate unobserved component models of GDP**

One of the drawbacks of Watson’s model is that despite carrying the label of a structural economic model, it is in fact atheoretical in that it embodies no behavioral economic relationships. Various authors have attempted to add behavioral content to Watson’s model by employing multivariate UC models that link the unobserved trend and cyclical components not only to observed log GDP, but also to observed price inflation \( \pi \). This is typically done by adding a so-called Phillips curve to Watson’s univariate model, which provides a link between changes in the level of price inflation and changes in the cyclical component of GDP.

For example, Gerlach and Smets (1999) (hereafter GS) add the following Phillips curve:

\[
\Delta \pi_t = \alpha_1 \Delta \pi_{t-1} + \alpha_2 \Delta \pi_{t-2} + \alpha_3 \Delta \pi_{t-3} + \gamma (r_{t-1} - \bar{r}_{t-1}) + \varepsilon_{\pi_t},
\]

Their relationship allows for rich dynamics in the evolution of changes in the rate of inflation \( \Delta \pi_t = \pi_t - \pi_{t-1} \) through the autoregressive coefficients \( (\alpha_1, \alpha_2, \alpha_3) \). In general, \( \gamma \) is assumed to be positive, so that a widening gap between actual GDP and the trend, captured by the cyclical component \( c_t \), leads to higher price inflation.

GS also modify the specification of the cyclical component of the model by incorporating information on the real federal funds rate (difference between the level of the Federal Reserve’s target interest rate and the average level of price inflation):

\[
c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \lambda (r_{t-1} - \bar{r}_{t-1}) + \varepsilon_{c_t},
\]

where \( r_t \) denotes the nominal U.S. federal funds rate and \( \bar{r}_t \) denotes the average level of price inflation over the previous four quarters. GS argue that this
modified cyclical equation is essentially a reduced-form aggregate demand function with $\lambda$ assumed to be negative, so that a rising real interest rate decreases aggregate demand. The other equations in the GS model are the same as the constant and time-varying drift UC models described above.

One of the key observations driving speculation that the U.S. had experienced a significant increase in its trend growth rate of GDP was that inflation was considered to be relatively low for such a rapidly expanding economy. To understand this argument, we must examine the relationship between inflation and trend growth embodied in the Phillips curve. An increase in the trend growth rate of GDP would, other things being equal, raise the level of the trend component and lower the level of the cyclical component, which would in turn imply a lower rate of increase in price inflation through the Phillips curve relationship. In light of this, GS’s model is particularly well suited to exploring the existence of the new economy, since it allows inflation also to affect the measurement of the trend component.

Moving onto the trend growth rate estimates, table 1 shows that the estimated multivariate UC constant drift model has an underlying trend growth rate of 3.32 percent, which is marginally higher than the estimated constant trend growth rate from the univariate UC model. The multivariate UC discrete-jump model parameters, on the other hand, suggest that while there was significant change in the trend growth rate in the earlier part of the sample, there was not a statistically significant increase in the trend growth rate of GDP in the 1990s. A similar picture emerges from the multivariate UC unit-root drift estimates reported in figure 9 (dark gray line). Based on these results, the trend growth rate changed little from the end of the productivity slowdown era through to the new economy period. In fact, the unit-root drift estimates suggest that after rising slightly in the late 1990s, the trend growth rate actually fell below the levels recorded in the productivity slowdown period.

Parameters governing the cyclical components are virtually identical across all three multivariate UC models. This finding is echoed by the similarity of the multivariate UC cyclical components under the three trend growth rate assumptions plotted in figure 10. These cyclical components suggest that, other things being equal, if the Fed had relied on
these multivariate UC models to estimate the cyclical component of GDP, its monetary policy response would have been invariant to its trend growth rate assumption over the late 1990s.

Overall, these multivariate UC models estimates provide no evidence in favor of the new economy, so what was the factor underlying the strong growth of GDP? Turning to estimates of the multivariate UC random trend components in figure 11, we see even stronger evidence than in the univariate UC model case that there is a clustering of positive random fluctuations to the level of the trend component of GDP in the multivariate model in the late 1990s. Again, this suggests that the high GDP growth rates of the late 1990s were the result of a level shift in the trend component, which was driven by a sequence of relatively large, positive random fluctuations that had a permanent effect on the level of the trend, but no permanent effect on the trend growth rate.

Conclusion

New economy advocates argue that the high productivity growth rates of the second half of the 1990s ushered in a permanent increase in the trend growth rate of U.S. GDP. I test formally whether there was a statistically significant change in the trend growth rate of GDP over the late 1990s. Using a number of widely used approaches to estimate the trend component of GDP, I find that there was variation of the trend growth rate of GDP over the latter half of the 1990s, but it was not statistically different from zero. I conclude that, at least in terms of GDP data, the U.S. was the same old economy in the late 1990s.

NOTES

1See Taylor (1993) for details.

2Details of data sources and dates used for estimation are reported in appendix A.

3I estimate the time-varying \( \lambda \) and \( \mu \) using standard linear spline and knot regression techniques, which restrict the estimated trend to be a continuous line; see Greene (1990) pp. 248–251 for details.

4At 2.94, my estimate of the productivity slowdown growth rate is slightly higher than the typical trend growth rate based on productivity growth rate estimates. As Seskin (1999) shows, this upward revision to the trend growth rate can be explained by the shift to current chain-weighted data, which also incorporates revisions to the measurement of business fixed investment that raised the average growth rate of GDP over the entire sample by roughly 0.3 percentage points.

5Orphanides and van Norden (2002) highlight problems in measuring the cyclical component of GDP in real time. The methods used in this article are subject to their critique of real-time estimates of the cycle. However, my main objective here is not to estimate a real-time cycle, but to document whether the trend growth rate of GDP changes over the latter half of the 1990s using the best available data, so their criticism is not relevant for the trend growth rate results presented here.

6Trend properties of the data used in this article are reported in appendix A. In particular, table A1 reports augmented Dickey–Fuller tests for log GDP. According to these tests, log GDP has a stochastic trend.

7Details of the techniques used to estimate the UC models are reported in appendix B. I use the following conventions when reporting parameter estimates or plotting unobserved components: The cycle is expressed as percentage deviations from the trend, while the underlying growth rate of the trend is expressed as annualized percentage rates. Plots of the unobserved cycle and trend components refer to the two-sided estimates generated by the Kalman smoother. Appendix C reports estimates of the other parameters of the unobserved component models.
APPENDIX A: TREND PROPERTIES OF GDP AND INFLATION DATA

An important assumption in Watson’s (1986) and Gerlach and Smets’ (1999) models is that the log of GDP has a unit root. Gerlach and Smets’ model goes one step further in assuming that price inflation also has a unit root. This section reports the results of augmented Dickey–Fuller (ADF) tests for nonstationarity using quarterly U.S. real chain-weighted gross domestic product (GDP) and consumer price index (CPI) data from 1961:Q1 to 2003:Q4. Note, the quarterly CPI data are calculated as averages of the three months in the quarter.

The left-hand side of table A1, panel A reports ADF t-statistics for cases with a constant and time trend, using various lags of the change in the log of GDP ($\Delta x_t = x_t - x_{t-1}$). These test statistics do not allow me to reject the null of a unit root in the log of GDP at typical levels of significance.

A potential time-varying model of the trend growth rate of GDP is a unit root without drift. A time-varying trend growth rate with a unit root would require a unit root in the change in the log of GDP $\Delta x_t$. The right-hand side of table A1, panel A reports ADF t-statistics for the change in the log of GDP. I am able to reject the null of a unit root in the growth rate of GDP at conventional levels of significance when a constant is included in the regression. This implies that the trend growth rate of real GDP is a stationary process. However, Stock and Watson (1998) argue that if the variance of the innovation to the trend growth rate is small, the growth rate of GDP will have a nearly unit moving average (MA) root. It is well-known that unit-root tests have a high rejection rate in the presence of large MA roots, which means that the reported ADF statistics are consistent with a model that has a unit root in the trend growth rate with a small innovation variance.

<table>
<thead>
<tr>
<th>TABLE A1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-root tests</td>
</tr>
<tr>
<td><strong>A: Real GDP</strong></td>
</tr>
<tr>
<td>Lags</td>
</tr>
<tr>
<td>Constant, trend</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td><strong>B: CPI inflation</strong></td>
</tr>
<tr>
<td>Lags</td>
</tr>
<tr>
<td>Constant, trend</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

Source: Author’s calculations based on GDP and CPI data from 1961:Q1 to 2003:Q4.

Panel B of table A1 repeats these experiments for inflation $\pi_t$, measured as the change in the log of the CPI. The left-hand side suggests that at conventional levels of significance, I cannot reject the null of a unit root in inflation. The right-hand side suggests that I can reject the null of a unit root in the change in the level of inflation ($\Delta \pi_t = \pi_t - \pi_{t-1}$) at conventional levels of significance. This implies that $\Delta \pi_t$ is a stationary process.
APPENDIX B: ECONOMETRIC ISSUES

I estimate all models using maximum likelihood. In each case the likelihood function is evaluated using the Kalman filter on the model’s state space representation. I simplify the estimation by transforming the models so that they are specified in first differences rather than levels of real GDP. For example, I estimate the univariate model with constant drift or drift with discrete jumps using the following structure:

\[
\Delta x_t = \mu_{\Delta t} + \Delta c_t + \varepsilon_t \quad \text{for} \quad i \leq t \leq j
\]

\[
c_t = \Phi c_{t-j} + \phi c_{t-j} + \varepsilon_t
\]

The advantages of this data transformation are twofold. First, the computation costs are lower because the state vector \( \xi_0 \) is reduced to current and lagged values of the GDP gap:

\[
\xi_0 = [c_j, c_{j-1}, \ldots, c_0]'
\]

Second, these components are assumed to be stationary, so I can use the exact likelihood to estimate the model by specifying the initial values of the state vector \( \xi_0 \) and the initial covariance matrix of the associated estimation error \( P_0 \) in terms of the population moments of the state vector:

\[
\xi_0 = 0 \quad \text{and} \quad P_0 = \sigma_e^2 \begin{bmatrix} 1 & \Phi/(1-\phi_1) \\ \phi_1/(1-\phi_1) & 1 \end{bmatrix},
\]

where \( \sigma_e^2 = \frac{\sigma^2}{1-\phi_1^2-\phi_2^2-2\phi_1\phi_2/(1-\phi_1)} \). This avoids the many problems associated with estimating models in levels, such as the unobserved component estimates depending critically on initial values.

The set up of the unit-root models is more complicated. The state vector is expanded to include current and lagged values of the growth rate:

\[
\xi_0 = [c_j, c_{j-1}, \mu_j, \mu_{j-1}]'
\]

and I must use the conditional likelihood to estimate the model’s parameters. I follow Harvey (1993) by explicitly using \( \Delta x_t \) as an estimator of \( \mu_{\Delta t} \) and noting that the variance of the associated estimation error is \( E[(\Delta x_t - \mu_{\Delta t})^2] = E[(\Delta c_t + \varepsilon_t)^2] \), which implies the following initial state vector and associated estimation error covariance matrix:

\[
\xi_0 = [0 \quad \Delta x_0 \quad \Delta x_{-1}]'
\]

\[
P_0 = \sigma_e^2 \begin{bmatrix} 1 & \phi_1/(1-\phi_2) & (1-\phi_1)/(1-\phi_2) & 2\phi_1 - \phi_1^2 - \phi_2 \\ \phi_1/(1-\phi_2) & 1 & \phi_1/(1-\phi_2) - 1 & (1-\phi_1)/(1-\phi_2) \\ (1-\phi_1)/(1-\phi_2) & \phi_1/(1-\phi_2) - 1 & 2(1-\phi_1)/(1-\phi_2) + \sigma_e^2 & (2\phi_1 - \phi_1^2 - \phi_2)/(1-\phi_2) + \sigma_e^2 \\ (2\phi_1 - \phi_1^2 - \phi_2)/(1-\phi_2) & (1-\phi_1)/(1-\phi_2) & (2\phi_1 - \phi_1^2 - \phi_2)/(1-\phi_2) & (1-\phi_1)/(1-\phi_2) + \sigma_e^2 \end{bmatrix}
\]

Fortunately, the innovations to the inflation equation are stationary variables, so I can follow the same approach to initializing the multivariate models.
## Table C1
### Univariate unobserved component model

#### A: Constant drift

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>SE  Q(16) LLF</td>
</tr>
<tr>
<td>3.30</td>
<td>(0.24) 20.49 -207.33</td>
</tr>
</tbody>
</table>

#### B: Discrete jump

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>SE  Q(16) LLF</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>(0.13) 2.92 1.28</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>(0.14) 0.37 0.37</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>(0.25) 1.66 -0.72</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>(0.12) 2.55 0.37</td>
</tr>
</tbody>
</table>

#### C: Unit-root drift

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>SE  Q(16) LLF</td>
</tr>
<tr>
<td>1.68</td>
<td>(0.13) 0.81 21.82</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>(0.13) 0.81 20.35</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>(0.13) 0.08 0.08</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>(0.14) 0.36 0.08</td>
</tr>
<tr>
<td>$\phi_5$</td>
<td>(0.12) 0.23 0.06</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses. SE denotes equation standard error. Q(n) is the Box-Ljung test for randomness of the errors distributed χ2n. LLF denotes the log of the likelihood function. Source: Author's calculations based on GDP data from 1961:Q1 to 2003:Q4.
### Table C2

#### Multivariate unobserved component model

##### A. Constant drift

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$SE$</td>
</tr>
<tr>
<td>3.32</td>
<td>0.24</td>
</tr>
<tr>
<td>(0.03)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.76</td>
<td>-0.85</td>
<td>2.78</td>
<td>0.25</td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.17)</td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\sigma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.59</td>
<td>-0.05</td>
<td>0.47</td>
<td>1.41</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

#### B: Discrete jump

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{73Q1:95Q4}$</td>
<td>$SE$</td>
</tr>
<tr>
<td>3.00</td>
<td>0.32</td>
</tr>
<tr>
<td>(0.49)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.12</td>
<td>-0.08</td>
<td>1.74</td>
<td>-0.83</td>
</tr>
<tr>
<td>(0.77)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.17)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\sigma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.60</td>
<td>-0.05</td>
<td>0.47</td>
<td>1.42</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

#### C: Unit-root drift

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Summary statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{73Q1:95Q4}$</td>
<td>$SE$</td>
</tr>
<tr>
<td>3.00</td>
<td>0.32</td>
</tr>
<tr>
<td>(0.49)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>$\sigma_3$</th>
<th>$\sigma_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.76</td>
<td>-0.85</td>
<td>2.73</td>
<td>0.24</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.16)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\sigma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.59</td>
<td>-0.05</td>
<td>0.47</td>
<td>1.41</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses. SE denotes equation standard error. $Q(n)$ is the Box-Ljung test for randomness of the errors distributed $\gamma_n$. LLF denotes the log of the likelihood function. Source: Author's calculations based on GDP data from 1961:Q1 to 2003:Q4.
REFERENCES


On May 4–6, 2005, the Federal Reserve Bank of Chicago will hold its 41st annual Conference on Bank Structure and Competition at the Westin Hotel in Chicago. Since its inception, the conference has encouraged an ongoing dialogue and debate on current public policy issues affecting the financial services industry. Each year the conference brings together several hundred financial institution executives, regulators, and academics to examine current policy issues.

The 2005 conference will address issues related to the evolution of lending. For most of the twentieth century, lending channels resembled silos: Businesses borrowed at commercial banks, home buyers borrowed at thrift institutions, and consumers borrowed at finance companies and credit unions. Loans were held on-balance-sheet, and the interest income they generated was the fundamental source of lender profits. Lenders had comparative advantages at gathering financial information, in part because they also provided savings, investment, and payment vehicles for their customers. Only the very largest corporations could borrow in public debt markets.

Advances in information processing, communications, and financial technologies have drastically altered the lending landscape. Lenders face more competition and borrowers have more choices. Banks play a key role in this mix, but their role is changing. Increasingly, both business borrowers and household borrowers are benefiting from
market financing—either directly as clients of investment banks or indirectly via asset securitization. As interest income from these arrangements flows to investors in corporate bonds and asset-backed securities, lenders have become more reliant on fee income from originating and servicing loans, providing back-up credit facilities, arranging loan syndications, and underwriting debt securities.

These ongoing changes have important implications for the structure and performance of the financial industry, for the amount of credit created and its distribution, and for macroeconomic growth and stability. For example, will improvements in information flows and securitization practices erode the need for relationship-based portfolio lending to small businesses and the primary role of small banks? Will the scale economies inherent in automated, transactions-based credit provision make the economics of lending untenable for small institutions? At the other extreme, will increasing financial market efficiency further erode the market for syndicated portfolio lending to large businesses? Are large banking companies consolidating in pursuit of size-based efficiencies or market power? Have commercial and investment banks already begun to exploit market power by ‘tying’ underwriting and lending for large corporate customers? Is there an endgame in which a handful of global banking companies will dominate international, or even domestic, credit markets? These and related public policy issues will be discussed at the 2005 conference.

As in past years, much of the program will be devoted to the primary conference theme, but there will also be a number of sessions on current industry issues. Some of the highlights of the conference include:

- Keynote addresses by Federal Reserve Board Chairman Alan Greenspan and House Financial Services Committee Chairman Michael G. Oxley (OH).

- Two panels discussing the conference theme: one addressing business lending and one concerned with retail lending. Participants on these panels include John F. Bovenzi, Chief Operating Officer, Federal Deposit Insurance Corporation; Robert Tetenbaum, Executive Vice President, First Manhattan Consulting Group; Allan Tubbs, President, Manquoketa State Bank; Michael D. Sharkey, President, LaSalle Business Credit; Jeff L. Plagge, President and Chief Executive Officer, First National Bank of Waverly, Iowa; Rick Spitler, Founding Partner, Novantas LLC; Tom Okel, Head of Syndicated Capital Markets, Banc of America Securities; Michael Frow, Senior Credit Officer, Harris Bank; Cathy Lemieux, Senior Vice President, Federal Reserve Bank of Chicago; and Sam Scott, President, Austin Bank of Chicago.

- Special luncheon presentation by Jerry Grundhofer, Chairman and Chief Executive Officer, U.S. Bancorp, Inc.

- Panel discussion on problems in pension funds, including Social Security reform. Panelists include Bradley Belt, Executive Director, Pension Benefit Guarantee Corporation; Randall Kroszner, University of Chicago; Zvi Bodie, Boston University; and Dennis Logue, University of Oklahoma.


As usual, the Wednesday sessions (May 4th) will showcase more technical research that is of primary interest to research economists in academia and research agencies. The Thursday and Friday sessions are designed to address the interests of a broader audience.

If you are not currently on our mailing list or have changed your address and would like to receive an invitation and registration forms for the conference please contact:

Ms. Regina Langston
Conference on Bank Structure and Competition
Research Department
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, Illinois 60604-1413

Telephone: 312-322-5641
email: rlangston@frbchi.org
The cost of business cycles and the benefits of stabilization

Gadi Barlevy

Introduction and summary

During the past half century, policymakers in the U.S. have consistently sought to chart a stable course for economic growth. The importance accorded to this goal does not merely owe to the views of select policymakers, but is mandated by law. In 1946, Congress passed the Employment Act, which encouraged the federal government to adopt policies that would lead to maximum employment and price stability. Evidently dissatisfied with the fulfillment of these goals, some 30 years later Congress passed the Full Employment and Balanced Growth Act in 1978 (also known as the Humphrey–Hawkins Act after its two coauthors), which strengthened the original Employment Act. Among other things, the 1978 law mandated that the Federal Reserve should specifically aim to maintain economic growth in line with the economy’s potential to expand. That is, policymakers were instructed to steer the economy in such a way as to ensure steady output growth, fast enough to maintain full employment but not so fast as to ignite inflation.

In stark challenge to the conventional wisdom that inspired such legislation, Robert Lucas argued in his influential 1987 monograph *Models of Business Cycles* that deviations from stable growth over the post-WWII (postwar) period in the United States were actually a minor concern that did not merit the high priority accorded to them under the law. More precisely, Lucas asked how much individuals should be willing to give up in principle to live in a world not subject to the degree of macroeconomic volatility the U.S. witnessed during this period. Assuming preferences that many economists view as a reasonable benchmark, he calculated that individuals would sacrifice at most 0.1 percent of lifetime consumption, prompting him to conclude that there would be little benefit to “devising ever more subtle policies to remove the residual amount of business cycle risk.”

Not surprisingly, Lucas’s results have attracted quite some controversy, and various researchers have revisited his calculation since his monograph was published. This article reviews the literature prompted by Lucas’s original observation, with an emphasis on two questions. First, does the subsequent literature confirm that postwar macroeconomic volatility is as minor a problem as Lucas’s original calculation suggests? And second, what do these estimates tell us about the inherent benefits from further pursuing stabilization policy?

I argue that the work that followed Lucas’s original calculation suggests his estimate significantly understates the true cost of postwar macroeconomic volatility. But at the same time, the mere fact that postwar business cycles were costly need not imply that attempting to neutralize them would have been highly desirable; that depends on what shocks were responsible for this volatility and whether they could have been effectively offset, questions economists have yet to fully resolve. As such, Lucas’s conclusion that there was little to gain from more aggressive stabilization may be correct. But even if there is little benefit from further stabilization, it need not follow that macroeconomic stabilization *per se* is unimportant. Society might have been much worse off had policymakers not pursued stabilization to the extent they did during the postwar era, and avoiding even greater volatility over this period should have ranked as a high priority.

The original Lucas calculation

In calculating the cost of business cycles, Lucas (1987) reasoned that people’s concern about

Gadi Barlevy is a senior economist and economic advisor at the Federal Reserve Bank of Chicago and a faculty research fellow of the National Bureau of Economic Research. The author is grateful to Craig Furfine, Jeff Campbell, Eric French, and Merritt Lyon for their thoughtful comments.
macroeconomic fluctuations is primarily due to how these fluctuations affect the amount of goods and services they get to consume. He then argued that we can view aggregate consumption expenditures each year as the amount of resources that can be used to satisfy such needs. Since aggregate consumption expenditures fluctuate over the business cycle, Lucas attributed the cost of business cycles to the fact that individuals are forced to contend with volatile and unpredictable consumption rather than stable and predictable consumption growth.

To be more precise, Lucas assumed consumption can be decomposed into a part that grows systematically over time and a part that fluctuates with prevailing economic conditions. Let us refer to the systematic part as trend consumption and denote its value in year $t$ by $C_t^*$. Actual consumption in year $t$, denoted $C_t$, will deviate from trend by a random percentage $\epsilon_t$, that is,

$$C_t = (1 + \epsilon_t)C_t^*.$$

The random deviation $\epsilon_t$ is assumed to have a zero mean and to be independent across time. That is, consumption $C_t$ will be equal to trend consumption $C_t^*$ on average, although in any given year it may be higher or lower than the trend, independent of what happened to consumption in previous years. Figure 1 shows log per-capita consumption from 1948 to the present, together with an estimate for trend consumption $C_t^*$ as Lucas suggested constructing it.

Lucas further assumed that the way individuals value consumption can be summarized with a simple utility function that assigns a value to every sequence of consumption expenditures \{\(C_t, C_{t+1}, C_{t+2}, \ldots\)\}. Let \(U(C_t, C_{t+1}, \ldots)\) denote the value a typical individual assigns to the corresponding consumption sequence. To quantify the cost of volatility, Lucas asked by what fraction we would need to increase lifetime consumption to make an individual with this utility function just as happy as in a world where consumption never deviated from trend, that is, where the individual could consume $C_t^*$ each year. Formally, Lucas calculated the value of $\mu$ for which

$$U((1+\mu)C_t, (1+\mu)C_{t+1}, \ldots) = U(C_t^*, C_{t+1}^*, \ldots).$$

The exact details of Lucas’s calculation are provided in box 1. Under his assumptions, the cost of business cycles can be approximated by the formula

$$\mu = \frac{1}{2} \gamma \sigma_{\epsilon}^2,$$

where $\gamma$ measures how averse an individual is toward risk and $\sigma_{\epsilon}^2$ denotes the variance of deviations from trend consumption. Thus, business cycles are more costly the more volatile is consumption (that is, the higher is $\sigma_{\epsilon}^2$) and the more averse individuals are to consumption volatility (that is, the higher is $\gamma$).

Using empirically plausible values for $\gamma$ and $\sigma_{\epsilon}^2$, Lucas arrived at a cost of 0.008 percent. That is, individuals would be willing to sacrifice no more than one-hundredth of 1 percent of their consumption to achieve macroeconomic stability. While acknowledging that his calculation abstracts from many important issues, Lucas argued it was unlikely that the cost of macroeconomic volatility would exceed 0.1 percent. A quick glance at figure 1 reveals why: Since aggregate consumption is not especially volatile, $C_t$ and $C_t^*$ are not dramatically different, and individuals will be close to indifferent between the two paths.

In the next few sections, I discuss the ways subsequent authors have criticized the above calculation. These are summarized in table 1 on page 34. The table is organized according to which feature of Lucas’s calculation each article modifies, and provides the range of cost estimates each paper presents as plausible.
### TABLE 1

**Alternative calculations for the cost of business cycles**

#### Panel A: Modify preferences and/or persistence of shocks

<table>
<thead>
<tr>
<th>Article</th>
<th>Cost (%)</th>
<th>Preference specification</th>
<th>Nature of consumption fluctuations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstfeld (1994)</td>
<td>0.02-0.5</td>
<td>Epstein-Zin preferences</td>
<td>Independent over time</td>
</tr>
<tr>
<td></td>
<td>0.01-1.8</td>
<td></td>
<td>Permanent</td>
</tr>
<tr>
<td>Dolmas (1998)</td>
<td>0.04-0.7</td>
<td>Epstein-Zin preferences</td>
<td>Serially correlated with autocorrelation of 0.98</td>
</tr>
<tr>
<td>Tallarini (2000)</td>
<td>2.1-12.6</td>
<td>Epstein-Zin preferences (but a much higher risk-aversion)</td>
<td>Serially correlated with autocorrelation of 0.99</td>
</tr>
<tr>
<td>Pemberton (1996)</td>
<td>0.01-1.1</td>
<td>First order risk-aversion</td>
<td>Independent over time</td>
</tr>
<tr>
<td>Dolmas (1998)</td>
<td>0.05-2.4</td>
<td>First order risk-aversion</td>
<td>Serially correlated with autocorrelation of 0.98</td>
</tr>
<tr>
<td></td>
<td>0.4-22.9</td>
<td>First order risk-aversion</td>
<td>Permanent</td>
</tr>
<tr>
<td>Otrok (2001)</td>
<td>0.004</td>
<td>Time non-separable preferences</td>
<td>Moderately persistent but not permanent</td>
</tr>
<tr>
<td>Alvarez and Jermann (2000)</td>
<td>&lt; 0.3</td>
<td>Estimated non-parametrically</td>
<td>Moderately persistent but not permanent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>from asset price data</td>
<td></td>
</tr>
</tbody>
</table>

#### Panel B: Calibrated risk to match household data rather than aggregate data

<table>
<thead>
<tr>
<th>Article</th>
<th>Cost (%)</th>
<th>Assumed effect of stabilization</th>
<th>Other remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imrohoroglu (1989)</td>
<td>0.30</td>
<td>Workers less likely to be unemployed for long periods</td>
<td></td>
</tr>
<tr>
<td>Atkeson and Phelan (1994)</td>
<td>0.02</td>
<td>Interest rates become less volatile</td>
<td></td>
</tr>
<tr>
<td>Krusell and Smith (1999, 2002)</td>
<td>-0.66</td>
<td>Earnings and interest rates both less volatile</td>
<td>Also match wealth distribution (so some households do not save)</td>
</tr>
<tr>
<td>Average across all households</td>
<td>3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Households with no wealth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storesletten et al. (2001)</td>
<td>0.6-2.5</td>
<td>Earnings and interest rates both less volatile</td>
<td>Reduction in earnings volatility is calibrated differently from Krusell and Smith</td>
</tr>
<tr>
<td>Average across all households</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Households with no wealth</td>
<td>1.5-7.4</td>
<td>Earnings less volatile</td>
<td>Earnings shocks assumed permanent</td>
</tr>
<tr>
<td>Krebs (2001)</td>
<td>7.5</td>
<td>Earnings less volatile</td>
<td></td>
</tr>
<tr>
<td>Beaudry and Pages (2001)</td>
<td>1.4-4.4</td>
<td>Earnings less volatile</td>
<td>Cost is for households with no wealth</td>
</tr>
</tbody>
</table>
**Alternative ways of implementing Lucas's calculation**

Even if one accepts the approach that underlies Lucas's calculation, it is still possible to quibble with the particular assumptions Lucas used to arrive at his estimate. I begin by reviewing criticisms that are in this spirit.

One problem concerns the particular function \( U(\cdot) \) Lucas used. Although this utility function is common in applied macroeconomics and has some empirical support, it has a difficult time accounting for attitudes toward certain types of risks, and as such might understate how much individuals dislike consumption risk. For example, individuals whose preferences correspond to Lucas's assumptions would be quite willing to invest in risky equity, while the large premium on stocks over bonds suggests that in practice individuals are more risk averse given they require a hefty return to invest in equity. One way to fix this is to allow for a higher degree of risk aversion. For example, whereas Lucas focused on the case where \( \gamma = 1 \), Obstfeld (1994) and Dolmas (1998) argue that a value of \( \gamma \) as high as 20 may be plausible, which would increase the costs relative to those Lucas reported by a factor of 20; but since the cost Lucas calculated was so small, the implied cost of business cycles is still no more than 0.5 percent of lifetime consumption. Both authors also consider a more general utility function advocated by Epstein and Zin (1991) that can be more easily reconciled with data on asset prices. This alternative specification suggests consumption volatility is not very costly, unless fluctuations in consumption are highly persistent, which for reasons I discuss below may not correspond to what we usually think of as business cycle volatility.

Tallarini (2000) uses the same generalized utility from Epstein and Zin (1991), but argues that far greater values of risk aversion are needed to accord with the premium on risky equity. As a result, he estimates the cost of business cycles to be much larger, between 2 percent and 12 percent of lifetime consumption. Pemberton (1996) and Dolmas (1998) consider a different utility specification known as first-order risk aversion. The implied cost of business cycles for this specification is only slightly larger than the one Obstfeld and Dolmas report and, for reasonable parameter

---

**TABLE 1 (CONTINUED)**

<table>
<thead>
<tr>
<th>Panel C: Stabilization affects long-run growth rather than leaving it unchanged</th>
<th>Article</th>
<th>Assumed effect of stabilization</th>
<th>Cost (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramey and Ramley (1991)</td>
<td>Increases output by avoiding mismatch between technology and economic conditions</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>DeLong and Summers (1988)</td>
<td>Decreases temporary declines in output (implied cost based on calculations in box 2)</td>
<td>1.6-1.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Cohen (2000)</td>
<td>Avoids temporary declines in consumption</td>
<td>1.0</td>
<td>0.3-0.8</td>
</tr>
<tr>
<td>Gall et al. (2003)</td>
<td>Lower distortions in the economy (higher consumption from given amount of labor)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Stabilization affects long-run growth rather than leaving it unchanged</th>
<th>Article</th>
<th>Assumed effect of stabilization</th>
<th>Cost (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matheson and Maury (2000)</td>
<td>Increases/decreases investment and long-run growth (cost only reflects this effect)</td>
<td>0.3-0.5</td>
<td>0.1-0.3</td>
</tr>
<tr>
<td>Epuial and Pommeret (2003)</td>
<td>Increases/decreases investment and long-run growth (cost only reflects this effect)</td>
<td>0.3</td>
<td>0.0-0.3</td>
</tr>
<tr>
<td>Barley (2004a)</td>
<td>Increases/decreases investment and long-run growth (cost only reflects this effect)</td>
<td>7.5-8.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Barley (2004b)</td>
<td>Increases/decreases investment and long-run growth (cost only reflects this effect)</td>
<td>0.3</td>
<td>0.0-0.3</td>
</tr>
</tbody>
</table>

---

Federal Reserve Bank of Chicago
Lucas’s calculation

Lucas’s calculation begins by assuming $C_t^* = \lambda C_{t-1}^*$, where $\lambda > 1$ measures the average growth rate for consumption during the post-WWII period. Actual consumption $C_t$ is then set equal to $(1+\varepsilon_t)C_t^*$, where $(1+\varepsilon_t)$ for all $t$ are independent and identically distributed lognormal random variables with mean 1 and variance $\sigma^2$. The standard deviation $\sigma$ can be computed from the standard deviation of $\ln(C_t/C_t^*) = \varepsilon_t$. Rather than estimate a trend, consistent with his specification for $C_t^*$, Lucas used the Hodrick–Prescott filter of aggregate consumption as his measure for $C_t^*$, from which he estimated $\sigma = 1.3\%$.

For his utility function, Lucas used the constant relative risk aversion utility function

$$U(C_t) = E_0 \left[ \sum_{t=0}^{\infty} \frac{(C_t^*)^{1-\gamma} - 1}{1-\gamma} \right].$$

Here, $\beta$ denotes the rate at which utility is discounted over time and $\gamma$ is equal to the coefficient of relative risk-aversion, that is, the higher is $\gamma$, the more reluctant the individual is to face a volatile consumption path. Lucas sets $\beta$ to 0.95 and $\gamma$ to 1, parameters that many macroeconomists would view as reasonable benchmarks. Standard arguments can be used to show that for $\gamma = 1$, the function $\frac{C_t^* - 1}{1-\gamma}$ reduces to $\ln C_t^*$.

A little algebra reveals that the solution to the equation

$$E \left[ \sum_{t=0}^{\infty} \frac{(1+\mu_0)(1+\varepsilon_t)(\beta\lambda)^t C_0}{1-\gamma} \right] = \sum_{t=0}^{\infty} \frac{(\beta\lambda)^t C_0}{1-\gamma}$$

yields the approximate formula $\mu = \frac{1}{2} \gamma \sigma^2$. For the coefficient of relative risk-aversion, the implied cost is thus $\mu = \frac{1}{2} (1)(0.013)^2 = 0.00008$, that is, less than one-hundredth of 1 percent.

values, does not exceed 1 percent of lifetime consumption. Otrok (2001) proposes still another specification for utility, but finds that plausible parameter values yield even more negligible costs.

Thus, most of the papers that propose alternative utility formulations continue to find small costs of business cycles, although a few argue the costs are significantly larger. So which of these specifications best captures individual preferences? Fortunately, Alvarez and Jermann (2000) developed an approach that does not require imposing a utility function, but infers one indirectly from a variety of asset prices, including the return on equity. They argue that asset prices reveal that individuals strongly dislike fluctuations in trend consumption growth, not cyclical fluctuations in consumption. To appreciate this point, consider figure 1. The growth rate of trend consumption $C_t^*$ varies over time: Per capita consumption grew at roughly 3 percent per year in the 1960s, compared with about 2 percent per year in the remaining postwar period. Alvarez and Jermann infer that the reason individuals require a high premium to hold stocks is that the return on stocks tended to be low in those periods when trend consumption growth was low. But the fact that households are so concerned with slow trend growth does not mean they are equally alarmed about temporary deviations from trend. Indeed, Alvarez and Jermann calculate that individuals would be willing to sacrifice at most 0.3 percent of lifetime consumption to eliminate only business cycle volatility in consumption, although they would sacrifice a lot more to avoid fluctuations in trend consumption growth. The preferences that are most consistent with data on asset prices therefore suggest the cost of business cycles is fairly small.

Another objection to Lucas’s calculation concerns his assumptions regarding deviations from trend consumption. Lucas assumed that the fact that consumption is below trend this year says nothing about whether it will be above or below trend next year. In practice, though, if consumption is below trend this year it is also likely to be below trend next year. Depending on how persistent shocks are and which utility function one assumes, this can affect the implied cost of consumption volatility. As evident in table 1, even if shocks are likely to persist for several years, the cost
of business cycles is typically less than 1 percent for most utility specifications. But when Obsfeld (1994) assumes shocks are permanent, so a fall in consumption today is expected to persist indefinitely, he finds that the cost of cycles can be as much as 1.8 percent. Dolmas (1998) shows that the cost of business cycles can be even larger—over 20 percent of lifetime consumption—when shocks are permanent and individuals have preferences that exhibit first-order risk aversion. Yet these permanent shocks are essentially changes in trend consumption growth, which presumably reflect changes in the economy’s potential, rather than temporary deviations from trend that policymakers can try to offset. The fact that the cost of permanent fluctuations in consumption can be so large thus mirrors the findings of Alvarez and Jermann that what individuals particularly dislike are fluctuations in trend consumption. Although society would be much better off if these permanent shocks were avoided, this is not a cost that could be avoided by conventional stabilization policy.

Using individual-level data: Preliminary results

A potentially more compelling criticism of Lucas’s estimate concerns its reliance on aggregate data. To see why using aggregate data might be problematic, suppose there was a small fraction of the population whose consumption was highly volatile, while consumption for everyone else was constant. Average consumption across the entire population would not appear very volatile; but for the unlucky few whose consumption is volatile, fluctuations will be quite costly. More generally, suppose that the small declines in aggregate consumption during recessions are driven by large declines in the consumption of a small but randomly chosen number of individuals, reflecting the fact that it is hard to predict exactly where the effects of downturns will be most severe. Since any individual runs the risk of a dramatic fall in his consumption, eliminating cyclical fluctuations might make all individuals much better off. In essence, focusing on aggregate consumption understates the volatility of consumption \( \sigma_x^2 \) that individuals face and, as such, understates the cost of business cycles.

Unfortunately, there is no time series on consumption at the level of households with which to carry out Lucas’s calculation.8 Instead, estimates of the cost of business cycles based on household data rely on more readily available observations on earnings. More precisely, researchers use individual earnings data to estimate a stochastic income process for a typical household, and then use theory to predict the consumption of a household facing this income process. They then calculate the cost of business cycles from the household’s predicted as opposed to actual consumption.

An important assumption in this line of work is that credit markets are “incomplete,” that is, that credit markets provide only limited protection against income risk. Households facing volatile incomes would naturally try to borrow when their incomes are low to maintain a constant level of consumption. Such borrowing will not allow them to escape consumption volatility altogether, since in recessions there will be more low-income households that wish to borrow and fewer high-income households willing to lend, raising interest rates and making it too costly to keep consumption constant. Still, with unlimited access to credit, one can show that individuals will be able to limit the volatility of their consumption to that of aggregate consumption, in which case Lucas’s original calculation would be applicable. But his calculation would not be applicable if households were limited in their borrowing, as is often the case in practice.

Formally, let \( y \) denote the annual labor income for a given individual in year \( t \). We begin by constructing a stochastic income process whose realizations mimic the incomes we observe for different households. For example, suppose income fluctuations were primarily due to periodic episodes of unemployment. We can then capture income fluctuations with a simple process whereby the income of an individual household can take on two values, one that corresponds to the average earnings of employed workers and one that corresponds to the average earnings of unemployed workers (for example, unemployment benefits). We can then estimate the transition probabilities between employment and unemployment from individual observations. A more sophisticated approach would also take into account the possibility that workers earn more on their jobs in boom times than they do in recessions.

Let \( q \) denote the net value of the individual’s asset holdings in year \( t \), and let \( r \) denote the interest rate paid on assets held between year \( t \) and year \( t + 1 \). Likewise, let \( c \) denote the individual’s consumption expenditures in year \( t \). Individuals are assumed to choose consumption expenditures to maximize utility \( U(q, c_{t+1}, \ldots) \), given the process for \( y \), and subject to the constraint that \( q_{t+1} = (1 + r)q_t + y_t - c_t \). This constraint states that the value of the assets an individual has at the beginning of year \( t + 1 \) is just the sum of the value of the assets he held in year \( t \), the interest he earned on these assets, and the wage income he earned, minus whatever he spent on purchases in year \( t \). To capture the limited ability of households to borrow,
we can add the restriction that \( a_t \geq 0 \) for all \( t \), that is, individuals are not allowed to carry any debt. A weaker restriction would allow for some amount of debt, so the lower bound on assets would be a negative number rather than zero. Solving this maximization problem yields a predicted sequence for consumption \( \{c_t, c_{t+1}, \ldots \} \).

Next, we use economic theory to forecast how the income process would change once aggregate fluctuations are stabilized. Denote the process for income in a stable world by \( \{y_t, y_{t+1}, \ldots \} \), so an asterisk denotes the value of a variable once aggregate fluctuations are eliminated. Once again, we can solve for the consumption decisions \( c_t^*, c_{t+1}^*, \ldots \) of an individual facing the constraints \( a_{t+1}^* = (1 + r_t^*)a_t^* + y_t^* - c_t^* \) and \( a_t^* \geq 0 \). Given the two consumption paths, we can once again ask how much we need to increase consumption in the world with volatility to make an individual as happy as when aggregate volatility is eliminated, that is, what value of \( \mu \) would ensure \( U((1 + \mu)c_t, (1 + \mu)c_{t+1}^*) = C(c_t^*, c_{t+1}^*) \).

The various papers that pursue this hypothesis differ in how they each chose to model the income process \( y_t^* \). Atkeson and Phelan (1994) argue that as long as income while employed and income while unemployed do not vary with the business cycle, the income process \( y_t^* \) should be identical to \( y_t \). To see why, suppose the probability an individual will be unemployed is 3 percent in a boom and 9 percent in a recession and that each year is equally likely to be a recession or a boom. From an individual’s perspective, then, the probability of being unemployed in some year in the future is \( \frac{1}{2} \times 3\% + \frac{1}{2} \times 9\% = 6\% \). Now, consider a stabilization policy where the government hires workers in recessions but not in booms to keep the probability of being unemployed constant at 6 percent. Each worker now faces the same earnings risk once as before, namely a 6 percent probability of being unemployed in any given year. But this does not mean individuals are not affected by stabilization. Without government intervention, demand for borrowing will be higher in recessions when more people are unemployed and, consequently, the equilibrium interest rate \( r_t \) will be higher as well. By contrast, in the stable environment, the interest rate \( r_t^* \) will be constant over time. Stabilization thus eliminates variations in the rate at which an individual can borrow or lend. For this reason, the consumption choices \( c_t^* \) in the stable economy may differ from \( c_t \). But when Atkeson and Phelan ask how much individuals would need to be as happy as when they get to consume \( c_t^* \), the answer is only 0.02 percent of lifetime consumption.

By contrast, Imrohoroglu (1989) argues that stabilization does affect earnings risk, although at the same time she ignores the interest rate risk that Atkeson and Phelan emphasize. Her argument relies on the observation that unemployment spells are typically short in booms but long in recessions, whereas in a stable environment unemployment durations would presumably be of average length. The virtue of stabilization is that it allows individuals to avoid long spells of unemployment, which are hard to save for. While stabilization also eliminates short unemployment spells, borrowing-constrained households do not suffer as much from eliminating short spells as they benefit from eliminating long ones. When Imrohoroglu computes the cost of business cycles assuming individuals cannot borrow and earn zero real interest on their savings, she finds a cost of business cycles of 0.3 percent. When she also allows individuals to borrow at a real rate of 8 percent (while saving at a rate of zero), the cost falls to a mere 0.05 percent. While her analysis ignores fluctuations in the interest rate over the cycle, recall that Atkeson and Phelan find these to be negligible. Thus, preliminary work on the cost of business cycles with incomplete markets appeared to reaffirm Lucas’s original conclusion.

**More recent work using individual-level data**

More recent work, however, has questioned this conclusion. The reason for the small cost of business cycles above is that interest rates are not particularly volatile over the cycle, nor are unemployment spells in the U.S. very long, even in recessions. Since households could easily save enough to sustain them through short periods of unemployment, the papers cited above conclude that business cycles are not especially costly. Yet there are two problems with this conclusion. First, fluctuations can contribute to earnings risk beyond just unemployment risk. For example, since wages are procyclical, workers who are laid off in recessions will re-enter the work force at lower wages that may remain low for far longer than the duration of a typical unemployment spell. Second, even though individuals could save for bad times, evidence on the distribution of wealth suggests a significant number of them do not. More recent work has taken these observations into account and suggests more significant costs of business cycles.

Consider first the work of Krussell and Smith (2002). They allow the interest rate \( r_t \) to vary over the cycle, so individuals face interest rate risk as described by Atkeson and Phelan (1994). At the same time, they follow Imrohoroglu (1989) in assuming that stabilization will allow individuals to avoid long spells of unemployment. But they also introduce two new features: 1) they assume stabilization has a more
significant effect on earnings risk than in Imrohoroglu’s formulation, in line with empirical evidence; and 2) they modify the model to accord with the observation that a considerable fraction of all households hold very little wealth.

Turning first to the effects of stabilization on earnings risk, Krussell and Smith incorporate Imrohoroglu’s observation that stabilization allows individuals to avoid long spells of unemployment. But they introduce two additional features. First, they assume that the wages households earn while employed vary over the cycle but would remain constant under stabilization, so \( \dot{y}^* \) would be less volatile than \( y \), even for households that avoid unemployment. Second, they assume stabilization lowers the risk of becoming unemployed. This can be motivated by the observation that some jobs that are profitable in booms turn unprofitable in recessions. Workers employed on those jobs would earn high wages in booms, but would be immediately laid off in the next recession. If these jobs remain profitable after stabilization, workers on these jobs would no longer have to fear unemployment from a downturn. However, workers would earn lower wages on these jobs under stabilization, since they would no longer earn the high wages they used to earn in booms.

In addition to changing the way stabilization affects earnings risk, Krussell and Smith modify Imrohoroglu’s model to accord with evidence on the distribution of wealth across households, specifically with the observation that wealth is highly concentrated. To do this, they allow for heterogeneity in discount rates across individuals. Households that are more patient save more and, as such, account for a disproportionate share of total wealth. Similarly, households that are more impatient hold very little wealth. While this leaves them vulnerable to periods of low consumption while unemployed, they are too impatient to cut back on their current consumption and save for when their income is low. By choosing the distribution of discount rates appropriately, Krussell and Smith are able to reconcile their model with the empirical distribution of wealth.

For households that are unemployed and have exhausted their borrowing capacity, Krussell and Smith estimate that eliminating fluctuations would be worth almost 4 percent of lifetime consumption. However, the cost of fluctuations for other individuals in the economy is much smaller and even negative for households with moderate savings (these households are not concerned about earnings volatility given their savings, and they like the fact that in the cyclical environment, wages are high precisely when they are more likely to be employed). Wealthy households do have a strong preference for stabilization, although this has nothing to do with volatility directly; rather, eliminating fluctuations would lead other households to cut back their precautionary savings, causing the supply of loanable funds to shrink and interest rates to rise, which obviously benefits those with high levels of assets. On the whole, Krussell and Smith find that the majority of households would be made worse off under stabilization, and averaging over all individuals implies business cycles are socially beneficial on net, although mildly so. As such, their findings hardly point to stabilization as a pressing social concern. But their results do illustrate that business cycles are costly for households with few assets.

Subsequent work has argued that Krussell and Smith themselves underestimate the degree of earnings risk individuals face. For example, although Krussell and Smith allow wages to fluctuate over the cycle, the degree to which they let wages vary with economic conditions depends on the predictions of a model rather than on direct evidence on earnings. When Storesletten, Telmer, and Yaron (2001) look at reported household earnings, they find that the standard deviation of earnings across households more than doubles in recessions, far more than implied by Krussell and Smith’s model. Moreover, Storesletten et al. find that earnings shocks are highly persistent, so that when a household’s income falls this year, for whatever reason, its earnings are likely to be low for far longer than in Krussell and Smith’s model. Using the same utility function Lucas considered, Storesletten et al. estimate that eliminating fluctuations would be worth 0.6 percent of lifetime consumption, while households with little savings (which in their model are young households that have yet to accumulate any wealth) would be willing to sacrifice 1.5 percent of their consumption. For somewhat higher degrees of risk aversion, but still within the range Lucas considered, they estimate the cost for the population as a whole at 2.5 percent of lifetime consumption, while those without any savings would be willing to sacrifice 7.4 percent.

Although Storesletten et al. assume earnings shocks are highly persistent, households can still protect themselves fairly well against these shocks by saving. This is because earnings are persistent, but not permanent. Krebs (2003) considers a similar model where shocks are permanent, so a fall in income today will lead expected income in all future years to fall by the same amount. In this case, individuals will not be able to borrow to offset negative shocks to their income, even when credit markets operate perfectly; after all, who would lend to an individual to cover earnings losses
that are never expected to be recovered? Krebs estimates that, overall, individuals with the same preferences as Lucas assumed would be willing to sacrifice 7.5 percent of lifetime consumption to eliminate fluctuations in this case. But it is hard to tell from the data whether earnings shocks are permanent or just highly persistent, and the cost of cycles is considerably smaller in the latter case.\textsuperscript{10}

Beaudry and Pages (2001) also consider the case where individuals do not protect themselves against earnings shocks. However, rather than allow for permanent shocks, they assume individuals have no incentive to save at the equilibrium interest rate. Moreover, rather than estimating earnings volatility from evidence on earnings dispersion as Storesletten et al. and Krebs do, they use data on the cyclicality of starting wages. Their logic is that, just as in earlier work, layoffs contribute to much of the earnings risk individuals face. However, unlike previous work, this is not because of the earnings workers forego while unemployed, but because laid-off workers typically re-enter the work force at a lower wage than they previously earned. While it is never a good thing to be laid off and have to start from scratch, it is particularly bad if you have to do so in a recession. They calibrate their model to data on the volatility of starting salaries over the cycle and, using Lucas's original utility function, estimate that individuals would be willing to sacrifice 1.4 percent of consumption to eliminate fluctuations in starting salaries over the cycle. When they allow for more risk aversion as in Storesletten et al., they estimate a cost of 4.4 percent. However, this cost is only borne by workers; employers in their model are assumed not to care about volatility, and the implied cost of business cycles for the population as a whole is smaller.\textsuperscript{11}

In sum, once we take into account evidence on the low savings rates of many households, as well as the fact that cyclical fluctuations can lead to persistent earnings declines, postwar business cycles start to matter; specifically, there is a core of households that are disinclined to save and as such would be willing to sacrifice between as much as 4 percent and 7 percent of lifetime consumption to avoid such volatility. Remaining households are likely to suffer less from cyclical fluctuations and may even benefit from them. The overall cost of cycles is thus more modest, but can still run as much as 2.5 percent of lifetime consumption.

The effects of volatility on the level of consumption

A separate problem with Lucas's calculation is his assumption on how stabilization affects the level of consumption. Lucas asserted that stabilization would eliminate deviations from trend, implying consumption would revert to its average level. But as various economists have since noted, the level of consumption might change in response to stabilization, so that stabilization might increase average consumption relative to the volatile economy.

The papers described in the previous section using household income data are immune to this criticism, since they derive consumption $c^*_t$ as the solution to a household problem rather than setting it to the average of observed consumption. However, they still abstract from some of the ways that stabilization can affect the level of consumption, and as such can still understake the true cost of business cycles. As in most of the literature that explores this hypothesis, my discussion will focus on aggregate data.

One critique along these lines comes from DeLong and Summers (1988). They argue that rather than steadying economic activity at its average level, stabilization would prevent economic activity from falling below its maximum potential, in line with the mandates of the Full Employment and Balanced Growth Act of 1978. Thus, stabilization policy would "fill in troughs without shaving off the peaks." While their discussion is couched in terms of output, one can easily adapt their argument for consumption. Let $C^*_t$ denote the level of consumption that would prevail in year $t$ in the counterfactually stable economy. Previously, $C^*_t$ also reflected the average of consumption; but now the two series are no longer assumed to be the same. Let $\varepsilon_t$ denote the percent deviation of actual consumption in year $t$ from $C^*_t$, i.e., $C_t = (1 + \varepsilon_t)C^*_t$. If consumption in the stable economy represents the maximum level consumption can attain, $\varepsilon_t$ must be less than or equal to zero. The average value of $\varepsilon_t$ is therefore negative, as opposed to zero. Consequently, the consumption path in the stable economy $C^*_t$ exceeds the average level of consumption in the volatile economy.

Just as Lucas used the assumption that $\varepsilon_t$ is zero on average to recover $C^*_t$ from data on $C_t = (1 + \varepsilon_t)C^*_t$, DeLong and Summers propose a way to recover $C^*_t$, from $C_t$ when $\varepsilon_t \leq 0$. Their approach is described in box 2. Alternatively, we can use data on business cycle peaks to isolate years when $\varepsilon_t = 0$, and then interpolate between these points to recover $C^*_t$. In particular, the National Bureau of Economic Research (NBER) has attempted to identify peaks and troughs in economic activity ever since 1850, which we can use to identify years in which $\varepsilon_t$ was equal to 0. This approach is also detailed in box 2. Both series are illustrated in figure 2 overleaf, together with the original data on aggregate consumption from figure 1. The
average deviation $\bar{\varepsilon}$ is 1.9 percent using DeLong and Summers’ approach and 1.6 percent using the series interpolated from NBER peaks. The cost of business cycles turns out to be roughly equal to this average, so these magnitudes also represent the amount individuals would sacrifice to attain $C^*_t$. In closely related work, Cohen (2000) finds a slightly smaller cost of business cycles of 1 percent, still much larger than the cost Lucas calculated.

The difference between Lucas’s estimate and the one that emerges from DeLong and Summers’ analysis stems from their different views of stabilization. Which of these is more compelling? Each imposes what it views as reasonable assumptions on the deviation $\bar{\varepsilon}$ between actual consumption and its level after stabilization to estimate $C^*_t$. But a more compelling approach would be to derive $C^*_t$ using economic theory, rather than imposing ad-hoc restrictions on $\bar{\varepsilon}$, to recover $C^*_t$ and let the theory dictate whether the level of consumption following stabilization will be higher than average consumption in the cyclical economy.

One explanation for why stabilization should increase consumption is that shocks affect the economy asymmetrically: Positive shocks boost economic activity less than negative shocks dampen it. Mankiw (1988) and Yellen and Akerlof (2004) sketch out such an argument and cite evidence that unemployment responds asymmetrically to changes in inflation, suggesting that if the Federal Reserve were able to stabilize inflation at its average level, unemployment would fall and more output could be produced and consumed. Mankiw estimates that stabilization should increase output on average by about 0.5 percent per year, while Yellen and Akerlof’s estimates suggest output would increase by between 0.5 percent and 0.8 percent. On a similar theme, Gali, Gertler, and
Lopez-Salido (2003) develop a formal model in which market frictions imply that welfare (and under certain assumptions, consumption) responds asymmetrically to employment fluctuations. They find that a policy that stabilizes employment would increase welfare by an amount equivalent to increasing lifetime consumption by between 0.30 percent and 0.75 percent.

Ramey and Ramey (1991) suggest another explanation for why stabilization ought to increase the average level of consumption. Their argument is based on the notion that firms need to pre-commit to a specific technology before they commence production. In an uncertain environment, firms may end up with a technology that is inappropriate for the scale of production they would have to undertake. Thus, volatile environments are more likely to involve inefficient production, resulting in lower average output. Ramey and Ramey estimate that fluctuations lower output by 1.7 percent on average, although they also note that if households are risk-averse they will sacrifice slightly more than this to avoid fluctuations. This is on par with the magnitudes suggested by DeLong and Summers.13

A third reason stabilization might change the level of consumption concerns its effect on capital accumulation. If individuals accumulate more capital in the stable environment, there will be more inputs available for production in the long run and, thus, average output will eventually be higher than in the volatile environment. However, as I discuss in more detail in the next section, the theoretical effects of stabilization on the capital stock are ambiguous; investment can either rise or fall in response to stabilization. For now, I simply note that the welfare effects associated with such changes are negligible and would not contribute much to the cost of business cycles. But the other explanations for why stabilization ought to increase average consumption suggest a cost of business cycles as large as 2 percent.

**The effects of volatility on consumption growth**

The previous section focused on scenarios in which eliminating fluctuations increases the level of consumption. Graphically, this implies that stabilization induces a parallel shift up in consumption from the path Lucas assumed, which is displayed in figure 1. But eliminating fluctuations may also affect the growth rate of consumption. I now discuss work that explores this possibility.

The most commonly cited reason stabilization should affect consumption growth concerns its effect on investment. The intuition for this is as follows: Since firms are likely to be more cautious about investing in uncertain environments, eliminating fluctuations should lead firms to accumulate capital more rapidly. This allows firms to produce more output, enabling households to enjoy more consumption and, presumably, making them better off. However, as I now explain, this line of reasoning turns out to be misleading.

First, eliminating volatility can just as plausibly discourage investment as encourage it. For example, recall that in the face of volatility, households choose to maintain precautionary savings to sustain them through periods of low earnings. Stabilization would mitigate the need for such savings. As savings become scarcer, interest rates would rise and might discourage firms from investing.14 But even if stabilization encourages investment, the resulting increase in consumption growth comes at a cost. This is because investment uses up resources that would otherwise have been used to produce consumption goods, so households get to enjoy less initial consumption. Whether households are better off under faster growth is therefore ambiguous.

To put it another way, the effects of stabilization on investment do not reflect a simple change in the rate at which consumption grows; rather, they involve changes in the trade-off between present and future consumption. In a well-functioning economy...
where households act in their own best interest, changes in this trade-off ought to reflect the preferences of households and, as such, make them better off. Hence, assuming trend consumption remains unchanged once the economy is stabilized ignores an implicit benefit from stabilization. But this benefit is likely to be modest, given that households already chose their consumption optimally in the volatile environment. In fact, when Matheron and Maury (2000) and Epaulard and Pommeret (2003) calculate the welfare cost of business cycles due to their effects on investment, they find effects of no more than 0.5 percent.

The reason an increase in investment has such a small effect on welfare is that most of the benefits from the faster growth it gives rise to are offset by lower initial consumption. But Barlevy (2004a) argues that eliminating fluctuations can increase consumption growth even when initial consumption is unchanged. This is because changes in investment affect growth asymmetrically; an increase in investment increases growth less than a similar decrease in investment decreases growth, reflecting among other things the inability of firms to undertake too many investment projects at once. In this case, simply eliminating fluctuations in investment without ever changing the level of investment should increase growth. Estimates reported in that paper suggest that if stabilization would steady investment at its average level, the growth rate of per-capita consumption would increase from 2 percent per year to about 2.35 percent per year, which is well within the range of historical variation in trend consumption growth.

Figure 3 illustrates how trend consumption $C^*_t$ from figure 1 would change if, in addition to no longer fluctuating around its trend, consumption grew by an additional 0.35 percentage points per year. Although the effect on growth is modest, its cumulative effects are large, and households would presumably significantly prefer this new consumption path. Indeed, Barlevy (2004a) estimates the cost of cycles due to their effect on growth at 7.5–8.0 percent of lifetime consumption, much larger than the cost of business cycles described so far.

Note that figure 3 assumes stabilization has no effect on average investment. But recall that stabilization might also lead to a change in the level of investment, so consumption may be steeper or flatter than captured by the figure. However, as noted earlier, in a well-functioning economy, changes in the trade-off between present and future consumption will only be to the benefit of households. In that case, households should be at least as well off without cycles as with the consumption path depicted in figure 3, even if stabilization causes investment to fall by enough to lead to a lower overall growth rate. What matters is not whether consumption actually grows more rapidly in the absence of fluctuations, but that stabilization makes it possible to grow more rapidly from the same amount of resources.

In the opposite direction, various papers have argued that business cycles facilitate rather than depress growth. One hypothesis relies on the idea of intertemporal substitution; firms can take advantage of the fact that productivity is lower in recessions to undertake growth-enhancing activities without having to sacrifice as much output. While there is some truth to this, Barlevy (2004b) argues that one of the main inputs into productivity growth, research and development, is concentrated precisely when its opportunity cost in terms of foregone output is highest, that is, in booms. Thus, at least with respect to research and development, business cycles force society to trade off present and future consumption less favorably, not more favorably, imposing a social cost that is estimated to equal 0.3 percent of lifetime consumption. This reinforces rather than contradicts the view that business cycles retard the economy’s growth potential, in this case by increasing the amount of foregone output required to achieve growth.
Shleifer (1986) offers a separate argument for why volatility might be essential for growth. His reasoning is that firms invest in developing new technologies to earn excess profits. If stabilization eliminates periods of high profits, it may discourage investment and growth. Shleifer develops an illustrative example in which the absence of fluctuations leaves the economy stagnant. Since the economy operates inefficiently in his model, the argument that changes in investment make households better does not apply, and the falloff in investment makes households worse off. However, recall from the previous section that stabilization is also likely to increase the level of economic activity and with it average profits. This partly mitigates the concern that stabilization would suppress the incentives to innovate.

Finally, Jovanovic (2004) argues that volatility is an unavoidable byproduct of growth, so stabilization may curtail growth. His argument is that growth involves experimentation: Firms try out new ideas, some of which fail spectacularly. If the only way to stabilize the economy is to preclude such experimentation, stabilization may lead to stagnation. However, it is not obvious that stabilization would necessitate suspending experimentation, as opposed to moderating the negative consequences of failure. Indeed, in Jovanovic’s model, reducing the volatility that results from experimentation would both facilitate growth and make society better off.

**Taking stock: How costly is postwar volatility?**

Taken together, the research that followed up on Lucas’s original insight regarding the cost of postwar U.S. business cycles has raised important shortcomings in his approach. On the one hand, Lucas correctly pointed out that aggregate consumption does not fluctuate very much over the business cycle, so an individual household whose consumption mirrored aggregate consumption would not be much better off if these fluctuations were smoothed out. This conclusion proves to be robust. But in a world with imperfect credit markets, the consumption of individual households may be far more volatile than aggregate consumption, and as such they would benefit more from eliminating macroeconomic volatility. Even when we take into account wealthier households that are not much affected by business cycles, the average cost to society can be as large as 2.5 percent of aggregate consumption per year.

Beyond the direct cost of consumption volatility, there is evidence that business cycles impose an even larger indirect cost through their effect on the level and growth rate of economic activity. That is, living in a volatile world not only forces households to contend with unpredictable consumption, but also to consume less than they would otherwise. These costs are not mutually exclusive of the cost of higher uncertainty, so the true cost of business cycles relative to a world with no fluctuations should be the sum total of these costs. The final tab comes to over 10 percent of lifetime consumption, an unquestionably large cost.

The costs are based entirely on the way business cycles impact consumption. But as various commentators have noted, business cycles might be costly in other ways as well. For example, they may force households to work a different number of hours each, something they may be just as reluctant to do as varying their consumption over time. Likewise, business cycles may make households anxious about the prospect of earnings losses, even those whose incomes are spared. There is probably some truth to these arguments. However, one can easily fall into the trap of adopting an utopian view of what stabilization can achieve. By restricting attention to the fairly conventional and, more importantly, measurable ways by which business cycles affect consumption, the work surveyed above makes a compelling case that postwar business cycles were quite costly after all.

**Policy implications: Is stabilization an important priority?**

The fact that postwar business cycles were so costly raises two immediate questions for policymakers. First, should policymakers have acted more aggressively to stabilize the economy during this period than they actually did? And second, is stabilization an important priority that should guide policymakers, as current law dictates? I now argue that despite the apparently large costs of business cycles over the postwar period, it is far from obvious that society would have been much better off if policymakers had pursued a more aggressive stabilization, since at least some of the shocks that were responsible for cyclical fluctuations over this period could not have been easily offset. At the same time, the fact that even modest amounts of volatility can impose such a large social cost reaffirms that stable growth should be an important goal for policymakers. In other words, even if it is not possible to defend against all sources of volatility, including potentially those responsible for much of the volatility during the postwar period, preventing the economy from becoming even more volatile should certainly rank as a high priority.

In his original monograph, Lucas reasoned that since the cost of business cycles is so small, there is little to be gained from further stabilization. In revisiting his estimates, some of the papers cited above have argued that the inverse is also true, that is, the fact that the implied cost of business cycles is so
large implies that the benefits to more aggressive stabilization must be substantial. But just because business cycles are costly does not automatically imply that stabilization is desirable; instead, that depends on what causes business cycle fluctuations, what tools are available to policymakers, and whether these tools can effectively offset the underlying shocks. Even if Lucas’s original calculation understates the cost of business cycles, his conclusion that further stabilization is unwarranted may very well hold true.

In his recent review article, Lucas (2003) argues that evidence on the nature of cyclical fluctuations over the postwar period suggests there was very little scope for policymakers to pursue stabilization more aggressively. Reviewing the evidence on the sources of output volatility during the postwar period, he finds that at most one-third of the variation in output can be attributed to monetary shocks, which the Federal Reserve presumably has the best chance of offsetting. The remaining 70 percent of output volatility is due to changes in real economic variables. For example, one shock to real economic variables during this period was the sharp increase in oil prices in the 1970s. A dramatic run-up in the price of oil raises production costs and affects the economy’s potential for producing goods in the short run, that is, as long as existing production technologies are still in place. In this case, there is probably little that policymakers can do to successfully stabilize the economy. At best, they can try to offset the shock by lowering other aspects of production costs, but such intervention can easily do more harm than good by distorting firms’ incentives to abandon more costly energy-intensive technologies. In fact, one can formally show that, at least under certain assumptions, policymakers should not try to offset exogenous fluctuations in real economic variables. Assuming these assumptions were met, policymakers would have at best been able to reduce macroeconomic volatility by one-third, and the benefits to pursuing more aggressive stabilization would be far more modest than the implied cost of aggregate fluctuations.

However, one has to be careful in interpreting evidence on the source of fluctuations. For example, consider fluctuations in aggregate productivity over the business cycle. These would be counted as fluctuations in real as opposed to monetary factors. As pointed out above, if these changes are driven by technological considerations, for example, changes in the economic environment that affect the viability of existing technologies such as a change in the relative price of a key input like oil, there may be little policymakers can do. But fluctuations in aggregate productivity might instead reflect fluctuations in variables that policymakers could affect. For example, Benhabib and Farmer (1994) develop a model in which if firms are optimistic about economic conditions, they will choose to operate at a larger scale, which in turn contributes to raising aggregate productivity and reaffirms their decision to operate at a larger scale. But if firms are pessimistic about economic conditions, they will choose to operate at a smaller scale, resulting in lower aggregate productivity. In this case, policymakers might be able to credibly announce policies that dissuade firms from being pessimistic; for example, they might pledge to pursue an accommodative policy if productivity were low. If firms find it optimal to expand their scale under easy monetary policy, such a policy would preclude the economy from settling at a low level of productivity. Policymakers could then stabilize fluctuations by affecting expectations, a point Benhabib and Farmer themselves allude to. The extent to which the large cost of postwar business cycles could have been avoided through prudent policy thus depends on what forces were responsible for this volatility in the first place.

Without further research as to the underlying source of business cycle fluctuations, then, we cannot reject Lucas’s conclusion that there was little to be gained from pursuing a more aggressive stabilization over this period. Nevertheless, the fact that even small amounts of volatility are of such great consequence suggests that the answer to our question whether stabilization should rank as a high priority for policymakers is yes. Lucas himself was careful in his original monograph to argue that while there is little to gain from eliminating residual risk above and beyond whatever stabilization policies were already being pursued at the time, this does not invalidate the potentially grave importance of existing stabilization policies. For example, he readily acknowledged in his monograph that “fluctuations at the pre-Second World War level, especially combined as they were with an absence of adequate programs for social insurance, were associated with large costs in welfare.” This is confirmed in recent work by Chatterjee and Corbae (2001), who show that the same calculation by Imorhoroglu (1989) that yields such small costs of business cycles for the postwar period implies individuals should have been willing to sacrifice more than 6 percent of lifetime consumption to avoid prolonged episodes such as the Great Depression, since very long unemployment spells are very costly. Incorporating the other features described in this survey would magnify this cost even more. To the extent that the alternative to the stabilization policies that
were pursued in the postwar period was the risk of another Great Depression, there can be no dispute that prudent policies that keep the economy relatively stable are an important priority, especially given the argument advanced by some that it was bad policies that either exacerbated or prolonged the Depression.15

That said, one does not need the extreme of the Great Depression to appreciate the benefits inherent in stabilization policy. As the work surveyed in this article reveals, even a modest amount of macroeconomic volatility can impose significant social costs. The fact that there are some shocks policymakers are unable to do much about, and that such shocks may have accounted for a significant share of the macroeconomic volatility during the postwar period, does not take away from the observation that households are likely to be significantly better off in stable environments than in volatile ones. Even if policymakers were not in a position to stabilize much more aggressively than they did during the postwar period, they may still have played an important role in safeguarding the economy from any additional shocks that would have made output even more volatile.

**Conclusion**

Economists are split as to whether postwar business cycles were costly. On the one hand, there are those who accepted Lucas’s original conclusion, a view reinforced by early work that appeared to confirm his results even after accounting for greater degrees of risk aversion and the fact that credit markets provide only incomplete protection against earnings risk. On the other side are those who from the outset dismissed Lucas’s conclusion as implausible and remained convinced that stabilization is an important policy goal, even if they didn’t always offer much to directly counter his argument. This article argues that more recent work that explores particular features absent from Lucas’s calculation reveals that postwar business cycles were in fact costly, but that this does not necessarily imply that more aggressive stabilization during this period was warranted. Determining whether policymakers should have acted more aggressively requires a better understanding of what forces are ultimately responsible for business cycle fluctuations, a difficult question that economists are slowly but surely making progress on. But even if ultimately there wasn’t much more that policymakers could have done to further insulate the economy from cyclical shocks during this period, maintaining a stable growth path as mandated by the Full Employment and Balanced Growth Act of 1978 does appear to be a highly desirable goal. To the extent that policymakers prevented the economy from being even more volatile during this period, then, they deserve great credit.
NOTES

1Two other recent surveys are Lucas (2003) and Yellen and Akerlof (2004). Each reaches a somewhat different conclusion than the present survey on at least one of these questions.

2Subsequent work has argued for omitting expenditures on durables, since utility depends on the total outstanding stock of durable goods rather than on the amount of durable goods purchased in the current year. The implied cost of volatility using nondurable consumption is not dramatically different.

3That is, \( C^* \) is the Hodrick–Prescott filter of the original consumption series \( C \). Since I estimate this from annual data, I use a weight of 100. Lucas’s original calculation was based on quarterly data.

4Campbell and Cochrane (1995) similarly argue the equity premium implies a large cost of business cycles.

5DiTella, MacColloch, and Oswald (2003) and Wolfers (2003) propose using survey data on how happy people feel as another way of estimating the cost of cycles without imposing a particular utility function. For example, Wolfers regresses well-being data on the mean and variance of unemployment to arrive at a tradeoff between the two. One could do the same with the mean and variance of consumption; however, while consumption grows over time, average reported well-being does not. This incongruity suggests either individuals do not strongly prefer more consumption to less or, more likely, that well-being measures are not directly comparable over time.

6However, it is possible to disaggregate consumption at the level of individual states, as in Rober and Pallage (2002). They find that retail sales at the state level are more volatile than at the national level, suggesting the effects of macroeconomic shocks are concentrated among a subset of states. Accordingly, the cost of volatility they find is somewhat larger than Lucas computed from total U.S. data.

7Atkeson and Phelan do not deny that unemployment duration varies over the cycle; rather, they argue stabilization makes long spells less likely to occur at the same time others experience long spells, rather than less likely to occur at all. Which view is more reasonable depends on the underlying model and the nature of stabilization.

8The 2002 paper is a revised version of their 1999 paper; my discussion is based on the 2002 version.

9There appears to be some confusion about this in the literature. Several papers claim that Storesletten et al. assume earnings shocks are permanent, when in fact they do not.

10Turnovsky and Bianconi (2005) also consider a model where shocks are permanent. But they assume stabilization reduces the average level of volatility rather than its variation over time. Moreover, they allow households to vary their labor supply in response to shocks. Their estimate for the cost of cycles is about 2 percent.

11Several papers claim Beaudry and Pages obtain large costs because they assume stabilization eliminates all idiosyncratic earnings risk. While it is true that workers in their model face no risk in the stable economy, Beaudry and Pages calibrate the earnings loss workers suffer in their model to the extra amount workers lose when they are laid off in recessions as opposed to booms, not the (much larger) amount workers lose whenever they are laid off. In particular, workers who are laid off in booms in their model experience no wage losses. Thus, their welfare estimates only reflect the gains from eliminating the cyclical part of idiosyncratic risk, not the gains from eliminating all idiosyncratic earnings risk.

12Technically, this asymmetry corresponds to the notion that consumption is a concave function of whatever variable is being stabilized.

13Porter and Puch (2004) make a similar point, although in their framework firms commit to a price rather than to a technology. While they demonstrate that this commitment magnifies the cost of business cycles, they view their model as too stylized to yield informative estimates for the true cost of business cycles.

14Even ignoring precautionary savings, uncertainty may encourage rather than discourage firms from investing. With more volatility, profits will be higher if uncertainty is resolved favorably but no lower if uncertainty is resolved unfavorably, as long as firms can cut their losses by shutting down or adjusting their labor hiring. While this point has long been recognized in the investment literature, it has not figured much in work on the cost of business cycles, where the notion that firms can cut their losses is typically ignored.

15Chatterjee and Corbae’s estimates assume policy did not change between the postwar and prewar period. However, since their results assume downturns of the magnitude of the Great Depression are rare, given that they failed to occur in the postwar period, their 6 percent would represent a lower bound on the true cost of eliminating these crises.
REFERENCES


A stable money demand: Looking for the right monetary aggregate

Pedro Teles and Rui lin Zhou

Introduction and summary
The stability of a money demand relationship has been a major concern in monetary economics for the last 50 years. It is conventional to call the relationship between real money, a nominal interest rate, and a measure of economic activity a money demand relationship. A stable relationship between these variables helps answer important questions such as the following: What is the average growth rate of money that is consistent with price stability, given the average growth of the economy and a stable nominal interest rate? Knowledge about the response of money demand to changes in the nominal interest rate may also help quantify the welfare gains from a low average inflation rate.

In an essay in honor of Allan Meltzer, Lucas (1988) reassesses the evidence on the stability of the money demand estimated by Meltzer (1963) and justifies that stability not only on empirical grounds but also on theoretical ones. He shows that there is a theoretical equilibrium relationship between real money, a nominal interest rate as a measure of the opportunity cost of money, and gross domestic product (GDP) as a measure of transactions that is not exactly a money demand, but that is indeed stable. He estimates that equilibrium relationship using the monetary aggregate M1 as the measure of money with data up to 1985 and argues that there is a stable relationship between those variables with a unitary income elasticity and with a strong negative response of real balances to the nominal interest rate (see box 1 for definitions of the different monetary aggregates).

The relationship estimated by Lucas (1988) holds very well until the mid-1980s but not well at all after that. This could be because the demand for money is not a stable relationship after all, contrary to what the simple model would suggest. Another conclusion, which is our view, is that the measure of money is not a stable measure. In particular, we argue that technological innovation and changes in regulatory practices in the past two decades have made other monetary aggregates as liquid as M1, so that the measure of money should be adjusted accordingly. We show that once a more appropriate measure of money is taken into consideration, the stability of money demand is recovered.

Banking deregulation in the 1980s and 1990s and financial innovation in the 1990s associated with the development of electronic payments indeed suggest we need to reconsider the measure of transactions demand for money. Until the end of the 1970s, the transactions demand for money was well approximated by M1. Since then, however, a series of sweeping regulatory reforms and technological developments in the banking sector have significantly changed the way banks operate and the way people use banking services and conduct transactions. First, the Depository Institutions Deregulation and Monetary Control Act of 1980 abolished most of the interest rate ceilings that had been imposed on deposit accounts since the Banking Act of 1933 and authorized nationwide negotiable orders of withdrawal accounts (NOWs), which are interest-bearing checking accounts classified in M1. Furthermore, the Garn–St Germain Depository Institutions Act of 1982 authorized money market deposit accounts (MMDAs), interest-bearing savings accounts that can be used for transactions with some restrictions. MMDAs are classified in M2...
behavior use reclassify with Federal Meltzer century, Figure Lucas gate the growinghouse growing electronic tertiary for and demand sake and the difficulty showing Lucas’ breakdown (see box 1). These two major banking reforms blurred the traditional distinction between the monetary aggregates M1 and M2 in their transactions and savings roles. Second, the rapid development of electronic payments technology and, in particular, the growing use of credit cards and the automated clearinghouse (ACH) as means of payment, reinforced the effect of the banking reforms in slowing down the growth of M1. Both credit cards and ACH transactions can be settled with MMDAs and, therefore, with M2 rather than M1. Third, the widespread adoption of retail sweep programs (discussed in detail later) by depository institutions since 1994, which reclassify checking account deposits as saving deposits overnight, reduced the balances that were classified in M1 by almost half.

These fundamental changes in the regulatory environment and the transactions technology justify the use of a different measure of money after 1980. The measure MZM (money zero maturity) includes balances that can be used for transactions immediately at zero cost and was initially proposed by Motley (1988) and Poole (1991) as a more appropriate measure of the transactions demand for money (see box 1). We show that changing the monetary aggregate measure from M1 to MZM from 1980 onward preserves the long-run relationship between real money, the opportunity cost of money, and economic activity up to a constant factor.

In the next section, we show evidence of the difficulty in explaining the behavior of M1 with the behavior of GDP and the nominal interest rate. Then, we discuss why MZM, rather than M1, is an appropriate measure of the transactions demand for money in the past two decades. Finally, we estimate a money demand equation derived from a simple transaction technology model, using M1 as the monetary aggregate before 1980 and MZM after 1980 and obtain evidence in support of the stability of money demand.

An unstable demand for M1

Figures 1 and 2 reproduce figures 1 and 4 in Lucas (2000), extending the data through 2003. Figure 1 suggests that, over the course of the past century, movements in the ratio of M1 to nominal GDP have been inversely related to movements in the short-term nominal interest rate. Following Meltzer (1963), Lucas (1988) uses data up to 1985 to estimate a money demand equation, using M1 as the measure of money and a short-term nominal interest rate as the measure for the opportunity cost of money, and confirms Meltzer’s result that the income elasticity is about 1.0 and the interest elasticity is high. Lucas (1988) reports an interest rate semi-elasticity between 0.05 and 0.1, which for an interest rate of 4 percent corresponds to an interest elasticity between 0.2 and 0.4. Using data from 1900 through 1994, Lucas (2000) reports an interest elasticity of 0.5, consistent with a shopping time model for money demand. The money demand equation derived in Lucas (2000) is

\[ \frac{M_t}{P_t} = \alpha Y_t^{\gamma} i_t \]

where \( M_t \) is the monetary aggregate measured by M1, \( P_t \) is the price level, \( Y_t \) denotes the aggregate output level, \( i_t \) is the short-term nominal interest rate, and the interest elasticity is \( \gamma = 0.5 \), while \( \alpha \) is a constant term. Real money responds to output with a unitary elasticity and negatively to the nominal interest rate with a relatively large elasticity, so that in response to a 1 percent increase in the nominal interest rate,
response to the interest rate movements over time would be less and less pronounced. The constant term also changes across the three periods, corresponding to the increased inability to explain the low growth in M1 with movements in economic activity and the nominal interest rate. Ball (2001) argues that the data after 1987 represent evidence against a stable money demand. He estimates a linear relationship between the logarithm of real money, the logarithm of output, and a nominal interest rate for subperiods of 1903–94. For the period 1903–87 the evidence is consistent with a stable relationship with a unitary income elasticity and a relatively high interest elasticity, as shown by Lucas (1988) and Stock and Watson (1993). However, the need to account for the low reaction of M1 to lower interest rates and higher output after 1980 lowers both the estimated interest elasticity and income elasticity. The relatively low income and interest elasticity in the postwar period (1947–94) are significantly different from the unitary income elasticity and relatively high interest elasticity in the prewar period (1903–45), leading Ball to argue against a stable long run money demand.1

Figure 1 plots the actual and estimated real balances using the money demand equation above with an interest elasticity γ = 0.5. Clearly, one would expect a larger reaction of real balances to the lower interest rates in the 1980s and 1990s. A lower elasticity of 0.32, instead of 0.5, would still not get close to being consistent with the actual low growth in M1.

This is apparent from Figure 3, where we plot the logarithm of the ratio of M1 to nominal GDP against the logarithm of the nominal interest rate for the period 1900–2003. Figure 3 indicates that there could be a different money demand relationship for each of the three periods 1900–79, 1980–94, and 1995–2003. The solid line corresponds to the estimated elasticity of 0.32 for the entire 1900–2003 period. The interest elasticity for the three subperiods would be 0.26, 0.12, and -0.07, respectively, so that the

---

1. Ball (2001)

---

Figure 1: M1/nominal GDP and the nominal interest rate, 1900–2003

Figure 2: Actual and estimated real balances M1/P, 1900–2003 (interest elasticity of 0.5)
Measuring money used for transactions

In this section, we argue that M1 was a good measure of money used for transactions before major developments in banking regulation and financial innovation starting in the early 1980s. Since then, a measure such as MZM has become more appropriate.

Figure 4 (p. 54) shows the trend growth of all four monetary aggregates: M1, M2, M3, and MZM since 1959 (data for MZM are available since 1974). In particular, since 1980, M1 has grown at a low rate (5.1 percent) and flattened after 1994. In contrast, average MZM growth has been 9 percent since 1980. The rapid expansion in MZM is evident in the figure; its value surpassed that of M2 in 2001.

Before 1980, M1, consisting of currency, non-interest-bearing demand deposits, and a very small amount of interest-bearing checkable deposits (see figure 5 (p. 54) and discussion in the next section) was the primary transaction monetary aggregate. The main components of M2, other than M1, were savings deposits, mostly passbook savings accounts on which checks could not be written, and small time deposits. Neither could be directly used for transactions. The other component of M2, retail money market mutual funds (MMMFs), a nonbank financial instrument (some have restricted check-writing capacity) developed in the mid-1970s and remained very small, as shown in figure 5. Therefore, there was a clear distinction between M1 and the components of M2 other than M1 before 1980. The former could be used for transactions at zero cost and did not bear interest, while the latter were interest-bearing instruments that could not be directly used for transactions. Since then, this distinction has become
less clear-cut. Three major developments in banking regulation and financial innovation are responsible for the change.

Financial innovation and regulatory reform since 1980

Banking deregulation

The banking deregulation that ensued in the late 1970s and early 1980s changed the banking industry landscape from a highly regulated one into a fairly competitive one. An unavoidable consequence of the deregulation was the blurring of the various components of M1 and M2 as transaction/saving instruments.

The reform started in the 1970s when many commercial banks and depository institutions were struggling to survive in the high inflation, high interest rate environment, with their hands tied by many regulations, in particular, Federal Reserve Regulation Q. This regulation prohibited interest payment on demand deposits and imposed interest-rate ceilings on time and savings deposits. The first move toward deregulation was the authorization granted by several northeastern states to state-chartered mutual saving banks, and later other depository institutions, to offer NOWs, an interest-bearing transaction account. Other products or services designed to provide consumers with more efficient cash management tools developed at the same time. For example, commercial banks and thrifts were able to provide prepaid ranged automatic transfer services (ATS) from consumers’ savings accounts to their checking accounts, customers could transfer their savings balances to checking remotely, and federally chartered credit unions were allowed to issue share drafts. These innovations were officially sanctioned by the Depository Institutions Deregulation and Monetary Control Act in 1980. More specifically, the act eliminated most of the interest rate ceilings on time deposits and savings accounts and authorized the use of checkable NOW accounts and other interest-bearing accounts (such as ATS and share draft accounts at credit unions) by individuals and non-profit organizations. The privilege was extended to all levels of government agencies in 1982. The only exception is demand deposits of corporations, on which the 1933 prohibition of interest payment remains in effect today. These regulatory changes allowed depository institutions to compete more effectively for funds; they also removed the impediments for depositors to earn the market rate of return on their transaction balances. The direct consequence of the act is the prevalent use of interest-bearing checking accounts.

A second major regulatory banking reform was the Garn–St Germain Depository Institutions Act of 1982. It authorized the creation of money market deposit accounts (MMDAs) to compete with MMMFs. Classified as an M2 account, an MMDA...
is an interest-bearing account that carries no reserve requirements. The account offers limited transaction capacity: no more than six withdrawals by check or pre-authorized transfer per month, but no limit on deposits or number of withdrawals from an ATM, by mail, or at a branch. This act led to a substantial increase in the use of checkable savings accounts for transactions.

The deregulatory measures of the early 1980s, allowing for interest payments on checking accounts and checking privileges on savings accounts, blurred the distinction between transaction and saving deposits, consequently blurring the distinction between M1 and M2.

**Electronic payments**

Following the banking deregulation in the 1980s, the rapid development of electronic payments in the 1990s also fostered the use of components of broader monetary aggregates for transaction purposes. Credit cards are particularly responsible for this.

Credit cards are often used as a substitute for cash, check, and debit card transactions. Monthly balances on a credit card can be paid with an automated clearing house (ACH) transaction or a check written on a checking account or checkable savings account. The fact that there is a single payment at a certain date reduces the need to maintain high daily balances in checking accounts to meet the uncertain sequence of transaction and payment flows. This reduction is reinforced by the fact that it is possible to use checkable savings accounts to pay for credit card balances. The total number of credit and debit card transactions almost tripled in 1990s, from 10.8 billion in 1990 to 30 billion in 2000 (Humphrey, 2002).

The ACH is another important development in electronic payments. ACH is a nationwide mechanism that processes electronically originated batches of credit and debit transfers. ACH credit transfers include direct deposit payroll payments and payments to contractors and vendors. ACH debit transfers include consumer payments on insurance premiums, mortgage loans, and other kinds of bills. This form of electronic bill payment is a substitute for checks. A share of these transactions is from checkable savings accounts, classified in M2, instead of checking accounts. The Federal Reserve Banks operate the nation’s largest ACH operation, which in 2000 processed more than 80 percent of commercial interbank ACH transactions.

In 1991, the Federal Reserve processed 1,119 million commercial (not including government) ACH transactions (valued at $5,549 billion), while in 2003 the number jumped to 5,588 million transactions ($13,952 billion), an annual increase of 14.3 percent in volume (8 percent in value).

**Retail sweep programs**

A third important development leading to the confounding roles of M1 and M2 for transactions and savings was the adoption of retail sweep programs that reclassify checking account deposits as savings deposits overnight. Since 1994, commercial banks have started using deposit-sweeping software to dynamically reclassify the balances in checking accounts above a certain level as MMDAs and to reclassify them back when the balances on the checking accounts are too low. By adopting the practice, depository institutions avoid reserve requirements on the reclassified portion of the checking account (the reserve requirement on demand deposits, ATS, NOW, and other checkable deposits can be as high as 10 percent, depending on the size of the institution). The software effectively creates a shadow MMDA for every checking account, based on the customer’s payment patterns, subject to the constraint that the number of “transfers” (reclassifications) from an MMDA to a checking account does not exceed six each month. The shadow account is included in M2, but not in M1.

More and more banks are adopting the retail sweep programs. As indicated by figure 6, the total amount of sweeps of transaction deposits into MMDAs has been rising steadily since 1994, from zero to an

![Figure 6](image-url)

**Figure 6**

*Retail sweep programs, 1994–2003*
amount nearly equal to transaction deposits in M1. According to the Federal Reserve Board’s estimates, as of December 2003, the sweeps of transaction deposits into MMDAs were approximately $575.5 billion, while total transaction deposits in published M1 were $621.3 billion. The widespread use of retail sweep programs substantially affected the growth of M1. The nominal value of M1 has been almost flat since 1994.

**MZM as a better measure of transaction balances since 1980**

As a result of the financial innovations and regulatory reforms since 1980, components of the “transactions” aggregate M1 bear interest, and components of the “savings” aggregate M2 are used for transactions. These changes call for a reconsideration of the measure of transactions demand for money and its opportunity cost. More specifically, if there is to be a stable, long-run relationship between real money, its opportunity cost, and transactions, a different measure of money and its opportunity cost may be necessary to sustain the relationship.

Motley (1988) and Poole (1991) argue that the present classification of monetary aggregates (M1, M2, M3) is inherently arbitrary, in particular in light of the banking industry developments discussed above. They believe that the important distinction should be whether the deposit has a specified term to maturity. For example, NOW accounts in M1 and MMDAs in M2 are nonterm deposits, but small and large denomination time deposits in M2 and M3 are term assets. Nonterm deposits can be readily converted into transaction balances, or in other words, are fully liquid. On the other hand, term deposits that have to be liquidated before maturity incur the cost of an early withdrawal penalty. In an environment free of government regulation, and within the limits of technology...
constraints, agents’ portfolio decisions depend on their liquidity preferences and the return on the assets. The term/nonterm distinction of monetary aggregates is aligned with private agents’ incentives.

Motley proposed classifying all non-term deposits, money that can be accessed without notice and at par, as a new monetary aggregate. Poole coined the name MZM (money zero maturity) for the measure. Specifically, MZM is defined as

\[ MZM = M2 − \text{Small denomination time deposits} + \text{Institutional MMMFs}. \]

Institutional MMMFs, currently classified in M3, are interest-bearing checkable accounts that allow holders to get around the zero-interest demand deposits restriction.

**The demand for money**

In appendix 1, we show that it is possible to derive from a simple stochastic general equilibrium monetary model the equilibrium relationship

\[ \frac{M_t}{P_t} = \alpha Y_t (i_t - i^n)_t^\gamma, \]

which is a variant of equation 1 that accounts for the fact that money may earn interest. This is an exact equilibrium relationship of observable economic variables. As pointed out in Lucas (2000), this is reason to think that the empirical analog to that relationship, which will have to account for measurement error, is a stable relationship. The equilibrium relationship in equation 2 is not exactly a money demand function, computed from the decision by households on how much money to hold, given economic variables out of their control, namely the prices of goods and assets and endowments. It does, however, look like the money demand functions that are commonly estimated.

In this section, we estimate the empirical counterpart of the money demand equation above using ordinary least squares (OLS). First, like Lucas (1988, 2000), we use M1 as the measure of money and a short-term nominal interest rate as its opportunity cost. As mentioned before, the estimated interest elasticity is 0.32, lower than the 0.5 reported by Lucas (2000) for the period 1900–94. If we estimate the elasticity for three subperiods, 1900–79, 1980–94, and 1995–2003, the interest elasticities are lower, 0.26, 0.12, –0.07, respectively. It would also be apparent that the curves would be shifted down.

Next, we estimate equation 2 using M1 as the measure of money for the period 1900–79 and MZM for the period 1980–2003. Because components of M1 bore no interest before 1980 (mostly cash and demand deposits) and components of MZM are interest-bearing after 1980 (NOWs, MMDAs, MMMFs), we assume that \( i^n = 0 \) before 1980 and we set \( i^n \) to MZM’s own rate after 1980. MZM’s own rate is a weighted average of the returns on the different components of MZM. We allow different intercepts for the two periods, because it is not reasonable to impose the coincidence of the two series, M1 and MZM, in 1980, but we do impose a common interest elasticity. The estimated money demand equation is as follows,

\[ 1900–79: \ln \left( \frac{M_t}{PY_t} \right)_t = -2.07 - 0.24 \ln i_t; \]

\[ 1980–2003: \ln \left( \frac{MZM_t}{PY_t} \right)_t = -1.8 - 0.24 \ln (i_t - i^n). \]

If we allowed for separate interest elasticity for the two periods in the regression, the two elasticities would be 0.26 and 0.2, respectively, for the first and second periods.
Figure 7 plots the logarithm of M1/nominal GDP for the period 1900–79 and that of MZM/nominal GDP for the subsequent period 1980–2003 against the logarithm of the opportunity cost of using these balances, along with the linear regression lines. The roughly common elasticity across the two periods suggests that the response of the money aggregate to changes in its opportunity cost, in percentage terms, has remained stable over the last century, as long as one uses the appropriate definition of monetary aggregate and its opportunity cost. The upward shift of the function (smaller intercept) reflects the fact that MZM and M1 include different liquid assets, even if all are zero maturity. In figure 8, we plot the actual and estimated real money demand using M1 for the period 1900–79 and MZM for the period following the deregulation and financial innovation.

**Conclusion**

While real M1 has increased very little in the last quarter century, nominal interest rates have come down considerably. If the interest elasticity were the one reported by Lucas (2000), we would expect a substantial increase in M1 that did not occur. This could indicate that the money demand relationship estimated by Meltzer (1963) and Lucas (1988), among many others, is not a stable equilibrium relationship. Instead, we argue that M1 is not the appropriate measure of money, following the regulatory reforms and innovation in electronic payments since the early 1980s. If we use an alternative, more appropriate measure of money, that is, MZM or money zero maturity, the long-run relationship between money and its opportunity cost is preserved. We estimate the interest elasticity to be 0.24, so that a 1 percent increase in the opportunity cost of holding money induces a 0.24 percent decline in real money balances.

Why do we care about estimating a stable money demand at the cost of an unstable measure of money? In addition to the theoretical interest of this issue, there is also a practical aspect to it. It is a worthy objective of a monetary authority to provide elastic liquidity at stable prices. A stable estimate of money demand, whatever the appropriate monetary aggregate might be, is an important tool in performing this task.
NOTES

1To be able to make comparisons, we use the same data as Lucas (2000) for relevant data analysis and figures. In particular, M1, real GDP, the price deflator, and the nominal interest rate are constructed, as in Lucas (2000), from different data sources for different periods. See the appendix for a detailed description of the data used in this article.

2The income and interest elasticities measure the percentage increase in real money in response, respectively, to a 1 percent increase in real GDP and a 1 percent decline in the nominal interest rate. The semi-elasticity measures the percentage increase in real money induced by a decline in the interest rate of 100 basis points.

3In a shopping time model, there is a transactions technology relating the volume of transactions to time and money used in performing those transactions.

4The relatively low income elasticity is indistinguishable from a time trend in money demand.

5The NOWs were first introduced in Massachusetts and New Hampshire in 1972, then Connecticut, Maine, Rhode Island, and Vermont in 1976, followed by New York in 1978. See Laporte (1979).

6Business customers have several ways to minimize the loss of interest on demand deposits. One way is the sweep programs developed during the 1960s and 1970s that allow business demand deposits to be swept overnight into interest-bearing accounts such as repurchase agreements and money market mutual funds.

7The term checking account is used to mean demand deposits and other checkable accounts, such as NOW accounts, classified in M1. Checkable savings accounts are accounts classified in M2 that have checking privileges.

8The Federal Reserve Board makes monthly estimates available on the nationwide change in NOW accounts attributable to the implementation of sweeps during the month. These are not the current amounts being swept, and no data are available regarding the aggregate volume of deposits currently affected by sweep programs. Depositories do not report to the Federal Reserve the size of their sweep programs.

9The MZM data and the data on the rate of return on MZM are provided by the Federal Reserve Bank of St. Louis.

10Our results are consistent with those of Carlson et al. (2000), who find a stable cointegrating relationship between real MZM, an opportunity cost measure, and a measure of economic activity, using data for the period 1976-98. The income elasticity is not different from one.

11Elastic currency is the wording used in the 1913 Federal Reserve Act that established the Federal Reserve System.
APPENDIX 1: MONETARY MODEL

Here, we consider a simple transaction technology monetary model and derive an equilibrium relationship between real money, the opportunity cost of money, and output that holds exactly. That stable relationship justifies on theoretical grounds the stability of the empirical money demand equation estimated in this article.

The economy consists of an infinitely lived representative household/firm and a government. Production uses labor according to the linear technology
\[ Y_t = A P_t, \]
where \( Y_t \) is output and \( n_t \) is time used for production. \( A_t \) is a stochastic technological parameter realized in the beginning of period \( t \). The history of these shocks up to period \( t \) (or state at \( t \)) is denoted by \( A' \). The initial realization \( A_0 \) is given.

Households have preferences over consumption \( c_t \), described by the utility function:
\[ E_\infty \sum_{t=0}^{\infty} \beta^t \left( \frac{c_t^{1-\sigma} - 1}{1-\sigma} \right), \]
where \( \beta \) is a discount factor.

Households conduct transactions according to the Cobb–Douglas transaction technology
\[ c_t = \xi(A_t s_t)^{\frac{M_t}{P_t}}, \]
where \( M_t \) is money balances, \( P_t \) is the price of the good in units of money, and \( s_t \) is the time used for transactions. The technology parameter is the same for the two technologies, production of the good, and transactions.

The total amount of time used for transactions and for the production of the good is normalized to one.
\[ s_t + n_t = 1. \]

The government issues money \( M_t^{s} \) and makes transfers to the households \( T_t \).

In the beginning of period \( t \), households enter an assets market where they purchase money balances \( M_t \) that pay net interest \( i_t^n \) in the following period, as well as nominal bonds \( B_t \) that pay interest \( i_t \) and \( Z_{t+1} \) units of state-contingent nominal securities, with price \( z_{t+1} \), normalized by the probability of occurrence of state \( A_t^{t+1} \), in units of currency at \( t \) that pay one unit of money at the beginning of period \( t+1 \) in a particular state \( A_t^{t+1} \). Subsequently, they enter a goods market where they purchase consumption with \( M_t \) according to the transaction technology in equation 4. They also receive total income \( P_t A_t(1-t) \), as well as nominal transfers, net of taxes, \( T_t \). The period by period budget constraints are
\[ 5) \quad M_t + B_t + E_t z_{t+1} Z_{t+1} \leq (1 + i_t^n) M_{t-1} - P_t A_t c_{t-1} + (1 + i_t) B_{t-1} \]
\[ + Z_t + P_t A_t (1 - s_{t-1}) + T_t. \]

A competitive equilibrium is a set of prices and quantities such that a) households choose \( \{c_t, s_t, B_t, M_t, Z_{t+1}\}_{t=0}^{\infty} \) to maximize utility in equation 3 subject to the restrictions in equations 4 and 5 together with a no-Ponzi games condition on the holdings of assets, given \( \{P_t, i_t, n_t, z_t, A_t\}_{t=0}^{\infty} \) and \( \{T_t\}_{t=0}^{\infty} \); b) the government satisfies \( M_t^s = (1 + i_t^n) M_{t-1}^s + T_t \); and c) markets clear, so that
\[ 6) \quad B_t = 0, \]
\[ 7) \quad Z_{t+1} = 0, \]
\[ 8) \quad M_t = M_t^s, \]
\[ 9) \quad c_t = A_t (1 - s_t). \]

We could derive a money demand equation using the first order conditions of the households’ problem. That equation, however, would be a function of all the prices, including the prices on the state contingent nominal debt, as well as unobservable shocks, and therefore, could not be directly estimated using simple econometric methods. Instead, the first order conditions can be used to derive the following relationship
\[ 10) \quad \frac{m_t}{c_t} = \alpha \left( i_t - i_t^n \right)^{-\gamma} t \geq 0, \]
where \( m_t \) denotes real money balances, \( m_t = \frac{M_t}{P_t} \), and
\[ \alpha = \left( \frac{1 - \gamma}{\gamma} \right)^{\frac{1}{\gamma - 1}}. \]
As pointed out by Lucas (1988), this equation is not exactly a money demand, rather it is an equilibrium relationship between real money, consumption, and the opportunity cost of holding money that holds exactly in this stochastic environment. Given
APPENDIX 1: MONETARY MODEL (CONTINUED)

that in this simple model consumption coincides with output, \( c_t = Y_t \), equation 10 can be rewritten as

\[
\frac{M_t}{P_t} = \alpha Y_t \left( i_t - i^* \right)^{-\nu},
\]

with interest elasticity equal to the Cobb–Douglas transactions technology parameter \( \nu \).¹ Note that the derived income elasticity is one.

The assumptions on the homogeneity of the transaction technology and technology progress in the two sectors, as well as assumptions on the utility function, imply that the long-run income elasticity is one. Alternative assumptions could imply a trend in money demand. Empirically, this could be captured by a time trend or by an income elasticity different from one, as in Ball (2001). Instead, we argue that the evidence is consistent with a stable long-run money demand with a unitary income elasticity and no time trend, if the monetary aggregate is appropriately defined to capture the technological and regulatory innovations since 1980.

¹Lucas (2000) reports the interest elasticity to be \( \nu = 0.5 \). He justifies this result by arguing that equation 10 with \( \nu = 0.5 \) is an approximation to the equilibrium relationship when the transaction technology is Baumol–Tobin. In fact, if the transaction technology was Baumol–Tobin, \( s_t = \eta \left( \frac{c_t}{m_t} \right) \), the money demand equation 10 would be, \( \frac{m_t}{c_t} = \omega \left( \frac{A_t}{c_t} \right)^s (i_t - i^*)^{-s} \),

where \( \omega = \eta^s \). The approximation amounts to ignoring the term \( \left( \frac{A_t}{c_t} \right)^s \).
APPENDIX 2: DATA USED IN FIGURES AND REGRESSIONS

The following is a list of data used in the figures and regressions for this article. Unless explicitly specified, all monetary aggregates are in billion of dollars and are not seasonally adjusted annual data (we take the December value of each year’s value).^1

**M1**

**M2**

**M3**

**MZM**

**Other checkable deposits** (quarterly frequency)
FRB data available through Haver Analytics (FMOTN in USECON).

**MMMFS** (quarterly frequency)
FRB data available through Haver Analytics (FMGMN in USECON).

**Institutional money market mutual funds** (quarterly frequency)
FRB data available through Haver Analytics (FMIMON in USECON).

**Transaction deposits swept into MMDAs** (Cumulative)
FRB data available through Haver Analytics (FMSWEEP in USECON).

**Demand deposits**
FRB data available through Haver Analytics (FMIDN in USECON).

**Price deflator**
1929–2003 (2000 = 100): BEA data available through Haver Analytics (DAGDP in USECON or USNA).^2

**Real GDP**
1900–28 (millions of 1929 dollars), Kendrick (1961, Table A-III).
1929–2003 (in chained 2000 dollars), BEA data available through Haver Analytics (GDPHA in USECON or USNA).^3

**Nominal interest rate**
1900–69: Friedman and Schwartz (1982, table 4.8, column 6), defined as short-term commercial paper rate.
1970–2003: three-month commercial paper, FRB data available through Haver Analytics, FFP3 in USECON.

**Opportunity cost**
1900–79: M1’s opportunity cost is defined as the nominal interest rate for this period.
Three-month T-bill rate: FRB data available through Haver Analytics (FTBS3 in USECON).

---

^1We follow Lucas (2000).
^2These two series overlap in 1929 and using the ratio of the two series’ values in 1929, we construct a new price deflator that goes from 1900 to 2003, with 2000 = 1.0.
^3From these two series, we construct a new real GDP in 2000 dollars using the new price deflator (2000 = 1.0).
REFERENCES


Poole, William, 1991, testimony before the U.S. Congress, Committee on Banking, Finance, and Urban Affairs, subcommittee on Domestic Monetary Policy.
