Federal Reserve Bank
of Chicago

First Quarter 1998

Economic perspectives

Economic perspectives

Economic Perspectives is published by the Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the management of the Federal Reserve Bank.

Single-copy subscriptions are available free of charge. Please send requests for single- and multiple-copy subscriptions, back issues, and address changes to the Public Information Center, Federal Reserve Bank of Chicago, P.O. Box 834, Chicago, Illinois 60690-0834, telephone 312-322-5111 or fax 312-322-5515.

Economic Perspectives and other Bank publications are available on the World Wide Web at http://www.frbchi.org.

ISSN 0164-0682

Contents

First Quarter 1998, Volume XXII, Issue 1

Lessons from the history of money

François R. Velde

Introduction and summary

The use of money began in the sixth century B.C. in what is now western Turkey, when lumps of gold found in rivers were melted and turned into pieces of uniform size imprinted with a stamp. For almost all of the time since then, the common monetary system has been *commodity money*, whereby a valuable commodity (typically a metal) is used as a widely accepted medium of exchange. Furthermore, the quantity of money was not under anyone's control; private agents, following price incentives, took actions that determined the money supply.

Today, the prevalent monetary system is that of *fiat money*, in which the medium of exchange consists of unbacked government liabilities, which are claims to nothing at all. Moreover, governments have usually established a monopoly on the provision of fiat money, and control, or potentially control, its quantity. Fiat money is a very recent development in monetary history; it has only been in use for a few decades at most.

Why did this evolution from commodity money to fiat money take place? Is fiat money better suited to the modern economy or was it desirable but impractical in earlier times? Were there forces that naturally and inevitably led to the present system?

Fiat money did not appear spontaneously, since government plays a central role in the management of fiat currency. How did governments learn about the possibility and desirability of a fiat currency? Did monetary theorizing play any role in this evolution?

In this article, I will argue that the evolution from commodity to fiat money was the result of a long process of evolution and learning. Commodity money systems have certain advantages, in particular in providing a natural anchor for the price level. But they also have certain disadvantages, manifested in particular in the difficulty of providing multiple denominations concurrently. These problems arose early on, in the fourteenth century, in the form of money shortages. Societies tried to overcome these disadvantages, and this led them progressively closer to fiat money, not only in terms of the actual value of the object used as currency, but also in terms of the theoretical understanding of what fiat money is and how to manage it properly.

In the process, societies came to envisage the use of coins that were worth less than their market value to replace the smaller denominations that were often in short supply. These coins are very similar to bank notes; they are printed on base metal, rather than paper, but the economics behind their value is the same. What governments learned over time about the provision of small change is thus directly applicable to our modern system of currency.

In his *A Program for Monetary Stability* (1960), Milton Friedman begins with the question: Why should government intervene in monetary and banking questions? He answers by providing a quick history of money, which he describes as a process inevitably leading to a system of fiat money monopolized by the government (p. 8):

> These, then, are the features of money that justify government intervention: the resource cost of a pure commodity currency and hence its

François R. Velde is an economist at the Federal Reserve Bank of Chicago. This article draws on joint work with Thomas J. Sargent, University of Chicago and Hoover Institution.

> tendency to become partly fiduciary; the peculiar difficulty of enforcing contracts involving promises to pay that serve as medium of exchange and of preventing fraud in respect to them; the technical monopoly character of a pure fiduciary currency which makes essential the setting of some external limit on its amount; and finally, the pervasive character of money which means that the issuance of money has important effect on parties other than those directly involved and gives special importance to the preceding features. ... The central tasks for government are also clear: to set an external limit to the amount of money and to prevent counterfeiting, broadly conceived.

This article will find much to validate this view. It turns out that the problem of counterfeiting, identified as central by Friedman, provided obstacles that were overcome only when the appropriate technology became available. As technology changed and offered the possibility of implementing a form of fiduciary currency, various incomplete forms of currency systems were tried, with significant effects on the price level. These experiments led to the recognition that quantity limitation was crucial to maintaining the value of the currency. The need for a government monopoly, however, does not emerge from our reading of the historical record, and we will see that the private sector also came up with its own solutions to the problem of small change, thereby presenting alternatives to the monetary arrangements we have adopted.[1]

Commodity money and price stability

Among the desirable features of a monetary system, *price stability* has long been a priority, as far back as Aristotle's discussion of money in *Ethics*. In the words of the seventeenth century Italian monetary theorist Gasparo Antonio Tesauro (1609), money must be "the measure of all things" (*rerum omnium mensura*) (p. 633). Aristotle also noted that commodity money, specifically money made of precious metals, was well suited to reach that goal: "Money, it is true, is liable to the same fluctuation of demand as other commodities, for its purchasing power varies at different times; but it tends to be comparatively constant" (Aristotle, *Ethics*, 1943 translation).

The commodity money system delivers a nominal anchor for the price level. The mechanism by which this takes place can be described in the context of a profit-maximizing mint, which was how coins were produced in the Middle Ages and later.[2] Suppose there is a way to convert goods into silver and silver into goods at a constant cost (in ounces of silver per unit of goods), which can be thought of as either the extraction cost of silver and the industrial uses of the metal or the "world price" of silver in a small country interpretation. Silver is turned into coins by the mint; the mint (which really represents the private sector) also decides when to melt down existing coins.

The government's role is limited to two actions. It specifies how much silver goes into a coin, and it collects a seigniorage tax[3] on all new minting.

When the mint is minting new coins, its costs are the cost of the silver content, the seigniorage tax, and the production cost;[4] its revenues are the market value of the coins, which is the inverse of the price level. Similarly, when the mint is melting down coins, its costs are the market value of the coins, and its revenues are the value of the silver contained in them.

Whether the mint will produce new coins or melt down existing coins will thus depend on how the price level relates to the parameters: silver content of the coins, production costs, and seigniorage rate. The price level cannot be too low (or the purchasing power of the coins too high) or the mint could make unbounded profits by minting new coins and spending them. Similarly, the price level cannot be too high (or the purchasing power of the coins too low), or the mint would make profits by melting down the coins. The absence of arbitrage for the mint places restrictions on the price level, which is contained in an interval determined by the minting point and the melting point (figure 1).

This system, which prevailed until the late nineteenth century, has some noteworthy features. The quantity of money is not controlled directly by the government; rather, additions to or subtractions from the money stock are made by the private sector, on the basis of incentives given by the price level. The incentives operate so as to make the system self-regulating. If coins become too scarce, their value increases and the price level falls until it reaches the minting point, when more coins are added to the stock. If coins

FIGURE 1

Constraints on the price level caused by arbitrage

become too numerous, on the other hand, their market value reaches their intrinsic value and it becomes worthwhile for the mint to melt them down. The commodity nature of the currency places bounds on the price level, but does not determine the price level within that interval.

Within the interval, the price level depends on how the quantity of money relates to the volume of transactions, according to Irving Fisher's famous quantity theory equation.[5] As long as the price level is inside the interval, the stock of coins, or quantity of money, is fixed. Variations in the volume of transactions or in income would shift the price level up or down, unless such variations were so severe as to push the price level up to the melting point or down to the minting point. In that case, the mint would enter into action and modify the quantity of money in the appropriate way.

Consider now the interval in figure l. Its position on the real line is determined by the world price of silver and the silver content of a dollar coin. Any reduction in the number of ounces of silver per dollar, that is, any debasement of the currency, shifts the interval to the right; the price level is therefore higher. But the width of the interval is determined by production costs and the seigniorage tax. We may take production costs as a technological given, but the seigniorage tax is chosen by the government. In principle, the government could make the tax a subsidy; it could even subsidize the production costs completely. In that case, the interval in figure 1 would be reduced to a point, the minting point and melting point would coincide, and the price level would be completely tied to the world price of silver. This would eliminate any fluctuations in the price level due to the quantity theoretic effects described above. The only variations would be due to fluctuations in the world price of silver. In western European practice, however, the seigniorage rate was positive in almost all countries.

Although governments considered minting a fiscal prerogative, they were constrained in their choice of the seigniorage rate. High rates, a form of monopoly rent, were possible only if the government could effectively prevent competition. But in medieval Europe, all manner of coins circulated in all places and individuals were quite willing to take their metal to the mint of a nearby lord or king, subject to transportation costs, if they found the local seigniorage rate too high. Also, the technology for making coins was rather crude and available to any jeweler or goldsmith, so that counterfeiters would also be tempted by high seigniorage rates. In practice, then, the width of the interval was rather small, and production costs with seigniorage were on the order of 1 percent to 2 percent for gold and 5 percent to 10 percent for silver (the latter being ten times less valuable, transport costs were higher).

Multiple denominations and token coinage

This simple commodity system lacks one feature: multiple denominations. Although it is always possible to express any price in pennies, in practice it is necessary to have a range of coins of various denominations.[6]

In its last incarnation (the so-called classical gold standard), the commodity money system handled multiple denominations in a straightforward way, which is described in textbooks, for example, John Stuart Mill (1857).

The standard formula

The method that Cipolla (1956) calls the standard formula, consists of choosing a principal (large) denomination, which continues to be provided as before at the initiative of the private sector, thus continuing to provide a nominal anchor for the price level. The provision of lower or subsidiary denominations relies on three key elements: 1) monopolization of coinage by the government, 2) issue of *token* coins, and 3) peg of the token coins by having the government convert them on demand into the larger denominations. The intrinsic content of token coins was somewhat or much smaller than the face value at which they circulated. Some authors call such coins partly fiduciary. The opposite of a token coin is a *full-bodied coin*.[7]

In the case of the gold standard, the larger denominations were gold coins, and currencies (the U.S. dollar, the British pound, and the French franc) were defined by the number of ounces of gold per currency unit. The subsidiary coinage consisted of silver and bronze coins, which were token. The government's willingness to peg, say, the silver quarter at 1 / 40 of a gold eagle was implemented by the U.S. Treasury.[8]

Thus, in the standard formula, tokens play the same role as convertible notes issued by the central bank. As with notes, a mechanism serves to regulate the quantity outstanding: Excess quantities of token quarters are turned in at the treasury in exchange for gold eagles, while needed tokens are sold by the mint.[9]

The advantages of a token coinage are the same as the advantages of a representative money system, as pointed out by a long line of writers, including Adam Smith, John M. Keynes, and Milton Friedman. Resources that had been spent forming and maintaining that part of the stock of metallic currency were freed up for other purposes. To quote the French monetary official Henri Poullain, writing in 1612: "In a card game, where various individuals play, one avails oneself of tokens, to which a certain value is assigned, and they are used by the winners to receive, and by the losers to pay what they owe. Whether instead of coins one were to use dried beans and give them the same value, the game would be no less enjoyable or perfect" (Poullain, 1709, p. 68).

Another advantage, from the point of view of the government, is that the issue of tokens is quite profitable. To the extent that tokens circulate for more than their intrinsic value plus the costs of minting, they represent a pure profit, the seigniorage in the medieval and modern sense of the word.

These two advantages (social savings and government revenues) have been understood for centuries, and, as Friedman points out, have provided impetus for the development of money away from a strict, full-bodied commodity version. However, these two motivations do not determine clearly in which direction money will develop; perhaps, in fact, each pushes in a different direction. The tension will be illustrated in the historical process I describe.

Prerequisites of a token coinage

Whatever its advantages, the implementation of the standard formula depended on some prerequisites. With a token coinage, the profits to the issuer are large, and, as Friedman says (1960, p. 6), "In fraud as in other activities, opportunities for profit are not likely to go unexploited." The government's ability to maintain its monopoly on token issue is thus dependent on the prevention of counterfeiting.[10] While nowadays counterfeiting may seem to be a significant but not overwhelming nuisance, which suitable technology can always remedy (such as that embodied in the recently issued $100 and $50 bills), in the past it presented an insuperable obstacle to the development of the standard formula.

One way to prevent counterfeiting is to impose high costs of entry to counterfeiters. Law enforcement provides a second method; as the Italian economist Montanari wrote in 1683, "A die which costs the prince 3 to make, will cost a counterfeiter 8 or 12; because he who works at the mint does not risk his life, and receives only the wage commensurate to his activity; but if a goldsmith has to make a coin at the risk of his whole being, he will not be persuaded if not with a lot of gold." The death penalty[11] for counterfeiters adds a risk premium to the counterfeiters' wage costs, which may or may not be sufficient to wipe out their potential profits. A third method is to make the government currency difficult to imitate, for example, if it is produced with a technology that is not accessible to the private sector in some way; either the government can make better coins or the same coins more cheaply.

If such a cost or technology advantage is not available to the government, then attempts at issuing token coinage will be plagued by counterfeiting or competition from neighboring currencies. Ultimately, the gross seigniorage rate will be driven down to the production costs (common to both government and counterfeiters). Thus, without the appropriate technology, only full-bodied coins can be used for small denominations.

The big problem of small change

This seemingly trifling aspect of the monetary system turns out to have bedeviled Western societies for centuries. Nowadays, the only problem most people see with small change is that we have too many pennies around, but for students of monetary history, the "big problem of small change" (a phrase coined by Carlo Cipolla in 1956) refers to recurrent coin shortages that were prevalent before the adoption of token coinage. The last time the U.S. experienced a shortage of small change was in 1965–66, when quarters and dimes still contained silver; the Coinage Act of 1965 made them completely token (Spengler, 1966).

Full-bodied small change

The medieval technology for making coins was very simple. Metal was melted and beaten into sheets, the sheets were cut with shears into blanks, and the blanks were placed between two hand-held dies. The upper die was struck with a hammer and the blank imprinted. Dies were made by goldsmiths using ordinary tools, and the design on coins could easily be copied by any goldsmith. Thus, the government and the private sector had access to the same technology.

Around 800 A.D., Charlemagne unified most of Western Europe and created a uniform currency. Until the twelfth century, Europe only had one coin, the silver penny, initially minted identically across Charlemagne's empire. Thus the commodity money system was in its simple, one-coin form. Around the year 1200, large improvements in the European economy, improved safety, and economic expansion led to greater volumes of trade and the need for larger denominations than the penny. This led to the appearance of silver coins of about five to ten times the content of a penny, called *grossi*. Over time, the denomination structure became richer, with the addition of gold coins in the mid-thirteenth century. Coins throughout the denomination structure remained close to full-bodied.

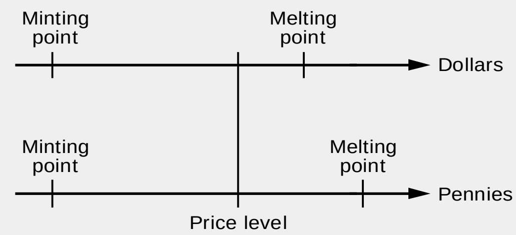
However, the commodity money system acquires unexpected complications when multiple denominations are introduced. To see this, let us return to the mint's problem, and suppose we have two currencies, dollars and pennies. The same reasoning as before will apply to both coins separately. As a result, the requirement that there be no arbitrage left for the mint will now place two sets of restrictions on the price level, which we can represent by two intervals, as in figure 2.

In order to make the two intervals comparable, the lower one (which corresponds to pennies) is scaled by the market exchange rate between the two coins (expressed in dollars per pennies). This simply means that the mint's calculations about minting or melting pennies are computed in dollars.

The intervals must overlap, of course. Recall that the position of a coin's interval on the real line is linked to the intrinsic content of that coin, so that a smaller intrinsic content of the dollar corresponds to a higher price level. With two coins, the ratio of intrinsic contents must be reasonably close to the intended parity between denominations, although it need not coincide with that parity. But that is not enough: A coin is produced only when the price level reaches the minting point. Therefore, if the lower ends

FIGURE 2

Constraints on the price level of two coins



of the intervals do not coincide, one type of coin is never minted. Equating the lower ends of the intervals (by the government's choice of the intrinsic contents and the seigniorage rates) makes the mint stand ready to buy silver for the same price, whether it pays in pennies or dollars.

On the other hand, if the upper ends of the intervals do not coincide, one coin might be melted, but the price level could still rise further and the other coin remain in use. Equating the upper ends of the intervals makes the ratio of metal contents in the two coins equal the exchange rate, in which case pennies are strictly full-bodied. If the melting point for pennies is higher than the melting point for dollars, pennies are relatively light.

Thus, if pennies are not full-bodied, a sufficient rise in the price level will make large coins disappear. If the mint prices differ, a sufficient fall in prices will prompt minting of only one of the two coins. The perpetual coexistence of both coins in the face of price fluctuations requires that pennies be full-bodied and that equal mint prices prevail for both coins; that is, the intervals must coincide and the sum of the seigniorage rates and the production cost must be equal for the two coins.

The state of the technology creates yet another difficulty. We have seen that government had little freedom to choose the seigniorage rate: It had to be positive and could not be large. But making small coins was much more expensive than making large coins, because making a small coin or a large coin involves essentially the same process, independent of the size or content of the coin. In the extreme, if it costs the same to make a penny or a dollar, then the production costs for 100 pennies is 100 times the cost per dollar for the same value of output (the coins). Historical data shows that the cost of making a coin fell with the denomination, but not fast enough. Figure 3 plots the production costs as a function of coin size for various European countries.

This technological constraint presented the mints with a dilemma: provide only full-bodied coins and see pennies never minted, or offer the same price for bullion in pennies or in dollars and face the risk of seeing the price level increase and large coins disappear. Thus, the commodity money system with full-bodied denominations has the potential for either shortages or gluts of small change.

In fact, shortages of small change were a common complaint, running through centuries of monetary history all over Europe and also (in the early nineteenth century) in the U.S. The above argument, although limited to the supply side, shows how vulnerable the commodity money system was to such shortages, given the technology available. An analysis of the demand side reveals even more trouble.

If we think of pennies and dollars as required for consumption purchases (a feature called a cash-in-advance constraint), but we assume that large coins cannot be used in small transactions, whereas small coins can be used in large transactions, it emerges that, within the overlapping intervals of figure 2, there is a certain indeterminacy of the exchange rate between dollars and pennies or the ratio at which pennies enter into the total money stock $M = M_1 + eM_2$ (where M is the total stock in dollars, M_1 is the number of dollar coins, e is the market exchange rate, and M_2 is the number of pennies). As long as there are enough pennies to carry out small transactions (not just in the physical sense M_2 but in terms of their total value eM_2), there can be more or fewer pennies or they can be worth more or less. If, for some reason, the relative share of small transactions changes and more pennies are needed, more pennies will be provided only if the minting points are lined up correctly *and* the price level falls enough. But for the general price level to fall, the shock must affect the volume of *all* transactions, and it is not hard to imagine situations where the existing stock of pennies is insufficient, yet no new pennies are minted.

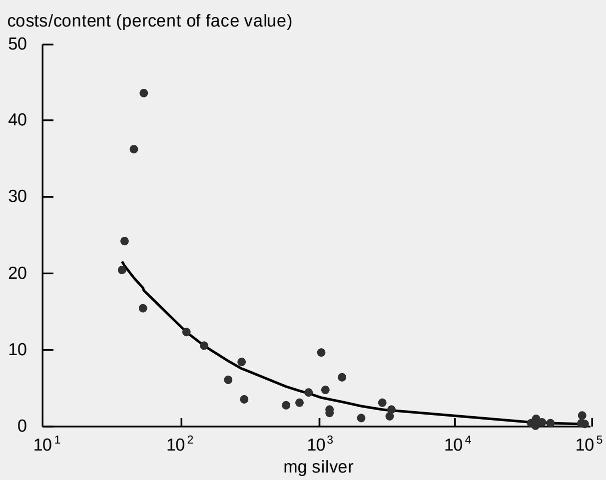
These shortages of small change have a curious feature: In a decentralized economy, agents choose how many pennies to hold. In order for them to hold too few pennies, there needs to be a price incentive for them to economize on pennies. This occurs through a rate of return dominance, that is, the return on holding pennies is lower than the return on holding dollars. In other words, the market exchange rate, e (in dollars per pennies), falls. But this means that the share of pennies in the total stock of coins shrinks further, accentuating the shortage of small change. Furthermore, a fall in the exchange rate shifts the lower interval of figure 2 to the left, making it likelier that the price level will hit the *upper* bound of the interval for pennies, the melting point.

Thus, shortages of small change push the economy in a vicious cycle, by making the shortage even more severe through a depreciation of the smaller denominations, and ultimately bringing about a melting down of pennies, once they have depreciated to the value of their intrinsic content.

Within the confines of the available technology, one partial remedy is for the government to counteract the leftward shift of the interval due to the fall in e by reducing the intrinsic content of pennies, which shifts the interval to the right. Figure 4 plots the evolution of the mint equivalent (the inverse of the intrinsic content) for two medieval Florentine silver coins, the *picciolo* (a penny, or 1d) and the *grosso* (worth 4d), during the

FIGURE 3

Production costs of coins in late medieval Europe, 1350–1500

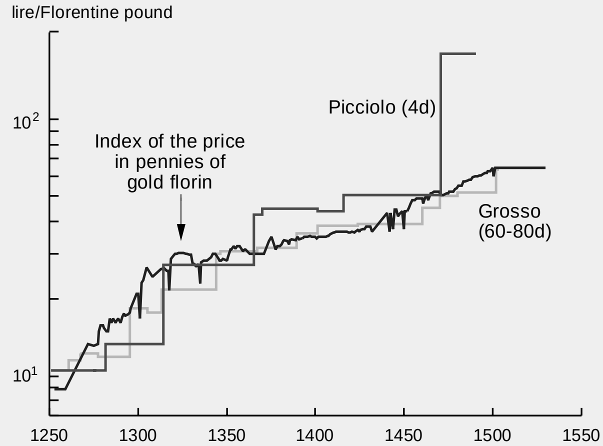


Note: Production costs were calculated by the percent of the face value of the coin as a function of the coin size in milligrams of silver.

Source: T. Sargent and F. Velde, 1997, "The evolution of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-13.

FIGURE 4

The intrinsic content of two Florentine coins, 1250–1530



Notes: The data have been plotted inversely. The coins, picciolo and grosso, are measured in Italian denaro (d), the equivalent of a penny.

Source: T. Sargent and F. Velde, 1997, "The big problem of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-8.

Middle Ages (the gold florin's intrinsic content remained constant). A pattern of recurrent debasements is apparent. The graph also displays the price of the gold florin in terms of silver pennies. This corresponds to the exchange rate of pennies per dollar or $1/e$ (the florin ranged from 240d in 1250 to 1,680d in 1530). One way to interpret this graph is that the periodic debasements, evident as upward steps, occurred to remedy the upward drift in the price of the florin, as our model predicts.

This version of the model takes the price of silver in terms of real resources as constant. In fact, this cost could be taken as variable over time, embodying a variety of shocks (changes in the technology to mine silver, including new discoveries, and changes in the demand for silver in industrial activities). Furthermore, the model assumes that large and small coins are made of the same metal; but small and medium coins being made of silver and large coins being made of gold, the intervals of figure 2 shift around due to changes in the relative price of gold and silver. Depending on the width of the intervals, small shocks might be accommodated, but larger variations lead to the same problems outlined above, unless e is allowed to change. The difficulties in providing multiple denominations render bimetallism (the simultaneous use of two metals in legal tender currencies with a fixed exchange rate) a fragile system.

The evolution of monetary doctrine

These shortcomings of the commodity money system were a result of the state of minting technology until 1550 or so. Moving toward the standard formula, or toward fiduciary coinage, required a better technology. However, the technology would have gone unexploited had monetary doctrine not weakened its attachment to the concept of full-bodied coinage. This evolution of monetary doctrine can be traced in the writings of medieval jurists.[12] This doctrine arose from their efforts to understand observed price patterns and devise ways to deal with the legal consequences for private contracts (the problem of the standard of deferred payments).

Because medieval Europe had begun with the penny and later added larger coins, the tradition was that prices were denominated in pennies, dozens of pennies (shillings), and scores of shillings (pounds).[13] Many nominal debts and contracts were thus expressed in pounds of the small coin, whose constant debasement led to the long-term inflation that is apparent in figure 4.

When the penny was the only coin, monetary doctrine was straightforward. In modern terms, it applied standard price theory to money, treating it as a commodity like any other. When a loan of 100 pennies came due, 100 pennies were owed, irrespective of any fluctuations in the purchasing power of pennies. The Neapolitan jurist Andrea d'Isernia (1220–1316) wrote: "If I lend you a measure of wheat in May when it is expensive and is worth perhaps 3 tarini, and I reclaim it in July after the harvest when it is worth perhaps 1 tarino, it is enough to return the measure of the same wheat in kind, even though it is worth less; likewise if it is worth more, for example if I lent it in July and demanded it in the following May ... the same reasoning applies for money as it does for wheat and wine" (d'Isernia, 1541).

From Charlemagne's reform around 800 A.D. (which restored a uniform currency in Western Europe) to the twelfth century, the penny changed content at various rates, through the action of wear and tear and debasements. Such changes in the intrinsic content of a penny were also treated by jurists in a similar way. The jurist Azo (d. 1220) formulated a simple rule: "The same money or measure is owed that existed at the time of the contract" (in Stampe, 1928, p. 36).

With the appearance of larger denomination coins and the existence of time-varying rates of exchange between denominations, the legal problems grew more challenging, and jurists began to diverge in their answers. A distinction was made between the "intrinsic quality" of a coin (its metal content) and the "extrinsic quality," taken to mean either its purchasing power (the inverse of the price level) or its rate of exchange with other coins. The general consensus prevailing in the fourteenth and fifteenth centuries called for adjusting debt repayments for variations in the intrinsic quality, but ignoring variations in extrinsic quality; and small coins were considered legal tender to the degree that they were full-bodied and interchangeable with large coins.

However, jurists also observed the existence of positive seigniorage rates (the width of the interval in figure 1), and realized that money's purchasing power could be greater than its intrinsic value. In other words, they discovered that the price level could move above the minting point. One strand of the legal literature insisted that seigniorage should be set close or equal to 0. Others, who argued that precious metals as bullion and in the form of coins should afford the same utility, recommended that the state subsidize the mint completely (in particular, the jurist Bartolo da Sassoferrato, 1313–1357). As jurists, they tried to define rules for repayment of monetary debts. They correctly perceived that their proposal would eliminate some fluctuations in the standard of value.

In practice, the jurists realized that governments were unwilling to subsidize mints and were tempted to increase seigniorage revenues as much as they could. A small tax was considered acceptable and a larger tax under very specific circumstances, such as a fiscal emergency (paying for a sudden war or the king's ransom). Some even argued that, in the words of Gabriel Biel (d. 1495) a large seigniorage rate "is the easier way to collect quickly the required funds without fraud and undue exactions from the subjects. It is, moreover, felt less and for this reason more easily borne without protest and without the danger of a rebellion on the part of the people. It is the most general form of taxation embracing all classes, clergy, laity, nobility, plebeians, rich and poor alike" (Biel, 1930 translation, p. 35).

Some jurists like d'Isernia even went further. D'Isernia probably observed episodes such as the siege of Faenza in 1241, when the Emperor Frederic II ran out of money and paid his troops with leather money that he redeemed into gold after the successful conclusion of the siege. D'Isernia argued that, under the specific circumstances already identified by the current doctrine, money could be made of worthless material, like lead or leather, as long as it was redeemed after the end of the emergency into good money. This was the basis for the concept of deficit financing, which would play an important role in the development of fiat money. By the late sixteenth century, these notions were commonly held. The widely cited René Budel (1591) held it "to be indubitable that a Prince in the midst of costly wars, and therefore in great necessity, can order that money be made out of leather, bark, salt, or any material he wants, if he is careful to repair the loss inflicted thereby on the community with good and better money" (Budel, 1591, chapter 1, paragraph 31).

In other words, the intrinsic content could be set to 0, as long as some measure of convertibility, either immediate or in the near future, was implied. In 1481, a small town in Catalonia carried out an experiment to solve its problem of small change: it was authorized by the king of Aragon to issue pure copper coins,[14] whose intrinsic value was about 25 percent of their face value, as long as "the city be known to pledge, and effectively pledge to receive said small money from those who might hold it, and to convert it and return for it good money of gold or silver, whenever and however much they be asked" (in Botet y Sisó, 1911, p. 328). This experiment was imitated by a number of other Catalonian cities, although they were plagued by counterfeiting, which the state of technology made relatively easy.

Technological change and policy experiments

These developments in monetary doctrine, and the early Catalonia experiment, show that technology remained the real barrier to the implementation of a standard formula for small change. The technology did change, in two major waves; and each wave opened up new possibilities that governments exploited.

Recall that the standard formula incorporates several ingredients: monopolization of coinage, issue of tokens, and convertibility of the tokens. The ingredients are logically distinct. The period between the first and the second wave of technological change (1550 to 1800) saw a wide variety of experiments, in which some but not always all ingredients were proposed or implemented. The variety of outcomes offered a rich mine of lessons in monetary doctrine.

Mechanization and the Age of Copper

The first major shift in minting technology took place around 1550. In southern Germany, two processes were independently developed to mechanize the minting process, using machines rather than tools to cut uniform blanks and impress them with a design. One technology (the screw-press) proved to be better than the other (the cylinder-press), but also more expensive, and only prevailed in the late seventeenth century. Until then, the other proved popular in a number of countries, including the various German states and Spain.

The king of Spain heard about the cylinder-press technology from his cousin the count of Tirol, who had been the first to install the new machines in his state mint. The machines were imported and set up in Segovia in 1582, and applied to the silver coinage of pieces of eight. The coins produced in Segovia were much more uniform and round, and more sharply imprinted, than anything done using the old hand tools. The Spanish government soon realized the potential in this technology, and decided in 1596 to produce all small denominations in pure copper with the new machines. King Philip II explained his reasons in an edict:

> We have been advised by people of great experience, that the silver which is put in those billon coins[15] is lost forever and no profit can be drawn from it, except in their use as money, and that the quantity of silver which is put to that use for the necessities of ordinary trade and commerce in this kingdom is large. We have also been advised that, since we have established a new machine in the city of Segovia to mint coins, if we could mint the billon coinage in it, we would have the assurance that it could not be counterfeited, because only a small quantity could be imitated and not without great cost if not by the use of a similar engine, of which there are none other in this kingdom or the neighboring ones. And it would thus be possible to avoid adding the silver (in Rivero, 1919, p. 150).

Until then, copper, silver, and minting costs each represented a third of the face value of billon coinage. With Philip II's decree, the silver was withheld and the copper content reduced.

Philip II had efficiency in mind. He ordered that the new copper coins be issued only to retire existing small denomination coins (M_2) with token coinage and that the mechanism with its melting and minting points be preserved for providing large denomination silver coins (M_1). Retaining the mechanism for supplying M_1 would keep the price level within the appropriate melting and minting points so long as some large denomination coins continued to circulate.

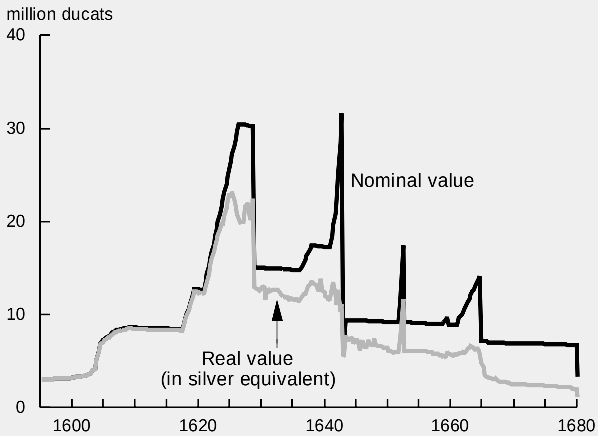
But Philip II's successors, Philip III (1598–1621) and Philip IV (1621–64), saw that the cylinder press offered opportunities to enhance revenues. A first experiment in 1602, whereby the copper content of coins was reduced by 50 percent with no resulting effect on the price level, convinced the government that the intrinsic value of the coins could be made much lower and the seigniorage rate much more lucrative. Another experiment, carried out in 1603, further reinforced the point that individuals did not care about the composition of their money balances. After the 1602 reduction, two kinds of pennies circulated, one twice as heavy as the other; it was decided that all old (heavy) pennies were to be brought to the mint, stamped with a "2" and one two-cent coin returned for every two old pennies presented. The operation was successful and all old pennies were presented, affording the government 50 percent seigniorage on the stock of pennies.

From that point on, the Castilian government knew no restraint, and enormous quantities of *vellón* (as these copper pennies were called) were minted and used to finance government consumption. Figure 5 shows the path of nominal and real balances of vellón in that period; note that the total money stock before 1600 was around 20 to 30 million ducats.

Recall that we express the total quantity of money as $M = M_1 + eM_2$, where M_1 represents the stock of large denomination (silver) coins, and M_2 represents small denomination (copper) coins. The exchange rate between the two types of coins is e, and M_2 is expressed in dollars. The policy followed by the Castilian government consisted in increasing M_2 to the point at which it completely replaced M_1, all the while with no inflation (real and nominal balances coincide). In terms of the total money stock, $M_1 + eM_2$, a

FIGURE 5

Nominal and real values of vellón, 1595–1680



Source: F. Velde and W. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.

progressive displacement of M_1 by vellón is consistent with no change in *e*, and, other things being equal, an unchanged money stock will correspond to a constant price level. However, once M_1 has disappeared, the money stock consists only of copper coins M_2, and all further increases in M_2 result in increases in the price level, as is apparent in figure 5. Once the figure of about 20 million ducats was reached, nominal and real balances diverged, and inflation set in with a vengeance. The disappearance of silver released the price level from the constraints imposed by the melting / minting points for the dollar interval, and unleashed the quantity theory with copper as the determinant of the price level.

The only way to return the price level to its bounds was to engineer a reappearance of the silver coins, either by decreasing M_2 or by decreasing *e*. The Castilian government toyed with the idea of decreasing M_2 by an open-market operation (selling bonds to buy back the copper coinage), but in the end decided to halve *e* overnight, in 1628.

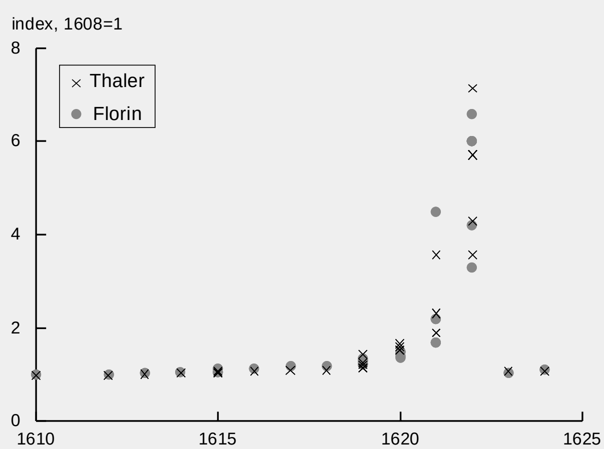
The rest of the movements in vellón balances are due to repetitions of the earlier operations of vellón issue, restamping (multiplying the face value of existing coins by *N* and extracting a seigniorage of $(N–1) / N$) and overnight devaluations. As figure 5 shows, Castilians grew weary of the manipulations, which were less successful as balances of vellón fell over time.

The Spanish experience unleashed unprecedented "man-made" inflation, which made the Price Revolution of the sixteenth century (price level increases due to the inflow of American gold and silver) look tame. It was among the first large-scale experiments in inconvertible fiat currency (although the coins were accepted at face value in payment of taxes). It demonstrated the ease with which token coinage could overtake the money stock, the workings of the quantity theory, the need for the issuer of inconvertible token coinage to restrain issues, and the strength of the temptation created by high seigniorage rates for a government unwilling or unable to raise other taxes.

The Spanish experiment was not the only one at the time. During the Thirty Years War, which started in 1618, many German states concurrently debased their small denominations (all the while maintaining silver coinage intact) and issued large amounts of copper coinage to raise revenues through seigniorage. The results are shown in figure 6, which tracks the exchange rate between large denomination coins and small denomination coins and makes it clear why the Germans called this *die große Inflation* (the great inflation), at least until a similar experiment exactly 300 years later (the famous

FIGURE 6

Prices of the gold florin and silver thaler, in Bavarian kreutzers



Source: H. Altmann, 1976, *Die Kipper- und Wipperinflation in Bayern*, München: Neue Schriftenreihe des Stadtarchivs München, pp. 272–273.

German hyperinflation of 1922–23 under the Weimar Republic). Poland and Russia also underwent copper inflations in the 1650s, as did the Ottoman empire in the 1690s. This is why the seventeenth century has earned the name the Age of Copper.

Lessons from the Castilian inflation

The lessons were not lost on contemporary observers. The Spanish episode was discussed not only by writers in Spain, but also in Italy, France, and elsewhere, leading to a consensus on quantity limitations and limited legal tender for small coins.

One of the more famous commentators was the Jesuit Juan de Mariana (1536–1624), who wrote a treatise on the vellón coinage between 1603 and 1606, as the experiment was beginning and inflation had not yet taken off. He lays out arguments pro and con, and thus provides a window on the debates among policymakers around the Spanish king.

The advantages vaunted by proponents of the copper coinage are not limited to the social savings mentioned by Philip II in his edict. Proponents claimed that without a stock of silver coins as a potential reserve to settle trade deficits, Spain would be forced to maintain surpluses and resort to import substitution, thereby stimulating Spanish industry; they also claimed that the copper money was lighter and easier to transport, and that its cheap provision would lower the rate of interest and stimulate agriculture and industry. In other words, arguments were made that, beyond the social savings from forsaking commodity money, increases in the quantity of money could stimulate output.[16]

Mariana was conscious that incentives for counterfeiting created by the overvaluation of copper coins could be resolved by the new machines in the Segovia mills. He was doubtful of the arguments on balance of trade and stimulus of the economy, which could be made to go the other way through an anticipated inflation effect. He predicted that copper coinage would drive out silver, lead to an increase in prices, and induce the government to set price controls that would either be ignored or counterproductive, at which point the government would be forced to reduce the face value of the coins, as indeed happened in 1628. Mariana saw the projected sequence of inflation and deflation as disruptive to trade and contracts and, therefore, to the king's tax revenues. He also viewed the high seigniorage rates of 50 percent in the restamping operations as immoral, because in his view the king has no right to tax his subjects without their explicit consent. Mariana noted that such high tax rates would never be tolerated on any other tax base. The worst consequence he predicted was general hatred of the government. Quoting Tacitus, he recalled that "everyone claims prosperity for himself, but adversity is blamed on the leader" (1994, p. 104).

The Frenchman Poullain, quoted earlier, concluded that token coins could replace other coins for domestic transactions and that this was precisely why their quantity should be limited. Poullain, as a monetary official, successfully fought back various plans to issue copper on a large scale. Only twice, in 1640 and 1653, did France come close to embarking on a Spanish-style inflation, in both cases at times of fiscal emergency.

The Italian Montanari, also quoted above, wrote: "It is clear enough that it is not necessary for a prince to strike petty coins having metallic content equal to their face value, provided he does not strike more of them than is sufficient for the use of his people, sooner striking too few than striking too many. If the prince strikes only as many as the people need, he may strike of whatever metallic content he wishes" (Montanari, 1804, p. 109). Various other writers stressed quantity limitations, as well as limited legal tender for small coins. The latter measure uncouples the two stocks of money in the equation $M_1 + eM_2$, which was critical in the Spanish experience.

Monopoly versus laisser-faire

English coins had always been made of sterling silver, and shortages of small change became particularly acute when pennies and farthings ceased to be minted altogether in the sixteenth century. From that point until 1817, English policy alternated between three regimes for the supply of small change: private monopolies of inconvertible token coinage, government monopoly of full-bodied coinage, and laisser-faire (that is, the absence of government intervention).

Private monopolies (1613–44) were created by royal charter, which granted various individuals in turn (usually well-connected aristocrats) the exclusive right to issue token coinage, although these were never made legal tender and their quantities were limited by the terms of the charter. A government monopoly was asserted in 1672, making private tokens illegal, and the Royal mint issued copper coins, intermittently and insufficiently, until 1754. Although mechanization had been adopted in 1660, England remained committed to full-bodied copper coinage.[17]

The laisser-faire regime (mid-sixteenth century to 1613, 1644 to 1672, and the late eighteenth century) was characterized by the absence of government-issued small denominations and by the issue of tokens by private parties or local governments. In the late sixteenth century, up to 3,000 London merchants issued tokens. In the period from 1644 to 1672, over 12,700 different types of tokens have been catalogued, issued in 1,700 different English towns. From the 1740s on, trade tokens took over when official coinage ceased. Some of these issues were authorized by government. The city of Bristol sought and secured permission to issue farthings in 1652, and went through three different issues over the next 20 years. The Bristol farthings, furthermore, were officially convertible into large denominations. They are also known to have been counterfeited. The government put an end to the laisser-faire regime twice, in 1672 and by the Act of Suppression in 1817; each time, it did so immediately after adopting a new technology.

France's experiences were somewhat parallel. In the early seventeenth century, private monopolies were instituted for brief periods of time. France also had a brief experience with free token issue in 1790–92. The government had decided in September 1790 to issue substantial amounts of large-denomination paper currency backed by a land sales scheme. Soon, thousands of private and municipal banks emerged to intermediate the government's notes with their own small denomination notes and, in some cases, coins. Initially, the government abstained from regulating the industry, which operated on fractional or 100 percent reserves, depending on the institution. But soon the government moved to eliminate its competitors in the business of issuing currency. The government decided to issue medium-sized notes (equivalent to silver coins) in June 1791, followed in December 1791 by small-denomination notes. Technical difficulties postponed the first issue of small notes to August 1792. The government could now impose a monopoly. Within a few weeks, all private banks were forbidden to issue their notes and private coins were outlawed, amid unproven allegations of wildcatting and fraud.

These episodes present parallels with the Free Banking Eras of eighteenth century Scotland and the nineteenth century U.S. One of the ingredients of the standard formula is monopolization of coinage; it is not clear, on theoretical or historical grounds, that this ingredient is needed if the other two (issue and convertibility of tokens) are present. Of the two advantages of the token coinage system, social savings and government revenues, the latter clearly provides an impulse toward monopolization that the former does not.

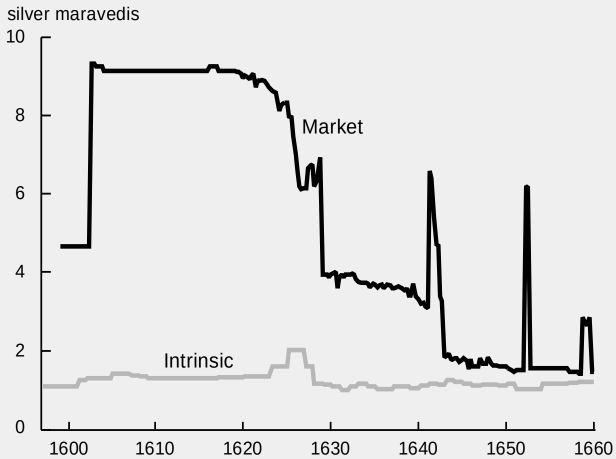
The steam engine and the gold standard

The second major technological innovation following the mechanization of minting around 1550 was the adaptation of the steam engine to minting. In 1787, Matthew Boulton, partner of James Watt, produced trade tokens for the Anglesey Copper Company. A few years later he was producing copper coins for private issuers across England and even in France. In Paris, the most popular token coins in 1790–92, issued by the firm of Monneron, were minted in Birmingham by Boulton's steam presses. The British government contracted with him to produce official copper coinage in 1797, then bought the technology, and in 1817 eliminated its competitors by making private coins illegal. The new steam-driven presses were used to mint the new silver coins, which, under the Coinage Act of 1816, were for the first time issued as partly fiduciary coins, whose intrinsic value was significantly lower than their face value. It took a decade and a half before an implicit agreement was reached between the Bank of England and the Treasury for the convertibility of the silver coinage into gold upon demand. By 1830, the standard formula had been fully implemented.

England's implementation of the standard formula in 1816 applied both to bronze or copper coinage and to silver coins, leaving gold as the single anchor for the price level (the gold standard) and officially abandoning bimetallism. It took other countries some time to follow suit: Germany in 1871, France and the Latin Monetary Union (Belgium, Switzerland, Spain, and Italy) in 1873, the Netherlands in 1875, and the U.S. in 1873–79 (the so-called Crime of 1873). Recently, researchers (Friedman 1990 and Flandreau 1997) have argued that this abandonment was a mistake, and bimetallism was better suited to stabilizing the price level than the gold standard. Nevertheless, there was no substantial difference between applying the standard formula to silver and applying it to copper coinage, and the

FIGURE 7

Market and intrinsic values of a vellón cuarto coin, 1597–1659



Source: F. Velde and W. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.

forces identified by Friedman (1960) and those leading to coin shortages seemed to lead to the outcome effectively adopted by most countries.

Conclusion

The questions raised by Friedman (1960) about the necessary ingredients for an efficient and well-managed currency are old questions indeed. The big problem of small change led monetary thinking on the path to fiduciary currency, at least in the form of intrinsically trifling but convertible tokens; policy followed only after the right technology became available. As technology changed and experience accumulated, various elements of the standard formula were tried separately, including irredeemable copper money. The resulting inflation led to the recognition that a form of quantity theory was at play, and led governments to formulate various ways of limiting the quantity—through convertibility and through monopoly.

Of the main ingredients of the standard formula, the historical trend points clearly to token coinage. Monopolization is less obvious an outcome, especially given the prolonged Free Token Eras of England. Friedman argues that government needs a monopoly on fiduciary currency because free entry into the issue of irredeemable paper would drive down currency to its intrinsic value (namely, 0). As figure 7 shows, this is what happened in seventeenth century Spain, as the market value of copper coinage was driven down to its intrinsic value. Arguably, counterfeiting was widespread, but judging by figure 5, government issues are enough to account for the phenomenon. Surely, experiences with fiat money in the twentieth century (a century replete with hyperinflation) show that governments can drive the value of a paper currency they monopolize to its intrinsic value with great efficacy.

Perhaps it is not surprising that seventeenth century Spain was under an autocratic regime, as was contemporary France (which came close to the same outcome). England, where counterweights to the executive were at least apparent at the time and constitutionally set in 1688, maintained a different policy. Nor perhaps is it a surprise that the standard formula was first implemented in Britain, the most advanced democracy in Europe at the time. The policy was implemented in 1816, just as Britain was emerging from a successful use of inconvertible paper money to finance 20 years' worth of wartime expenditures (in contrast to France's similar attempt in 1790–97, which proved less durable). Irredeemable currency for deficit financing was already a centuries-old idea; the Catalonian town of Gerona used a coin issued as siege money to start a convertible-token system in 1481. Success with deficit financing was probably a good predictor of success with subsidiary coinage; both may have something to do with the degree of accountability of policymakers.

NOTES

[1]Much of the material presented here derives from the work in Sargent and Velde (1997a, 1997b).

[2]The model sketched here is developed fully in Sargent and Velde (1997a).

[3]Seigniorage is literally the lord's right to collect a tax, and is derived from the French term for lord, *seigneur*.

[4]This cost is exclusive of the coin's content. It represents the costs of transforming metal into coins, and is to some degree independent of the content.

[5]The equation is $pY = vM$, where p is the price level, Y is income or the volume of transactions, v is velocity, and M is the quantity of money.

[6]The optimal denomination structure is an unstudied problem; however, see Telser (1995).

[7]In the numismatic sense, token means something that is not officially money, but used as money; numismatists will speak of full-bodied tokens. From an economic viewpoint, the distinction between official and unofficial money is somewhat arbitrary.

[8]Act of June 9, 1879: "Be it enacted ... that the holder of any of the silver coins of the United States of smaller denominations than one dollar, may, on presentation of the same in sums of $20, or any multiple thereof, at the office of the Treasurer or any assistant treasurer of the United States, receive therefor lawful money of the United States" (Statutes at Large 21 [1879]: 7).

[9]The status of silver dollars remained uncertain between the Bland–Allison Act of 1878 and the final defeat of the pro-silver forces after 1896. Only after 1900 did the silver dollar become no different in nature from other subsidiary coins.

[10]Note that the ability to maintain a monopoly on full-bodied coinage is dependent on the same.

[11]The punishment for counterfeiters was particularly severe. In medieval France, they were boiled alive (not poached). A document from 1311 details the costs of executing two counterfeiters, including the price of a large cauldron and the cost of adding iron bars to the cauldron, a detail that suggests a rather long process (Saulcy, 1879–92, Vol. 1, p. 180).

[12]An anthology of their writings is in Velde (1997).

[13]This did not preclude the denomination of many prices in terms of the gold coin and fictitious subdivisions thereof.

[14]Interestingly, the coins were modeled on a currency issued some years earlier as emergency money during a siege and later left in circulation.

[15]A mixture of silver, to give value, and copper, to give bulk, commonly used for small denominations.

[16]Another famous proponent of similar arguments was the Scot John Law (1671–1729), whose experiment in setting up a paper currency in France went spectacularly awry in 1720, during the Mississippi Bubble.

[17]The proclamation of 1672 stated that small coins "cannot well be done in silver, nor safely in any other metal, unless the intrinsic value of the coin be equal, or near to that value for which it is made current." Sir Isaac Newton, master of the mint, wrote in 1720: "Halfpence and farthings (like other money) should be made of a metal whose price among Merchants is known, and should be coined as near as can be to that price, including the charge of coinage. ... All which reasons incline us to prefer a coinage of good copper according to the intrinsic value of the metal" (Shaw, 1896, pp. 164–165).

REFERENCES

Altmann, Hans Christian, 1976, *Die Kipper- und Wipperinflation in Bayern*, München: Neue Schriftenreihe des Stadtarchivs München.

Aristotle, 1943, *Ethics*, H. Racham (trans.), Cambridge, MA: Harvard University Press.

Biel, Gabriel, 1935, *Treatise on the Power and Utility of Money*, Robert B. Burke (trans.), Philadelphia, PA: University of Philadelphia Press.

Botet y Sisó, Joaquim, 1911, *Les Monedes Catalans*, Vol. 2, Barcelona: Institut d'Estudis Catalans.

Budel, René, 1591, *De Monetis et Re Numaria*, Book 1, Cologne: Schann Nackel.

Cipolla, Carlo M., 1956, *Money, Prices and Civilization in the Mediterranean World, Fifth to Seventeenth Century*, New York: Gordian Press.

D'Isernia, Andrea, 1541, *Commentaria in Usus Librum Feudorum*, Lyon, France: J. Giunta.

Flandreau, Marc, 1997, "The gradient of a river: Bimetallism as an implicit fluctuation band," Paris: l'Observatoire Français des Conjonctures Economiques, manuscript.

Friedman, Milton, 1990, "The crime of 1873," *Journal of Political Economy*, Vol. 98, December, pp. 1159-1194.

__________, 1960, *A Program for Monetary Stability*, New York: Fordham University Press.

Mariana, Juan de, 1994, *De Monetae Mutatione*, Josef Falzberger (ed.), Heidelberg: Manutius Verlag.

Mill, John Stuart, 1857, *Principles of Political Economy*, Fourth Edition, London: John W. Parker and Son.

Montanari, Geminiano, 1804, "Della moneta: Trattato mercantile," in *Scrittori Classici Italiani di Economia Politica*, parte antica, Vol. 3, Milan: G.G. Destefanis.

Poullain, Henri, 1709, *Traité des Monnoies*, Paris: Frédéric Léonard.

Rivero, Casto Maria del, 1919, *El Ingenio de la Moneda*, Revista de Archivos, Bibliotecas y Museos, Vol. 40.

Sargent, Thomas J., and François R. Velde, 1997a, "The big problem of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-8.

__________, 1997b, "The evolution of small change," Federal Reserve Bank of Chicago, working paper, No. WP-97-13.

Saulcy, Ferdinand de, 1879–92, *Recueil de Documents Relatifs à l'Histoire des Monnaies*, 4 vols., Paris: Imprimerie Nationale.

Shaw, William Arthur, 1896, *The History of Currency, 1252 to 1894*, London: Wilsons & Milne.

Spengler, Joseph J., 1966, "Coin shortage: Modern and premodern," *National Banking Review*, Vol. 3, pp. 201–216.

Stampe, Ernst, 1928, "Das Zahlkraftrecht der Postglossatorenzeit," *Abhandlungen der Preußischen Akademie der Wissenschaften*, philosophisch-historische, Klasse 1, Berlin: Verlag der Akademie.

Telser, Lester G., 1995, "Optimal denominations for coins and currency," *Economic Letters*, Vol. 49, pp. 425–427.

Tesauro, 1609, "Tractatus de monetarum augmento et variatione," in *De Monetarum Augmento Variatione et Diminutione Tractatus Varii*, Turin, Italy.

Velde, François R., 1997, "An anthology of writings on money, 12th–17th century," Federal Reserve Bank of Chicago, manuscript.

Velde, François R., and Warren E. Weber, 1997, "Fiat money inflation in 17th century Castile," Federal Reserve Bank of Chicago, manuscript.

The decline of job security in the 1990s: Displacement, anxiety, and their effect on wage growth

Daniel Aaronson and Daniel G. Sullivan

Introduction and summary

The news media frequently suggest that American workers have suffered a significant decline in job security during the 1990s. Of course, planned and actual employment reductions at major corporations such as AT&T, IBM, and General Motors have been important stories. But articles such as those in the 1996 *New York Times* book *The Downsizing of America* go beyond reporting individual cases of layoffs to suggest that there has been a fundamental change in the employment relationship. According to such articles, workers in general have suffered a loss of job security, and long-term employment relationships are a thing of the past. Moreover, such articles claim that this decreased job security has left workers feeling more anxious about their futures.

The perception of declining job security is shared by many policymakers and other analysts, who believe worker anxiety to be a major reason wage inflation in the 1990s has remained modest in the face of historically low levels of unemployment. Perhaps most famously, Federal Reserve Board Chairman Alan Greenspan testified to Congress in February 1997 that "atypical restraint on compensation increases has been evident for a few years now and appears to be mainly the consequence of greater worker insecurity."[1] Former U.S. Labor Secretary Robert Reich recently made much the same point when he wrote, "Wages are stuck because people are afraid to ask for a raise. They are afraid they may lose their job."[2]

Labor economists, however, have often been skeptical of claims of widespread declines in job stability and security. They note that media accounts are long on anecdotes and short on evidence based on nationally representative survey data. Moreover, the most carefully executed studies using scientifically designed survey data collected through the early 1990s often reached conclusions quite at odds with media reports. For instance, Diebold, Neumark, and Polsky (1997) concluded that "aggregate job retention rates have remained stable." Similarly, Farber (1998) found that "there has been no systematic change in the overall distribution of job duration over the last two decades."

More recently, however, researchers have begun to analyze survey data from the mid-1990s, and conclusions somewhat more in line with media reports are emerging. For instance, Neumark, Polsky, and Hansen (1997) reported that "there is some evidence that job stability declined modestly in the first half of the 1990s. Moreover, the relatively small aggregate changes mask rather sharp declines in stability for workers with more than a few years of tenure." Similarly, Farber (1997b) concluded that "after controlling for demographic characteristics, the fraction of workers reporting more than ten and more than 20 years of tenure fell substantially after 1993 to its lowest level since 1979." Thus job stability—the tendency of workers and employers to form long-term bonds—seems to be declining somewhat. Moreover, evidence of significant change is especially apparent in more direct measures of worker security, such as Farber's (1997a) tabulations of the number of workers reporting involuntary job loss. The extent of changes in job tenure, turnover, and displacement reported in these more recent studies is much too modest to justify

Daniel Aaronson is an economist and Daniel G. Sullivan is a senior economist and vice president at the Federal Reserve Bank of Chicago. *The authors would like to thank Ann Ferris for her very capable assistance.*

the most sensationalistic news reports. Nevertheless, some decline in job security, especially for workers who have attained significant seniority, now seems reasonably clear.

In this article, we review some of the findings of this research on job stability and job security. We then present some new tabulations of rates of job loss for high seniority workers based on the Bureau of Labor Statistics' (BLS) *Displaced Worker Surveys* (DWS). Next, we look directly at workers' own perceptions of their job security using data from the National Opinion Research Center's *General Social Survey* (GSS). Finally, we attempt to relate our measures of displacement and worker anxiety to wage growth by examining time-series data for the nine U.S. census divisions.

Our tabulations of annual displacement rates from the DWS focus on workers with five or more years of tenure. We find that among such workers, job loss due to "shift or position abolished," which among the surveys' possible reasons for job loss comes closest to capturing the notion of "downsizing," increased quite dramatically from annual rates of two tenths or three tenths of a percent throughout the 1980s to a range of six tenths or seven tenths of a percent in the mid-1990s. Determining the trend in displacement more generally is complicated by changes in the DWS. However, our preferred estimates suggest that overall displacement rates were higher in 1995, the most recent year for which we have data, than at any time since the data began in 1979. We estimate a 1995 displacement rate of about 3.4 percent for workers with five or more years of tenure. By comparison, the rate for 1982, which was in the middle of a severe recession, was only about 2.5 percent. We consider this a substantial increase in the risk of displacement for high-seniority workers.

We also find that displacement has become somewhat more "democratic" in the 1990s. Previously, high-seniority workers who were highly educated, were in white-collar jobs, or were employed in the service producing industries were relatively immune to displacement. More recently, however, displacement rates for these groups have risen especially fast, while those for some groups who had high rates of displacement in the 1980s, such as those with at most a high school education, those in blue-collar occupations, and those working in manufacturing, rose less or even fell relative to their peaks in the early 1980s. As a result of this increased democratization of displacement, many more workers may now consider themselves at risk for job loss.

The GSS data suggest that workers' own perceptions of their job security have also declined. The fraction of workers not responding "very unlikely" to the question, "How likely is it that you will lose your job in the next year?," rose from about 31 percent in 1989 to about 40 percent in 1996, the most recent year for which data are available. The 1996 figure approximately matches the highest reading since this question began to be asked in 1977. The 1996 reading is especially remarkable given that unemployment was generally below 6 percent, while in 1982, when such a level of anxiety was previously reached, unemployment was nearly 10 percent. One should not, however, exaggerate the extent to which workers' anxiety over job loss has increased. The main change has been an increase in the number of workers responding that it is "not too likely" rather than "very unlikely" that they will lose their jobs. The percentages of workers responding that it is "fairly likely" or "very likely" have risen more modestly.

With a few exceptions, the groups of workers who have experienced the largest increases in displacement rates have also had the largest increases in reported probabilities of job loss. For instance, an increase in perceived likelihood of job loss has been especially great among white-collar workers. Perceived job security has actually increased for blue-collar workers. Another interesting finding concerns the relationship between workers' perceptions of their job security and the use of computers in their industry. In the early 1980s, workers in industries with greater computer usage felt more secure on average than other workers. By the mid-1990s, however, the relationship had reversed, with workers in industries with greater computer usage feeling less secure.

Finally, we attempted to judge to what extent our findings of an increase in displacement rates and workers' perceptions of their chances of job loss are related to changes in aggregate wages. Standard short-run Phillips curve analyses such as Gordon (1997) have tended to predict higher levels of wage inflation than have actually occurred in the last two or three years. Though significant forecast errors are nothing new for such models,[3] the importance of the question to policymakers has led to a great deal of speculation as to why wage inflation has remained subdued. Our findings and those of other researchers, which suggest that job security has declined in recent years, add some plausibility to the case that worker anxiety has played a role in restraining wage inflation.

However, one could easily point to other recent changes in labor markets or elsewhere in the economy that might be affecting wage growth.[4] Why should one particular change—that towards reduced job security—be considered the key factor? To make a more convincing case for the importance of worker insecurity, one would want to observe that in the past, when displacement or anxiety was high relative to unemployment, wage inflation had also been subdued. Moreover, to obtain any sense of the quantitative importance of worker insecurity in restraining wage inflation, one needs to examine historical evidence. Unfortunately, with annual measures of displacement and insecurity that go back only to the late 1970s, there is not much hope of extracting such information from the U.S. time-series data.

Our strategy is to look cross-sectionally as well as over time and ask whether census regions that have had higher displacement rates or worker perceptions of insecurity have tended to have lower wage growth. Such a strategy parallels the "wage curve" analyses of Blanchflower and Oswald (1994), although, as in Blanchard and Katz (1997), we employ the traditional Phillips curve specification in which the change in wages, rather than their level, is related to unemployment and measures of job security.

We pool separate data from the nine census regions to estimate the effect of displacement rates or perceptions of insecurity on forecasts of wage inflation. We find that, holding constant the unemployment rate, higher values of both displacement and worker insecurity are associated with lower wage growth. However, even with the additional source of data variation that comes from pooling the separate census divisions, our estimates of the magnitude of these effects are imprecise. Indeed, if we allow for the possibility that there may be some permanent, unmeasured characteristics of regions that are associated with different levels of wage growth, we cannot reject the hypothesis that the true effect of job security on wages is zero and that estimates the size of those we obtain could have arisen by chance. Nonetheless, our best estimates suggest that increases in displacement rates and workers' own anxiety about their job security could be responsible for restraining wage growth by about three tenths to seven tenths of a percentage point per year during the mid-1990s. Such an effect would explain all or most of the puzzle of lower than expected wage inflation.

Previous research on turnover and displacement

The *New York Times* describes its book on downsizing as putting "a human face on a historic predicament that is as ubiquitous as it is painful."[5] A large body of research demonstrates that job loss is painful, at least for workers who have attained significant tenure.[6] For example, Jacobson, LaLonde, and Sullivan (1993c) found that even six years after job loss, earnings losses among a sample of Pennsylvania workers displaced in the early 1980s were still equal to about 25 percent of their predisplacement earnings levels. What has been somewhat less clear to researchers is whether job loss has become any more ubiquitous in recent years.

It is helpful to divide the relevant research into two parts—that on job stability and that on job security. By stability we mean the tendency for workers and firms to develop long-term relationships. Research on job stability questions the many media accounts claiming that such long-term employment relationships have gone the way of buggy whips. By security we mean workers' ability to remain in employment relationships as long as their own performance is satisfactory. Research on job security asks whether there has been an increase in involuntary job loss due to reasons beyond workers' control. Job stability depends on workers' own choices, in addition to the factors that influence job security. For instance, if a group of workers increase their commitment to the labor force or to their particular employers, then their job stability may rise even if they are increasingly subject to threats of displacement. As we shall see, research suggests a larger 1990s decline in job security than in job stability.

In our view, trends in job security are much more relevant to the discussion of whether special factors might be restraining wage inflation than are trends in job stability. Indeed, if declines in job stability are less dramatic than declines in job security, it must largely be because workers are less likely to leave jobs voluntarily, and a decreased tendency to quit jobs may itself signal worker insecurity. Nevertheless, we begin with a short account of research on job stability.

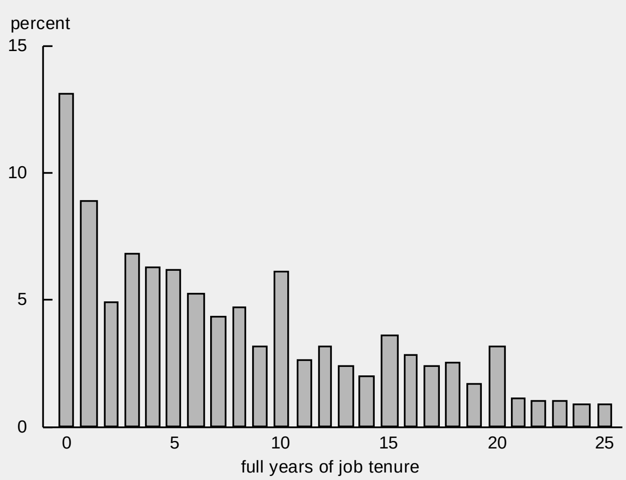
The starting point for much of the research on job stability is the distribution of job tenure.

Most of what is known about this distribution derives from a series of supplements to the *Current Population Survey* (CPS).[7] As an illustration, figure 1 displays the distribution of job tenure for employed men between the ages of 35 and 44. These data were collected from the most recent *Mobility Supplement* to the CPS, which was conducted in February 1996. The figure shows that the most common tenure levels are the shortest—for example, less than one year and between one and two years—with a roughly monotonic decline in the number of workers with successively longer tenure. Nevertheless, there are many workers with substantial levels of job tenure. For men in the 35 to 44 age group, the median tenure is about 6.1 years. Moreover, about 33 percent of such workers have been in their current jobs at least ten years and about 22 percent have been in their current jobs at least 20 years.

Figure 2 shows how median job tenure has changed over time for men and women in three age groups, 25 to 34, 35 to 44, and 45 to 54. These data are derived from CPS *Mobility Supplements* conducted in January of 1963, 1966, 1968, 1973, 1978, 1983, 1987, and 1991 and February of 1986.[8] Not surprisingly, older workers typically have longer job tenures than younger workers. Also, men typically have longer tenures than women, who are more likely to have interrupted their careers for family reasons. However, our primary interest is in the aggregate trends in these data.

FIGURE 1

Distribution of job tenure for men aged 35–44, February 1996



Source: Authors' calculations based on data from U.S. Department of Commerce, Bureau of the Census, *Current Population Survey, Mobility Supplement*, February 1996.

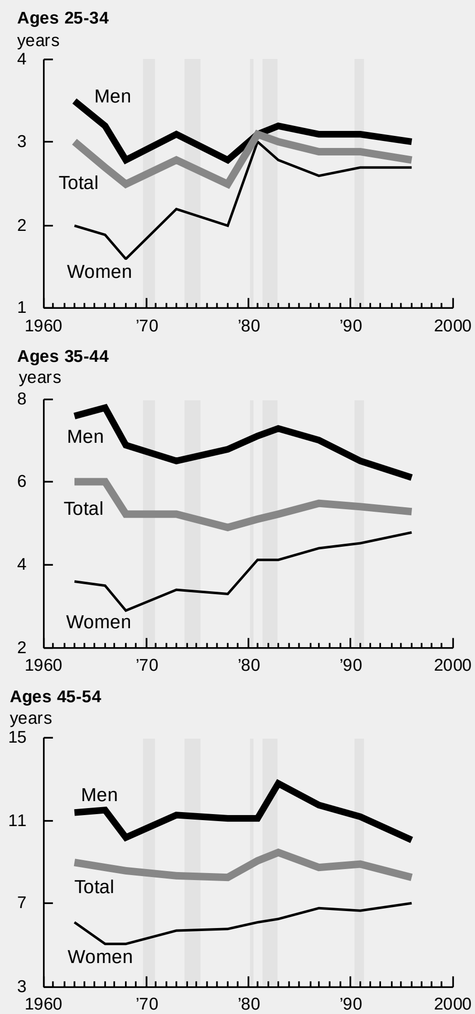
For men, especially those in the two highest age groups, median job tenures declined from 1991 to 1996, which is consistent with claims of decreased job stability. However, women's job tenure rose for all age groups. So, overall, there has been relatively little change in median job tenure during the 1990s. Moreover, the drop in male tenure for the two oldest groups seems to be mainly a continuation of a trend that was evident throughout the 1980s. Thus, it is difficult to conclude that job stability has suffered more than a modest decline in the 1990s.[9]

Farber (1997b) shows that similar, though somewhat more dramatic, changes took place during the 1990s at the high end of the tenure distribution. In particular, he shows that the percentages of workers reporting more than ten and 20 years of tenure declined significantly between 1991 and 1996. The proportion of workers aged 35 to 64 with ten or more years of tenure declined from 38.3 percent to 35.4 percent. For men the decline was more dramatic, from 44.3 percent to 40.0 percent, while among women, the decline was from 31.4 percent to 30.3 percent. Similar drops were reported for workers with different educational levels. However, the occupational groups that have historically had the highest long-term employment levels, such as managerial, professional and technical, and blue-collar workers, had the largest declines in the 1990s. Similarly, the declines were greatest in industries, such as transportation, communications, and public utilities, in which long-term employment had been most common. As a result, the frequency of long-term employment is now more similar across occupations and industries.

Farber's (1997b) results, as well as the trends in median job tenure shown in figure 2 suggest that job stability among men has declined modestly during the 1990s. For women, job stability has either declined very modestly or continued to rise, depending on whether it is measured by median tenure or the proportion of workers with high tenure levels. For both sexes, the changes appear too modest and gradual to support the sensationalistic media reports proclaiming the end of long-term employment relationships.

FIGURE 2

Median tenure, ages 25–54



Note: Shaded areas indicate recessions.

Source: Authors' calculations based on data from U.S. Department of Labor, "Employee tenure in the mid-1990s," January 30, 1997, and Farber (1998).

To reach a tenure of ten years, a worker must survive from the first year of a job into the second, from the second to the third, and so on for ten years. Thus, the distribution of job tenures and, in particular, the fraction of workers with ten or more years of tenure can be thought of as depending on a sequence of survival probabilities going back many years. This means that if the probability of remaining in a job for another year were to have suddenly dropped sometime in the early 1990s, it would take a number of years for this change to show up fully in the tenure distribution. In this case, results such as Farber's (1997a) and those shown in figure 2 might not reveal the full extent of change. For this reason, it is of interest to examine job survival or retention probabilities.[10]

Diebold, Neumark, and Polsky (1997) and Neumark, Polsky, and Hansen (1997) have carefully analyzed the trend in retention rates from the mid-1980s to the mid-1990s, using data on tenure distributions from CPS supplements for every four years from 1983 to 1995. The authors estimate four-year retention rates—the probability that a worker will remain in a job an additional four years—by dividing the number of workers with a given set of characteristics and a certain number of years of tenure in one survey by the number of workers with those characteristics but four fewer years of tenure in the survey four years earlier. Adjustments are made for a number of potential problems, including nonresponse to the survey, the tendency of workers to round their tenure to a multiple of five years, the differing levels of unemployment at the time of the surveys, and the special nature of the tenure data derived from the February 1995 CPS *Supplement on Contingent Work*.

Table 1 contains some representative results from Neumark, Polsky, and Hansen (1997).[11] Evidently, the trend over time in retention rates depends to a great extent on workers' initial level of tenure. For workers with less than two years of initial tenure, four-year retention probabilities are estimated to have increased from 32.9 percent for the 1983–87 period to 34.6 percent for the 1987–91 period to 39.1 percent for the 1991–95 period. However, for workers with two to less than nine years of tenure, rates first declined then rose slightly. The strongest evidence of a decline in job stability comes from the group of workers who initially had between nine and 15 years of tenure. The retention rate for these workers declined from 81.6 percent for 1987–91 to 74.8 percent for 1991–95. Retention rates also declined sharply between 1987–91 and 1991–95

TABLE 1

Four-year job retention rate estimates

Initial tenure	1983–87	1987–91	1991–95
0 to < 2	32.9%	34.6%	39.1%
2 to < 9	58.6	54.8	56.4
9 to < 15	82.7	81.6	74.8
15 plus	63.0	70.2	63.3
Weighted average	53.9	53.6	54.4

Source: Derived from Neumark, Polsky, and Hansen (1997).

for workers with 15 or more years of tenure, but returned to rates observed in the 1983–87 period. The weighted average rate was quite stable, falling just 0.3 percentage points from 1983–87 to 1987–91 and then increasing 0.8 percentage points from 1987–91 to 1991–95.

The results on retention probabilities are consistent with those on tenure levels in suggesting some modest declines in job stability for workers with several years of tenure. Several researchers have reported more dramatic declines in job stability during the 1980s and / or 1990s. For example, Boisjoly et al. (1994), Rose (1995), and Marcotte (1996) report evidence of declining job stability from the *Panel Study of Income Dynamics* (PSID) data. Similarly, Swinnerton and Wial (1995) reported significant declines in job retention rates in the 1991–95 period. However, in our view the combined results of Diebold et al. (1996), Swinnerton and Wial (1996), and Jaeger and Stevens (1997) show that the more dramatic declines reported in the literature were largely the result of researchers failing to take account of occasional changes in survey question wording. The most careful analyses of job stability trends imply that there have been at most modest declines in stability in the late 1980s and 1990s.[12]

Research suggests, we believe, larger declines in measures of job security. Farber (1997a) analyzes data from the seven *Displaced Worker Surveys* (DWS), CPS supplements that are described in the next section. He finds that "rates of job loss are up substantially relative to the standard of the last decade, particularly when some consideration is given to the state of the labor market." He also finds that displacement rates increased most for several groups, such as the more educated and those in white-collar occupations, that have traditionally had relatively low levels of displacement, which implies that displacement has become somewhat more democratic. Changes in the reasons workers give for their job loss also point to especially large increases in what the media might mean by "downsizing." Finally, Farber finds that the consequences of displacement in terms of time spent unemployed and reduced wage rates upon reemployment appear to be mainly a function of the business cycle. The consequences of displacement were worse during the recessions of the early 1980s and 1990s, but there is little evidence of a secular increase in the seriousness of displacement.

Valetta (1997) also finds an increasing number of dismissals in data from the PSID, which is consistent with Farber's results (1997a). Moreover, Valetta's finding that the increase is concentrated among workers with higher levels of tenure is consistent with our finding below that the increase in displacement rates for workers with five or more years of tenure has been especially dramatic.

Displacement trends for high-seniority workers

Below, we present new measures of the rate of job displacement for workers with five or more years of tenure. We had two main goals. First and most important, we wanted the measures to be comparable over time so that we could accurately judge whether displacement was increasing. Given changes in the underlying survey methodology, this is not completely straightforward and, despite our best efforts, it is possible that certain of our measures change over time for reasons that have nothing to do with actual changes in the rate of job displacement. Our second goal was to create annual time series, the highest frequency possible, so as to be better able to examine the relationship between displacement and wage inflation.

Our measures of displacement are based primarily on the Bureau of Labor Statistics' DWS. These surveys were conducted as supplements to the CPS in January of even years from 1984 to 1992 and in February 1994 and 1996.[13] For the purposes of the survey, displacement is defined as involuntary job loss not related to a worker's performance. Thus, displacement excludes quits and cases in which workers are discharged for poor performance.[14] The surveys are retrospective, asking individuals whether they have experienced job loss any time over the last five years in the case of the 1984 to 1992 surveys and over the last three years in the case of the 1994 and 1996 surveys. Thus, our earliest information on displacement is for 1979 and our latest is for 1995.

For workers who report that they were displaced in the relevant time period, the DWS asks for the specific reason for their displacement. The possible responses are:

- Plant or company closed down or moved,
- Insufficient work,
- Position or shift abolished,
- Seasonal job completed,
- Self-operated business failed, and
- Some other reason.

This list of reasons is less than ideal. For example, insufficient work might be the reason why one of the other events occurred. A plant may have closed because there was insufficient work to do. Position or shift abolishment is probably supposed to cover instances of "corporate downsizing," but it is possible that those in nine to five work environments will be confused by the reference to shifts. In any case, it lumps together instances of complex "re-engineering" exercises, which presumably reflect long-run organizational changes, with closings of shifts in factories, which are more likely to be associated with short-run declines in demand. The seasonal job and self-employment categories don't correspond to many people's conception of job displacement and, in fact, make up only a trivial fraction of the job loss that we consider. Finally, perhaps because of some of the ambiguities of the preceding categories, "other" is a common response. In fact, growth in the "other" category is responsible for a large percentage of the total growth of displacement of high-seniority workers.

The first difficulty we face in constructing a consistent measure of worker displacement is that the DWS only collects information, such as the year of displacement, the worker's tenure, and other characteristics of the lost job, for at most one incident of displacement over the relevant period. If workers were displaced twice or more in the same period, they are instructed to answer the additional questions for the lost job on which they had the highest tenure. This inevitably leads to some undercounting of incidents of displacement. Moreover, as Farber (1997a) notes, the change in the length of the period over which the DWS asks workers to report on displacement creates a problem of comparability over time, since the undercounting problem is more severe when the interval covered is five years.

Farber's approach to this problem is to examine only displacement that occurred in the last three years of the five-year periods covered by the 1984 to 1992 surveys. As he notes, these rates are still not comparable to rates computed from the three-year intervals of the 1994 and 1996 surveys, because some workers may lose a job in year one or two of the five-year period before the survey and then lose another job in year three, four, or five. If the workers had accumulated less tenure on the second lost job than they had on the first, they would be recorded as losing a job in the 1994 and 1996 surveys, but not in the last three years before the 1984 to 1992 surveys. Farber's solution is to use PSID data to quantify the frequency of job loss patterns and adjust rates in the DWS to offset them.[15]

Our approach is to restrict our analysis to incidents of job displacement in which the affected workers had five or more years of tenure. Obviously, it is not possible to lose two such jobs in one three- or five-year interval, so the number of such job loss incidents should be correctly tallied no matter whether the year is part of a three- or five-year interval in the DWS. Of course, we will miss *all* displacement incidents in which workers had less than five years of tenure. However, the consequences of job loss are not likely to be particularly great for workers with little tenure and, thus, our measure may capture the most important forms of job displacement.

The DWS gives us estimates of the number of workers with five or more years of tenure who are displaced in a particular year. To calculate a displacement rate, we need to divide this estimate of the number of high-tenure displaced workers by the number of high-tenure workers who were at risk in that year. We derive the latter figure as the product of the level of total employment and the fraction of total employment accounted for by workers with five or more years of tenure.

Our estimated displacement rate is $r_t^5 = \frac{d_t^5}{n_t f_t^5}$,

where d_t^5 is the number of workers with five or more years of tenure displaced in year t, n_t is total employment, and f_t^5 is the fraction of employment accounted for by workers with five or more years of tenure.

As noted above, we derive estimates of d_t^5 from the DWS. To estimate n_t, we use the CPS outgoing rotation files. The outgoing rotations are those CPS members who are in the fourth and eight months of their eight-month participation, about 25 percent in a given month. Pooling the outgoing rotations for all 12 months of the year yields a large data set that can be used to estimate employment levels quite precisely. To estimate f_t^5, we use the CPS tenure supplements described earlier. As noted, these were conducted in 1981, 1983, 1987, 1991, and 1996. To compute displacement rates for 1979 and 1980, we use the value of f_t^5 from 1981. For other years in which there was no supplement, we interpolate linearly from the preceding and succeeding tenure supplements. Because the fraction of workers with five years of tenure changes very slowly relative to the number of displaced workers, this interpolation causes no problems.

There is another problem with the 1994 and 1996 DWS. The follow-up questions on the details of the displacement episode are not asked if workers do not give one of the first three standard reasons for displacement. This is unfortunate because, as we have already noted, a nontrivial and growing number of workers report "other" as their reason for displacement. To ignore workers not responding with one of the three standard reasons would, we feel, significantly skew our results.[16] However, for workers giving a nonstandard reason, we do not know whether they had five years of tenure or in what year they lost their job.

To deal with this problem, we estimated statistical models to gauge the percentage of displaced workers giving nonstandard reasons who had five years of tenure and, of those, the percentage who were displaced in each of the three years covered by the surveys. The details of our procedure are contained in box 1. The idea is to use the displaced workers giving nonstandard reasons in 1992 to determine which worker characteristics reported in the basic CPS were associated with having five years of tenure and then to use those characteristics to estimate the percentage of displaced workers reporting nonstandard reasons in the 1994 and 1996 surveys who had five years of tenure. Similarly, we used the workers reporting standard reasons and five years of tenure in the 1994 and 1996 surveys to determine the characteristics associated with being displaced in each of the three years covered by the surveys and then used those characteristics to predict which year workers reporting nonstandard reasons were displaced.

Table 2 shows our basic results for overall displacement rates. The rows of the table correspond to the years in which displacement occurred. The first five columns of the table correspond to the number of years after the displacement year that the displacement rate was measured. For example, the only information on the 1979 displacement rate comes from the 1984 DWS, which was conducted with a lag of five years. The estimated rate, 0.96 percent, is thus shown in the column headed five-year lag. For the majority of years, we have multiple measures of the displacement rate. For example, for 1985 we have estimates from the 1986, 1988, and 1990 DWS. These rates, shown in the columns for one-, three-, and five-year lags, are estimated to be 2.29 percent, 1.79 percent, and 1.43 percent, respectively.

TABLE 2

Percent displaced among workers with five or more years tenure: Alternative measurement lags

Year	One-year lag	Two-year lag	Three-year lag	Four-year lag	Five-year lag	Overall estimate
1979					0.96	1.46
1980				1.23		1.69
1981			1.59		1.23	1.92
1982		2.27		1.84		2.52
1983	2.41		1.66		1.52	2.25
1984		1.75		1.22		1.80
1985	2.29		1.79		1.43	2.22
1986		1.99		1.56		2.18
1987	1.83		1.34		1.34	1.84
1988		1.44		1.18		1.61
1989	1.71		1.62			1.85
1990		1.90				2.11
1991	2.76		2.35			2.83
1992		2.40				2.67
1993	2.92		2.31			2.89
1994		2.21				2.46
1995	3.44					3.44

Source: Authors' calculations from data of the U.S. Department of Labor, Bureau of Labor Statistics.

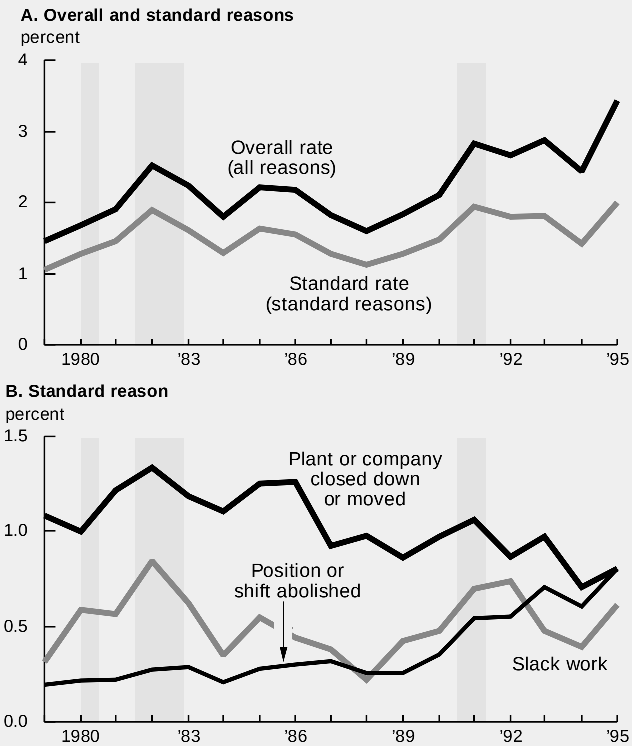
The results for 1985 illustrate the final difficulty we face in constructing an annual measure of job displacement for workers with five or more years of tenure. That is, displacement rate estimates tend to drop as the time since the survey increases. Workers seem to forget incidents of displacement as time passes, a phenomenon noted previously by Topel (1990) and others. As a result, it is inappropriate to simply average the various measures to arrive at an overall displacement rate for that year. For instance, if we were to directly compare the single estimate for 1979 with the single estimate for 1995, we would be comparing a rate measured with a five-year lag with a rate measured with a one-year lag. Thus, the comparison would reflect not only differences in actual displacement rates between the years, but also the tendency of rates measured with a greater lag to be lower.

Table 2 reveals that estimated displacement rates tend to drop on average by about 11 percent for each additional year that the survey lags the year of displacement. Our solution to this problem, which is described in detail in box 1, is essentially to adjust rates based on lags greater than one year upward by about 11 percent for each additional year that the survey lags the year of displacement. Our final estimates of the annual displacement rates, which are shown in the last column of table 2, are averages of all the adjusted rates for the year in question. For instance, the estimated 11 percent annual decline suggests that if the rate for 1979 had been measured in 1980, it would have been 1.46 percent, rather than 0.96 percent. Thus, our final estimate for 1979 is 1.46 percent. In a year with multiple measurements, the measures are adjusted by different amounts, depending on how long after the year of displacement the survey was taken. For instance, the estimate for 1985 obtained with a one-year lag is left at 2.29 percent, but the rate obtained with a three-year lag is adjusted up from 1.79 percent to 2.18 percent to reflect the additional two years since the survey; the rate with a five-year lag is adjusted from 1.43 percent to 2.17 percent to reflect the additional four years since the survey. The adjusted rates of 2.29 percent, 2.18 percent, and 2.17 percent are then combined to obtain the final estimate of 2.22 percent.[17]

The final results are plotted over time as the black line in figure 3, panel A. The overall displacement rate for workers with five years of tenure rose during the recessions of the early 1980s from 1.5 percent in 1979 to a peak of 2.5 percent in 1982. It then declined during the economic expansion that followed to a low of about 1.6 percent in 1988. Then, in the 1990s, it rose rather dramatically. It is not surprising that the rate should have risen during the recession of 1990–91, but the 1991 rate, at over 2.8 percent, was 0.3 percentage points higher than in 1982, even though by most measures the 1982 recession was much more severe. More noteworthy is the failure of the displacement rate to decline during the expansion of the mid-1990s. Indeed, in 1995 the rate shot up to 3.4 percent, its highest ever reading. The high overall displacement rates that we estimate for the mid-1990s are consistent with the view that job security declined significantly for workers with five or more years of tenure.

FIGURE 3

Displacement rates, workers with 5 years tenure



Note: Shaded areas indicate recessions.
Source: Authors' calculations based on data from U.S. Department of Labor, Bureau of Labor Statistics, *Displaced Worker Survey*, 1984–96.

BOX 1

Constructing an annual index of displacement

To estimate the fraction of workers reporting nonstandard reasons for displacement in the 1994 and 1996 DWS who had five or more years of tenure, we estimated a logistic regression model using the sample of such workers in the 1992 DWS.[1] The dependent variable in this model was an indicator for having five years of tenure and the independent variable consisted of dummy variables for the nine census regions, sex, ten-year age categories, race, marital status, education less than high school, high school graduate, some college, and college degree, as well as part-time status, one-digit occupation, and one-digit industry of the person's job as reported in the main CPS. We then used the estimates of the parameters of this model, along with the equivalent characteristics for workers reporting nonstandard reasons for displacement in 1994 and 1996 to form an estimate of the probability that such workers had five or more years of tenure at the time of their job loss.

We estimated the fraction of such workers that were displaced in each of the three possible years covered by the 1994 and 1996 surveys by estimating a multinomial logistic regression model on the sample of 1994 or 1996 displaced workers reporting standard reasons for displacement. In this model, the dependent variables were indicators for the year of displacement and the independent variables were the same as in the model above. The parameter estimates were then used to estimate the probability that workers were displaced in each of the three years covered by the 1994 and 1996 surveys.

In computing displacement rates based on the 1994 and 1996 surveys, we then counted all workers reporting nonstandard reasons for displacement as displaced in all three possible years. However, we multiplied the weights for such individuals by the estimated probabilities of having five or more years of tenure and of being displaced in the year in question. This procedure should provide estimates of displacement rates among those with five years of tenure that are consistent over time if 1) the relationship between the probability of five-year tenure and the independent variables remains constant from 1992 to 1994 and 1996, and 2) the distribution of year of displacement conditional on the independent variables is the same for workers displaced due to standard and nonstandard reasons.

The final task in computing annual displacement rates is to combine rates measured for the same year by different surveys into a single overall rate. We did this by estimating the following simple statistical model:

$$\log r_{st} = \alpha_t + \gamma(s - t - 1) + \varepsilon_{st},$$

where r_{st} is the displacement rate for year t measured by the survey in year s, and ε_{st} is an error term assumed to have constant variance and to be uncorrelated across observations. The parameter γ measures the rate at which estimates of displacement rates decline as time between displacement and the survey increases. Its estimate corresponds to an approximately 11 percent rate of decline. The overall rate is captured by the year of displacement effects, α_t. Specifically, the estimate of the rate corresponding to a one-year lag between displacement and the survey is $\exp(\alpha_t)$. These are the estimates shown in the final column of table 2 and plotted in figure 3.

In order to compute estimates for separate demographic groups, we expanded the above model to

$$\log r_{dst} = \alpha_{dt} + \gamma(s - t - 1) + \varepsilon_{dst},$$

where r_{dst} is the rate for demographic group d in year t as measured by the DWS of year s. The demographic specific rates are then $\exp(\alpha_{dt})$. We also computed estimates of displacement rates adjusted for changes in the age and sex distribution. These were based on models of the form

$$\log r_{dstk} = \alpha_{dt} + \beta_k + \gamma(s - t - 1) + \varepsilon_{dstk},$$

where r_{dstk} is the rate for the age and sex group k. The presence of the β_k controls for changes in the age and sex distribution that might affect estimates of overall rates. However, the adjusted rates were similar enough to the unadjusted rates that we only report the latter.

[1]See, for example, Maddala (1983) for an explanation of the logistic regression model discussed below.

The colored line in figure 3, panel A shows the rate of displacement due to the first three standard reasons in the survey. Comparing the two lines, it is clear that a large part of the significant mid-1990s increase is due to an increase in the number of displaced workers giving “other” as their reason for displacement. However, even the colored line, which is not dependent on the imputations of tenure and year of displacement described in box 1, suggests that there has been some decline in security, especially given the level of unemployment. The rate of displacement for standard reasons is estimated to be higher in 1995 than it was in 1982, even though the unemployment rate was below 6 percent during most of 1995, while it was nearly 10 percent in 1982. Thus, even when limited to displacement for standard reasons, our results suggest a noticeable decline in job security.

Figure 3, panel B displays separate displacement rates for the three standard reasons. Evidently, the rate due to firms or plants closing or moving has declined somewhat in the 1990s, while the rate due to slack work has remained relatively high, given the state of the business cycle. However, the most notable feature of figure 3, panel B is the sharp increase beginning in 1990 of the displacement rate due to shifts or positions being abolished. This rate, which probably comes the closest to capturing corporate downsizing, was between 0.2 percent and 0.3 percent from 1979 to 1989, but rose to more than 0.8 percent in 1995. This two hundred or three hundred percent increase seems to represent a rather significant break from history.

Figure 4, panel A shows the overall displacement rate for men and women. For most of the period covered by our data, women were less subject to displacement than men, with the typical gap in rates being five tenths or six tenths of a percentage point. In the last three years, however, the gap has been much smaller, about one tenth of a percentage point. Thus, by our measure, women have suffered a larger decline in job security than men. This finding highlights the difference between the displacement rates estimated here and the trends in median tenure discussed earlier. Median tenure has generally been increasing for women relative to men. However, tenure levels are measures of stability, reflecting workers’ own commitment to the labor force and individual employers in addition to forces beyond workers’ control, such as displacement. In the comparison of male and female tenure levels, workers’ own choices are likely the more important factor. For this reason, we would argue that displacement rates are the better measure of worker insecurity.

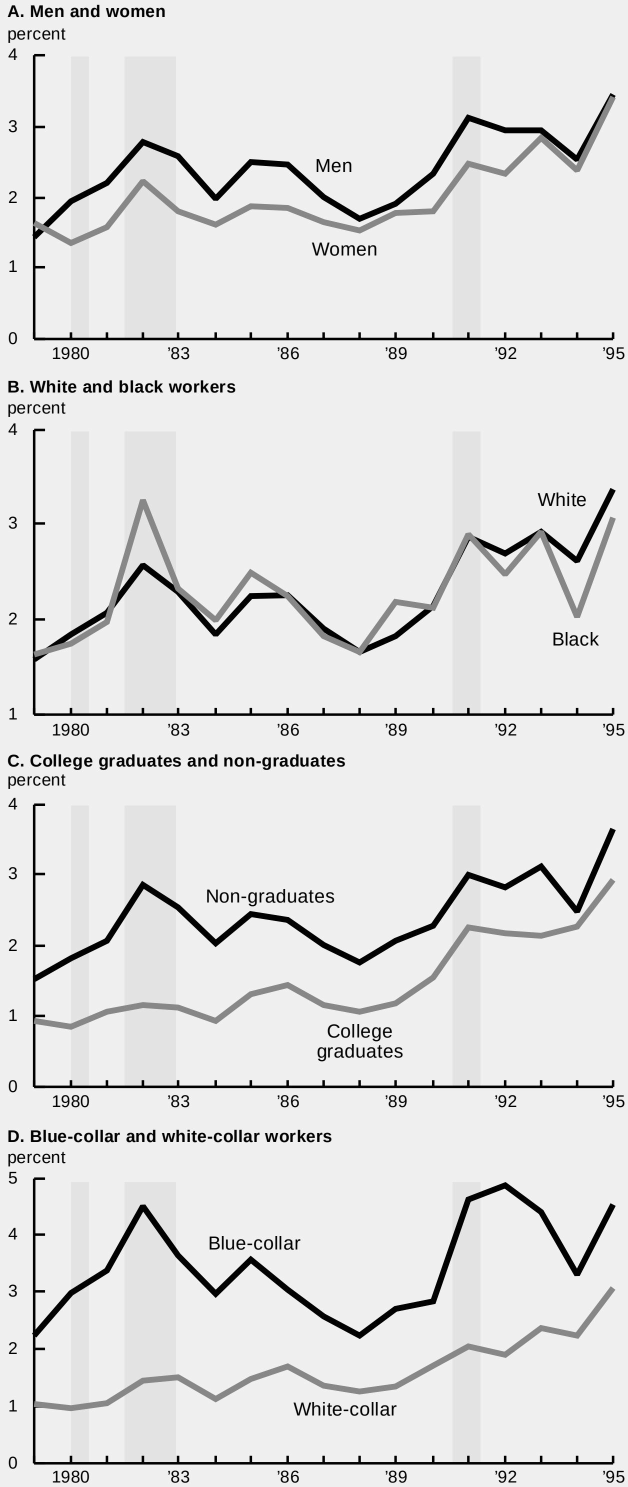
Figure 4, panel B displays displacement rates for white and black workers. Once we restrict the sample to workers with five or more years of tenure, the difference between the races is relatively minor. Still, there have been some changes over time. Early in the period covered, especially during the recession of the early 1980s, blacks had noticeably higher displacement rates. However, by the end of the period, whites had higher rates of displacement.

Figure 4, panel C shows the breakdown between those with a college degree and those without a college degree. Although displacement rates for college graduates remain much lower than those for workers without college degrees, the gap has narrowed considerably in the 1990s. Until 1990, displacement rates for college graduates never exceeded 1.3 percent and the gap between them and non-graduates was often a percentage point or more. In the 1990s, displacement rates for college graduates rose especially sharply, to levels of more than twice their previous peak. Thus, the gap in displacement rates between those with a college degree and those without has narrowed considerably, though rates for college graduates remain significantly lower.

Figure 4, panel D shows displacement rates for blue-collar and white-collar workers. Though displacement rates for high-tenure blue-collar workers remain about a percentage point higher than those for white-collar workers, the gap has clearly shrunk during the 1990s. For instance, even the recessions of the early 1980s had little effect on displacement rates for high-tenure white-collar workers, but since 1988, their rates of job loss have approximately doubled. By contrast, the recessions of the early 1980s caused a major increase in blue-collar displacement to levels only slightly lower than in recent years. Even more dramatic differences in displacement trends are observed between more narrowly defined occupations. For example, 1995 displacement rates for laborers are significantly lower than in 1982, while those for professional and technical workers are approximately three times higher.

FIGURE 4

Displacement rates among workers with five or more years tenure, by demographic category



Note: Shaded areas indicate recessions.

Source: Authors' calculations based on data from U.S. Department of Labor, Bureau of Labor Statistics, *Displaced Worker Survey*, 1984–96.

Figure 5 shows estimated displacement rates for workers in goods producing and service producing industries. Again, a large gap in rates in the 1980s has narrowed appreciably. Displacement rates for those in goods producing industries are still significantly lower than in 1982, but rates for those in the service producing industries are about two and half times greater. Even so, workers in goods producing industries remain significantly more at risk for displacement than those in service producing industries. More dramatic changes can be identified for certain industries. For instance, displacement rates for workers in the finance industries rose from about 0.5 percent to 1.0 percent in the 1980s to 2.8 percent in 1995.

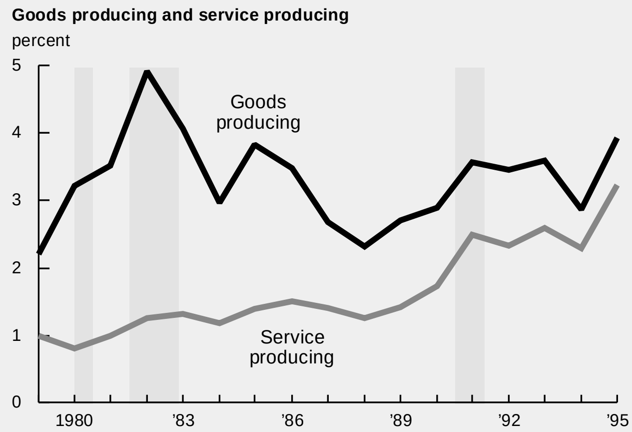
The results in figures 4 and 5 all point to the general increase in high-tenure displacement rates having been accompanied by a kind of democratization, in which those who had been relatively immune to job displacement have seen the fastest increase in displacement. Previously, those with a college education, in white-collar jobs, or in service producing industries might have considered themselves immune to job loss. Given the increase in displacement rates that we have estimated for these groups, this is probably no longer the case for many such workers. Thus, the number of workers who feel at risk may have increased even more than the increase in the displacement rate would suggest.

Workers' perceptions of job security: The NORC-GSS

In a series of recent papers, Manski (1990, 1993) has observed that researchers know a great deal about the outcomes that individuals or groups experience but much less about the outcomes that they expect. This assertion is particularly relevant for job security research, which to date has focused on the measurement of displacement rates, tenure distributions, and other measures of actual employment outcomes. However, a primary issue in this literature concerns measuring perceptions of

FIGURE 5

Displacement rate, by industry



Note: Shaded areas indicate recessions.
Source: Authors' calculations based on data from U.S. Department of Labor, Bureau of Labor Statistics, *Displaced Worker Survey*, 1984–96.

the risk of future economic harm. Therefore, these measures are indirect, in the sense that expectations about risk, which are subjective in nature, must be inferred from individual or group realizations.[18]

The *General Social Survey* (GSS) data set allows us to address the perceptions question directly. Up to now, this data set has received some attention in the popular press but little among researchers studying job security.[19] The GSS is a nationally representative annual survey conducted by the National Opinion Research Center (NORC). The survey asks a series of demographic and employment questions, including, in most years since 1977, two questions about job security. Respondents are asked 1) "Thinking about the next 12 months, how likely is it that you will lose your job or be laid off—very likely, fairly likely, not too likely, or not at all likely?" and 2) "About how easy would it be for you to find a job with another employer with approximately the same income and fringe benefits that you now have? Would you say very easy, somewhat easy, or not easy at all?"

Data are available for 13 years between 1977 and 1996 covering roughly 10,000 individuals. Several years (1980, 1984, 1987) are missing because NORC did not ask the job security questions, and other years (1979, 1981, 1992) are missing because the GSS was not conducted. The sample includes all respondents who are currently employed, English-speaking, and aged 18 to 64. It is important to note that the sample makes no restriction on tenure because such information is not given in the GSS. Therefore, the job security perceptions sample is not strictly comparable to the displacement rate sample discussed earlier. This probably accounts for some different trends among subsamples of the population.

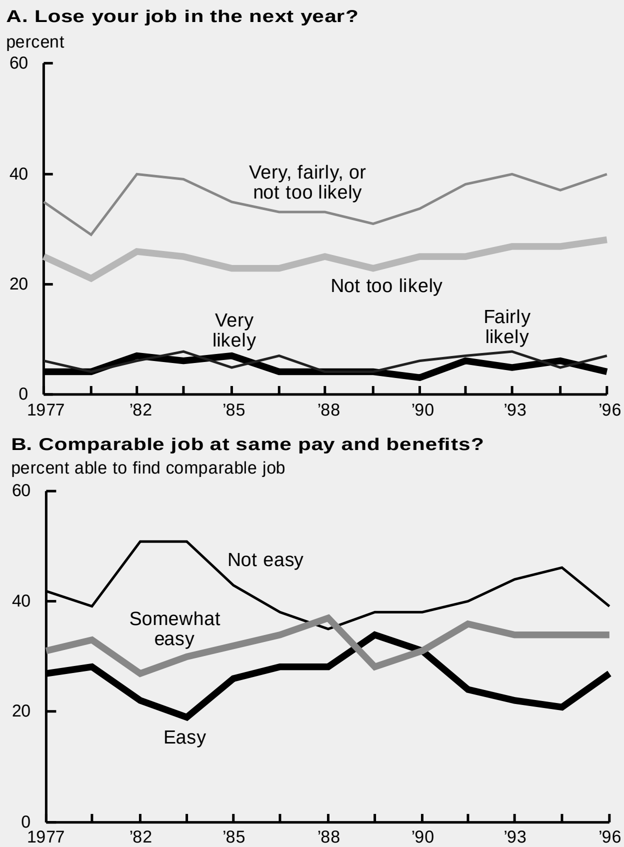
The GSS does have some important limitations.[20] Most noteworthy is that each GSS survey year consists of an independently drawn nationally representative sample of the population. Thus, unlike other national surveys such as the PSID, the GSS does not allow us to observe the same individuals across time. Surveys that follow individuals allow the use of panel data techniques to control for unmeasured individual-specific characteristics, such as ability or ambition, that change across the business cycle and are correlated with the other variables in the model. Such a survey format would allow us to investigate the future employment dynamics of workers and examine whether job anxiety predicts future job displacement or wage loss.

The easiest way to see how perceptions of job security have changed over time is to graphically examine the responses to the GSS questions. Figure 6, panels A and B show the distribution of responses to the two questions from 1977 to 1996. Each line represents a separate response except the highest line in panel A, which is the sum of the very, fairly, and not too likely responses. Between 30 percent and 40 percent of workers feel some degree of insecurity about losing their job in the next year, although only 10 percent of respondents feel very or fairly sure that job loss will occur. Between 35 percent and 50 percent of workers respond that it would not be easy to find a comparable job.

As with the displacement rates, the responses are fairly cyclical through the early 1990s. Using the job loss likelihood question, job security declined during recessions in the early 1980s and 1990s and increased during the expansion of the 1980s. But since 1991, the percentage of workers who answer that they are not at all likely to lose their job has fallen, despite the strong and widely felt expansion of the economy. Amazingly, in 1996, the fraction of workers who answered that they had some concern about their job's future

FIGURE 6

Potential for job loss and comparable reemployment



Source: Authors' calculations based on data from the National Opinion Research Center, *General Social Survey*, 1997.

was equal to the percentage that answered this way during the severe 1982–83 recession. However, most of this is due to an increase in the percentage of workers who answer they are not too likely to lose their job. Therefore, while there has been a noticeable shift in worker anxiety during this expansion, most of the change is due to workers acknowledging some, albeit a slight, likelihood of losing their job over the next year.

The job comparability question also tends to be cyclical, but showed signs of breaking this trend during the initial phase of the 1990s expansion. Beginning in 1988, the percentage of workers who answered that it would not be easy to find a comparable job at the same pay and benefits monotonically increased, peaking at almost 46 percent in the 1994 survey. However, in the 1996 survey, the "not easy to find a comparable job" response declined and the percentage answering it was easy to find a comparable job increased.

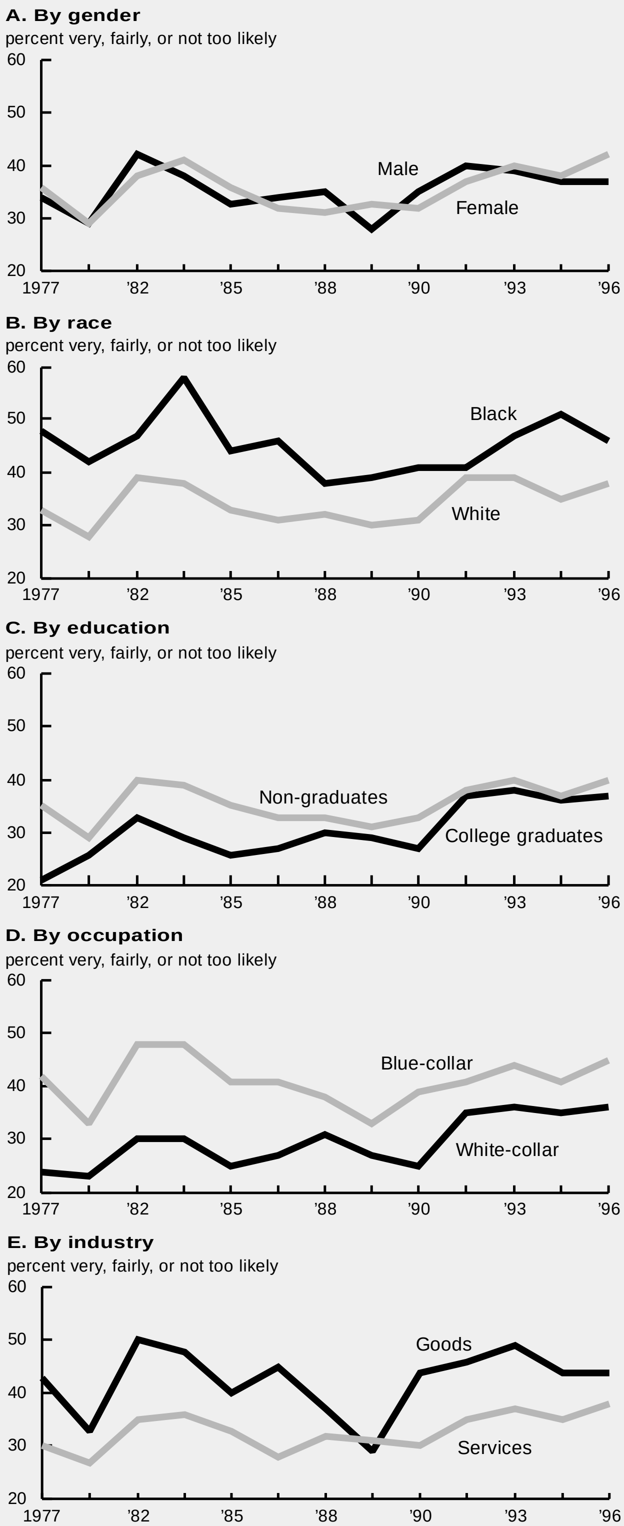
Therefore, through 1996, workers seemed somewhat less concerned about their chances of finding a comparable job, but somewhat more concerned about the likelihood of losing their current job.[21]

Figures 7 and 8 show the trends in these two series by gender, race, education, industry, and occupation. The two primary differences between the GSS results and the displacement rate results are exhibited in figure 7, panels A and B. Panel A shows that there is no male–female gap in perceptions of job security throughout the sample period, while the displacement rates showed a large male–female gap that narrowed over time. Figure 7, panel B displays a large black–white gap in worker anxiety that has narrowed somewhat over time, while figure 4 showed no significant difference in displacement rates by race, except during the 1982 recession. We believe that a substantial portion of these differences may be due to the different tenure restrictions in the samples. That is, while the results on displaced workers come from a sample of workers with five years of tenure, we cannot make comparable restrictions on the GSS sample. The importance of this restriction is evident in other research. For example, Fairlie and Kletzer (1997) use the DWS to estimate displacement rate gaps between black and white workers but make no sample restrictions based on tenure. They find a 30 percent gap in displacement rates between the races from 1982 and 1991. Likewise, between 1982 and 1991, the black–white gap in the GSS job loss data is 29 percent.

On the other hand, figure 7, panels C and D look quite similar to the displacement rate results, pointing again to a democratization of job insecurity. White-collar and college-educated workers were relatively immune to job anxiety during the 1970s and 1980s, but have experienced substantial increases in job insecurity during the 1990s. The change has been large enough to basically eliminate the gap in job insecurity between college graduates and non-graduates. Blue-collar workers still feel less secure than white-collar workers, but the gap is less than half what it was in the 1970s and early 1980s.

FIGURE 7

Likelihood of job loss, by demographic category



Source: Authors' calculations based on data from the National Opinion Research Center, *General Social Survey*, 1997.

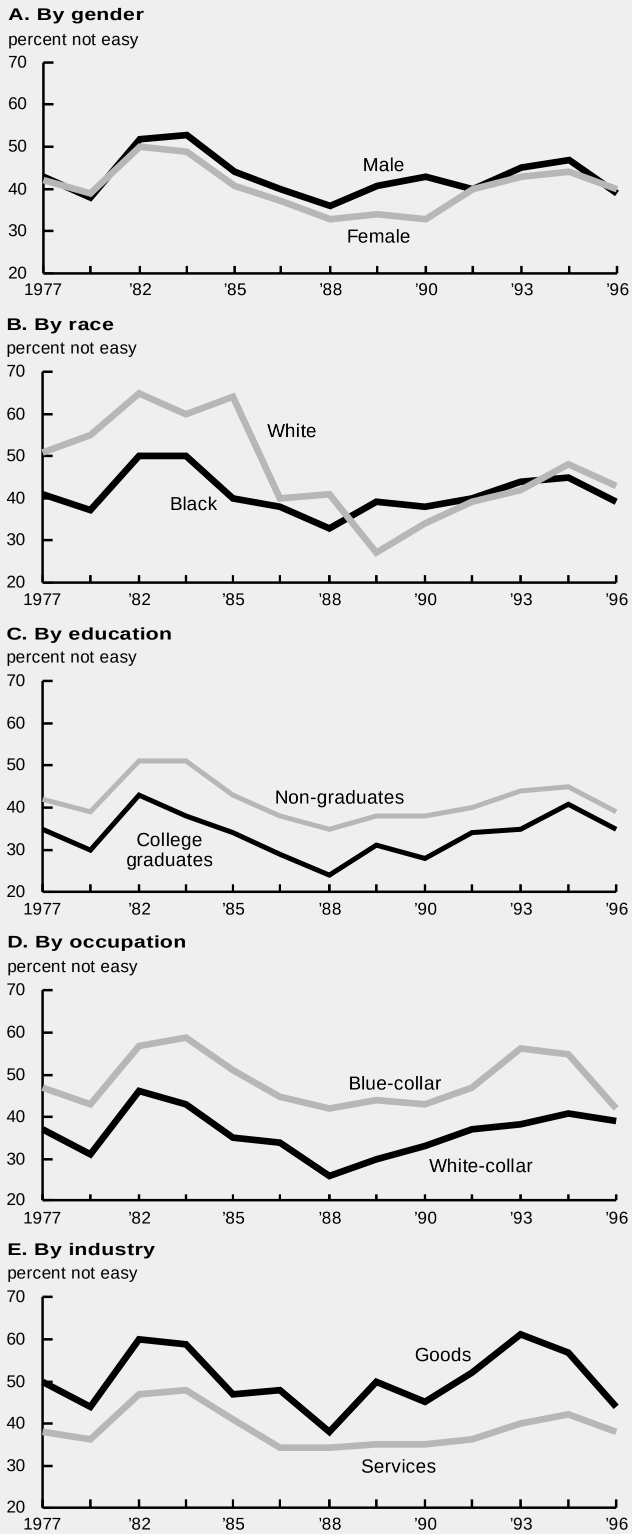
As shown in figure 7, panel E, job security has declined during the 1990s in the service sector but has remained relatively flat (other than a temporary drop in 1993) in the goods sector. Most of the decline in the service industry arises from the services, finance, insurance, and real estate (FIRE), and government sectors. Analogous to the displacement rate findings, perceptions of job security have dropped substantially in FIRE, with roughly 50 percent fewer workers saying that they are not at all likely to lose their job in the next year. The lack of movement among goods producing sectors hides some variance between specific industries. In particular, job insecurity (measured by the probability of losing your job) in manufacturing has doubled since 1989, surpassing the level of anxiety witnessed in 1982. In 1996, job insecurity in the manufacturing sector was substantially higher than in all other major industries. The goods sector has not increased because agriculture and construction workers have experienced corresponding declines in job insecurity over the past few years.

Figure 8, panel A shows the percentage of male and female workers who believe it is not easy to find a comparable job with the same pay and benefits. This graph shows a small but persistent male–female gap that is eliminated in 1996. Using this job security measure, most of the 1990s increase in anxiety appears to be due to female workers. Figure 8, panel B shows that a large black–white gap during the 1970s and early 1980s had disappeared by the end of the 1980s.

Figure 8, panels C and D again display the diminishing gap in worker anxiety between college-educated and non-college-educated workers and white- and blue-collar workers, respectively. Panel D shows that the historically large difference between white- and blue-collar employees all but vanished in 1996. White-collar workers are among the few groups that did not experience a sharp drop in anxiety

FIGURE 8

Likelihood of finding comparable employment, by demographic category



Source: Authors' calculations based on data from the National Opinion Research Center, *General Social Survey*, 1997.

about finding a comparable job in 1996, reflecting increased anxiety among professional workers. In 1996, 42 percent of professional workers responded that it was not easy to find a comparable job, up from 30 percent in 1989, matching the percentage that answered that way during the 1982 recession.

Finally, the 1996 drop in anxiety about finding a comparable job is mainly from the goods sector (figure 8, panel E). Anxiety about finding a comparable job for service sector employees peaked in 1994, but remains slightly above the levels seen during the last expansion. Nearly every group believed it would be easier to find a comparable job in 1996 than in 1994, exceptions being government employees and professional and sales workers.

Controlling for population characteristics

The results presented thus far are based on raw data. However, other changes in the work force during the last 20 years, including shifts in the age and educational distribution of U.S. workers, may be confounding the time trends in worker anxiety. Is the trend in worker anxiety by industry, occupation, or education the same even after simultaneously controlling for multiple characteristics of the population? To find out, we estimate "ordered probit" regressions, an appropriate statistical technique for this problem because it accounts for the discrete and ordered nature of the job security questions. The details of the estimation procedure are described in box 2.

Table 3 reports the coefficients, standard errors, and marginal effects from a specification that uses the likelihood of losing your job as the dependent variable and industry, occupation, year, gender, race, age, marital status, education, and region dummies as controls.[22] The marginal effects measure the impact of a change in some variable, say whether the individual is a sales worker, on the probability

BOX 2

Ordered probit regressions

The ordered probit model is based on a latent regression such as

$$1)\qquad y_i^* = \beta x_i + \varepsilon_i,$$

where y_i^* is the unobserved job insecurity of person i, x_i are demographic and other individual characteristics of person i and ε_i is a person-specific error term. The parameter β is a vector of coefficients that measure the average impact of the demographic variables on the level of job security. While we do not observe y_i^*, we do observe the k possible answers allowed by the survey, as represented by y_i:

$$\begin{aligned} y_i &= 0 \quad \text{if} \quad y_i^* \leq 0 \\ y_i &= 1 \quad \text{if} \quad 0 \leq y_i^* < \mu_1 \\ y_i &= 2 \quad \text{if} \quad \mu_1 \leq y_i^* < \mu_2 \\ &\vdots \\ y_i &= k \quad \text{if} \quad \mu_{k-1} \leq y_i^*. \end{aligned}$$

For example, in the GSS likelihood of losing your job question, $y_i = 0$ corresponds to answering "not at all likely to lose my job," while $y_i = 3$ corresponds to the "very likely to lose my job" answer. The μ_i's are unknown intercept parameters to be estimated in the model.

Assuming a normal distribution in the error term, we can calculate the probability of each of the k answers as

$$2)\qquad \text{Prob}(y = j|x) = \begin{cases} \Phi(\mu_0 + \beta x) & \text{if } j = 0 \\ \Phi(\mu_j + \beta x) - \Phi(\mu_{j-1} + \beta x) & \text{if } 0 < j \leq k-1, \\ 1 - \Phi(\mu_{k-1} + \beta x) & \text{if } j = k \end{cases}$$

where Φ is the standard normal cumulative distribution function. From equation 2, we can calculate the marginal effect (the impact of a change in the x variable on the probability of event y occurring) by

$$3)\qquad \begin{aligned} \frac{\partial \text{prob}(y=0)}{\partial x} &= \phi(\mu_0 + \beta x)\beta \\ \frac{\partial \text{prob}(y=j)}{\partial x} &= (\phi(\mu_j + \beta x) - \phi(\mu_{j-1} + \beta x))\beta \\ \frac{\partial \text{prob}(y=k)}{\partial x} &= (1 - \phi(\mu_{k-1} + \beta x))\beta, \end{aligned}$$

where ϕ is the standard normal density function and the x variables are measured at their mean value. Equation 3 essentially calculates the effects of changes in the covariates on the cell probabilities.

It should be noted that many of the independent variables in our models are 0–1 indicators, such as whether the individual is a college graduate. In this case, the marginal effect is calculated as the difference between the cell probabilities when the event occurs (a college graduate) and when the event does not occur (not a college graduate):

$$\text{Prob}(y = j|x',1) - \text{Prob}(y = j|x',0),$$

where $x',1$ is the vector of covariates where the college graduate variable is set to 1 and $x',0$ is the vector of covariates where the college graduate variable is set to 0.

TABLE 3

Likelihood of losing your current job in the next year: Ordered probit analysis

	Coefficient	Standard error	Marginal effect on base case probability: Not at all likely	Not too likely	Fairly likely	Very likely
Base case probability[a]			0.697	0.225	0.046	0.032
Agriculture	0.069	0.097	0.024	–0.014	–0.005	–0.005
Construction	–0.322	0.070*	–0.120	0.062	0.027	0.031
Manufacturing	–0.234	0.054*	–0.086	0.046	0.019	0.021
Transportation, communications, and utilities	–0.058	0.066	–0.021	0.012	0.004	0.004
Wholesale trade	–0.021	0.084	–0.007	0.004	0.002	0.002
Retail trade	–0.128	0.057*	–0.046	0.026	0.010	0.010
Finance, insurance, and real estate	0.060	0.069	0.021	–0.012	–0.004	–0.004
Services	–0.079	0.050	–0.028	0.016	0.006	0.006
Professional, technical	0.148	0.047*	0.049	–0.030	–0.010	–0.009
Managerial	0.264	0.049*	0.085	–0.053	–0.017	–0.015
Sales	0.119	0.056*	0.040	–0.024	–0.008	–0.008
Craftsman	–0.014	0.052	–0.005	0.003	0.000	0.000
Operative or laborer	–0.165	0.047*	–0.060	0.033	0.013	0.014
Service worker	0.126	0.048*	0.042	–0.026	–0.009	–0.008
1978	0.125	0.062*	0.042	–0.026	–0.009	–0.008
1982	–0.194	0.061*	–0.071	0.039	0.016	0.017
1983	–0.236	0.060*	–0.087	0.046	0.019	0.021
1985	–0.089	0.060	–0.032	0.018	0.007	0.007
1986	–0.018	0.062	–0.006	0.004	0.001	0.001
1988	0.034	0.069	0.012	–0.007	–0.002	–0.002
1989	0.062	0.070	0.021	–0.013	–0.004	–0.004
1990	0.022	0.070	0.008	–0.004	–0.002	–0.002
1991	–0.147	0.067*	–0.053	0.029	0.012	0.012
1993	–0.184	0.066*	–0.067	0.037	0.015	0.016
1994	–0.121	0.057*	–0.044	0.024	0.009	0.010
1996	–0.151	0.056*	–0.055	0.030	0.012	0.013
Female	–0.035	0.029	–0.012	0.007	0.003	0.003
Black	–0.265	0.039*	–0.098	0.052	0.022	0.024
Other race	–0.095	0.070	–0.034	0.019	0.007	0.007
Age 18–24	0.031	0.043	0.011	–0.006	–0.002	–0.002
Age 44–65	0.130	0.029*	0.044	–0.027	–0.009	–0.008
Never married	–0.111	0.034*	–0.040	0.022	0.009	0.009
Divorced	–0.024	0.037	–0.009	0.005	0.002	0.002
Separated	–0.198	0.064*	–0.072	0.039	0.016	0.017
Widowed	–0.181	0.076*	–0.066	0.036	0.014	0.015
High school dropout	–0.124	0.038*	–0.044	0.025	0.010	0.010
College graduate	0.024	0.038	0.008	–0.005	–0.002	–0.002
Graduate school graduate	0.055	0.056	0.019	–0.011	–0.004	–0.004
New England	0.120	0.064	0.040	–0.024	–0.008	–0.008
Mid-Atlantic	0.008	0.046	0.003	–0.002	–0.000	–0.000
East North Central	0.093	0.044*	0.032	–0.019	–0.006	–0.006
East South Central	0.040	0.059	0.014	–0.008	–0.003	–0.003
South Atlantic	0.045	0.044	0.016	–0.009	–0.003	–0.003
West North Central	0.077	0.055	0.026	–0.016	–0.005	–0.005
West South Central	–0.032	0.052	–0.011	0.006	0.002	0.002
Mountain	–0.116	0.058*	–0.042	0.023	0.009	0.009
Intercept 1[b]	0.517	0.077*				
Intercept 2[b]	1.420	0.078				
Intercept 3[b]	1.849	0.079*				
Log likelihood		–18,316				
Sample size		9,935				

* = significant at the 5% level.

[a]Base case is a white, married male, aged 25 to 44, high school graduate, who worked a clerical government job in the Pacific region in 1977. Industry, occupation, region, and year dummies are relative to government (industry), clerical (occupation), Pacific (region), and 1977 (years).

[b]Each response has its own intercept. See box 2 for details. The three intercept terms are used to compute the marginal effects for the four categories of responses (final four columns).

Note: Dependent variable is the likelihood of losing your job in the next year. The possible answers are not at all likely, not too likely, fairly likely, and very likely.

Source: Authors' calculations based on data from the National Opinion Research Center, *General Social Survey*, various years.

of the individual responding to the job security questions in a particular way. The results are reported relative to a base case white, married, male, high school graduate, aged 25 to 44, who worked in a clerical job in the government in 1977. The first row shows that the probability of the base case person responding that he is not at all likely to lose his job is 69.7 percent. The third row, Construction, reveals that the probability of a not at all likely response from a clerical worker in the construction industry is 12.0 percentage points lower (or 57.7 percent) than for a clerical worker in the government in 1977.

Overall, the table shows that many of the characteristics that look significant in the univariate graphs, such as occupation and race, remain significant indicators, even after controlling for the demographic and employment variables. This is also true of ordered probit regressions where the comparable job question is the dependent variable. Table 3 gives further detail on the specific industry and occupational groups that traditionally experience higher levels of job anxiety, including workers in the construction and manufacturing sectors and operatives and laborers in all industries. While the probability of managerial workers responding that they are not at all likely to lose their job is 78.2 percent, the same probability for an operative or laborer is 63.7 percent. These industry and occupational differences are statistically significant.

We can test whether job security has changed over this expansion by calculating time trend effects within the ordered probit framework. The rows labeled 1978 to 1996 in table 3 show the results of such an exercise. As with the simple univariate graphs, perceptions of job security have been quite low since 1991 when measured by the likelihood of job loss. Controlling for demographic, industry, and occupation shifts cannot explain the recent high insecurity felt by workers. Job anxiety remains on the order of that seen during the last two recessions.

Also, we stratified the sample by gender, race, education, occupation, and industry and ran separate ordered probit regressions for different categories of workers. The purpose of this exercise is to see whether the time trends reported in the graphs still exist after controlling for other demographic, industrial, and occupational structural shifts. By running separate regressions for each demographic group, we allow the parameters on other covariates to change across groups. This flexible specification allows, say, the effect of being married to exert a different influence on perceptions of job security for high school dropouts and college graduates. However, the main inferences from these results (not shown) do not change much. The recent trend in increased job anxiety arises primarily from better educated and white-collar workers. On the other hand, workers who are high school dropouts are more secure about their job in 1996 than at any other time since 1977, with the exception of 1989, the end of the 1980s expansion. Managerial and professional workers have witnessed increases in job insecurity, while there is no statistical trend apparent in other detailed occupations. Increased anxiety appears in manufacturing, services, and government, while construction workers, who have traditionally had a high probability of job loss because of the seasonal nature of the work, have seen an increase in job security during the 1990s.

Lastly, we looked at the perceptions of job loss among a few nonstandard groups of workers. Table 4 reports the coefficients, standard errors, and marginal effects from additional variables that are asked (sometimes periodically) in the GSS or are computed from other data sources. The first group of variables is other work characteristics, including union membership, the size of the employee's work site, whether the firm pays fringe benefits, whether the organization has gone through a merger or reorganization during the last five years, computer usage in the industry, and employment conditions in the industry and region. The second group of variables lists several hardships that individuals have recently experienced, including poverty, unemployment, work problems, financial problems, other hardships, and health problems. The letter at the beginning of each row indicates a separate regression that includes all of the demographic and employment variables reported in table 3. The sample sizes from each of these regressions are reported in column 1. Since some questions begin to be asked after 1977, all marginal effects are calculated relative to the same base case person in the first year that the question is included in the GSS.

The first row shows that union members are likely to be more insecure about their future job prospects than nonunion members even after controlling for compositional differences in occupation and industry between the groups. This finding may be confounded by the choice to join a union (workers who are more insecure about their future employment are more likely to join unions) and therefore suffers from what econometricians call endogeneity bias. It is a bit surprising, because much of the research on unions suggests that union workers are less sensitive to business cycles because wages and employment are set in multiyear contracts. However, union wages have been growing slower than nonunion wages recently, suggesting that workers are concerned enough about job security that they are willing to trade off wage growth for more security. Furthermore, the decline in union membership over the last few decades could signal reduced bargaining power of union employees.

In a way, the union results are similar to some limited evidence on workers in organizations undergoing change. In 1991, the GSS included information on whether a respondent's firm has gone through a merger or a reorganization. Because we only have one year of data, the precision of the point estimates is low. Nevertheless, the magnitude of the marginal effects is consistent with stories about restructuring and downsizing leading to more insecurity during the 1990s. Unfortunately, because the sample is redrawn each year, we cannot test whether these workers have faced greater job loss frequencies in subsequent years.

In regressions that add the size of the employee's work site, the results suggest that those who work at smaller sites are less likely to be concerned about their job. However, the question asks the size of the work site not the size of the organization.

TABLE 4

Effect of other variables on likelihood of losing your current job in the next year: Ordered probit analysis

				Marginal effect on base case probability			
Other variables	Sample size	Coefficient	Standard error	Not at all likely	Not too likely	Fairly likely	Very likely
Work characteristics							
a. Union member	4,761	–0.124	0.049*	–0.040	0.025	0.008	0.007
b. Current organization merged	550	–0.209	0.132	–0.080	0.039	0.019	0.022
c. Current organization reorganized	552	–0.123	0.120	–0.046	0.024	0.011	0.012
d. No fringe benefits	463	–0.491	0.196*	–0.193	0.077	0.042	0.074
Size of work site							
e. 1–9 employees	3,079	0.158	0.079*	0.060	–0.032	–0.013	–0.015
e. 10–49 employees	3,079	0.107	0.078	0.041	0.021	–0.009	–0.011
e. 100–499 employees	3,079	0.076	0.080	0.029	–0.015	–0.006	–0.008
e. 500+ employees	3,079	0.031	0.081	0.012	–0.006	–0.003	–0.003
f. Region unemployment rate[a]	9,935	–0.086	0.019*	–0.026	0.016	0.005	0.005
g. Industry unemployment rate[a]	9,935	–0.029	0.020	–0.010	0.006	0.002	0.002
h. Industry computer use[b]	9,529	–0.193	0.078*	–0.070	0.039	0.015	0.015
Work and other problems							
i. Below the poverty line	7,620	–0.352	0.051*	–0.132	0.066	0.031	0.035
j. Unemployment spell in last 5 yrs.	3,946	–0.448	0.045*	–0.135	0.090	0.026	0.020
k. Unemployment spell in last 10 yrs.	5,752	–0.403	0.035*	–0.135	0.083	0.029	0.023
l. Problems at work	281	0.004	0.184	0.001	–0.000	–0.000	–0.000
l. Financial problems	281	–0.124	0.216	–0.043	0.028	0.010	0.006
l. Other hardships	281	–0.535	0.219*	–0.201	0.110	0.052	0.039
m. Not healthy	5,292	–0.229	0.047*	–0.087	0.043	0.020	0.025

* = significant at the 5% level.

[a]Industry and regional unemployment rates are calculated from the March 1977–96 *Current Population Survey*.

[b]Industry computer use is from Autor, Katz, and Krueger (1997). They calculate the share of computer users by three-digit Standard Industrial Classification codes from the October 1984, 1989, and 1993 *Current Population Survey*. The computer use data are linearly interpolated between 1984 and 1993, set at 1984 levels in year prior to 1984, and at 1993 levels in years post 1983.

Notes: The letter at the beginning of each row indicates a separate regression that includes all of the control variables listed in table 3. The base case is the same as table 3. If the variables listed in this table were not reported in 1977, the base case is the first year the question was asked. Except for the unemployment rates and the computer usage variable, the marginal effects are reported as the difference between the base case where the characteristic is not present (say, the respondent is not a union member) and where the characteristic is present (is a union member). The size of work site coefficients are relative to a company with 50–99 employees. See box 2. The base case probabilities, which are not reported, are available upon request.

Source: Authors' calculations based on data from the National Opinion Research Center, *General Social Survey*, 1997.

When we add industry and regional unemployment rates to the statistical model, we find, not surprisingly, that workers in regions and, to a much smaller extent, industries that are experiencing higher unemployment are less secure about their own prospects. (The unemployment rates are calculated from the March CPS.)

Finally, we add a variable that measures the share of computer users in each individual's three-digit SIC industry. The data are compiled by Autor, Katz, and Krueger (1997) using the October 1984, 1989, and 1993 CPS. As part of the *Education Supplements*, the three CPS surveys asked workers whether they used a computer at work, where a computer is defined as a desktop terminal or PC and not a hand-held data device or electronic cash register. We interpolate computer shares by industry between the 1984 and 1993 end dates and hold years before 1984 and after 1993 constant at the 1984 and 1993 levels. Surprisingly, the results suggest that workers in industries that are more computer intensive are less secure about their jobs, after controlling for demographics, time, industry, and occupation.[23] When the computer usage variable is interacted with the time dummies, it becomes apparent that this computer industry–job insecurity correlation is driven by the 1993 to 1996 period. Prior to the 1990s, there is a positive relationship between working in a computer-intensive industry and job security. Unfortunately, we have no data on whether the individual respondents are computer users.

The bottom of table 4 reports the parameters from the "problem" variables. Many of these coefficients are significant and negative, suggesting that job insecurity goes hand-in-hand with other work- and non-work-related problems. Furthermore, the large effects from previous unemployment suggest that these workers are more prone to insecurity than those who have not experienced a spell of unemployment in the last ten years. This result suggests that past job loss may be a reasonable indicator of future anxiety and is consistent with studies that use displacement rates as an indicator of job security. However, the results on past unemployment may be driven by unobserved characteristics, such as ability.

Job security and wage growth

While a number of papers have measured recent trends in job security or stability, none that we are aware of attempt to link these trends to wage growth. Yet the allegedly slow rise in compensation during this tight labor market expansion is one of the driving forces behind public policy concerns about job security. Many analysts argue that workers have sacrificed wage growth for a more secure relationship with their current employers.

To investigate this question, we follow an estimation strategy pursued in Blanchflower and Oswald (1994), among others. In particular, we look cross-sectionally as well as over time and ask whether census regions that have had higher displacement rates or worker perceptions of insecurity have tended to have lower wage growth. The regressions that we use are similar to the original Phillips curve, which posits a negative relationship between the rate of wage change and the contemporaneous unemployment rate.[24]

We use three wage measures, annual, weekly, and hourly earnings, that are computed from the 1977 to 1996 March CPS. However, our preferred wage measure is hourly earnings, because it does not confound changes in wages with changes in hours worked. This is an important distinction because annual hours are highly correlated with the job security measures (as well as unemployment rates). Therefore, we have to be careful to distinguish a wage effect from a labor supply (hours) effect. This is probably less of a concern with the weekly measure.

We include controls for one of two job security measures (a security index calculated from the GSS and a displacement rate calculated from the DWS), the contemporaneous unemployment rate (calculated from the March CPS), and time- and location-specific indicator variables. The time- and location-specific variables account for unexplained characteristics of wages that are common across time and regions (essentially, they allow the intercept term to vary over time and region). For example, the time variables will account for changes in productivity growth and expectations of inflation that are common across regions within the U.S. Box 3 gives the technical details of our estimation procedure.

Table 5 highlights some of our findings. Panels A and B are from separate regressions. The security measure is the insecurity index in panel A, and the log displacement rate in panel B.

The wage effect is reasonably consistent across the two job security measures. The coefficients (since the variables are measured in logs, the coefficients are elasticities) suggest that, using the annual earnings measure, a 10 percent increase

BOX 3

The impact of job security on wage growth

The wage-job security relationship is estimated from a regression of the form:

$$4) \qquad w_{rt} = \alpha y_{rt} + \lambda U_{rt} + \varphi w_{rt-1} + v_r + v_t + v_{rt},$$

where w_{rt} is the log wage, y_{rt} is a measure of the level of job security, and U_{rt} is the log unemployment rate. The variables are aggregated into a market r at time t. We aggregate individuals into the nine census regions since geographical labor markets smaller than regions are not available in the GSS. The wage and unemployment measures are computed from the 1977 to 1996 March *Current Population Surveys* (CPS). The annual earnings measure is a sum of all income earned in the previous year. The weekly earnings measure is calculated as annual earnings divided by the number of weeks worked in the previous year. The hourly earnings measure is calculated as annual earnings divided by the number of weeks worked in the previous year times the number of hours worked per week in the previous year.

There are several ways to estimate equation 4. Perhaps the simplest way is to average all individuals in market r at time t and use the cell means as the observation unit. This is essentially what was done for the displacement rates (see box 1). However, the standard errors will be biased downward because common unmeasured factors of individuals may be attributed to local employment conditions (Moulton, 1990). Instead, for the job loss likelihood index, we use a "two-step" procedure. In the first step, we estimate ordered probit regressions like equation 1 in box 2 but augment them to allow a calculation of a region-specific security index for each year. In particular, we regress the loss likelihood responses, y_i^*, on demographics (X_{it}), and year interacted with region dummies ($R_{it}v_t$):

$$y_{it}^* = \beta X_{it} + \delta R_{it} v_t + \varepsilon_{it}.$$

The vector of region–year dummy coefficients (δ) is equivalent to the mean residuals by year and region and can be interpreted as indexes of job insecurity, after controlling for differences in education, gender, age, income, and marital status of workers in particular areas. We also run the final wage equations with job security indexes that are not demographically adjusted and find that this adjustment does not make a significant difference to the inferences.

The wage variables are estimated from a log wage equation of the form:

$$\ln w_{irt} = \beta X_{irt} + \mu_{irt}.$$

Again, the X matrix controls for education, marital status, and other standard human capital controls. We use the mean residual by region and year ($\bar{\mu}_{rt}$) as a measure of the wage adjusted for these demographics.

In the second step, we estimate ordinary least squares regressions of the mean residual from the first stage wage equation ($\bar{\mu}_{rt}$) on the contemporaneous unemployment rate (U_{rt}), the lagged dependent variable, the security index (δ_{rt}), and region and year indicator variables (or fixed effects):

$$\bar{\mu}_{rt} = \alpha \delta_{rt} + \lambda U_{rt} + \varphi \bar{\mu}_{rt-1} + v_r + v_t + v_{rt}.$$

The sample size varies depending on the job security index that is used. The regressions that include the displacement rate are run on 17 years (17 years times nine regions equals 153 observations) and the regressions that include the security index are run on 13 years (117 observations).

in the job security measure results in a 0.2 percent decline in wage growth. This is statistically significant at conventional levels. However, the hourly wage coefficient implies that about half of this decrease is a wage effect and the other half is an hours-worked effect. Furthermore, the wage effect is imprecisely enough estimated that we cannot reject the hypothesis that the true effect is zero. However, if this effect is real, the impact on nominal wage growth during the 1990s has been fairly large. Referring to figure 3, displacement rates rose from around 2.0 percent in the 1980s

TABLE 5

Relationship between wages and job security

	Change in log annual earnings		Change in log weekly earnings		Change in log hourly earnings	
	No region controls	Region controls	No region controls	Region controls	No region controls	Region controls
A. (sample size = 117)						
Insecurity index	–0.018	–0.010	–0.013	–0.007	–0.013	–0.008
	(0.012)	(0.013)	(0.012)	(0.013)	(0.010)	(0.011)
Log unemployment rate	–0.029	–0.048	–0.029	–0.045	–0.024	–0.040
	(0.012)	(0.015)	(0.012)	(0.015)	(0.010)	(0.013)
B. (sample size = 153)						
Log displacement rate	–0.021	–0.018	–0.020	–0.017	–0.010	–0.008
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Log unemployment rate	–0.028	–0.052	–0.028	–0.051	–0.026	–0.044
	(0.009)	(0.011)	(0.009)	(0.012)	(0.008)	(0.011)

Notes: All regressions include year controls. See text for explanation of variables and sample. The unit of observation is the nine census regions.

Sources: Authors' calculations based on data from the Bureau of Labor Statistics, *Current Population Survey*, 1996; the Bureau of Labor Statistics, *Displaced Worker Survey*, 1984–96; and the National Opinion Research Center, *General Social Survey*, 1997.

to 2.75 percent in the early 1990s to almost 3.5 percent in 1995. Using a job insecurity wage elasticity estimate of –0.01, this suggests that job insecurity lowered wage growth by 0.3 percentage points a year in the early 1990s and roughly 0.7 percentage points in 1995, relative to what would have happened if displacement rates had stayed at the 1980s level. The job anxiety index grew approximately 25 percent during the 1990s, suggesting a 0.3 percentage point decline in wages per year from the results in table 5, panel A. However, our estimates for hourly wages cannot reject the possibility that these effects arose purely by chance.

Our analysis is just a first step in estimating the impact of job security on wage inflation. There is much more work to be done on this question. First, we plan to explore micro-data-based techniques to solve technical problems associated with having two measures, such as wage growth and job security, that are jointly determined. Second, as it is currently measured, the security index encompasses a fair amount of noise or measurement error. This measurement error leads to a downward bias in the wage–security relationship. Finally, a key question is causation. Does high job security cause high wages or vice versa? This question could be examined by estimating vector autoregressive models, which allow a flexible relationship between wages, unemployment, and job security.

Conclusion

Our review of the literature and our new results on displacement for high-tenure workers reveal a modest decline in job stability and a larger decline in job security, especially for workers with higher levels of job tenure. Apparently, some of the increases in displacement that have been observed in the 1990s have been offset by declines in quit rates. The higher displacement rates suggest that workers have more reason to be worried about their job security in the 1990s, and the lower quit rates suggest they may be less confident about their job prospects. Consistent with these findings, our tabulations of workers' evaluations of their chances of job loss reveal a noticeable increase in the proportion of workers who feel that they are at least at some risk of job loss.

When we relate variations in displacement rates and anxiety levels over time and across census divisions to the corresponding variations in wage growth, we find estimates of the effect of insecurity on wages that would be large enough to explain all or most of the puzzle of slow wage growth in the 1990s. Of course, these estimates are rather imprecise and may even have arisen by chance. Still, we believe that these results add to the case for worker insecurity having restrained wage growth and justify further research on the topic.

NOTES

[1]Greenspan (1997).

[2]Reich (1997).

[3]See, for example, Staiger, Stock, and Watson (1997).

[4]The manner in which workers and employers are matched to each other has changed quite noticeably in the 1990s. The process may have been made more efficient by the rapid expansion of the temporary services industry. (See Segal and Sullivan, 1995, 1997). Also, Internet job postings may make interregional job search more efficient. Such developments may reduce the likelihood of bottlenecks and spot labor shortages that contribute to inflationary pressures. Or, in the language of the short-run Phillips curve, we can argue that they have reduced the natural rate of unemployment independently of any increase in worker anxiety.

[5]*New York Times* (1996).

[6]Notable contributions to this literature include Podursky and Swaim (1987), Kletzer (1989), Topel (1990), Ruhm (1991), and Jacobson et al. (1993a, 1993b, 1993c). Fallick (1996) and Kletzer (1997) provide recent surveys of this literature.

[7]The CPS is a monthly mini-census of about 45,000 households that is the source for such familiar statistics as the unemployment rate. When appropriately weighted to account for the scientifically designed sampling procedures employed by the BLS, the CPS yields nationally representative estimates.

[8]The figures are taken from U.S. Department of Labor (1997) and Farber (1998).

[9]Tenure data were also collected as part of the CPS *Pension and Benefits Supplements* of May 1979 and April 1993. The latter, which found higher median tenure for most age and sex groups than the *Mobility Supplement* of January 1991, supported the conclusion of Farber (1998) that job durations were relatively stable in the 1980s and 1990s. However, the tenure data from the *Pension and Benefits Supplements* are based on a slightly different question than those from the *Mobility Supplements.* The latter asks how long workers have been *continuously* employed by their current employers, while the former simply asks how long workers have been employed by their current employers. Omitting the condition that the employment be continuous could raise the tenure estimates. Suppose a worker was employed by a firm for five years, left for two years, and then returned for another five. In the *Mobility Supplements,* in which the question refers to continuous employment, the worker is likely to report a tenure of five years. However, in the *Pension and Benefits Supplements,* in which the question simply asks workers how long they have worked for their employers, the worker is likely to report a tenure of ten years. Thus, it is possible that some or all of the higher tenure reading in the 1993 survey was due to the omission of the word continuous in the key question rather than an actual increase in job stability. Changes in question wording also complicate the interpretation of the trends in figure 2. Before 1983, the *Mobility Supplements* asked workers when they started working for their current employers. Tenure was then calculated based on workers' responses. Since 1983, workers have been asked how long they have continuously worked for their current employers, which yields the tenure information directly. Of course, if workers correctly answered all questions, it would make no difference whether tenure was solicited directly or calculated from the start date of their jobs. But workers do not always report accurately; figure 1 shows that workers have a tendency to report tenures that are multiples of five years. In the earlier *Mobility Supplements*, there was a tendency to report start years that were multiples of five, such as 1960, 1965, and so on. This change compromises the comparability of the data over time.

[10]Hall (1982) studied retention probabilities using data from a single cohort which, as shown by Ureta (1992), requires a stable rate of job beginnings, as well as a stable set of retention probabilities.

[11]Neumark, Polsky, and Hansen (1997) report several alternative estimates of retention probabilities. Those shown in table 1 are, we believe, their preferred estimates.

[12]The lack of a major decline in job stability is also consistent with the work of Stewart (1997), who analyzed the March CPS annual demographic files and found no increase in the rate of job change from the previous calendar year.

[13]See Hipple (1997).

[14]Of course, since this is a survey of individuals, there may be instances in which respondents misreport by saying, for example, that they were displaced when they were actually fired. Such mismeasurement is possible with any household survey.

[15]The PSID has a number of advantages for studying turnover and displacement. Unfortunately, the sample size is too small to estimate disaggregated rates. Thus, Farber computes a single set of adjustment factors that he applies to all workers.

[16]Tabulations, such as those in Hipple (1997), that do not count workers displaced for nonstandard reasons do not show an increase in the current period comparable to what we find below.

[17]Our procedure yields what we believe to be consistent comparisons across years. It is possible, however, that the overall level of the estimates could be off by some constant percentage. Suppose, for example, that the reason the displacement rates decline by 11 percent for each additional year that the survey lags the year of displacement is not that the rates measured with a one-year lag are correct and the other years reflect forgetting, but that the rates measured five years later are correct and earlier surveys reflect "spurious remembering." (In our opinion, this is a much less likely scenario but one that we obviously can't rule out.) Then our estimates will all be too high by about 52 percent (11 percent compounded for four years), but the pattern across years will be unaffected.

[18]Some observers (for example, Neumark and Polsky, 1997) argue that attitudinal questions about job security may not provide convincing evidence of actual job loss if perceptions are formed from misinformation. They point out that much of the reporting on job security relies on anecdotal evidence and, therefore, is not based on random sampling. It is always possible to find someone who is struggling, even in a booming economy. During the early 1990s, the recession hit journalists and editors, as well as other white-collar workers especially hard, perhaps resulting in more stories about displacement than were warranted. Since press reports may help form perceptions of the chance of job loss among readers, there is the danger that we might observe an increase in perceptions of job insecurity that has little to do with actual job loss.

[19]An exception is Schmidt and Thompson (1997). Several polling agencies, such as Gallup and Yankelovich, and survey organizations, such as the University of Michigan Survey Research Center (SRC), have also been soliciting perceptions of worker security over the last two decades. See Otoo (1997) for an analysis of the SRC data. Dominitz and Manski (1997) describe the new *Survey of Economic Expectations*, which asks respondents their level of concern about losing their job, losing part of their income, losing their health insurance, and being victimized by a burglary. However, this survey began in 1994 and therefore provides no information on longer-term trends in job security.

[20]See Dominitz and Manski (1997) for a criticism of the wording of the GSS questions.

[21]The Gallup poll also found that the number of workers answering "not at all likely" to lose their job decreased during the 1990s. On the other hand, Yankelovich, which asks whether losing your job worries you, found little change in response between 1992 and 1995. But both polls found some job security differences across the 1990s by education or income. Likewise, Otoo (1997) found significant increases in job anxiety between 1988 and 1995 using the SRC micro data.

[22]Additional results, including an analogous table for the comparable job question, are available from the authors upon request.

[23]With no such controls, the correlation between industry computer usage and job security is positive and highly significant.

[24]Blanchflower and Oswald estimate models that relate the *level* of wages to unemployment. But using wage levels implies that the coefficient on a lagged wage variable is less than one, whereas most of the literature (see Card, 1995, and Blanchard and Katz, 1997) has found that it is close to one. When we ran level wage equations, we also found the coefficient on the lagged dependent variable to be one, suggesting that the relationship is really between unemployment and security and changes in wages. The change specification also avoids technical problems associated with having a lagged dependent variable on the right-hand side with a regional fixed effect and serial correlation in the error term.

REFERENCES

Autor, David, Lawrence Katz, and Alan Krueger, 1997, "Computing inequality: Have computers changed the labor market?," National Bureau of Economic Research, working paper.

Blanchard, Olivier, and Lawrence Katz, 1997, "What we know and do not know about the natural rate of unemployment," *Journal of Economic Perspectives*, Winter, pp. 51–72.

Blanchflower, David, and Andrew Oswald, 1994, *The Wage Curve*, Cambridge, MA: The MIT Press.

Boisjoly, Johanne, Greg J. Duncan, and Timothy Smeeding, 1994, "Have highly-skilled workers fallen from grace? The shifting burdens of involuntary job losses from 1968 to 1992," Northwestern University, working paper.

Card, David, 1995, "The wage curve: A review," *Journal of Economic Literature*, June, pp. 785–799.

Diebold, Francis, David Neumark, and Daniel Polsky, 1997, "Job stability in the United States," *Journal of Labor Economics*, Vol. 15, No. 2, pp. 206–233.

_________, 1996, "Comment on Kenneth A. Swinnerton and Howard Wial, 'Is job stability declining in the U.S. economy?'," *Industrial and Labor Relations Review*, Vol. 49, No. 2, pp. 348–352.

Dominitz, Jeffrey, and Charles Manski, 1997, "Perceptions of economic insecurity: Evidence from the *Survey of Economic Expectations*," *Public Opinion Quarterly*, Vol. 61, No. 2, Summer, pp. 261–287.

Fairlie, Robert, and Lori Kletzer, 1997, "Jobs lost, jobs regained: An analysis of black / white differences in job displacement in the 1980s," University of California, Santa Cruz, working paper.

Fallick, Bruce, 1996, "A review of the recent empirical literature on displaced workers," *Industrial and Labor Relations Review*, Vol. 50, No. 1, pp. 5–16.

Farber, Henry, 1998, "Are lifetime jobs disappearing: Job duration in the United States 1973–93," *Labor Statistics Measurement Issues*, John Haltiwanger, Marilyn Manser, and Robert Topel (eds.), Chicago: University of Chicago Press, forthcoming.

_________, 1997a, "The changing face of job loss in the United States, 1981–95," *Brookings Papers on Economic Activity: Microeconomics*.

_________, 1997b, "Trends in long term employment in the United States, 1979–96," paper prepared for the third public German-American Academic Council Federation Symposium, "Labor Markets in the USA and Germany," held at Bonn, Germany, June 10–11.

Gordon, Robert, 1997, "The time-varying NAIRU and its implications for economic policy," *Journal of Economic Perspectives*, Vol. 11, Winter, pp. 11–32.

Greenspan, Alan, 1997, "Monetary policy," testimony and report before the U.S. House Committee on Banking, Housing, and Urban Affairs, 104th Congress, 1st session, February 26.

Hall, Robert, 1982, "The importance of lifetime jobs in the U.S. economy," *American Economic Review*, Vol. 72, pp. 716–724.

Hipple, Steven, 1997, "Worker displacement in an expanding economy," *Monthly Labor Review*, Vol. 120, No. 12, December.

Jacobson, Louis, Robert LaLonde, and Daniel Sullivan, 1993a, "Earnings losses of displaced workers," *American Economic Review*, Vol. 72, September, pp. 685–709.

_________, 1993b, *The Costs of Worker Dislocation*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

_________, 1993c, "Earnings losses of high seniority displaced workers," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 17, No. 6, November / December, pp. 2–20.

Jaeger, David, and Ann Huff Stevens, 1997, "Is job stability in the United States falling? Trends in the *Current Population Survey* and *Panel Study of Income Dynamics*," Yale University, working paper.

Kletzer, Lori, 1998, "Job displacement: What do we know, what should we know?," *Journal of Economic Perspectives*, forthcoming.

_________, 1989, "Returns to seniority after permanent job loss," *American Economic Review*, Vol. 79, June, pp. 536–543.

Maddala, G. S., 1983, *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

Manski, C., 1993, "Adolescent econometricians: How do youth infer the returns to schooling?," in *Studies of Supply and Demand in Higher Education*, C. Clotfelder and M. Rothschild (eds.), Chicago: University of Chicago Press.

_________, 1990, "The use of intentions data to predict behavior: A best case analysis," *Journal of the American Statistical Association*, Vol. 85, pp. 934–940

Marcotte, David, 1996, "Has job stability declined? Evidence from the *Panel Study of Income Dynamics*," Northern Illinois University, working paper.

Moulton, Brent, 1990, "An illustration of a pitfall in estimating the effects of aggregate variables on micro units," *Review of Economics and Statistics*, Vol. 72, No. 2, May, pp. 334–338.

National Opinion Research Center, 1997, *General Social Survey*, retrieved from the Internet at www.icpsr.umich.edu / gss, September.

Neumark, David, and Daniel Polsky, 1997, "Changes in job stability and job security: Anecdotes and evidence," Michigan State University, working paper.

Neumark, David, Daniel Polsky, and Daniel Hansen, 1997, "Has job stability declined yet? New evidence for the 1990s," paper prepared for February 1998 Russell Sage Foundation Conference on Changes in Job Stability and Job Security.

New York Times, 1996, *The Downsizing of America*, New York: Random House.

Otoo, Maria Ward, 1997, "The sources of worker anxiety: Evidence from the Michigan survey," Federal Reserve Board of Governors, working paper.

Podursky, Michael, and Paul Swaim, 1987, "Job displacement earnings loss: Evidence from the *Displaced Worker Survey*," *Industrial and Labor Relations Review*, Vol. 41, October, pp. 17–29.

Reich, Robert, 1997, *Locked in the Cabinet*, New York: Alfred Knopf.

Rose, Stephen, 1995, "Declining job security and the professionalization of opportunity," National Commission for Employment Policy, research report, No. 95-04.

Ruhm, Christopher, 1991, "Are workers permanently scarred by job displacements?," *American Economic Review*, Vol. 81, No. 1, pp. 319–324.

Schmidt, Stefanie, and Christopher Thompson, 1997, "Have workers' beliefs about job security been keeping wage inflation low? Evidence from public opinion data," Milken Institute, working paper.

Segal, Lewis, and Daniel Sullivan, 1997, "The growth of temporary services work," *Journal of Economic Perspectives*, Vol. 11, Spring, pp. 117–136.

__________, 1995, "The temporary work force," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 19, No. 2, March / April, pp. 2–19.

Staiger, Douglas, James Stock, and Mark Watson, 1997, "The NAIRU, unemployment, and monetary policy," *Journal of Economic Perspectives*, Vol. 11, Winter, pp. 33–49.

Stewart, Jay, 1997, "Has job stability increased? Evidence from the *Current Population Survey*: 1975–1995," U.S. Department of Labor, Bureau of Labor Statistics, working paper.

Swinnerton, Kenneth, and Howard Wial, 1996, "Is job stability declining in the U.S. economy? Reply to Diebold, Neumark, and Polsky," *Industrial and Labor Relations Review*, Vol. 49, No. 2, January, pp. 352–355.

__________, 1995, "Is job stability declining in the U.S. economy?," *Industrial and Labor Relations Review*, Vol. 48, No. 2, January, pp. 293–304.

Topel, Robert, 1990, "Specific capital and unemployment: Measuring the costs and consequences of job loss," in *Carnegie Rochester Conference Series on Public Policy*, Vol. 33, Allan H. Meltzer and Charles I. Plosser (eds.), Amsterdam: North-Holland, pp. 181–214.

U.S. Department of Commerce, Bureau of the Census, 1996, *Current Population Survey*, Washington, DC, February.

U.S. Department of Labor, 1997, "Employee tenure in the mid-1990s," Washington, DC, press release, January 30.

U.S. Department of Labor, Bureau of Labor Statistics, 1984–96, *Displaced Worker Survey*, Washington, DC.

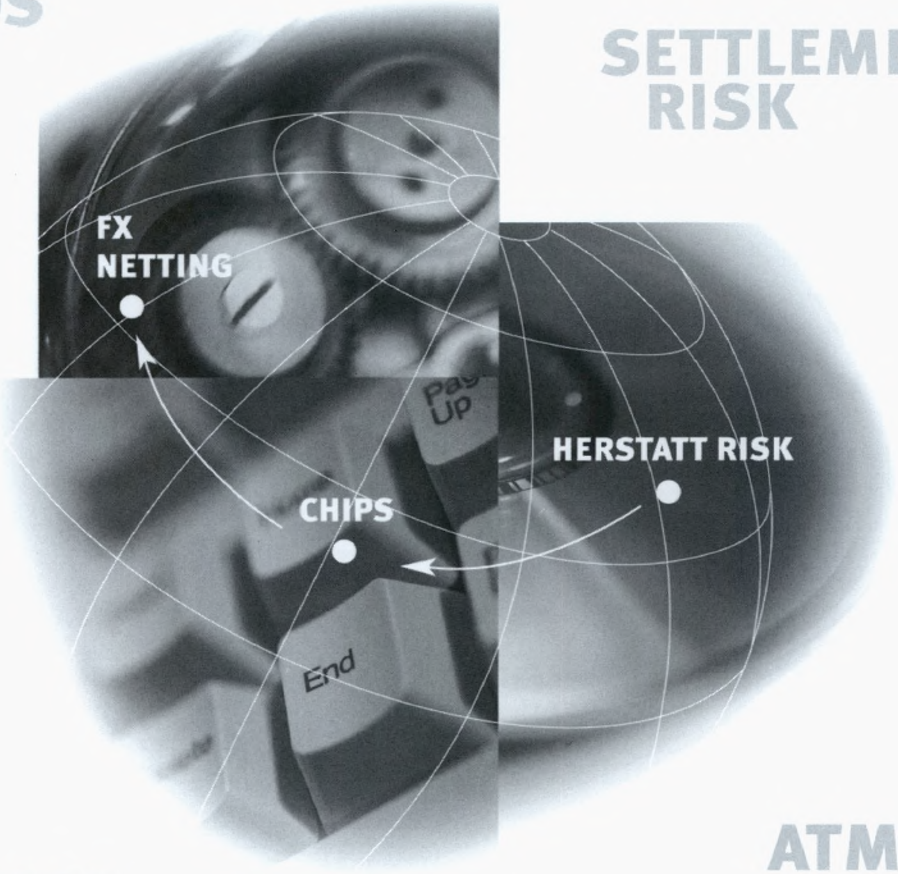
Ureta, Manuelita, 1992, "The importance of lifetime jobs in the U.S. economy, revisited," *American Economic Review*, Vol. 82, March, pp. 322–335.

Valletta, Robert, 1997, "Declining job security," Federal Reserve Bank of San Francisco, working paper.

DEBIT CARDS

IMAGING

SETTLEMENT RISK



ACH

ATM

RTGS

NETWORK ECONOMIES

Payments Systems in the Global Economy: Risks and Opportunities

34th Annual Conference on Bank Structure and Competition

May 6-8, 1998

On May 6–8, 1998, the Federal Reserve Bank of Chicago will hold its 34th annual Conference on Bank Structure and Competition at the Westin Hotel in Chicago. The theme of the conference will be an evaluation of developments in payments systems.

As payments technology and the structure of the financial services industry are changing, there is a need to evaluate how the payments system is evolving and what public policy issues are raised by this evolution. For example, what is the current state of payments activity and the resulting levels of risk? What are the different problems for paper-based, small-value electronic, and large-value electronic payments systems? Why do some industry participants argue for the expansion of real-time gross settlement arrangements, while others emphasize the efficiency of netting arrangements? What is the evidence of systemic risk in the interbank market? Is there a need for regulatory oversight of this market? Is this risk growing with the increase in cross-border transactions? To prevent such problems, is there a need for international coordination by supervisory agencies? How can liquidity constraints best be addressed in the settlement of international transactions? Do the appropriate means to address these issues differ between developed and transition economies? In the U.S., the Federal Reserve has recently undertaken a thorough analysis of the appropriate role the central bank should have in payments (the Rivlin Committee). What arrangements exist internationally? Similarly, why are there significant differences in payment modes across countries?

The 1998 conference will feature discussions of these and related issues by some of the most prominent members of the financial services industry, both domestic and international. This elite group will include Alan Greenspan, Chairman of the Board of Governors of the Federal Reserve System; Andrew Crockett, General Manager of the Bank for International Settlements; and Edward E. Crutchfield, Chairman and Chief Executive Officer, First Union Corp.

The theme panel will focus on recent development in payments activity including alternative settlement mechanisms for both domestic and international transactions, the role of the central bank in payments and payment initiatives taken by industry participants. The panel will feature Catherine A. Allen, Chief Executive Officer, Bankers Roundtable's Banking Industry Technology Secretariat; Alice Rivlin, Vice Chair, Board of Governors of the Federal Reserve System; Jill Considine, President, New York Clearing House and CHIPS; David Roscoe, CLS Services; and Martin Mayer, Brookings Institution.

The conference will also include sessions on the following topics:

- Retail Payment Developments
- International Comparison of Payment Systems
- Financial Instability
- Credit Access
- The Impact of Consolidation on Lending
- The Federal Safety Net and the Role of Firewalls
- Risk Management and Default
- Liquidity Constraints
- Supervisory Examination Information and Market Discipline

Invitations to the conference will be mailed in March. If you are not currently on the conference mailing list or have changed your address and would like to receive an invitation, please contact the Meeting Services Department of the Federal Reserve Bank of Chicago at 312-322-5186 or 322-5641.

**You may also send your request to
Meeting Services
Bank Structure Conference, 12th floor
Federal Reserve Bank of Chicago
P.O. Box 834, Chicago, Illinois 60690-0834
or via E-mail to rlangston@frbchi.org.**

Are international business cycles different under fixed and flexible exchange rate regimes?

Michael A. Kouparitsas

Introduction and summary

By the year's end, Europe will have taken the final step in the most ambitious monetary experiment of the postwar era by establishing a common currency area (the European Monetary Union [EMU]), an extreme form of fixed exchange rate regime in which member countries use the same currency. There is a widespread belief that countries tied to a fixed exchange rate regime are more susceptible to foreign disturbances, particularly monetary disturbances. In other words, there is a belief that flexible exchange rates offer greater insulation from foreign disturbances. A major concern surrounding the EMU and fixed exchange rate regimes, in general, is that business cycles of member countries may become more volatile under a common currency or fixed exchange rate because they are not only subject to domestic shocks but also have increased sensitivity to foreign disturbances.

This conventional view of fixed versus flexible exchange rate regimes stems more from anecdotal evidence than statistical evidence. Two recent events support this view. First is the experience of the United Kingdom (UK) and its continental counterparts in the 1990s. Member countries of the European Exchange Rate Mechanism (ERM), which stayed tied to the German mark (DM) after German reunification, were forced to tighten monetary policy and suffered a severe and persistent economic downturn that is only now abating. The UK chose to leave the ERM in 1992 and devalue against the DM rather than raise domestic interest rates to maintain its currency peg with the DM. Unlike its continental counterparts, the UK experienced a strong recovery in the early 1990s, which has carried through to the present. Second, severe economic downturns in Mexico in 1994 and Asia in 1997 came about because of massive capital outflows and banking collapses that flowed from currency crises involving a U.S. dollar exchange rate peg that was inconsistent with the market's desired level. Looking to the past, monetary historians like Eichengreen (1992) frequently argue that countries that abandoned the gold exchange standard experienced an economic downturn that was far less severe than that of countries which stayed pegged to the United States' currency during the depression of the 1930s.

One empirical observation that seems to be at odds with this view is the emergence of a stronger international business cycle after the abandonment of the fixed exchange rate regime (which had been established by the Bretton Woods agreement in July 1944) in the early 1970s. The key stylized fact supporting this is the observed higher correlations between national output fluctuations of the U.S. and other G7 (Group of Seven) countries in the flexible exchange rate period from 1973 to the present, or the post-Bretton Woods (PBW) period, relative to the Bretton Woods (BW) fixed exchange rate period from 1949 to 1971. This evidence works against the conventional view of fixed versus flexible regimes because cross-country correlations of output fluctuations rise if the importance of global or foreign shocks rises. Moreover, it questions the

Michael A. Kouparitsas is an economist at the Federal Reserve Bank of Chicago. *The author would like to thank Jonathan Siegel for suggesting this topic, useful discussions, and valuable research assistance, Charles Evans for useful discussions and providing his RATS code and various data series, and Hesna Genay and David Marshall for valuable comments on an earlier draft.*

insulation properties of flexible exchange rates over fixed exchange rates. This evidence also suggests that the behavior of international business cycles may be intimately related to the exchange rate regime.

This article offers an exploratory analysis of the link between exchange rate regimes and the behavior of international business cycles. I estimate statistical models of the U.S. and its G7 counterparts over the postwar fixed and flexible exchange rate periods. I use these empirical models to get a better sense of the factors underlying the higher degree of business cycle comovement between the U.S. and the other G7 nations in the PBW period. There are essentially four factors that would lead to higher correlations of U.S. and G7 industrial production: 1) an increase in the volatility of global disturbances (such as oil prices); 2) an increase in the volatility of U.S. disturbances that affect the rest of the G7 and an increase in the volatility of G7 disturbances that affect the U.S.; 3) increased sensitivity to G7 disturbances for the U.S. and increased sensitivity to U.S. disturbances for the rest of the G7; and 4) a change in U.S. and G7 responses to global or foreign disturbances, so that they became more alike.

My empirical results suggest that higher comovement emerged in PBW due to a combination of factors 2 and 4. First, the sensitivity to U.S. monetary policy shocks for the rest of the G7 remained unchanged over the fixed and flexible exchange rate regimes, but the volatility of shocks to U.S. monetary policy increased significantly in the flexible exchange rate period. This made U.S. monetary policy disturbances a more important source of variation for G7 industrial production and, in the process, raised the correlation between U.S. and G7 output fluctuations. Second, the responses of the G7 to all shocks, global and domestic, changed in the flexible regime so that they were more alike than in the fixed exchange rate period. One of the important findings of this study is that G7 sensitivity to foreign and domestic monetary policy shocks remained unchanged over the fixed and flexible exchange rate periods. This result questions conventional wisdom, which argues that flexible exchange rates insulate countries against foreign monetary shocks. It also suggests that the domestic impact of monetary policy is invariant to the exchange rate regime.

Overview of U.S.–G7 exchange rate regimes

In July 1944, representatives from 44 countries met in Bretton Woods, New Hampshire, to draft and sign the Articles of Agreement that established the International Monetary Fund.[1] The system set up by the Bretton Woods agreement called for fixed exchange rates against the U.S. dollar and an unvarying dollar price of gold of $35 an ounce. Member countries held their official international reserves in gold or dollar assets and had the right to sell dollars to the Federal Reserve for gold at the official price. The system was thus a gold exchange standard, with the dollar as its principal reserve currency.

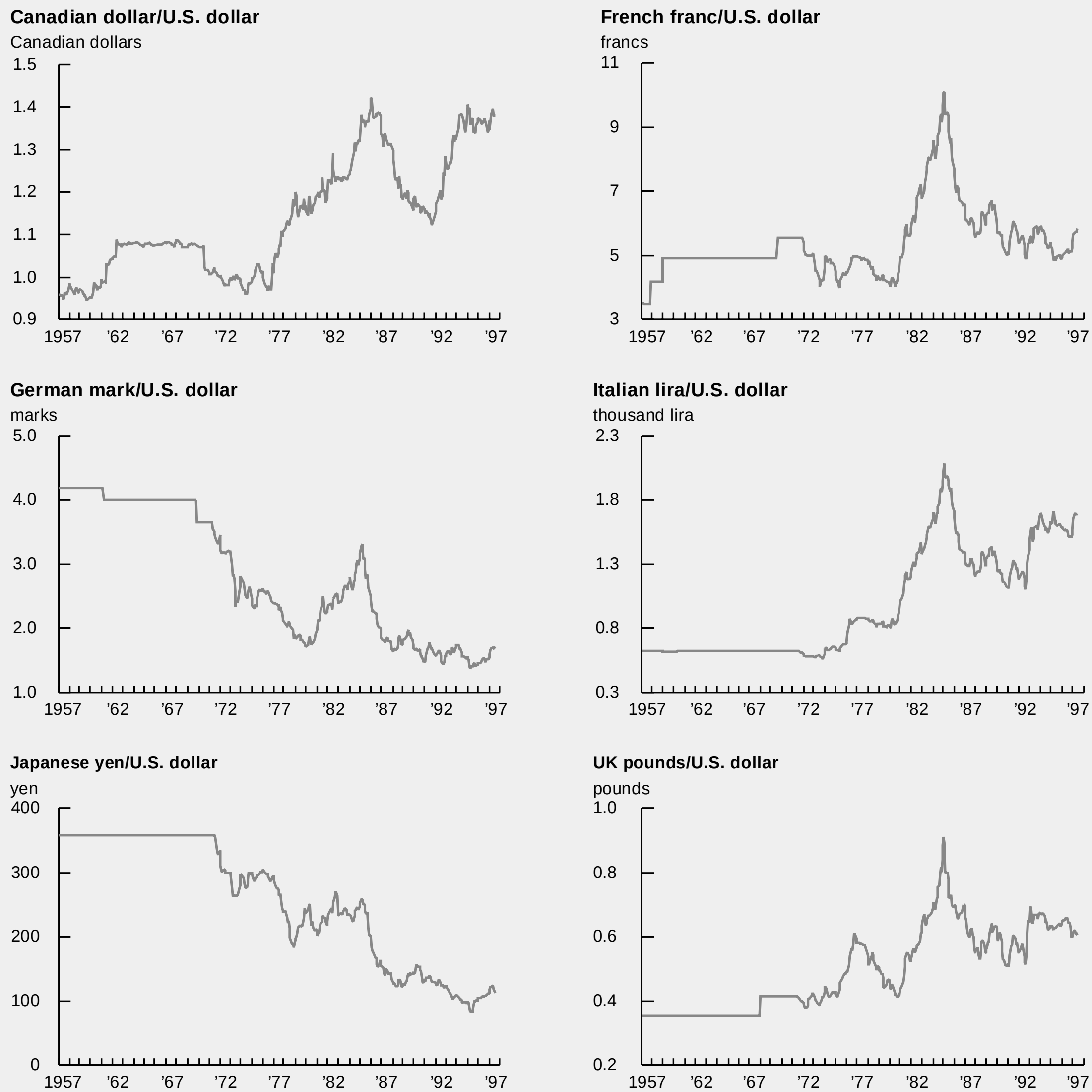
The earliest sign that BW was near collapse came in early 1968 when central bankers announced the creation of a two-tier gold market, with one private tier and the other official. Private traders traded freely on the London gold market and the gold price set there was allowed to fluctuate. In contrast, central banks would continue to transact with others in the official tier at the fixed price of $35 dollars an ounce. This came about because of speculation of a rise in the official gold conversion rate following the British pound's devaluation in November 1967. The gold exchange standard was intended to prevent inflation by tying down gold's dollar price. By severing the link between the supply of dollars and a fixed market price of gold, central bankers had removed the system's built-in safeguard against inflation.

The U.S. experienced a widening current account deficit in early 1971. This set off a massive private purchase of the DM, because most traders expected a revaluation of the DM against the dollar. By August 1971, the markets forced the U.S. to devalue the dollar and suspend gold convertibility with other central banks. Under the Smithsonian agreement in December 1971, the U.S. dollar was devalued roughly 8 percent against all other currencies. An ever-widening U.S. current account deficit led to further speculative attacks against the dollar in February 1973. By March, the U.S. dollar was floating against the currencies of Europe and Japan. This marked the official end of the fixed exchange period for the U.S., although one could argue that the U.S. abandoned fixed exchange rates in August 1971. In my analysis, I treat August 1971 as the end of the fixed exchange rate period and the period following January 1974 as the flexible exchange rate period, because all industrial countries had moved to flexible exchange systems by this date.

Over the last 100 years, the U.S. has participated in nine different exchange rate regimes with other G7 countries.[2] Many of these exchange arrangements emerged because of the disruption to currency markets caused by the two world wars. Exchange rate regimes are generally characterized as either fixed or floating. These labels are misleading as they suggest that fixed or floating regimes are perfectly homogeneous. In a fixed exchange rate system, currencies are pegged to some reserve currency. The pegged currency in the case of BW was the U.S. dollar. Alternatively, floating exchange rate regimes allow currencies to move freely against all currencies. Historically, exchange rate regimes have been somewhere in between these extremes. Figure 1 shows how the foreign currency / U.S. dollar rates of the UK, Germany, France, Italy, Japan, and Canada varied during the BW era. It is clear that the Canadian dollar / U.S. dollar rate was allowed to vary considerably over the period, while the other currencies were allowed large discrete devaluations / revaluations. Similarly, the regimes following BW were not pure floating regimes. What is immediately obvious from figure 1 is that exchange rate movements at all frequencies have been

FIGURE 1

Fixed versus flexible foreign exchange rates



Source: International Monetary Fund.

considerably more volatile under the flexible exchange rate regime.

Analyzing exchange rate regimes and international business cycles

There is a wealth of empirical research documenting the properties of macroeconomic time series from the postwar fixed and flexible exchange rate eras. For example, Baxter and Stockman (1989) investigate the differences in time-series behavior of key economic variables during the BW and PBW periods. Figure 2 shows selected data from Baxter and Stockman. In contrast to Baxter and Stockman, I find that the cross-country correlations of cyclical movements in U.S. and G7 industrial production are considerably higher in the flexible exchange rate period (see upper panel of figure 2).[3,4] The obvious exception is Canada. The correlation between Canadian and U.S. industrial production is roughly constant over the BW and PBW periods. Volatility statistics reported in the lower panel of figure 2 are similar to Baxter and Stockman's in suggesting that industrial production was more volatile in G7 countries in the flexible exchange rate period.

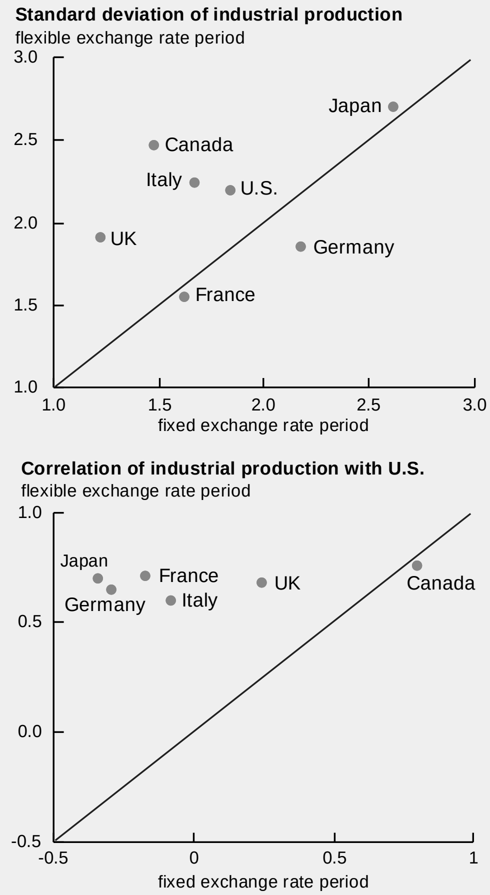
Given the relatively small sample size for the industrial output data, the correlations in the PBW period may be driven by one or two influential data points. I explore this issue in figure 3 by plotting cyclical fluctuations in G7 industrial production series over the fixed and flexible regimes. The low correlation between the U.S. and other G7 country industrial production (excluding Canada) is obvious in the BW period, the period before the solid vertical line. More importantly, the high correlation in the PBW period seems to be linked to the 1973–75 period, which coincides with the first oil price shock, and the 1979–83 period, which coincides with the second oil price shock and the period when the U.S. Federal Reserve experimented with direct targeting of monetary aggregates.

My findings add to similar results in the literature using other empirical techniques, such as frequency domain analysis. For example, Gerlach (1988) and Bowden and Martin (1995) find that the correlation between national output of the U.S. and that of various European countries has increased over so-called business cycle horizons ranging from one and a half years to eight years. Their analysis also suggests that the volatility of national output rose in the flexible period.

As many researchers have noted, work like that of Gerlach, Bowden and Martin, and Baxter and Stockman leaves open the question of whether the increased interdependence observed under the flexible exchange rates is attributable to a change in the response to underlying disturbances (which may have flowed from the move to a flexible exchange rate regime) or the changing nature of the underlying disturbances themselves. This question has been the focus of two different quantitative literatures.

FIGURE 2

International business cycles under fixed and flexible exchange rate systems



Note: Industrial production data filtered using the monthly business cycle band-pass filter described in Baxter and King (1995).

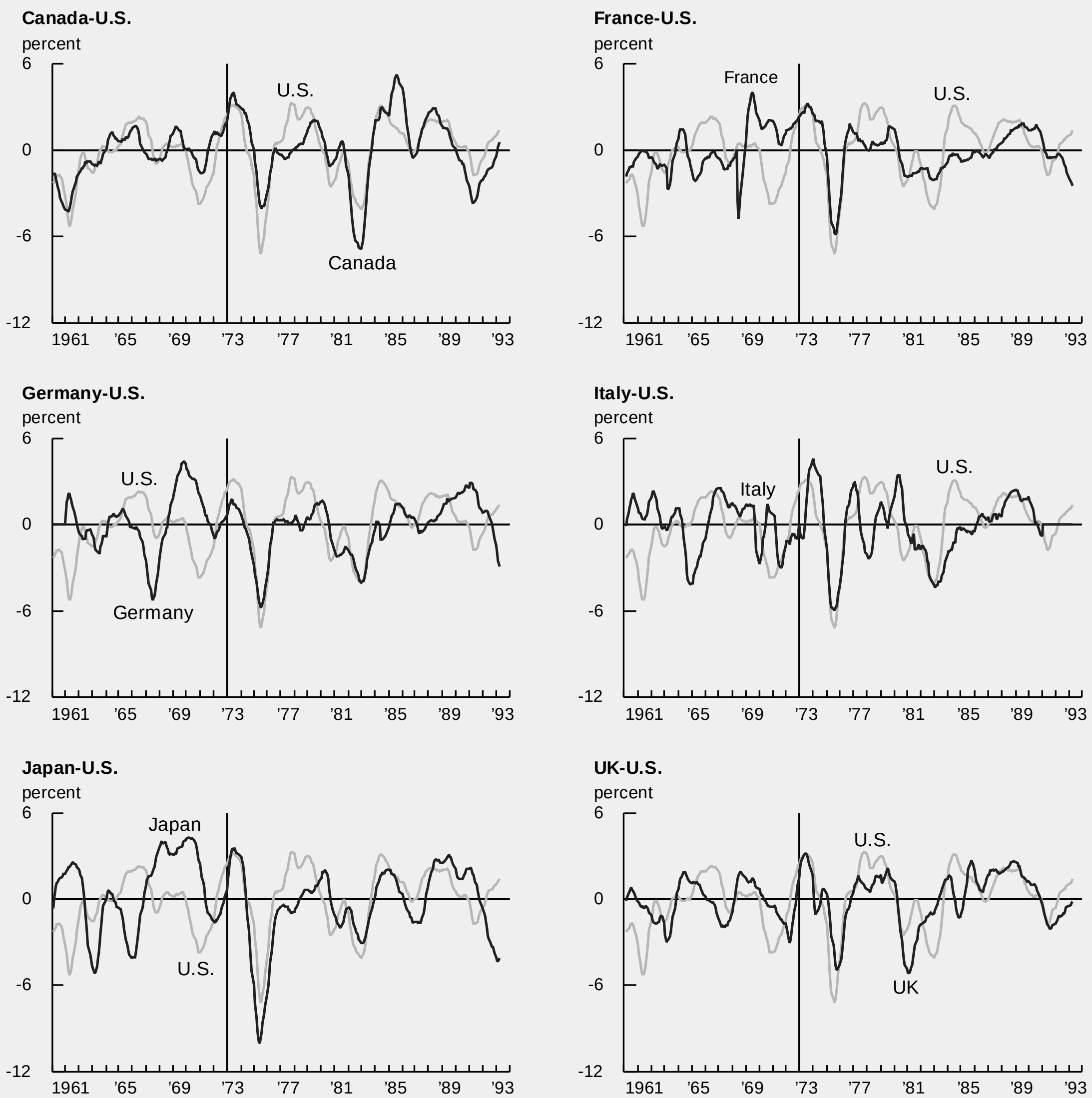
Source: Author's calculations using data from the International Monetary Fund.

Theoretical research on international business cycles

One branch has attempted to explain the international business cycle through quantitative theoretical models of international trade. So far these models are "real" in the sense that there is

FIGURE 3

Cyclical movements of U.S. and G7 industrial production



Note: Industrial production data filtered as in figure 2.
Source: Author's calculations using data from the International Monetary Fund.

no role for monetary disturbances. They completely ignore monetary aspects of the international business cycle by relying wholly on international business cycle transmission through real routes such as goods and asset trade. This literature was recently surveyed by Baxter (1995) and Backus, Kehoe, and Kydland (1995). They report that models that allow for realistic trade in capital are unable to generate international comovement. In contrast, less realistic models that ignore trade in capital goods, such as Stockman and Tesar (1995), have been shown to generate international comovement. Others (including Kouparitsas [1996]) have been successful at explaining positive output correlations between developing and industrial countries by allowing for trade in capital and intermediate goods. Unfortunately, the business cycle transmission mechanisms at work in these industrial and developing country trade models are absent in international trade between industrial countries. This analysis suggests that monetary or nominal factors may be an important component in explaining international business cycles of industrial countries.

Empirical research on international business cycles

Others have approached the issue by studying international business cycles within the context of structural econometric models.[5] For example, Hutchinson and Walsh's (1992) individual country analysis studies U.S.–Japanese business cycles over the fixed and flexible regimes. In addition, multicountry analyses, such as Ahmed et al. (1993) and Bayoumi and Eichengreen (1994), study U.S.–aggregate G7 business cycles. A common finding among these studies is that the nature of underlying disturbances changed over the fixed and flexible periods. In particular, global shocks became more volatile relative to national shocks. There is some disagreement over whether there was any change in the way the U.S. and G7 responded to these underlying disturbances when they shifted from fixed to floating rates. Ahmed et al. (1993) argue that there was no change in the response to shocks under the flexible regime. Hutchinson and Walsh (1992) and Bayoumi and Eichengreen (1994) argue that there were changes in the response to shocks in the flexible period. Hutchinson and Walsh find that flexible exchange rates afforded Japan some additional insulation from foreign disturbances, while Bayoumi and Eichengreen argue that the shift to flexible exchange rates steepened the aggregate demand curve of the G7, which tended to make prices more, and output less, sensitive to supply shocks.

My analysis is essentially a multicountry version of Hutchinson and Walsh (1992). I look at the behavior of U.S.–G7 business cycles by studying bivariate models for the six U.S.–G7 pairs. I adopt a slightly different structural model of the U.S. and G7 by drawing on the approach of Eichenbaum and Evans (1995), developed in their work on the link between monetary policy disturbances and exchange rate movements. Despite this difference, my results suggest that the findings from Hutchinson and Walsh's (1992) U.S.–Japan analysis extend to other G7 countries.

Methodology and data

One way of summarizing interactions among a set of variables is through a vector autoregression (VAR). A VAR is a statistical method that allows one to estimate how an unpredictable disturbance (or change) in one variable affects other variables in the economy. For example, one of the questions that is raised by theoretical research is whether a change in foreign monetary policy has a weaker effect on domestic industrial production under flexible exchange rates. A VAR can be used to answer this type of question, since it allows one to estimate the way that an unpredicted change in monetary policy affects domestic industrial production under a fixed or flexible exchange rate regime.

The choice of variables that one includes in a VAR depends on the questions one wants answered. There is a wide range of variables one can use in analyzing U.S.–G7 business cycles. I follow Hutchinson and Walsh (1992) by limiting my analysis of U.S.–G7 business cycles to six VARs, which essentially study interaction between the U.S. and a foreign country of interest, in this case Canada, France, Germany, Italy, Japan, or the UK. Each VAR is designed to study how unpredicted changes in world oil prices, U.S. and foreign industrial production, and U.S. monetary policy (ratio of nonborrowed reserves to total reserves) affect U.S. and foreign industrial production.[6]

One of the challenges facing researchers is that data for the BW period typically date back to 1960, which leaves a small sample of just under 12 years. I use January 1974 as the start date of the flexible period, because all of the G7 countries had moved to a flexible exchange rate system by then. PBW data run through to the present, so the sample size is over 20 years. Following Eichenbaum and Evans, I overcome these data limitations by using monthly data and restricting the VARs, so that they estimate relationships between the four variables with data from the previous six months. In other words, I estimate the link between movements in industrial production and oil price movements that occurred within the last six months.[7]

With these models in hand, I am able to address whether the higher degree of business cycle comovement between the U.S. and the other G7 nations in the PBW period is due to 1) an increase in the volatility of global disturbances (such as oil prices); 2) an increase in the volatility of U.S. disturbances that affect the rest of the G7 and an increase in the volatility of G7 disturbances that affect the U.S.; 3) increased sensitivity to G7 disturbances for the U.S. and increased sensitivity to U.S. disturbances for the rest of the G7; or 4) a change in U.S. and G7 responses to global or foreign disturbances, so that they became more alike. For instance, consider estimates of the VAR over the fixed and flexible exchange rate regimes. In this setting, differences in the relative volatility of disturbances across the two periods will be reflected in changes in the ratio of the standard deviations of unpredicted movements in oil prices, output, and monetary policy across the two periods. Differences in the way U.S. and foreign industrial production react to various disturbances will be embodied in the estimated parameters of the VAR and revealed through changes in the shape and size of the model's impulse response function. For a description of the methodology in greater detail, see the technical appendix.

Empirical results

I break my empirical analysis into three parts. First, I determine the sources of variation in industrial production of the U.S. and other G7 countries in the BW and PBW periods. Second, I highlight changes in the underlying disturbances by studying the variance of disturbances. Finally, I analyze whether the response to the disturbances changed over the BW to PBW period by comparing the shape of the impulse response functions from BW and PBW models.

Were foreign or global disturbances more important in the flexible exchange rate period?

Table 1 reports decompositions of forecast errors of industrial production for various U.S.–G7 pairs. These decompositions indicate the share of the error attributable to a particular disturbance for a given forecast horizon. The forecast error variance decompositions suggest that there was a change in the

TABLE 1

Forecast error variance decompositions for industrial production of G7 countries

U.S.–Canada model

U.S.

Months ahead	Oil prices		U.S. industrial production		Canadian industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

Canada

Months ahead	Oil prices		U.S. industrial production		Canadian industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	3	1	17	14	80	85	1	0
6	5	1	18	18	77	80	1	2
12	29	1	11	17	50	74	10	8
24	34	4	3	11	42	55	20	30
36	38	7	2	8	45	40	14	45
60	41	9	3	6	49	29	7	57

U.S.–Japan model

U.S.

Months ahead	Oil prices		U.S. industrial production		Japanese industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

Japan

Months ahead	Oil prices		U.S. industrial production		Japanese industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	4	1	0	8	95	91	0	0
6	2	0	0	12	97	88	1	1
12	1	1	1	11	98	83	0	6
24	2	6	7	6	89	57	2	31
36	4	9	23	4	62	39	12	48
60	5	11	31	4	38	30	26	55

Notes: The first column in each block refers to the number of months (s = 3, 6, ..., 60) ahead for the forecast. The upper panel of a block describes the decomposition of the forecast error for U.S. industrial production, while the lower panel of a block describes the decomposition of the forecast error for foreign country industrial production. Columns in the upper and lower panels indicate the percentage of the s step ahead forecast error arising from a particular structural disturbance in the fixed and flexible exchange rate models.

Source: Calculations from author's statistical model, using the following monthly data series: International Monetary Fund, world crude oil prices and G7 industrial production; and Federal Reserve Board of Governors, nonborrowed reserves and total reserves.

(Cont. on following page)

TABLE 1 (Cont.)

Forecast error variance decompositions for industrial production of G7 countries

U.S.–Germany model

U.S.

Months ahead	Oil prices		U.S. industrial production		German industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

Germany

Months ahead	Oil prices		U.S. industrial production		German industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	4	2	1	2	95	93	0	3
6	6	4	2	3	92	90	1	3
12	12	6	2	4	85	81	1	9
24	15	4	1	3	81	62	2	31
36	16	6	1	2	81	43	2	49
60	17	9	1	2	79	26	3	63

U.S.–UK model

U.S.

Months ahead	Oil prices		U.S. industrial production		UK industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

UK

Months ahead	Oil prices		U.S. industrial production		UK industrial production		U.S. monetary policy	
	Percentage of forecast error due to:							
	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible	Fixed	Flexible
3	0	3	2	5	97	91	0	1
6	1	5	3	5	96	89	0	1
12	11	5	9	7	77	80	3	9
24	12	6	11	4	60	54	17	35
36	12	8	12	3	57	38	19	51
60	12	11	15	3	50	25	23	62

(Cont. on following page)

relative importance of the various disturbances during the BW and PBW periods at forecast horizons of one to five years. The findings appear to be uniform across the six sets of bilateral pairs. From the perspective of the other G7 countries, foreign disturbances seem to play a larger role in the flexible exchange rate period. In particular, domestic industrial production disturbances clearly dominate shocks to oil prices, U.S. industrial production, and U.S. monetary policy in the fixed exchange rate period, but are a less important source of variation in the flexible exchange rate period. A similar result emerges for U.S. industrial production. Disturbances to U.S. industrial production are also a less important source of variation to U.S. industrial production in the flexible exchange rate period. The most striking result is the increased importance of U.S. monetary disturbances under the flexible exchange rate regime. Finally, in contrast to prior beliefs, the role of oil price disturbances is little changed across the two regimes. Overall, these results suggest that a greater share of the fluctuations in G7 industrial production seem to be driven by common sources of disturbance in the flexible exchange rate period. These findings are similar to those of Hutchinson and Walsh's (1992) Japanese study.

Forecast error variance decompositions point to the sources of variation in industrial output, but they do not answer the question of whether the changing character of the relative variance of disturbances or of the response to these disturbances is at the heart of the increased comovement of national outputs. To answer this question, I need to look at the variance of

TABLE 1 (Cont.)

Forecast error variance decompositions for industrial production of G7 countries

U.S.–France model

U.S.

Months ahead	Oil prices: Fixed	Oil prices: Flexible	U.S. industrial production: Fixed	U.S. industrial production: Flexible	French industrial production: Fixed	French industrial production: Flexible	U.S. monetary policy: Fixed	U.S. monetary policy: Flexible
	Percentage of forecast error due to:							
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

France

Months ahead	Oil prices: Fixed	Oil prices: Flexible	U.S. industrial production: Fixed	U.S. industrial production: Flexible	French industrial production: Fixed	French industrial production: Flexible	U.S. monetary policy: Fixed	U.S. monetary policy: Flexible
	Percentage of forecast error due to:							
3	1	4	1	6	96	90	2	0
6	2	8	1	8	93	82	4	2
12	4	11	1	11	91	71	4	7
24	5	9	5	10	85	54	4	27
36	7	11	9	7	76	40	8	42
60	8	15	14	5	65	29	14	52

U.S.–Italy model

U.S.

Months ahead	Oil prices: Fixed	Oil prices: Flexible	U.S. industrial production: Fixed	U.S. industrial production: Flexible	Italian industrial production: Fixed	Italian industrial production: Flexible	U.S. monetary policy: Fixed	U.S. monetary policy: Flexible
	Percentage of forecast error due to:							
3	5	0	92	98	1	2	2	1
6	4	0	93	90	1	8	2	2
12	5	2	77	63	9	10	9	25
24	18	10	27	26	19	4	36	60
36	26	12	15	16	29	2	30	70
60	36	13	7	10	42	3	15	74

Italy

Months ahead	Oil prices: Fixed	Oil prices: Flexible	U.S. industrial production: Fixed	U.S. industrial production: Flexible	Italian industrial production: Fixed	Italian industrial production: Flexible	U.S. monetary policy: Fixed	U.S. monetary policy: Flexible
	Percentage of forecast error due to:							
3	1	1	1	3	95	96	3	0
6	2	3	1	6	94	91	3	0
12	14	3	5	10	76	83	5	4
24	21	4	9	9	58	63	12	24
36	23	6	10	6	54	45	13	42
60	25	10	11	5	49	28	15	57

Note: See table 1, page 52 for notes and sources.

the disturbances and the impulse response functions of the models estimated over the fixed and flexible exchange rate periods.

Were global disturbances more volatile in the flexible exchange rate period?

Table 2 reports the ratio of the standard deviations of the various disturbances under the fixed and flexible exchange rate regimes. As expected, unexpected changes to oil prices are roughly ten times more volatile over the flexible period. In contrast, unexpected changes to U.S., Canadian, German, Italian, Japanese, and UK industrial production display roughly the same level of volatility across the periods, while unexpected changes to industrial production in France are considerably lower in the flexible period. Finally, unexpected changes to U.S. monetary policy are roughly twice as volatile in the flexible exchange rate period. Based on these findings, it is clear that for G7 countries (excluding the U.S.), foreign sources of disturbance became relatively more volatile in the flexible exchange rate period.[8] The question that remains is whether the G7's response to these disturbances changed with the move from fixed to flexible exchange rates.

Are responses to disturbances different under fixed and flexible exchange rate regimes?

Figures 4–7 compare the responses of the G7 countries over the fixed and flexible exchange rate periods to the four underlying disturbances—changes in oil prices, U.S. industrial production, G7 industrial production, and U.S. monetary policy. Note that the models' responses are standardized so that each figure plots the response to a 1 percent increase in a given disturbance. This allows

TABLE 2

Estimated percentage standard deviations of structural disturbances

Canada–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	Canadian industrial production	U.S. monetary policy
Fixed	1.0	0.7	0.9	0.7
Flexible	11.2	0.7	1.2	1.3
Ratio	11.0	1.0	1.3	1.8

France–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	French industrial production	U.S. monetary policy
Fixed	1.0	0.7	3.8	0.7
Flexible	11.2	0.7	1.3	1.3
Ratio	10.9	1.0	0.3	1.8

Germany–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	German industrial production	U.S. monetary policy
Fixed	1.0	0.7	1.7	0.7
Flexible	11.2	0.7	1.5	1.3
Ratio	11.1	1.0	0.9	1.8

Italy–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	Italian industrial production	U.S. monetary policy
Fixed	1.0	0.7	2.3	0.7
Flexible	11.3	0.7	2.3	1.3
Ratio	10.9	1.1	1.0	1.8

Japan–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	Japanese industrial production	U.S. monetary policy
Fixed	1.0	0.7	1.0	0.7
Flexible	11.3	0.7	1.1	1.3
Ratio	11.2	1.0	1.2	1.8

UK–U.S. model

Period	Structural disturbance: Oil prices	U.S. industrial production	UK industrial production	U.S. monetary policy
Fixed	1.0	0.7	1.1	0.7
Flexible	11.0	0.7	1.4	1.3
Ratio	10.7	1.0	1.2	1.8

Notes: The first (second) row in each block refers to the standard deviation of the structural disturbance in the fixed (flexible) exchange rate model. The third row is the ratio of the standard deviation of the structural disturbance in the flexible to fixed period (values exceeding 1 indicate an increase in the variance of the structural disturbance).

Source: Calculations from author's statistical model, using the following monthly data series: International Monetary Fund, world crude oil prices and G7 industrial production; and Federal Reserve Board of Governors, nonborrowed reserves and total reserves.

me to compare the shape and size of the response under fixed or flexible exchange rates.

Oil price disturbances

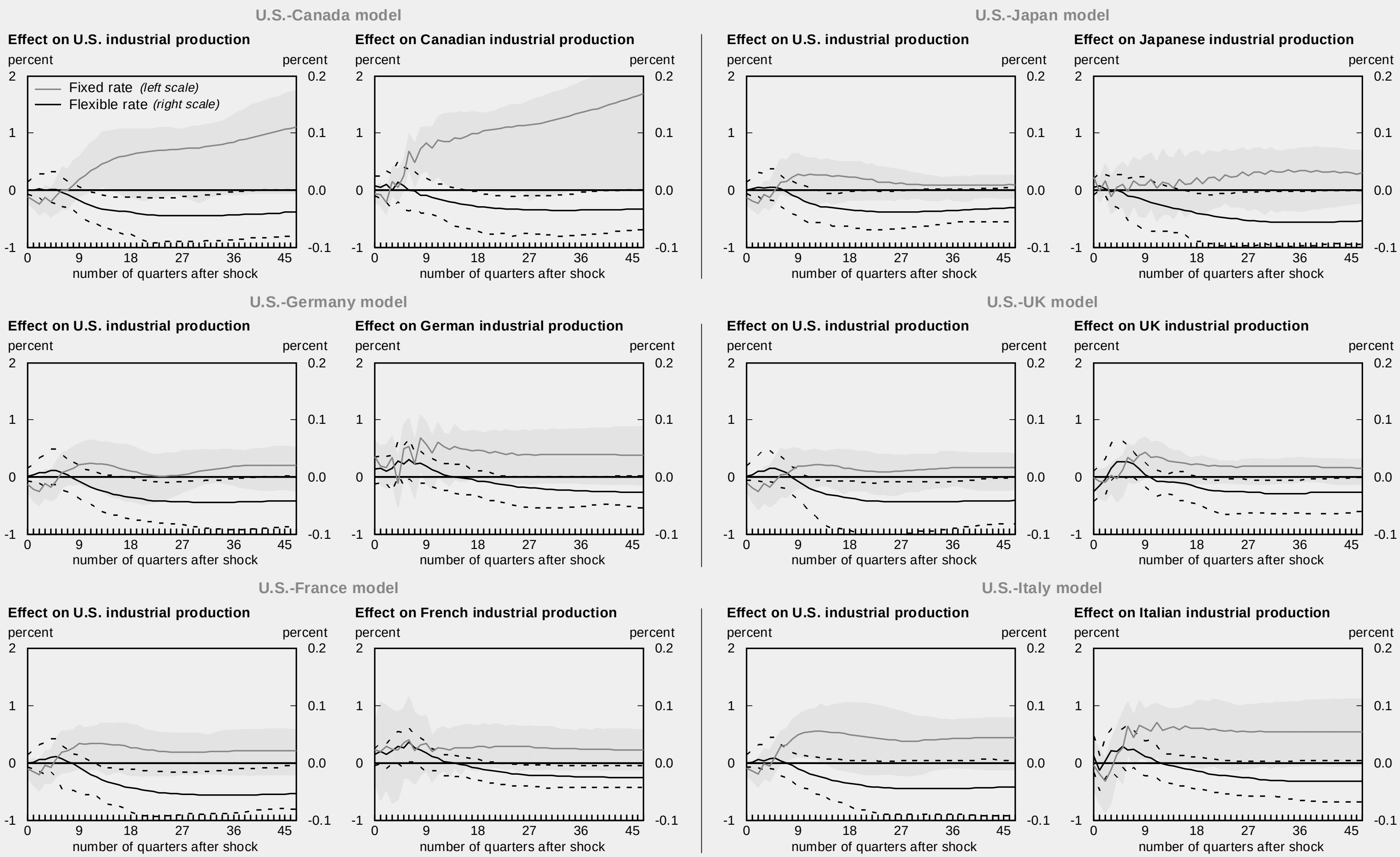
Figure 4 plots responses to oil price changes in the fixed and flexible periods. Note the scale for the response function for the flexible period is one-tenth that of the fixed period response. It is obvious that a 1 percent shock to oil prices had a smaller impact on U.S. and G7 industrial production in the flexible period in both the short and long run. The response functions for oil price changes also have quite different shapes over the two periods. The impact effect of oil prices varies across G7 countries for the fixed exchange rate period, while the long-run effect is consistently positive. In contrast, during the flexible exchange rate period, the impact effect of oil prices is positive, while the long-run effect is negative for all G7 countries. The previous section suggests that oil price movements were generally no more significant a source of variation in the flexible period. Figure 4 suggests that oil price changes were a source of the increased comovement, because G7 countries started responding in a similar way to these common shocks in the flexible period.

U.S. industrial production

A similar result emerges for shocks to U.S. industrial production (see figure 5). In the fixed exchange rate period, shocks to U.S. production generally had a negative immediate impact on other G7 countries, which changed to a positive long-run effect. In the flexible period, the other G7 countries' response to U.S. industrial production shocks changed to a hump-shaped pattern. This pattern is similar to the response function of U.S. industrial production in both the fixed and flexible periods. This suggests that the

FIGURE 4

Impulse response functions: Shock to world oil prices

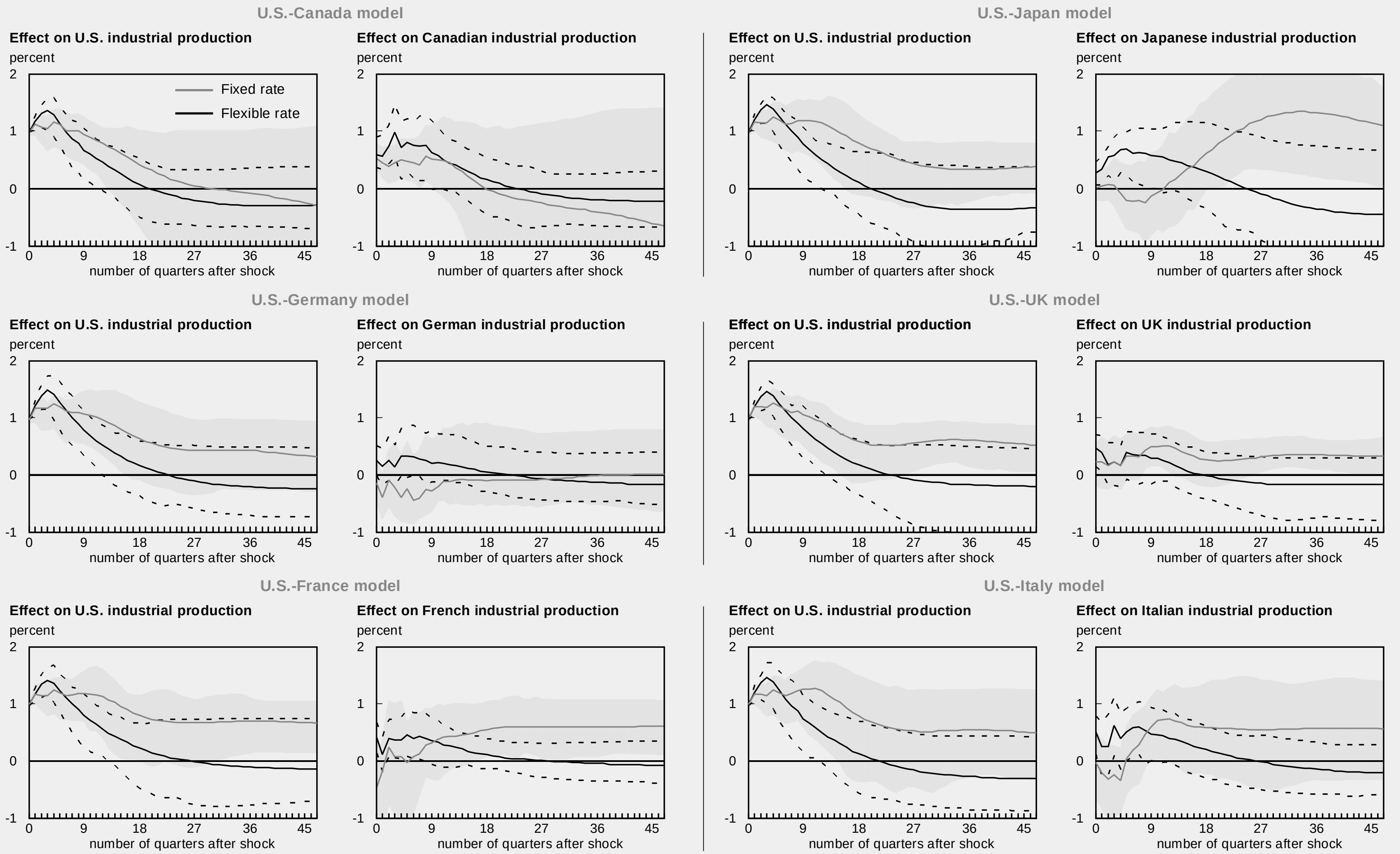


Notes: All figures report percentage changes in U.S. and other G7 industrial production following a 1 percent shock to the world price of oil in the fixed or flexible exchange rate period. The solid color (black) lines represent the point estimate for the fixed (flexible) exchange rate period. The color shaded areas (dashed lines) are the 95 percent confidence bands, computed by Monte Carlo simulation using 1,000 independent draws for the fixed (flexible) exchange rate model.

Source: Calculations from author's statistical model, using the following monthly data series: International Monetary Fund, world crude oil prices and G7 industrial production; and Federal Reserve Board of Governors, nonborrowed reserves and total reserves and U.S. industrial production.

FIGURE 5

Impulse response functions: Shock to U.S. industrial production

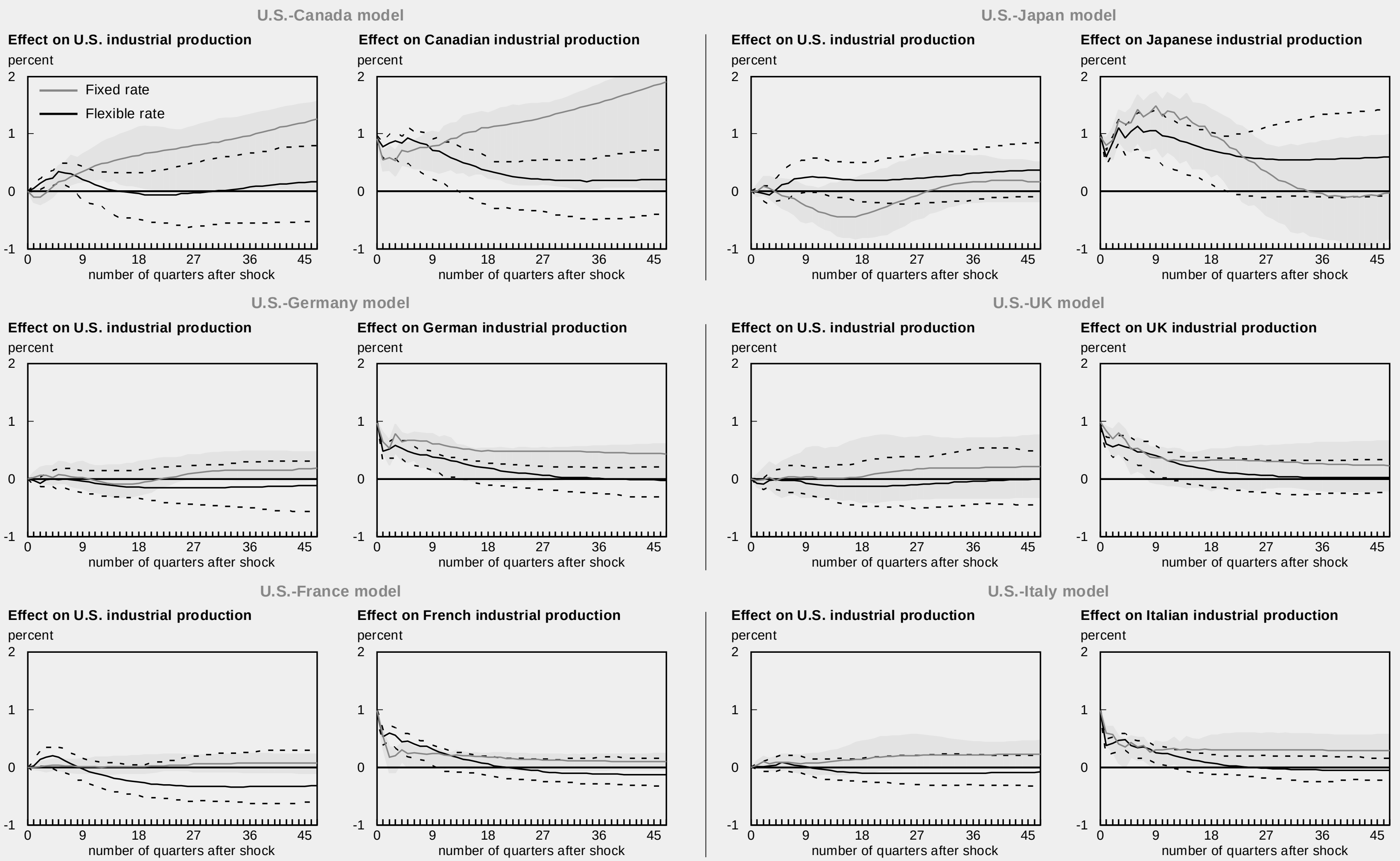


Notes: All figures report percentage changes in U.S. and other G7 industrial production following a 1 percent shock to U.S. industrial production in the fixed or flexible exchange rate period. The solid color (black) lines represent the point estimate for the fixed (flexible) exchange rate period. The color shaded areas (dashed lines) are the 95 percent confidence bands, computed by Monte Carlo simulation using 1,000 independent draws for the fixed (flexible) exchange rate model.

Source: See figure 4.

FIGURE 6

Impulse response functions: Shock to G7 industrial production

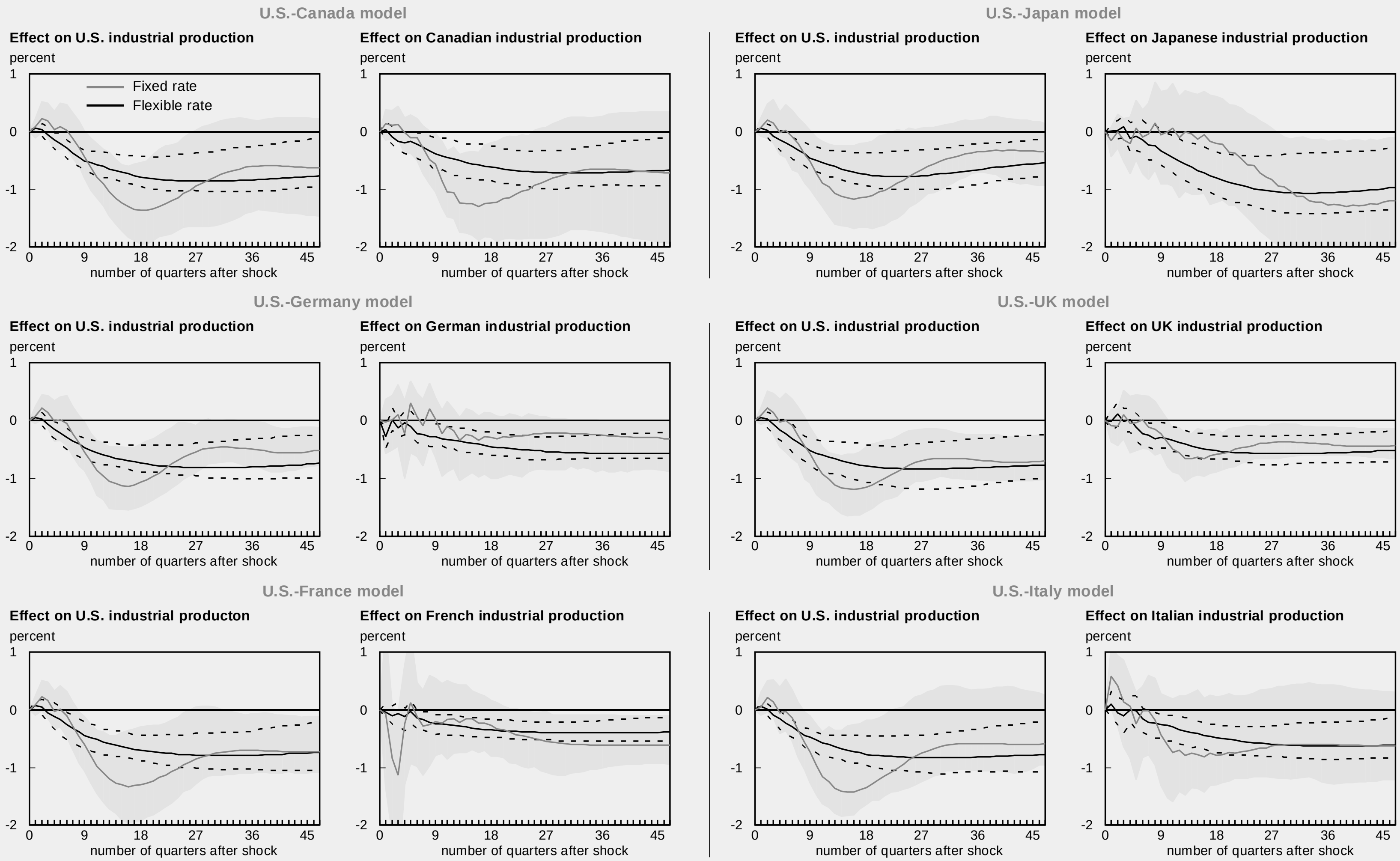


Notes: All figures report percentage changes in U.S. and other G7 industrial production following a 1 percent shock to foreign industrial production in the fixed or flexible exchange rate period. The solid color (black) lines represent the point estimate for the fixed (flexible) exchange rate period. The color shaded areas (dashed lines) are the 95 percent confidence bands, computed by Monte Carlo simulation using 1,000 independent draws for the fixed (flexible) exchange rate model.

Source: See figure 4.

FIGURE 7

Impulse response functions: Shock to U.S. monetary policy



Notes: All figures report percentage changes in U.S. and other G7 industrial production following a 1 percent shock to U.S. monetary policy in the fixed or flexible exchange rate period. The solid color (black) lines represent the point estimate for the fixed (flexible) exchange rate period. The color shaded areas (dashed lines) are the 95 percent confidence bands, computed by Monte Carlo simulation using 1,000 independent draws for the fixed (flexible) exchange rate model.

Source: See figure 4.

transmission of U.S. production shocks changed significantly in the latter period. One clear exception to this is Canada, which had roughly the same hump-shaped response to U.S. industrial production shocks in the two estimation periods. Just as in the case of oil prices, the common responses to U.S. industrial production shocks in the flexible exchange rate period are also a source of the increased comovement of U.S.–G7 industrial production.

Foreign industrial production

In contrast to the earlier results, figure 6 suggests that U.S. industrial production's response to foreign industrial production shocks is largely unchanged over the two periods. Except for Canada, foreign industrial production innovations have an insignificant impact on U.S. industrial production in the short and long run in both the fixed and flexible periods. In the case of shocks to Canada's industrial production, the U.S. and Canada share similar shaped response functions under the two regimes, so this is a possible source of comovement for the U.S. and Canada.

U.S. monetary policy

Finally, I look at unexpected changes in U.S. monetary policy. The monetary indicator used here is the ratio of nonborrowed reserves to total reserves. An exogenous increase in the ratio would indicate a tightening of monetary policy. Figure 7 shows that, historically, shocks to U.S. monetary policy (higher ratios of nonborrowed to total reserves) are associated with a contraction in U.S. and G7 industrial production. Textbook open economy macroeconomic models suggest that a standardized foreign monetary policy shock will have a smaller impact on countries that maintain flexible exchange rates. That also appears to be the consensus view from anecdotal evidence on the abandonment of the gold exchange standard and the UK's recent exit from the ERM. Figure 7 reveals that G7 countries responded to U.S. monetary disturbances is a similar way in the BW and PBW periods. In particular, the impulse response functions of these countries to U.S. monetary disturbances display the same shape, with a significant negative long-run effect. These results suggest that for other G7 countries, flexible exchange rates offer no greater insulation against foreign monetary disturbances. This result is clearly at odds with the consensus viewpoint.

Recall the finding from the previous section that U.S. monetary disturbances became more volatile in the flexible period. Combining this with the fact that the response to these shocks is common, we can see why U.S. monetary policy disturbances became a greater source of variation in G7 industrial production.

Summary

These experiments suggest that the correlation of U.S. and G7 output fluctuations rose in the flexible exchange rate period because of two factors. First, the G7's response to various structural disturbances became more alike in the flexible exchange rate period. Second, global or foreign shocks, such as U.S. monetary policy, became more volatile in the flexible exchange rate period.

Conclusion

This article sheds light on the link between exchange rate regimes and international business cycles. The key stylized fact is that the correlation of cyclical fluctuations in industrial output of the U.S. and other G7 countries rose quite dramatically in the flexible exchange rate period. This calls into question conventional wisdom, which argues that flexible exchange rates increase the degree to which national economies are insulated from the effects of foreign / global disturbances. By estimating a series of bilateral models of the U.S. and its G7 counterparts over the postwar fixed and flexible exchange rate periods, I was able to determine that higher comovement emerged in the PBW period due to a combination of two factors. First, the sensitivity to U.S. monetary policy shocks among the rest of the G7 countries remained unchanged over the fixed and flexible exchange rate regimes, but the volatility of shocks to U.S. monetary policy increased significantly in the flexible exchange rate period. This made U.S. monetary policy disturbances a more important source of variation for G7 industrial production and, in the process, raised the correlation of U.S. and G7 output fluctuations. Second, the responses of the G7 to all shocks, global and domestic, changed in the flexible regime so that they were more alike than in the fixed exchange rate period. One of the important findings of this study is that G7 sensitivity to foreign and domestic monetary policy shocks remained unchanged in the flexible exchange rate period.

There is much debate in the popular press and academic circles about the desirability of pursuing a common currency area in Europe.

The debate is an old one. Early examples include work by Mundell (1961), who argued that the desirability of a common currency area depends on the nature of disturbances and the economies' response to these shocks. Highly correlated disturbances and similar responses to disturbances were argued by Mundell to be essential elements in the desirability of a common currency area. Here, I use empirical techniques that uncover the nature of disturbances and the responses to these shocks with a view to understanding why fluctuations in national outputs of countries are highly correlated. My results suggest that G7 countries respond to shocks in a similar way and that common global shocks explain a large share of the variance of national output fluctuations. In the light of these empirical findings and Mundell's theoretical results, it would seem that the G7 would gain from the move to a common currency.

TECHNICAL APPENDIX

This appendix describes my methodology in greater technical detail. To isolate the various exogenous shocks, including U.S. monetary policy shocks, I use the vector autoregression (VAR) procedure developed by Christiano, Eichenbaum, and Evans (1994a, 1994b). Let Z_t denote the 4×1 vector of all variables in the model at date t. This vector includes changes in the log of world oil prices (*POIL*), log levels of U.S. industrial production (*USIP*), log levels of industrial production for another G7 country (*FORIP*), and the ratio of U.S. nonborrowed to total reserves (*NBR*), which I assume is the U.S. monetary policy indicator. The order of the variables is:

1) $Z_t = (POIL_t, USIP_t, FORIP_t, NBR_t)$.

I assume that Z_t follows a sixth-order VAR:

2) $$Z_t = A_0 + A_1 Z_{t-1} + A_2 Z_{t-2} + \ldots + A_q Z_{t-6} + u_t,$$

where A_i, $i = 0,1, \ldots , 6$ are 4×4 coefficient matrices, and the 4×1 disturbance vector u_t is serially uncorrelated. I assume that the fundamental exogenous process that drives the economy is a 4×1 vector process $\{\varepsilon_t\}$ of serially uncorrelated shocks, with a covariance matrix equal to the identity matrix. The VAR disturbance vector u_t is a linear function of a vector ε_t of underlying economic shocks, as follows:

$$u_t = C\varepsilon_t,$$

where the 4×4 matrix C is the unique lower-triangular decomposition of the covariance matrix of u_t:

$$CC' = E[u_t u_t'] .$$

This structure implies that the jth element of u_t is correlated with the first j elements of ε_t, but is orthogonal to the remaining elements of ε_t.

Following Christiano et al., I assume that in setting policy, the U.S. Federal Reserve both reacts to the economy and affects the economy; I use the VAR structure to capture these cross-directional relationships. In particular, I assume the feedback rule can be written as a linear function Ψ defined over a vector Ω_t of variables observed at or before date t. That is, if I let NBR_t denote the ratio of U.S. nonborrowed to total reserves, then U.S. monetary policy is completely described by:

3) $$NBR_t = \Psi(\Omega_t) + c_{4,4}\varepsilon_{4t},$$

where ε_{4t} is the fourth element of the fundamental shock vector ε_t, and $c_{4,4}$ is the (4,4)th element of the matrix C. (Recall that NBR_t is the fourth element of Z_t.) In equation 3, $\Psi(\Omega_t)$ is the feedback-rule component of U.S. monetary policy, and $c_{4,4}\,\varepsilon_{4t}$ is the exogenous U.S. monetary policy shock. Since ε_{4t} has unit variance, $c_{4,4}$ is the standard deviation of this policy shock. Following Christiano et al., I model Ω_t as containing lagged values (dated $t-1$ and earlier) of *all* variables in the model, as well as time t values of those variables the monetary authority looks at contemporaneously in setting policy.

In accordance with the assumptions of the feedback rule, an exogenous shock ε_{4t} to monetary policy cannot contemporaneously affect time t values of the elements of Ω_t. However, lagged values of ε_{4t} can affect the variables in Ω_t.

I incorporate equation 3 into the VAR structure of equations 1 and 2. Variables *POIL*, *USIP*, and *FORIP* are the contemporaneous inputs to the monetary feedback rule. These are the only components of Ω_t that are not determined prior to date t. With this structure, we can identify the right-hand side of equation 3 with the fourth equation in VAR equation 1: $\Psi(\Omega_t)$ equals the fourth row of $A_0 + A_1Z_{t-1} + A_2Z_{t-2} + ... + A_6 Z_{t-6}$, plus $\sum_{i=1}^{3} c_{4i}\varepsilon_{it}$ (where c_{4i} denotes the $(4,i)$th element of matrix C, and ε_{it} denotes the ith element of ε_t). Note that NBR_t is correlated with the first four elements of ε_t. By construction, the shock $c_{4,4}$ ε_{4t} to U.S. monetary policy is uncorrelated with the monetary policy feedback rule Ω_t.

I estimate matrices A_i, $i = 0,1, ... , 6$ and C by ordinary least squares. The response of any variable in Z_t to an impulse in any element of the fundamental shock vector ε_t can then be computed by using equations 1 and 2.

The standard error bounds in figures 4 through 7 are computed by taking 1,000 random draws from the asymptotic distribution of A_0, A_1, ... , A_6, C, and, for each draw, computing the statistic whose standard error is desired. The reported standard error bounds give the 95 percent confidence bands from 1,000 random draws.

NOTES

[1]This section draws heavily on material in Krugman and Obstfeld (1994), chapter 19.

[2]Researchers such as Grilli and Kaminsky (1991) argue that during this period the U.S. was involved in four fixed exchange rate regimes, the gold standard from January 1879 to June 1914, the gold exchange standard from May 1925 to August 1931, the wartime control period from September 1939 to September 1949, and finally the Bretton Woods system from October 1949 to August 1971. With the exception of the wartime control period, these regimes involved a fixed rate of exchange between the U.S. and other currencies in addition to a fixed dollar price of gold. The intervening years and the period following abandonment of the Bretton Woods system have been characterized by various floating exchange rate regimes.

[3]In general time series, data are nonstationary. Nonstationary data do not have well-defined standard deviations or correlations. One way of overcoming this problem is to filter the data using a filter that removes nonstationary components and renders the data stationary. Baxter and Stockman report statistics for two different filters, a linear time trend and first difference filter. In subsequent work, Baxter (1991) argued that these filtered data highlight frequencies of the data that are uninteresting for policy analysis. Baxter and King (1995) responded to this by developing a filter that is designed to isolate components of the data corresponding to frequencies policy analysts are interested in, the so-called business cycle frequencies of one and a half to eight years. I use a Baxter–King business cycle filter to isolate cyclical movements in industrial production. However, filtering industrial production with a linear time trend or first difference filter yields the same conclusion. This suggests that Baxter and Stockman's (1989) figure 4 is mislabeled.

[4]Backus, Kehoe, and Kydland (1995) study the cyclical properties of a broader set of national output data, for a smaller set of countries (Canada, Japan, the UK, and the U.S.), over the fixed and flexible periods. Using a similar business cycle filter, developed by Hodrick and Prescott (1997), they also find that the volatility of gross domestic product (GDP) and the correlation between foreign and U.S. GDP rose in the flexible period.

[5]Other empirical attempts have relied on cross-sectional econometric methods. For example, Canova and Dellas (1993) study the relationship between trade interdependence and business cycle comovement. They argue that comovement in the PBW period seems to be due to common shocks rather than changes in the international transmission of business cycles.

[6]Adding an indicator of foreign monetary policy had no impact on the analysis.

[7]Before I can shed light on the issue of whether increased comovement in national output occurred because of changes in the relative volatility of global versus national disturbances and/or changes in the response to national and global disturbances, I need to impose some structure on the system of equations described by the VAR. There are numerous forms of indentifying restrictions in the literature. In their work on Japan, Hutchinson and Walsh (1992) impose long-run restrictions on the data. Identification in Ahmed et al. (1993) and Bayoumi and Eichengreen (1994) comes from different theoretical models. I use a recursive structure popularized by Sims (1972). This approach imposes restrictions on the covariance function of the disturbances of the model. In particular, structural disturbances are identified by imposing a recursive information ordering. Throughout the analysis, I impose the following information ordering: world oil prices; U.S. industrial production; foreign industrial production; and indicator of U.S. monetary policy. This approach assumes, as in Eichenbaum and Evans (1995), that the U.S. monetary authority chooses the value of the mone-

tary instrument after observing contemporaneous movements in oil prices and U.S. and foreign industrial production. In this setting I can conveniently refer to the structural disturbances as an oil price or global shock, U.S. output shock, foreign output shocks, and U.S. monetary policy shock.

[8]Ahmed et al. (1993), Bayoumi and Eichengreen (1994), and Hutchinson and Walsh (1992) also find that foreign or global shocks became relatively more volatile in the flexible exchange rate regime. This is a noteworthy result because each study uses a different structural identification, but essentially ends up with the same general conclusion about the changing source of disturbances in the international economy.

REFERENCES

Ahmed, S., B. W. Ickes, P. Wang, and B. S. Yoo, 1993, "International business cycles," *American Economic Review*, Vol. 83, pp. 335–359.

Backus, D. K., P. J. Kehoe, and F. E. Kydland, 1995, "International business cycles: Theory and evidence," in *Frontiers of Business Cycle Research*, T. F. Cooley (ed.), Princeton, NJ: Princeton University Press.

Baxter, M., 1995, "International trade and business cycles," in *Handbook of International Economics*, Volume III, G.M. Grossman and K. Rogoff (eds.), Amsterdam: North-Holland.

_________, 1991, "Business cycles, stylized facts, and the exchange rate regime: Evidence from the United States," *Journal of International Money and Finance*, Vol. 10, pp. 71–88.

Baxter, M., and R. G. King, 1995, "Measuring business cycles: Approximate band pass filters for economic time series," National Bureau of Economic Research, working paper, No. 5022.

Baxter, M., and A. C. Stockman, 1989, "Business cycles and the exchange-rate regime: Some international evidence," *Journal of Monetary Economics*, Vol. 23, pp. 377–400.

Bayoumi, T., and B. J. Eichengreen, 1994, "Macroeconomic adjustment under Bretton Woods and the post-Bretton-Woods float: An impulse response analysis," *The Economic Journal*, Vol. 104, pp. 813–827.

Bowden, R. J., and V. L. Martin, 1995, "International business cycles and financial integration," *Review of Economics and Statistics*, Vol. 77, No. 2, pp. 305–320.

Canova, F., and H. Dellas, 1993, "Trade interdependence and the international business cycle," *Journal of International Economics*, Vol. 34, pp. 23–47.

Christiano, L. J., M. Eichenbaum, and C. Evans, 1994a, "The effects of monetary policy shocks: Evidence from the flow of funds," Federal Reserve Bank of Chicago, working paper, No. 94–2.

_________, 1994b, "Identification and the effects of monetary policy shocks," Federal Reserve Bank of Chicago, working paper, No. 94–7.

Eichenbaum, M., and C. L. Evans, 1995, "Some empirical evidence on the effects of shocks to monetary policy on exchange rates," *Quarterly Journal of Economics*, Vol. 110, pp. 975–1009.

Eichengreen, B. J., 1992, *Golden Fetters: The Gold Standard and the Great Depression 1919–1939*, New York: Oxford University Press.

Gerlach, H. M. S., 1988, "World business cycles under fixed and flexible exchange rates," *Journal of Money, Credit, and Banking*, Vol. 20, No. 4, pp. 621–632.

Grilli, V., and G. Kaminsky, 1991, "Nominal exchange rate regimes and the real exchange rate: Evidence from the United States and Great Britain, 1885–1986," *Journal of Monetary Economics*, Vol. 27, pp. 191–212.

Hodrick, R. J., and E. C. Prescott, 1997, "Post-war U.S. business cycles: An empirical investigation," *Journal of Money, Credit, and Banking*, Vol. 29, pp. 1–16.

Hutchinson, M., and C. E. Walsh, 1992, "Empirical evidence on the insulation properties of fixed and flexible exchange rates: The Japanese experience," *Journal of International Economics*, Vol. 32, pp. 241–263.

Kouparitsas, M.A., 1996, "North–South business cycles," Federal Reserve Bank of Chicago, working paper, No. WP-96-9.

Krugman, P. R., and M. Obstfeld, 1994, *International Economics: Theory and Policy*, Third Edition, New York: Harper Collins.

Mundell, R. A., 1961, "A theory of optimal currency areas," *American Economic Review*, Vol. 51, pp. 657–665.

Sims, C. A., 1972, "Money, income, and causality," *American Economic Review*, Vol. 62, pp. 540–552.

Stockman, A. C., and L. L. Tesar, 1995, "Tastes and technology in a two country model of the business cycle: Explaining international comovements," *American Economic Review*, Vol. 85, pp. 168–185.

Effects of personal and school characteristics on estimates of the return to education

Joseph G. Altonji

Introduction and summary

Hundreds of studies have shown that more educated workers receive higher wages and earnings than less educated workers.[1] This earnings gap has varied over time but has always been substantial. Recent research by Murphy and Welch (1992) shows that the difference in the average wages of college graduates and high school graduates increased substantially during the 1980s. Rosenbaum (1997) reports an earnings gap of more than 60 percent in the 1990s. However, there is much disagreement on the extent to which the earnings difference is due to the education difference. Does college make people better workers, or are better workers simply more likely to attend college? The wisdom of expanding the higher education system hinges in part on the relative importance of these two explanations of the college / high school wage differential.

There are two main channels through which a spurious correlation between education and wages might arise. First, family background, primary and secondary school quality, and ability might affect both postsecondary schooling and the wage level independent of postsecondary schooling. Second, family background, ability, and primary and secondary school characteristics may affect the rate at which students learn. Students who are more able, from better family backgrounds, or from better schools may choose more postsecondary education than the less advantaged because they receive a larger payoff to a year in college. In this case, the difference in earnings between high school graduates and college graduates will exceed the gain in earnings that a typical high school graduate would receive if he or she had chosen college. See Siebert (1985), Willis (1987), and Griliches (1977) for discussions of these issues.

The empirical evidence on whether controlling for family background and ability reduces estimates of the financial return to education is inconclusive. Much of this literature uses a statistical technique called ordinary least squares (OLS) regression to hold constant other factors while comparing the earnings of people with different levels of education. Many studies show a reduction in the estimated return, but some that have paid attention to the fact that mismeasurement of education becomes a more serious problem when one controls for ability or family background find somewhat smaller levels of bias and, in some cases, obtain higher estimates of the return to education. (See Griliches, 1979, and Siebert, 1985, for surveys.) Ashenfelter and Krueger (1991) and Angrist and Krueger (1991) find that conventional OLS regression estimates, if anything, understate the return to education.[2] These papers and other related recent work have led some to argue that failure to control for ability and background may lead to a substantial *underestimate* of the return to education. As Lang (1993) notes, if well-educated parents push their children to obtain education beyond the point of diminishing returns, then regression estimates of the return to education could be understated.

Joseph G. Altonji is a professor of economics and acting director of the Institute for Policy Research at Northwestern University, a research associate at the National Bureau of Economic Research, and a consultant to the Federal Reserve Bank of Chicago. *The author gratefully acknowledges research support from the National Center on Education and Employment, the Teachers College at Columbia University, under a research contract from the U.S. Department of Education, the Spencer Foundation, and the Institute for Policy Research at Northwestern University. The author thanks Thomas Dunn, Alan Gustman, Bruce Meyer, James Spletzer, and Barbara Wolfe for comments on an earlier draft.*

In contrast to the extensive literature on family background and ability measures, there has been little work on whether failure to control for school quality, secondary school curriculum, and community characteristics leads to bias in estimates of the return to postsecondary education. Most of the data sets that have been used to study the returns to education contain relatively little information about school curriculum and the community. Furthermore, it is hard to envision a data set that would contain measures of all of the relevant school and community characteristics. There are substantial differences across schools in parental and school characteristics that I do observe. (See appendix table 1). One naturally suspects that there are unobserved differences among high schools and communities that influence both education and wages.[3]

Data from the National Longitudinal Survey of the High School Class of 1972 (NLS72) and a matching postsecondary transcript survey (PETS) provide an opportunity to make some progress on this issue. Because the NLS72 contains several students from a large number of high schools, it is possible to statistically control for *all* observed and unobserved characteristics common to students from the same high school. One may also control for characteristics common to students in the same program (that is, academic or nonacademic track) within a given high school. In addition, the data set contains information on parental background, high school curriculum, and test scores. Consequently, I am able to control for a much richer set of factors than previous studies. At the same time, I am able to deal with potential downward bias in estimates of the return to education that would be induced by misreporting of college attendance. I do this by using information on education from PETS along with the sample members' reports of education.

My main conclusion is that controlling for family background leads to a substantial reduction in estimates of the rate of return to postsecondary education, which is defined as the percentage increase in wages that results from a year of college. The OLS estimate of the return to post-secondary academic education falls from 8.2 percent when one does not control for family background to 6.5 percent when one does. The results using the PETS data indicate that measurement error is not responsible for the reduction. Similar reductions are found among the samples of students in high school academic programs and those in nonacademic programs. I conclude that OLS estimates without detailed controls for family background and ability are overstated by about one fourth. It is important to point out, however, that the earnings gap between high school and college graduates has risen since the NLS72 data were collected. Even if the earnings gap between high school graduates and college graduates substantially overstates the return to going to college, that gap has grown so large in recent years that my results imply that college is currently a good financial investment for most people.

My other conclusions are as follows. First, estimates of the rate of return to postsecondary academic education for academic and nonacademic track high school students are remarkably similar. This is true despite the fact that students from academic programs earn substantially more than those from nonacademic programs, even after controlling for observed family background characteristics and achievement and aptitude measures. Second, controlling for high school curriculum does not have much effect on the education coefficients. Third, controlling for the specific high school the student attended has only a modest effect on the rate of return to education. For the combined sample, controlling for these factors reduces estimates of the percentage increase in earnings from a year of college by about 0.5 percentage points (for example, from 6.0 to 5.5). This suggests that failure to control for differences in high school variables does not lead to serious biases in studies of education and wages. This is good news because few data sets permit one to control for these factors.

Below, I present the wage equation that underlies most of the econometric analysis and the econometric methodology used to estimate it. Next, I discuss the data and present estimates of the return to education.

Econometric framework

The empirical analysis is based on a regression model that says that the natural logarithm of the real wage of an individual is determined by years of education, a set of other factors that I observe and can control for statistically, and a set of other factors for which I do not have data. The model takes the form

$$1) \quad W = S\rho + \textit{Effects of Control Variables} + \textit{Error Term},$$

where S is a measure of postsecondary education, such as years of schooling, obtained by a particular individual, ρ is a regression coefficient, and W is the natural logarithm of the real average hourly wage rate in a particular year of a particular person who attended a particular high school. The error term captures the influence of a potentially large number of factors that affect the wage that I do not know about. These factors include characteristics of the high school and the community that are the same for all persons who attended the same high school. Box 1 provides more detail about the form of control variables and the error term of the model.

I wish to estimate the coefficient ρ, where ρ is the effect of an extra year of school on the wage for a randomly selected person. Because wages are measured in natural logarithms, the percentage increase in wages induced by a unit increase in education is approximately equal to 100 *ρ when ρ is smaller than 0.1. The standard approach to estimating ρ is to estimate the effect of an additional year of schooling by OLS regression. OLS estimates of ρ will be biased if the unobserved factors that influence the wage also influence S. S is likely to be positively related to variables that increase the productivity of higher education, lower the direct costs to the student or lower the discount rate, or raise the nonmonetary benefits of education. Consequently, one would expect family background, ability and achievement, course of study in high school, and other high school and community factors to affect not only wage rates but also postsecondary schooling. The evidence for the NLS72 is that they affect both the wage and schooling. (See Altonji, 1988). If one does not adjust for these factors by including them in the set of control variables in equation 1, then S will "get credit" for them when one uses OLS to estimate the

BOX 1

The wage regression

The log wage rate is determined by

$$2)\quad W_{iht} = X_{ih}B_1 + C_{ih}B_2 + S_{ih}\rho + Z_h G + \omega_{iht} + S_{ih}\rho_i + S_{ih}\rho_h,$$

where I have suppressed controls for labor market experience and the year.

In equation 2, W_{iht} is the log of the real average hourly wage rate of person i from high school h in year t. The vector X_{ih} contains controls for whether the individual is female, black and / or Hispanic, a set of family background characteristics, location, and a set of aptitude and achievement measures. The elements of C_{ih} are measures of the high school curriculum taken by person i. S_{ih} is a measure of postsecondary education, such as years of schooling, and Z_h is a vector of observed high school and community characteristics. The vectors B_1, B_2, and G and the variable ρ are regression coefficients.

The composite error component ω_{iht} is

$$3)\quad \omega_{iht} = \upsilon_i + \upsilon_h + m_h + \varepsilon_{iht};$$
$$\upsilon'_i = \upsilon_i + \upsilon_h,$$

where υ'_i is an index of student and family specific factors that affect W_{iht} independently of the high school and community environment, υ_h is the mean in the high school of υ'_i, υ_i is the difference between υ'_i and υ_h for student i, m_h is an index of high school and community factors that affect W_{iht}, and ε_{iht} is a transitory error component that is assumed to be uncorrelated with all explanatory variables in the wage equation and with the other error components. The component υ_i is uncorrelated with υ_h and m_h by construction.

There are two additional error components in the wage equation. The rate of return to education $\rho + \rho_i + \rho_h$ varies across individuals and depends on an individual-specific component ρ_i and a high school component ρ_h, where ρ_i and ρ_h are uncorrelated by construction and have means of 0. The unobserved term $S_{ih}(\rho_i + \rho_h)$ is treated as part of the wage equation error in estimation. Below, I allow ρ to depend on whether a student is in an academic or nonacademic program by estimating separate equations for these groups. My econometric methods assume that variation in ρ_i and ρ_h is unrelated to S_{ih}.

effect of a change in S on the wage, and the estimate of ρ will be too large. In fact, many studies of the return to education have few controls for ability, family background, curriculum in high school, and other characteristics of the high school and community. Even when one uses a rich data set such as the NLS72, the fact both education and wages are influenced by *observed* measures of family background, student achievement, and the high school environment suggests that unobserved determinants of education are correlated with the wage error term. This is because the observed measures are likely to be incomplete or unreliable.

In the empirical work below, I systematically add controls for family background, curriculum in high school, aptitude and achievement, and observed high school characteristics to the wage equation and examine the sensitivity of estimates of the return to education to choice of control variables. I also use a statistical procedure called ordinary least squares-fixed effects (OLS-fixed effects) to control for the influence of unobserved factors that are common to students who attended the same high school. Specifically, I add a set of indicator variables ("dummy" variables) to the set of control variables, one for each high school in the sample. The indicator variable for a particular high school takes on the value 1 if the individual attended that high school and 0 otherwise. The indicator variables will absorb the effects of all factors that are common to students who attended the same high school.[4] Essentially, the OLS-fixed effects procedure estimates the effect of education on wages by relating differences in wages to differences in education across individuals who attended the *same* high school. I present separate estimates for students who were in the academic track and for students who were in the nonacademic track in high school, as well as for the combined sample. The OLS-fixed effects estimates for a specific track relate differences in wages among students who were in the same track in the same high school to differences in their postsecondary education.

Unfortunately, the use of high school fixed effects does not eliminate all of the factors that could lead to biased estimates of ρ. Even after one controls for observed measures of family background and aptitude and achievement, unobserved ability differences among students from the same high school may affect both S and the log wage. Furthermore, the quality of instruction and peer group experiences of students probably varies substantially even within a track in a given high school, so the fixed effect analysis does not control for all high school characteristics that influence particular students.[5] However, this study goes further than previous studies by controlling for high school and high school track-specific observed and unobserved variables and for high school curriculum.

The fact that people sometimes misreport years of schooling poses an additional problem.

BOX 2

The instrumental variables estimator

The mechanics of the IV estimator are as follows. First, I regress the person's report of S on the PETS measures of education and the control variables in the wage model. When high school indicators are included in the wage equation, I include them in the first stage regression for S along with the transcript measures. Then I use the predicted values from this first stage regression as the measure of S when I estimate the wage model. I use the transcript information as instrumental variables rather than as direct measures of education because PETS was not successful in obtaining transcripts for all students who claimed to have attended postsecondary schools, in some cases due to lack of cooperation from the schools. Consequently, the PETS measure of postsecondary education will also differ from actual schooling. If the measurement errors in the PETS data are uncorrelated with the information on years of schooling and degree attainment provided by the student, then the use of the predicted measure of S will eliminate the bias from measurement error.[1]

[1]Students were asked during each follow-up survey to identify any schools that they were attending or had attended. Correlated measurement errors could arise if a student attended college but said that he or she did not. In this case the student would not provide the name of the postsecondary school attended and no transcript would be found. I assume that people do not hide the fact that they attended college if they attended college for a significant period of time.

Measurement error in S will bias the estimate of ρ toward 0. This is because the "noise" in S will reduce the sample correlation between wages and S. The inclusion of controls for the high school, curriculum, family background, and test scores may exacerbate downward bias in the education coefficient arising from measurement error in education, because much of the true variation in schooling will be correlated with these controls while the measurement error will not. I address the measurement error issue by using the independent information about educational attainment in the PETS that accompanied the NLS72 to create "instruments" for the education measures and then estimate the wage model by the method of instrumental variables (IV) instead of OLS, as described in box 2.

Data: National Longitudinal Survey of the High School Class of 1972

The NLS72 is a Department of Education survey of individuals who were high school seniors during the spring of 1972. Thus, high school dropouts are excluded. The individuals were resurveyed in 1973, 1974, 1976, and 1979. A subsample was resurveyed in 1986.[6]

TABLE 1

Definitions of key variables

Variable	Definition
W	Natural logarithm of the real hourly wage rate.
YRSACD79	Years of postsecondary academic education completed by 1979.
YRSVOC79	Years of postsecondary vocational education by 1979.
VOC79	Indicator variable that equals 1 if a person attended a postsecondary vocational education program and did not attend a postsecondary academic program.
SOC1479	Indicator variable that equals 1 if a person has less than two years of college (regardless of whether the person also attended vocational school) and 0 otherwise.
SOC1579	Indicator variable that equals 1 if a person attended college for two or more years but did not receive a four-year degree and 0 otherwise.
COLL79	Indicator variable that equals 1 if a person received a four-year degree but did not receive an advanced degree and 0 otherwise.
ADV79	Indicator variable that equals 1 if a person received a graduate degree and 0 otherwise.

The key variables used in the study are listed in table 1. Note that the indicator variables VOC79, SOC1479, SOC1579, COLL79, and ADV79 are mutually exclusive. I also construct a set of education measures from PETS to use as instruments.[7]

The control variables for region and city size, family background, aptitude and achievement measures, high school curriculum (semester hours in each of eight subjects), and high school characteristics are listed in the footnotes to the tables. Descriptive statistics and variable definitions are provided in appendix table 1. Only the education coefficients are shown in tables 2 and 3.

Estimates of the return to education

Table 2 presents OLS estimates of the effects of YRSACD79 and YRSVOC79. Columns 1–4 do not include dummy variables for each high school, while columns 5–7 do. All equations contain controls for race, sex, experience, and the year the wage data refer to.[8] The column headings indicate whether controls for region and city size (region), family background and achievement and aptitude measures (family / achievement), and high school curriculum and high school characteristics (high school) are included. The high school indicators absorb the effect of any variables that are constant within the high school, and so region and city size and fixed high school characteristics are implicitly controlled for in columns 5–7.

The returns to academic education

The coefficients in the table for YRSACD79 are estimates of the average amount that the log wage rises in response to an extra year of academic postsecondary education. For example, the coefficient on YRSACD79 is .0817 when only the basic controls are included (column 1). This coefficient implies that spending an extra year in college raises the log wage by .0817. This translates into an increase in the wage of about 8 percent. This is typical of estimates from other data sets for the year 1980, which is in the middle of the time period that the wage data are drawn from. The coefficient falls to .0653 when family background and ability and aptitude

TABLE 2

Effect of education on wages: OLS estimates (Dependent variable: log wage)

	OLS				OLS-fixed effects (Constants for each high school)		
	Basic controls	Region	Family/ achievement, region	High school, family/ achievement, region	Region	Family/ achievement, region	High school, family/ achievement, region
				Combined sample			
YRSACD79	.0817	.0790	.0653	.0644	.0749	.0605	.0598
	(.0028)	(.0028)	(.0035)	(.0036)	(.0029)	(.0036)	(.0036)
YRSVOC79	.0145	.0150	.0133	.0135	.0173	.0163	.0154
	(.0053)	(.0052)	(.0052)	(.0052)	(.0053)	(.0052)	(.0052)
				Students in academic programs			
YRSACD79	.0731	.0734	.0636	.0637	.0663	.0568	.0567
	(.0043)	(.0042)	(.0047)	(.0047)	(.0046)	(.0050)	(.0051)
YRSVOC79	.0133	.0152	.0165	.0166	.0183	.0201	.0180
	(.0076)	(.0074)	(.0075)	(.0075)	(.0081)	(.0081)	(.0080)
				Students in nonacademic programs			
YRSACD79	.0689	.0670	.0572	.0563	.0651	.0547	.0550
	(.0046)	(.0046)	(.0056)	(.0057)	(.0053)	(.0065)	(.0065)
YRSVOC79	.0196	.0192	.0162	.0178	.0138	.0121	.0116
	(.0074)	(.0073)	(.0073)	(.0074)	(.0078)	(.0078)	(.0078)

Notes: Region = NO.CENTRAL, SOUTH, WEST, SMLTOWN, MED.CITY, BIGCITY, HUGECITY, MED.SUBURB, BIGSUBURB, HUGESURB, COLL-PROX.

Family/achievement = FATHER-ED, MOTHER-ED, LOWSES, ED-MONEY, MOTHER-WORK, BLUECOLF, ENGLISH, FATH-COLL, MOTH-COLL, DISC-PLANS, PAR-INTEREST, PAR-INFL, IMPTAVER, COLLEGE-ABILITY, TEACHER-ASSESSMENT, VOCABULARY, PICTURE.NUMB, READING, LETTER.GROUP, MATH, MOSAIC.COMP, HOMEWORK, and dummy variables for whether data were missing for FATH-COLL, MOTH-COLL, or BLUECOLF.

High school characteristics include controls for the level and square of the fraction of the student body who are black, the student/teacher ratio, whether the school is private or parochial, the number of grades in the high school, the daily attendance rate, the dropout rate, the teacher turnover rate, the fraction of teachers with master's or Ph.D. degrees, the availability of advanced science courses, the number of students in the school, and the means across students of the number of courses taken between tenth and twelfth grade in science, foreign language, social studies, English, mathematics, industrial arts, commercial arts, and fine arts.

The coefficients in the table are estimates of the effect of additional years of education on the log wage. The combined sample contains 38,595 person-year observations on 9,239 students from 897 high schools. The academic sample contains 18,653 person-year observations on students from the academic programs in 858 high schools. The nonacademic sample contains 19,942 person-year observations on students from the vocational or general programs in 864 high schools. Summary statistics and variable definitions are given in appendix table 1.

All equations include BLACK, HISP, CSEX, a quadratic in years of work experience, and a quadratic in the calendar year that the wage measure refers to.

Variables that do not vary across high schools, such as the region variables and the high school variables noted above, are implicitly controlled for in the equations with high school dummies.

"White" standard errors in parentheses account for arbitrary forms of heteroscedasticity and correlation across observations on students from a given high school.

Source: Author's calculations based on data from the *National Longitudinal Survey of the High School Class of 1972* (U.S. Department of Education, 1972–86).

measures are added, a decline of .0164. This reduction is consistent with the findings of most other studies that have used detailed controls for family background and ability or made use of sibling pairs.[9] On the other hand, adding controls for the student's courses and a set of high school characteristics lowers the YRSACD79 coefficient by only .0009 to .0644.

Does the fact that almost all studies of the economic value of college fail to control for unobserved high school and community characteristics matter? The answer is that there is only a small upward bias without these controls. For example, when one adds a separate constant term (or fixed effect) for each high school to the specification in column 2, which does not

TABLE 3

Effect of education on wages: Instrumental variables estimates
(Dependent variable: log wage)

	Instrumental variables estimator: Basic controls	Instrumental variables estimator: Region	Instrumental variables estimator: Family/ achievement, region	Instrumental variables estimator: High school, family/ achievement, region	Instrumental variables estimator with fixed effects (Constants for each high school): Region	Instrumental variables estimator with fixed effects (Constants for each high school): Family/ achievement, region	Instrumental variables estimator with fixed effects (Constants for each high school): High school, family/ achievement, region
Combined sample							
YRSACD79	.0817	.0793	.0582	.0572	.0770	.0560	.0551
	(.0031)	(.0031)	(.0037)	(.0037)	(.0033)	(.0039)	(.0040)
YRSVOC79	.0254	.0080	−.0127	−.0151	.0184	−.0104	−.0133
	(.0170)	(.0167)	(.0169)	(.0169)	(.0171)	(.0173)	(.0175)
Students in academic programs							
YRSACD79	.0765	.0765	.0618	.0615	.0699	.0570	.0568
	(.0058)	(.0056)	(.0062)	(.0062)	(.0065)	(.0071)	(.0072)
YRSVOC79	.0550	.0379	.0338	.0327	.0283	.0336	.0314
	(.0340)	(.0331)	(.0333)	(.0329)	(.0379)	(.0379)	(.0378)
Students in nonacademic programs							
YRSACD79	.0724	.0722	.0589	.0578	.0758	.0636	.0641
	(.0054)	(.0055)	(.0061)	(.0062)	(.0064)	(.0073)	(.0073)
YRSVOC79	.0180	.0033	−.0074	−.0095	−.0041	−.0143	−.0173
	(.0212)	(.0213)	(.0212)	(.0213)	(.0225)	(.0225)	(.0226)

Notes and source: See table 2. In addition, for columns 1–4 the instruments consist of dummies for whether the individual had a postsecondary transcript, a transcript from a vocational school, a transcript from a two-year public college, a four-year public college, a private college, dummies for whether the individual's highest degree was a license or certificate, an associate degree, a bachelor's degree, or an advanced degree, and a count of the number of transcripts for the individual. The instrumental variables estimator with fixed effects includes dummy variables for each high school in both the instruments and the wage equation.

contain controls for family background, aptitude and achievement, or courses taken, the coefficient on YRSACD79 falls from .0790 to .0749 (see column 5). That is, the estimate of the percentage change in wages induced by an extra year of education falls from 7.9 percent to 7.49 percent. When one controls for background and achievement, the comparable coefficients without and with high school dummies are .0653 and .0605, respectively. When one controls for curriculum and observed high school characteristics, adding the high school constants reduces the coefficient on YRSACD79 from .0644 to .0598. Thus, failure to control for high school differences leads to an upward bias of .005 in the education coefficient, which (multiplying by 100) is an upward bias of 0.5 percentage points in the rate of return to education.

Similar results are obtained for students from academic and nonacademic programs. The coefficients for the two subgroups are remarkably similar. They are also a bit below the coefficients for the combined sample. This reflects the fact that both the wage level and YRSACD79 are positively correlated with whether one is in an academic high school program, even after controlling for background, aptitude and achievement, and semester hours by subject area.

Appendix table 2 reports OLS estimates of the effects of academic education when the dummy variables VOC79, SOC1479, SOC1579, COLL79, and ADV79 are used to parameterize the model. The coefficients on the education variables are all relative to a high school graduate. The results are qualitatively consistent with those based upon the linear specification in table 2.

Instrumental variables estimates

For the combined sample, the use of the transcript measures of education as instruments for the person's report of education has no effect (to four digits) on the estimated return to YRSACD79 when one does not control for family background and test scores. It leads to a slight *reduction* (relative to OLS) in estimates of the return to academic education when one controls for family background and test scores. This implies that the reduction in the education slope from about .079 with only regional controls to .058 when family background and test scores are added is not an artifact of measurement error in the education variable. There is only a small drop in the IV estimate (from .058 to .056) when high school fixed effects are added to the equation with family background and test scores (table 3). The IV results confirm the earlier OLS finding that failure to control for high school and community variables leads to only a small bias in estimates of the return to education.

The use of IV in place of OLS does not significantly change the conclusions for the academic and nonacademic groups.[10] The IV estimates of models that use five dummy variables for education outcomes indicate that controlling for high school makes almost no difference for academic education and, if anything, leads to an increase in the estimated return to vocational education. (See appendix table 3.)

The returns to vocational education

Tables 2 and 3 report OLS and IV estimates of the effect of years of vocational education (YRSVOC79) on wages for the combined sample and the academic and nonacademic subgroups. The mean of YRSVOC79 is .5110 for the combined sample and .5031 and .5183 for the academic and nonacademic subsamples, respectively, which says that the average high school graduate from the class of 1972 obtained about a half year of postsecondary vocational education. For the combined sample, the OLS results for the linear specification indicate a much lower return for vocational education than for academic education, with a coefficient of .0145 in the absence of controls (table 2, column 1), and .0154 when one controls for background, aptitude and achievement, high school curriculum, and the high school (table 2, column 7). These estimates imply that the financial return to spending a year in postsecondary vocational education is only about 1.5 percent. The estimates are similar for students who took an academic program in high school and students who took a nonacademic program. The fact that these estimates *rise* when one adds more detailed control variables is consistent with abundant evidence that less advantaged individuals tend to pursue vocational education.

However, the low estimates of the return to a year of vocational education should be treated cautiously for two reasons. First, vocational education is a very heterogenous category and programs lasting just a few months may be coded as lasting a year. (See Grubb, 1993.) This would lead to downward bias. Second, it is possible that the value of vocational education is lower if one has also obtained academic postsecondary education. This would make sense if the skills acquired in vocational education are not used by students who later pursue academic education. The wage models in appendix tables 2 and 3 that use the indicator variables VOC79, SOC1479, SOC1579, COLL79, and ADV79 as the education measures shed some light on this issue. This is because the vocational education variable, VOC79, excludes individuals who obtained both academic and vocational postsecondary education. It is 1 if the person obtained some vocational education and did not obtain any academic education and 0 otherwise. As a result, the mean of VOC79 is much lower for the academic high school track sample than for the nonacademic track sample, despite the fact that the mean of YRSVOC79 is similar for the two groups. For the combined sample, the OLS coefficient on VOC79 implies that vocational education raises wages by 4.8 percent to 6.5 percent, depending upon what one controls for. I suspect there are differences in the content of postsecondary vocational education for academic track versus nonacademic track students, and these differences may underlie the larger coefficient on VOC79 for the academic sample.

The IV estimates for YRSVOC79 and VOC79 follow the same general pattern as the OLS estimates, but are imprecise, particularly for the academic sample. Some of the point estimates for YRSVOC79 are negative but not statistically significant. However, for the combined sample the coefficient on VOC79 is quite substantial (.1179) when one controls for the high school, family background, curriculum, and test scores, although the standard error is .064 (appendix table 3). A possible explanation (other than sampling error) is that the returns to vocational programs that are sufficiently well established to lead to a transcript and / or a license or certificate are larger than the returns to other programs. The IV estimates give more weight to such programs than the OLS estimates do. Grubb's (1993) analysis of NLS72 suggests substantial heterogeneity in vocational programs. A key policy issue is how to enhance the labor market skills of persons who are not well suited for or interested in academic postsecondary education. The results suggest that some vocational training programs have substantial labor market value for students who specialize in vocational education after high school.

The impact of controlling for high school and community characteristics and for family background and achievement measures on estimates of the return to vocational education is sensitive to whether one uses OLS or IV, to the form of the education variables, and to whether the student was in an academic or nonacademic program in high school. I will not discuss the detailed results in the tables. Part of the problem is that the IV coefficient estimates for VOC79 are very imprecise, particularly for the academic sample.

Conclusion

The OLS and IV estimates with high school fixed effects indicate that only modest biases result from the failure of previous studies to control for differences in high schools and for differences in primary school and community characteristics common to students from the same high school. This is good news for researchers, because few data sets permit one to study clusters of students from the same high school. On the other hand, in contrast to several recent studies, I find that failure to control for family background and aptitude and achievement measures leads one to overestimate the rate of return to college education by about one fourth.

NOTES

[1]Siebert (1985) and Willis (1986) provide surveys of the link between education and earnings.

[2]Ashenfelter and Krueger obtain a 16 percent return to education when they contrast wages of identical twins with different schooling levels and use an instrumental variables scheme based on a twin's estimate of his / her sibling's schooling to deal with measurement error. However, Ashenfelter and Rouse (1997) use a larger sample of twins and obtain estimates closer to those obtained here.

[3]The evidence from Akin and Garfinkel (1977), Morgan and Sirageldin (1968), and Johnson and Stafford (1973) collectively suggests a positive link between school quality proxies and labor market outcomes. Card and Krueger (1992) find that school quality proxies that are related to educational attainment are also related to education slopes.

[4]The standard errors for both the OLS and instrumental variables regressions with and without high school fixed effects allow for arbitrary high school-specific forms of heteroscedasticity, serial correlation, and correlation across students from the same high school.

[5]There is information on tracking in the NLS72, and in future work it would be interesting to use a fixed effect to control for observed and unobserved characteristics that are common to students from the same track in high school. In terms of the model in box 1, the use of fixed effects controls for the high school error component $\upsilon_h + m_h$. It does not eliminate potential bias from the correlation between S_{ih} and the individual error component υ_i or between S_{ih} and the component ρ_i and ρ_h of the rate of return to education.

[6]I restrict the sample to the 16,683 individuals from the schools that participated in the base year survey. The sample is reduced to 15,680 by eliminating observations with missing high school test information and to 12,980 by eliminating individuals who did not respond to all of the first four follow-ups. Information from the 1986 follow-up was then added for persons who were in the earlier sample of 12,980. The yearly wage observations are created using information on earnings divided by hours for 1977, 1978, and 1979, and information on the wage at the beginning and end of each job held between 1980 and 1986 up to a maximum of the four most recent jobs. An observation for 1977 is included if 1) the individual was not a full-time student in October 1976 or October 1977, 2) the number of hours worked in 1977 was greater than 1,040, *and* 3) the log of the 1977 real wage was between $.50 and $75 in 1967 dollars. Observations for 1978 and 1979 were included if they met the corresponding three criteria for 1978 and 1979, respectively. Data for beginning and ending job dates (1980–86) were included if 1) the number of hours worked in the appropriate year was greater than 1,040, and 2) the log of the real wage was between $.50 and $75 in 1967 dollars. Restriction of the sample to cases with complete data on the variables used in the wage analysis reduced the sample size to 38,595 observations on 9,239 individuals from 897 high schools. The subsample of students in academic programs contains 18,653 person–year observations from 858 high schools. The corresponding figures for the nonacademic (general and vocation tracks) subsample are 19,942 and 864.

[7]The variables constructed from the PETS survey include the number of transcripts found for each student and nine indicator variables for whether the student had the following transcript combinations: 1) at least one transcript; 2) a transcript from a nonacademic institution; 3) a transcript from a two-year public academic institution; 4) a transcript from a four-year public ac-

ademic institution; 5) a transcript from a private academic institution; 6) a license or certificate but no academic degree; 7) an associate degree but no bachelor's or advanced degree; 8) a college degree but no advanced degree; and 9) an advanced degree. The PETS survey contains at least one transcript for 83 percent of the sample members who reported some postsecondary education by 1979, 74.8 percent of those who reported vocational education or some college but no degree, and 96.16 percent of those who reported a college or advanced degree. Transcript evidence of a college or advanced degree was found for 82.29 percent of the sample members who reported a college or advanced degree. Transcript evidence of a college or advanced degree was found for 3.16 percent of the sample who did not report a college or advanced degree by 1979. Also, transcript evidence of an advanced degree was found for 8.13 percent of the persons who reported college as their highest degree in 1979, which may in part be due to completion of their advanced degrees after 1979.

[8] I include a quadratic in years of work experience and a quadratic in the calendar year that the wage measure corresponds to.

[9] See Griliches (1979) and Olneck (1979) for discussions of alternative estimates of the return to education based on sibling data.

[10] In a study conducted after the initial drafts of this article were completed, Kain and Rouse (1995) use the NLS72 and PETS and also find that controlling for family background and ability measures leads to a substantial reduction in OLS estimates of the returns to two- and four-year colleges. However, they obtain higher estimates of the return when they use distance from the college and tuition as instrumental variables for college attendance.

APPENDIX

APPENDIX: TABLE 1

Means and standard deviations of wage and education variables

	Combined sample				Academic		Nonacademic	
Explanatory variables	Mean	Standard deviation	Standard deviation in high school	Fraction of variance across high schools	Mean	Standard deviation	Mean	Standard deviation
Wages								
LOGWAGE, log of real average hourly wage, 1967 dollars	.9196	.4635	.4402	.0980	.9916	.4746	.8523	.4425
Education								
YRSACD79, years of postsecondary academic education by 1979	1.988	1.843	1.666	.1829	2.936	1.753	1.101	1.439
YRSVOC79, years of postsecondary vocational education by 1979	.5309	.7641	.7183	.1163	.5228	0.794	0.5385	.7347
VOC79, 1 if some vocational, no college	.0851	–	–	–	.0381	–	.1290	–
SOC1479, 1 if less than 2 years college	.1803	–	–	–	.1413	–	.2168	–
SOC1579, 1 if more than 2 years college, no degree	.1738	–	–	–	.1995	–	.1497	–
COLL79, 1 if college degree, no advanced degree	.3022	–	–	–	.4928	–	.1240	–
ADV79, 1 if advanced degree	.0285	–	–	–	.0534	–	.0053	–
Gender and race/ethnicity								
BLACK	.0892	–	–	–	.0663	–	.1106	–
HISP	.0366	–	–	–	.0213	–	.0509	–
FEMALE	.4916	–	–	–	.4780	–	.5043	–
Family background								
FATHER-ED, father's education	12.75	2.545	2.171	.2723	13.49	2.632	12.06	2.250
MOTHER-ED, mother's education	12.43	.098	1.854	.2191	12.96	2.158	11.93	1.190
LOWSES, 1 if low SES	.2340	–	–	–	.1343	.3410	.3273	.4692
ED-MONEY, 1 if worry over money interfered with high school education	.2891	–	–	–	.2261	.4183	.3480	.4764
MOTHER-WORK, 1 if mother worked while in elementary school	.4021	–	–	–	.3804	.4855	.4223	.4939
BLUECOLF, 1 if father blue collar	.3216	–	–	–	.2925	.4549	.3488	.4766
ENGLISH, 1 if English spoken at home	.9207	–	–	–	.9183	.2739	.9230	.2666
FATH-COLL, 1 if father wants college or grad school	.5787	–	–	–	.7814	.4133	.3890	.4875
MOTH-COLL, 1 if mother wants college or grad school	.6140	–	–	–	.8183	.3856	.4229	.4940
DISC-PLANS, 1 if often discussed plans with parents	.7902	–	–	–	.8510	.3561	.7332	.4423
PAR-INTEREST, 1 if uninterested parents interfered with high school	.2019	–	–	–	.1272	.3332	.2718	.4449
PAR-INFL, 1 if parents influenced post high school plans a great deal	.4397	–	–	–	.5037	.5000	.3799	.4854
Geographic variables								
SMLTOWN	.2953	–	–	–	.3027	–	.2883	–
MED.CITY	.0832	–	–	–	.0863	–	.0802	–
MED.SUBURB	.0487	–	–	–	.0596	–	.0386	–
BIGCITY	.1020	–	–	–	.0955	–	.1080	–
BIGSUBURB	.1046	–	–	–	.1171	–	.0930	–
HUGECITY	.0785	–	–	–	.0890	–	.0686	–
HUGESURB	.0950	–	–	–	.1127	–	.0784	–
NO.CENTRAL	.2916	–	–	–	.2750	–	.3072	–
SOUTH	.3176	–	–	–	.2737	–	.3585	–
WEST	.1678	–	–	–	.1461	–	.1880	–
COLL-PROX	1.785	–	–	–	1.701	–	1.863	–
Aptitude and achievement measures								
IMPTAVER, grades	15.64	7.586	3.469	.7909	14.44	7.437	16.76	7.553
COLLEGE-ABILITY, 1 if definitely college material; 5 if definitely not	1.843	.9659	.8984	.1349	1.477	.6991	2.186	1.052

(Cont. on following page)

APPENDIX: TABLE 1 (cont.)

Means and standard deviations of wage and education variables

Explanatory variables	Combined sample: Mean	Combined sample: Standard deviation	Combined sample: Standard deviation in high school	Combined sample: Fraction of variance across high schools	Academic: Mean	Academic: Standard deviation	Nonacademic: Mean	Nonacademic: Standard deviation
TEACHER-ASSESSMENT, 1 if teacher expectation high; 5 if low	2.085	.8701	.8227	.1060	1.816	.7948	2.337	.8621
VOCABULARY	52.31	9.896	8.640	.2377	56.34	9.328	48.54	8.872
PICTURE.NUMB, associative memory	51.57	9.680	8.938	.1474	53.75	9.138	49.54	9.729
READING	52.31	9.424	8.458	.1945	56.07	8.321	48.80	9.033
LETTER.GROUP, inductive reasoning	52.33	8.878	8.034	.1811	55.33	7.082	49.54	9.456
MATH, quantitative comparisons (basic competence in math)	52.50	9.539	8.486	.2086	57.06	7.939	48.25	8.924
MOSAIC.COMP, perceptual speed and accuracy	51.46	9.187	7.407	.3499	53.11	8.681	49.92	9.377
HOMEWORK, hours on homework per week	4.467	3.278	3.018	.1523	5.315	3.442	3.674	2.899

Notes: Means and standard deviations of variables used in the wage equations for the full sample, the academic sample, and the nonacademic sample. The combined wage sample contains 38,595 observations on 9,239 individuals from 897 high schools. The academic (nonacademic) sample contains 18,653 (19,942) observations on 4,292 (4,947) individuals from 858 (865) high schools. The table also reports the standard deviation of each variable within a high school, and the fraction of the sample variance that is across high schools. The standard deviations and the variance decomposition in the table refer to the cross section–time series sample, to which individuals contribute different numbers of observations. Consequently, they provide only a rough indication of relative importance of variation within the high school and variation across high schools in wages, education, and background characteristics. (See Altonji, 1988, for a more thorough treatment of this issue.) However, the results indicate that there is substantial variation across high schools in background characteristics, aptitude and achievement measures, and curriculum. Note also that there are substantial differences in the means for the academic and nonacademic samples.

Source: See text table 2.

APPENDIX: TABLE 2

Effect of postsecondary education on wages: OLS estimates (Dependent variable: log wage)

	OLS estimator				OLS-fixed effects (Constants for each high school)		
	Basic controls	Region	Family/ achievement, region	High school, family/ achievement, region	Region	Family/ achievement, region	High school, family/ achievement, region
				Combined sample			
VOC79	.0634	.0615	.0479	.0479	.0658	.0528	.0521
	(.0137)	(.0134)	(.0135)	(.0134)	(.0138)	(.0137)	(.0137)
SOC1479	.0773	.0574	.0241	.0251	.0479	.0149	.0146
	(.0113)	(.0111)	(.0119)	(.0120)	(.0118)	(.0125)	(.0127)
SOC1579	.1880	.1682	.1191	.1160	.1521	.1017	.0486
	(.0120)	(.0117)	(.0133)	(.0134)	(.0125)	(.0141)	(.0143)
COLL79	.3437	.3285	.2571	.2542	.3130	.2390	.2354
	(.0126)	(.0126)	(.0155)	(.0158)	(.0132)	(.0159)	(.0162)
ADV79	.5057	.4908	.4101	.4040	.4658	.3781	.3736
	(.0291)	(.0289)	(.0310)	(.0313)	(.0299)	(.0317)	(.0319)
				Students in academic programs			
VOC79	.1525	.1507	.1342	.1297	.1840	.1629	.1568
	(.0310)	(.0314)	(.0330)	(.0325)	(.0329)	(.0338)	(.0337)
SOC1479	.0876	.0701	.0462	.0477	.0711	.0492	.0440
	(.0244)	(.0245)	(.0258)	(.0257)	(.0269)	(.0270)	(.0274)
SOC1579	.1933	.1757	.1427	.1403	.1560	.1254	.1185
	(.0237)	(.0236)	(.0259)	(.0260)	(.0271)	(.0294)	(.0294)
COLL79	.3319	.3221	.2677	.2667	.2975	.2450	.2398
	(.0237)	(.0238)	(.0271)	(.0272)	(.0265)	(.0294)	(.0296)
ADV79	.4734	.4658	.3995	.3997	.4378	.3711	.3665
	(.0347)	(.0341)	(.0369)	(.0366)	(.0383)	(.0407)	(.0480)
				Students in nonacademic programs			
VOC79	.0388	.0379	.0293	.0306	.0302	.0243	.0254
	(.0149)	(.0146)	(.0144)	(.0144)	(.0159)	(.0157)	(.0156)
SOC1479	.0681	.0495	.0294	.0310	.0414	.0210	.0235
	(.0135)	(.0133)	(.0139)	(.0139)	(.0146)	(.0155)	(.0156)
SOC1579	.1610	.1473	.1187	.1165	.1310	.1004	.0998
	(.0159)	(.0156)	(.0167)	(.0169)	(.0172)	(.0184)	(.0186)
COLL79	.2889	.2844	.2407	.2388	.2836	.2379	.2400
	(.0194)	(.0194)	(.0219)	(.0221)	(.0222)	(.0254)	(.0253)
ADV79	.4882	.4809	.4253	.4223	.3848	.3318	.3327
	(.1436)	(.1438)	(.1456)	(.1479)	(.1461)	(.1477)	(.0475)

Notes and source: See text table 2. In addition, the indicator variable VOC79 is 1 if an individual never attended college but did attend a postsecondary vocational school and 0 otherwise. SOC1479 is 1 if a person has less than two years of college (regardless of whether the student also attended vocational school) and 0 otherwise. SOC1579 is 1 if a person attended college for two or more years but did not receive a four-year degree and 0 otherwise. COLL79 is 1 if a person received a four-year degree but did not receive an advanced degree and 0 otherwise. ADV79 is 1 if a person received a graduate degree and 0 otherwise. The coefficients are estimates of difference in the log wage of a high school graduate and a person whose highest education level is in the particular category.

APPENDIX: TABLE 3

Estimates of the return to education: Instrumental variables (Dependent variable: log wage)

	Instrumental variables estimator				Instrumental variables estimator with fixed effects (Constants for each high school)		
	Basic controls	Region	Family/ achievement, region	High school, family/ achievement, region	Region	Family/ achievement, region	High school, family/ achievement, region
				Combined sample			
VOC79	.0404	.0611	.0573	.0625	.1195	.1202	.1179
	(.0617)	(.0613)	(.0614)	(.0605)	(.0641)	(.0645)	(.0642)
SOC1479	.0640	.0352	.0171	–.0177	.0353	–.0136	– .0145
	(.0293)	(.0287)	(.0289)	(.0290)	(.0296)	(.0296)	(.0297)
SOC1579	.1968	.1828	.1112	.1032	.1871	.1142	.1072
	(.0275)	(.0271)	(.0278)	(.0280)	(.0074)	(.0288)	(.0292)
COLL79	.3134	.3024	.2056	.2029	.3049	.2097	.2052
	(.0209)	(.0205)	(.0216)	(.0217)	(.0215)	(.0277)	(.0227)
ADV79	.6264	.6160	.4904	.4779	.6177	.4883	.4837
	(.0643)	(.0643)	(.0646)	(.0642)	(.0659)	(.0665)	(.0663)
				Students in academic programs			
VOC79	.0858	.1572	.0505	.0411	.2551	.2225	.2179
	(.1590)	(.1564)	(.1589)	(.1569)	(.1725)	(.1727)	(.1721)
SOC1479	.00925	–.0247	–.0731	–.0804	–.0283	–.0659	–.0779
	(.0743)	(.0716)	(.0713)	(.0714)	(.0761)	(.0755)	(.0754)
SOC1579	.1931	.1669	.1142	.1060	.1668	.1187	.1090
	(.0567)	(.0562)	(.0577)	(.0576)	(.0625)	(.0645)	(.0644)
COLL79	.2722	.2534	.1773	.1726	.2607	.1893	.1815
	(.0547)	(.0536)	(.0552)	(.0551)	(.0595)	(.0610)	(.0609)
ADV79	.5433	.5333	.4162	.3983	.5443	.4291	.4198
	(.0811)	(.0797)	(.0816)	(.0810)	(.0889)	(.0921)	(.0918)
				Students in nonacademic programs			
VOC79	.0272	.0666	.0616	.0677	.0281	.0286	.0276
	(.0644)	(.0641)	(.0634)	(.0639)	(.0698)	(.0704)	(.0701)
SOC1479	.0923	.0609	.0321	.0312	.0690	.0409	.0414
	(.0304)	(.0298)	(.0303)	(.0304)	(.0332)	(.0342)	(.0342)
SOC1579	.1491	.1490	.1102	.1027	.1373	.1006	.0966
	(.0345)	(.0345)	(.0348)	(.0355)	(.0375)	(.0384)	(.0386)
COLL79	.2877	.2909	.2324	.2283	.2921	.2377	.2383
	(.0299)	(.0300)	(.0310)	(.0313)	(.0340)	(.0364)	(.0362)
ADV79	.7301	.7546	.6786	.7012	.8399	.7895	.8112
	(.3376)	(.3425)	(.3429)	(.3485)	(.3740)	(.3731)	(.3749)

Notes and source: See text tables 2 and 3 and appendix table 2.

REFERENCES

Akin, John S., and Irwin Garfinkel, 1977, "School expenditures and the economic returns to schooling," *Journal of Human Resources*, Vol. 12, Winter, pp. 460–477.

Altonji, Joseph G., 1988, "The effects of family background and school characteristics on education and labor market outcomes," Northwestern University, unpublished paper, December.

Angrist, J., and A. Krueger, 1991, "Does compulsory schooling attendance affect schooling and earnings?," *Quarterly Journal of Economics*, Vol. 106, No. 4, November, pp. 979–1014.

Ashenfelter, Orley, and Alan Krueger, 1991, "Estimates of the economic return to schooling from a new sample of twins," National Bureau of Economic Research, working paper, No. 4143.

Ashenfelter, Orley, and Cecilia Rouse, 1997, "Income, schooling, and ability: Evidence from a new sample of identical twins," National Bureau of Economic Research, working paper, No. 6106, July 1.

Card, David, and Krueger, Alan, 1992, "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *Journal of Political Economy*, Vol. 100, No 1, February, pp. 1–40.

Griliches, Zvi, 1979, "Sibling models and data in economics: Beginnings of a survey," *Journal of Political Economy*, Vol. 87, October, pp. S37–64.

__________, 1977, "Estimating the returns to schooling: Some econometric problems," *Econometrica*, Vol. 45, January, pp. 1–22.

Grubb, W. Norton, 1993, "The varied economic returns to postsecondary education," *Journal of Human Resources*, Vol. 20, No 2, Spring, pp. 363–382.

Hanushek, Eric, 1986, "The economics of schooling," *Journal of Economic Literature*, Vol. 24, No. 3, September, pp. 1141–1177.

Johnson, George, and Frank Stafford, 1973, "Social returns to quantity and quality of schooling," *Journal of Human Resources*, Vol. 8, Winter, pp. 139–155.

Kain, Thomas, and Cecilia Rouse, 1995, "Labor market returns to two- and four-year colleges," *American Economic Review*, Vol. 85, No. 3, June, pp. 600–615.

Lang, Kevin, 1993, "Ability bias, discount rate bias, and the return to education," Boston University, Department of Economics, unpublished paper, May.

Morgan, James, and Ismail Sirageldin, 1968, "A note on the quality dimension in education," *Journal of Political Economy*, September / October, pp. 1069–1077.

Murphy, K., and F. Welsh, 1992, "The structure of wages," *Quarterly Journal of Economics*, Vol. 7, February, pp. 284–326.

Olneck, Michael R., 1979, "The effects of education," in *Who Gets Ahead?: The Determinants of Economic Success in America*, C. Jencks (ed.), New York: Basic Books.

Rosenbaum, Dan T., 1997, "Ability, schooling ranks, and labor market trends: Controlling for changes in the education distribution when making comparisons between groups over time," Northwestern University, unpublished paper, November.

Siebert, W. Stanley, 1985, "Developments in the economics of human capital," in *Labour Economics*, London and New York: Longman.

U.S. Department of Education, National Center for Education Statistics, 1972–86, *National Longitudinal Survey of the High School Class of 1972.*

Willis, Robert J., 1986, "Wage determinants: A survey and reinterpretation of human capital earnings functions," in *Handbook of Labor Economics*, O. Ashenfelter and R. Layard (eds.), Amsterdam: North-Holland, pp. 525–602.