The Frequency and Costs of Individual Price Adjustment

Alexander L. Wolman

The concept of "sticky prices" has been one of the most common explanations for the perceived importance of monetary policy since at least as far back as the 1930s. Simply put, if nominal prices for individual goods do not continuously adjust to economic conditions, then it is natural to think that monetary policy can influence the level of real economic activity through its ability to determine the nominal quantity of money. In evaluating whether this channel for monetary policy is important, two sets of research questions are relevant. First, do individual prices indeed change infrequently, and if so, why? Second, within macroeconomic models, what are the aggregate implications of the pricing behavior found in the data, and are those implications consistent with aggregate economic data? This article reviews research on the first set of questions in the hope of deriving lessons useful for improving the macroeconomic models that can address the second set.[1]

Weiss (1993) and Wynne (1995) have written surveys on similar topics. Weiss promotes the importance of infrequent price adjustment, whereas Wynne is a skeptic. This article differs from their work in that it covers the many papers that have appeared since 1995 and provides a history of thought perspective on theories of infrequent price adjustment. Much of my previous research has involved sticky price models, so I have a stake in what the evidence reveals.

I received helpful comments from Mike Dotsey, Martin Gervais, Andreas Hornstein, and Thomas Humphrey. The ideas expressed in this paper are the author's and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] I do not consider infrequent nominal wage adjustment. Loosely, the parts of my paper that refer to repeated relationships can be thought of as applying to wages as well as prices.

Table 1 Summary of Empirical Studies

Object of Study	Authors
Retailers	Lach and Tsiddon; Rátfai; Slade; Levy et al.; Dutta et al.
Periodicals	Mussa; Weiss; Cecchetti; Willis
Industrial Products	Carlton; Zbaracki et al.
BLS Averages for Wholesale Prices	Mills; Means
Goods in Catalogs	Kashyap
Large Number of Firms (through surveys)	Blinder et al.; Buckle and Carlson; Carlson; Hall, Walsh, and Yates

Research on price stickiness has involved continual interplay between theory and empiricism. The early empirical studies discussed below approach "pure" empirical exercises. Research was not conducted in a vacuum, but these studies seem to have only broad theoretical motivation, and—initially—the results were not used to support particular theories. Subsequent theories of pricing behavior were developed and refined; the theory of explicit "menu" costs of price adjustment has been most sharply refined. Most of the empirical work I survey was conducted with this theory as its organizing framework. However, I also describe recent empirical work that takes a more naive approach, studying the pricing process at a large industrial firm. Together with two surveys of firms' pricing behavior, this recent work relates to several less-refined theories of infrequent price adjustment. Table 1 lists the empirical studies I survey.

Prices do change infrequently for many retail transactions. Furthermore, price adjustment behavior appears to be consistent with explicit, direct costs of changing prices. Evidence of infrequent price changes also exists for nonretail transactions, but the costs associated with price adjustment are not as easy to pin down. New evidence, supplementing years of conjecture, suggests that these costs involve the repeated nature of many buyer-seller relationships. The main challenges ahead are to improve both measurement, so that we better understand the nature of buyer-seller relationships, and theory, to study the macroeconomic implications of such relationships. It is not clear that conclusions about monetary policy based on direct costs of price change will carry over to models where infrequent price change results partly from repeated relationships.

1. MILLS'S DATA AND "ADMINISTERED PRICES"

Frederick Mills (1927) published what may be the first study of the frequency of price changes. He documented the behavior of wholesale price quotations for more than two hundred goods, using data from the Bureau of Labor Statistics' wholesale price bulletins covering the period from 1890 to 1926. The broad range of goods in Mills's book makes it a valuable source: he covers everything from cotton yarn ("Carded, white, mulespun, northern, cones, 10/1") to doorknobs ("Steel, bronze-plated"). A drawback to Mills's data, however, is that it is stated at a monthly frequency, with the monthly observations either taken one day per month or as an average of daily or weekly observations. Some of the measurements of price-change frequency are thus inaccurate, although prices that change less than once per month must be fairly accurately represented. Furthermore, the data on many of the goods were *averages* of observations from multiple price reporters—BLS field workers who recorded prices.

Mills's basic finding is that the frequency of price changes varies widely across goods: Excluding the years 1914–1921, roughly one-fifth of the goods changed price in less than 10 percent of months, while another one-fifth changed price in more than 90 percent of months. Mills pointed out that the distribution of price-change frequencies was bimodal (U-shaped), but did not speculate on the causes or implications of his findings. In reviewing Mills's book, John Maynard Keynes (1983 [1928], p. 226) wrote, "[i]t is the peculiarity of Mr. Mills that he starts without any theories and ends without any, being content to set out his material for the benefit of those who have less taste than he has for laborious investigation, and more taste for theorising." Keynes surely was one who had a taste for "theorising," but his review simply summarizes parts of the book, emphasizing dispersion in relative prices more than the frequency of price changes.

It was up to Gardiner Means, in 1935 an adviser to the U.S. Secretary of Agriculture, to impose an interpretation on Mills's material. Means's data is similar to that studied by Mills except that it is for the years 1926–1933—a period covering the early part of the Great Depression. Not surprisingly, that data has the same bimodal distribution for the frequency of price change noted by Mills. Means provided an interpretation that relied heavily on this feature of the data: "There are two essentially different types of market in operation—the traditional market in which supply and demand are equated by a flexible price and the administered market in which production and demand are equated at an inflexible administered price" (Means 1935). Means went on to argue that inflexible, administered prices had grown in importance as the economy had become more industrialized and were largely responsible for the severity of the Great Depression.

The "administered price" thesis spurred a voluminous literature, and Means became known as "one of the most influential economists in the history of this country" (Stigler and Kindahl 1973, p. 717). Although Means's theory evolved over the years, it never clearly defined administered prices. Stigler and Kindahl wrote that "Means' theory has indeed become difficult to refute or confirm. . . . The implications have become so broad as to be almost uselessly vague" (1973, p. 719). Even the fundamental observations that drove Means's theorizing were shown to be suspect or false. First, Tucker (1938) used other data sources to argue that prices changed less frequently in the nineteenth century than during the 1929–1933 period. Second, Scitovszky (1941) pointed out that the U-shaped distribution of price-change frequencies was an artifact of the data format. Both Mills's and Means's data were monthly, so any prices that changed more than once per month were recorded as changing monthly; had they been recorded at, say, daily frequency, one end of the distribution would have flattened out. A similar argument can be made for the other end of the distribution, although Scitovszky did not discuss it. Any price that did not change during the sample period is treated identically, whereas if the sample were long enough, many of these items would be seen to have different frequencies of price change.

Although the bimodal distribution was spurious, Means was correct in pointing out that many prices change infrequently (he found this to be the case even for many goods that had more than one reporter). Furthermore, while Means did not clearly define "administered prices," it may be a useful term. The prices of many goods are changed only as a result of conscious decision by an individual or group of individuals, that is, by administration. And one would expect that "administrative costs" might lead to infrequent price change. In the next section I summarize a mature literature that makes explicit one narrow form of administrative costs. Subsequently, I will discuss a nascent literature that can be interpreted as resurrecting a richer notion of administered prices and studying the process of administration.

2. EXPLICIT COSTS OF PRICE ADJUSTMENT

The best-developed theory of infrequent price changes is also the simplest. This theory combines some monopoly power on the part of firms with explicit physical costs of changing nominal prices. Because these costs are usually assumed to be fixed (independent of the magnitude of the price change), they are often referred to as "menu costs." Firms with monopoly power have leeway to choose the price of their products. The physical costs of changing prices will make firms choose to change prices infrequently as long as the costs have a significant component that is fixed, or sufficiently concave in the size of the price change.[2] The idea behind this theory is straightforward, so

[2] If there are costs of changing prices, but those costs are convex in the size of the price change, then firms will choose to change their prices a little bit every chance they get rather than waiting to adjust. See Rotemberg (1983).

there is a large body of work analyzing its implications. The theoretical work has led to empirical work studying both the length of time prices are fixed and the costs of changing prices. As a form of administrative costs, menu costs are distinguished in that they are mechanical: firms can be thought of as contracting out the printing of menus, knowing exactly what the cost will be.

The Theory

Robert Barro (1972) was the first to study a formal theoretical model of a firm facing explicit fixed costs of changing its price. According to Barro, "Shifts in price involve direct administrative costs to the producer (seller). . . . The administrative costs associated with price changes are straightforward, and can reasonably be described as a lump sum amount, independent of the size or direction of adjustment" (1972, p. 21). Sheshinski and Weiss (1977, 1983, 1992) made important contributions to understanding the behavior of firms facing fixed costs of price adjustment, as did Danziger (1987, 1988) and Kuran (1983). Caplin and Spulber (1987) initiated the formal analysis of the effects of monetary policy in models with fixed costs of price adjustment, and surprisingly they found that although costs of price adjustment made firms change their prices infrequently, the price level moved one-for-one with the money supply, implying monetary neutrality. Subsequent developments by Caplin and Leahy (1991, 1997), Caballero and Engel (1993), Dotsey, King, and Wolman (1999), and Danziger (1999) suggest that Caplin and Spulber's finding is not robust, i.e., that monetary policy can and generally does have real effects in menu-cost models.[3] While these articles have extended Barro's analysis in important ways, they have all maintained the basic assumption that there is a direct cost to firms of changing the nominal price of products they sell.

Although no one until Barro incorporated the idea into a formal model, economists discussed explicit costs of price adjustment as early as the 1930s. Note that Barro uses the term "administrative costs." Explicit costs of price adjustment do not appear in Means's written work, but in a 1936 article John Kenneth Galbraith referred to private communication with Means that indicates Means—and Galbraith—thought about menu-type costs:

> Professor Gardiner C. Means has drawn my attention to the cost of making a price change under modern conditions as an incentive to the holding of prices constant. A concern with nation-wide sales outlets must make certain that dealers are informed of the change; it must distribute new price schedules and provide safeguards against "leaks" as well as risk a temporary cessation of business in case there is such a "leak." It must

[3] This is not an exhaustive list of references.

> also recast its advertising to acquaint the public with the change. All of these things cost money and all of this expenditure is avoided if prices are allowed to stay where they are. (p. 470)

Galbraith describes a rich array of costs in this paragraph. The cost to "distribute new price schedules" corresponds most closely to the menu-type cost incorporated in most theoretical work since Barro; other costs Galbraith describes have a different flavor, and I return to them below. Another early reference to costs of price adjustment can be found in Scitovszky (1941). He writes that "the adjustment of price policies and of production often involves high subjective costs—not only for administrative and technical but also for political reasons." He then cites a 1935 survey article on monopoly theory by John Hicks as a source for the idea of administrative and technical costs. The relevant passage from Hicks reads, in part, "[T]he variation in monopoly profit for some way on either side of the highest profit output may often be small . . . and if this is so, the subjective costs involved in securing a close adaptation to the most profitable output may well outweigh the meagre gains offered" (1935, p. 8). Hicks's argument is framed in terms of output, but it is clear that a similar argument could be made based on price.

Hicks anticipates arguments made by Akerlof and Yellen (1985a, b) and Mankiw (1985). Like Hicks, these authors note that the nature of optimal behavior often implies that small deviations from optimality will have a small effect on a monopolist's profits. A small cost of adjusting price may then make it optimal for a firm to keep its price fixed for several periods. This is precisely the idea that has been formalized in the literature initiated by Barro (1972), and Hicks came close to elucidating it in 1935. The unique contribution of Akerlof and Yellen and Mankiw was to show that under certain conditions, the implications of small costs to individual firms could be large for the economy as a whole.

Models in which there are explicit costs of price adjustment have three main empirical implications. First, trivially, there are explicit costs to changing individual prices, and because of these costs prices will not be changed continuously. Second, individual prices will change more frequently when the benefits to adjusting price are high or the costs are low. And third, the frequency of price adjustment should be high (low) when the inflation rate is high (low) because inflation causes the benefit to changing prices to rise over time. Many empirical studies have been motivated by menu-cost-type models of costly price adjustment, and they focus primarily on these main implications.

Empirical Evidence from High Inflation Episodes

Several authors have studied individual price data generated under conditions where there is great variation in the inflation rate. Such data are especially interesting for evaluating menu-cost models because of the predicted positive relationship between inflation and the frequency of price change.

Mussa (1981) and Weiss (1993) describe newspaper prices during 1920s German hyperinflation. Between 1920 and 1923, as the German inflation rate rose to more than 10,000 percent per month, the cover prices of three newspapers changed with increasing frequency; however, even in the months of highest inflation those prices appear not to have changed every day.[4] Weiss also describes the behavior of an Israeli newspaper's price during a period of rising inflation from 1972 to 1984, and one sees the same qualitative features as in the German data. A drawback to looking at data on daily newspaper prices is that single copy sales of newspapers are not the only source of revenue associated with a newspaper. It is possible that subscription prices were indexed to inflation, and advertising may have been an important source of revenue. Presumably, though, some copies of the newspaper were indeed purchased at the printed price, so the behavior of that price does provide support for the existence of costs of price adjustment.

In a series of papers beginning in 1992, Saul Lach and Daniel Tsiddon study the behavior of the prices of 26 foodstuffs at grocery stores in Israel, during three subperiods between 1978 and 1984. As with newspapers in Israel during roughly the same period, food prices were changed infrequently, and the frequency responded to the overall inflation rate. As inflation rose from a 4.9 percent monthly rate in 1978–1979 to a 6.6 percent monthly rate during the 1981–1982 subperiod, the average duration of a price fell from 2.2 months to 1.5 months. These basic facts on the duration of price quotations are presented in Lach and Tsiddon (1992); in that article and their 1996 papers the authors also provide some deeper insights into the implications of their data. Among the features of the data they discuss is the distribution of relative prices. Caplin and Spulber's early menu-cost model generates a uniform distribution of prices, whereas Lach and Tsiddon find a unimodal distribution. This finding does not mean that Lach and Tsiddon's data are inconsistent with menu-cost models; more general models, such as that in Dotsey, King, and Wolman (1999), generate nonuniform distributions. However, even Dotsey, King, and Wolman's model would need to be extended—for example, with additional firm-specific shocks—to generate the type of distributions found by Lach and Tsiddon.[5] For some goods, they find that there are small price changes, but

[4] Weiss's data are in the form of a time series plot, for the newspapers *Germania* and *Tageblatt.* These data end in July 1923. Mussa displays the dates of price changes for the *Neue Preussische Zeitung*, for the period ending November 17, 1923, close to the date of monetary reform.

[5] The difficulty is that Lach and Tsiddon find nearly symmetric distributions, whereas Dotsey, King, and Wolman's model predicts that in an inflationary world the mode should be at the highest price charged. Eden (forthcoming) also analyzes Lach and Tsiddon's data. He argues that the relationship between relative price variability and inflation in the data is inconsistent with Dotsey, King, and Wolman's model.

the average price change for a given store is not small (1996b). Evidence of many small price changes would typically be seen as conflicting with menu-cost models, but Lach and Tsiddon argue that this evidence is consistent with models where the cost incurred is for changing the entire *menu* of a store's prices. And indeed, they find that price changes are generally synchronized within a firm, but not for a given good across firms (1996a).

Rátfai (2000) studies the behavior of meat prices in Hungary. His data cover 14 specific products sold in 8 stores from 1993 through 1996. Observations are at the monthly frequency. During this period annual CPI inflation in Hungary varied between 16 percent and 30 percent. The average duration of the meat price quotations was three months. Another measure of price change frequency is that in 58 percent of the observations, price was unchanged from the previous month. Rátfai provides data on both the overall inflation rate and different measures of the frequency of price adjustment over time. The data show a positive relationship, as predicted by theory. Rátfai also provides a detailed description of the distribution of price changes by size. His data appear consistent with fixed costs of price adjustment in that one does not observe many tiny adjustments.

Empirical Evidence from Moderate Inflation Eras

Menu-cost models make sharp predictions about periods in which the average inflation rate changes dramatically. Periods when inflation is moderate and relatively steady lend themselves to different analysis. One can still measure the frequency of price adjustment and the distribution of price changes. Menu-cost models predict that price changes will be infrequent and large rather than frequent and small, but this prediction needs to be qualified in light of Lach and Tsiddon's analysis of multiproduct firms. In effect, their work teaches us that it will be difficult to reject menu-cost models; evidence of small price changes does not necessarily mean that menu costs are unimportant. In addition to documenting the behavior of individual prices, researchers have also attempted to measure menu costs, both directly and indirectly.

Of all the studies of price adjustment frequency from data generated under moderate inflation, Anil Kashyap's (1995) analysis of prices in catalogs covers the longest period of time, from 1953 to 1987. Kashyap traces the prices of 12 items in total from L.L.Bean, Orvis and REI catalogs. Some of the items were sold during the entire 35-year period, while others were sold for as little as 12 years of the sample. All of these merchants fix their catalog prices for a minimum of six months at a time, but prices of the 12 items in Kashyap's study changed even less frequently. The Orvis binoculars changed price most frequently—every 11.2 months on average—and the Orvis fishing fly changed price least frequently—every 30.4 months. Even the items that changed prices relatively frequently on average had lengthy episodes without a price change.

The aforementioned binoculars went for 42 months from 1968 to 1971 without a price change, and the L.L.Bean camp moccasin—which had a price change every 11.5 months on average—went for 78 months without a price change from 1959 to 1965. Kashyap also finds that there are many examples of small price changes. Given Lach and Tsiddon's findings regarding synchronization across and within firms, it is interesting to know what Kashyap's data reveal along these lines. Overall, he finds little evidence of synchronization in price changes; however, there are too few goods per catalog in his study to reach robust conclusions about synchronization within a catalog. Kashyap's findings are somewhat troubling with respect to standard menu-cost models. Each of the firms chooses to incur a menu cost (print a new catalog) twice per year. Standard models would predict the cost being incurred only when optimal. It seems clear, though, that the complicating factor for catalog retailers is that the menu cost is incurred in conjunction with marketing activities, and marketing, as opposed to price changes, may be the dominant factor in determining the semiannual frequency.

John A. Carlson (1992) uses data from a quarterly survey conducted by the National Federation of Independent Business, covering the years 1979–1990. He presents both prospective and retrospective information on firms' price-changing behavior. The prospective information is in the form of responses to the question, "In the next three months, do you plan to change the average selling price of your goods and/or services?"[6] Retrospectively, firms are asked, "How are your average selling prices now compared to three months ago?" In the first two years of the sample, when inflation was relatively high, between 30 percent and 45 percent of firms reported that their prices had remained constant and had been expected to remain constant over the previous three months. As predicted by the existence of menu costs, in the lower inflation period from 1983 to 1990, these numbers stabilized at a higher level: between 55 percent and 75 percent of firms reported constant actual and expected prices over the prior three months.

Robert A. Buckle and Carlson (2000) report on a similar survey conducted by the New Zealand Institute of Economic Research. The *Quarterly Survey of Business Opinion* asks executives of manufacturing and building firms whether the direction of price change in the last three months has been up, same, or down. For all firms in the survey, the average duration of a price is reported to be 6.7 months. The authors also split up their sample by firm size. They find that the smallest firms in the sample keep their prices fixed on average for 50 percent longer than the largest firms. Those estimates are probably imprecise. Survey respondents are generally considering the price of more than one product, so their responses may reflect these products' average behavior. The

[6] See Carlson (1992, p. 325).

category "unchanged" may be chosen for average changes that are considered small. Nonetheless, it is notable that small firms seem to keep their prices fixed longer. Buckle and Carlson argue that this suggests that costs of price adjustment do not vary much across firms: large firms reap a greater benefit from adjusting their price and hence adjust more often.

Stephen Cecchetti (1985, 1986) documents the behavior of 38 U.S. magazine cover prices between 1953 and 1979. The prices are measured annually, with the first issue of each year. From year to year, the number of magazines that changed their price ranges from 3 percent of the total, in 1953, to 50 percent of the total, in 1974. In the 1950s (1953–1959), 12 percent of the magazines changed their price in an average year, while in the 1970s—a period of higher inflation—30 percent of the magazines changed their price in an average year. Cecchetti's evidence is thus both consistent with the evidence on newspaper prices in Germany and Israel and, because magazines sell subscriptions and advertisements, subject to the same caveat. Cecchetti goes beyond simple descriptive statistics to estimate a complicated price-setting model. While his estimates are consistent with the existence of menu costs, he concludes that menu costs are not the whole story. The total costs of price adjustment depend on the size of the price change. This finding is consistent with the idea of "customer costs" of price adjustment, which are discussed below.

Willis (2000) uses Cecchetti's magazine price data to estimate the menu costs of price adjustment. His estimation method begins by specifying an explicit intertemporal optimization problem for each magazine. Next, reduced-form hazard functions are estimated, which represent the probability of a price change as a function of a small number of aggregate and magazine-specific variables. Then, an indirect inference procedure is used to match up the reduced-form estimates with structural parameters. Willis finds that the average adjustment cost is 4 percent of revenues. I have already discussed the fact that magazines generate revenue from subscriptions and advertising, as well as single-issue sales. It is difficult to interpret Willis's results because ad marketing and price discrimination between subscribers and newsstand purchasers may influence the cover-pricing decision. Nonetheless, Willis's method is a promising one.

Slade (1998) combines a dynamic model of optimal price setting with sophisticated econometric methods to estimate the costs of price adjustment for particular goods. She uses weekly data for 1984–1985 on saltine prices and sales at ten supermarkets. Price setters face profit maximization problems similar to those studied in the literature following Barro (1972) except that there is a role for "goodwill" in their decisions. Goodwill enters the demand function faced by firms, and it rises and falls according to whether price is above or below a "normal" level. Slade's estimation framework allows for both fixed and variable components of price adjustment costs, and the costs being estimated represent all costs of price adjustment. That is, the theoretical framework maintains that there are fixed and variable costs of price adjustment, and it does not break down those costs according to whether they are menu costs or anything else. Slade reports that price remains unchanged about 80 percent of the time. She estimates that the average cost of a price change for a box of saltines is $2.72, with 94 percent of those costs fixed with respect to the size of the price change. When broken down by store, the average fixed costs range from $2.17 to $3.09, and when broken down by brand, they range from $2.15 to $2.68. A national chain has the lowest costs, an independent store has the highest costs, and two regional chains are in the middle.[7]

Two recent papers on individual price adjustment are concerned mainly with measuring the costs of price adjustment, although both provide some information on the frequency of price adjustment for individual goods. Levy et al. (1997) and Dutta et al. (1999) study price adjustment at supermarkets and drugstores, respectively. In both cases, the data source is a company that sells electronic shelf-label systems. As part of its sales effort, the company has undertaken detailed studies of the pricing practices of some of its potential customers. Included in the studies is information on both the frequency and the costs of price adjustment.

Levy et al. find that four supermarket chains change the prices of approximately 16 percent of their products each week, while a fifth chain only changes the prices of 6 percent of its products each week. The chain with fewer price changes also happens to be located in a state that requires each item to carry its own price tag—which raises the cost of price adjustment. Because the sales pitch of the electronic shelf-label system's manufacturer promises in part to save retailers money on their pricing process, we might expect the manufacturer's estimates of the direct costs of price adjustment to be too high. In each case, however, the estimates were reviewed by the retailers themselves, who apparently did not strongly disagree with the manufacturer.[8] On the other hand, we should not expect good estimates of any costs of price adjustment other than menu costs from these studies: the manufacturer of electronic labeling systems has no incentive to measure additional costs. To the extent that other costs of price adjustment are high, the benefits of reducing the direct costs of price adjustment will be smaller. The retailer would save on the existing amount of price adjustment, but might not choose to significantly increase the frequency of its price adjustment.

Levy et al. report that the physical costs of price adjustment average 0.70 percent of store revenues. Profit margins are small in the supermarket

[7] In Slade's paper, it is the manufacturers that are setting price. The statement that "a national chain has the lowest costs," for example, means that the manufacturer's cost of changing price at that chain are lower than at other chains.

[8] The electronic shelf-label system illustrates that technology affects the costs of price adjustment.

industry—the authors assume an average of 2 percent—meaning that menu costs are estimated to be 35 percent of profits. They also estimate that changing the price of a specific item costs 52¢ on average for four of the chains in their sample and $1.33 for the fifth chain. As mentioned above, the fifth chain is subject to laws requiring that all items carry price tags, and it changes prices much less frequently than the others. More than 70 percent of the costs of changing price are labor costs. The source data for this study do not include information about the head office managerial costs of price adjustment. However, based on their own knowledge of this industry, the authors estimate these managerial costs at 6.8 percent of the total costs of price adjustment. It is notable that estimated cost per price change is much lower for Levy et al. than for Slade. There are at least three potential explanations for this discrepancy. First, saltines may be items with particularly high menu costs; Levy et al. do not provide data on individual goods. Second, Levy et al. may not be measuring certain forms of price adjustment costs. Finally, Slade's model may be misspecified, resulting in cost estimates that are too high.

With respect to the drugstore chain, Dutta et al. find that prices change on 7.5 percent of the products in an average week. They report similar estimates of the costs of changing prices for drugstores to those found for supermarkets, with the direct costs of price adjustment estimated to be 0.74 percent of revenue and 27 percent of profits. Changing the price of a specific item is estimated to cost 42¢ on average. Based on these numbers, it is not clear whether one would describe menu costs as higher for supermarkets or drugstores. It is worth noting, then, that the frequency of price change is much lower in drugstores than in supermarkets. It must be that either unmeasured costs of price adjustment are higher for drugstores or the standard notion of profits is less sensitive to price for drugstores than for supermarkets.

A Skeptical View

The papers discussed in this section generally support the idea that prices are changed infrequently because of fixed costs of price adjustment. Like many of these studies, Carlton (1986) presents evidence that individual prices change infrequently. However, his preferred interpretation is not that there are physical costs to price adjustment.

Carlton's data set was compiled by George Stigler and James Kindahl. They used the data for a 1970 book that was in large part a critical response to Gardiner Means. The principal drawback to the BLS data analyzed by Mills and Means is that much of it is averaged over multiple sellers. For their book, Stigler and Kindahl gathered data from individual purchasers in industries similar to those studied by Means. They chose, however, to concentrate on the "goods-level" data rather than the transactions-level data. Carlton studied this same data at the level of the individual purchaser. The goods are intermediates used in manufacturing. There are 11 different product groups, including steel, truck motors, glass, and cement. The average duration of price rigidity across the product groups ranges from 5.9 months (household appliances) to 19.2 months (chemicals). These data thus show strong evidence of infrequent price changes.

Long average intervals between price changes do not necessarily imply that fixed costs of price adjustment are important. Carlton shows that in the Stigler-Kindahl data there are many instances of small price changes, which is inconsistent with simple menu-cost models. Richer menu-cost models allow fixed costs of changing price to vary across firms, and Carlton acknowledges that this may be one factor in explaining the data. However, he also emphasizes a fundamentally different explanation, that "firms and buyers differ in their need to rely on the price system to achieve allocative efficiency" (1986, pp. 648–49). Carlton spells out this idea further in a 1989 essay, arguing that price adjustment alone is a good way to clear markets for some homogeneous goods. For other goods, especially in the presence of long-term relationships, efficient allocations may be achieved at lower cost by varying other characteristics, such as quality or delivery time. From this perspective, there is nothing special about adjusting price versus adjusting other characteristics. The observation that prices do not change frequently in a particular market tells us something about the nature of that market, but it does not necessarily imply that there are high physical costs of price adjustment.[9] Carlton's analysis is important to bear in mind when confronted with data about the frequency of price adjustment alone; however, his analysis must be weighed against those that actually measure price adjustment costs.

3. ADMINISTERED PRICES REHABILITATED

Menu costs are one form of administrative price adjustment cost. Economists have been drawn to studying menu costs because it has been relatively straightforward both to measure them (in principle) and to incorporate them in rich dynamic models. However, the idea of menu costs overlooks the administrative *process,* which may be costly. It also ignores the possibility that firms may face *indirect* costs of changing their prices, related to the effect of a price change on consumer behavior.[10] A recent empirical paper by Zbaracki et al. (2000) studies costs of the administrative process and indirect

[9] The implications for monetary policy of Carlton's explanation for rigid prices are unclear. In the context of labor markets, Barro (1977) has argued that a similar explanation for infrequent (wage) adjustment overturns the idea that monetary policy has real effects. Without understanding the precise features of the market that lead nominal prices to go unchanged, however, this conclusion seems tenuous.

[10] These changes in consumer behavior represent something other than movement along a smooth demand curve, as will be seen below.

costs associated with changing customer behavior, as well as menu costs. As I describe this empirical work, I will give some more detailed background on the theoretical basis for indirect costs of price adjustment.

Zbaracki et al. report on an intensive analysis of a large industrial corporation that sells more than a thousand products. The authors studied this corporation for two years with the objective of learning about the firm's price-setting process. Their study follows a "'discovery oriented' perspective" and is geared toward understanding what the authors refer to as managerial costs and customer costs of price adjustment, as well as physical costs (2000, p. 4). In their terminology, physical costs correspond roughly to menu costs—they are costs of printing and distributing price lists, both hard copies and electronic versions. Managerial costs are those costs incurred during the long process—four months for this corporation—of choosing list prices for the upcoming year. Customer costs are divided into "hard" and "soft" costs; hard customer costs are similar to managerial costs, except that the former are incurred after list prices have been set. These costs involve communicating with customers regarding new prices. Soft customer costs are described as "potential harm to customer relationships and company reputation" (2000, p. 19).

The idea of soft customer costs recalls one of the costs listed by Galbraith, that of the risk of lost business. These costs are indirect, and the "risk of lost business" is the risk that even a small price change will cause a discrete shift in demand. Absent physical costs of changing price, such factors would still make it optimal for firms to keep their prices fixed in the face of small enough changes in demand and supply conditions. Although the idea of indirect costs was perhaps implicit in Means's work and the literature he inspired, clear statements can first be found in Sweezy (1939) and Hall and Hitch (1939). These authors developed the "kinked demand curve" theory of oligopoly. This theory posits that a decrease in price will not gain a firm very many customers, but an increase in price will cause it to lose many customers. Faced with this situation, a firm will choose to maintain a fixed price in response to modest changes in market conditions. Stigler (1947) cast doubt on the simple kinked demand theory but could not kill it. A large recent literature studies conditions under which models with consumer search imply that firms' demand curves will be endogenously kinked at some initial price.[11] Informal discussions of this idea can be found in Okun (1981); see Stiglitz (1989) for a discussion of the more technical literature. Because models of consumer search are relatively complicated, theorists have made less progress in studying the macroeconomic implications of "customer costs" than direct costs; see, however, Woglom (1982), Goodfriend (1997) and Benabou (1988, 1992). As for empirical work

[11] A major stumbling block in this literature has been the issue of how the firm's "initial price" is determined.

studying indirect costs of price adjustment, Zbaracki et al. is one of the few examples.

The firm studied by Zbaracki et al. has an annual price-setting cycle. Each month it sends out an updated pricing sheet, but very rarely does that sheet list changed prices. Although the goods at the industrial corporation of Zbaracki et al. are very different from the catalog items studied by Kashyap, in both cases the firms often choose not to change prices even when new "price lists" are distributed. The authors estimate the physical costs of price adjustment to be 0.04 percent of revenue and 0.7 percent of net margin. Managerial costs of price adjustment are 0.28 percent of revenue and 4.6 percent of net margin. Finally, hard customer costs of price adjustment are 9.1 percent of revenue and 15 percent of net margin.

Zbaracki et al. do not attempt to quantify soft customer costs, but they do provide a qualitative discussion of these costs, which is consistent with modern versions of kinked demand theory. According to a salesperson at the firm, "every time you have one of these price changes you have to go in there and you are opening a Pandora's box" (2000, p. 26). In other words, enacting a price change can have severe consequences for customer relationships.

Note that in Zbaracki et al., physical costs are only one-seventh as large as managerial costs, which sharply contrasts with the estimates for supermarkets, where Levy et al. report that physical costs are roughly nine times as large as managerial costs. On one level, it is perhaps not surprising that managerial costs are relatively higher for the industrial corporation. Supermarket chains have one headquarters supporting hundreds of retail outlets, and the physical costs of price changes must be incurred at every retail outlet. At the industrial corporation, there are no retail outlets, so physical costs are small. This explains why the ratio of managerial costs to physical costs should be larger for the industrial corporation than for the supermarkets. However, it does not explain why managerial costs should be so much larger *as a percentage of revenues* at an industrial corporation. This differential suggests that pricing is more difficult at the industrial company than at supermarkets. To the extent that the supermarket has numerous competitors with readily observable prices, perhaps it is not surprising that pricing is more difficult for the industrial corporation.

The most striking aspect of these numbers is the magnitude of the customer costs of price adjustment. Hard customer costs of price adjustment are 28 times as large as physical costs and managerial costs combined. Recall that there was no mention of customer costs in the studies of supermarkets and drugstores. In the latter industries, there are many customers, each of whom is small. Hence, essentially no resources are expended communicating with customers on an individual basis.[12] In contrast, the corporation studied by Zbaracki et al. has only about 1,400 customers. Almost three hundred of the largest customers receive some individualized attention as the new price lists are introduced, and any of the customers will receive some attention from a salesperson if they complain about a price change.[13]

4. SURVEY EVIDENCE

Blinder et al. (1998) and Hall, Walsh, and Yates (1997) have conducted large surveys of firms' pricing practices. These surveys sought quantitative information about the frequency of price adjustment, and qualitative information about the costs of price adjustment. In addition, the authors included questions in their surveys about whether different theories of price adjustment were deemed relevant.

Blinder et al.'s book describes the results of interviews the authors conducted with roughly two hundred companies.[14] The interviews were survey-like in that they consisted of a common set of questions for all firms. Many of the questions related to factors that influence firms' pricing behavior. Blinder et al. also asked the firms how often they changed their prices. This question seems to have been framed in a way that avoids the problem associated with Buckle and Carlson's (2000) survey data.[15] Blinder et al. asked, "How often do the prices of your most important products change in a typical year?" The median number of price changes per year is 1.4; 10 percent of firms change their prices less than once per year, 29 percent of firms change their prices between two and four times per year, and 22 percent of firms change their prices more than four times per year.

Blinder et al. also include questions aimed at ascertaining the nature of any costs of price adjustment. They describe 12 theories of price stickiness and ask firms to rate the importance of each one for their own pricing practices. Some of their theories, such as procyclical elasticity of demand, cost-based pricing, and constant marginal cost, relate not to infrequent price adjustment but to "small" price adjustment. The most notable finding is that 65 percent of firms report that they have implicit contracts with customers to keep prices constant. "Implicit contracts" is a term usually used in the context of employer-employee relationships. Blinder et al. use it to refer to the customer costs that they trace to Okun (1981). That so many firms in the survey emphasize implicit contracts highlights the importance of repeated relationships

[12] The Internet may be changing this, in conjunction with the large databases now available.

[13] Of course, supermarkets also respond to customer complaints, but it seems unlikely that a significant number of complaints are related to price changes.

[14] The titles of the interviewees ranged from Manager to CEO.

[15] In Buckle and Carlson's data, respondents may have been reporting averages across products.

in determining price behavior, for implicit contracts only arise when there are repeated relationships. Repeated relationships show up in other responses to the study as well: 38 percent of firms report having *explicit* nominal contracts with customers. Furthermore, one-quarter of the firms report that they were inhibited from changing prices because they would antagonize customers and thereby lose future sales.[16] As for the direct menu costs of price adjustment, these are stated to be important by more than 25 percent of the firms surveyed.[17]

Simon Hall, Mark Walsh, and Anthony Yates (1997) conducted a survey similar to Blinder et al.'s for the United Kingdom. They obtained written responses to a questionnaire from 654 companies. The question about frequency of price change was, "In the last twelve months, how many times have you actually changed the price of your main product?" Roughly 6 percent of firms responded that they had not changed their price in the last year, 44 percent changed their price between two and four times, and 14 percent changed their price more than four times.[18] Hall, Walsh, and Yates's survey includes some of the same questions asked by Blinder et al. about theories of pricing. However, they only ask firms whether various general theories are important for their own pricing practices, not about details related to specific theories. In comments on a preliminary paper by Blinder, Blanchard (1994, p. 150) criticizes "theory recognition" questions: "The image . . . that recurs throughout my reading of the results is, that confronted with the twelve statements, the firms often had the reaction: 'Now that you say it, yes, maybe that *is* kind of what we do.'" If one is sympathetic to Blanchard's view, then some of Hall, Walsh, and Yates's results are quite striking. Several of the theories mentioned by Hall, Walsh, and Yates were recognized as relevant by *less than one-quarter* of firms, and the least recognized theory was that of physical menu costs, with only 7.3 percent recognition. As Blanchard notes, however, and as I have discussed above, theoretical work has shown that menu costs that are small from the firm's perspective may nonetheless have significant macroeconomic implications. Therefore, it is probably unwise to ignore evidence of small but widespread direct price adjustment costs gleaned from intensive studies solely because of the survey evidence.

5. CONCLUSIONS

The empirical research surveyed here leaves no doubt that the prices of many goods change infrequently (i.e., are sticky) relative to the frequency of changes

[16] The distinction Blinder et al. make between fear of antagonizing customers and implicit contracts is not clear.

[17] One quarter of the firms stated that there were physical costs to price adjustment. Many firms also reported managerial costs, but it is not possible to determine the amount of overlap between these two groups of firms.

[18] These numbers are approximate, as Hall, Walsh, and Yates provide graphical but not numerical frequency distributions.

in market conditions. I have, however, made no attempt to quantify a notion such as the GDP-weighted frequency of price adjustment, which would be a worthwhile endeavor. That prices do change only infrequently results from one form or another of costs of price adjustment.[19] The simplest type of price adjustment costs, physical menu costs, are easy to identify and straightforward to measure (though obtaining access to data may not be easy). These costs are nontrivial for retailers, but limited evidence for the industrial corporation studied by Zbaracki et al. suggests that menu costs may be quite small in other sectors. For the same industrial corporation, other administrative costs of price adjustment are quite large; these can be summarized as costs of the time it takes managers to choose prices. More observations about these costs from other firms would be valuable. The final form of price adjustment costs I have identified is more nebulous. Indirect customer costs involve a discrete shift in demand when a firm changes a price. There is a long theoretical tradition of studying such phenomena, but that tradition has not yet led to appealing macroeconomic models. Again, the work by Zbaracki et al. suggests that such costs do exist and that more work, both theoretical and empirical, is called for.

What do these findings tell us about macroeconomic modeling? It is an oversimplification to model all goods as symmetric with respect to price adjustment frictions. This has been the standard approach in recent research using sticky price models. Studying models where price adjustment is systematically less costly for some goods than others is straightforward. Incorporating long-term relationships in a way that can generate endogenous price rigidity, however, is a more important and less straightforward modeling extension.

REFERENCES

Akerlof, George A., and Janet L. Yellen. “Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?” *American Economic Review*, vol. 75 (September 1985), pp. 708–20.

____________. “A Near-rational Model of the Business Cycle, with Wage and Price Inertia,” *Quarterly Journal of Economics*, vol. 100 (Supplement 1985), pp. 823–38.

Backman, Jules, ed. *Price Practices and Price Policies, Selected Writings*. New York: The Ronald Press Company, 1953.

[19] Even Carlton’s theory involves costs or price adjustment in a sense: It is more costly to vary price than other characteristics of a product.

Barro, Robert J. "Long-Term Contracting, Sticky Prices, and Monetary Policy," *Journal of Monetary Economics*, vol. 3 (July 1977), pp. 305–16.

______________. "A Theory of Monopolistic Price Adjustment," *The Review of Economic Studies*, vol. 39 (January 1972), pp. 17–26.

Benabou, Roland. "Inflation and Efficiency in Search Markets," *The Review of Economic Studies*, vol. 59 (April 1992), pp. 299–329.

______________. "Search, Price Setting and Inflation," *The Review of Economic Studies*, vol. 55 (July 1988), pp. 353–76.

Blanchard, Olivier J. "On Sticky Prices: Academic Theories Meet the Real World: Comment," in N. Gregory Mankiw, ed., *Monetary Policy*. Chicago: University of Chicago Press, 1994, pp. 150–54.

Blinder, Alan, Elie R. D. Canetti, David E. Lebow, and Jeremy B. Rudd. *Asking About Prices: A New Approach to Understanding Price Stickiness*. New York: Russel Sage Foundation, 1998.

Buckle, Robert A., and John A. Carlson. "Menu Costs, Firm Size and Price Rigidity," *Economics Letters*, vol. 66 (January 2000), pp. 59–63.

Caballero, Ricardo J., and Eduardo M. R. A. Engel. "Heterogeneity and Output Fluctuations in a Dynamic Menu-Cost Economy," *Review of Economic Studies*, vol. 60 (January 1993), pp. 95–119.

Caplin, Andrew, and John Leahy. "Aggregation and Optimization with State-Dependent Pricing," *Econometrica*, vol. 65 (May 1997), pp. 601–25.

______________. "State-Dependent Pricing and the Dynamics of Money and Output," *Quarterly Journal of Economics*, vol. 106 (August 1991), pp. 683–708.

Caplin, Andrew, and Daniel Spulber. "Menu Costs and the Neutrality of Money," *Quarterly Journal of Economics*, vol. 102 (November 1987), pp. 703–25.

Carlson, John A. "Some Evidence on Lump Sum versus Convex Costs of Changing Prices," *Economic Inquiry*, vol. 30 (April 1992), pp. 322–31.

Carlton, Dennis W. "The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?" in Richard Schmalensee and Robert D. Willig, eds., *Handbook of industrial organization,* vol. 1. Amsterdam: North-Holland, 1989, pp. 909–46

______________. "The Rigidity of Prices," *American Economic Review*, vol. 76 (September 1986), pp. 637–58.

Cecchetti, Stephen G. “The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines,” *Journal of Econometrics*, vol. 31 (April 1986), pp. 255–74.

______________. “Staggered Contracts and the Frequency of Price Adjustment,” *Quarterly Journal of Economics*, vol. 100 (supplement 1985), pp. 935–59.

Danziger, Leif. “A Dynamic Economy with Costly Price Adjustments,” *American Economic Review*, vol. 89 (September 1999), pp. 878–901.

______________. “Costs of Price Adjustment and the Welfare Economics of Inflation and Disinflation,” *American Economic Review*, vol. 78 (September 1988), pp. 633–46.

______________. “Inflation, Fixed Costs of Adjustment and Measurement of Relative Price Variability: Theory and Evidence,” *American Economic Review*, vol. 77 (1987), pp. 704–13.

Dotsey, Michael, Robert G. King, and Alexander L. Wolman. “State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output,” *Quarterly Journal of Economics*, vol. 114 (May 1999), pp. 655–90.

Dutta, Shantanu, Mark Bergen, Daniel Levy, and Robert Venable. “Menu Costs, Posted Prices, and Multiproduct Retailers,” *Journal of Money, Credit and Banking*, vol. 31 (November 1999), 683–703.

Eden, Benjamin. “Inflation and Price Adjustment: an Analysis of Micro Data,” *Review of Economic Dynamics* (forthcoming).

Galbraith, John Kenneth. “Monopoly Power and Price Rigidities,” *Quarterly Journal of Economics*, vol. 50 (May 1936), pp. 456-75.

Goodfriend, Marvin. “A Framework for the Analysis of Moderate Inflations,” *Journal of Monetary Economics,* vol. 39 (June 1997), pp. 45–65.

Hall, R. L., and C. J. Hitch. “Price Theory and Business Behavior,” *Oxford Economic Papers* (May 1939), pp. 12–45.

Hall, Simon, Mark Walsh, and Anthony Yates. “How Do U.K. Companies Set Prices?” Bank of England Working Paper No. 67, July 1997.

Hicks, John R. “Annual Survey of Economic Theory: the Theory of Monopoly,” *Econometrica*, vol. 3 (January 1935), pp. 1–20.

Kashyap, Anil K. “Sticky Prices: New Evidence from Retail Catalogs,” *Quarterly Journal of Economics*, vol. 110 (February 1995), pp. 245–74.

Keynes, John Maynard. 1928 review of Frederick C. Mills, “The Behaviour of Prices,” reprinted in *The Collected Writings of John Maynard Keynes*, vol. XI. London: Macmillan, 1983, pp. 225–28.

Kuran, Timur. "Asymmetric Price Rigidity and Inflationary Bias," *American Economic Review*, vol. 73 (June 1983), pp. 373–82.

Lach, Saul, and Daniel Tsiddon. "Staggering and Synchronization in Price Setting: Evidence from Multiproduct Firms," *American Economic Review*, vol. 86 (December 1996), pp. 1175-96.

______________. "Small Price Changes and Menu Costs." Manuscript. January 30, 1996.

______________. "The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data," *Journal of Political Economy*, vol. 100 (April 1992), pp. 349–89.

Levy, Daniel, Mark Bergen, Shantanu Dutta and Robert Venable. "The Magnitude of Menu Costs: Direct Evidence from a Large U.S. Supermarket Chain," *Quarterly Journal of Economics*, vol. 112 (August 1997), pp. 791–825.

Mankiw, N. Gregory. "Small Menu Costs and Large Business Cycles: A Macroeconomic Model," *Quarterly Journal of Economics*, vol. 100 (May 1985) pp. 529–38.

Means, Gardiner C. "Industrial Prices and their Relative Inflexibility," S. Doc. 13, 74 Cong. 1 Sess. Washington: Government Printing Office, 1935.

Mills, Frederick C. *The Behavior of Prices*, New York: NBER, 1927.

Mussa, Michael. "Sticky Individual Prices and the Dynamics of the General Price Level," *Carnegie-Rochester Conference Series on Public Policy*, vol. 15 (1981), pp. 261–96.

Okun, Arthur M. *Prices and Quantities: A Macroeconomic Analysis*. Washington: The Brookings Institution, 1981.

Rátfai, Attila. "The Size, Frequency and Synchronization of Price Adjustment: Microeconomic Evidence." Manuscript. July 2000.

Rotemberg, Julio J. "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies,* vol. 49 (October 1982), pp. 517–31.

Scitovszky, T. De. "Prices Under Monopoly and Competition," *Journal of Political Economy*, vol. 49 (October 1941), pp. 663–85.

Sheshinski, Eytan, and Yoram Weiss. "Staggered and Synchronized Price Policies under Inflation: The Multiproduct Monopoly Case," *Review of Economic Studies,* vol. 59 (April 1992), pp. 331–59.

______________. "Optimum Pricing Policy under Stochastic Inflation," *Review of Economic Studies,* vol. 50 (July 1983), pp. 513–29.

______________. "Inflation and Costs of Price Adjustment," *Review of Economic Studies,* vol. 44 (June 1977), pp. 287–303.

Slade, Margaret E. "Optimal Pricing with Costly Adjustment: Evidence from Retail Grocery Prices," *Review of Economic Studies*, vol. 65 (January 1998), pp. 87–107.

Stigler, George J. "The Kinky Oligopoly Demand Curve and Rigid Prices," *Journal of Political Economy*, vol. 55 (October 1947), pp. 432–49.

______________, and James K. Kindahl. "Industrial Prices, as Administered by Dr. Means," *American Economic Review*, vol. 63 (September 1973), pp. 717–21.

______________. *The Behavior of Industrial Prices*. New York: National Bureau of Economic Research, 1970.

Stiglitz, Joseph E. "Imperfect Information in the Product Market," in Richard Schmalensee and Robert D. Willig, eds., *Handbook of industrial organization,* vol. 1. Amsterdam: North-Holland, 1989, pp. 771–847.

Sweezy, Paul M. "Demand Under Conditions of Oligopoly," *Quarterly Journal of Economics*, vol. 47 (August 1939), pp. 568–73.

Tucker, Rufus S. "The Reasons for Price Rigidity," *American Economic Review,* vol. 28 (March 1938), pp. 41–54.

Weiss, Yoram. "Inflation and Price Adjustment: A Survey of Findings from Micro Data," in Eytan Sheshinski and Yoram Weiss, eds., *Optimal Pricing, Inflation, and the Cost of Price Adjustment*. Cambridge, Mass: MIT Press, 1993, pp. 3–18.

Willis, Jonathan L. "Estimation of Adjustment Costs in a Model of State-Dependent Pricing," Research Working Paper 00-07. Kansas City: Federal Reserve Bank of Kansas City, December 2000.

Woglom, Geoffrey. "Underemployment Equilibrium with Rational Expectations," *Quarterly Journal of Economics,* vol. 97 (February 1982), pp. 89–107.

Wynne, Mark A. "Sticky Prices: What Is the Evidence?" Federal Reserve Bank of Dallas *Economic Review*, 1995 Q1, pp. 1–12.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen. "The Managerial and Customer Dimensions of the Cost of Price Adjustment: Direct Evidence from Industrial Markets." Manuscript. June 2, 2000.

Is Money Useful in the Conduct of Monetary Policy?

Michael Dotsey, Carl D. Lantz, and Lawrence Santucci

How useful monetary aggregates are for the conduct of monetary policy is a long-standing question. We analyze this question by examining their role as information variables in situations where the monetary authority uses an interest rate instrument. Monetary aggregates may be useful in that context if they contain information about the underlying contemporaneous state of the economy by helping to predict imperfectly observed variables that appear in the policymaker's reaction function. This use of money generally requires that the demand for money be well behaved and that random movements in the money demand function do not severely reduce the signal content of money. Alternatively, if the policy rule involves expectations of future variables, then money may be useful for predicting those variables.[1] Analyzing money's usefulness requires a very different statistical analysis for each of these two roles. The first role deals with the stability of the money demand relationship and the precision with which the money demand curve can be estimated, while the second role deals with the usefulness of money for forecasting.

While in practice monetary authorities do use monetary aggregates as information variables, their use varies over institutions and over time. For example, Hetzel (1981) indicates that the behavior of M1-influenced Federal Reserve policy decisions over part of the 1970s and Dotsey (1996) provides

We have benefited from a number of helpful discussions with Yash Mehra and Mark Watson. Bob Hetzel, Pierre Sarte, and Alex Wolman made many useful suggestions. The views expressed herein are the authors' and do not represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] For recent discussions of the role of forecasts in monetary policy, see Svensson (1999), Woodford (1999), and Amato and Laubach (2000).

evidence that money played a role during the early 1980s. The behavior of money, however, does not always enter into policy deliberations. Currently, most FOMC participants pay little attention to the growth rate of the monetary aggregates.[2] The time-varying use of money could be related to its time-varying usefulness. We here explore how money's behavior and predictive content have changed over time.

We do this in two ways. First, we analyze both the long-run and short-run behavior of M1 and M2 in Sections 1 and 2. We find that the parameters of the money demand function are time-varying and that our ability to explain money demand also varies over time. In Section 3, we look at how useful money is for forecasting nominal income, real output, and inflation. Our analysis indicates that M1 has been periodically useful in helping to forecast economic activity, but that its usefulness has waned. M2, on the other hand, has fairly consistently helped forecast nominal GDP and on occasion has been useful in improving the forecasts of real GDP. Section 4 concludes.

1. LONG-RUN RELATIONSHIP

We first examine the long-run relationship between money, income, prices, and interest rates. This investigation is important because it indicates the correct statistical specifications needed for the analysis in the rest of the article. If money, income, prices, and interest rates are cointegrated, then the empirical work that analyzes the demand for money and the predictive content that money has for future output growth and inflation must take account of the cointegrating relationship. Failure to do so will result in an improper specification of the empirical model.

The first step in any such investigation is to determine the order of integration of the relevant variables. These variables are: nominal M1; nominal M2; nominal GDP; real GDP; inflation as measured by changes in the GDP deflator; the three month Treasury bill rate; and the opportunity cost of holding M2, which is given by the difference between the T-bill rate and the own rate paid on M2 balances, real M1 balances, and real M2 balances. All variables with the exception of inflation and the two interest rate measures are measured in logs, and our sample goes from 1959:II through 2000:I. Other than the opportunity cost, all variables are nonstationary in levels. The stationarity of the opportunity cost reflects the cointegration between the T-bill rate and M2's own rate. It is not surprising that these two variables would exhibit a long-run relationship.

[2] A reading of recent policy discussions summarized in the regularly released minutes of FOMC meetings indicates that very little weight is placed on the behavior of money in the setting of policy.

Table 1 ADF Test Results

variable	test includes constant	test includes trend	test includes trend squared
M1	−3.07		−4.74
M2	−2.71	−3.01	
Y	−3.35	−3.76	−4.19
y	−6.35		−6.02
π	−2.34		
R	−5.52		
$m1$	−4.39		
$m2$	−4.27	−4.26	
5 percent critical value	−2.91	−3.45	−3.89

We then examine whether first differences of the variables are stationary or if the variables are integrated of order one. The results of augmented Dickey-Fuller (ADF) tests are displayed in Table 1.[3] Values of the test statistic that are less than the critical value indicate rejection of the null hypothesis that the variable is integrated. The lag lengths were chosen by the step-down method advocated in Ng and Perron (1995). When a trend or quadratic trend variable is significant in the regressions, test statistics are included for that specification. With the exception of nominal M2 growth and inflation, all the variables seem to be integrated of order one. Importantly, real M1 ($m1$) and real M2 ($m2$) are integrated of order one, and these variables will be used to investigate cointegration.

The results of our unit root tests, displayed in Table 1, are fairly standard. It is, however, worth presenting them since our sample size is somewhat larger than most reported studies. For example, given the recent move of many monetary authorities to explicitly or implicitly target inflation, one would expect inflation to eventually exhibit stationary behavior. It is worth checking to see if the professed change in emphasis on controlling inflation has shown up in the statistical characterization of nominal variables.

Cointegration

We now wish to look at the cointegrating relationship between real money balances, real income, and nominal interest rates. The two behavioral equations that inform our investigation are fairly standard specifications of the long-run relationship between real money balances, income, and interest rates:

[3] All unit root and cointegration tests were performed using the ADF, CADF, and PS procedures in the Gauss module, *Coint* written by Ouliaris and Phillips (1994–1995). These procedures produce the value of the relevant test statistic and its critical values.

$$m1_t = a + by_t - cR_t + e_t \tag{1.1}$$

and

$$m2_t = \alpha + \beta y_t - \gamma R_t - \delta(R_t - R_t^{M2}) + \varepsilon_t. \tag{1.2}$$

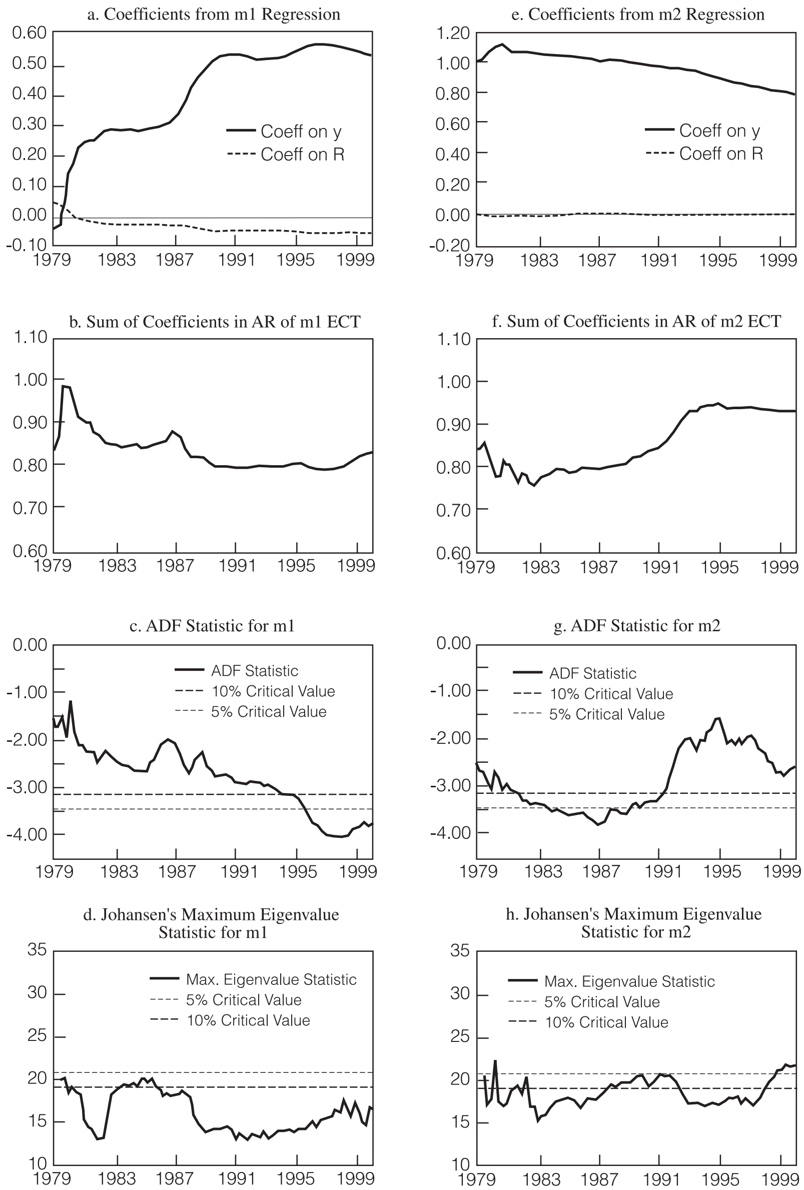
Equation (1.1) displays a simple demand function for real M1 balances as a function of real GDP and the nominal interest rate. Equation (1.2) depicts the demand for real M2 balances as a function of these same variables, as well as the opportunity cost of holding balances that are in M2 but not in M1. As mentioned, the money variables and output are in logs.

Before formally testing for cointegration we perform a heuristic exercise to examine the autoregressive behavior of the series. First, we recursively estimate a dynamic OLS regression of the respective real monetary aggregate on real GDP and the nominal interest rate. We use dynamic OLS, which includes leads and lags of first differences of the explanatory variables, to correct for correlation between the residual in the cointegrating relationship and the residuals in the processes generating the explanatory variables. The errors from the regression are computed as $m_t - \hat{a} - \hat{b}y_t + \hat{c}R_t$ for each definition of money. We then look at the sum of coefficients on a fourth order autoregression of this error; this sum can be thought of as the $\hat{\rho} - 1$ part of the Dickey-Fuller test statistic, $T(\hat{\rho} - 1)$. This sum is plotted in panels b and f of Figure 1. The sum is informative because it indicates the size of ρ, although no confidence intervals are calculated. One can see that the autocorrelation of the M1 residual declines over much of the sample, and as the sample size increases, it is likely that M1 will be judged to be cointegrated. The opposite is true of M2.

There are a number of issues involved in the various tests for cointegration proposed in the literature. Because the effect of the interest rate in money demand equations is generally small (as indicated in panels a and e of Figure 1), the presence of cointegration largely involves money's behavior with respect to output. Output is partially governed by a trend and partially governed by a nontrend nonstationary component. The various tests emphasize only one component of output in determining whether money and output are cointegrated. That is, the critical values of the tests are derived based on whether the asymptotic distributions are dominated by a time-trend or a random walk component. In reality both components are important, and for this reason the plots of sum of the coefficients in a fourth order autocorrelation are informative.

When testing for cointegration, one must take a stand on what portion of output is most important. In conducting augmented Dickey-Fuller tests, we assume the trend is the most important portion of output and follow the methodology advocated in Hamilton (1994, p. 597). First we perform auxiliary regressions of the interest rate and money on output. Of these two

Figure 1 Cointegrating Results



regressions, we then take the residual from the second (money) regression, and regress it on a constant, the residual from the auxiliary interest rate regression, and a time trend. Using the residual from this regression, we conduct an ADF test (see panels c and g of Figure 1). A test statistic that is less than the critical value indicates rejection of the null of no cointegration. Here we see that as the sample size increases, cointegration cannot be rejected for M1, but M2 appears to be cointegrated only over the first part of our sample. We should point out that the critical values for the tests are not uniform critical values for a sequence of random variables, but instead represent critical values that are appropriate for an individual test with a specific end date. Our tests only show what a researcher testing for cointegration at a specific date would find.

An alternate test for cointegration commonly performed in the literature is the Johansen (1988) maximum eigenvalue test (see panels d and h of Figure 1).[4] Here a test statistic above the critical value indicates that the variables are cointegrated. For both M1 and M2 the test produces results that are somewhat at odds with the ADF tests. For example, the Johansen test indicates that M1 was only cointegrated in the mid-1980s and is not cointegrated at present. It also indicates cointegration in the late 1970s. Given the behavior of the autocorrelation coefficient, the results appear counterintuitive. The autocorrelation coefficient for the error correction term has been relatively low in the 1990s, which should increase the likelihood that no cointegration among the variables will be rejected. Regarding M2, the Johansen test indicates cointegration, but shows that cointegration was not nearly so uniformly present in the 1980s. Both tests do, however, indicate cointegration in the late 1980s and early 1990s.

Stability

In our analysis of the long-run relationships' stability, the recursive estimates of the coefficients in the dynamic OLS regression on M1 seem to settle down as the sample size increases. More formal tests for parameter stability are conducted using the SupF and MeanF statistical tests developed in Hansen (1992).[5] For both tests the null hypothesis is that the coefficients are constant. The SupF test tests against the alternative of a single structural break at an unknown break date, while the MeanF test tests against the alternative: that the coefficients follow a martingale. The SupF test performs an F test for a structural break at each point on an interior interval of the data sample. The interval is chosen to allow sufficient sample size for constructing the F test. We can calculate the distribution of the supremum of the F test and derive a test statistic. Similarly, we can derive a distribution for the mean of the F-statistics. The SupF test rejects stability at the 1 percent significance level and the MeanF test rejects at the 10 percent significance level. The rejections

[4] The Johansen tests were conducted using the SJ procedure in the Gauss *Coint* package with a specification of a trend and six lags.

[5] We wish to thank Bruce Hansen for making available the code for performing these tests.

occur largely because of a sharp spike in the F-statistic in late 1980 and early 1981.

The coefficient on income in the M2 specification seems to be drifting downward while the coefficient on the T-bill has been increasing. It currently is positive, which makes little theoretical sense. Stability is, however, only rejected by the MeanF test at the 10 percent significance level.

Comparability of Results

Our results on cointegration are in agreement with a number of studies in this area. Stock and Watson (1989) reject cointegration for M1 using monthly data over the sample 1960:2 to 1985:12, which is consistent with our results since only after 1996 with the ADF test do we find cointegration at the 5 percent significance level. Our results are also consistent with the findings in Friedman and Kuttner (1992), who do not find that M1 is cointegrated over a sample ending in 1990:4, and with their finding for cointegration for M1 over the sample 1960:2 through 1979:2 if one employs Johansen's procedure, which they do. Miyao (1996), however, indicates that the Johansen test may overstate the finding of cointegration.

Regarding M2, Friedman and Kuttner find cointegration over their shorter sample, but only find evidence for cointegration for M2 at the 10 percent level over their entire sample, which ends in 1990:4. Given that they employ Johansen's method, their results are broadly consistent with ours. Like us, Miyao finds no evidence of cointegration for M2 over his entire sample, 1959:1 to 1993:4, but he also fails to uncover evidence for cointegration over his earlier subsamples when using both ADF and Johansen test statistics. His tests on earlier samples, which end in 1988:4 and 1990:4, are at odds with ours since we fail to reject the null at 10 percent significance levels. Our results are in greater agreement with those of Carlson et al. (2000), who find that M2 is cointegrated until about 1990. Swanson (1998), on the other hand, finds evidence for cointegration for both M1 and M2 over the period 1960:2 through 1985:12 using Johansen's methodology. His results are consistent with our M1 result, but not our M2 result. He uses monthly data, and it could be that sampling frequency is important for the test results, especially those involving M2. Lastly, our results are consistent with those of Feldstein and Stock (1994), who find that M2 velocity is cointegrated with the nominal interest rate. The coefficient on income elasticity is very close to one in the 1980s and early 1990s, so constraining it to be one as they do does not significantly affect the test results.

On the basis of our results and for conciseness, we choose to treat both M1 and M2 as cointegrated and include the estimated error correction term in the empirical work of the next two sections. We realize that the evidence in favor of cointegration is not overwhelming: that the evidence varies with sample periods, methodology, and data frequency. We therefore indicate those instances where our results are sensitive to the presence of an error correction term.

2. THE DEMAND FOR MONEY

We next investigate the time-varying behavior of the demand for money in order to shed light on whether the current behavior of money contains information useful to the monetary authority for controlling nominal income or inflation. This question is related to the desirability of monetary targeting. As emphasized by Friedman (1969), a well-defined and stable money demand curve is a necessary condition for monetary targeting to produce desirable economic outcomes, thus his emphasis on understanding the demand for money. Even if one does not wish to use money as an instrument or intermediate target, the current behavior of money may provide useful information about imperfectly observed variables such as current output or inflation. The usefulness of this information is related to understanding the demand for money, and we therefore share the same emphasis.

Lately the literature has moved away from this approach and has instead emphasized the notion of Granger causality. Recent examples include Friedman and Kuttner (1992), Estrella and Mishkin (1997), and Feldstein and Stock (1994). Those papers argue that in order for money to be useful in the conduct of monetary policy, it must have predictive content for some variable that the monetary authority cares about.

Money as a Signal

We believe the foregoing view is too restrictive. It neglects the signal value that money may have for contemporaneous and lagged values of economic variables that could plausibly be of interest to the central bank.[6] In reality, output and prices are not contemporaneously observable and are at best imperfectly observed with a lag. It may very well be that these variables, like the underlying shocks that impact the economy, may never be fully observed. In this case an optimizing monetary authority may find it desirable to use the economic information contained in money when setting its interest rate instrument. This point is made in Dotsey and Hornstein (2000), who consider the case of optimal time-consistent monetary policy. Their analysis would carry over to the study of optimal policy when the central bank is fully credible, or

[6] Furthermore, the notion of Granger causality involves general equilibrium considerations as pointed out in Dotsey and Otrok (1994). Since we are primarily concerned with money's usefulness when the central bank employs an interest rate rule we do not belabor these earlier points. Instead we concentrate on the contemporaneous signal value of money.

to a situation where the central bank was following a feedback rule that possessed desirable properties across a wide range of models. Using money as a signal of underlying state variables or of endogenous variables that may be part of some feedback rule could be helpful depending on how good a signal money is in practice. The value of that signal is directly related to the behavior of the demand for money.

To be more specific, consider a case where the monetary authority is following a rule in which the nominal interest rate target depends on output whose true value is never fully observed. Also, for simplicity assume that all variables are stationary and that output is the only endogenous variable not observed. That is, the price level, the interest rate, and nominal money are known. Simultaneously observing nominal M1, prices, and the nominal interest rate conveys the following signal,

$$s_m = (\hat{a} - a) + b(y_t - \bar{y}) + (\hat{b} - b)\bar{y} - (\hat{c} - c)\bar{R} + e_t,$$

where a bar over a variable indicates the variable's mean and a hat indicates an estimate of the parameter.[7] The monetary authority would in this case employ the Kalman filter to update its inference of output using the above signal. The precision of that estimate would depend on the variance of the money demand disturbance, which is directly related to how well money demand is behaved. It would also depend on the variance of the parameter estimates in the money demand regression.[8] In a case where the demand for money is stable, the variance of the parameters would get arbitrarily small as the sample size got larger. As more data were acquired, the estimation of the parameters would become more precise. Consequently, the signal content of money would then depend on whether one could well explain its current behavior. In a case where parameter estimates are time varying and unstable, the variance of the parameter estimates would not become arbitrarily small, and variability in the parameters would contaminate the signal value of money with respect to output.

The above explanation also applies to a situation where the variables are nonstationary and where perhaps all variables with the exception of the interest rate are observed with error. Whether money will be a useful signal of the level of income and prices will depend on how precisely it is measured and how precisely the cointegrating relationship is estimated. Thus, the stability properties analyzed in the previous section take on added significance apart from whether or not cointegration exists. The fact that the cointegrating vectors are unstable implies that money may provide a relatively poor signal of prices

[7] The signal is a first order linear approximation of the regression in equation (1).

[8] In the case where ouput is only observed with measurement error, the estimated coefficients will suffer from the effects of that measurement error as well.

and output. However, because the coefficients in the cointegrating relationship for M1 seem to be settling down and the rejection of stability was due to behavior in the early 1980s, the information contained in M1 may be more useful. In any event, how useful either monetary aggregate is will depend on the noise in its signal relative to the noise in other signals, such as reported output, that are available to the monetary authority.

An Error Correction Representation

The central bank may be interested not only in money as a signal, but also in the growth rate of output and prices, both past and present. Examining an error correction representation of the demand for money is therefore necessary if we are to ascertain money's usefulness in communicating the values of these variables. We now turn to that exercise.

The error correction money-demand equations that we estimate are

$$m1_t = a_0 + b_0(cv_{t-1}) + c(L)\Delta y_{t-1} + d(L)\Delta m1_{t-1} - e(L)\Delta R_{t-1} + u_t \tag{2.1}$$

for $m1$ and

$$\begin{aligned} m2_t \quad = \quad & \alpha_0 + \beta_0(cv_{t-1}) + \gamma(L)\Delta y_{t-1} + \delta(L)\Delta m1_{t-1} - \epsilon(L)\Delta R_{t-1} \\ & -\zeta(L)(R_t - R_t^{M2}) + u_t \end{aligned} \tag{2.2}$$

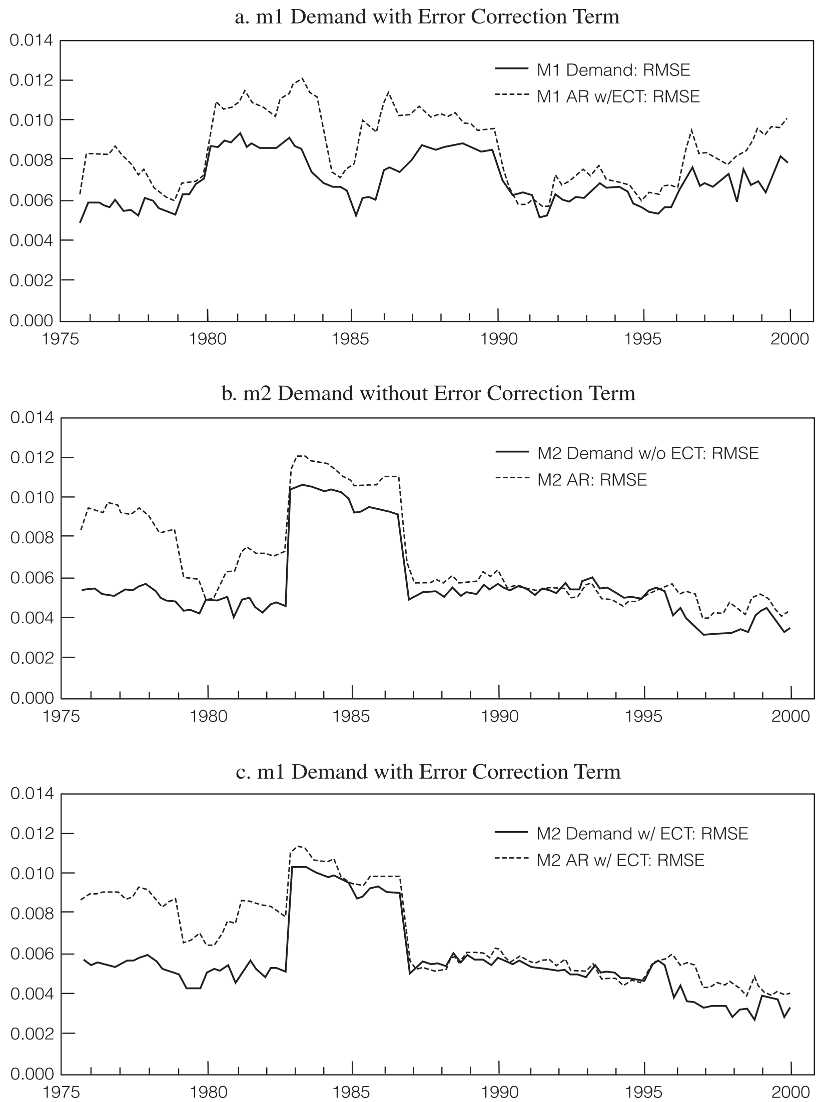
for $m2$, where cv_{t-1} is the error correction term. The $m2$ equation includes an additional term capturing the opportunity cost of holding balances in M2 that pay explicit interest. We also looked at the possibility of including polynomials in time, but they were found to be insignificant.

Using these equations we first ask if money demand was well explained at any given point in time. We do this by estimating 15-year rolling windows of money demand regressions and looking at the standard deviation of the residuals of those equations over 4 years.[9] We use rolling windows because of the voluminous amount of research indicating that these regressions are unstable over time. Later we confirm this instability. The results of this exercise are depicted in Figure 2, where the dates on the horizontal axis are the end dates of each sample period. Although we run the error correction models using rolling windows, we arrive at the estimates of the error correction terms, cv, recursively; the latter make use of all the available data up to the end date of the sample.

---

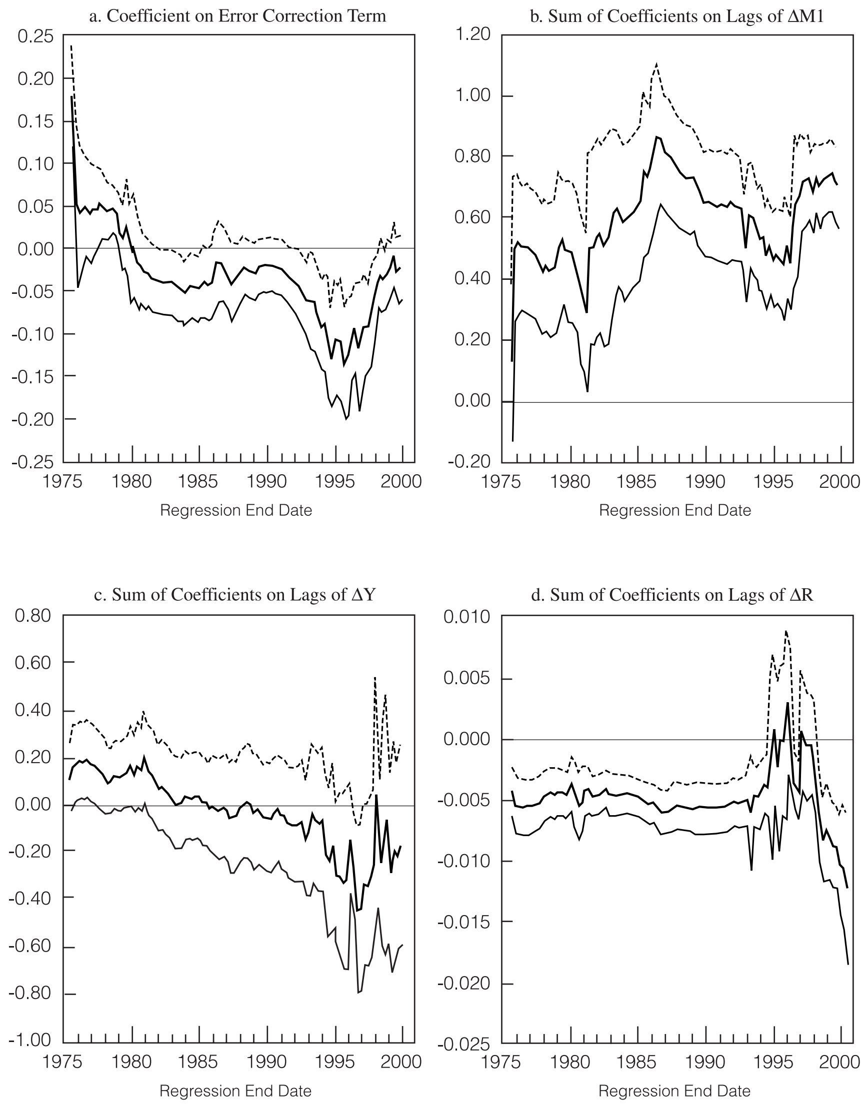
[9] All regressions are run using the robust errors routine in RATS, which corrects the standard errors of the regression coefficient when there is autocorrelation and heteroskedasticity in the errors.

Figure 2 In-Sample RMSE



This experiment shows how well the money demand regression explains the recent behavior of money. A benchmark is included that shows the errors occurring in a simple autoregression of money along with the error correction term. It is clear in the top panel that the ability of equation (3) to explain $m1$'s behavior varies over time with standard deviations ranging from approximately 40 basis points to 90 basis points. The early and mid-1970s reflect the best

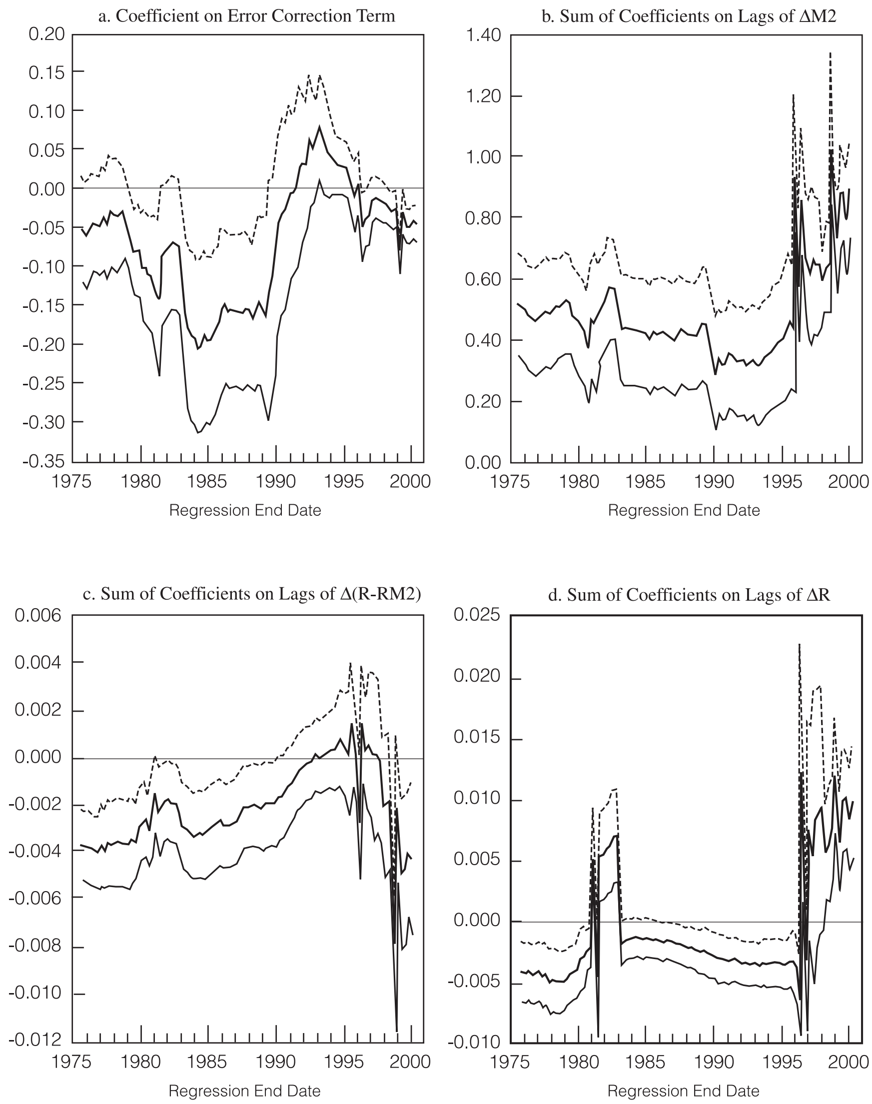
Figure 3 Coefficients in m1 Rolling Regression



performance of the regression and it is not surprising that this would be a period when monetary policy responded to M1 (see Hetzel [1981]).

Panel 2 of the figure examines M2's performance. Here the standard errors are slightly higher using $m2$ than $m1$. Also, the standard errors are relatively small at both the beginning and end of the sample, indicating that

Figure 4 Coefficients in m2 Rolling Regression



the $m2$ relationship was less variable in the 1970s and is currently fairly well behaved.

As we mention above, the signal content of money is related to the stability and the precision of the various coefficient estimates in the money demand regression. The value of the coefficients and their two standard error bands for the $m1$ regression are displayed in Figure 3. We do not display the constant since its value is small and insignificantly different from zero. If we exclude the end of the sample, the coefficients for the most part appear fairly stable. This stability was largely confirmed by the results of a time-varying parameter regression, but that regression did indicate statistically significant variation in the coefficient on the T-bill rate. We conduct a more formal test for stability in the presence of an unknown sample break using Andrews's (1993) sup Wald test. This test is basically similar to the SupF tests conducted in the previous section. To perform it, one constructs a Wald test for parameter constancy at each point on the interior of the data sample. A test statistic for the supremum of these values can be calculated, as can the statistic's critical values. In Figure 5, we graph the test statistic and the 5 percent critical value. The test rejects stability, with the rejection of stability arising from large values of the Wald statistic in the late 1960s and early 1970s. Between 1974 and 1993, the test statistic is below the 5 percent critical value.[10]

In Figure 4, we examine the behavior of coefficients in the M2 regression. The coefficients on the error correction term, the T-bill, $m2$, and the opportunity cost all show statistically significant variability. The coefficients on the last three variables fluctuate in the 1990s, but this high-frequency volatility did not have much influence on parameter estimates obtained using a time-varying parameter procedure. However, the Andrews test for stability (lower panel of Figure 5) does reject stability of the regression coefficients with the Wald statistic jumping above the 5 percent critical value in 1987.

The implications of this exercise for using money to help implement policy are decidedly mixed. For example, at times the demand for money appears to be well behaved, implying a close link between the behavior of money and the behavior of nominal output. At other times money demand is less predictable and the relationship appears unstable, implying that money may not be providing accurate information about the behavior of nominal income. Given this inconsistency and the desirability of following a simple and transparent rule of behavior, the central bank might reasonably decide not to use money in a feedback rule because the optimal response is likely to be time varying and difficult to explain.

The above findings do not imply that money serves no purpose. A number of economists recommend that the monetary authority respond to expectations of future variables such as expected future inflation.[11] In that regard money may communicate useful information about these variables. It is to this issue that we next turn.

---

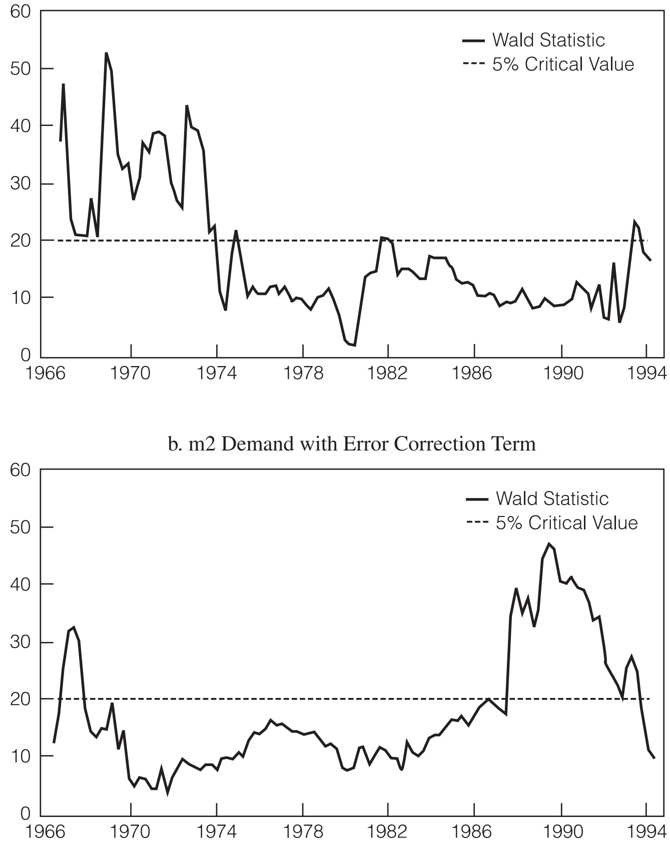
[10] Given instability and lack of significance of time in the full sample regression, we do not report any estimation using recursive procedures. It turns out that the in-sample errors using recursive regressions are similar to those of the rolling window regressions.

[11] Two recent articles that advocate such policies are Svensson (1999) and Amato and Laubach (2000).

Figure 5 Results of Andrews's Test for Structural Break

a. m1 Demand with Error Correction Term

b. m2 Demand with Error Correction Term



3. THE PREDICTIVE CONTENT OF MONEY

In this section we examine whether money has any useful predictive content for real GDP, nominal GDP, and inflation. As discussed in Dotsey and Otrok (1994), when the Fed uses an interest rate instrument that does not feed back on monetary variables, there may be a presumption against finding that money would Granger cause any of these variables. That presumption, however, is based on a number of restrictive assumptions, including the accurate observability of output and prices, that money balances do not serve in some buffer stock capacity, and that money demand shocks do not result from improvements in financial technology having significant effects on resource constraints. If observations on output and prices occur with significant lags and are subject to measurement error, then contemporaneous observation of money will be useful in solving the signal extraction problems faced by economic agents who are not completely informed. Therefore, observations on money will influence both agents' and the monetary authority's decisions and could help predict economic variables. Also, if agents accumulate money balances before engaging in expenditures, then large money balances today will indicate higher output in the future. Similarly, if changes in velocity are due to technological innovations that are persistent and affect resource availability, then observations on money will provide information about these innovations. An optimizing monetary authority should respond to these innovations, and hence money will have predictive content.[12]

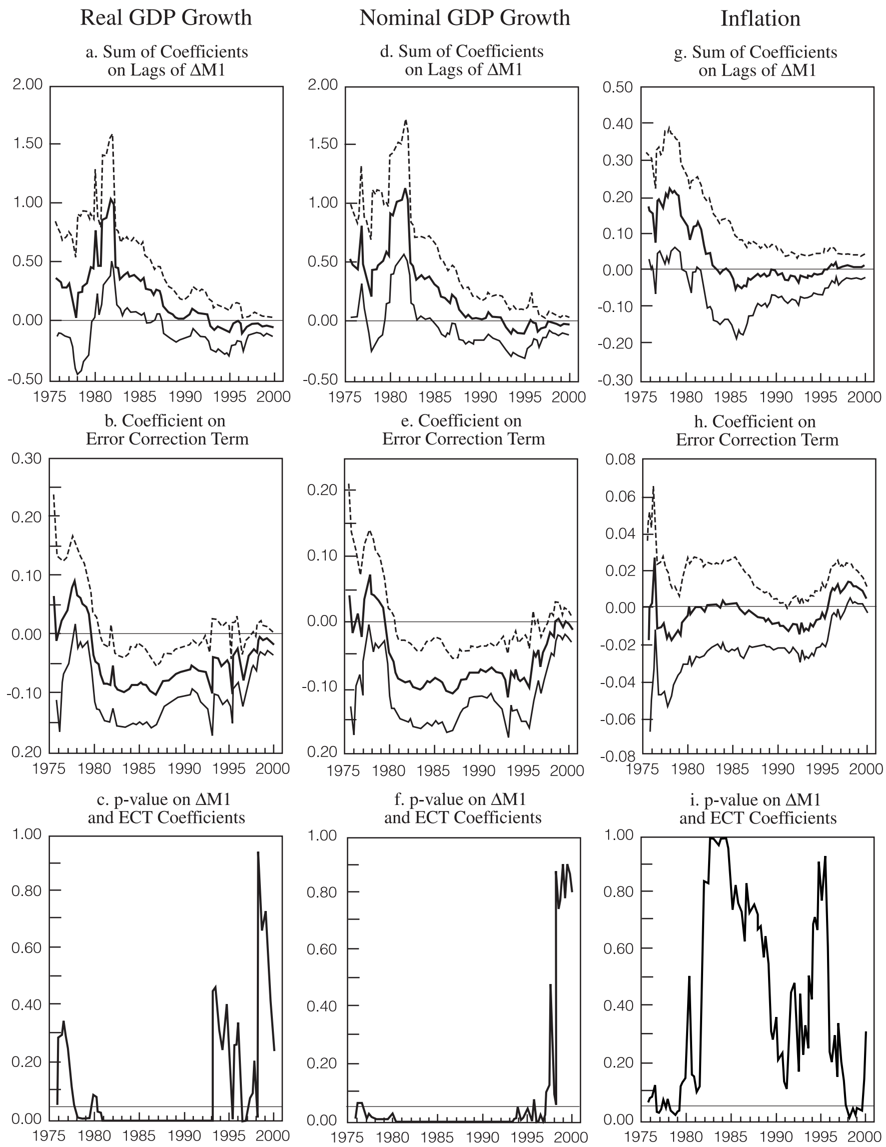
Figures 6 and 7 analyze the predictive content of M1, while Figure 8 investigates the predictive content of M2.[13] We should note that omitting the error correction term does at times worsen M2's forecasting ability. Figure 7 reports the same information regarding M1's predictive content, but also includes a time trend in the specification. This investigation follows from the recommendation of Stock and Watson (1989). The assumption that money is neutral in the long run implies that changes in trend money growth will not have any long-run consequences for output. In the short run the implications for changes in trend money growth could easily be quite different from those for cyclical changes. For example, in a model where firms change their prices only infrequently, the breakdown of how a change in money influences nominal income will in general depend on the persistence of the change in money growth (see Dotsey, King, and Wolman [1999]). If the change was perceived as either permanent or a change in trend, firms would be expected to aggressively change their prices, and the change in money growth would have a largely nominal impact. If the change was temporary or cyclical, the real effect could be significant. By putting a trend term in the forecasting equation, we are able to isolate the forecasting performance of cyclical changes in money growth.

We also conduct the analysis using 15-year rolling windows; as above, standard errors are corrected for the presence of heteroskedasticity and autocorrelation. We choose to use rolling windows based on evidence that the relationships are unstable. Our choice of a 15-year window is based on the results in Swanson (1998), who finds that 10-year rolling windows may be too short to give an accurate measure of the effect of money on industrial production. We also pick optimal lag lengths for each regressor using the Schwarz criteria.

[12] This is at least one of the theoretical messages in recent research by Dotsey and Hornstein (2000). Similarly, if money demand disturbances arise from shocks to preferences, the monetary authority will find it optimal to adjust the nominal interest in reaction to these disturbances or its best guess of these disturbances.

[13] The general forecasting model is an error correction specification where the growth rates of real and nominal GDP, as well as inflation, are regressed on a constant, an error correction term, lags of real GDP grouth, lags of money growth, lags of changes in the treasury bill rate, and lags of inflation.

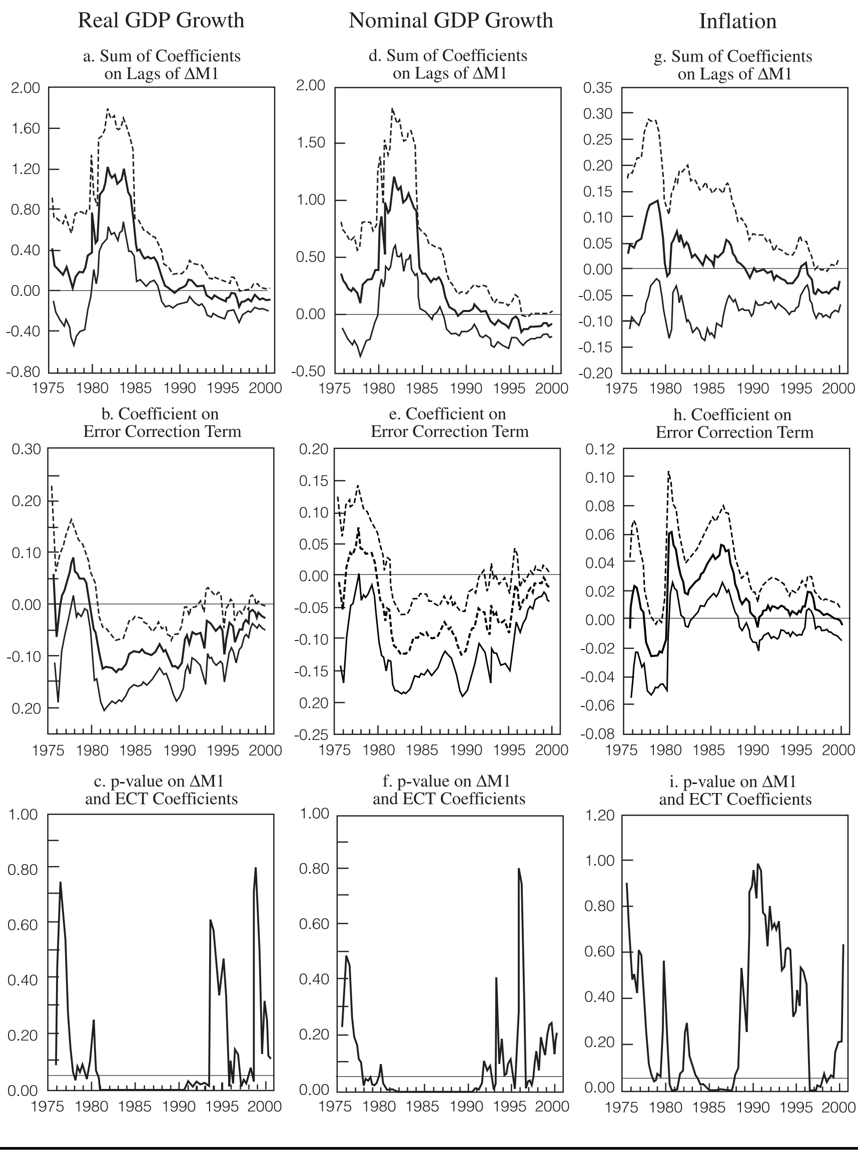
Figure 6 Predictive Content of M1



Results for M1

Figure 6 indicates that M1 had significant predictive content for real GDP and nominal GDP during the late 1970s and 1980s, but that it no longer helps forecast one quarter ahead movements in either of these variables. This

Figure 7 Predictive Content of M1 with Time Trend



finding is consistent with those of Estrella and Mishkin (1997) and the 6 lag specification of Stock and Watson (1989), but differs from the latter's 12 lag specification and from the results reported in Friedman and Kuttner (1992). With 12 lags, Stock and Watson do not find that nominal M1 Granger-causes real output over their sample 1960:2 to 1985:12. Friedman and Kuttner do not find evidence of Granger-causality over the sample 1960:2 to 1990:4; however, they do find predictive content for M1 over the subsample that ends in 1979:3. The difference between our results and those of Friedman and Kuttner is largely due to two main differences in our methodologies. One difference is that we find $m1$ and y to be cointegrated, and we therefore include an error correction term in the specification. The other is that we optimally select lag lengths; we generally end up with lags on M1 that are less than three quarters and often pick only one lag. Also, we look at rolling windows, but a recursive procedure produces results that are qualitatively similar. In the early part of the sample, much of the predictive content is coming from M1 growth, the sum of whose coefficients is positive and significantly greater than zero. In the 1980s much of M1's significance comes from the long-run or cointegrating relationship of real $m1$ with real output and interest rates. Interestingly this coefficient has a negative sign, which runs counter to the notion that M1 serves in a buffer stock capacity.

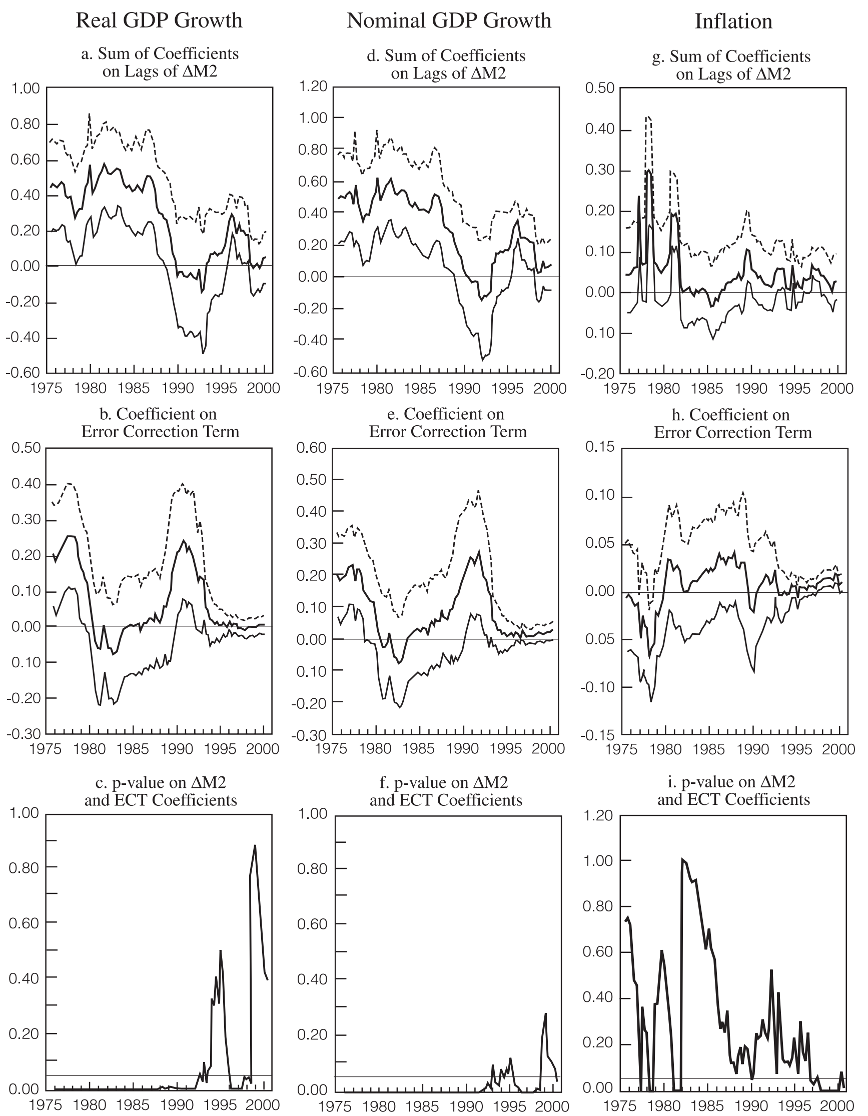
We also find that M1 helps predict nominal output through 1995 (see the middle column in Figure 6). This result is at odds with that reported in Feldstein and Stock (1994). We also observe that the behavior of M1 does not help forecast inflation (see the last column in Figure 6), which is consistent with the result reported in Cecchetti (1995).

Adding a time trend to the specification does not qualitatively have any impact on the results, which contrasts with the main message of Stock and Watson (1989). The contrast, however, could be due to lag length specifications because only Stock and Watson's 12 lag length specification produces the sharp differences in detrended versus raw money growth. Also, we include an error correction term, which would be picking up long-run relationships in both specifications. The inclusion of a trend term, therefore, may not have as much impact. Indeed, the coefficient on the trend term is insignificantly different from zero.

Results for M2

In Figure 8, M2 appears to have significant explanatory power in forecasting real GDP in the 1970s and 1980s, although it is no longer very helpful in that regard. It does Granger-cause nominal output over most of the sample, but it does not help predict inflation until the very end of the sample (see the last column of Figure 8). Furthermore, in the regressions on all three dependent variables, the sum of the coefficients on lagged M2 growth is positive. The coefficient on the error correction term is often insignificantly different from zero, but it happens to be significant in just those periods when the sum of the coefficients on lagged M2 growth is not. Thus, adding an error correction term provides overall help in predicting the three economic variables of interest. The general lack of statistical significance in the error correction term, however, indicates that there is no compelling evidence that broader money serves

Figure 8 Predictive Content of M2



as a buffer stock either. This last result is consistent with that of McPhail (1999), who analyzes Canadian data.

Our result that M2 is helpful in predicting the behavior of real and nominal GDP is consistent with that of Feldstein and Stock (1994), Dotsey and Otrok (1994), and Swanson (1998), but differs from that of Friedman and Kuttner (1992). It is also not consistent with the results in Estrella and Mishkin (1997), who find that M2 does not help predict nominal GDP over the sample 1979:10 to 1995:12 and that M2 does not Granger-cause inflation. They use monthly data, nine monthly lags, and the CPI deflator to measure inflation, while we use quarterly data, the GDP deflator, and varying lag lengths that are optimized for each sample. By looking at a comparable quarterly specification, we find that both the presence of an error correction term and the optimization over lag lengths are responsible for the difference in results.

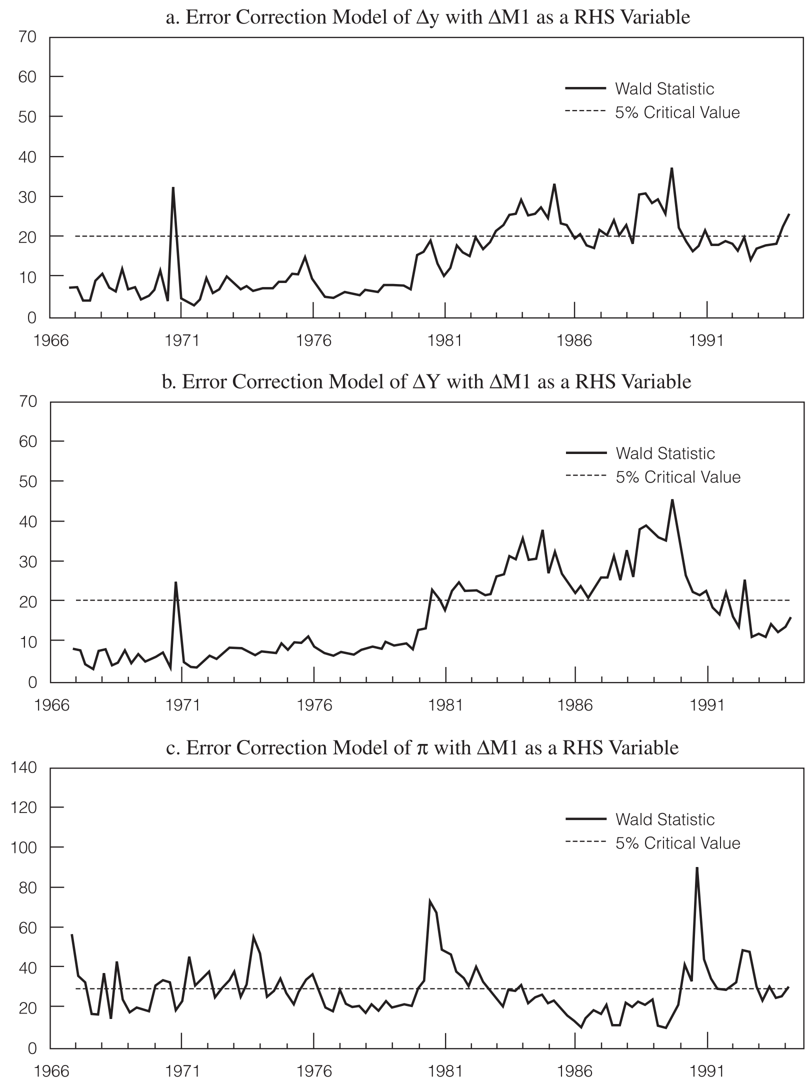
Our result that M2 does not help predict inflation is at first glance in conflict with the results presented by Cecchetti (1995) as well. He primarily looks at forecast horizons of a year and longer using monthly data, and he finds that M2 is significant for predicting inflation. He also finds evidence of instability in the relationship, with the worst predictive performance occurring between 1983 and 1989 although M2 is still significant at the 10 percent confidence level. If, however, we replace the GDP deflator with the PCE deflator, we find that M2 has significant predictive content for inflation over the 1990s, but fails to help predict inflation in the mid-1980s. One major difference between our study and that of Cecchetti is that the latter only includes M2 and lagged inflation in his specification, while the former also includes lagged interest rates and lagged output growth.

As with the results for M1, including a time trend does not appreciably affect the results of our study, so we do not report those results. There is, however, one particular change related to forecast horizon that makes a notable difference in our conclusions: In the context of predicting one-year-ahead nominal income growth using M2, M2 is always significant. The coefficient on the error correction term is large and significant in the late 1980s and early 1990s—just at a time when the coefficients on lagged M2 growth are insignificant. That is the only specification in which a monetary variable is uniformly informative about a potentially important macroeconomic variable. One should not get too excited about this result, however, because the coefficients move around a good deal and the relationship, while having good predictive ability, does not appear to be stable.

Stability

Feldstein and Stock (1994) use a battery of stability tests and find that the relationship between M2 and nominal income is largely stable, although there may be some parameter instability regarding the constant term. Feldstein and Stock also indicate that the M1–nominal income relationship is unstable; Figure 9 is consistent with that result. We again use the Andrews sup Wald test and graph the p-values for the test of a sample break at each date on the chart. Figure 10 indicates a rejection of stability for the relationship between M2 and the three dependent variables, and therefore our results differ from those of Feldstein and Stock.

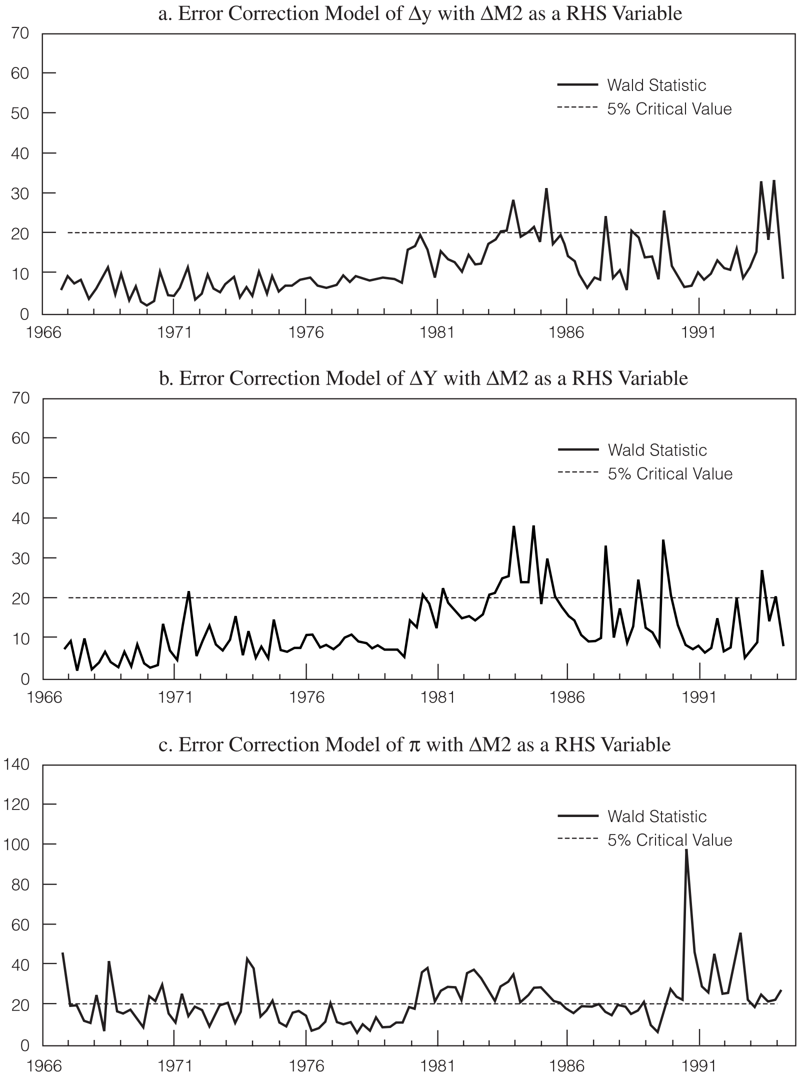
Figure 9 Results of Andrews's Test for Structural Break



4. SUMMARY

We have examined the behavior of both M1 and M2 with respect to their potential policy usefulness in providing information about contemporaneous but imperfectly observed variables or in helping to forecast future variables that may appear in an interest rate rule. As we show, the two notions are quite different and require different statistical investigations. By and large,

Figure 10 Results of Andrews's Test for Structural Break



the behavior of money itself is not reliable enough to advocate targeting either M1 or M2 or including them in a feedback rule. Their predictability varies substantially over time, and the coefficients in the various regressions we run do not appear to be stable. M1 and M2 do, however, seem to be useful in forecasting. Although their forecasting ability varies with time, the periods over which they often have significant predictive content can be prolonged enough to allow one to ascertain when those times occurred.

Even though the relationships we have investigated are not quite stable, much of their instability seems to be evolutionary in nature. That is, the changes in parameters appear to occur gradually. This fact suggests that a modeling strategy allowing the parameters to vary over time rather than holding them constant would better explain the behavior of the aggregates themselves and improve their forecasting ability. The biggest benefit to incorporating time variation might accrue from modeling the cointegrating relationship as evolving slowly over time. Using rolling windows and recursive estimation of the cointegrating relationship probably does not capture the behavior of money adequately. Financial innovations affect the behavior of money; these innovations are seldom radical and their adaptation is usually gradual. They are essentially an unobserved variable in the money demand regressions, and one hopes that future research will help account for their effects more thoroughly.

Furthermore, regulatory changes such as the elimination of regulation Q interest rate ceilings on personal checking accounts in 1981; allowing banks to offer MMDA accounts in 1983; changes in capital requirements that occurred in the late 1980s (see Lown et al. [1999]); and the relaxation of the use of sweep accounts in the 1990s have each had an impact on the demand for money. Some of these regulatory changes were no doubt reactions to technological changes that were taking place outside the banking sector, and thus they may be thought of as part of some endogenous process. Nevertheless, regulatory changes often have a discrete and uncertain impact on the demand for money. Policymakers are well aware of these changes, and modeling strategies can often be devised to incorporate them into the demand for money function; many, then, may view our investigation of money's usefulness as overly harsh. However, incorporating such regulatory changes formally into the behavior of money demand often requires a number of years of subsequent data, reducing the signal value of money during these episodes. For that reason, we refrain from accounting for the many regulatory changes occurring in the last 20 years. Nevertheless, we view our exploration of money's usefulness as a worthy exercise.

REFERENCES

Amato, Jeffrey D., and Thomas Laubach. "The Role of Forecasts in Monetary Policy," Federal Reserve Bank of Kansas City *Economic Review*, vol. 85 (second quarter 2000), pp. 21–32.

Andrews, Donald W. K. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, vol. 61 (July 1993), pp. 821–56.

Carlson, John B., Dennis L. Hoffman, Benjamin J. Keen, and Robert H. Rasche. "Results of a Study of the Stability of Cointegrating Relations Comprised of Broad Monetary Aggregates," *Journal of Monetary Economics,* vol. 46 (October 2000), pp. 345–84.

Cecchetti, Stephen G. "Inflation Indicators and Inflation Policy," in Ben Bernanke and Julio Rotemberg, eds., *NBER Macroeconomic Annual 1995.* Cambridge, Mass.: MIT Press, 1995.

Dotsey, Michael. "Changing Policy Orientation in the United States," in Horst Siebert, ed., *Monetary Policy in an Integrated World Economy*. Tubingen: Mohr, 1996.

____________, and Andreas Hornstein. "Is the Behavior of Money Useful to a Discretionary Policymaker?" Manuscript. 2000.

____________, and Christopher Otrok. "M2 and Monetary Policy: A Critical Review of the Recent Debate," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 80 (Winter 1994), pp. 41-59.

____________, Robert G. King, and Alexander L. Wolman. "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics,* vol. 114 (May 1999), pp. 655–90.

Estrella, Arturo, and Frederic S. Mishkin. "Is There a Role for Monetary Aggregates in the Conduct of Monetary Policy?" *Journal of Monetary Economics*, vol. 40 (October 1997), pp. 279–304.

Feldstein, Martin, and James H. Stock. "The Use of a Monetary Aggregate to Target Nominal GDP," in N. Gregory Mankiw, ed., *Monetary Policy*. Chicago: University of Chicago Press, 1994.

Friedman, Benjamin M., and Kenneth K. Kuttner. "Money, Income, Prices, and Interest Rates," *The American Economic Review*, vol. 82 (June 1992), pp. 472–92.

Friedman, Milton. "The Optimal Quantity of Money," in *The Optimal Quantity of Money and Other Essays.* Chicago: Aldine Publishing Company, 1969, pp. 1–50.

Hamilton, James D. *Times Series Analysis.* Princeton, NJ: Princeton University Press, 1994.

Hansen, Bruce E. "Test for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, vol. 10 (July 1992), pp. 321–35.

Hetzel, Robert L. “The Federal Reserve System and Control of the Money Supply in the 1970s,” *Journal of Money, Credit and Banking*, vol. 13 (February 1981), pp. 31–43.

Johansen, S. J. “Statistical Analysis of Cointegrating Vectors,” *Journal of Economic Dynamics and Control,* vol. 12 (1988), pp. 231–54.

Lown, Cara S., Stavros Peristiani, and Kenneth J. Robinson. “What Was Behind the M2 Breakdown?” Manuscript. 1999.

McPhail, Kim. “Broad Money: A Guide for Monetary Policy,” Bank of Canada. Manuscript. 1999.

Mehra, Yash P. “A Review of the Recent Behavior of M2 Demand,” Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 83 (Summer 1997), pp. 27–43.

Miyao, Ryuzo. “Does a Cointegrating M2 Demand Relation Really Exist in the United States?” *Journal of Money, Credit and Banking*, vol. 28 (August 1996), pp. 365–80.

Ng, Serena, and Pierre Perron. “Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag,” *Journal of the American Statistical Association*, vol. 90 (March 1995), pp. 268–81.

Ouliaris, Sam, and Peter C. B. Phillips. *Coint 2.0a: GAUSS Procedures for Cointegrated Regressions.* 1994–1995.

Saikkonen, Pentti. “Asymptotically Efficient Estimation of Cointegrating Regressions,” *Econometric Theory*, vol. 7 (March 1991), pp. 1–21.

Stock, James H., and Mark W. Watson. “Interpreting the Evidence on Money-Income Causality,” *Journal of Econometrics*, vol. 40 (1989), pp. 161–81.

Svensson, Lars E. O., “How Should Monetary Policy be Conducted in and Era of Price Stability?” in Federal Reserve Bank of Kansas City *New Challenges for Monetary Policy,* 1999, pp. 195–260.

Swanson, Norman R. “Money and Output Viewed Through a Rolling Window,” *Journal of Monetary Economics*, vol. 41 (June 1998), pp. 455–73.

Woodford, Michael. “Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?” in Federal Reserve Bank of Kansas City *New Challenges for Monetary Policy,* 1999, pp. 277–316.

The IT Revolution: Is It Evident in the Productivity Numbers?

Andreas Hornstein and Per Krusell

There is little doubt that we are witnessing a technological revolution. The question is, does this technological revolution have revolutionary economic consequences? In particular, is economic productivity growing at a much faster rate today, and if so, will it continue to do so in the future? In this article, we review recent literature on the measurement of productivity growth in the United States. We find considerable evidence that the internet technology (IT) revolution has had an impact on productivity.

In order to understand the effects of IT on today's economy, one should look at the past century. When we consider postwar U.S. productivity movements, two events stand out: the impressive productivity growth performance from the end of World War II up to the early seventies, and the ensuing productivity slowdown, which lasted until the mid-nineties. Labor productivity growth, which averaged about 2 percent per year from the fifties on, suddenly decreased to nearly 0 percent, and then seemed to settle at a rate around 1 percent. Moreover, this postwar productivity pattern is observed not only for the U.S. but throughout the western world.

This productivity slowdown remains quite poorly understood (for an overview and detailed data, see Hornstein and Krusell [1996]). One interpretation of the productivity data is that the fast postwar growth was a transitional period that made up for the losses during the Great Depression, and the post-1974 period of low productivity growth rates is really the normal state of the

Parts of this paper have been prepared for the European Commission, Directorate General for Economic and Financial Affairs, Seminar on "Economic Growth in the EU." Any opinions expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System. Andreas Hornstein: Research Department, Federal Reserve Bank of Richmond, andreas.hornstein@rich.frb.org. Per Krusell: Department of Economics, University of Rochester, pekr@troi.cc.rochester.edu, and consultant to the Research Department of the Federal Reserve Bank of Richmond.

economy. We are not convinced that this view is correct. From our perspective, the productivity slowdown is interesting because it occurred at the same time that IT applications became more widespread in the economy. The paradox is that new technology developments since that time have been associated with a productivity slowdown, and not an upturn, at least until quite recently. In other words, economists have had legitimate reasons to challenge those talking about a technology revolution on economic grounds: for it to have had significance, productivity growth (or economic welfare measured in some other way) ought to have gone up.

This article points to a number of reasons why the technology revolution may have had a significant impact on the economy's production structure despite its apparent insignificance in aggregate productivity statistics. First, we emphasize a number of methodological issues that may have prevented standard accounting procedures from detecting increases in productivity. Second, we take the view that the technology revolution may have affected the production structure in a quite asymmetric form, mainly showing its economic impact through changes in relative prices. Given this view, we investigate the hypothesis that the technology revolution, after all, has had important consequences on productivity: (1) it has led to radical changes in productivity among different sectors/factors of production; and (2) a number of factors related to the technology developments themselves have resulted in measures of *aggregate* performance that do not accurately reflect the (positive) effects on the economy. We find that while there is some support for this hypothesis, the evidence is not conclusive.

In Section 1 we review recent methodological and measurement advances in standard growth accounting, also known as "total factor productivity" accounting, which uses some basic economic theory to account for changes in productivity. The central question is the extent to which we have been, and are still, witnessing a technology revolution that has a large impact on the productivity of our economies. The IT boom has radical implications not just as an example of rapid structural change, but also from a measurement perspective. In particular, it seems to have brought about, and promises to bring many more, large changes in a range of products used as both inputs and outputs. Existing measurement methods may quickly become obsolete as products change and new products are introduced, and substantial work to improve these methods, both theoretically and with new forms of data collection, becomes of first-order importance. An increasingly large part of the output of our economy now has a quality-improvement aspect to it that is nontrivial to capture quantitatively. Quality mismeasurement is a long recognized problem, especially in the service sector (where output is often measured directly by input), and we can learn from attempts by economists to deal with this sector. In this article we discuss—at a broad level—what the main problems are and what advances have been implemented so far. We then present the most recent estimates of aggregate productivity change for the United States using improved methods.

In Section 2 we review some recent literature on factor-specific productivity (FSP) accounting. The FSP approach imposes additional theoretical structure on the measurement of productivity and attempts a more detailed account of the sources of the productivity change. The motivation for moving from TFP to FSP measurement is based on the large changes in various relative prices that we have observed during the past few decades. First, equipment prices have fallen at a rapid rate; the seminal work by Gordon (1990) documents these developments in detail, based on careful quality measures of a large range of durable goods. Second, relative prices of skilled and unskilled labor have gone through large swings, most recently with a large increase in the relative wage of educated workers. These relative price changes suggest that there are factor-specific productivity changes. More theory is needed—that is, more assumptions need to be made—in order to gain more precise insights into the nature of the technology changes. We therefore spend time developing some theory necessary to shed light on these issues on a conceptual level, before discussing some recent practical applications.

1. TOTAL FACTOR PRODUCTIVITY

Concept

Standard economic theory views production as the transformation of a collection of inputs into outputs. We are interested in how this production structure is changing over time. In this section we derive the basic concepts used in productivity accounting.

We keep things simple and assume that there is one output, y, and two inputs, capital k and labor n. The production structure is represented by the production function, F: $y = F(k, n, t)$. Since the production structure may change, the production function is indexed by time t. Productivity changes when the production function shifts over time, i.e., there is a change in output that we cannot attribute to changes in inputs. More formally, the marginal change in output is the sum of the marginal changes in inputs, weighted by their marginal contributions to output (marginal products), and the shift of the production function $\dot{y} = F_k \dot{k} + F_n \dot{n} + F_t$.[1] This is usually expressed in terms of growth rates as

$$\hat{y} = \eta_k \hat{k} + \eta_n \hat{n} + \hat{z}, \text{ with } \hat{z} = F_t / F,$$

[1] The marginal change of a variable is its instantaneous rate of change over time; that is, if we write the value of a variable at a point in time as $x(t)$, then the marginal change is the time derivative $\dot{x}(t) = \partial x(t)/\partial t$. Nothing is lost in the following if the reader interprets $\dot{x}(t)$ as the change of a variable from year to year, that is, $x(t) - x(t-1)$.

where hats denote growth rates and the weight on an input growth rate is the elasticity of output with respect to the input: $\eta_k = F_k k/F$ and $\eta_n = F_n n/F$. Alternatively, if we know the elasticities we can derive productivity growth as output growth minus a weighted sum of input growth rates. Indeed, it was Solow's (1957) important insight that under two assumptions we can replace an input's output elasticity, which we do not observe, with the input's share in total revenue, which we do. First, we assume that production is constant returns to scale, i.e., if we double all inputs, then output will double. This implies that the output elasticities sum to one: $\eta_k + \eta_n = 1$. Second, we assume that producers act competitively in their output and input markets, i.e., they take the prices of their products and inputs as given. Profit maximization then implies that inputs are employed until the marginal revenue product of an input is equal to the price of that input. In turn, this implies that the output elasticity of an input is equal to the input's revenue share. For example, for the employment of labor, profit maximization implies that $p_y F_n = p_n$, which can be rewritten as $\eta_n = F_n n/F = p_n n/p_y y = \alpha_n$ (p_i stands for the price of good i). With these two assumptions we can calculate productivity growth, also known as total factor productivity (TFP) growth, as

$$\hat{z} = \hat{y} - (1 - \alpha_n)\,\hat{k} - \alpha_n \hat{n}.$$

Implementation of the Solow growth accounting procedure thus requires reliable information on the prices and quantities of inputs and outputs. We discuss below some of the issues that arise in productivity accounting and how they are affected by the current advances in information technologies.

Implementation and Weaknesses

In this section we discuss issues of aggregation, changes of quality versus quantity, missing inputs, and available observations on prices. Finally, we briefly discuss the underlying assumptions of constant returns to scale and perfect competition.

Aggregation

Any modern economy produces a large variety of commodities and uses an equally large variety of commodities as inputs in production. In order to make useful statements on the overall performance of the economy, we have to define broad commodity aggregates. The theoretically preferred aggregation procedure is the construction of Divisia indexes (see, for example, Jorgenson, Gollop, and Fraumeni [1987] or Hulten [1973]).[2] In practice we approximate a Divisia index with a chain-linked price and quantity index.

[2] We define the Divisia index below.

As an example, consider the production structure described above, but assume that there are two types of labor: unskilled labor n_u and skilled labor n_s, which trade at prices w_u and w_s. Suppose that skilled and unskilled labor combine and generate the labor aggregate $n = G(n_u, n_s)$. The aggregator function G is constant returns to scale. Using the same arguments as above for the Solow growth accounting procedure, we can write aggregate labor growth as a cost-share-weighted sum of the skilled and unskilled labor growth rates

$$\hat{n} = \omega_u \hat{n}_u + \omega_s \hat{n}_s,$$

where $\omega_u = w_u n_u / (w_u n_u + w_s n_s)$ and $\omega_s = w_s n_s / (w_u n_u + w_s n_s)$. Notice that the aggregator function is time invariant, i.e., the productivity of skilled and unskilled labor does not change over time. Essentially, this is an identification assumption: given our assumptions we can only make statements about aggregate productivity, not about factor-specific productivity. We will return to this issue below. In general, if we have prices and quantities for a collection of commodities $\{(p_i, q_i) : i = 1, \ldots, m\}$ we assume that there is a constant returns-to-scale aggregator function $\bar{q} = Q(q_1, \ldots, q_m)$, and define the growth rate of the aggregate quantity index as

$$\widehat{\bar{q}} = \sum_{i=1}^{m} \frac{p_i q_i}{\bar{p}\bar{q}} \hat{q}_i \text{ with } \bar{p}\bar{q} = \sum_{i=1}^{m} p_i q_i$$

and $\bar{p}$ is the implicit aggregate price index. The expression for the growth rate of aggregate output is also the definition of the Divisia quantity index for this particular collection of commodities.

The level of aggregation depends on the focus of the research. Recent research on the effects of IT has tried to establish how much of output growth can be attributed to IT capital accumulation and if the spread of IT has affected TFP growth in various industries differentially. For this purpose, researchers have constructed separate aggregates for IT related capital, such as computing and communications equipment and other capital goods. In this context, Divisia indexes have the nice property that they aggregate consistently. Therefore, we can first construct industry TFP growth from industry output, capital, and labor growth, then use the industry data to construct aggregate output, capital, and labor growth, and finally obtain aggregate TFP growth, which is also a weighted sum of industry productivity growth rates (see Jorgenson et al. [1987]).

Price, Quantity, and Quality

Until now we have worked according to the assumption that it is easy to obtain prices and quantities for any particular commodity. Yet the commodity structure of an industrial economy is not static: existing commodities are improved upon or replaced by new commodities. Very often the distinction between improvement and replacement is just a matter of degree, and it is more useful to think of products as having certain quality properties that are relevant to the purpose for which the commodity is used, be it in consumption or production. For example, the average car today is very different from the same car 20 years ago in terms of its performance characteristics, its maintenance requirements, etc. One way to measure the car production in the economy would simply be to count the number of units produced, but this method clearly does not reflect the superior quality of today's car relative to yesterday's car, and it would lead one to underestimate output growth and thereby productivity growth. The appropriate procedure is then to adjust a car for its quality content, i.e., a car is weighted according to its characteristics. Not accounting for quality change in the production of cars would lead one to underestimate output and productivity growth.

When commodities differ according to quality, we have to construct quality-adjusted price measures. In order to see if a broadly defined commodity has become expensive over time, we do not compare the price of two similar commodities at different points in time, but we compare the quality-adjusted prices. Only an increase in the quality-adjusted price represents a true price increase. The use of quality-adjusted prices, or hedonic prices, was pioneered by Griliches (1961). Gordon's (1990) work on durable goods shows that the relative price of durable goods declines at a substantially faster rate once one accounts for quality change.

The counterpart of this observation is that, using Gordon's (1990) price deflator, the quantity of durable goods produced increases at a faster rate. For the case of investment goods, disregarding quality change leads one not only to underestimate output growth, but also to underestimate input growth because investment is used to construct capital stocks. The effect on measured productivity growth can be ambiguous since we underestimate both output and input growth. Obviously, adjusting for quality change is more important for new products in innovative industries such as IT. Recently this approach has been successfully applied to the construction of computer price indexes in the United States (see Cole et al. [1986]). This application was successful because there was a well-defined and easily measured set of characteristics describing the performance of computers.

When it is difficult to apply hedonic pricing—e.g., when it is impractical or conceptually hard to distinguish products' characteristics—there can be other ways to make quality adjustments. In a recent innovative paper by Bils and Klenow (1999), one such method is developed and put to work. Loosely speaking, the alternative to brute-force measurement of quality components of products proposed by Bils and Klenow is to use theory. Consider, for instance, vacuum cleaners. Most households own one. However, there are many brands and qualities of vacuum cleaners, and by looking at detailed household data one can find out, on average, to what extent additional household income translates into a more expensive vacuum cleaner. After using the cross section to find out how income translates into quality, Bils and Klenow (1999) turn to the time series and use aggregate changes in household income to predict the added quality component in vacuum cleaner purchases over time. This method is applied to a broad set of products, and the results can be summarized in a downward revision of the growth of the official Bureau of Labor Statistics price index by over 2 percentage points per annum (with a corresponding upward revision in real output growth).

The problems of quality adjustment are well recognized but hard to resolve. With the explosion of new IT-related products, quality mismeasurements are likely to be more severe. Fortunately, IT advances are likely to ease data collection in the future, but the brute-force method surely needs to be complemented with alternative methods, such as that of Bils and Klenow (1999).

In the service sector, a failure to account for changes in quality may explain the poor productivity performance in a variety of industries. A prominent example in the United States is the banking sector, where until recently bank output was extrapolated based on bank employment. As heavy users of IT, the financial industries have substituted equipment capital for labor. Given the measurement procedure, the capital-labor substitution had the unfortunate consequence of lowering measured output while maintaining or increasing the use of inputs. An incorrect measure of output thus results in apparent productivity declines. Recent revisions define bank output as an employment-weighted sum of the number of different transactions performed (see Moulton [2000]). However, this still does not correct for changes in the quality of transactions, such as convenience, reliability, and speed. One might expect that future IT developments will allow for further improvements along these quality dimensions.

Another sector whose contribution to the aggregate economy has increased and where quality aspects are very important is the provision of health services. For this sector one would have to construct medical diagnosis price indexes that would account for the accuracy of the diagnosis and the inconvenience to the patient, as well as treatment price indexes that would account for the success rate, intrusiveness, side effects, etc. of the treatment (see Shapiro, Shapiro, and Wilcox [1999]).

Finally, to our knowledge, all productivity accounting exercises assume that the quality of different types of labor remains constant over time. That is, the overall quality of the aggregate labor force may increase because the economy employs more skilled than unskilled labor, but the quality of skilled and unskilled labor is assumed to remain the same. We return to this issue in the next section.

Missing Inputs and Outputs

Closely related to the measurement of quality is the problem of missing inputs and outputs, which can bias measured rates of productivity growth. This problem is evident in the treatment of expenditures on Research and Development (R&D) and software and in the construction of capital stock series from investment and depreciation.

Progress can be made toward solving the problem of missing inputs and outputs. The recent National Income and Product Account (NIPA) revision in the United States now includes one previously missing capital input, namely computer software (see Moulton, Parker, and Seskin [1999]). Before the NIPA revision, software was not treated as an investment good, but as an intermediate good; it therefore did not add to final demand. The inclusion of software investment has contributed about 0.2 percentage points to total U.S. GDP growth from the early 1990s on (see Seskin [1999]).[3] Jorgenson and Stiroh (2000) find that software investment contributes about 5 percent to output growth and that software capital makes up about one-ninth of total capital accumulation. It is not clear if R&D spending should be treated the same way as spending on software. Since R&D spending generates knowledge, similar to "organizational" capital acquired by firms when they learn how to use new IT, one could interpret this knowledge capital as a missing input. On the other hand, these inputs—knowledge and organizational capital—are not traded commodities so their accumulation might as well be captured as productivity improvements.

Related to the question of missing inputs and outputs are Kiley's (1999) and Whelan's (2000) discussions of capital stock measurement. Research has shown that the perpetual inventory model with geometric depreciation is a reasonable approximation to observed depreciation patterns for durable goods. If the depreciation rate, δ, of a class of durable goods is constant over time, then the net increase of the capital stock, k, is investment, x, minus depreciation, δk:

$$\dot{k} = x - \delta k.$$

Kiley (1999) argues that for high rates of capital accumulation, such as those observed for IT over the last 20 years, observed investment expenditures do not capture all resource costs associated with capital accumulation. Resource costs that are incurred in the process of new capital formation and are in addition to the observed investment expenditures are called "adjustment costs." Adjustment costs can affect productivity measurement in two ways. On the

[3] The 1999 NIA revisions increased real GDP by about 0.4 percentage points for the period 1992–1998. Revised price index numbers contributed about 0.1 percentage points. The remaining increase was mainly due to the revised treatment of software expenditures, but also reflects the effects of the revised output measurement in the banking sector (Seskin 1999).

one hand, the presence of adjustment costs could mean that standard measures of output underestimate true output because they are net of adjustment costs. On the other hand, we could say that standard procedures overestimate true capital accumulation because the marginal product of investment in the production of capital goods is not constant at one, but declining. Whelan (2000) argues that for computers, current procedures overestimate depreciation rates because they confuse physical depreciation—which lowers the effective capital stock—with economic depreciation, which does not affect the effective capital stock.[4]

Finally, we want to raise the perennial favorite issue of market versus nonmarket transactions, or how to value household production. One of the presumed improvements generated by the Internet is the additional convenience it provides to consumers: the Net makes product search and price comparisons easier, provides access to certain services, and reduces overall the time households have to spend on transactions. While one can interpret this problem as one of missing inputs and outputs, to us it appears to be a quality measurement problem. From the point of view of the household, transactions have certain characteristics, and a commodity or service obtained through the Net has different characteristics than the same commodity or service obtained in a store. Very little work has been done to date to assess the value of these changes in characteristics.

Input Cost Shares

For the Solow growth accounting, we identify an input's output elasticity with the input's revenue share. We therefore need measures of the prices or rental rates of all inputs. This measurement is apparently not a problem for the calculation of aggregate productivity growth, for we have data on payments to labor and can treat payments to capital as the residual. There is an issue,

---

[4] Usually the depreciation rate for a particular type of capital good is estimated from a cross section of quality-adjusted used capital goods prices. Older capital goods are less efficient due to depreciation, and this is reflected in their lower prices. Assuming geometric depreciation, the slope of the age-price line for the different vintages of used capital goods reflects the depreciation rate. Suppose in addition not only that capital goods depreciate, but that their lifetime is finite because at some age it is no longer profitable to operate a capital good (economic depreciation). In this case the slope of the age-price line also incorporates the effects of a declining remaining service life for the capital good. For capital goods with long service lives this effect is not very important, but for computers, which currently have quite short service lives, this effect may dominate the physical depreciation.

Whelan (2000) suggests that for computer equipment, current procedures identify all of the slope of the price-age line with physical depreciation, even though most of it is due to a finite service life. If this is true, then this procedure may underestimate the computer capital stock during the early phase of computer capital accumulation because it depreciates computer capital too fast. On the other hand, this procedure also assumes that capital is around for a much longer time period than the actual service life of capital, that is, it overestimates the capital stock in the later phase of computer capital accumulation. One would expect that after some time, these two effects would balance and the measured capital stock would be about right.

however, about how to allocate proprietor's income; part of it is payment for labor services and part represents capital income, but this issue appears to be minor. Our problem is not completely solved, however, since we still have to construct aggregate capital and labor series. If we define aggregate capital and labor as an equally weighted sum of the different types of capital and labor in the economy, we are done. These measures of capital and labor, however, are not the theoretically preferred Divisia indexes as discussed above, and for the construction of a Divisia index we need to know the amount of the payment made to each component of the index, not just the sum of all payments. This requirement creates a problem for the construction of capital aggregates: Because capital is usually owned by producers, the services of capital are not traded on spot markets, and we therefore have no observations on rental rates for different types of capital. The usual procedure is to make an assumption on the rate of return on capital and then calculate the rental rate of capital implied by an asset pricing equation (see Jorgenson, Gollop, and Fraumeni [1987]). The asset pricing relationships are derived under the assumption that there is no uncertainty, which is not an innocuous assumption if one believes that the economy can experience a technological revolution. During a revolutionary period, one would expect that first, overall uncertainty would increase, and second, that not all types of capital would be equally affected by the increased uncertainty.

Solow's Assumptions

Let us now consider the two theoretical assumptions underlying Solow growth accounting: production is constant returns to scale and producers are competitive. In the context of business cycles analysis, it has been argued that Solow growth accounting systematically mismeasures true changes in productivity because the underlying assumptions do not hold, namely that there are increasing returns to scale and that producers do not equate price to marginal cost (see Hall [1988]). Extensive research in this area has not completely resolved all issues, but it is apparent that constant returns to scale and competition are reasonable approximations (Basu and Fernald 1999). A remaining problem for short-run productivity accounting appears to be a missing input: unobserved factor utilization, which is essentially a theoretical construct and is very difficult to measure. We do not expect this problem to affect medium- to long-term growth accounting since theory suggests that there is no trend growth to this particular activity.

We usually attribute changes in TFP to technological improvements, but clearly TFP also reflects the effects of government regulations, market structures, firm organization, prevailing work rules, etc. In a nice case study, Schmitz (1998) describes how in the 1980s changing competitive pressures on the U.S. and Canadian iron-ore mining industry induced mines to change their work practices in a way that increased their productivity without actually changing their technology. Standard Solow growth accounting registers the change in TFP, but it does not explain why it occurs. It is probably also true that productivity changes of this variety will result in one-time improvements, and cannot account for sustained productivity growth. Nevertheless, an explanation of this observation would be useful because many discussions of IT suggest that it makes the environment more competitive and induces firms to respond with technological and organizational improvements. A possible explanation might be based on a framework where (1) managers, workers, and capital owners somehow share the surplus from operating a firm, (2) the relative surplus shares depend on the relative bargaining strengths of the different parties, and (3) changes in competitive pressures have a differential impact on the parties' bargaining positions. This framework would represent a radical departure from the standard Solow growth accounting assumptions.

Recent Evidence on IT and Productivity Growth

We now review the most recent work on IT and aggregate productivity growth using standard Solow growth accounting. The papers of Jorgenson and Stiroh (2000), Oliner and Sichel (2000), Whelan (2000), and Kiley (1999) focus on the contribution of IT capital to aggregate output growth and changes in aggregate productivity growth rates during the 1990s. There are some differences in the studies' particular definitions of IT capital, but the studies share the conclusion that in the 1990s the contribution of IT capital accumulation to output growth increased and productivity growth increased. Hall (1999, 2000) proposes a method for measuring unobserved capital accumulation associated with the diffusion of IT and studies the implications for TFP growth.

Aggregate TFP

Jorgenson and Stiroh (2000) identify IT capital with computer hardware and software and with communications equipment. They find that from 1973–1990 to 1995–1998, the contribution of IT to aggregate growth doubles and productivity growth triples. For the earlier time period, IT investment accounted for one-tenth of output growth, and IT capital accumulation made up one-fifth of total capital accumulation and about one-tenth of total output growth. In the more recent period, IT investment accounts for about one-fifth of output growth, and IT capital accumulation makes up two-fifths of total capital accumulation and about one-fifth of total output growth. At the same time, productivity growth increases from an annual rate of 0.3 percent to 1 percent, a rate that is about as high as the golden era of the 1950s and 1960s. Oliner and Sichel (2000), using the same definition of IT capital and a somewhat narrower definition of output, find a similar increase in productivity growth and contribution of IT capital accumulation to output growth. Both studies find the increase in productivity growth rates to be limited to the post-1995 period.

Whelan (2000) identifies IT with computing equipment and argues that standard measures of depreciation overestimate the physical depreciation rates of computing equipment. His estimates of computing equipment stocks in 1998 exceed standard values by almost 50 percent, which would indicate an enormous measurement error. The implied faster growth rates and higher revenue shares for computing equipment double the contribution of IT capital accumulation to output growth. Since aggregate output growth is not affected by the redefinition of depreciation rates, the higher contribution of IT capital accumulation is offset by a corresponding decline in the contribution of other capital and overall TFP growth.

In sum, the most recent studies find important productivity improvements for the very last part of the 1990s, but the productivity slowdown still appears to be a mystery. It is possible that the productivity slowdown may simply be indicative of structural change and that an increasing share of the economy is badly mismeasured. This hypothesis is discussed in some detail in Hornstein and Krusell (1996), although the main conclusion from that work—and any other work we are aware of—is that there is suggestive but not much hard evidence at this point that would allow us to reassess the 1973–1995 period. Structural change can have important implications in economies with a number of adjustment costs in the form of learning, reorganization, etc., and we have noted the mismeasured quality problem in our discussion of methods.

In this context, Kiley (1999) also identifies IT with computing equipment and performs a growth accounting exercise that allows for adjustment costs to capital accumulation.[5] He finds that in the 1970s and 1980s high IT capital accumulation rates actually reduced observed net-output growth. Unlike other studies, which use standard Solow growth accounting, he finds that aggregate productivity growth has remained constant from the 1970s to the 1990s. Adjustment costs are a theoretical concept used to motivate why short-run movements in investment are less volatile than predicted by the standard growth model; it is very difficult to obtain direct evidence on them. Since adjustment costs are used to study short-run dynamics, they are usually normalized such that they are zero when the investment-capital stock ratio is at its long-run average. The available evidence on adjustment costs seems to relate to local deviations from long-run averages. Obviously, for a new product like IT, the investment-capital stock ratio will be very different from its long-run value, which means that adjustment costs might be substantial and can have a

[5] Solow growth accounting is based on standard economic theory, as represented by the growth model, and it does not include adjustments costs. The introduction of adjustment costs forces Kiley (1999) to deviate from the usual nonparametric Solow growth accounting procedure: He has to specify a functional form for the adjustment cost function.

measurable impact on net-output growth. On the other hand, we do not know if the local properties of adjustment costs apply when investment-capital stock ratios are far from their long-run averages.

Finally, Hall (1999, 2000)—or eHall, according to his latest paper—starts out with the assumption that in the 1990s the spread of IT was associated with the accumulation of a new type of capital, e-capital for short, and that e-capital is not measured by standard National Income and Product Accounting. This assumption reflects the observation that the market value of private corporations relative to the replacement cost of their physical capital increased from a factor of one at the beginning of the 1990s to a factor of three at the end of the 1990s. eHall extends the standard growth accounting framework and assumes that measured output is produced with physical capital, e-capital, skilled labor, and unskilled labor and that new e-capital is accumulated through the employment of skilled labor.[6] Using data on the market value of firms in addition to the usual series on quantities and factor rental rates employed in growth accounting, he constructs a series for e-capital. He finds that with e-capital, the contribution of other inputs and TFP to output growth is substantially reduced: without e-capital, TFP accounts for two-fifths of total output growth; with e-capital the combined contribution of e-capital and TFP accounts for three-fourths of total output growth, and most of it is due to e-capital. eHall's approach has the undesirable property that large stock market reevaluations imply the creation/destruction of large amounts of e-capital, since equity makes up a large portion of the market value of firms. This property does not fit well with our understanding of capital as a durable good. We can account for large equity market reevaluations if we assume that a substantial fraction of a firm's assets are indeed not reproducible, but that market values reflect current and future production opportunities. This is an important point to consider in an environment where new technologies change the way we see the future.

Disaggregated TFP

Finally, we may want to know where technical change takes place. Is it concentrated in particular industries, or do we see a general increase in TFP for all industries? Is industry TFP growth related to the use of IT? Evidence on these points is mixed.

Jorgenson and Stiroh (2000) report TFP growth rates for a range of two-digit industries. They find that TFP growth varies widely: the best-performing industries include Trade, Electronic and Electric Equipment, Agriculture, Industry Machinery and Equipment, Transport and Warehouse; the worst-

[6] eHall also relates the interaction of e-capital and skilled and unskilled labor to the idea of skill-biased technical change. We discuss this issue in the sections on multifactor productivity growth below.

performing industries with negative TFP growth are Services and Finance, Insurance, and Real Estate (FIRE). Two observations are applicable. First, the results concerning the relative ranking of industries and the fact that large parts of the economy (FIRE, Services) show negative TFP growth rates are similar to those of previous studies, i.e., accounting for IT has apparently not had an impact.[7] Second, the impact of IT on particular industries appears to be mixed. Given the advantages IT provides to inventory control and production planning, industries such as Trade and Transportation and Warehousing should have benefited from the diffusion of IT. On the other hand, although a substantial fraction of IT investment is going to Services and FIRE, the productivity performance of these industries has not improved at all. Above, we have suggested that for these industries IT diffusion may simply worsen the output measurement problem.

More specifically related to the production of computing equipment is Oliner and Sichel (2000), who use changes in the relative price of computers and semiconductors to evaluate the contribution of the sector producing computing equipment to aggregate TFP growth.[8] They find that despite the relatively small revenue share of the computing equipment sector, that sector accounts for about half of total TFP growth.

Microstudies

We have noted that it is difficult to find any clear relationship between the utilization of IT and the resulting TFP growth at the industry level. However, Brynjolfsson and Hitt (2000a) argue that more evidence on the impact of IT applications on productivity is available for firm level data. Brynjolfsson and Hitt (2000b) estimate the impact of computing equipment on TFP growth at the firm level using a variation of Solow growth accounting. Essentially they argue that TFP growth at the firm level is positively correlated with the growth of computer capital in the firm. They also suggest that the benefits from investment in computer capital are delayed, which can be interpreted as being due to needing to learn to use IT or to the accumulation of IT-related organizational capital.

Bresnahan, Brynjolfsson, and Hitt (1999) study the interaction of IT capital accumulation, firm-specific human capital accumulation, and organizational change at the firm level. They use survey data on firms to construct an index of human capital (average education levels, skill levels as perceived by management, and occupational mix), an index of human capital investment (training and screening activities), and an index of how "decentralized" the

[7] Unfortunately, Jorgenson and Stiroh have not yet calculated industry TFP numbers for post-1995.

[8] See also our discussion of Greenwood, Hercowitz, and Krusell (1997) in the next section.

firm's organizational structure is (measures of team orientation). They find that their measures of IT capital, human capital, and work organization are all positively correlated. In particular, (1) firms that have a better educated work force and a more decentralized work organization tend to use more IT capital, and (2) firms that use more IT capital tend to spend more on training their work force.

Finally, Brynjolfsson and Yang (1999) study the relation between the stock of computer capital in a firm and the market value of that firm. They argue that one dollar of computer capital in the firm raises the market value of a firm by more than one dollar. They suggest that this markup reflects other "unmeasured" capital which is complementary to computer capital. In this context they point out that when a firm implements a new information management system, the biggest cost component is consulting, training, and software development, not hardware expenditures.[9]

2. FACTOR-SPECIFIC PRODUCTIVITY

At times of significant technological change, the relative importance of different inputs, or factors of production, may change substantially because of specific technological innovations. For example, during the last 30 years we have witnessed striking changes in the relative prices of new equipment capital and in the premiums paid to highly educated workers (skilled workers for short). These changes likely reflect factor-specific technology movements, that is, technological advances that have enabled some factors to enjoy large increases in marginal productivity while others have seen none or have decreased. In other words, factor-specific productivity measurements may capture the economic signs of a technological revolution even when TFP measurements show tranquility. The increased productivity of one factor may lead to changes in the provision of factors—by changing the amount of hours worked of different kinds of labor or by causing changes in the accumulation of physical and human capital—in such a way that TFP does not change much.

This hypothesis has been described and compared to past technology revolutions, such as the introduction of electricity in the beginning of the 20th century, in Greenwood and Yorukoglu (1997). In this article, we only discuss this possibility on a broad level; future work will explore it in more detail. We point out the advantage of these multidimensional productivity measures, and we underline their shortcomings. We also present some recent examples of empirical work aimed at factor-specific productivity (FSP) measurement.

[9] The recent reclassification of software expenditures as investment should ameliorate this problem in the NIPA.

Introducing Factor-Specific Productivity

Solowian productivity accounting shows how the productivity of all inputs changes while it imposes a minimal amount of theoretical structure. Factor-specific productivity accounting imposes more theoretical structure than the Solowian method does, but also allows us to evaluate productivity changes for individual inputs. Factor-specific productivity accounting recognizes that TFP growth may not be exogenous, but in fact can depend on the relative use of different inputs. This recognition has important consequences for policy evaluation. For example, whether or not government taxes or subsidizes the accumulation of equipment capital affects capital stock accumulation, which in turn may affect measured TFP growth. In order to evaluate the effect of a tax or subsidy on equipment investment, one would need to know more about the *nature* of the productivity improvements. In particular, it may be that technological change interacts asymmetrically with different inputs and that the effect of capital accumulation on TFP growth depends on the nature of these interactions. We leave the discussion of policy and other counterfactual experiments to a future paper. Here, we will show some useful ways of allowing technological change to interact differently with different inputs: we will characterize productivity change multidimensionally.

To illustrate the difference between TFP growth and factor-specific productivity growth, consider the basic neoclassical production structure with output a function of capital and labor (see Section 1 above). Now impose the additional assumption that output is a *time-invariant* function of *efficiency units* of each input

$$y = F(A_k k, A_n n),$$

where k is the number of machines and A_k is a machine-specific productivity factor that changes over time (and similarly for labor). The factor-specific productivities are assumed to be exogenous. The marginal change in output is $\dot{y} = F_1(\dot{A}_k k + \dot{k} A_k) + F_2(\dot{A}_n n + \dot{n} A_n)$. Assuming constant returns to scale and perfect competition we can write the growth rate of output as

$$\hat{y} = (1 - \alpha_n)\,(\hat{A}_k + \hat{k}) + \alpha_n(\hat{A}_n + \hat{n}),$$

where α_n is the labor income share, and the components of TFP growth are

$$\hat{z} = (1 - \alpha_n)\hat{A}_k + \alpha_n \hat{A}_n.$$

Since, with the exception of a Cobb-Douglas production function, the labor income share depends on the input ratio, TFP growth will depend on inputs. If the elasticity of substitution between inputs was greater than one, the labor income share would decrease with an increase in the stock of capital. In this case, we would expect that a subsidy to capital increases TFP growth if and only if $\hat{A}_k > \hat{A}_n$. In addition, TFP growth is not a goal in itself. Rather, a government's objective might be a higher output growth rate, in which case the relative importance of A_k and A_n clearly matters again.

The disadvantage of factor-specific productivity accounting is that it imposes substantially more structure than the assumption of constant returns to scale and marginal-product pricing, as is sufficient for Solow's TFP accounting.[10] As we will discuss below, much more structure is typically needed in order to draw inference about factor-specific productivity. The example above assumes that the production technology is invariant over time with the inputs measured in "efficiency units," such that technological change can take only the form of increases in the efficiency factors.[11] The quantitative results depend on the form of the F function (on the elasticity of factor substitution especially).

Finally, we would like to address the potential quantitative importance of considering FSP measures. One potential objection to FSP is based on the observation that over the last century, aggregate labor and capital shares have been remarkably stable in the United States. These stable factor shares suggest that the aggregate production function for the United States is well approximated by the Cobb-Douglas function with constant factor shares and unit elasticity of substitution. But with unit elasticity FSP and TFP accounting are equivalent. There are several other observations, however, that indicate substantial variation of factor income shares. First, stable shares are mainly observed for the very broadest aggregates: labor versus capital. For breakdowns of the labor input—into different skill (educational) groups—shares have had strong trends and swings around trend. For the capital income share, there is information that for our object of interest—new equipment/IT-related capital—the cost share has increased dramatically. Second, if one looks at a cross section of countries, especially including countries at a lower level of development, then one sees that there is variation in the labor share with development; at the very least, countries do not seem to have the same labor shares. Third, and on a related point, some developed countries have had much larger swings in the aggregate shares than what has been observed in the United States; one example is the dramatic increase of profit shares in France in the 1980s.

---

[10] There is a sense in which factor-specific productivity (FSP) accounting imposes less structure than TFP accounting. In the example above we back out FSP for capital, given a functional form for production. But FSP essentially represents a change in quality, and a more structured procedure would try to obtain measures of quality, as discussed above for TFP.

[11] There are alternative, tractable structural approaches to how technology changes over time. For example, one could specify a CES function with elasticities changing exogenously over time. More structurally still, Cordoba (2000) describes the sequential adoption of output technologies with higher and higher capital shares. He then shows, with closed-form solutions, that this form of "structural change" implies increases in the capital-output ratio while allowing the interest rate to remain constant; these properties seem to well approximate the development path of many countries.

We now turn to factor-specific productivity measurement for the four factors mentioned most often: capital equipment and structures and skilled and unskilled labor.

Investment-Specific Technological Change

The recent decades of technological change and the current focus on IT are often described as examples of how technology is "embodied" in new equipment. The idea is that productivity improvements occur in the sector producing equipment investment, and these productivity improvements are transmitted to the rest of the economy through new equipment investment. Today's higher productivity in the investment-goods producing sector then effectively enhances the production possibilities for consumption in the future, through increases in the capital stock. The embodiment question has a long history; see, for example, Solow (1959) and Jorgenson (1966), as well as a recent discussion and evaluation in Hercowitz (1998). Greenwood, Hercowitz, and Krusell (1997) argue that for the postwar United States, especially after the 1970s, most productivity growth was of the embodied variety.[12]

In its most basic version, capital-embodied technological change represents changes in factor-specific productivity: currently produced new capital goods are relatively more productive than previously produced capital goods. On the other hand, once we measure capital in terms of efficiency units, we can interpret capital-embodied technological change in terms of product-specific changes in TFP, namely the productivity of the economy's investment goods sector relative to the consumption goods sector. We follow this second interpretation, but we discuss capital-embodied technological change in the current section on changes in FSP, as opposed to in the previous section on changes in TFP, because we impose considerably more theoretical structure when we derive measures of capital-embodied technological changes.

We first provide a general discussion of TFP accounting in a simple two-sector model of the economy and how it relates to the usual measures of aggregate TFP accounting. We follow this route in order to show how the assumptions used by Greenwood et al. (1997) allow them to interpret their results in terms of aggregate productivity in a one-sector economy. Finally, we present Greenwood et al.'s (1997) results on the relative contributions of sectoral TFP growth to aggregate growth.

Goods—consumption c and new capital x—are produced using the factors capital and labor as inputs to constant-returns-to-scale technologies

$$c = z_c F_c(k_c, n_c) \text{ and } x = z_x F_x(k_x, n_x);$$

[12] This finding relates to the observations by Jorgenson and Stiroh (2000) and Oliner and Sichel (2000) on sectoral TFP growth discussed in the previous section.

total factor inputs can be freely allocated across sectors,

$$k_c + k_x = k \text{ and } n_c + n_x = n;$$

and investment is measured in efficiency units. The technologies may differ across sectors because of different factor substitution properties (F_c may differ from F_x), and technological improvements may occur at different rates (z_c may grow at a different rate than does z_x). We normalize productivity relative to the consumption goods sector, $z_c = z$ and $z_x = qz$, and q is the relative productivity advantage of producing new capital goods. The evolution of the capital stock is described by

$$\dot{k}(t) = x(t) - \delta k(t).$$

One could now proceed and calculate sector-specific TFP growth as described in the previous section. Greenwood et al. (1997) choose an alternative route and use this setup to calculate and interpret TFP growth within an aggregate growth accounting framework.

In the first section of the present article, we postulated—in accordance with Solow's assumptions—a measure y of aggregate output as equaling an aggregate production function of capital, labor, and time. It was not made explicit in that framework what output was made up of. In the one-sector neoclassical growth model, aggregate output is by definition equal to the sum of consumption and investment, $y = c + x$, but what is aggregate output in a multisector economy, such as the two-sector model just described? Rather than starting with an aggregate output concept, we will first summarize the production possibilities in the two-sector economy described above by the transformation function $G(c, x, k, n, t) = 0$. The function G tells us what combinations of consumption and investment goods c and x the economy can produce, given its total factor inputs k and n. Since production is constant-returns-to-scale, the transformation function is homogeneous of degree one in outputs and inputs.

Following our discussion of Divisia indices in Section 1, we can define an aggregate output index and a measure of aggregate TFP growth based on this output measure:

$$\hat{y}^D \equiv s_c\hat{c} + (1 - s_c)\hat{x} \text{ and } \hat{z}^D \equiv \hat{y}^D - \alpha\hat{k} - (1 - \alpha)\hat{n}.$$

The growth rate of the Divisia index of aggregate output is a weighted average of the output growth rates in the two sectors, where the weights are the revenue shares of consumption and investment, $s_c = p_c c/(p_c c + p_x x)$.[13] The growth rate of aggregate TFP is then defined analogously to the one-sector economy as the difference between the aggregate output growth rate and the weighted average of the aggregate input growth rates, where the weights are the aggregate income shares of capital and labor, $\alpha = p_k k/(p_k k + p_n n)$.

[13] See our definition of Divisia indices in the section "Aggregation," above.

Divisia indices allow us to perform aggregate productivity accounting, but there is no particular theoretical justification suggesting that a Divisia index is the unique aggregator function for the economy.[14] In fact, one can show that for multi-sector models there does not exist an output aggregator; that is, in general no function exists that relates some measure of aggregate output to measures of aggregate inputs (Hall 1973). However, let us now assume that G is separable so that it is possible to find an output aggregator:

$$G(c, x, k, n, t) = H(c, x, t) - F(k, n, t).$$

Here, we interpret F as the aggregate production function and H as the aggregate output function, and both functions are homogeneous of degree one. In particular, aggregate output is defined as $y = H(c, x, t) = F(k, n, t)$. Notice that we must in general allow both of these functions to depend on time in order to allow technological change of a general kind.

Now this setup can be specialized further to illustrate different kinds of technological change: rather than allowing the variable t to have a general influence on H and F, consider instead $F(zk, zl)$ and $H(c, x/q)$, where z and q are time-dependent processes. That is, technological change is expressed only through z and q, with z representing neutral technological change and q investment-specific technological change.[15] Can we obtain measures for the two types of technological change for this aggregate specification of the economy? We might want to proceed as in the case of the Solow residual, and define the productivity growth rates based on the equations which relate output growth to input or expenditure growth rates

$$\hat{y} = \alpha\hat{k} + (1-\alpha)\,\hat{n} + \hat{z} = s_c\hat{c} + (1-s_c)\left(\hat{x} - \hat{q}\right).$$

But here we face a problem: Although our theory suggests that there exists an output aggregate, we do not have a measure of that output aggregate. In order to construct the measure of aggregate output, we need to know the functional form of H and F and the values of the productivity levels z and q. One way to proceed is to assume that there is no investment-specific technological change, that is, q is constant. With this assumption we have identified the aggregate output index. In particular, the growth rate of aggregate output is equal to the Divisia index growth rate $\hat{y}^D$ defined above. On the other hand, we have defined our problem away: There no longer is any investment-specific technical change.[16]

[14] As stated in the first section, Divisia indices have certain nice properties in terms of aggregation (they are revenue-weighted sectorial indices, and this property applies to output, input, and productivity indices), but this does not mean that they are in any sense the “true” aggregators for an economy.

[15] This is a slightly more general version of the model Greenwood et al. (1997) analyze (they use specific functional forms for H and F), but cast in a one-sector form.

[16] An interesting question is, Is this model potentially consistent with the falling relative price of investment? In principle the answer is “yes” because a falling relative price can be obtained

Since Greenwood et al. (1997) want to study the role of investment-specific technological change, they have to make other assumptions in order to identify z and q. They assume that the factor substitution properties in the two sectors of the economy are the same, that is, $F_c = F_x = F$. With this restriction one can show that $H(c, x/q) = c + x/q$, and that $1/q$ is the price of investment goods relative to consumption goods. Greenwood et al. (1997) can recover H by deflating nominal GDP with the consumption goods deflator; that is, they define aggregate output in terms of consumption goods, $y^{GHK} = [p_c c + p_x x]/p_c = c + x/q$. This is an unusual definition of aggregate output—it does not coincide with the Divisia measure—but it is justified within the confines of the model. In fact, it is rather natural given that consumption is the ultimate source of welfare in the model.

Suppose we next calculate aggregate TFP based on this definition of aggregate output; we would then obtain

$$\hat{z}^{GHK} = \hat{y}^{GHK} - \alpha\hat{k} - (1-\alpha)\hat{n} = \hat{z}.$$

Greenwood et al.'s (1997) definition of TFP growth indeed recovers exogenous productivity changes that are not contaminated with the endogenous response of the economy to these productivity changes. On the other hand, their definition of aggregate TFP actually recovers productivity in the consumption goods sector and not in the "aggregate economy." Given the assumptions they make, what does the aggregate TFP index based on the Divisia output index recover? Using the definition of $\hat{z}^D$, we can show that

$$\hat{z}^D = s_c\hat{z} + (1-s_c)\left(\hat{z} + \hat{q}\right).$$

That is, the Divisia-based residual is a revenue share-weighted aggregation of the sector-specific residuals. Moreover, the relative importance of $\hat{q}$ for $\hat{z}^D$ is measured by $(1-s_c)\hat{q}/(\left[\hat{z} + (1-s_c)\,\hat{q}\right]$. Note that the displayed Divisia measure of aggregate TFP growth mixes the exogenous sectoral productivity growth rates with the economy's endogenous response to these growth rates as reflected in the consumption and investment share.[17]

When Greenwood et al. (1997) implement their approach for the U.S. economy, they use Robert Gordon's (1990) quality adjustments to construct the quality-adjusted inverse relative price of new investment goods $1/q$. With the quality-adjusted investment series they construct the capital stock, and with Solowian growth accounting methods for their consumption-based output measure they construct a series for z. Their method implies a growth rate of investment-specific technology of around 3 percent per annum, with growth in neutral technology of around 1 percent per annum. Moreover, consistent

by appropriate assumptions on H. As the economy grows, the isoquants may significantly change shape so as to produce a continuous relative price change. Although this perspective seems rather academic, it is a logical possibility.

[17] In "Introducing Factor-Specific Productivity" we discussed why this might be undesirable.

with the hypothesis of an equipment-led technology revolution, the growth rate of the q series increased, already beginning in the mid-1970s, by about half a percentage point. This finding is also consistent with McHugh and Lane's (1987) finding, based on cross-section evidence, that adjacent vintages show significantly smaller productivity differences prior to the mid-1970s. Of course, neutral technology slowed down considerably at around the same time—the TFP version of the productivity slowdown.[18]

Building on this measurement, Greenwood et al. (1997) attribute, in terms of the model and as an entirely structural exercise, a substantial part of long-term output growth to investment-specific technological change (growth in q) rather than neutral technological change (growth in z).

How Technology Affects Skilled and Unskilled Labor

U.S. data over the last couple of decades reveal substantial changes in the returns to education, the skilled wage premium. Moreover, typical wage regressions show large increases in residual variance: wage variance that cannot be attributed to observed characteristics such as age, experience, education, race, or gender. Katz and Murphy (1992) and earlier observers speak of "skill-biased technical change" as the explanatory factor behind these wage developments. Alternative explanations have been proposed, such as a decrease in union activity and increased foreign competition for unskilled labor. Bound and Johnson (1992), however, conclude that the lion's share of the changes in relative wages reflect relative changes in factor-specific productivity. In a recent paper, Krusell, Ohanian, Ríos-Rull, and Violante (2000) provide a more structural explanation of the wage premium. They argue that it is not factor-specific technological progress that increases the relative productivity of skilled labor, but rather the rapid accumulation of equipment capital together with a skilled labor-capital complementarity that determines the wage premium.

Skill-Biased Technological Change

We will next discuss how factor-specific productivity measurements have been used to rationalize changes in the wage premium. Consider a production function with capital and the two types of labor of the following kind:

$$y = F\left[k, G\left(A_s n_s, A_u n_u\right)\right],$$

where G is a CES function with substitution elasticity parameter ν and we have abstracted from capital-specific productivity (alternatively, capital

[18] A very similar exercise is conducted in Gort, Greenwood, and Rupert (1999), where quality-adjusted data are used not only for equipment but also for structures, thus identifying both a q_e and a q_s series.

embodies technological change and k is measured in efficiency units, as in the previous section on investment-specific technological change). As above, we assume that F and G are time-invariant, so that any technological change comes through (A_s, A_u). When the two types of labor are paid their marginal products, the production structure implies that

$$\log \frac{w_s}{w_u} = \frac{\nu - 1}{\nu} \cdot \log \frac{A_s}{A_u} - \frac{1}{\nu} \cdot \log \frac{n_s}{n_u}.$$

Katz and Murphy (1992) use this structure to interpret their leading finding that in a regression of (the log of) relative wages on (the log of) relative factor supplies and a time trend, using 1963–1987 aggregate U.S. annual data, one obtains:

$$\log \frac{w_s}{w_u} = 0.033 \cdot t - 0.71 \cdot \log \frac{n_s}{n_u}.$$

Katz and Murphy (1992) conclude from this wage regression that (1) the input elasticity ν is about $\sqrt{2}$; and (2) the productivity of skilled labor relative to that of unskilled labor increased on average by almost 12 percent per year over the period.[19] More interestingly for our purposes, the wage premium first rose during the 1960s, fell over the early 1970s, and finally rose sharply beginning in the late 1970s. The latter increase continued unabated through the end of the 1990s. What do these relative swings tell us about technology?

We first observe that a large range of different data sets (time series as well as cross-section) and methods also yield an input elasticity of $\sqrt{2}$. This suggests that assuming a stable production function with $\nu = \sqrt{2}$ is reasonable, and we can thus back out the entire sequence of factor-specific technology ratios using the same methodology as in the example with investment-specific technological change analyzed in Greenwood et al. (1997). For this data set, one observes that (1) the overall increase of the wage premium is due to a rise of the relative productivity of skilled labor A_s/A_u, (2) the fall of the wage premium in the early 1970s is due to the significant increase of the relative supply of skilled labor n_s/n_u, and (3) the relative productivity of skilled labor started to rise sharply in the late 1970s.

The approach just described allows us to recover factor-specific productivities, conditional on assumptions about the relative factor substitutabilities. In particular, it assumes that skilled and unskilled labor are equally substitutable with capital. Going back to a well-known paper by Griliches (1969), it has long been argued that most production technologies exhibit "capital-skill complementarity." That is, capital and skilled labor are more complementary than are capital and unskilled labor.

[19] In the wage regression, the coefficient on the log of relative factor supplies represents $1/\nu$, that is $\nu = 1/0.71$. Furthermore, if the relative productivity of skilled labor grows at the rate $(1+\gamma)$, $(A_s/A_u)_t = (1+\gamma)(A_s/A_u)_{t-1}$, then in the wage regression the coefficient on time represents $\nu - 1/\nu \log(1+\gamma)$, that is $\gamma \approx .033 \cdot \nu/\nu - 1$.

Capital-Skill Complementarity

Krusell et al. (2000) argue that the higher wage premium is actually due to capital accumulation since skilled labor is relatively more complementary with capital than is unskilled labor.[20] They capture the differential complementarity between capital and skilled and unskilled labor using the following nested CES production technology

$$y = F\left[A_u n_u, G\left(k, A_s n_s\right)\right],$$

where F and G are CES functions. This structure, unlike the one studied in the previous section, allows capital-skill complementarity. If the factor elasticity between capital and skilled labor is denoted μ, and that between unskilled labor and either skilled labor or capital is denoted ν, then we have capital-skill complementarity if $\mu < \nu$. Capital-skill complementarity means that the relative wage will change when the capital stock changes, even if labor inputs *and* labor-specific productivity levels do not change. Krusell et al. (2000) show that with an estimate of μ in line with the findings from the labor demand literature (see, for example, Hamermesh [1993]), a ν around $\sqrt{2}$, and a measure of quality-adjusted capital, the relative wage movements in the data can be quite closely tracked without *any* change in the relative labor productivity A_s/A_u.[21] When Krusell et al. (2000) relax the assumption of constant relative labor productivity, they find that the relative productivity of skilled labor grows at a modest 3 percent per year.

Notice how the results of Krusell et al. (2000) stand in sharp contrast to the conclusion based on Katz and Murphy's (1992) work. When relative wage changes are driven by changes in relative labor productivities alone, a different capital accumulation behavior would have no effect on wages. On the other hand, from the perspective of Krusell et al. (2000), there would have been no rapid increase in the skill premium in recent years had it not been for the faster growth rate of capital. In sum, it appears plausible that equipment-specific technological change, possibly accompanied by some additional, independent skill-biased technological change unrelated to equipment, lies behind the large movements in relative wages of the last 30 years. That is, relative wage data can be usefully employed to understand the nature and evolution of aggregate technology in the economy.

---

[20] Similar points, but in different theoretical structures, have been made in Greenwood and Yorukoglu (1997) and Caselli (1999).

[21] Krusell et al. (2000) emphasize the relative complementarities between equipment capital and skilled and unskilled labor. The quality-adjusted equipment capital stock is again based on the work of Gordon (1990) and subsequent updates, especially for IT technology.

Factor-Specific Productivity and Relative Factor Endowments

Our discussion of factor-specific productivity so far has assumed that it evolves exogenously. In a recent study Caselli and Coleman (2000) apply the methods of Krusell et al. (2000) to a cross section of countries. Some of their results seem to suggest that countries choose among a menu of skilled-unskilled labor productivities and that these choices depend on the countries' relative factor endowments.

Caselli and Coleman (2000) obtain measures of capital and skilled- and unskilled-labor input measures for a large cross section of countries. The authors then assume marginal-product pricing and estimate a nested CES technology, which is common in form across countries. Given the estimated CES parameters, they back out a set of factor-specific productivity levels, one for each country. Although the results are preliminary as of this moment, three interesting conclusions appear from our perspective. First, capital-skill complementarity receives support. Second, countries appear to have very different mixes in factor-specific productivity levels. In particular, there seems to be a negative correlation between A_s and A_u in the cross section: Countries with high skilled-labor productivities tend to have low unskilled-labor productivities and vice versa.[22] This correlation suggests the existence of a "productivity possibility frontier," but the results also indicate that the choices along the (A_s, A_u) frontier still leave much to be explained: many countries are significantly inside the frontier and thus are operating inefficiently. Third, countries with relatively more skilled labor tend to have relatively high skilled labor productivities and vice versa. This tendency suggests that a country's technology choice depends on its factor endowments, a point which has been made by Acemoglu (see, for example, Acemoglu [2000] and Acemoglu and Zilibotti [1998]).

3. A UNIFIED VIEW OF THE LAST QUARTER CENTURY?

What emerges from the sections on TFP and FSP measurement is a view of technological change in the United States that is based on major improvements in equipment production, with major effects on both aggregate and factor-specific productivities. Several questions arise, however, regarding the effects of TFP and FSP change. The first of these relates to the productivity slowdown. Among the most significant productivity movements over the entire century is the large slowdown in TFP starting around 1973. Can the asserted improvements in equipment-producing technologies be made consistent

[22] It is an open question whether these findings are robust to better measurement of capital; with appropriate quality adjustments, large differences can be observed in capital stocks, and the capital-skill complementarity hypothesis then has implications for the factor-specific productivity measurements.

with the productivity slowdown? Some recent papers advance the hypothesis that learning problems associated with the use of the new equipment may have been responsible for the aggregate slowdown.[23] However, empirical assessments of learning costs in implementing new technology are inherently difficult (see, for example, the recent arguments in Atkeson and Kehoe [2000] and Hornstein [1999]) and we are far from being able to reach a conclusion regarding this hypothesis. Nevertheless, the possibility remains an interesting one.

Second, changes in technology are unlikely to occur only in the United States; insights about the efficiency of different production methods, new blueprints, and capital travel relatively easily across borders. Does international data support the above productivity analysis? The European data tell a different story about labor markets than the U.S. data. Whereas unemployment stayed low in the United States, it increased dramatically in Europe, and European relative wages did not move nearly as much as U.S. relative wages. However, the unemployment response in Europe occurred concurrently with the relative wage response in the United States, so a common underlying explanation should not be ruled out.

In particular, it seems quite plausible that differences in labor market institutions, and one common shock, can yield quite different responses in two economies. Our hypothesis, thus, is that more heavily regulated and unionized labor markets can make skill-biased technological change lead to increases in the rate of unemployment instead of increases in the skill wage. This hypothesis has been explored theoretically in Ljungqvist and Sargent (1998), Marimon and Zilibotti (1999), and recently in Hornstein, Krusell, and Violante (2000).[24] As labor market theory with frictions—allowing a nontrivial role for unemployment—has not advanced as far quantitatively as has neoclassical theory, it is too early yet for a firm evaluation of this hypothesis. As for data on equipment relative prices, we do not know of European data comparable to the U.S. data by Gordon (1990) on the more recent revisions for some equipment categories in the NIPA. Improvements in equipment price measurement should, in our view, be placed high on the agenda in the United States and even higher in Europe.

A unified view of the macroeconomic productivity and labor market performances during the last quarter century in the western world is a very interesting one that ought to be explored in much further detail. It can be viewed as a "third industrial revolution," placing advanced equipment and IT on center stage. We believe that careful reassessments of input and output measurements, together with theory developments aimed at structural evaluation of the main hypothesis, would be most productive.

---

[23] See, for example, Hornstein and Krusell (1996) and Greenwood and Yorukoglu (1997).

[24] Related work, focusing more on wages than on unemployment, is found in Violante (1999) and Aghion, Howitt, and Violante (1999).

REFERENCES

Acemoglu, Daron. "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature* (forthcoming).

______________, and Fabrizio Zilibotti. "Productivity Differences," *Quarterly Journal of Economics* (forthcoming).

Aghion, Philippe, Peter Howitt, and Giovanni Violante. "Technology, Knowledge, and Inequality." Working Paper. University of California, Berkeley, 1999.

Atkeson, Andrew and Patrick J. Kehoe. "The Transition to a New Economy Following the Second Industrial Revolution," Federal Reserve Bank of Minneapolis Working Paper 606, September 2000.

Basu, Susanto, and John G. Fernald. "Why Is Productivity Procyclical? Why Do We Care?" International Finance Discussion Working Paper 638. Board of Governors of the Federal Reserve System, June 1999.

Bils, Mark, and Peter J. Klenow. "Quantifying Quality Growth." Mimeo, 1999.

Bound, John, and George Johnson. "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations," *American Economic Review,* vol. 82 (June 1992), pp. 371–92.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence," *Quarterly Journal of Economics* (forthcoming).

Brynjolfsson, Erik, and Lorin M. Hitt. "Computing Productivity: Firm-Level Evidence," Working Paper, 2000b.

______________. "Beyond Computation: Information Technology, Organizational Transformation and Business Performance," *Journal of Economic Perspectives,* vol. 14 (Fall 2000a), pp. 23–48.

Brynjolfsson, Erik, and Shunkyu Yang. "The Intangible Costs and Benefits of Computer Investments: Evidence from the Financial Markets." Proceedings of the International Conference on Information Systems, Atlanta, Georgia, December 1997 (revised April 1999).

Caselli, Francesco. "Technological Revolutions," *American Economic Review,* vol. 89 (March 1999), pp. 78–102.

______________, and John Coleman. "The World Technology Frontier," Working Paper, University of Chicago, 2000.

Cole, Roseanne, Y. C. Chen, Joan A. Barquin-Stelleman, Ellen Dulberger, Nurhan Helvacian, and James H. Hodge. "Quality-Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment," *Survey of Current Business,* vol. 66 (January 1986), pp. 41–50.

Cordoba, Juan Carlos. "Transient Growth and Structural Change." Mimeo, University of Rochester, 2000.

Gordon, Robert J. *The Measurement of Durable Goods Prices*. Chicago, IL: University of Chicago Press, 1990.

Gort, Michael, Jeremy Greenwood, and Peter Rupert. "Measuring the Rate of Technological Progress in Structures," *Review of Economic Dynamics*, vol. 2 (January 1999), pp. 207–30.

Greenwood, Jeremy, and Mehmet Yorukoglu. *Carnegie-Rochester Conference on Public Policy* XLVI, 1997 [1974], pp. 49–95.

__________, Zvi Hercowitz, and Per Krusell. "Long-run Implications of Investment-Specific Technological Change," *American Economic Review,* vol. 87 (June 1997), pp. 342–62.

Griliches, Zvi. "Capital-Skill Complementarity," *Review of Economics and Statistics,* vol. 51 (November 1969), pp. 465–68.

__________. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change," in *The Price Statistics of the Federal Government*. New York: National Bureau of Economic Research, 1961, pp. 137–96.

Hall, Robert E. "e-Capital: The Link between the Stock Market and the Labor Market in the 1990s," Working Paper, 2000.

__________. "The Stock Market and Capital Accumulation," NBER Working Paper 7180, 1999.

__________. "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy,* vol. 96 (Oct. 1988), pp. 921–47.

__________. "The Specification of Technology with Several Kinds of Output," *The Journal of Political Economy,* vol. 81 (July-August 1973), pp. 878–92.

Hamermesh, Daniel S. *Labor Demand*. Princeton, NJ: Princeton University Press, 1993.

Hercowitz, Zvi. "The 'Embodiment' Controversy: A Review Essay," *Journal of Monetary Economics,* vol. 41 (February 1998), pp. 217–24.

Hornstein, Andreas. "Growth Accounting with Technological Revolutions," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 85 (Summer 1999), pp. 1–22.

______________, and Per Krusell. "Can Technology Improvements Cause Productivity Slowdowns?" in Ben S. Bernanke and Julio J. Rotemberg, eds., *Macroeconomics Annual 1996*. Cambridge, Mass.: MIT Press, 1996, pp. 209–59.

______________, and Giovanni Violante. "Technology, Policy, and the Inequality-Unemployment Tradeoff." Mimeo, 2000.

Hulten, Charles R. "Divisia Index Numbers," *Econometrica,* vol. 41 (November 1973), pp. 1017–26.

Jorgenson, Dale W. "The Embodiment Hypothesis," *Journal of Political Economy,* vol. 74 (February 1966), pp. 1–17.

______________, Frank Gollop, and Barbara Fraumeni. *Productivity and U.S. Economic Growth*. Cambridge, Mass.: Harvard University Press, 1987.

Jorgenson, Dale W., and Kevin J. Stiroh. "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity* (forthcoming).

Katz, Lawrence, and Kevin Murphy. "Changes in Relative Wages: Supply and Demand Factors," *Quarterly Journal of Economics*, vol. 106 (February 1992), pp. 35–78.

Kiley, Michael T. "Computers and Growth with Costs of Adjustment: Will the Future Look Like the Past?" Finance and Economics Discussion Working Paper 36. Board of Governors of the Federal Reserve System, July 1999.

Krusell, Per, Lee Ohanian, José-Víctor Ríos-Rull, and Giovanni Violante. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica,* vol. 68 (September 2000), pp. 1029–53.

Ljungqvist, Lars, and Thomas J. Sargent. "The European Unemployment Dilemma," *Journal of Political Economy,* vol. 106 (June 1998), pp. 514–50.

Marimon, Ramon, and Fabrizio Zilibotti. "Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits," *Economic Journal,* vol. 109 (April 1999), pp. 266–91.

McHugh, Richard, and Julia Lane. "The Role of Embodied Technical Change in the Decline of Labor Productivity," *Southern Economic Journal,* vol. 53 (April 1987), pp. 915–24.

Moulton, Brent R. "Measurement of Banking Services in the U.S. National Income and Product Accounts: Recent Changes and Outstanding Issues." Presented at a meeting of the BEA Advisory Committee, May 5, 2000.

———, Robert P. Parker, and Eugene Seskin. “A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts: Definitional and Classificational Changes,” *Survey of Current Business,* vol. 79 (October 1999), pp. 7–20.

Oliner, Stephen D., and Daniel E. Sichel. “The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?” Finance and Economics Discussion Working Paper 20. Board of Governors of the Federal Reserve System, March 2000.

Schmitz, James A., Jr. “Reorganizing Work to Increase Productivity: The Dramatic Recovery of the United States and Canadian Iron-ore Industries.” Mimeo, Federal Reserve Bank of Minneapolis, 1998.

Seskin, Eugene P. “Improved Estimates of the National Income and Product Accounts for 1959–98: Results of the Comprehensive Revision,” *Survey of Current Business,* vol. 79 (August 1999), pp. 15–43.

Shapiro, Irving, Matthew D. Shapiro, and David W. Wilcox. “Quality Improvement in Health Care: A Framework for Price and Output Measurement,” *American Economic Review,* vol. 89 (May 1999), pp. 333–37.

Solow, Robert M. “Investment and Technological Progress,” in Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes, eds., *Mathematical Methods in the Social Sciences 1959*. Stanford, CA: Stanford University Press, 1960, pp. 89–104.

———. “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics,* vol. 39 (August 1957), pp. 312–20.

Violante, Giovanni. “Technological Acceleration, Skill Transferability and the Rise in Residual Inequality.” Mimeo, University College London, 1999.

Whelan, Karl. “Computers, Obsolescence, and Productivity,” Finance and Economics Discussion Working Paper 6. Board of Governors of the Federal Reserve System, January 2000.