Short-Run Effects of Money When Some Prices Are Sticky

Lee E. Ohanian and Alan C. Stockman

Much of the literature in macroeconomics is concerned with the effects of monetary disturbances on the real economy, particularly the role of money in business cycles. Monetary shocks can have important real effects in "Keynesian" models because this class of models generally involves nominal rigidities in prices or wages. In sharp contrast, a broad neoclassical tradition in macroeconomics (including real business cycle theory) typically assumes prices are completely flexible, although there have been some recent attempts to combine these traditions, as in Kydland (1987), Cho and Cooley (1990), and King (1990).

While there is much evidence that certain types of goods have sticky nominal prices, there is also evidence of frequent price changes for other types of goods, such as the relatively homogeneous commodities sold on near-auction markets, food, automobiles (transactions prices), and computers. Typically, Keynesian macroeconomic studies postulate a sticky price level, so that a change in the nominal money supply is (in the short run) a change in the real money supply. These studies generalize from the evidence that *some* prices are sticky to the hypothesis that the *general price level* is sticky. In contrast, neoclassical studies often assume flexible prices, so that the price level adjusts immediately to changes in the nominal money supply. These studies typically ignore or discount the evidence that certain prices are sticky.

Studies presenting evidence that certain nominal prices are "sticky" include Stigler and Kindahl (1970), which found evidence of infrequent changes in transactions prices in product markets, and Carlton (1986), which extended the Stigler-Kindahl study and documented slow changes in nominal transactions

Lee E. Ohanian is assistant professor of economics at the University of Pennsylvania, and Alan C. Stockman is the Marie C. and Joseph C. Wilson Professor of Economics at the University of Rochester. They would like to thank Tim Cook, Mike Dotsey, Tom Humphrey, and Peter Ireland for very helpful comments, and Mark Gertler for very helpful discussions on an earlier draft. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

prices for producers' goods even without long-term relationships between buyers and sellers and showed that delivery lags and other product characteristics often change before or in place of changes in nominal prices. Carlton's result (1989, p. 921) that the degree of price rigidity differs greatly across industries (with the average period between price changes ranging from 5.9 months for household appliances to 19.2 months for chemicals) is one motivation for our assumption below that sectors differ in their speed of price adjustment. Other papers include Cecchetti's (1986) study of stickiness in nominal magazine prices, Kashyap's (1991) study showing substantial price sluggishness in three major mail-order catalogs (even when new catalogs are published), Rees's (1961) study providing evidence that catalog prices and retail-store prices have similar properties, and Blinder's (1991) survey that found that most firms change nominal prices one time or less in a typical year. Other evidence for nominal price sluggishness includes the well-known fact that prices are seldom formally indexed to a price index and the fact that real exchange rates (exchange rate-adjusted ratios of price indexes across countries) vary much more under floating exchange rate systems than under pegged exchange rate systems (see Stockman [1983], Mussa [1986], and Baxter and Stockman [1989]). This evidence strongly suggests that the exchange rate system affects international relative-price variability, a fact that is easy to explain with models in which some prices are sticky and much harder to explain in flexible-price models. Related evidence appears in Engel (1991). On the other hand, many sectors of the U.S. economy appear to have very flexible prices—with nominal prices that often change weekly, daily, or every few minutes.

This article studies a hybrid model in which some nominal prices are sticky and others are flexible. This model turns out to have several interesting properties. Unexpected changes in the money supply change the relative prices of sticky-price and flexible-price goods, so the real effects of monetary disturbances can differ across sectors. With certain parameter configurations, the model has the ability to produce *endogenous* price sluggishness in the flexible-price sector because the equilibrium response of those prices to a change in money is small in the short run. With other parameter configurations, the response of flexible nominal prices to a monetary disturbance is sufficiently large that the change in real money balances is small, as are monetary effects on the real economy working through the standard Keynesian transmission mechanism. In that case, however, a monetary disturbance has large effects on relative prices and induces different responses of output in different sectors of the economy. Monetary shocks, in this way, may contribute to sectoral shifts in the economy. Nominal price sluggishness also affects the short-run response of the economy to real disturbances (e.g., to changes in technology), even in sectors of the economy with flexible prices.

Because there is currently no well-established theory to explain nominal price stickiness, we follow Svensson (1986), Lucas (1991), Lucas and

Woodford (1992) and Cho and Cooley (1990) in assuming that nominal prices in the sticky-price sector are set one period in advance.[1] (We assume that the *implications* of sluggish price adjustment are largely independent of the *source* of that sluggishness.) In contrast to those models, however, the economy we study also has a flexible-price sector with trading at Walrasian prices. An interesting feature of our more general theory (to coin a phrase) is that it encompasses the standard Keynesian model (in one of its guises) and the flexible-price neoclassical model as special cases.

We study two versions of the model. The first version is developed in the spirit of Barro and Grossman (1976): when nominal prices cannot adjust to clear markets, some agents are rationed and output is determined by the minimum of the quantity demanded and the quantity supplied. In this case positive money shocks result in excess demand, with households rationed, and negative money shocks result in excess supply, with producers constrained. The second version we study assumes that output is determined by the quantity demanded. This version of the model is more consistent with recent sticky-price literature such as Blanchard and Kiyotaki (1987), who assume monopolistic competition so that small, positive money shocks leave the sticky-price sector in a situation of demand-determined output and do not imply rationing of buyers as in the Barro-Grossman model. (Instead, firms supply more of the good as long as price exceeds marginal cost.) We show that the effects of positive money shocks differ across the two versions of the model, though the effects of negative money shocks are similar in both versions.

This article does not attempt to match closely the implications of the model with data. Instead, its purpose is to analyze the properties of a simple model with sticky- and flexible-price sectors and to examine how its properties depend on basic parameter values. Consequently, the analysis we present below focuses on the effects of isolated, exogenous monetary disturbances.

1. A SIMPLE COMPETITIVE EQUILIBRIUM MODEL

We begin with a simple flexible-price equilibrium model that we have also examined in Ohanian and Stockman (1994) and (in a two-country framework) in Stockman and Ohanian (1993). The model has two consumption goods, X and Y, and labor. We introduce money through a cash-in-advance constraint, intended to stand in for a more general transactions model of money. We assume, for simplicity, that there are complete asset markets. The representative household maximizes utility:

[1] In contrast to Lucas and Woodford, we simply assume the level at which nominal prices are predetermined in the sticky-price sector rather than deriving an endogenous distribution of prices.

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{(1-\rho)} \left[\alpha x_t^{(\sigma-1)/\sigma} + (1-\alpha) y_t^{(\sigma-1)/\sigma} \right]^{[\sigma/(\sigma-1)]\cdot(1-\rho)} - v(L_{Xt} + L_{Yt}) \right] \tag{1}$$

subject to the two constraints

$$n_{t-1} + \tau_t + P_{X,t-1} k_{X,t-1} L^{\delta}_{X,t-1} + P_{Y,t-1} k_{Y,t-1} L^{\delta}_{Y,t-1} - M_t + \nu_t (q_t + d_t) - \nu_{t+1} q_t = 0 \tag{2}$$

and

$$M_t - P_{Xt} X_t - P_{Yt} Y_t \geq 0 \tag{3}$$

each period. Equation (2) is a budget constraint for period t asset markets and (3) is the cash-in-advance constraint which applies to period t product markets (which immediately follow period t asset markets as in Lucas [1982]). The terms x and y refer to consumption of goods X and Y, L_X and L_Y refer to labor hours producing goods X and Y, $0 \leq \delta < 1$ is a parameter of the production function, k_X and k_Y are exogenous productivity parameters, n_{t-1} refers to the household's money holdings at the end of period $(t-1)$ product markets (which is the slack in inequality [3] from the previous period and equals *zero* in our equilibrium), τ refers to a lump-sum transfer of money to the household from the government, P_X and P_Y are nominal prices, M_t is the nominal money the household chooses as it leaves period t asset markets and enters period t product markets, and ν_t is a vector of other assets the household owns at the beginning of period t, with dividend vector d and ex-dividend price-vector q.[2]

Several important parameters that we will focus on later appear in (1) and (2). First, α is a parameter describing tastes. Because α helps determine the equilibrium share of good X in total output, we will vary it in "The Size of the Sticky-Price Sector" subsection of Section 2 to discuss changes in the relative sizes of the X and Y industries. Next, ρ is the inverse of the intertemporal elasticity of substitution; an increase in ρ means households are less willing to trade current consumption for future consumption (that is, they are willing to

[2] One can also think of k_X and k_Y as fixed levels of the capital stock. However, adding capital accumulation to the model would change its implications in several ways. The most obvious change would occur in the dynamics of adjustment to equilibrium following a disturbance. In addition, changes in the rate of capital accumulation would provide an additional margin of substitution for the economy that could tend to smooth consumption over time and thereby reduce the response of interest rates to exogenous disturbances. By abstracting from capital accumulation, the current article greatly simplifies the analysis. The benefit of this simplicity is that it facilitates understanding; the cost is that it may lead to slightly different quantitative results than a more complicated model with capital accumulation. We are currently extending the model to include capital and will report on the results in a forthcoming paper.

pay more for a more constant consumption stream). The subsection "The Size of Intertemporal Substitution" explains how the size of ρ affects our results. Third, σ is the elasticity of substitution between goods X and Y; a larger σ means the goods are better substitutes. The impact of the size of σ on our results is the subject of the subsection "The Size of Intratemporal Substitution." Finally, δ determines the curvature of the production function, with lower values of δ indicating higher degrees of diminishing returns to labor; the subsection "The Degree of Curvature in Production" discusses the impact of this parameter on our results.[3]

We assume that the cash-in-advance constraint (3) holds as an equality, $k_X = k_Y = 1$ for all t, and that $\tau \equiv 0$. The *flexible-price perfect foresight equilibrium* for this simple production economy satisfies

$$M_t^s = P_{Xt}L_{Xt}^{\delta} + P_{Yt}L_{Yt}^{\delta}, \tag{4}$$

$$P_{Xt}\lambda_t = \left[\alpha L_{Xt}^{\delta(\sigma-1)/\sigma} + (1-\alpha)L_{Yt}^{\delta(\sigma-1)/\sigma}\right]^{(1-\rho\sigma)/(\sigma-1)} \alpha L_{Xt}^{-\delta/\sigma}, \tag{5}$$

$$P_{Yt}\lambda_t = \left[\alpha^{(\sigma-1)/\sigma} L_{Xt}^{\delta} + (1-\alpha)L_{Yt}^{\delta(\sigma-1)/\sigma}\right]^{(1-\rho\sigma)/(\sigma-1)} (1-\alpha) L_{Yt}^{-\delta/\sigma}, \tag{6}$$

$$v = \beta P_{Xt}\delta L_{Xt}^{\delta-1}\lambda_{t+1}, \tag{7}$$

and

$$v = \beta P_{Yt}\delta L_{Yt}^{\delta-1}\lambda_{t+1}, \tag{8}$$

where M_t^s is the (exogenous and constant, because $\tau = 0$) money supply at the end of period t asset markets and λ is the current-value Lagrange multiplier on constraint (2). (It is easy to show that $\lambda = \gamma$, the multiplier on the cash-in-advance constraint, because of the first-order condition for the choice of M_t.) It is also easy to show that the nominal interest rate on a one-period nominal asset satisfies the usual pricing condition:

$$1 + i_t = \frac{\lambda_t}{\beta\lambda_{t+1}}. \tag{9}$$

2. EQUILIBRIUM WHEN SOME PRICES ARE STICKY

This section examines the implications of the basic model when prices in one sector are predetermined (for one period) at the expected market-clearing level. We assume for now that output is determined by the minimum of quantity demanded and quantity supplied. We return to this assumption later and modify it so that output is always demand determined.

[3] The other parameters in equation (1), β and v, have no important effects on our results.

We introduce short-term price stickiness into the model by assuming that sellers must choose $P_{X,t}$ one period in advance (that is, at the end of period $t-1$). We assume, however, that the nominal price of Y, P_Y, adjusts instantaneously to clear markets at each date. We examine the effects of a permanent, unanticipated change in the money supply starting from a nonstochastic steady-state equilibrium in which the money supply is constant and P_X is fixed at its expected equilibrium level. The money supply change occurs at the beginning of period t. Real variables dated at $t+1$ and later are unaffected by this change, but real variables at date t are affected because $P_{X,t}$ is predetermined.

First consider the excess-supply case. Suppose the money supply falls permanently by 1 percent at date t, with $P_{X,t}$ fixed for one period. When $P_{X,t}$ is above its equilibrium level, the quantity of X supplied exceeds the quantity demanded, so output of X will be *demand determined*. As a result, equation (7) (describing the supply of X) does not hold. That is, people would like to work more in the X industry and sell more of product X, but the price is predetermined at a level above the equilibrium, so the quantity demanded is insufficient to satisfy supply. Instead, sellers are rationed (equally in equilibrium). So we have equations (4)–(6) and (8) in the four variables $L_{X,t}, L_{Y,t}, P_{Y,t}$, and λ_t, (with λ_{t+1} taking its new steady-state value).

Because a change in the money supply has no steady-state effect on x, y, or L_X, equation (5) implies that the change in money has no effect on $P_{X,t+1}\lambda_{t+1}$ in the new equilibrium. But the fall in the money supply reduces $P_{X,t+1}$ by 1 percent, so it must raise λ_{t+1} by 1 percent.

Our first result is the following: a fall in the money supply reduces P_Y, and the percentage fall in P_Y is less than the percentage fall in the money supply if and only if the elasticity of substitution in consumption, σ, exceeds one. This means that if X and Y are good substitutes (relative to the Cobb-Douglas case of $\sigma = 1$), then exogenous price stickiness in the X sector causes *endogenous* price stickiness in the Y sector. The overall price level also adjusts sluggishly in this case.

Our second result is that a fall in the money supply causes a rise in the one-period nominal interest rate if and only if the degree of relative risk aversion, ρ, exceeds one, that is, if and only if the elasticity of intertemporal substitution $(1/\rho)$ is less than one.

Next, consider the excess-demand case. Suppose the money supply rises permanently by 1 percent at date t, with $P_{X,t}$ fixed for one period. When $P_{X,t}$ is below its equilibrium level, the quantity of X demanded exceeds the quantity supplied, so output of X is *supply determined*. As a result, equation (5) (describing the demand for X) does not hold. Instead, buyers are rationed (equally in equilibrium) and we have equations (4) and (6)–(8) in the (same) four variables $L_{X,t}, L_{Y,t}, P_{Y,t}$, and λ_t, (with λ_{t+1} taking its new steady-state equilibrium value).

Our first result for the excess-demand case is that the equilibrium response of P_Y to a rise in the money supply, is:

$$\frac{d \ln P_Y}{d \ln M} = 1 + \frac{s_X}{1 - s_X}, \tag{10}$$

where s_X is the share of good X in total spending. This means that a rise in the money supply raises the nominal price of Y by *more* than it would if the price of X were fully flexible and that this "overshooting" of P_Y is increasing in the share of the economy with sticky prices.

Our second result in the excess-demand case is that the overshooting of P_Y *necessarily* implies an inverse effect of money on interest rates. To see why, consider the pricing relationship for a one-period nominal bond:

$$\frac{1}{(1+i)} = \beta \, \frac{U_{y,t+1}}{U_{y,t}} \, \frac{P_{y,t}}{P_{y,t+1}}, \tag{11}$$

where $U_{y,t}$ denotes marginal utility of good y at date t. Two factors affect the nominal interest rate: the marginal rate of substitution between Y today and Y tomorrow and the rate of change of the nominal price of Y. In the excess-demand case, both factors tend to decrease the interest rate. First, note that households are rationed in purchasing X, so substitution into Y today raises the marginal rate of substitution, which reduces the interest rate. Second, $P_{y,t}$ overshoots the new equilibrium level, $P_{y,t+1}$, which results in (expected) deflation in the Y sector, which also tends to reduce the nominal interest rate.

As long as $\delta < 1$, which means that there are diminishing returns to labor, the changes in labor inputs in response to a positive money shock are

$$\frac{d \ln L_Y}{d \ln M} = \left(\frac{s_X}{1 - s_X}\right)\left(\frac{1}{1-\delta}\right) \tag{12}$$

and

$$\frac{d \ln L_X}{d \ln M} = \left(\frac{-1}{1-\delta}\right). \tag{13}$$

In this case a positive monetary disturbance moves labor from the sticky-price sector (X) to the sector with the rising relative price (Y). Because output is supply determined, it is interesting to note that the elasticity of substitution between the two goods does not affect the sectoral reallocation of labor between the X and Y industries.

Effects of a Fall in the Money Supply

The analytic results available for this model are limited, so we now turn to a numerical analysis of the model. Consider a permanent 1 percent fall in the money supply (from 10 to 9.9), starting from a steady-state equilibrium. Table 1 shows the results when $\alpha = 0.5,\ \sigma = 2,\ \delta = 0.64,\ v = 1,\ \beta = 0.96,$

Table 1 Baseline Model

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.4869	0.4761	0.4869	−2.21
Labor in Y	0.4869	0.489	0.4869	0.43
Total labor	0.9738	0.9651	0.9738	−0.89
Output of X	0.6309	0.6219	0.6309	−1.42
Output of Y	0.6309	0.6326	0.6309	0.27
GNP	1.262	1.255	1.262	−0.58
Price of Y	7.925	7.858	7.846	0.15
CPI	7.925	7.891	7.846	0.58
Interest rate	4.167	4.771	4.167	0.60

and $\rho = 2$.[4] (We analyze permanent changes in the money stock to eliminate labor-supply responses that reflect temporary inflation tax considerations.)

The first column of Table 1 shows the variables: labor inputs in the X and Y industries, total labor, output in each industry (x and y) and total real GNP (evaluated at equilibrium prices and production shares), the nominal price of good Y (the nominal price of X in the old steady state and the short run equals the old steady-state price of good Y, and the new steady-state prices are also equal), the economy's consumer price index, and the nominal interest rate (in percent per period). The next column shows the old steady-state ("Old SS") levels of the variables, before the change in money. The "SR" column shows the short-run effects of the fall in money (while the nominal price of X is fixed at its previous level). The "New SS" column shows the new steady state. The column labeled "ratio" shows the percentage by which each variable in the short run exceeds its new steady-state level—except for the interest rate row in which the "ratio" shows the absolute *difference* between the interest rate in the short run and in the long run.

With the parameter values in Table 1, the sticky-price sector (X) represents one-half of output in the economy. Half of all labor is employed in the sticky-price sector. A permanent 1 percent fall in the money supply is neutral in the long run (with a 1 percent fall in nominal prices and no effects on real variables). But in the short run, with p_X predetermined, real GNP falls about 0.58 percent. This fall in total GNP masks major differences across sectors:

[4] The value of β in Table 1 is appropriate if prices in the X sector are sticky for about one year. If, instead, they are sticky for about one quarter, then a more appropriate level of β is 0.99. An unexpected change in the money supply of about 1 percent *per quarter* with prices sticky for one quarter has nearly the same effects as an unexpected change of about 1 percent *per year* when prices are sticky for a year.

output in the sticky-price sector falls 1.4 percent, while output in flexible-price industries rises 0.27 percent.[5] The fall in money reduces the nominal price of Y, which raises the relative price of X. This leads to a fall in the quantity of X demanded and creates excess supply in the X industry. Output of X is determined by the minimum of the quantity demanded and the quantity supplied, so output of X falls. But consumers substitute (partly) into purchases of Y, so output of Y rises. Notice that the nominal price of Y falls by almost exactly the amount it would fall if the price of X were flexible (it falls by almost 1 percent—to about one-seventh of 1 percent above its new steady-state level). Because the nominal price of Y responds almost proportionally to the change in the money supply while the nominal price of X is fixed and because each sector represents one-half of the economy's output, the CPI falls about halfway to its new long-run level.

As in standard Keynesian models, the fall in the money supply has a short-run "liquidity effect" on the nominal interest rate. In Table 1, the interest rate rises 60 basis points from 4.17 percent to 4.77 percent in the short run. This increase is slightly higher than the estimates reported by Christiano and Eichenbaum (1992), who estimate that a lower bound for the liquidity effect is that a 1 percent fall in the money supply raises the federal funds rate by about 27–53 basis points (within one to two quarters). Because expected inflation is negative (the CPI is expected to fall another 0.6 percent), this represents a rise in the real interest rate (measured in terms of the output bundle) of about 120 basis points. Notice that the liquidity effect occurs despite the introduction of money through a cash-in-advance constraint, which (when binding as in these examples) builds in a zero interest elasticity of the demand for money. Ohanian and Stockman (1994) examine the question of how much price stickiness is necessary to generate a liquidity effect of money on interest rates of realistic size and find that only a small sticky-price sector can be sufficient to produce interest rate effects of the magnitude found in the data.

Table 1 provides an initial answer to one of our central questions: Are nominal prices in flexible-price sectors "sluggish" in response to monetary and real disturbances—so that relative prices remain close to their equilibrium levels—or do nominal prices in flexible-price sectors change more than proportionally to monetary disturbances—so that the overall price level adjusts to equate the supply of and demand for money? The answer provided by Table 1 is a compromise between these two possible responses: P_Y is not endogenously sluggish, but neither does it change more than proportionally to the monetary disturbance. As a result, the overall price level exhibits a degree of sluggishness

[5] Because the capital stock and technology are fixed in this experiment, the marginal product of labor rises in the sticky-price sector (as employment in that sector falls) and rises in the flexible-price sector (as employment in that sector rises).

at the same time the monetary disturbance contributes to a change in relative prices.[6]

The Size of Intertemporal Substitution

Raising the degree of relative risk aversion from two to three (i.e., lowering the elasticity of intertemporal substitution from one-half to one-third) raises labor effort and output (and lowers nominal prices) in the steady state. However, Table 2 shows that the *responses* of the economy to a fall in the money supply are virtually unchanged, except for a larger liquidity effect on the interest rate. With $\rho = 3$, a 1 percent fall in the money supply raises the nominal interest 121 basis points in the short run, roughly double the response when $\rho = 2$. The other responses of the economy are virtually unaffected. A reduction in the elasticity of intertemporal substitution raises the size of the liquidity effect for a simple reason. A fall in the money supply raises the interest rate in the short run because households become cash constrained: with the smaller money supply, P_X fixed at its old level, and P_Y roughly at its new equilibrium level, households cannot afford to buy as many consumption goods as they did before the fall in money or as many as they will buy after all nominal prices adjust. Households attempt to smooth consumption intertemporally by borrowing. The equilibrium real interest rate is bid up as all households attempt to borrow. The higher real interest rate induces households to accept the temporary reduction in consumption. However, the smaller the elasticity of intertemporal substitution, the larger the increase in the real interest rate required to induce households to accept the temporarily low level of consumption. So increases in ρ raise the interest rate response to a monetary disturbance.

The Size of Intratemporal Substitution

We have assumed that outputs of the two sectors are substitutes in the sense that the elasticity of substitution σ exceeds one. Now suppose that $\sigma = 0.5$ rather than 2. Table 3 shows that reducing the elasticity of substitution from two to one-half has several effects on the economy's response to a monetary disturbance. First, output in the flexible-price sector now *falls* along with

[6] The effects of increases in productivity in this model differ from the effects in either standard Keynesian models or neoclassical models. Suppose productivity rises permanently by 1 percent in each sector: output is 1 percent higher for each level of labor input. In the long run, this reduces nominal prices and employment in each sector and raises output in each sector, with no permanent effect on the interest rate. (Labor input is constant in response to a productivity change if we assume log utility, in which case income and substitution effects are offsetting.) But with the nominal price P_X fixed in the short run, the relative price of X rises as the nominal price of Y falls. As a result, an economy-wide rise in productivity can *reduce* output in the sticky-price sector in the short run. It also raises output in the flexible-price sector more in the short run than in the long run and temporarily raises nominal and real interest rates.

Table 2 Baseline Model, $\rho = 3$

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.718	0.7078	0.718	−1.42
Labor in Y	0.718	0.7199	0.718	0.27
Total labor	1.436	1.428	1.436	−0.58
Output of X	0.5959	0.5827	0.5959	−2.21
Output of Y	0.5959	0.5985	0.5959	0.43
GNP	1.192	1.181	1.192	−0.89
Price of Y	6.964	6.905	6.894	0.15
CPI	6.964	6.934	6.894	0.58
Interest rate	4.17	5.38	4.17	1.21

Table 3 Baseline Model, $\sigma = 0.5$

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.4869	0.4814	0.4869	−1.13
Labor in Y	0.4869	0.4856	0.4869	−0.28
Total labor	0.9738	0.967	0.9738	−0.70
Output of X	0.6309	0.6263	0.6309	−0.73
Output of Y	0.6309	0.6298	0.6309	−0.18
GNP	1.262	1.256	1.262	−0.45
Price of Y	7.925	7.838	7.846	−0.10
CPI	7.925	7.881	7.846	0.46
Interest rate	4.167	4.64	4.167	0.47

output in the sticky-price sector. Second, output in the sticky-price sector falls much less when $\sigma = ½$ than when $\sigma = 2$. The reason for these differences is straightforward. When a fall in the money supply reduces the nominal price of Y but not the price of X, households substitute out of consumption of X into consumption of Y. When X and Y are good substitutes, there is a large increase in the demand for Y and a large decrease in the demand for X, which raises equilibrium output of Y and causes output of X to fall by a large amount. If, however, the goods are complements in the sense that an increase in consumption of one of the goods raises the marginal utility of the other good, then the fall in equilibrium consumption of Y reduces the marginal utility of consuming X. Instead of rising, the demand for X falls and equilibrium output of X also falls. The fall in demand for Y is smaller in this case, and equilibrium output of Y falls less than it would if X and Y were good substitutes. This also explains why, with $\sigma = ½$, the price of Y falls more (overshooting its new equilibrium level), whereas if $\sigma = 2$, P_Y falls only partway to its new equilibrium level.

The Degree of Curvature in Production

Table 4 presents the results of the same experiment as in Table 1, but with $\delta = 0.9$ rather than $\delta = 0.64$. This means that the economy experiences only a small degree of diminishing returns to labor. The assumption that $\delta = 0.64$ is more appropriate based on long-run studies of aggregate production functions, but some time-series estimates suggest a higher value of δ in the short run. While the steady state of the economy with $\delta = 0.9$ differs from that presented in Table 1, the *response* of the economy to a monetary disturbance is similar. Total employment falls less, because changes in employment do not so quickly result in diminishing returns. Employment in the sticky-price sector falls less for the same reason. Because output of the flexible-price good rises more in this case, the price P_Y falls less.

Table 4 Baseline Model, $\delta = 0.9$

Variable	Old SS	SR	New SS	Ratio
Labor in *X*	0.6429	0.6324	0.6429	−1.64
Labor in *Y*	0.6429	0.646	0.6429	0.48
Total labor	1.286	1.278	1.286	−0.58
Output of *X*	0.6719	0.662	0.6719	−1.47
Output of *Y*	0.6719	0.6748	0.6719	0.43
GNP	1.344	1.337	1.344	−0.53
Price of *Y*	7.441	7.37	7.367	0.05
CPI	7.441	7.406	7.367	0.53
Interest rate	4.167	4.715	4.167	0.55

The Size of the Sticky-Price Sector

Now consider changing the relative sizes of the two sectors of the economy. Table 5 takes the same parameter values as in Table 1 except $\alpha = 0.2$ rather than 0.5. This implies that the sticky-price sector is about 21 percent of aggregate GNP and accounts for 11 percent of employment. A permanent 1 percent fall in money reduces the nominal price of *Y* by almost 1 percent immediately and reduces employment in the sticky-price sector by 2.86 percent and output by 1.8 percent in the short run. Real GNP falls 0.1 percent and total employment falls 0.23 percent, as output in the flexible-price sector rises 0.07 percent. The liquidity effect (inverse effect of money on interest rates) in Table 5 is smaller than in Table 1, but reaches the lower end of the range estimated by Christiano and Eichenbaum (1992) if the elasticity of intertemporal substitution is reduced to one-third instead of one-half (that is, if $\rho = 3$ rather than 2), in which case the interest rate rises 31 basis points in the short run. The fall in α also raises the percentage response of labor in the sticky-price sector to a monetary

Table 5 Baseline Model, $\alpha = 0.2$

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.1015	0.09862	0.1015	−2.86
Labor in Y	0.7797	0.7806	0.7797	0.11
Total labor	0.8813	0.8792	0.8813	−0.23
Output of X	0.2313	0.2271	0.2313	−1.84
Output of Y	0.8528	0.8534	0.8528	0.07
GNP	1.025	1.024	1.025	−0.10
Price of Y	10.380	10.280	10.270	0.04
CPI	9.754	9.665	9.656	0.10
Interest rate	4.167	4.324	4.167	0.16

disturbance, because it reduces the absolute size of that sector. Similarly, it reduces the percentage response of labor in the flexible-price sector because it raises the absolute size of that sector. With a smaller sticky-price sector, households are less cash constrained by a fall in the money supply, so the interest rate response is smaller. And the smaller the sticky-price sector, the smaller the effect of that sector on the nominal price P_Y. Ohanian and Stockman (1994) examine these issues in greater detail and show that a change in the money supply can have a substantial "liquidity effect" on nominal and real interest rates in the short run even if only a small fraction of the economy has sluggish prices.

Costly Time-to-Move Labor Across Industries

The results discussed above involve substantial short-run movements of labor across industries in response to monetary shocks. Because labor is often costly to reallocate across industries in the short run, we now modify the model so that rapid labor mobility is costly. We assume it takes one period to move labor across sectors unless the worker pays a utility cost of moving equal to

$$v_2\left(\frac{ly}{ly+lx} - \frac{lyss}{lyss+lxss}\right)^2 + v_2\left(\frac{lx}{ly+lx} - \frac{lxss}{lyss+lxss}\right)^2, \tag{14}$$

where *lxss* and *lyss* are the original steady-state levels of employment in the X and Y industries and v_2 is a nonnegative parameter. We assume $v_2 = 10$ and the same parameters as in Tables 1 and 3. Tables 6 and 7 present the results of this experiment.

First, compare Table 6 to Table 1: the costly time-to-move assumption results in a much smaller increase in employment and output in the flexible-price sector; output in that sector is roughly constant (rising only 0.02 percent). Total labor supply falls more than in Table 1, as does GNP. Nominal prices in the flexible-price sector fall less than in Table 1: P_Y remains 0.32 percent above its new equilibrium level in the short run. If the *equilibrium* price response in the flexible-price sector is small, as in this case, studies such as Blinder (1991) that search for "menu costs" or similar reasons for "price-stickiness" in these markets would fail to uncover them because that price sluggishness would reflect an equilibrium response to the fact that *other* nominal prices are sticky. In fact, it is interesting that Blinder's survey found little or no evidence of "menu costs" in changing prices. Instead, firms reported that the reason they change nominal prices infrequently is that they are concerned about their product price relative to those of their "competitors." If we interpret "competitors" to include goods in the sticky-price sector, this observation may be consistent with the results in Table 6.[7]

Next, compare Table 7 to Table 3: the costly time-to-move assumption results in a much smaller fall in employment and output in the flexible-price sector; output in that sector is roughly constant (falling only 0.02 percent rather than 0.18 percent). Total labor supply falls less than in Table 3, as does GNP. Nominal prices in the flexible-price sector fall more than in Table 3: P_Y falls 1.3 percent and overshoots its new equilibrium level by 0.3 percent. The time-to-move assumption in this case reduces the response of interest rates by about 10 basis points. As before, the size of the liquidity effect is governed by the size of intertemporal substitution: if it is one-third rather than one-half, the nominal interest rate rises twice as much as in Tables 6 and 7.

Sticky Wages with Flexible Prices

We have assumed up to now that nominal stickiness in the X sector occurs mainly in product markets. We now modify the model so that nominal rigidities in the X sector have their origin in labor markets. We assume nominal wages in the X sector are predetermined for one period, while nominal product prices in both sectors (and nominal wages in the Y sector) are flexible. Table 8 presents the results of the same experiment as in Table 1 (with the same parameter values), but with sticky wages rather than sticky prices. We assume the nominal wage rate in the X sector is set one period in advance equal to the expected nominal marginal product of labor, which equals the steady-state marginal product multiplied by the *original* steady-state nominal price, P_X. In this sticky-wage economy, employment in the X industry is demand determined in the case of negative money shocks and supply determined in the case of positive money shocks.

[7] It is interesting that most of the reasons given by firms for sluggish nominal prices in Blinder's study deal with *relative* prices. For example, some firms said that price is only one component of an overall package that matters to buyers; others spoke of implicit contracts and so on. These reasons by themselves are not sufficient to explain sticky *nominal* prices. But if some nominal price is exogenously fixed, then these reasons could help explain the "spread" of stickiness to other nominal prices.

Table 6 Baseline Model, $v_2 = 10$

Variable	Old SS	SR	New SS	Ratio
Labor in *X*	0.4869	0.4768	0.4869	−2.08
Labor in *Y*	0.4869	0.4871	0.4869	0.03
Total labor	0.9738	0.9638	0.9738	−1.03
Output of *X*	0.6309	0.6225	0.6309	−1.34
Output of *Y*	0.6309	0.631	0.6309	0.02
GNP	1.262	1.253	1.262	−0.66
Price of *Y*	7.925	7.871	7.846	0.32
CPI	7.925	7.898	7.846	0.67
Interest rate	4.167	4.859	4.167	0.69

Table 7 Baseline Model, $\sigma = 0.5$, $v_2 = 10$

Variable	Old SS	SR	New SS	Ratio
Labor in *X*	0.4869	0.4818	0.4869	−1.05
Labor in *Y*	0.4869	0.4868	0.4869	−0.03
Total labor	0.9738	0.9686	0.9738	−0.54
Output of *X*	0.6309	0.6267	0.6309	−0.67
Output of *Y*	0.6309	0.6308	0.6309	−0.02
GNP	1.262	1.257	1.262	−0.35
Price of *Y*	7.925	7.822	7.846	−0.31
CPI	7.925	7.873	7.846	0.35
Interest rate	4.167	4.531	4.167	0.36

The results in Table 8 differ slightly from those in Table 1, but the main differences are quantitative rather than qualitative. Wage and price stickiness have similar results because the main effect of wage stickiness is to keep the marginal cost of production constant in nominal terms in the short run. This reduces the effect of money on equilibrium nominal product prices in the *X* sector. As a result, the economy resembles a sticky-product-price economy but with *some* nominal price movement, and equilibrium responses to money are smaller than in the sticky-product-price economy. Wage stickiness results in less aggregate labor movement—and less sectoral reallocation than does price stickiness. Total labor falls 0.50 percent in the sticky-wage economy, while it fell by 0.89 percent in the sticky-price economy. Employment in each sector changes by only about half as much in the sticky-wage case as in the sticky-price case. The sticky wages endogenously generate sluggish nominal prices: P_Y, the price of output in the sticky-wage sector, falls immediately by an amount equal to 44 percent of its long-run fall. In this sense, wage stickiness induces

Table 8 Baseline Model with Sticky Wages

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.4869	0.4809	0.4869	−1.23
Labor in Y	0.4869	0.4881	0.4869	0.24
Total labor	0.9738	0.969	0.9738	−0.50
Output of X	0.6309	0.6259	0.6309	−0.79
Output of Y	0.6309	0.6319	0.6309	0.15
GNP	1.262	1.258	1.262	−0.32
Price of X	7.925	7.89	7.846	0.56
Price of Y	7.925	7.852	7.846	0.09
CPI	7.925	7.871	7.846	0.32
Interest rate	4.167	4.502	4.167	0.3352

partial price stickiness. This induced price stickiness is even more pronounced if goods X and Y are less substitutable; Table 9 shows the results of a 1 percent fall in money in the sticky-wage model when $\sigma = 0.5$ (as in Table 4). In this case the short-run fall in P_X is only 28 percent of its long-run fall.

As $\delta \rightarrow 1$, the sticky-price and sticky-wage economies become equivalent. This occurs because a linear production function (with marginal-cost pricing of factors) implies that competitive payments to labor exhaust production. The relative similarity between the sticky-price and sticky-wage economies is consistent with Lucas' (1989) conjecture that the effect of money shocks on sticky-price and sticky-wage economies should be similar.

Table 9 Baseline Model with Sticky Wages, $\sigma = 0.5$

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.4869	0.483	0.4869	−0.81
Labor in Y	0.4869	0.486	0.4869	−0.20
Total labor	0.9738	0.9689	0.9738	−0.50
Output of X	0.6309	0.6277	0.6309	−0.52
Output of Y	0.6309	0.6301	0.6309	−0.13
GNP	1.262	1.258	1.262	−0.32
Price of X	7.925	7.902	7.846	0.72
Price of Y	7.925	7.84	7.846	−0.07
CPI	7.925	7.871	7.846	0.32
Interest rate	4.167	4.503	4.167	0.34

Increases in the Money Supply

Increases in the money supply have qualitatively different effects on the economy because it generates *excess demand* in the sticky-price sector (because increases in nominal prices in the flexible-price sector reduce the relative price of sticky-price goods). As a result, output in the sticky-price sector is determined by the *quantity supplied* rather than the quantity demanded. Table 10 shows the effects of a permanent 1 percent rise in the money supply (from 10 to 10.1), starting from a flexible-price steady state and with the same parameter values as in Table 1.

Table 10 Baseline Model with Money Rising One Percent

Variable	Old SS	SR	New SS	Ratio
Labor in *X*	0.4869	0.4736	0.4869	−2.73
Labor in *Y*	0.4869	0.5002	0.4869	2.73
Total labor	0.9738	0.9738	0.9738	0.00
Output of *X*	0.6309	0.6199	0.6309	−1.75
Output of *Y*	0.6309	0.6419	0.6309	1.74
GNP	1.262	1.262	1.262	0.01
Price of *Y*	7.925	8.082	8.004	0.97
CPI	7.925	8.004	8.004	−0.01
Interest rate	4.167	2.304	4.167	−1.86

Table 10 shows that a 1 percent rise in the money supply causes the nominal price of *Y* to overshoot its new steady-state value. The price of *Y* rises nearly 2 percent in the short run in response to the 1 percent increase in money. This overshooting of the price of *Y* leads the overall price level to respond rapidly to the increase in the money supply: nearly all of the long-run response of the CPI to the increase in money occurs immediately. The model therefore implies that the overall price level responds much more rapidly to a rise in the money supply than a fall (even though the price of *X* is assumed to be sticky upwards as well as downwards). It is interesting to note that this result is consistent with empirical work presented by Fischer (1981) that inflation is positively associated with periods of high relative-price variability. Also, Cody and Mills (1991) (among others) find that spot commodity prices are important predictors of future inflation in U.S. data. This is consistent with the two-sector model economy in that the immediate sharp increase in the price of flexible goods is a "leading indicator" of future changes in prices of sticky goods. Finally, this asymmetric response of prices is consistent with the widely held view that prices are more sticky in a downward direction than in an upward direction. Our model, however, generates this result even though individual prices are either flexible or sticky in *both* directions. This shows how an observer who looks only at the overall price level rather than individual prices could mistakenly conclude that some underlying friction allows prices to rise but not to fall.

While an increase in money raises output and employment in the flexible-price sector, it reduces output and employment in the sticky-price sector.

Aggregate GNP rises only slightly and total labor supply is fixed: the increase in money induces sectoral reallocation of employment and output. The increase in money also causes a large fall in the interest rate: the interest rate falls 186 basis points in response to the 1 percent rise in money. This result suggests that there may be an asymmetry in the size of the liquidity effect of money on interest rates depending on whether the money supply rises or falls. There is also an asymmetry with respect to the effects of intertemporal substitution on the size of the liquidity effect: when the money supply rises, the elasticity of intertemporal substitution has very small effects on the size of the liquidity effect.

Table 11 shows the effects of adding time-to-move labor to Table 10, with $v_2 = 10$. In this case a 1 percent increase in money has little effect on output or employment in either sector, though it again causes considerable overshooting of P_Y and a large (though smaller) liquidity effect: the nominal interest rate falls 105 basis points in response to a 1 percent rise in money. Notice that this interest rate response is smaller than without the time-to-move assumption; in contrast, we found above that adding the time-to-move assumption reduced the inverse effect of money on interest rates in response to negative monetary disturbances.

Table 11 Baseline Model with Money Rising One Percent, $v_2 = 10$

Variable	Old SS	SR	New SS	Ratio
Labor in X	0.4869	0.4864	0.4869	−0.10
Labor in Y	0.4869	0.4874	0.4869	0.10
Total labor	0.9738	0.9738	0.9738	−0.00
Output of X	0.6309	0.6305	0.6309	−0.06
Output of Y	0.6309	0.6313	0.6309	0.06
GNP	1.262	1.262	1.262	0.00
Price of Y	7.925	8.083	8.004	0.99
CPI	7.925	8.004	8.004	−0.00
Interest rate	4.167	3.114	4.167	−1.05

Increases in the Money Supply with Monopolistic Competition

A number of recent papers have studied economies with sticky prices and firms that have market power (Blanchard and Kiyotaki 1987; Mankiw 1985; Svensson 1986). In these papers, an increase in the money supply does not necessarily lead to rationing of consumers (as in our version of the Barro-Grossman model with excess demand). Instead, firms willingly supply output equal to the quantity demanded, provided that the fixed product price exceeds the marginal cost of production. (If marginal cost does exceed price, consumers are rationed and output is supply determined.) While the implications of a negative money shock are similar in both setups (in either case, firms would like to sell more than is demanded at the fixed price), the effect of an unexpected increase in the money supply differs across these two sticky-price models.

This section discusses the effects of an increase in money in our model economy modified so that output in the X industry is always determined by quantity demanded. This amounts to assuming that the X sector is monopolistically competitive, with price exceeding marginal cost. Rather than solving for the explicit equilibrium of a monopolistic-competition model, we can consider an experiment in which steady-state P_X is set above its market-clearing level. In this case, a small positive money shock leads to an expansion in output in the X industry. The results of this experiment are nearly identical, with a change in sign, to the results reported earlier for reductions in the money supply (creating a case of excess supply).

Increases in money have strikingly different effects on the economy depending on whether output in the sticky-price sector is demand determined (as it is here) or supply determined (as in the experiments reported earlier). In both cases, there is a sectoral reallocation of labor. In the supply-determined case discussed earlier, a positive money shock causes labor to move from the sticky-price sector (X) to the flexible-price sector (Y), which has a rising nominal (and relative) price. With monopolistic competition (where output is demand determined even in the case of a positive money shock), labor moves in the opposite direction: labor flows from the flexible-price sector to the sticky-price sector. In addition, fluctuations in output and interest rates due to monetary shocks in the monopolistic competition case are symmetric (in contrast to the asymmetric results reported above) because demand determines production regardless of the sign of the disturbance.

Without taking a stand on the market structure in sticky-price industries in the U.S. economy, these models suggest some interesting tests. Do monetary disturbances affect industries identically? Do they have asymmetric effects on the economy, depending on whether the disturbances are positive or negative? What are the characteristics of the sectoral flow of labor over the business cycle? Kretzmer (1989) provides evidence related to the first question: he finds that monetary shocks affect different industries differently. Using unanticipated-money regressions for individual industries, his evidence suggests that output and employment initially *decline* in response to a positive money shock in almost 40 percent of the industries.[8] This finding could be consistent with either of the market structures discussed above; to distinguish between the two structures, we would need to know whether the industries that contract in response to a positive money shock are sticky-price or flexible-price industries. Cover (1992) and Thoma (1992) present aggregate evidence on the second question:

[8] See Kretzmer's Table 2 (p. 288). He reports that 13 of 30 impact coefficients in his hours equations and 11 of 30 impact coefficients in his output equations are negative.

they find that positive and negative money shocks affect the economy differently. In particular, it appears that negative money shocks precede economic declines, but positive money shocks do not suggest significant increases in future output. This is consistent with our two-sector model with output determined by the short side of supply and demand, but not with the two-sector monopolistic-competition model. (Ball and Mankiw [1992] modify a one-good monopolistic-competition model so that positive trend inflation combined with menu costs triggers price adjustments that depend on the sign of the monetary shock.)

Comparison of Results to a One-Sector Model

It is interesting to contrast the implications of this two-sector model to those of a one-sector model. Recall that a one-sector model is a *special case* of our model, with the parameter $\alpha = 1$. The utility function becomes

$$U = \frac{c^{1-\rho} - 1}{1 - \rho} - vL, \tag{15}$$

where L is labor supply. Perhaps the most interesting comparisons are the effects of money on interest rates. If money falls unexpectedly and permanently, output is demand determined and (assuming the cash-in-advance constraint binds) output and consumption fall proportionally with the decrease in money. The interest rate is

$$\left[\frac{P_t}{P_{t+1}}\right]^{\rho-1} \frac{1}{\beta} = 1 + i. \tag{16}$$

With $\rho = 2$, the interest rate rises one for one with the fall in money. That is, a 1 percent fall in the money supply raises the nominal interest rate 1 percent, which is more than in the two-sector model. In the excess-demand case (resulting from an unexpected permanent increase in money), the differences are even more striking. In the one-sector model, labor supply falls in percentage terms by $\frac{1}{1-\delta} d \ln m$, which implies that output also falls proportionally to the change in money in the excess-demand case. Because money has increased, the cash-in-advance constraint clearly does not bind. As a result, the interest rate falls to zero for *any* increase in money.

3. CONCLUSIONS

There is considerable evidence that the nominal prices of some goods change very infrequently, while nominal prices of other goods change on a daily basis. This article presents a simple two-sector monetary economy with production in which the degree of price flexibility differs by sector. In the excess-demand/excess-supply setup, the model predicts that unanticipated monetary shocks cause (1) asymmetric changes in output and employment that depend on whether money has increased or decreased, (2) changes in relative prices over the cycle, (3) sectoral reallocation of labor, and (4) a significant liquidity effect of money on interest rates. In a related paper (Ohanian and Stockman 1994) we show that only a small amount of price stickiness is needed in this economy to generate a liquidity effect of reasonable size. The asymmetric effects of money in the model are consistent with recent empirical studies (Cover 1992; Thoma 1992), and the model is also consistent with the finding reported by Fischer (1981) that periods of significant inflation are associated with high relative-price variability. The model also reproduces the empirical finding documented by Cody and Mills (1991) that changes in the prices of flexible-price goods (spot commodities) predict future aggregate price changes and the empirical finding by Kretzmer (1989) that output and employment in a significant fraction of U.S. industries decline initially in response to a positive money shock. In the case that *X* and *Y* are good substitutes, the model also has the interesting implication that the price level is "sticky" downwards as an equilibrium phenomenon.

In our continuing work, we hope to show that even very short-term price stickiness can set into motion forces that lead to longer-lasting effects on real interest rates, output and employment, and nominal price changes. To study this, we are currently studying the effects of adding capital to the basic model described here. Our extension of this model to a two-country world in Stockman and Ohanian (1993) examines the effects of monetary disturbances on domestic and world interest rates, exchange rates, and domestic and foreign output, and shows that the effects of monetary disturbances are highly non-linear in open economies. In future work we intend to use the model to study the effects of alternative exchange rate systems, devaluations, and optimal currency areas.

REFERENCES

Ball, Lawrence, and N. Gregory Mankiw. "Asymmetric Price Adjustment and Economic Fluctuations," Working Paper. Boston: Harvard University, 1992.

Baxter, Marianne, and Alan Stockman. "Business Cycles and the Exchange-Rate Regime: Some International Evidence," *Journal of Monetary Economics,* vol. 23 (May 1989), pp. 377–400.

Barro, Robert J., and Herschel I. Grossman. *Money, Employment, and Inflation.* New York: Cambridge University Press, 1976.

Blanchard, Olivier, and Nobu Kiyotaki. "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review,* vol. 77 (September 1987), pp. 647–66.

Blinder, Alan. "Why Are Prices Sticky? Preliminary Results from an Interview Study," *American Economic Review,* vol. 81 (May 1991, Papers and Proceedings, 1990), pp. 89–96.

Carlton, Dennis W. "The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?" in R. Schmalensee and R. Wilig, eds., *Handbook of Industrial Organization,* Vol. 1. Amsterdam: North-Holland, 1989.

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

Cecchetti, Steve. "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," *Journal of Econometrics,* vol. 31 (August 1986), pp. 255–74.

Cho, Jang-Ok, and Thomas Cooley. "The Business Cycle with Nominal Contracts," Working Paper 260. Rochester: Rochester Center for Economic Research, 1990.

Christiano, Larry, and Martin Eichenbaum. "Liquidity Effects, Monetary Policy and the Business Cycle," Working Paper. Evanston, Ill.: Northwestern University, 1992.

Cody, Brian J., and Leonard O. Mills. "The Role of Commodity Prices in Formulating Monetary Policy," *The Review of Economics and Statistics,* vol. LXIII (May 1991), pp. 358–65.

Cover, James P. "Asymmetric Effects of Positive and Negative Money Supply Shocks," Working Paper. Tuscaloosa, Ala.: University of Alabama, 1992.

Engel, Charles. "Is Real Exchange Rate Variability Caused by Relative Price Changes? An Empirical Investigation," Working Paper. Seattle: University of Washington, 1991.

Fischer, Stanley. "Relative Shocks, Relative Price Variability, and Inflation," *Brookings Papers on Economic Activity,* 2:1981, pp. 381–431.

Kashyap, Anil. "Sticky Prices: New Evidence from Retail Catalogs," Working Paper. Chicago: University of Chicago, 1991.

King, Robert. "Money and Business Cycles," Working Paper. Rochester: University of Rochester, 1990.

Kretzmer, Peter. "The Cross-Industry Effects of Unanticipated Money in an Equilibrium Business Cycle Model," *Journal of Monetary Economics,* vol. 23 (September 1989), pp. 275–96.

Kydland, Finn. "The Role of Money in a Competitive Model of Fluctuations," Working Paper. Pittsburgh: Carnegie-Mellon University, 1987.

Lucas, Robert E., Jr. "The Effects of Monetary Shocks when Prices Are Set in Advance." Unpublished paper. University of Chicago, 1989.

———. "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics,* vol. 10 (November 1982), pp. 335–60.

———, and Michael Woodford. "Real Effects of Monetary Shocks in an Economy with Sequential Purchases." Unpublished paper. University of Chicago, 1992.

Mankiw, N. Gregory. "Small Menu Costs and Large Business Cycles: A Macroeconomic Model," *Quarterly Journal of Economics,* vol. 100 (May 1985), pp. 529–39.

Mussa, Michael. "Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates," *Carnegie-Rochester Conference Series on Public Policy,* vol. 25 (Autumn 1986), pp. 117–214.

Ohanian, Lee E., and Alan C. Stockman. "How Much Price Stickiness Is Necessary for Reasonable Liquidity Effects?" Working Paper. Rochester: Rochester Center for Economic Research, 1994.

Rees, Albert. "Alternative Retail Price Indexes for Selected Non-Durable Goods, 1947–59," in G. Stigler, ed., *The Price Statistics of the Federal Government.* New York: National Bureau of Economic Research, 1961.

Stigler, George, and J. Kindahl. *The Behavior of Industrial Prices.* New York: National Bureau of Economic Research, 1970.

Stockman, Alan. "Real Exchange Rates Under Alternative Nominal Exchange Rate Systems," *Journal of International Money and Finance,* vol. 2 (August 1983), pp. 147–66.

———, and Lee E. Ohanian. "Short-Run Independence of Monetary Policy Under Pegged Exchange Rates and Effects of Money on Exchange Rates and Interest Rates," Working Paper 4517. New York: National Bureau of Economic Research, 1993.

Svensson, Lars E. O. "Sticky Goods Prices, Flexible Asset Prices, Monopolistic Competition, and Monetary Policy," *Review of Economic Studies,* vol. 3 (July 1986), pp. 385–405.

Thoma, Mark A. "Subsample Instability and Asymmetries in Money-Income Causality," Working Paper. Eugene, Ore.: University of Oregon, 1992.

A Shift-Adjusted M2 Indicator for Monetary Policy

Robert Darin and Robert L. Hetzel

The Federal Reserve System influences the economy through its control of fiat money (currency and bank reserves) and the monetary aggregates. This influence is more predictable and is easier to observe when there is a stable relationship between the monetary aggregates and the public's dollar expenditure or output. The monetary aggregate M2 in particular has exhibited a long-term stable relationship with dollar output (Hetzel 1989, 1992; Ireland 1993). In the early 1990s, however, this relationship apparently disappeared. Over the three-year period 1990Q4 through 1993Q4, nominal GDP grew at an annualized rate of 5.2 percent, while M2 grew at an annualized rate of only 2.0 percent.

We first review regulatory and technological changes affecting financial intermediation that could be reducing the public's demand for M2. Specifically, we review the events that have encouraged bank depositors to place their funds in bond and stock mutual funds, which are not part of M2. We then investigate whether a version of M2 that adjusts for net flows into bond and stock mutual funds can reestablish the previous stable relationship with nominal output growth. This latter aggregate, "shift-adjusted" M2, consists of regular M2 plus cumulative net inflows from households into bond and stock mutual funds. The article concludes with a discussion of the likely future stability of money demand.

The views expressed are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1. MUTUAL FUNDS

From 1990Q4 through 1993Q4, total bank deposits declined by $25.1 billion, while bond and stock mutual funds rose by $506.7 billion. (All the figures here are for open-end mutual funds, that is, funds whose shares are continuously issued and redeemed.) A regular flow of news stories provides anecdotal evidence that individuals are taking funds out of small retail CDs, which are included in M2, and placing them in bond and stock mutual funds, which are not included in M2. For example, a story in the *American Banker* (3/22/93) states:

> Banks that sell mutual funds face a major test of their marketing mettle next month when a huge batch of certificates of deposit reaches maturity. As much as $110 billion in CDs will be up for grabs, and with interest rates at the lowest level in a generation, many customers won't be looking to roll over the investments. If recent history is any indication, depositors will pour much of that money into mutual funds. . . . [C]ustomers are becoming convinced that rates will remain low for a while and are ready to seek alternatives. . . . Now they are saying, "I've got to get some income." (P. 1)

Money Market Mutual Funds

The recent growth of bond and stock mutual funds is reminiscent of the growth of money market mutual funds in the late 1970s. Competition from the mutual fund industry for bank deposits began in earnest with the cyclical pickup in money market rates in 1977. Prior to that time, a large rise in market rates would cause Reg Q ceilings on the rates financial institutions could pay on time and savings deposits to become binding. Holders of small deposits had difficulty fleeing to money market instruments like commercial paper because of the large denominations of those instruments. By 1977, however, the availability of money market mutual funds, which pool funds from numerous individuals for investment in short-term financial assets, allowed depositors to avoid Reg Q and still hold assets that were available in small denominations and that could be bought and sold with low transactions costs. A good example was Merrill Lynch's Cash Management Account, a checkable money market account introduced in 1977.

Money market mutual funds not only collected deposits from investors, but also bought the commercial paper of corporations. Large corporations, often with better credit ratings than banks, found raising funds in this way cheaper than borrowing from banks. Intermediation that formerly went through banks now bypassed them completely. The ability of investors to circumvent Reg Q made inevitable its elimination, beginning with the introduction of all-savers certificates in June 1978 and ending with the elimination of the ceiling on savings deposits in April 1986. The introduction of money market deposit accounts (MMDAs) in 1982Q4 allowed banks to compete directly with money market funds. The decline in bank intermediation, however, led to a decline in the public's demand for M2 defined exclusive of money market mutual funds.

Through the early 1980s, most of the growth in mutual funds occurred in money market mutual funds. Like other M2 deposits, the shares of money market mutual funds are available in small denominations. Also, because these funds hold only short-term securities, their shares are redeemable at par. That is, they do not fluctuate in value with changes in interest rates. (The weighted-average maturity of money market mutual funds cannot be greater than 90 days.) As a consequence, including the shares of money market mutual funds in M2 was straightforward and restored the long-run stable relationship between M2 and the public's dollar expenditure and output.

Bond and Stock Mutual Funds

Table 1 shows net inflows into bond and stock mutual funds. These funds, as opposed to money market mutual funds, first began to grow significantly in 1984. They grew fairly strongly from the middle of 1985 through the middle of 1987, grew very little from the second half of 1987 through early 1990, and then began to grow rapidly toward the middle of 1991. The growth of bond and stock mutual funds in the mid-1980s is not associated with instability in the relationship between M2 and nominal output. Most of the growth in this period occurred in bond rather than stock funds and was heavily concentrated in mortgage-backed securities and, to a lesser extent, in junk bonds. Apparently, the investors financing this growth in bond funds were drawing funds from large CDs and money market instruments not included in M2.

Table 1 Annual Inflows
Fourth Quarter over Fourth Quarter

Year	Total Bank Deposits	Total Money Market Mutual Funds	Total Stock and Bond Mutual Funds
1984	228.38	39.44	20.63
1985	187.11	22.26	71.09
1986	184.70	53.11	135.54
1987	122.29	16.44	45.72
1988	177.04	19.71	−23.46
1989	97.60	91.02	3.07
1990	11.51	60.79	22.94
1991	−0.57	60.24	96.65
1992	−15.25	16.40	168.90
1993	−9.29	−18.47	241.12

Source: For data on bank deposits and money market mutual funds, the Board of Governors of the Federal Reserve System; for mutual fund data, the Investment Company Institute.

In the early 1990s, two forces combined to encourage the large-scale transfer of funds from bank deposits to bond and stock mutual funds. (See Duca [1992, 1993], Simpson and Scanlon [1993], and Reid and Small [1993].) First, the ongoing telecommunications and computer revolution continued to lower the cost to mutual funds of pooling cheaply the savings of investors. The resulting competition from mutual funds for bank deposits prompted a relaxation by regulators of the constraints imposed by Glass-Steagall, which prevents banks from underwriting securities. In the early 1990s, that relaxation allowed banks to market actively bond and stock mutual fund shares. The second force encouraging deposit outflows was the depressed state of the loan market, which prompted banks to pay low rates on their deposits, both absolutely and relatively to the returns available on stocks and bonds.

The first force, advances in communications and computer technology, reduced the cost of maintaining records on purchases and sales and on income and distributions. It also helped with the recordkeeping required to keep track of gains and losses for tax purposes. Vanguard Group introduced a series of U.S. government securities funds that charged a maximum annual operating fee of 15¢ per $100 (*Business Week*, 1/18/93), compared with the 23¢ per $100 of deposits charged for FDIC insurance alone. Also, mutual funds were offered as families of different kinds of funds, within which investors can easily switch by telephone. The increased ease in selecting mutual funds was exemplified in Charles Schwab's combined offering of almost 250 no-load mutual funds that do not charge brokers' fees. *The New York Times* (3/20/94) described this concentration of funds in one place as "a financial Wal-Mart that enables investors to trade funds as easily as stocks" (Sec. 3, p. 1).

In the last part of the 1980s, regulators, concerned about the ability of banks to compete for the public's savings, increasingly allowed banks to become involved in the marketing of mutual funds. The Glass-Steagall Act prevents banks from underwriting mutual fund shares. That is, banks cannot buy the underlying securities, repackage them in the form of mutual funds shares, and distribute those shares to the ultimate investors. By the early 1990s, however, banks or their affiliates had acquired the right to perform most of the other services needed to maintain a mutual fund. They could serve as the investment adviser. That is, they could select the particular stocks or securities specified by the fund's stated objectives. Banking organizations could also serve as transfer agent and custodian. That is, they kept records of ownership and of the collection and distribution of interest and dividend income. They also settled the accounts between buyers and sellers.

In the early 1980s, regulators allowed banks to establish a discount brokerage service not subject to the geographical limitations of the McFadden Act. In 1992, the Federal Reserve Board allowed bank holding companies to provide investment advice along with brokerage services. In 1993, the Office of the Comptroller of the Currency permitted Dean Witter and a subsidiary of

NationsBank to form a partnership to sell mutual funds and other securities in the branch offices of NationsBank. Also, in April 1993, the Federal Reserve Board allowed Mellon Bank Corporation, which had already teamed up with the mutual fund company Dreyfus, to buy The Boston Company, which provided administrative and advisory services to 84 different mutual funds.

In 1991, a fall in the rate of interest paid on bank deposits, combined with the ongoing technological and regulatory changes that were facilitating the creation of mutual funds, encouraged the large-scale transfer of funds from bank deposits to bond and stock mutual funds. Many retired investors used the income from bank CDs to support themselves. Especially with the sharp fall in short-term rates that began at the end of 1990, they moved out of CDs into bond and stock mutual funds, which promised a steadier cash flow (at the risk of capital fluctuation). *The Wall Street Journal* (2/12/93) wrote:

> [The yield on] Treasury bills plunged 37% last year. "That was the great T-bill crash of 1992," says Laurence Siegel. . . . Investors usually think of stocks as very risky and bonds as moderately risky. Meanwhile, T-bills, certificates of deposit, money market funds and other short-term debt instruments are seen as virtually risk-free. That's certainly the case, if all you care about is fluctuations in price. But if your concern is getting a steady stream of income, then holding T-bills and rolling them over as they mature is much more risky than holding stocks or bonds. (P. C1)

By fall 1992, the rate paid on six-month CDs had fallen to about to 3.25 percent, where it remained until early 1994. For all of 1992, the 30-year bond rate averaged 7.67 percent, while three-month CDs averaged 3.62 percent, an unusually wide difference of four percentage points. As of February 1994, savings deposits at commercial banks and savings banks paid on average 2.43 percent (Board of Governors of the Federal Reserve System statistical release H.6, "Monthly Survey of Selected Deposits," March 24, 1994). That level of short-term rates produced a transfer of funds out of bank deposits into higher-yielding bond and stock mutual funds.

The immediate cause of the disintermediation from financial institutions was the low rates paid on deposits, which in turn reflected weakness in loan demand. In addition, the need to rebuild capital forced many financial institutions to restrict their asset growth and, indirectly, their deposit growth. Finally, the well-publicized problems of financial institutions with debt defaults, especially in real estate, beginning in 1989 altered the perception of investors with small amounts of capital that bank deposits were the primary safe form of saving apart from savings bonds and Treasury bills. Mutual funds, in contrast, experienced no such bad publicity.

A survey by the Board of Governors (1993) documented the change in emphasis in the 1990s by banks from solely attracting deposits to retail marketing of mutual funds. The Board surveyed 56 large banks nationwide. All but four of these banks marketed mutual funds to their retail customers.

Three-quarters of the banks that marketed mutual funds had sales representatives at their branches. Forty percent of the banks had sales forces with over 50 people. By 1993, customers of Wells Fargo could buy and sell mutual funds through automated teller machines (*The Economist*, 9/4/93).

2. SHIFT-ADJUSTED M2

Shares in bond and stock mutual funds possess many of the characteristics of the deposits in M2. They are liquid and available in small denominations, and they can be bought and sold with low transactions costs. The existence of these common characteristics suggests adding shares in bond and stock mutual funds to M2 to create a more inclusive monetary aggregate that would be unaffected by transfers between these funds and M2. Bond and stock mutual funds, however, are not complete substitutes for the time deposits in M2. Fluctuation in their capital value presents a risk not present with bank time deposits. Also, they are not suitable for regular small transactions in that each sale of a mutual fund share creates a capital gain or loss that is taxable.

Fluctuation in capital value makes bond and stock mutual funds unsuitable for inclusion in a broad monetary aggregate. Consider, for example, the anomaly that could arise if bond and stock mutual funds were added to M2 to create a new, more inclusive monetary aggregate. Assume that a rise in money growth creates an expectation of a future rise in inflation. That expectation would produce a rise in bond yields, which would depress the value of bonds. The value of an inclusive definition of money would then fall and give a misleading message about the thrust of monetary policy. Money growth would fall when nominal output growth rose.

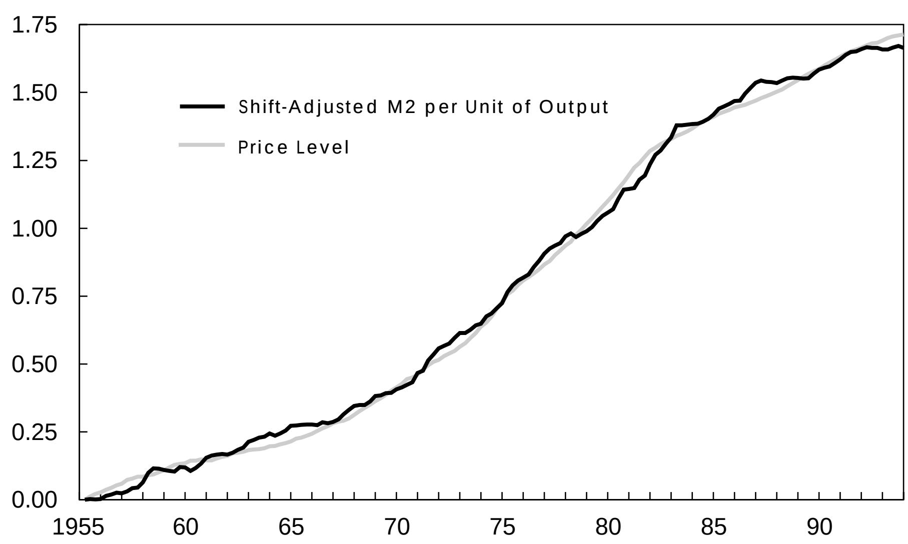
One way to offset the distortions in M2 produced by bond and stock mutual funds is to construct a shift-adjusted M2: M2 plus cumulative net inflows into bond and stock mutual fund shares not coming from institutional investors and not held in IRA/Keogh accounts. The exclusion of institutional holdings is consistent with the definition of M2, which excludes money market funds held by institutions. Similarly, the exclusion of IRA/Keogh accounts reflects the exclusion of these accounts in M2. The shift-adjusted measure also excludes reinvested dividends. The use of dollar inflows to adjust M2 avoids the problem of changes in the capital value of bonds and stocks.

This construct should not be thought of as a conventional monetary aggregate. The divergence between its growth and M2 growth, however, suggests a measure of the extent of shifts in the public’s demand for M2. By taking account of these shifts, it is possible that M2 could again be used as an indicator of the thrust of monetary policy. Shift-adjusted M2 is analogous to the shift-adjusted M1 the Fed used in 1981 (Bennett 1982). At that time, an adjustment to M1 was needed because the incorporation of interest-bearing NOW (negotiable order of withdrawal) accounts in the definition of M1 in 1980 and the introduction of these accounts nationwide in 1981 produced an inflow of funds into M1 from deposits not formerly included in M1. Shift-adjusted M1 subtracted an estimate of these inflows. Analogously, the suggested shift-adjusted M2 adds in an estimate of outflows from M2 into bond and stock mutual funds. Table 2 lists figures for conventional M2 and shift-adjusted M2. In 1991 the difference in their growth rates was only 1.4 percent, but that figure rose to 1.9 percent in 1992 and 2.4 percent in 1993.

3. SHIFT-ADJUSTED M2 AS AN INDICATOR

How well does shift-adjusted M2 predict the impact of monetary policy actions? Shift-adjusted M2 predicts better than unadjusted M2; however, it does not eliminate all of the unusual reduction in the public's demand for M2. Figure 1 uses shift-adjusted M2 per unit of output to predict the price level. (M2 is shift-adjusted starting in 1991.) Assuming that M2 velocity is stable over time, that is, the ratio of dollar output to M2 fluctuates around an unchanged value, M2 divided by real output should move with the price level over long periods of time. (See Humphrey [1989] for a history of the use by quantity

Figure 1 Money per Unit of Output as a Predictor of the Price Level



Notes: Data are quarterly observations of shift-adjusted M2 divided by real GDP and of the implicit price deflator. Both series are logarithms of the index numbers created by dividing each series by its 1955Q1 value. Shift-adjusted M2 is M2 plus cumulative net inflows starting in 1991 from households into bond and stock mutual funds (non-IRA/Keogh accounts).

Table 2 Growth of M2 and Shift-Adjusted M2

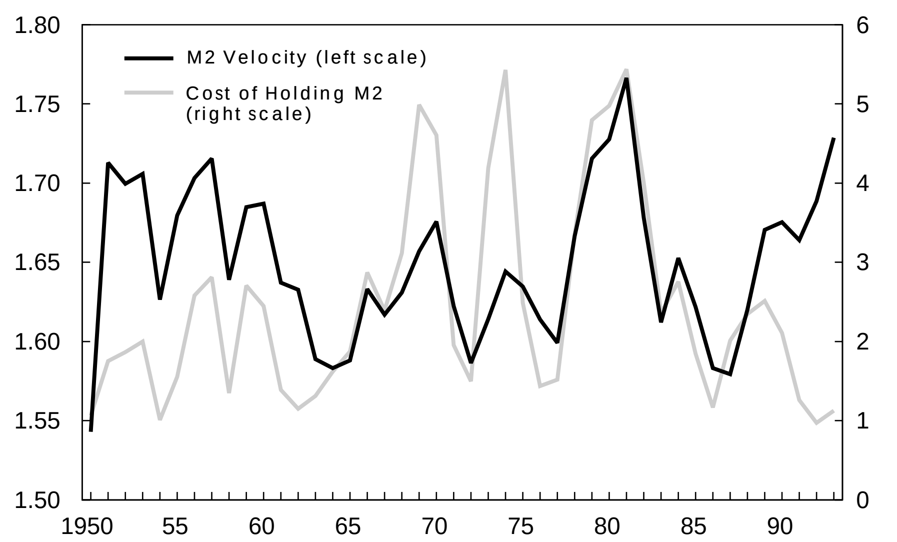
Year	Month	M2 (billions)	Shift-Adjusted M2 (billions)	Monthly Annualized Growth M2	Monthly Annualized Growth Shift-Adjusted M2	Quarterly Annualized Growth M2	Quarterly Annualized Growth Shift-Adjusted M2
		(1)	(2)	(3)	(4)	(5)	(6)
1991	1	3363.53	3365.40	3.85	4.54		
	2	3380.02	3384.65	6.04	7.08		
	3	3398.11	3405.88	6.62	7.79	3.89	4.47
	4	3409.00	3419.30	3.91	4.83		
	5	3418.86	3432.99	3.52	4.91		
	6	3426.64	3444.01	2.77	3.92	4.53	5.64
	7	3426.42	3447.19	−0.08	1.12		
	8	3427.40	3453.14	0.34	2.09		
	9	3427.51	3458.72	0.04	1.96	1.05	2.46
	10	3432.33	3469.06	1.70	3.65		
	11	3445.36	3486.11	4.65	6.06		
	12	3455.25	3501.98	3.50	5.60	2.02	3.84
1992	1	3464.09	3519.20	3.12	6.06		
	2	3483.61	3547.02	6.97	9.91		
	3	3486.28	3555.74	0.92	2.99	3.97	6.45
	4	3481.94	3554.52	−1.48	−0.41		
	5	3482.07	3560.23	0.04	1.94		
	6	3477.77	3562.72	−1.47	0.84	0.30	2.11
	7	3480.72	3572.35	1.02	3.29		
	8	3489.43	3587.62	3.05	5.25		
	9	3496.58	3599.00	2.49	3.87	0.96	3.09
	10	3507.46	3610.43	3.80	3.88		
	11	3510.53	3617.26	1.05	2.29		
	12	3509.04	3624.41	−0.51	2.40	2.32	3.51
1993	1	3502.78	3624.12	−2.12	−0.10		
	2	3494.21	3621.75	−2.90	−0.78		
	3	3494.83	3630.23	0.21	2.85	−1.33	0.89
	4	3497.99	3641.85	1.09	3.91		
	5	3521.86	3674.52	8.50	11.32		
	6	3528.70	3687.19	2.36	4.21	2.18	4.77
	7	3533.56	3700.11	1.66	4.29		
	8	3535.69	3712.65	0.73	4.14		
	9	3543.59	3727.09	2.71	4.77	2.46	5.05
	10	3545.09	3735.63	0.51	2.79		
	11	3556.19	3753.27	3.82	5.82		
	12	3563.11	3765.99	2.36	4.14	1.96	4.20

theorists of money per unit of output as a predictor of the price level.) As shown in Figure 1, M2 per unit of output and the price level do gravitate around each other. In 1992, however, M2 begins to underpredict the price level, even with the shift adjustment. By 1993Q4, the underprediction reaches 5 percent.

Figure 2 shows the normally positive relationship between M2 velocity and the financial market opportunity cost of holding M2. The latter variable is measured by the difference between the commercial paper rate and a weighted average of the explicit rates of return paid on the components of M2 (Hetzel 1989). An increase in the cost of holding M2 raises M2 velocity by lowering the demand for M2, and conversely. The shift-adjustment does not restore for the early 1990s the normal positive relationship between M2 velocity and the financial market opportunity cost of holding M2. Shift-adjusted M2 velocity should have fallen, but did not.

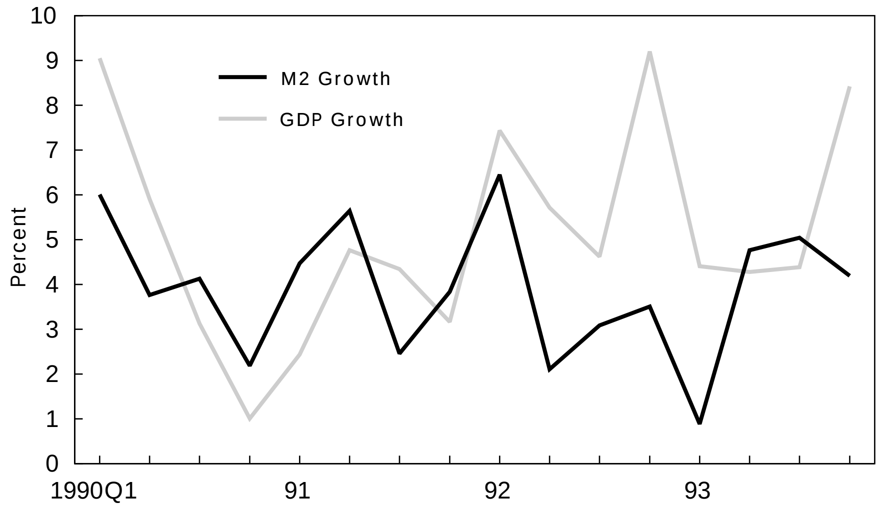
Figure 3 shows quarterly observations of quarterly rates of growth of shift-adjusted M2 and nominal output (GDP) over the recent period of instability in M2 demand. The two series exhibit some common fluctuations. The correlation between growth in nominal output and growth in M2 is .35. When M2 growth is lagged one quarter, the correlation is .41. The rise in nominal output growth

Figure 2 The Relationship Between Shift-Adjusted M2 Velocity and the Financial Market Opportunity Cost of Holding M2



Notes: The financial market opportunity cost of holding M2 is the difference between the rate on six-month commercial paper and a weighted average of the explicit rates of interest paid on the components of M2. Shift-adjusted M2 is M2 plus cumulative net inflows starting in 1991 from households into bond and stock mutual funds (non-IRA/Keogh accounts). Shift-adjusted M2 velocity is nominal GDP divided by shift-adjusted M2.

Figure 3 Shift-Adjusted M2 and Nominal Output Growth



Notes: Data are quarterly observations of quarterly annualized growth. Shift-adjusted M2 is M2 plus net household inflows into stock and bond mutual funds beginning in 1991.

during the economic recovery, however, is not matched by the usual rise in M2 growth. It appears that shift-adjusted M2 continues to contain some useful information about the effect of monetary policy actions on nominal output, but less than M2 does before 1991.

Shift-adjusted M2 appears to account for somewhat less than half of the unusual decrease in the demand for M2. Over the three-year period 1990Q4 through 1993Q4, the annualized rate of growth of nominal output was 5.2 percent, while the annualized rate of growth of M2 was only 2 percent. Over this same period, the cost of holding M2 (the financial market opportunity cost shown in Figure 2), fell about 1 percent. In the past, that decline would have produced somewhat faster growth in M2 than in nominal output (a fall in M2 velocity). Appendix B makes these figures more precise by comparing the prediction errors from a money demand regression estimated with conventionally defined M2 and with shift-adjusted M2. Using conventionally defined M2, the overprediction in the rate of growth of real M2 for the years 1991, 1992, and 1993 is 2.6, 3.2, and 4.2 percent, respectively. Using shift-adjusted M2, the overprediction in each year falls to about 2 percent.

It is possible that there is an explanation for the leftward shift in M2 demand that does not focus on the desire of the public to shift its savings from the deposits of financial institutions to capital market instruments. The explanations offered, however, have not proven satisfactory. One explanation offered was that the closing of thrifts by the Resolution Trust Corporation extinguished thrift deposits that had been included in M2. The Resolution Thrift Corporation, however, stopped closing insolvent thrifts after March 1992 because of lack of funds. Despite this fact, M2 continued to grow slowly relative to nominal output. The other explanation offered was that the public was using M2 balances to reduce its debt. Consumer installment credit, however, began to grow strongly in 1992Q3. Adding home equity loans to consumer installment credit results in a typical growth in consumer credit for a period of economic recovery. This growth implies that the public no longer considered its debt level excessive. M2 growth, however, did not subsequently revive.

It appears that the leftward shift in M2 demand derives from the public's increased desire to save with capital market instruments rather than bank deposits. This change in behavior is driven by the reduction in the transactions cost of buying and selling capital market instruments and by the availability of these instruments in small denominations made possible by the pooling of investors' savings in mutual funds. The failure of the shift-adjustment to account fully for the leftward shift in M2 demand evidently arises from a failure to account fully for the outflows of M2 deposits to other sources. Indirect confirmation of this conjecture comes from the Board of Governors' Surveys of Consumer Finances (Kennickell and Starr-McCluer 1993, p. 3). Between the 1989 and 1992 surveys, bank deposits as a fraction of households' financial assets fell from 31.9 to 26.1 percent, or 5.8 percentage points. Mutual fund holdings, however, rose only 3.2 percentage points, from 14.6 to 17.8 percent. The direct purchase of stocks and bonds apparently accounted for the remainder of the decline. Duca (1993) also points out that in 1992 sales of U.S. savings bonds held for under five years surged when money market rates fell below the floor of 4.16 percent paid on these instruments. It is likely that much of the inflows into these savings bonds came from M2 deposits.

Finally, shift-adjusted M2 does not account for the increase in the use of tax-sheltered forms of savings in the early 1990s. Governor Lindsey (1994) reports that in 1993 tax-sheltered forms of income, particularly pension fund and life insurance reserves, accounted for 70 percent of the net acquisition of financial assets. This tax shifting was probably undertaken in response to the increase in marginal tax rates in 1991 and 1993. Because the shift-adjustment made here to M2 does not include deposit inflows to IRA/Keogh accounts or deposit inflows from institutional investors, it does not capture the decline in M2 that occurs when an individual withdraws funds from a time deposit included in M2 and places them in a tax-sheltered investment. The increased importance of tax-sheltered savings can be seen by comparing mutual fund inflows in the mid-1980s and in the early 1990s. Over the three years 1985 to 1987, 43 percent of the inflows into bond and stock mutual funds went either into accounts held by institutional investors like life insurance companies or into IRA/Keogh accounts. This figure rose to 61 percent over the three-year period 1991 to 1993 (see Table 3 in Appendix A).

4. WILL M2 DEMAND BECOME STABLE AGAIN?

Will the behavior of the public's demand for M2 become predictable again? In particular, will the Fed again be able to set targets for M2 growth that can be reliably related to the desired rate of growth of the public's dollar expenditure? The admittedly equivocal answer is, "It could." The distressed condition of banks in the early 1990s and the associated low rates of interest on bank deposits are not likely to recur. After the completion of the current rechanneling of saving from the indirect intermediation provided by banks in favor of the direct intermediation of Wall Street, it is possible that mutual funds will grow steadily enough to avoid destabilizing M2 demand.

A somewhat different question is whether the Fed will be able to use M2 again as an indicator of the impact of its policy actions on the behavior of the public's dollar expenditure. Even if the public continues to shift funds between bank deposits and bond and stock mutual funds, a shift-adjusted M2, by offsetting the resulting fluctuations in M2 demand, could become a useful measure of the impact of monetary policy. Assuming that a primary reason that shift-adjusted M2 does not fully account for the leftward shift in M2 demand is the move toward tax-sheltered savings instruments, then, in the absence of major future changes in the tax code, shift-adjusted M2 should become a useful monetary indicator. It could become particularly useful in the event of a financial disturbance causing a large, sudden outflow from mutual funds. In that event, the resulting large changes in conventional M2 would be primarily noise.

The future usefulness of the monetary aggregates as targets or indicators also depends upon the tax and regulatory environment that banks will face. The computer technology that made mutual funds possible also makes it possible for banks to take deposits off their books in ways that can make it difficult to measure money accurately. Banks will have an incentive to pursue this technology as long as they face tax and regulatory obstacles to collecting deposits that are not incurred by other organizations competing for the public's savings. Two major handicaps that banks suffer in competing for the public's savings are the prohibition of payment of interest on demand deposits and the imposition of a tax in the form of noninterest-bearing reserve requirements. (Goodfriend and Hargraves [1983] discuss the role that reserve requirements have played as a tax.) These two institutional features create incentives for banks to lower the amount of demand deposits and other checkable deposits on their balance sheets in ways that distort measurement of the monetary aggregates.

In particular, sweep accounts allow banks to avoid the prohibition of the payment of explicit interest on demand deposits and the tax imposed by noninterest-bearing reserve requirements. For example, on March 21, 1994, a large bank advertised in *The Wall Street Journal* for an account that "automatically sweeps your excess cash into preselected investments daily." In one version, the bank sweeps balances from a NOW account above a specified amount into an MMDA account. Whenever the NOW account falls below a specified minimum level, the bank transfers funds from the MMDA back into the NOW account. (All funds are transferred back in with the sixth transfer to avoid exceeding the legal limit of six automatic transfers per month from an MMDA account.) Reid and Small (1993) state that "about a quarter of the banks selling mutual funds provide retail sweep accounts whereby funds in a depositor's account in excess of a predetermined amount are automatically invested in a money market mutual fund or some other uninsured investment vehicle" (p. 12).

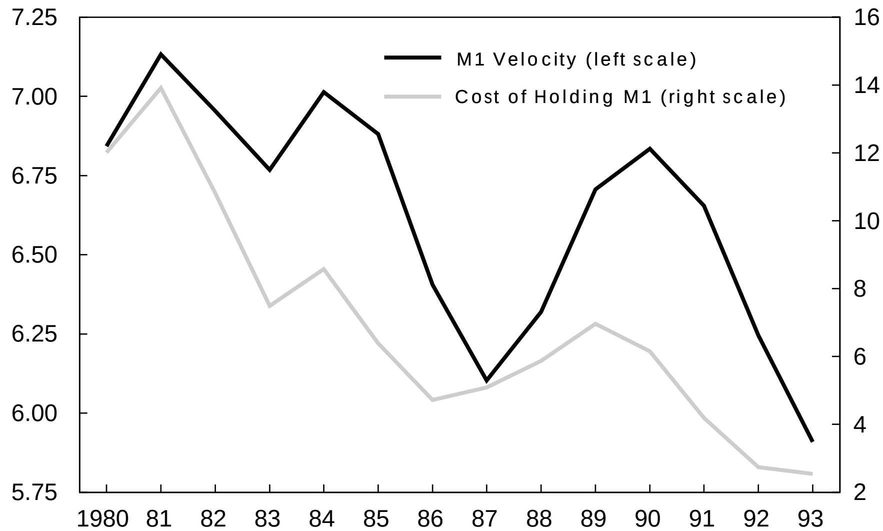
A group that helps banks design sweep accounts reports the benefits of offering sweep accounts (Treasury Strategies 1994, pp. 24 and 26):

> Reduced Reserves—Money market funds and trust sweeps move customer funds off balance sheet. Reserves of those deposits are eliminated and the amount of assets available for investment is increased.
>
> Reduced FDIC premiums—Again, by moving funds out of insured depository accounts, FDIC premiums are eliminated. This .23% savings is also a benefit in which both the customer and bank can share.
>
> Banks interested in encouraging fee payment for services and/or removing demand balances from their balance sheets set target balances and minimum sweep amounts at zero.

If a bank sweeps funds above a target balance into an MMDA, it does not affect the behavior of M2, but it does reduce measured M1. M1 has not been useful as a measure of money in the 1980s because of its high degree of sensitivity to changes in market interest rates (Hetzel and Mehra 1989). In the absence of a good measure of the interest sensitivity of the public's demand for M1 balances, when interest rates have changed, it has been hard to estimate how M1 was changing relative to the public's demand for it. The additional observations made possible by the passage of time could make this econometric problem manageable in the absence of sweep accounts. As Figure 4 shows, the variations in M1 velocity are related to the cost of holding it. In time, it might be possible to estimate again a reliable M1 demand function. In the future, with a high degree of substitutability between shares in bond and stock mutual funds and the time deposits in M2, a narrow aggregate like M1 could well be more stably related to the public's expenditure than M2.

Computer technology is making it easier to avoid the tax imposed by noninterest-bearing reserve requirements. In principle, the Fed could solve the problem by paying interest on required reserves. Such proposals were advanced in discussions of the 1980 Monetary Control Act, but they were politically unacceptable. Alternatively, the Board of Governors could reduce reserve requirements. It is, however, limited by the legal minimum reserve requirement of 8 percent.

Figure 4 Annual M1 Velocity and the Opportunity Cost of Holding M1



Notes: M1 velocity is nominal GDP divided by M1. The opportunity cost of holding M1 is the six-month commercial paper rate minus a weighted average of the rates paid on the components of M1 (zero for currency and demand deposits).

Depending upon where "swept" funds are placed, sweep accounts can reduce measured M1 or M2. To an extent, the kind of shift-adjustment proposed here could reduce the resulting mismeasurement of the monetary aggregates. However, the data necessary to make such adjustments are never likely to be complete and are available only with a lag. Bank intermediation bears the burden of a large variety of regulations not imposed on other forms of financial intermediation. It seems likely that in the future the ability to define monetary aggregates that are useful for monetary policy will depend upon whether banks have a continuing incentive to adapt to special taxes and regulations.

5. SUMMARY

The growth of mutual funds that began in earnest in 1978 has increasingly directed financial intermediation away from banks and directly into the money and capital markets. Initially, growth occurred in money market mutual funds. Because shares in these funds are redeemable at par, they could be included in the definition of M2. Redefined to include money market mutual funds, M2 retained its long-run stable relationship with nominal output. Beginning in the mid-1980s, growth in mutual funds has been concentrated in bond and stock mutual funds. Because shares in these funds exhibit fluctuation in capital value, they cannot be included in an expanded definition of M2 in a satisfactory way.

In the early 1990s, the combination of (1) low rates of return on bank deposits relative to capital market instruments and (2) the decreased cost of operating bond and stock mutual funds diminished the public's demand for saving in the form of bank deposits. M2 velocity rose as depositors redirected savings from time deposits to bond and stock mutual funds. In principle, a shift-adjusted M2, defined as M2 plus cumulative dollar inflows into bond and stock mutual funds, could maintain the same relationship to the public's dollar expenditure as conventional M2. The shift-adjustment, however, accounts for only about half of the unusual rise in M2 velocity that began in 1991. The remainder to the rise in M2 velocity is probably caused by the use of bank time deposits to purchase stocks and bonds directly and to make tax-sheltered investments.

APPENDIX A

CONSTRUCTION OF SHIFT-ADJUSTED M2

Tables 3, 4, and 5 detail the construction of shift-adjusted M2 from the Investment Company Institute (ICI) data contained in the monthly release "Trends in Mutual Fund Activity." As shown in Table 3, the first step is to subtract redemptions (including exchanges out of bond and stock mutual funds) from sales (including exchanges into bond and stock mutual funds but excluding reinvested dividends) of mutual funds to derive total net inflows. Subtracting net inflows due to institutions and net inflows into IRA/Keogh accounts then yields the net inflows due to households that are assumed to be coming out of deposits in M2. These net inflows are cumulated and added to conventionally defined M2 to derive shift-adjusted M2. Unfortunately, there are no direct measures of net inflows due to institutions or of net inflows into IRA/Keogh accounts. Tables 4 and 5 explain the derivation of these two series.

ICI publishes figures on the dollar values of institutional and IRA/Keogh accounts. Dollar inflows into these accounts can be calculated by subtracting capital gains from the changes in their dollar value. The capital gains (losses) for these accounts are assumed proportional to the capital gains (losses) for all types of mutual fund accounts. Figures on the level of mutual fund assets held in IRA/Keogh accounts are available monthly, while the figures for institutional accounts are only available for December of each year. Therefore, it is necessary to interpolate monthly asset levels for institutional accounts from year-end figures.

Table 3 Calculation of Household Net Inflows to Mutual Funds
Billions of dollars

Year	Qtr.	Sales of Mutual Funds (excludes reinvested dividends)	Redemptions	Total Net Inflow (1)−(2)	Net Inflow Due to Institutions	Net Inflow into IRA/ Keogh Accounts	Net Inflow Due to Households (3)−(4)−(5)	Annual Totals
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
1984	1	15.49	9.61	5.88	1.39	1.11	3.38	
	2	13.84	9.12	4.72	1.74	1.00	1.97	
	3	16.05	8.91	7.14	2.19	1.22	3.72	
	4	15.69	11.54	4.15	1.20	0.69	2.26	11.33
1985	1	24.35	12.48	11.87	2.01	2.03	7.84	
	2	32.36	13.41	18.95	3.64	3.29	12.02	
	3	35.22	15.86	19.36	4.58	1.69	13.09	
	4	44.11	16.17	27.94	8.07	2.74	17.13	50.08
1986	1	58.92	22.51	36.41	11.06	5.53	19.82	
	2	66.67	31.67	35.00	11.85	5.83	17.32	
	3	68.40	37.13	31.26	11.00	2.89	17.38	
	4	74.54	39.18	35.36	13.03	2.67	19.66	74.18
1987	1	94.17	46.57	47.60	10.66	7.10	29.84	
	2	76.20	68.37	7.83	0.35	7.31	0.16	
	3	61.88	63.99	−2.11	−2.99	0.63	0.25	
	4	48.03	69.61	−21.58	−8.16	−1.29	−12.13	18.12

1988	1	42.03	44.65	−2.62	−1.12	0.87	−2.37	
	2	41.28	44.83	−3.55	−1.33	0.43	−2.64	
	3	30.87	40.02	−9.14	−2.78	0.82	−7.19	
	4	33.23	41.09	−7.86	−2.41	0.49	−5.94	−18.13
1989	1	37.06	42.27	−5.21	−1.36	−0.52	−3.34	
	2	41.99	39.52	2.47	0.79	1.24	0.44	
	3	44.48	39.98	4.50	1.58	1.81	1.11	
	4	49.10	41.96	7.14	2.22	1.96	2.96	1.17
1990	1	52.58	45.73	6.85	4.63	4.36	−2.14	
	2	51.98	40.77	11.21	5.88	3.15	2.18	
	3	45.77	46.96	−1.19	1.64	0.44	−3.27	
	4	44.80	40.34	4.46	2.87	−0.46	2.04	−1.17
1991	1	61.82	45.87	15.95	5.87	1.95	8.13	
	2	71.68	48.53	23.16	9.41	4.83	8.92	
	3	77.92	46.27	31.65	10.42	4.72	16.51	
	4	94.14	58.12	36.01	12.75	6.84	16.42	49.98
1992	1	115.06	70.03	45.02	19.17	5.38	20.47	
	2	106.57	62.33	44.25	18.00	10.30	15.95	
	3	113.12	69.04	44.07	17.54	9.49	17.04	
	4	119.03	79.17	39.85	15.08	10.32	14.46	67.92
1993	1	148.38	87.61	60.77	27.15	12.90	20.73	
	2	150.13	88.99	61.14	28.35	12.58	20.21	
	3	161.92	98.02	63.90	30.02	7.26	26.62	
	4	171.71	114.01	57.70	29.41	8.42	19.87	87.43

Notes: All data are from the Investment Company Institute. Sales include exchanges into stock and bond mutual funds; redemptions include exchanges out of stock and bond mutual funds. Quarterly data are sums of monthly figures. Annual totals are sums of quarterly data from the previous column.

Table 4 Aggressive-Growth Stock Funds
Net Inflows into IRA/Keogh Accounts, 1991
Millions of dollars

		All accounts, including IRA/Keogh				IRA/Keogh accounts			
Year	**Month**	**Total**	**Monthly Change**	**Net Inflow**	**Capital Gain (Loss) (2)−(3)**	**Total**	**Monthly Change**	**Capital Gain (Loss) [(5)÷(1)]×(4)**	**Net Inflow (6)−(7)**
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1991	December	63287.3		2180.7		18528.3			
1992	January	65474.8	2187.5	1555.3	632.2	19425.1	896.8	187.6	709.2
	February	68050.4	2575.6	1872.0	703.6	19426.0	0.9	200.9	−200.0
	March	66136.0	−1914.4	1050.4	−2964.8	19691.8	265.8	−882.8	1148.6
	April	65941.1	−194.9	1075.0	−1269.9	19547.4	−144.4	−376.4	232.0
	May	67847.9	1906.8	951.1	955.7	19844.0	296.6	279.5	17.1
	June	66295.1	−1552.8	430.2	−1983.0	19292.7	−551.3	−577.1	25.8
	July	69545.6	3250.5	812.1	2438.4	20205.9	913.2	708.5	204.7
	August	67472.7	−2072.9	−214.7	−1858.2	19507.6	−698.3	−537.2	−161.1
	September	68160.4	687.7	414.7	273.0	19608.1	100.5	78.5	22.0
	October	72134.5	3974.1	543.3	3430.8	20769.8	1161.7	987.8	173.9
	November	78842.4	6707.9	2753.9	3954.0	22596.5	1826.7	1133.2	693.5
	December	83365.3	4522.9	2276.0	2246.9	23893.3	1296.8	644.0	652.8

Table 4 illustrates the estimation of net inflows into IRA/Keogh accounts for aggressive-growth stock mutual funds. We repeat these calculations for the other 18 categories of bond and stock mutual funds in order to arrive at aggregate data. Column (1) of Table 4 lists monthly figures for the dollar value of all aggressive-growth funds, and column (2) shows changes in column (1). Column (3) shows monthly net inflows into aggressive-growth funds. Capital gains (4) are the difference between the change in dollar value in column (2) and net inflows in column (3). Column (5) lists the monthly dollar values of aggressive-growth funds held in IRA/Keogh accounts, and column (6) shows changes in (5). Column (7) shows estimated monthly capital gains of IRA/Keogh accounts. It is derived by multiplying capital gains for all aggressive-growth funds from column (4) by the percentage of all assets held in IRA/Keogh accounts, which is column (5) divided by column (1). The resulting capital gain for IRA/Keogh accounts (7) is then subtracted from the change in the dollar value of IRA/Keogh accounts in column (6) to derive a monthly figure for the net inflow into aggressive-growth funds held in IRA/Keogh accounts, which is shown in column (8). Summing these figures across all types of mutual funds yields the figure in column (5), Table 3.

Table 5 illustrates the estimation of net inflows into institutional accounts. Part 1 shows year-end figures for the dollar value of total and institutional aggressive-growth funds. Column (c) shows the percentage held in institutional accounts. In Part 2, column (1) shows the dollar value of total aggressive-growth funds. The dollar amount held in institutional accounts, column (3), is estimated as the product of column (1) and the fraction held in institutional accounts, column (2), interpolated from the figures shown in column (c), Part 1. Column (4) shows monthly changes in these figures. Column (5), which is copied from column (4) of Table 4, shows capital gains for all bond funds. Capital gains for institutional accounts, column (6), is estimated by multiplying capital gains for all accounts, column (5), by the percentage of assets in institutional accounts, column (2). Net inflows into institutional accounts, column (7),

Table 5 Aggressive-Growth Stock Funds, Institutional Accounts—Part 1
Millions of dollars

Year	Month	All Accounts	Institutional Accounts	Percentage in Institutional Accounts
		(a)	(b)	(c)
1991	December	63287.3	21035.0	33.2%
1992	December	83365.3	24283.3	29.1%

Table 5 Aggressive-Growth Stock Funds, Institutional Accounts—Part 2
Millions of dollars

Year	Month	All Accounts	Interpolated Fraction Institutional Accounts	Institutional Accounts (1)×(2)	Change in Institutional Accounts	Capital Gain (Loss) All Accounts	Capital Gain (Loss) Institutional Accounts (2)×(5)	Net Inflows Institutional Accounts (4)−(6)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
1991	December	63287.3	33.2%	21035.0				
1992	January	65474.8	32.9%	21537.9	502.9	632.2	208.0	294.9
	February	68050.4	32.6%	22152.1	614.3	703.6	229.0	385.2
	March	66136.0	32.2%	21302.5	−849.6	−2964.8	−955.0	105.3
	April	65941.1	31.9%	21014.0	−288.5	−1269.9	−404.7	116.1
	May	67847.9	31.5%	21389.3	375.4	955.7	301.3	74.1
	June	66295.1	31.2%	20672.8	−716.5	−1983.0	−618.4	−98.1
	July	69545.6	30.8%	21448.3	775.5	2438.4	752.0	23.5
	August	67472.7	30.5%	20578.0	−870.3	−1858.2	−566.7	−303.6
	September	68160.4	30.2%	20554.4	−23.6	273.0	82.3	−106.0
	October	72134.5	29.8%	21505.8	951.5	3430.8	1022.8	−71.4
	November	78842.4	29.5%	23235.8	1729.9	3954.0	1165.3	564.6
	December	83365.3	29.1%	24283.3	1047.5	2246.9	654.5	393.0

is the difference between the change in the dollar amount in column (4) and capital gains for institutional accounts, column (6). Summing across all types of mutual funds yields the aggregate net inflow figure used in column (4), Table 3.

Because mutual fund data are end-of-month figures while M2 data are daily-average figures, the end-of-month figures are averaged to derive the monthly mutual fund data series.

APPENDIX B

AN ESTIMATED M2 DEMAND REGRESSION

The estimated M2 demand regression below uses shift-adjusted M2 for the years 1991, 1992, and 1993. It is a regression of percentage changes in real M2 on a dummy for the Korean War, percentage changes in real GDP, a contemporaneous and lagged value of changes in the financial market opportunity cost of holding M2, and second differences of percentage changes in nominal output. The last term is an estimate of the nominal rate of return on physical assets, which is used by Friedman and Schwartz in their money demand regressions. See Friedman and Schwartz (1982, Sec. 6.6.3) and Hetzel (1992). Estimation of the following regression using shift-adjusted M2 rather than conventionally defined M2 results in a reduction of the overprediction of real M2 of 23 percent in 1991, 55 percent in 1992, and 47 percent in 1993.

M2 Demand Regression, 1950 to 1993

$$\Delta \ln\ r\mathrm{M2}_t = \underset{(2.9)}{-3.0} * Korea + \underset{(11.9)}{.93}\ \Delta \ln\ rGDP_t - \underset{(4.8)}{1.1}\ \Delta(R_t - R\mathrm{M2}_t)$$

$$- \underset{(4.5)}{1.2}\ \Delta(R_{t-1} - R\mathrm{M2}_{t-1})\ - \underset{(6.3)}{.51}\Delta^2\ \ln GDP_t + \hat{e}_t$$

CRSQ = .64 SEE = 1.6 DW = 1.4 DF = 39

Notes: Observations are annual averages. M2 is shift-adjusted for the years 1991, 1992, and 1993. *r*M2 is per-capita M2 deflated by the implicit GDP deflator. *r*GDP is real per-capita gross domestic product. R is the four- to six-month commercial paper rate expressed as a decimal. *R*M2 is a weighted average of the own rates of return paid on components of M2. Korea is a shift dummy with a value of one in 1951, 1952, and 1953 and zero otherwise. Before 1959, M2 is M4 in Table 1 of Friedman and Schwartz (1970). From 1991 on, M2 includes inflows from households into non-IRA/Keogh bond and stock mutual funds. ln is the natural logarithm and Δ the first-difference operator. CRSQ is the corrected R-squared; SEE the standard error of estimate; DW the Durbin-Watson statistic; and DF degrees of freedom. Absolute value of t-statistics are in parentheses. Estimation is by ordinary least squares.

REFERENCES

American Banker. "Bank Mutual Funds Target Big Batch of Maturing CDs," March 22, 1993, p. 1.

Bennett, Barbara A. " 'Shift-Adjustments' to the Monetary Aggregates," Federal Reserve Bank of San Francisco *Economic Review,* Spring 1982, pp. 6–18.

Board of Governors of the Federal Reserve System. "Senior Financial Officer Survey on Retail Mutual Funds," May 12, 1993.

Business Week. "The Power of Mutual Funds," January 18, 1993, p. 66.

Duca, John V. "Should Bond Funds Be Included in M2?" Research Paper 9321. Dallas: Federal Reserve Bank of Dallas, June 1993.

__________. "The Case of the 'Missing M2,' " Federal Reserve Bank of Dallas *Economic Review,* Second Quarter 1992, pp. 1–24.

The Economist. "Hole in the Wall," September 4, 1993, p. 72.

Friedman, Milton, and Anna J. Schwartz. *Monetary Trends in the United States and the United Kingdom*. Chicago: University of Chicago Press, 1982.

__________. *Monetary Statistics of the United States*. New York: National Bureau of Economic Research, 1970.

Goodfriend, Marvin, and Monica Hargraves. "A Historical Assessment of the Rationales and Functions of Reserve Requirements," Federal Reserve Bank of Richmond *Economic Review,* vol. 69 (March/April 1983), pp. 3–21.

Hetzel, Robert L. "How Useful Is M2 Today?" Federal Reserve Bank of Richmond *Economic Review,* vol. 78 (September/October 1992), pp. 12–26.

__________. "M2 and Monetary Policy," Federal Reserve Bank of Richmond *Economic Review,* vol. 75 (September/October 1989), pp. 14–29.

__________, and Yash Mehra. "The Behavior of Money Demand in the 1980s," *Journal of Money, Credit, and Banking,* vol. 21 (November 1989), pp. 455–63.

Humphrey, Thomas M. "Precursors of the P-Star Model," Federal Reserve Bank of Richmond *Economic Review,* vol. 75 (July/August 1989), pp. 3–9.

Ireland, Peter N. "Price Stability Under Long-Run Targeting," Federal Reserve Bank of Richmond *Economic Quarterly,* vol. 79 (Winter 1993), pp. 25–45.

Kennickell, Arthur, and Martha Starr-McCluer. "Mutual Fund Ownership in the 1989 and 1992 SCF." Memorandum, Board of Governors of the Federal Reserve System, June 25, 1993.

Lindsey, Lawrence. "The Best of Times—the Worst of Times." Remarks made to the Baltimore Chapter of the Commercial Finance Association, March 7, 1994.

The New York Times. "The Next Giant in Mutual Funds?" March 20, 1994, Sec. 3, p. 1.

Reid, Brian, and David Small. "Bank Involvement in the Mutual Fund Industry." Mimeograph, Board of Governors of the Federal Reserve System, June 17, 1993.

Simpson, Thomas D., and Martha Scanlon. "Some Implications of the Growing Importance of Mutual Funds." Mimeograph, Board of Governors of the Federal Reserve System, June 25, 1993.

Treasury Strategies. "Treasury Strategies' 1994 Commercial Banking Sweep Account Survey Results." Chicago: Treasury Strategies, February 1994.

The Wall Street Journal. "T-Bill Trauma and the Meaning of Risk," February 12, 1993, p. C1.

The Role of Interest Rate Swaps in Corporate Finance

Anatoli Kuprianov

An interest rate swap is a contractual agreement between two parties to exchange a series of interest rate payments without exchanging the underlying debt. The interest rate swap represents one example of a general category of financial instruments known as derivative instruments. In the most general terms, a derivative instrument is an agreement whose value derives from some underlying market return, market price, or price index.

The rapid growth of the market for swaps and other derivatives in recent years has spurred considerable controversy over the economic rationale for these instruments. Many observers have expressed alarm over the growth and size of the market, arguing that interest rate swaps and other derivative instruments threaten the stability of financial markets. Recently, such fears have led both legislators and bank regulators to consider measures to curb the growth of the market. Several legislators have begun to promote initiatives to create an entirely new regulatory agency to supervise derivatives trading activity. Underlying these initiatives is the premise that derivative instruments increase aggregate risk in the economy, either by encouraging speculation or by burdening firms with risks that management does not understand fully and is incapable of controlling.[1] To be certain, much of this criticism is aimed at many of the more exotic derivative instruments that have begun to appear recently. Nevertheless, it is difficult, if not impossible, to appreciate the economic role of these more exotic instruments without an understanding of the role of the interest rate swap, the most basic of the new generation of financial derivatives.

The views expressed herein are those of the author and do not necessarily represent the views of either the Federal Reserve Bank of Richmond or the Board of Governors of the Federal Reserve System. The motivation for this article grew out of discussions with Douglas Diamond. Michael Dotsey, Jeff Lacker, Roy Webb, and John Weinberg provided thoughtful criticism and helpful comments.

[1] For a review of these stated concerns, recent policy initiatives, and pending legislation, see Cummins (1994a, 1994b), Karr (1994), and Rehm (1994).

Although the factors accounting for the remarkable growth of the swaps market are yet to be fully understood, financial economists have proposed a number of different hypotheses to explain how and why firms use interest rate swaps. The early explanation, popular among market participants, was that interest rate swaps lowered financing costs by making it possible for firms to arbitrage the mispricing of credit risk. If this were the only rationale for interest rate swaps, however, it would mean that these instruments exist only to facilitate a way around market inefficiencies and should become redundant once arbitrage leads market participants to begin pricing credit risk correctly. Thus, trading in interest rate swaps should die out over time as arbitrage opportunities disappear—a prediction that is at odds with actual experience.

Other observers note that the advent of the interest rate swap coincided with a period of extraordinary volatility in U.S. market interest rates, leading them to attribute the rapid growth of interest rate derivatives to the desire on the part of firms to hedge cash flows against the effects of interest rate volatility. The timing of the appearance of interest rate swaps, coming as it did during a period of volatile rates, seems to lend support to such arguments. Risk avoidance alone cannot explain the growth of the swaps market, however, because firms can always protect themselves against rising interest rates simply by taking out fixed-rate, long-term loans or by bypassing credit markets altogether and issuing equity to fund investments.

Recent research emphasizes that interest rate swaps offer firms new financing choices that were just not available before the advent of these instruments, and thus represent a true financial innovation. This research suggests that the financing choices made available by interest rate swaps may help to reduce default risk and may sometimes make it possible for firms to undertake productive investments that would not be feasible otherwise. The discussion that follows explains the basic mechanics of interest rate swaps and examines these rationales in more detail.

1. FUNDAMENTALS OF INTEREST RATE SWAPS

The most common type of interest rate swap is the fixed/floating swap in which a fixed-rate payer promises to make periodic payments based on a fixed interest rate to a floating-rate payer, who in turn agrees to make variable payments indexed to some short-term interest rate. Conventionally, the parties to the agreement are termed counterparties. The size of the payments exchanged by the counterparties is based on some stipulated notional principal amount, which itself is not paid or received.

Interest rate swaps are traded over the counter. The over-the-counter (OTC) market is comprised of a group of dealers, consisting of major international commercial and investment banks, who communicate offers to buy and sell swaps over telecommunications networks. Swap dealers intermediate cash flows between different customers, acting as middlemen for each transaction. These dealers act as market makers who quote bid and asked prices at which they stand ready to either buy or sell an interest rate swap before a customer for the other half of the transaction can be found. (By convention, the fixed-rate payer in an interest rate swap is termed the buyer, while the floating-rate payer is termed the seller.) The quoted spread allows the dealer to receive a higher payment from one counterparty than is paid to the other.

Because swap dealers act as intermediaries, a swap customer need be concerned only with the financial condition of the dealer and not with the creditworthiness of the other ultimate counterparty to the agreement. Counterparty credit risk refers to the risk that a counterparty to an interest rate swap will default when the agreement has value to the other party.[2] Managing the credit risk associated with swap transactions requires credit-evaluation skills similar to those commonly associated with bank lending. As a result, commercial banks, which have traditionally specialized in credit-risk evaluation and have the capital reserves necessary to support credit-risk management, have come to dominate the market for interest rate swaps (Smith, Smithson, and Wakeman 1986).

The discussion that follows largely abstracts from counterparty credit risk and the role of swap dealers. In addition, the description of interest rate swaps is stylized and omits many market conventions and other details so as to focus on the fundamental economic features of swap transactions. For a more detailed description of interest rate swaps and other interest rate derivatives, see Kuprianov (1993b). Burghardt et al. (1991) and Marshall and Kapner (1993) provide more comprehensive treatments.

Mechanics of a Fixed/Floating Swap

The quoted price of an interest rate swap consists of two different interest rates. In the case of a fixed/floating swap, the quoted interest rates involve a fixed and a floating rate. The floating interest rate typically is indexed to some market-determined rate such as the Treasury bill rate or, more commonly, the three- or six-month London Interbank Offered Rate, or LIBOR.[3] Such a swap is also known as a generic, or plain-vanilla, swap.

The basic mechanics of a fixed/floating swap are relatively straightforward. Consider an interest rate swap in which the parties to the agreement agree to

[2] An increase in market interest rates, for example, increases the value of a swap agreement to the fixed-rate payer, who will subsequently receive higher interest rate payments from the floating-rate payer.

[3] The London Interbank Offered Rate is the rate at which major international banks with offices in London stand ready to accept deposits from one another. See Goodfriend (1993) or Burghardt et al. (1991) for a detailed description of the Eurodollar market.

exchange payments at the end of each of T periods, indexed by the variable $t = 1, 2, \ldots, T$. Let $\overline{r^s}$ denote the fixed rate and $r^s(t)$ denote the floating interest rate on a fixed/floating swap. Payments between the fixed- and floating-rate payers commonly are scheduled for the same dates, in which case only net amounts owed are exchanged. The net cost of the swap to the fixed-rate payer at the end of each period would be $\overline{r^s} - r^s(t)$ for each \$1 of notional principal. If the swap's fixed rate is greater than the variable rate at the end of a period (i.e., $\overline{r^s} > r^s(t)$), then the fixed-rate payer must pay the difference between the fixed interest payment on the notional principal to the floating-rate payer. Otherwise, the difference $\overline{r^s} - r^s(t)$ is negative, meaning that the fixed-rate payer receives the difference from the floating-rate payer. The net cost of the swap to the floating-rate payer is just the negative of this amount. For the sake of notational convenience, the discussion that follows assumes that all swaps have a notional principal of \$1, unless otherwise noted.

Uses of Interest Rate Swaps—Synthetic Financing

Firms use interest rate swaps to change the effective maturity of interest-bearing assets or liabilities. To illustrate, suppose a firm has short-term bank debt outstanding. At the start of each period this firm refinances its debt at the prevailing short-term interest rate, $r^b(t)$. If short-term market interest rates are volatile, then the firm's financing costs will be volatile as well. By entering into an interest rate swap, the firm can change its short-term floating-rate debt into a synthetic fixed-rate obligation.

Suppose the firm enters into an interest rate swap as a fixed-rate payer. Its resulting net payments in each period $t = 1, 2, \ldots, T$ of the agreement are determined by adding the net payments required of a fixed-rate payer to the cost of servicing its outstanding floating-rate debt.

	Period t cost of servicing outstanding short-term debt	$r^b(t)$
$+$	Period t cost of interest rate swap payments	$\overline{r^s} - r^s(t)$
$=$	Period t cost of synthetic fixed-rate financing	$\overline{r^s} + [r^b(t) - r^s(t)]$

Thus, the net cost of the synthetic fixed-rate financing is determined by the swap fixed rate plus the difference between its short-term borrowing rate and the floating-rate index.

Banks often index the short-term loan rates they charge their corporate customers to LIBOR. Suppose the firm in this example is able to borrow at LIBOR plus a credit-quality risk premium, or credit-quality spread, $q(t)$. Suppose further that the swap's floating-rate index is LIBOR. Then,

$$\begin{aligned} r^b(t) - r^s(t) &= [\mathrm{LIBOR}(t) + q(t)] - \mathrm{LIBOR}(t) \\ &= q(t). \end{aligned}$$

The period t cost of synthetic fixed-rate financing in this case is just $\overline{r^s} + q(t)$, the swap fixed rate plus the short-term credit-quality spread $q(t)$.

Now consider the other side to this transaction. Suppose a firm with outstanding fixed-rate debt on which it pays an interest rate of $\overline{r^b}$ enters into a swap as a floating-rate payer so as to convert its fixed-rate obligation to a synthetic floating-rate note. The net period t cost of this synthetic note is just the cost of its fixed-rate obligation plus the net cost of the swap:

$$\text{Period } t \text{ cost of synthetic floating-rate note} = r^s(t) + (\overline{r^b} - \overline{r^s}).$$

The cost of synthetic floating-rate financing just equals the floating rate on the interest rate swap plus the difference between the interest rate the firm pays on its outstanding fixed-rate debt and the fixed interest rate it receives from its swap counterparty.

Thus, interest rate swaps can be used to change the characteristics of a firm's outstanding debt obligations. Using interest rate swaps, firms can change floating-rate debt into synthetic fixed-rate financing or, alternatively, a fixed-rate obligation into synthetic floating-rate financing. But these observations raise an obvious question. Why would a firm issue short-term debt only to swap its interest payments into a longer-term, fixed-rate obligation rather than just issue long-term, fixed-rate debt at the outset? Conversely, why would a firm issue long-term debt and swap it into synthetic floating-rate debt rather than simply issuing floating-rate debt at the outset? The next two sections explore the rationales that have been offered to explain the widespread use of interest rate swaps.

2. INTEREST RATE SWAPS, ARBITRAGE, AND THE THEORY OF COMPARATIVE ADVANTAGE

The rapid growth of the swaps market in recent years strongly suggests that market participants must perceive significant benefits associated with the use of such instruments. The rationale most frequently offered by market participants is that interest rate swaps offer users an opportunity to reduce funding costs.[4] Bicksler and Chen (1986) present what is perhaps the best-known exposition of this viewpoint, which is based on the principle of comparative advantage. In international trade theory, the principle of comparative advantage explains the economic rationale for international trade by showing how different countries facing different opportunity costs in the production of different goods can benefit from free trade with other countries. According to Bicksler and Chen, differential information in different markets, institutional restrictions, and transactions costs create "some market imperfections and the presence

[4] For example, see Rudnick (1987).

of comparative advantages among different borrowers in these markets" (p. 646). These market imperfections, according to Bicksler and Chen, provide the economic rationale for interest rate swaps.

The Quality-Spread Differential

All firms pay a credit-quality premium over the risk-free rate when they issue debt securities. These credit-quality premiums grow larger as the maturity of the debt increases. Thus, whereas a firm, call it firm A, might pay a credit-risk premium of 50 basis points over the risk-free rate on its short-term debt obligations, the credit-quality premium it is required to pay on longer-term debt, say ten-year bonds, might rise to 100 basis points.

Not surprisingly, firms with good credit ratings pay lower risk premiums than firms with lower credit ratings. Moreover, the credit-quality premium rises faster with maturity for poorer credits than for good credits. Thus, if firm B has a poorer credit rating than firm A, it might pay a credit-risk premium of 100 basis points on its short-term debt while finding it necessary to pay 250 basis points over the risk-free rate to issue long-term bonds. The quality spread between the interest rate paid by the lower-rated firm and that paid by the higher-rated firm is only 50 basis points in the short-term debt market, but rises to 150 basis points at longer maturities. The quality-spread differential, the difference in the quality spread at two different maturities, is 100 basis points in this example. Firm A has an absolute cost advantage in raising funds in either the short- or long-term debt markets, but firm B has a comparative advantage in raising funds in short-term debt markets.

To explore this line of reasoning in more detail, suppose firms A and B both need to borrow funds for the next two periods, $t = 1, 2$. Let $r^f(t)$ denote the period t short-term (one-period) risk-free interest rate and $\overline{r^f}$ the long-term (two-period) fixed risk-free rate. The period t cost of short-term debt to firm A is the short-term risk-free rate plus the credit-quality spread $q_A(t)$. To issue long-term fixed-rate debt, firm A would be required to pay $\overline{r^f} + \overline{q_A}$, where $\overline{q_A}$ denotes the long-term quality spread. Define $q_B(t)$ and $\overline{q_B}$ analogously. Assuming firm A has the better credit rating,

$$q_A(1) \leq q_B(1), \text{ and}$$

$$\overline{q_A} \leq \overline{q_B}.$$

An increasing quality spread means that

$$q_B(1) - q_A(1) < \overline{q_B} - \overline{q_A}.$$

Conditions Necessary for Arbitrage to Be Feasible

Under certain assumptions, both firms could lower their funding costs if firm A were to issue long-term debt, firm B were to issue short-term debt, and they swapped interest payments. To see how this would work, assume A and B enter into an interest rate swap with B as a fixed-rate payer and A as the floating-rate payer. As above, let $\overline{r^s}$ denote the fixed swap rate for a two-period agreement. To minimize the notational burden, assume that the swap floating rate is just the risk-free rate of interest, $r^f(t)$. The resulting period t ($t = 1, 2$) net cost of synthetic fixed-rate financing to firm B is:

	Period t cost of servicing short-term, floating-rate debt	$r^f(t) + q_B(t)$
$+$	Period t cost of interest rate swap	$\overline{r^s} - r^f(t)$
$=$	Period t cost of synthetic fixed-rate financing	$\overline{r^s} + q_B(t)$

The synthetic fixed-rate financing will be less costly for firm B than actual fixed-rate financing in each period t if and only if

$$\overline{r^s} + q_B(t) \leq \overline{r^f} + \overline{q_B},$$

which implies

$$\overline{r^s} - \overline{r^f} \leq \overline{q_B} - q_B(t).$$

The term on the left-hand side of the last expression is the swap fixed-rate credit-quality spread, or risk premium, over the risk-free long-term interest rate. Thus, the quality spread associated with the swap fixed rate must be less than the increase in the credit-risk premium firm B would need to pay to issue long-term debt. Otherwise, synthetic fixed-rate financing will not be cheaper than actual fixed-rate financing.

Now examine the transaction from the vantage point of firm A, the floating-rate payer. The cost of synthetic floating-rate financing is determined by the cost of servicing fixed-rate debt plus the net cost of the swap:

	Period t cost of servicing fixed-rate debt	$\overline{r^f} + \overline{q_A}$
$+$	Period t cost of swap	$r^s(t) - \overline{r^s}$
$=$	Period t cost of synthetic floating-rate financing	$r^s(t) + (\overline{r^f} + \overline{q_A} - \overline{r^s})$

Period t synthetic floating-rate financing will cost less than actual floating-rate financing for firm A if

$$r^s(t) + (\overline{r^f} + \overline{q_A} - \overline{r_s}) \leq r^s(t) + q_A(t),$$

which, in turn, requires that

$$\overline{q_A} - q_A(t) \leq \overline{r^s} - \overline{r^f}.$$

That is, the increase in the credit-quality premium firm A must pay when issuing long-term fixed-rate debt must be smaller than the risk premium it receives from the swap's fixed-rate payer.

Combining results, firm A will have a comparative advantage in issuing long-term debt and firm B in issuing short-term debt if

$$\overline{q_A} - q_A(t) \leq \overline{r^s} - \overline{r^f} \leq \overline{q_B} - q_B(t), \qquad t = 1, 2.$$

For the floating-rate payer, synthetic floating-rate financing is cheaper than actual short-term financing if the interest rate swap quality spread (which the floating-rate payer receives) is greater than the added interest expense of long-term debt. For the fixed-rate payer, synthetic fixed-rate financing is less costly than issuing long-term bonds if the premium of the fixed swap rate over the two-period risk-free rate is less than the difference between its long-term and short-term quality spreads. Both parties will enjoy gains from trade if the swap floating-rate payer charges the fixed-rate payer a smaller credit-quality spread than the fixed-rate payer would be forced to pay in the bond market.

The astute reader will notice that the conditions outlined above require the parties to the agreement to know future values of $q_A(t)$ and $q_B(t)$. Both firms know their current short-term quality spreads along with $\overline{q_A}$ and $\overline{q_B}$ at the start of period 1. But it is unrealistic to assume that firms will know their future short-term quality spreads with certainty. Bicksler and Chen (1986) implicitly assume that firms expect the above relations to hold (at least on average) based on the past behavior of the quality-spread differential.

There is empirical evidence that long-term quality spreads for lower-rated counterparties are lower in the interest rate swap market than in credit markets (Sun, Sundaresan, and Wang 1993). Smith, Smithson, and Wakeman (1988) and Litzenberger (1992), among others, note that the expected loss to a swap counterparty in the event of a default is much less than that associated with holding a bond because interest rate swaps are not funding transactions and involve no exchange of principal. Moreover, swaps receive preferential treatment under the Bankruptcy Code in the event of a default. Under these conditions it may not seem surprising to find that quality spreads do not increase as rapidly in the swap market and that the cost of synthetic fixed-rate financing often seems lower than that of actual long-term financing. But while interest rate swaps might offer firms a way around paying increasing quality-spread differentials, synthetic fixed-rate financing does not offer firms the proverbial "free lunch." As the following discussion will show, the risks responsible for increasing quality-spread differentials do not disappear when firms use interest rate swaps.

Criticisms of the Comparative Advantage Rationale

Smith, Smithson, and Wakeman (1986, 1988) argue that observed behavior in the swap market is not consistent with classic financial arbitrage of the type described by proponents of the comparative advantage rationale. The use of interest rate swaps to arbitrage quality-spread differentials, they argue, should increase the demand for short-term loans among firms with poor credit ratings while reducing demand for "overpriced" long-term loans. Eventually, such a process should reduce quality-spread differentials and therefore reduce demand for interest rate swaps. In fact, Bicksler and Chen (1986) did report evidence of declining quality-spread differentials as interest rate swaps came into widespread use. But trading activity in interest rate swaps has shown no sign of abating even as quality-spread differentials have declined. To the contrary, the market for interest rate swaps has grown exponentially since these instruments were first introduced in the early 1980s. According to the International Swap and Derivatives Association, the total notional principal amount of interest rate swaps outstanding has risen from $683 billion in 1987 to just over $3.8 trillion as of year-end 1992.

Smith, Smithson, and Wakeman (1986, 1988) observe that much of the apparent savings from the use of swaps can be attributed to the absence of a prepayment option on generic swaps. Fixed-rate bonds typically carry a prepayment option that allows the borrower to call and refund a debt issue should market interest rates fall. The cost of this option is incorporated into the interest rate the firm is required to pay on such bonds. In contrast, the generic interest rate swap carries no such prepayment option. Early termination of a swap agreement requires the value of the contract to be marked to market, with any remaining amounts to be paid in full. A borrower can buy a "callable" swap, which permits early termination, but must pay an additional premium for this option. Thus, to be fair, the cost of actual long-term debt should be compared to the cost of callable synthetic fixed-rate financing, which would reduce the measured cost advantage resulting from the use of interest rate swaps.

Another problem with the comparative advantage rationale, noted by Smith, Smithson, and Wakeman (1988), is that it does not address the underlying reason for the existence of quality-spread differentials between short- and long-term debt. Loeys (1985) notes that short-term creditors implicitly hold an option to refuse to refinance outstanding loans. He attributes the difference in quality spreads between short- and long-term debt to the value of that implicit option.[5] But while this option is valuable to lenders, it increases the risk of a future funding crisis to the borrowing firm, thereby increasing the risk of bankruptcy proceedings. The risk that lenders will refuse to refinance outstanding short-term debt is known as liquidity risk, or rollover risk. From the firm's perspective, added liquidity risk represents an implicit cost of short-term financing.

Bansal, Bicksler, Chen, and Marshall (1993) compare the cost of synthetic fixed-rate financing with the cost of actual fixed-rate financing when the hidden costs noted above are taken into account. They control for the cost of liquidity

[5] Wall and Pringle (1987) note that Loeys' hypothesis is only consistent with increasing quality-spread differentials if the ability of short-term debtholders to refuse to renew outstanding debt makes it easier to force reorganization of a financially distressed firm.

risk by adding in the expense of a bank standby letter of credit in which a bank guarantees that it will assume a firm's outstanding debt if the firm finds itself unable to roll over a commercial paper issue. To take account of the value of a prepayment option, they add the premium on a callable swap into the total cost of synthetic fixed-rate financing. Finally, they also take account of transactions and administrative costs. The cost advantage of synthetic fixed-rate financing disappears once these costs are taken into account. Bansal et al. conclude that "a significant part of the reputed gains from swaps . . . were illusory, stemming from the way the gains have been calculated in practice" (p. 91).

3. ALTERNATIVE EXPLANATIONS

Smith, Smithson, and Wakeman (1988) hypothesize that the rationale for interest rate swaps lies with their usefulness in creating new synthetic financial instruments for risk management. The early 1980s brought unprecedented interest rate volatility, exposing firms to the risk of fluctuating funding costs. Rawls and Smithson (1990) argue that these events led to an increased demand for risk-management services on the part of firms. Smith, Smithson, and Wakeman (1988) argue that the growth of the swaps market effectively increased market liquidity for forward interest rate contracts, citing rapidly falling bid-ask spreads for interest rate swaps as evidence.[6] Thus, they argue, trading in interest rate swaps has helped to complete forward markets and to lower the cost to firms of managing their exposure to interest rate risk.

The Role for Hedging in the Theory of Corporate Finance

The foregoing discussion has focused on increased volatility in financial markets as the major factor behind the growth of the derivatives market in recent years. That firms would wish to hedge against the risk of such volatility simply has been assumed. But as Smith, Smithson, and Wilford (1990) note, much of textbook portfolio theory suggests that not hedging might be a firm's best policy. The well-known Modigliani-Miller theorem states that a firm's financing decisions have no effect on its market value when (1) a firm's management and outside investors share the same information about the returns accruing to all investment projects; (2) transactions costs are negligible; (3) a firm's tax bill is not affected by its financing decisions; and (4) the costs of financial distress are inconsequential. Under these assumptions, portfolio theory holds that individual investors can efficiently diversify away volatility in individual firm profits at

[6] An interest rate swap can be viewed as a bundle of forward contracts (see Smith, Smithson, and Wakeman [1988]). Sun, Sundaresan, and Wang (1993) find that bid-ask spreads in the interest rate swap market are smaller than those in the underlying market for long-term, fixed-rate corporate debt.

least as well as the firms themselves. If so, there is no reason for firms to expend resources hedging against volatility in future cash flows.

When these assumptions are relaxed, however, financing decisions may affect a firm's value. First, a firm's managers can be expected to know more about the risks and returns to different investment projects than outside investors. Second, the existence of transactions costs makes some kinds of financing decisions more costly than others. Third, a volatile cash flow stream can make a firm more susceptible to financial distress and bankruptcy, which can be extremely costly as well as threatening management with loss of control. Fourth, existing tax laws favor certain forms of funding over others. Firms are permitted to treat interest payments on debt as a tax-deductible expense, but not dividend payments to shareholders. Moreover, tax laws sometimes favor the use of certain derivative instruments to restructure cash flows. For all these reasons, firms will sometimes have incentives to hedge their cash flows.

Agency Costs as a Rationale for Interest Rate Swaps: Incentives to Undertake Synthetic Fixed-Rate Financing

Miller (1977) stresses the tax advantages of debt to explain why firms finance their investments with a combination of debt and equity. As Jensen and Meckling (1976) note, however, firms issued debt long before corporate income taxes came into existence. As an alternative rationale for debt, Jensen and Meckling emphasize the difficulty outside investors face in evaluating the performance of managers. As defined by Jensen and Meckling, an agency relationship is "a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision making authority to the agent" (p. 308). If principals could always costlessly monitor the behavior of their agents, they could ensure that agents would always act in their best interests. Monitoring the behavior of agents is costly, however, and requires principals to expend resources. Thus, the agent might be required to incur certain bonding expenditures. Finally, if principals cannot ensure that agents will always act in their best interests despite monitoring and bonding, there may be some deadweight residual loss. Jensen and Meckling define "agency costs" as the sum of these expenditures. They show that debt finance can reduce overall agency costs for a firm, but their analysis does not consider the problem of interest rate volatility and the question of whether a firm should issue short-term or long-term debt.

Interest rate volatility would not affect the investment or financing decisions of firms if revenues were always perfectly correlated with changes in market interest rates, because revenues would vary along with debt servicing costs in this case. Revenues typically are not perfectly correlated with market interest rates, however. As a result, interest rate volatility can increase the risk of financial distress. If financial distress is costly (because of the administrative costs of bankruptcy proceedings), or if the firm's management values its right to exercise control over the affairs of the organization, management will have an incentive to mitigate such risks. Ideally, then, a firm would wish to schedule repayment of its capital financing costs to match the realization of revenues from its investments (Myers 1977). If a firm's revenues are completely uncorrelated with market interest rates, it could minimize the risk of future financial distress by funding long-term investments with long-term, fixed-rate debt and short-term investments with short-term debt.

Long-term lending carries substantial risks from an outside investor's viewpoint, however. A borrower's financial condition can deteriorate substantially over the term of the loan. Moreover, as Jensen and Meckling (1976) note, management has an incentive to take actions that benefit shareholders at the expense of creditors once a firm has received the proceeds of the loan. As an example, management can pursue high-risk strategies or otherwise attempt to dissipate the organization's assets by paying excessive dividends. Creditors could prevent such behavior if (1) they always knew as much about a firm's investment opportunities as its managers and (2) they could monitor management's behavior costlessly. But such actions are prohibitively costly for most creditors, if even feasible, for they would involve duplicating essentially all the functions of management. For these reasons, bondholders often demand loan covenants that limit management's discretion in deploying loan proceeds. Typically such covenants give creditors the right to exercise greater control over the firm when a condition of the loan is violated or in the event of a material deterioration in its financial condition. To be certain, enforcement of loan covenants still requires some monitoring on the part of creditors. Jensen and Meckling (1976) argue that these monitoring costs are ultimately borne by borrowers through higher interest rates.

Wall (1989) argues that the existence of agency costs is one reason that quality spreads widen with debt maturity. He notes that while established firms with good credit ratings and access to low-cost credit have incentives to limit risks, newer and smaller firms do not have the same incentives. Like Loeys (1985), Wall gives special emphasis to the influence creditors can exercise over borrowers when renegotiating short-term loans. Wall was among the first to observe that synthetic fixed-rate financing carries different incentives for borrowers than actual fixed-rate financing. To understand why this might be so, notice that the interest rate lenders charge a borrower when renewing a short-term loan can change for two reasons: (1) a change in market interest rates or (2) a change in the firm-specific credit-quality risk premium. Interest rate swaps compensate the borrower only for changes in market rates, and not for changes in the short-term quality spread. Thus, as noted earlier, the cost of synthetic fixed-rate financing is $\overline{r^s} + [r^b(t) - r^s(t)]$, the swap fixed rate plus the quality spread between the rate the firm pays on its short-term debt and the swap floating-rate index. A firm that chooses synthetic fixed-rate financing faces the risk that the quality spread $[r^b(t) - r^s(t)]$ might rise if lenders realize that management has increased the firm's riskiness. In extreme cases, the firm might even find itself unable to roll over its outstanding short-term debt and be forced into bankruptcy proceedings.

Wall's (1989) rationale for interest rate swaps lies with the observation that synthetic fixed-rate financing should discourage management from pursuing risky investment strategies.[7] According to this argument, interest rate swaps lower funding costs by controlling the adverse incentives a firm's management might have to increase the risk assumed by the firm to the detriment of creditors. Thus, interest rate swaps do make it possible for firms to reduce financing costs in Wall's theory. But the savings attributable to the use of swaps result from lower agency costs and do not constitute arbitrage in the sense that term is normally understood.

The Problem of Adverse Selection: More Incentives to Borrow Short and Swap into Fixed

Flannery (1986) and Diamond (1991) investigate the determinants of debt maturity by focusing attention on the incentives borrowers have to signal information to lenders about their financial condition. Their analysis is based on the assumption that outside investors have imperfect information about firms, and so are unable to discriminate perfectly between safe firms and relatively risky firms. If outside investors cannot perfectly discriminate between risky and safe firms, they will demand default-risk premiums on long-term debt that may appear excessively high to relatively safe borrowers. Conversely, the managers of a risky firm recognize that there is a high probability that the organization's financial condition will deteriorate, leading them to prefer long-term debt over short-term debt.

Firms that lenders can identify as risky borrowers have difficulty securing long-term loans and are forced to issue short-term debt that matures before the returns to an investment are realized. Often such firms must obtain their credit lines from banks, which specialize in credit evaluation and are well positioned to monitor the firm's activities. In the event that a firm's financial condition deteriorates, lenders can demand a higher interest rate upon refinancing, can further restrict the discretion of management, can engage in more intensive monitoring, can take some combination of these actions, or can even refuse to refinance outstanding debt. A firm that defaults on its outstanding debt obligations can be forced into bankruptcy proceedings.

If a firm's management believes that default premiums on long-term loans are excessive, it might choose a short-term funding strategy. By voluntarily

[7] Diamond (1984) makes a similar observation regarding the optimal hedging behavior of firms, although he does not discuss the rationale for interest rate swaps per se.

taking on liquidity risk, management can signal that it does not expect the firm's condition to deteriorate in the future. Over time a firm that consistently demonstrates its ability to meet its financial obligations develops a reputation as a safe firm. Thus, a safe firm might employ a short-term funding strategy until it can convince creditors to extend long-term loans on better terms (Diamond 1991).

One drawback to such a strategy is that it can leave the firm's cash flows unhedged. Arak, Estrella, Goodman, and Silver (1988) stress that interest rate swaps are not redundant securities, but offer firms new financing choices that were not previously available in credit markets. Like Wall (1989), Arak et al. note that synthetic fixed-rate financing requires the borrower to bear the risk of changes in the short-term credit-risk premium. Their hypothesized rationale for interest rate swaps differs somewhat from that of Wall, however. They hypothesize that firms may have an incentive to bear rollover risk when management is more optimistic about a firm's future prospects than the market. If a firm's management is optimistic about its financial condition, it may choose to issue short-term debt in the expectation that the quality spread will fall in the future. In effect, the firm speculates on its own quality spread while using swaps to immunize itself against market risk.

Titman (1992) and Minton (1993a) derive conditions under which a firm's best strategy is to use interest rate swaps in conjunction with short-term financing. Like Flannery (1986) and Diamond (1991), Titman and Minton emphasize that firms may have an incentive to bear the liquidity risk associated with short-term debt finance as a means of signaling management's belief that the firm's financial prospects will improve. Titman finds conditions under which synthetic fixed-rate financing gives firms an incentive to undertake safer investments. Minton finds that giving firms the option of using interest rate swaps can reduce default risk and, in doing so, increase the capacity of firms to undertake productive long-term investment. Both Titman and Minton find plausible conditions under which interest rate swaps reduce financing costs, albeit not through the channels of financial arbitrage.

Notice that the basic logic of the adverse selection rationale runs closely parallel to that of Wall's (1989) agency cost rationale. While borrowers in Titman and Minton's models choose short-term financing to signal management's belief that the firm is creditworthy, the act of taking on short-term debt mitigates incentives to take on added risk once loan proceeds are received, just as Wall predicts.

Incentives to Borrow Fixed and Swap into Floating

The preceding discussion has focused on the incentives firms might have to enter into a swap as a fixed-rate payer. But every swap agreement must also have a floating-rate payer. Wall (1989) and Titman (1992) hypothesize that floating-rate payers share in the gains fixed-rate payers receive from synthetic fixed-rate financing. Litzenberger (1992) notes at least two reasons why highly rated firms may be able to lower funding costs by issuing callable fixed-rate debt and then swapping into synthetic floating-rate debt. First, like Wall and Titman, he hypothesizes that floating-rate payers essentially act as financial intermediaries that earn income in return for managing a diversified portfolio of risky contractual obligations. The total exposure resulting from this activity is small, he argues, because (1) the credit risk associated with an interest rate swap is much smaller than that associated with actual lending; (2) most swap agreements take place among parties with at least single A credit ratings (lower-rated counterparties are rejected or required to post collateral); and (3) a diversified swap portfolio has little risk of a large credit loss.

Second, Litzenberger also notes that the highly rated AAA firms that typically become floating-rate payers often issue callable fixed-rate notes, and then sell the prepayment options on these notes by selling callable swaps to swap dealers. He argues that such transactions can create synthetic floating-rate financing at a modest savings in cost because the prepayment options attached to fixed-rate debt tend to be underpriced, probably because of a past history of non-optimal exercises on such options. Thus, Litzenberger attributes at least part of the incentive to become a floating-rate payer to arbitrage opportunities created by the mispricing of prepayment options for corporate bonds.

Smith, Smithson, and Wakeman (1988) emphasize that interest rate swaps can help to conserve on transactions costs. As an example, they note that it can be cheaper to sell an interest rate swap than to call and refund outstanding fixed-rate debt.

4. A COMPARISON OF INTEREST RATE FUTURES AND INTEREST RATE SWAPS

A discussion of the economic role of interest rate swaps would not be complete without at least some mention of interest rate futures. Interest rate futures can be used to create synthetic fixed-rate debt in much the same way as interest rate futures. In particular, selling a "strip," or sequence, of Eurodollar futures with successive maturity dates can be compared to buying an interest rate swap.

To see how interest rate futures can substitute for an interest rate swap, recall that the buyer (fixed-rate payer) of an interest rate swap receives a net payment from the seller whenever the floating-rate index exceeds the swap fixed rate. In the case of a generic swap with a floating rate indexed to some maturity of LIBOR, the buyer receives the difference in interest on the notional principal amount whenever the specified maturity of LIBOR exceeds the swap fixed rate. When LIBOR is below the fixed rate, the buyer must pay the difference in interest to the seller.

Selling a strip of Eurodollar futures creates a similar pattern of returns and payments. The seller of a Eurodollar contract receives the difference in interest on the notional principal ($1 million) when the futures rate negotiated at the outset of the agreement turns out to be less than the value of three-month LIBOR prevailing on the contract maturity date. Otherwise, the seller must pay the difference in interest to the buyer. Thus, selling a strip of Eurodollar futures produces a return stream comparable to that of a generic interest rate swap. Because of this similarity, an implied swap rate can be derived from Eurodollar futures rates.[8] Minton (1993b) finds evidence that the behavior of swap market rates is closely related to this implied swap rate.

These observations suggest that much of the rationale for interest rate swaps discussed above must also apply to interest rate futures—in particular, to Eurodollar futures. The foregoing discussion has focused on interest rate swaps because the growth of trading in Eurodollar futures in recent years appears to have been driven by the growth of the swap market. Although trading in Eurodollar futures predates the advent of the interest rate swap, trading was limited to contracts extending two years into the future at the time of the first widely publicized interest rate swap in 1982. As a result, Eurodollar futures were not as well suited for use in creating synthetic long-term financing as were interest rate swaps. More recently, the Chicago Mercantile Exchange has begun listing Eurodollar futures for delivery as far as ten years into the future. Burghardt et al. (1991) attribute the recent expansion of trading in Eurodollar futures to the growth of the interest rate swap market. Swap dealers in particular often use Eurodollar futures to hedge their commitments. Thus, although interest rate futures contracts can substitute for interest rate swaps, it was the growth of the swap market that had the greatest effect on corporate finance.

Kawaller (1990) and Minton (1993b) discuss the factors influencing the choice between interest rate futures and interest rate swaps. Kawaller emphasizes transactions costs and other practical considerations of managing a futures position as key factors influencing the choice between interest rate futures and interest rate swaps. The main benefit of a swap is that it can be custom-tailored to the needs of an individual firm, so that managing an interest rate swap is relatively easy compared to managing a futures market position. A firm that enters into an interest rate swap faces a set schedule for receiving or making its payments. As long as nothing happens to change the firm's underlying exposure to interest rates—that is, as long as nothing has happened to change the reasons the firm decided to create synthetic fixed- or floating-rate financing in the first place—managing an outstanding swap position merely requires the firm's treasurer to make or collect scheduled payments.

[8] For more detailed expositions, see Burghardt et al. (1991) and Kawaller (1990).

The principal disadvantage of interest rate swaps relative to interest rate futures lies with counterparty credit risk. Exchange-traded instruments such as interest rate futures are backed by a system of margin requirements, along with the guarantee of the exchange clearinghouse (which, in turn, is jointly backed by the paid-in capital of the clearinghouse member firms). This system of safeguards removes virtually all risk of default in the futures market. In contrast, a counterparty to an interest rate swap is exposed to the risk that the other counterparty might default. To be certain, most interest rate swaps take place between relatively creditworthy counterparties. Nonetheless, credit risk is a greater concern with interest rate swaps than with futures contracts.

The very factors that make interest rate futures safer also make managing a futures position somewhat more challenging than managing a swap commitment, however. First, a party to a futures contract is required to post margin before being permitted to buy or sell a futures contract. Second, the futures exchanges mark all outstanding positions to market at the end of each trading session, adding any realized gains or subtracting any realized losses from each trader's margin account. While this procedure minimizes default risk, it exposes any party with an open futures position to the risk of margin calls. As a consequence, payments are less predictable in the short run with a futures position than with an interest rate swap. Third, futures contracts are standardized agreements. Contract standardization, along with the clearinghouse guarantee, facilitates trading in futures contracts. Futures markets tend to be more liquid than OTC markets (and actual cash markets for that matter) as a result, lowering transactions costs. But while contract standardization facilitates trading, it also means that an interest rate futures contract will almost never be perfectly suited to the needs of any one trader.[9] Thus, an interest rate futures position requires greater monitoring and can be more difficult to execute unless a firm maintains a staff devoted to trading futures contracts.

While the factors that determine the choice between interest rate futures and interest rate swaps is of great interest to practitioners, it has not received a great deal of attention in the academic finance literature. Minton's (1993a) model of the hedging behavior of firms is a noteworthy exception. Minton finds that relatively safe firms—firms that expect their future credit-quality spreads to fall—may have an incentive to choose swaps over futures contracts so as to avoid the cost of margin requirements.

5. CONCLUDING COMMENTS

The reasons for the extraordinary growth of the swap market in recent years are not yet fully understood. But there seems to be a consensus that the market has

[9] For a more detailed description of futures exchanges and interest rate futures, see Kuprianov (1993a).

developed because interest rate swaps offer firms financing choices that were not available before the advent of these instruments. In this respect, interest rate swaps represent a true financial innovation.

The early rationale offered for the existence of the market—that firms used interest rate swaps to arbitrage credit market inefficiencies—cannot by itself explain the exponential growth of the market over the past decade. By the same token, it is unlikely that firms would use interest rate swaps if they did not lower financing costs in some way. Recent research suggests at least two reasons why firms use interest rate swaps. First, in cases where a firm's management expects its financial condition to improve, interest rate swaps make it possible for firms to hedge against changes in market interest rates while avoiding excessive fixed-rate quality-spread premiums. Second, interest rate swaps make possible financial arrangements that reduce the incentives of borrowing firms to take on added risk at the expense of creditors.

Conceived in the wake of unprecedented interest rate volatility brought about by a decade of accelerating inflation, the interest rate swap was born of necessity. In a period of low interest rate volatility, the choice between short- and long-term borrowing was primarily a choice between fixed and floating credit-quality spreads. With rising interest rate volatility, however, the ability to separate the effects of changes in market rates from changes in credit-quality spreads became more valuable, leading firms to experiment with alternative financing schemes. Based on the results of recent research, it appears that interest rate swaps have helped firms to weather the uncertainties of volatile financial markets by reducing default risk and facilitating increased productive investment.

REFERENCES

Arak, Marcelle, Arturo Estrella, Laurie S. Goodman, and Andrew Silver. "Interest Rate Swaps: An Alternative Explanation," *Financial Management,* Summer 1988, pp. 12–18.

Bansal, Vipul K., James L. Bicksler, Andrew H. Chen, and John F. Marshall. "Gains from Synthetic Financing with Interest Rate Swaps: Fact or Fancy?" Continental Bank *Journal of Applied Corporate Finance,* vol. 6 (Fall 1993), pp. 91–94.

Bicksler, James, and Andrew H. Chen. "An Economic Analysis of Interest Rate Swaps," *The Journal of Finance,* vol. 41 (July 1986), pp. 645–55.

Burghardt, Belton, Luce, and McVey. *Eurodollar Futures and Options*. Chicago: Probus Publishing Company, 1991.

Cummins, Claudia. “Gonzalez Moving Legislation on Derivatives to Front Burner,” *American Banker,* May 4, 1994a.

__________. “OCC Eyes Lid on Swaps,” *American Banker,* April 21, 1994b.

Diamond, Douglas W. “Debt Maturity and Liquidity Risk,” *The Quarterly Journal of Economics,* vol. CVI (August 1991), pp. 709–37.

__________. “Financial Intermediation and Delegated Monitoring,” *Review of Economic Studies,* vol. 51 (August 1984), pp. 393–414.

Flannery, Mark J. “Asymmetric Information and Risky Debt Maturity Choice,” *The Journal of Finance,* vol. 41 (March 1986), pp. 19–37.

Goodfriend, Marvin. “Eurodollars,” in Timothy Q. Cook and Robert K. LaRoche, eds., *Instruments of the Money Market,* 7th ed. Richmond: Federal Reserve Bank of Richmond, 1993, pp. 48–58.

Jensen, Michael C., and William H. Meckling. “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure,” *The Journal of Financial Economics,* vol. 3 (October 1976), pp. 305–60.

Karr, Albert R. “Bank Regulator Signals Move on Derivatives: Comptroller Voices Concern over For-Profit Trades and Internal Oversight,” *The Wall Street Journal,* April 21, 1994.

Kawaller, Ira G. “A Swap Alternative: Eurodollar Strips,” in Carl R. Beidleman, ed., *Interest Rate Swaps*. Homewood, Ill.: Business One Irwin, 1990, pp. 390–404.

Kuprianov, Anatoli. “Money Market Futures,” in Timothy Q. Cook and Robert K. LaRoche, eds., *Instruments of the Money Market,* 7th ed. Richmond: Federal Reserve Bank of Richmond, 1993a, pp. 188–217.

__________. “Over-the-Counter Interest Rate Derivatives,” in Timothy Q. Cook and Robert K. LaRoche, eds., *Instruments of the Money Market,* 7th ed. Richmond: Federal Reserve Bank of Richmond, 1993b, pp. 238–65.

Litzenberger, Robert H. “Swaps: Plain and Fanciful,” *The Journal of Finance,* vol. 47 (July 1992), pp. 831–50.

Loeys, Jan G. “Interest Rate Swaps: A New Tool for Managing Risk,” Federal Reserve Bank of Philadelphia *Business Review,* May/June 1985, pp. 17–25.

Marshall, John F., and Kenneth R. Kapner. *Understanding Swaps*. John Wiley and Sons: New York, 1993.

Miller, Merton H. “Debt and Taxes,” *The Journal of Finance,* vol. 32 (May 1977), pp. 261–75.

Minton, Bernadette. “Interest Rate Derivative Products and Firms’ Borrowing Decisions: The Case of Interest Rate Swaps and Short-Term Interest Rate Futures Contracts.” Manuscript. University of Chicago, December 1993a.

__________. “An Empirical Examination of U.S. Dollar Swap Spreads.” Manuscript. University of Chicago, November 1993b.

Myers, Stewart C. “Determinants of Corporate Borrowing,” *Journal of Financial Economics*, vol. 5 (November 1977), pp. 147–75.

Rawls, S. Waite III, and Charles W. Smithson. “The Evolution of Risk Management Products,” *Journal of Applied Corporate Finance,* reprinted in Robert J. Schwartz and Clifford W. Smith, Jr., eds., *The Handbook of Currency and Interest Rate Risk Management.* New York: New York Institute of Finance, 1990.

Rehm, Barbara A. “Trio of Heavyweights Sees No Swaps Crisis,” *American Banker,* May 11, 1994.

Rudnick, Linda T. “Discussion of Practical Aspects of Interest Rate Swaps,” in *Conference on Bank Structure and Competition.* Chicago: Federal Reserve Bank of Chicago, 1987, pp. 206–13.

Smith, Clifford W., Charles W. Smithson, and Lee Macdonald Wakeman. “The Market for Interest Rate Swaps,” *Financial Management,* vol. 17 (1988), pp. 67–73.

__________. “The Evolving Market for Swaps,” *Midland Corporate Finance Journal*, vol. 3 (Winter 1986), pp. 20–32.

Smith, Clifford W., Charles W. Smithson, and D. Sykes Wilford. “Financial Engineering: Why Hedge?” *Intermarket*, vol. 6 (1989), pp. 12–16.

Sun, Tong-Sheng, Suresh Sundaresan, and Ching Wang. “Interest Rate Swaps: An Empirical Investigation,” *Journal of Financial Economics*, vol. 34 (August 1993), pp. 77–99.

Titman, Sheridan. “Interest Rate Swaps and Corporate Financing Choices,” *Journal of Finance,* vol. 47 (September 1992), pp. 1503–16.

Wall, Larry D. “Interest Rate Swaps in an Agency Theoretic Model with Uncertain Interest Rates,” *Journal of Banking and Finance,* vol. 13 (May 1989), pp. 261–70.

__________, and John J. Pringle. “Alternative Explanations of Interest Rate Swaps,” in *Conference on Bank Structure and Competition.* Chicago: Federal Reserve Bank of Chicago, 1987, pp. 186–205.

John Wheatley's Theory of International Monetary Adjustment

Thomas M. Humphrey

Of the bullionist writers who advocated restoration of the gold convertibility of England's currency during the Bank Restriction period 1797–1821, few are as little known today as John Wheatley. Certainly his name is not as familiar as those of David Ricardo, Henry Thornton, Thomas Malthus, Francis Horner, William Huskisson, and other bullionists. Yet in some respects he was the most original of the group. His *Essay on the Theory of Money and Principles of Commerce* (1807) spelled out the logic and implications of the strict bullionist position more forcefully and systematically than any document before Ricardo's *High Price of Bullion: A Proof of the Depreciation of Bank Notes* (1810).

To Wheatley belongs much of the credit for expounding at least four hard-line bullionist propositions often attributed to Ricardo. First, money-stock changes have no effect on output and employment. Second, exchange rate depreciations, a high price of gold, and specie drains stem solely from an excess issue of currency. Third, being purely monetary phenomena, exchange rate changes, gold price movements, and specie drains are immune to real shocks operating through the balance of payments. Fourth, exchange rate depreciation and the excess of market over mint price of gold constitute proof and measure of overissue in inconvertible paper regimes. To these can be added a fifth contribution: his demonstration that monetary expansion and price inflation can continue indefinitely on a given gold base if all countries expand in step.

Wheatley derived these propositions from an analytical model characterized by sharp dichotomization of real and monetary sectors. He sought to show that monetary shocks do not affect real variables nor real shocks monetary

The author is greatly indebted to his colleagues Tim Cook, Mary Finn, Bob Hetzel, and Peter Ireland for their penetrating criticisms of earlier drafts of this article. The views expressed are those of the author and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

variables. To do so, he partitioned his real and monetary variables into separate compartments and allowed little or no interaction between the two. Neutrality, block exogeneity, and absence of reverse causality—these were the hallmarks of his analysis. They allowed him to contend that his monetary indicators were uncontaminated by real disturbances. As such, they signaled overissue pure and simple and so constituted an unambiguous measure of the need for monetary contraction to correct the excess issue. More than most economists before or since, he took the extreme position that monetary shocks affect only monetary variables and real shocks real variables.[1]

Despite Wheatley's originality, his work has suffered from neglect. Ignored in his own time because of a labored, archaic expository style and a hypercritical, vitriolic attitude toward his fellow economists, he has also been underrated in ours.[2] Modern commentators, when they mention him at all, typically focus exclusively on certain striking aspects of his work rather than on his complete analytical model. Thus Schumpeter (1954) concentrates on his crude version of the quantity theory of money. Einzig (1962), Frenkel (1978), Officer (1984), and Wu (1939) emphasize his purchasing power parity doctrine. Fetter (1942) and Viner (1937) spotlight his assumption of price and exchange rate invariance to real shocks. Metzler (1948), Morgan (1943), O'Brien (1975), and Viner (1937) accent his income-expenditure theory of unilateral transfers. Chipman (1983) criticizes his theory of gold price determination. None, however, mention his integration of these elements into a consistent theory of how an open economy responds to real and monetary disturbances.

The result is a gap in our knowledge of Wheatley's theory of the international adjustment mechanism. This gap is all the more regrettable because it contributes to the notion of a monolithic classical theory based on David Hume's account of the price-specie-flow mechanism. In fact, Wheatley's theory differs from Hume's. It emphasizes continuous purchasing power parity, gold market arbitrage, and unilateral payments accomplished through income changes rather than through price adjustments. It demonstrates that there is more than one classical theory of the international mechanism.

This article represents an effort to fill the gap and to give Wheatley his due. First, it specifies the basic building blocks of Wheatley's model. Second, it shows how he used these components to explain international adjustment to monetary disturbances under convertible and inconvertible currency regimes. Third, it shows how he modified his model to handle real disturbances and in so doing contributed to the theory of international transfers. Fourth, it outlines

[1] Wheatley's strong real-nominal dichotomy has its closest modern counterpart in the work of the contemporary real business cycle school founded by Kydland and Prescott (1982) and Long and Plosser (1983).

[2] On Wheatley's expository style and attitude toward his fellow economists, see Chipman (1983), pp. 7 and 49–50, and Hollander (1911), p. 464.

some policy implications of his analysis. Last, it evaluates his place in classical monetary thought.

1. BASIC BUILDING BLOCKS

The first task is to sketch the chief components of Wheatley's model. These consist of (1) the quantity theory of money, (2) the purchasing power parity doctrine, and (3) a theory of gold arbitrage. Together, they trace out a causal chain in which money determines prices, prices determine exchange rates, and exchange rate movements trigger specie drains under metallic and convertible currency regimes and currency depreciation under inconvertible paper regimes. Of the three components, the quantity theory employs closed-economy propositions; the parity and arbitrage doctrines, open-economy ones.

Quantity Theory

Wheatley adheres to a particularly strict or rigid version of the quantity theory. This version embodies the notions of (1) proportionality of money and prices, (2) money-to-price causality, (3) neutrality of money, (4) monetary rather than real theory of prices, and (5) exogeneity of the nominal stock of money.

On the proportionality postulate, Wheatley (1807) declares that "all prices are in proportion to the quantity of money." "This principle," he writes, "appears so obvious, that it would be superfluous to enter into the proof of its validity; and I shall assume it as a postulate that would be universally conceded" (p. 12). In symbols, $P = kM$, where P is the home country's price level, M is its money stock, and k is a constant coefficient equal to the ratio of the circulation velocity of money to real output—both variables (velocity and output) treated as fixed constants by Wheatley. A similar equation holds for the foreign country in Wheatley's two-country model. That is, $P^* = k^* M^*$, where the asterisks denote foreign country variables. Thus, "if the currency of one country is relatively greater than the currency of another, its price will be proportionally higher" (1819, p. 24), or $P/P^* = K(M/M^*)$, where K is the ratio of the constants k and k^*.

As for the notion of money-to-price causality, Wheatley endorses it in no uncertain terms. "Prices," he says, "are determined by the quantity of money" (1819, p. 24). Thus a money-stock expansion has "no other operation than to raise the price of produce" (1807, p. 38). It "has no other effect than to cause its own depression" in purchasing power $1/P$ (p. 37). In his view, money drives prices through a direct expenditure mechanism. A monetary expansion raises the ratio of cash to nominal transactions above the fraction people desire to hold. Cash holders spend the excess money immediately and mechanically in an effort to work off the unwanted balances. The increased spending raises prices and the volume of nominal transactions. The process ends when the desired cash ratio is restored and the new money is willingly held.

Having asserted direct causality, he implicitly rejects the notion of reverse causality running from prices to money. Reverse causality might arise in metallic currency regimes if an exogenous fall in the price level, by raising the goods value of gold, induces an increase in the supply of specie flowing either from the mines or from abroad through the balance of payments. It also might arise in a fiat paper regime if the central bank validates exogenous price increases with monetary expansion. Wheatley, however, says nothing about such mechanisms.

He does, however, expand at length on the quantity theory's neutrality proposition. He argues that money-stock changes exert no influence, temporary or permanent, on real output and employment. Perfect wage-price flexibility, he claims, ensures as much. Such flexibility means that nominal wages and prices adjust instantaneously and equiproportionally to monetary shocks. The result is that the real wage rate and thus the output and employment variables it determines are invariant to such shocks. Thus, contrary to the popular notion that "an increase of currency gives a stimulus to industry by the elevation of prices," the truth is that "no greater stimulus can in reality exist" (1807, p. 40). For "the wages of labour are augmented only in proportion to the increase [of money and prices], and purchase no greater quantity of produce after the addition than before it" (p. 40). Wheatley concludes: "[A]n increase of money . . . has no effect like an increase of produce to augment the wealth of a nation" (p. 37). It has "no other operation than to raise the price of produce, and augment the nominal incomes of all, without making any addition to their real opulence" (1807, p. 38). One can hardly find a clearer statement of the neutrality proposition in the entire classical literature.

Wheatley's neutrality proposition states only that monetary shocks affect prices and not real variables. It does not deny that real determinants might also drive the price level. To rule out this possibility, Wheatley asserts that price changes stem exclusively from monetary rather than real disturbances. "There is no other cause," he writes, "than a relative excess of currency which makes prices higher" (1819, p. 24). True, he notes that real shocks might depress output and so the demand for money, thereby rendering the existing money stock excessive. But he argues that any resulting price increase must be attributed to the monetary excess and not to the real shock. Likewise, he denounces as "imbecilic" Sir James Steuart's view that the same real forces of "demand and competition" that determine relative prices also drive the general price level (1807, p. ix). Not so, says Wheatley. Money-stock movements govern the general price level. He also asserts that secular inflations emanate from excessive paper money growth rather than from output contractions and a shortage of goods.

Finally, Wheatley posits money-stock exogeneity, which he sees as a corollary of the proposition that money is the independent causal factor governing prices. He realizes that if money were an endogenous variable responding passively to prior changes in the economy, he could not claim it plays the active initiating role in raising prices. For this reason, he treats gold coin and paper notes as exogenous variables emanating autonomously from “the fertility of mines, and the general publication of state and bank paper” (1807, p. 60). He ignores the feedback effect of prices on the profitability of mining and the production of gold. By dismissing such real determinants of commodity money, he essentially treats it as fiat currency.

Purchasing Power Parity Doctrine

The purchasing power parity doctrine forms the second building block in Wheatley’s model of international adjustment. In what is the clearest and most complete account of the doctrine before Gustav Cassel’s statement in the 1920s, he enunciates it in both its absolute and relative versions.

The absolute version says that the exchange rate E—defined as the domestic currency price of a unit of foreign currency—equals the ratio of domestic to foreign general price levels. This condition renders the purchasing power of money expressed in terms of a single currency everywhere the same. In symbols, $E = P/P^*$ or, equivalently, $P = EP^*$. As Wheatley (1807) puts it, “the course of exchange is exclusively governed by the relative state of prices, or the relative value of money, in the different countries between whom it is negotiated” (p. 85). It “approximates the price of their produce to a general level” (p. 45). He sees price-level parity $P = EP^*$ as emanating from the law of one price. That law, of course, states that abstracting from tariffs, transport costs, and other impediments to trade, the price of any given traded good is the same in all locations when quoted in the same currency. Since Wheatley assumes that all goods are traded and that identical commodities bear the same weight in each country’s price index and product mix, he essentially treats them as a single composite commodity. In his composite-commodity model, the law of one price applies to aggregate price levels as well as to the prices of individual goods. Therefore, the real exchange rate, or commodity terms of trade, between two nations equals unity, or $EP^*/P = 1$. In Wheatley’s words, “the state of exchange must uniformly coincide with the state of prices, or the interchange of produce could not be transacted on equal terms” (1819, p. 21). A unitary value of the terms of trade permits “any given quantity” of the composite commodity to “exchange for the same value in every part of the world” (1807, p. 46).

He gives an equally lucid statement of the doctrine’s relative version, according to which the percentage change of the exchange rate is the differential between the percentage changes of the price levels. “The exchange between London and Hamburgh,” he says, “is at any given moment five percent against London, only because the general prices at Hamburgh are at that time five percent lower than the general prices of London” (1807, p. 63). Here is his description of the relationship $e = p - p^*$, where the lowercase letters denote percentage changes in their uppercase counterparts.

Wheatley saw the parity doctrine as an extension of the quantity theory to the open economy. It is therefore not surprising to find him treating the exchange rate as a purely monetary phenomenon determined by relative national money stocks operating through national price levels. "Nothing can alter the state of the exchange," he wrote, "that does not alter the state of prices, and nothing can alter the state of prices, the quantity of produce being the same, that does not alter the state of currency" (1819, p. 27). In terms of the quantity theory equations presented above, Wheatley held that $E = P/P^* = kM/k^*M^* = K(M/M^*)$. From this equation derives his conclusion that "there is no other cause than a relative excess of currency, which makes the exchange unfavourable" (1819, p. 24). Thus "the course of exchange is the exclusive criterion of how far the currency of one country is increased beyond the currency of another" (1803, p. 207).

Wheatley comments at length on key propositions of the doctrine. Regarding causality, he asserts it runs unidirectionally from price levels to exchange rates. Regarding the transmission mechanism, he posits rational expectations. He argues that agents, observing price-level changes and anticipating the compensating exchange rate correction, incorporate those anticipations into the rate and so make them a reality.[3] The result is that exchange rates adjust instantaneously to price levels.[4] Regarding deviations from purchasing power parity, he denies their occurrence. Rational expectations maintain the exchange rate at parity equilibrium. Therefore, currencies cannot be temporarily over- or undervalued on the market for foreign exchange. Finally, regarding neutrality, he argues that because exchange rates are always at parity, their fluctuations cannot affect real trade balances or the terms of trade in the least. His treatment of these issues is among the more extreme and uncompromising in the classical literature.

[3] According to Wheatley, the exchange rate attains purchasing power parity equilibrium so swiftly as to bypass trade-balance effects. Other classical economists, notably Henry Thornton (1802, p. 198), had suggested that a rise in one country's prices relative to another's could only exercise its self-correcting influence on the exchange rate through the trade balance. In their view, a rise in home prices relative to foreign prices would, by spurring imports and checking exports, cause a trade deficit and a corresponding excess demand for foreign exchange to finance it. This excess demand in turn would bid up the exchange rate, thereby equalizing common-currency price levels. But Wheatley thought such intermediate steps unnecessary. In his view, the exchange rate adjusts immediately.

[4] "But immediately that the prices of one country become relatively higher than the prices of another, the course of exchange . . . will note the difference, and become unfavourable to the same extent" (Wheatley 1819, p. 24). "The instant that the variation occurs it is announced by the exchange, and a credit given in conformity to the difference" (Wheatley 1807, p. 64).

Gold Price Arbitrage

Wheatley's theory of gold arbitrage constitutes the third building block of his model. It explains how countries eliminate their excess or deficient money stocks under metallic and convertible currency regimes.

His theory is simplicity itself. It says that arbitrageurs will ship gold from where its monetary value is low to where its monetary value is high. Comparison of fixed domestic and foreign mint prices of gold quoted in a single currency at the rate of exchange reveals these respective monetary values. Let a rise in the exchange rate—that is, a depreciation of the home currency relative to the foreign currency—raise the common-currency value of gold abroad over its value at home. The result will be to precipitate a specie export. For arbitrageurs will find that they can convert domestic paper into gold at the fixed mint price, sell the gold abroad at its foreign mint price, convert the proceeds at the given exchange rate into more domestic currency than they started with, and thus buy more gold than they had before. They will continue to engage in this sequence of transactions as long as gold's geographical value differential yields arbitrage profits.

At this point, Wheatley introduces transport costs into the analysis. Instead of abstracting from such costs as in the case of goods, he argues that these costs are too substantial to neglect in the case of gold. In particular, he estimates that specie transport costs amount to at least 3 percent of gold's value. Consequently, arbitrageurs will find it profitable to export specie only when the exchange rate exceeds relative mint prices by more than enough to cover such costs. In symbols, $E > (1 + s)P_g/P_g^*$, where P_g and P_g^* are the domestic and foreign mint prices of gold and s is gold's shipping cost expressed as a percentage of the domestic mint price. Rewriting this condition as $EP_g^* - P_g > sP_g$ yields the profitability criterion for gold exports: such exports are profitable when the foreign-to-domestic gold price differential exceeds the cost of transit. It follows that "when the exchange becomes unfavourable to . . . the extent of five, ten, or fifteen percent, gold will find its way out." For arbitrageurs will reap "as a profit what remains after the charge of transit, which is . . . three percent" (1819, pp. 24–25).

Symmetrically, Wheatley argues that gold imports become profitable when the exchange rate falls below mint par by more than transit costs, or $E < P_g/P_g^*(1 + s)$. Then the home price of gold exceeds the foreign price by more than the cost of transit, or $P_g - EP_g^* > sEP_g^*$. Arbitrageurs, secure that their profits will not be eaten up by shipping costs, will import gold to realize the differential. In short, Wheatley specifies certain critical values of the exchange rate that trigger gold inflows and outflows. These values of course are the famous specie points of the trade literature.

2. ADJUSTMENT MECHANISMS

Armed with the foregoing concepts, Wheatley explains how an open economy adjusts to real and monetary shocks under alternative currency regimes. According to Wheatley, monetary shocks trigger specie redistributions under metallic and convertible currency regimes and exchange rate changes under inconvertible paper regimes. Adjustment ceases and equilibrium reigns when the common-currency prices of goods and gold are the same worldwide. By contrast, real shocks prompt changes in real incomes and expenditures but leave nominal exchange rates and specie flows unchanged. Adjustment ends when the income and spending changes re-equilibrate the real balance of payments.

In the following paragraphs, it will be useful to consider Wheatley's analysis of monetary shocks first. Such shocks take the form of exogenous increases in the domestic money stock. He examines the resulting responses in purely metallic, convertible currency, and inconvertible currency regimes. In all cases, money's influence is confined to nominal variables. No real variables change except for gold's relative price EP_g^*/P_g, which moves to its export point.

The Process of Adjustment to Monetary Shocks in a Purely Metallic Regime

Consider first a metallic or 100 percent reserve gold-standard regime. Let an exogenous increase in the quantity of specie attributable to "the fertility of the mines" occur in the home country. The monetary expansion produces an immediate, equiproportional rise in the domestic price level. With perfect wage-price flexibility, nominal wages rise in step with prices, leaving real wages and thus employment and output unchanged. Here is the first strict bullionist or Ricardian proposition, namely, that money is neutral in its effect on real activity.

Turning to the market for foreign exchange, Wheatley argues that agents there observe the rise in domestic prices and anticipate the compensating exchange rate depreciation. They incorporate those anticipations into the rate, which depreciates immediately. The resulting equiproportional rise in the price level and the exchange rate leaves the terms of trade or relative price of goods at home and abroad, EP^*/P, unchanged. With no terms-of-trade improvement to induce a trade deficit, none occurs.

Instead, adjustment occurs in the market for monetary gold. Given the fixed foreign and domestic mint prices of gold, P_g^* and P_g, it follows that the rise in the exchange rate raises gold's common-currency price abroad, EP_g^*, above its price at home, P_g, or $EP_g^* - P_g > 0$. Still, the exchange rate may not depreciate enough to raise the gold price differential $EP_g^* - P_g$ above the cost of transit sP_g. If so, nothing further happens.

But let the exchange rate rise above mint par by more than the cost of transporting gold and arbitrage becomes profitable. Agents then have an incentive to ship specie abroad to realize the gold price differential. By contracting the domestic money stock and deflating the domestic price level, the resulting loss of specie will restore the exchange rate to mint parity, thus putting a stop to further gold drains. In this way, the efflux of gold will have restored the natural distribution of specie required for national (and world) monetary equilibrium. The upshot in Wheatley's small-open-economy case is that (1) P and E are the same as before, (2) the specie increment is dispersed abroad where it is too small to affect the world money stock and world price level, and (3) P^* accordingly remains unchanged. Wheatley (1807) summarizes: "In every instance, therefore, where a relative excess of currency caused the same sum to measure a less value in one country than it measured in others, the course of exchange would become unfavourable, and by leading to the departure and general distribution of the surplus specie, maintain inviolable the level of money" (pp. 66–67).

Wheatley Versus Hume

Here was a new theory of the adjustment mechanism. It differed from David Hume's celebrated account of the price-specie-flow mechanism enunciated in his 1752 essays "Of Money" and "Of the Balance of Trade." In his first essay, Hume maintains that increases in the stock of metallic money temporarily stimulate real activity before raising prices proportionally. Wheatley's model permits no such temporary nonneutrality. And in the second essay, Hume implies that the rise in domestic prices produces no fully offsetting rise in the exchange rate. The result is an increase in the relative price of home to foreign goods, which, by rendering imports cheap and exports dear, precipitates a trade deficit. Wheatley's model allows no such terms-of-trade or trade-balance effects.

Moreover, since the exchange rate in Hume's model is fixed (or at least remains below the export point), it cannot depreciate sufficiently to produce the gold price differential that triggers specie arbitrage. True, gold moves abroad in his model to equalize commodity prices. But it does so passively just to finance the trade deficit rather than actively in search of a higher price abroad.

Thus, while Hume's model achieves the same equilibrium distribution of specie as Wheatley's model, it does so through a different process. In short, Hume's account features exchange rate fixity, terms-of-trade variation, trade deficits, continuous gold price parity, and no gold arbitrage. By contrast, Wheatley's account stresses exchange rate variation, continuous purchasing power parity, no trade deficits, and temporary gold price differentials that activate specie arbitrage. Evidently there is more than one classical theory of the adjustment mechanism.

Adjustment to Monetary Shocks in Convertible Currency Regimes

Wheatley notes that a process similar to that described above works to eliminate excess money supplies in convertible currency regimes wherein paper is freely convertible into gold at a fixed price upon demand. Now, however, correction is achieved through retirement of excess notes as well as through specie drains.

Let an exogenous increase in the "publication of state and bank paper" occur in the home country (1807, p. 60). Money, prices, and the exchange rate rise equiproportionally maintaining relative prices EP^*/P and the trade balance unchanged. When the exchange rate depreciates to the point where gold export becomes profitable, arbitrageurs present paper notes to banks for conversion into gold at the official mint price. The resulting specie drain obliges banks to contract the note issue to protect their gold reserves. Such contraction causes the price deflation that restores the exchange rate to mint par and puts a stop to specie drains.[5] Indeed, Wheatley contends that central bankers' knowledge of this process disciplined them to contract as soon as the exchange rate signaled overissue.[6]

Wheatley also pioneered the distinction between small and large open-economy models. He was the first to point out that a small economy's excess issue, being a negligible fraction of the world money stock, could have no perceptible influence on that stock or world prices. By contrast, a large economy's excess issue could affect both. The large economy would, after working off its excess balances, retain its relative share of the enhanced world money stock and its price level would be higher than before. By contrast, adjustment would leave the small economy's price level unchanged.

In this connection, he also observes that all nations expanding in concert could do what no single nation could do alone, namely, generate an unlimited rise in money and prices. "Paper," he writes, "might be increased in any given country to any extent, provided that the currency of other nations were augmented in a similar ratio to preserve the equivalency" (1807, p. 28). For if all countries expanded in step, none would be conscious of an excess of currency. Although money and prices would rise in each country, there would be no exchange rate depreciation, no rise in the market price of gold, no drain of gold reserves from one country to another to limit expansion. Each country's paper, no matter how greatly augmented, would retain its value relative to gold and to other currencies. Wheatley was the first economist to enunciate this point.

[5] "When the over-issue of paper made the prices of this country higher than the prices of others . . . and the course of exchange marked the difference, bullion, which is foreign money . . . sold at a correspondent premium . . . together with the charges of transit. . . . Bank paper, therefore, was pressed upon the Bank to be exchanged for guineas, that the guineas might be converted into bullion, to be . . . sent abroad. By this process, Bank paper was reduced in amount . . . and . . . its contraction lowered our prices to a level with prices of other countries . . . and restored the exchange to par" (1819, pp. 39–40).

[6] "The unfavourable exchange, which naturally resulted from a partial redundance, constituted the exclusive check to the issues of the Bank of England throughout the whole of the preceding century, and prevented the publication of a greater sum than that proportion, which was adequate to circulate the produce of this country at par with the produce of others" (1807, pp. 68–69).

Adjustment to Monetary Shocks Under Inconvertible Currency Regimes

Wheatley was also among the earliest to analyze the operation of an inconvertible paper currency, which he saw as introducing a new twist to his model. Under inconvertibility, money cannot leak out into foreign trade. An excess issue cannot be worked off through specie drains as it can under metallic and convertible currency regimes. Instead, the exchange rate eliminates the redundant currency by devaluing it in proportion to its excess.

As before, Wheatley begins his analysis by introducing an exogenous monetary disturbance into his model. He assumes the banking system, desiring to reduce its reserve ratio, injects additional paper notes into an economy initially in monetary equilibrium with the exchange rate at mint par. He then traces out the ensuing sequence of events.

As in the convertible currency case, the overissue of inconvertible paper generates immediate and proportional rises in the prices of goods and foreign exchange. Together, these increases operate to maintain the real terms of trade EP^*/P unchanged, thus forestalling trade deficits.

At the same time, the depreciating exchange rate raises gold's relative price abroad EP_g^*/P_g. But as long as that increase does not exceed the cost of transit, gold remains a non-traded good that sells domestically at its unchanged market (and mint) price.

Let the exchange rate rise by more than the cost of transit, however, and gold becomes a traded good and therefore subject to the law of one price. As a result, market and mint price diverge. No longer tied to the mint price by convertibility, gold now fetches a price equal to its common-currency price abroad, EP_g^*, minus the cost of transit.[7] And since that common-currency price varies one for one with the exchange rate, it follows that gold commands a premium over its old mint price equal to the percentage rate of depreciation of the exchange rate. As this percentage rate is also the rate of expansion of the money stock, Wheatley arrives at the Ricardian or strict bullionist proposition that the percentage premium on gold constitutes proof and measure of overissue under inconvertibility.

The gold premium also prohibits specie exports. For the same law-of-one-price condition that equalizes gold's value worldwide precludes arbitrageurs from making profits on shipping the metal. Wheatley's (1807, p. 70) exposition of this point is both seminal and definitive. Under inconvertibility, arbitrageurs cannot obtain gold from the central bank. If some coin remains in circulation, however, they can obtain it from domestic coin holders. To induce the latter to part with their gold, arbitrageurs must pay the asking price. But coin holders themselves have the option of shipping their gold abroad, selling it at the

[7] In Wheatley's words, specie annexes a premium and "resiliates to a level with its value abroad" (1807, p. 367).

foreign mint price, and converting the proceeds net of transit cost into domestic currency at the rate of exchange. Consequently, arbitrageurs must pay a per-ounce price of $P_g = EP_g^* - sP_g$ for domestic gold, to which must be added the cost of shipping it sP_g. But this sum $EP_g^* - sP_g + sP_g$ leaves no arbitrage profits to induce gold exports. For it just equals EP_g^*, exactly what gold fetches abroad.

Since no profits can be made by exporting gold, none is exported. Instead, exchange rate changes rather than gold drains eliminate excess money stocks. Exchange depreciation devalues money in proportion to its excess. Such devaluation keeps demand-adjusted money stocks everywhere the same, as international monetary equilibrium requires. Let k^*M^* be the foreign demand-adjusted money stock and kM/E its domestic counterpart measured in terms of a common currency. Then world monetary equilibrium requires that $k^*M^* = kM/E$. Any excess of kM will be offset by compensating rises in E to maintain the equality. In Wheatley's words, "the course of exchange has no other means" of working off an excess supply of inconvertible currency "than to reduce it to a discount in proportion to its excess" (1807, p. 69).

In short, Wheatley argues that exchange rates bear the full burden of adjustment under inconvertibility. Specie movements do not occur. To explain why specie does not move, he appeals to the law of one price. He also appeals to the idea of comparative cost. He argues that inconvertibility renders gold just another commodity whose price during inflation rises identically with all commodity prices. But identical rises in the prices of goods and gold imply that gold cannot be cheap in terms of goods. And not being the relatively cheap commodity, it cannot qualify for exportation on comparative cost grounds. Therefore it is not exported. For that reason, specie does not move when inconvertibility reigns. Instead, it leaves adjustment to the exchange rate.

3. ADJUSTMENT TO REAL SHOCKS

Having argued that exchange depreciation, gold price premia, and specie drains constitute proof and measure of overissue, Wheatley had to show that those same phenomena could not also arise from real shocks operating through the balance of payments. For if his monetary variables registered real disturbances as well as monetary overissue, they could hardly be unambiguous indicators of the latter alone. Candidate real shocks included domestic crop failures, subsidies and loans to Britain's allies in the war against Napoleon, and the expenses of maintaining troops on the continent. He had to show that these disturbances propagated their effects through non-monetary channels and could not affect his monetary variables in the least.

To do so, he posits a demand-shift, income-expenditure mechanism. In the case of domestic crop failures, he sees adjustment occurring through shifts in reciprocal demands. Jacob Viner explains. Wheatley, he says, insisted that "the demand of England and the rest of the world for each other's product would necessarily so immediately and completely adjust themselves . . . as to result under both a metallic and an inconvertible paper standard in the maintenance of equilibrium in the balance of payments without the aid of specie movements, changes in the relative level of prices in the two areas, or movements of the exchange rate" (Viner 1937, p. 142).

Let a home harvest failure depress domestic income. Imports, a function of income, therefore fall. The resulting decline in the foreign country's export sales induces it to cut back its purchases from the home country. Home exports consequently fall to match home imports. The trade balance remains unchanged, as do the exchange rate and the ratio of national price levels—provided, of course, that the central bank eradicates that portion of the money stock rendered redundant by the fall in income. In terms of Marshallian reciprocal demand schedules or offer curves, the curves of both nations shift inward by equal amounts to intersect the unchanged terms-of-trade vector at a smaller volume of trade. Exchange rate movements and specie flows are not required.

Wheatley uses the same demand-shift, income-adjustment mechanism to resolve the transfer problem. He argues that foreign remittances—loans and subsidies to Britain's allies plus military expenditures abroad—are effected by a transfer of goods without disturbing price levels, exchange rates, or the distribution of specie. Causation runs from remittances to incomes to import demands to the export surplus that transfers the goods. The home government, say, taxes domestic citizens and gives the proceeds to the foreign country as a subsidy. The subsidy reduces home income and raises foreign income by equal amounts. Imports as a function of income fall in the home country and rise in the foreign country. The result is a home-country export surplus that, if the propensities to import in the two countries just add up to one as Wheatley assumes, precisely equals the amount of the subsidy.[8] Here is Wheatley's special case in which income shifts accomplish the goods transfer with no help from

[8] By definition, the home country's real trade balance B is the difference between its real exports X and its real imports I, or $B = X - I$. Also by definition, home exports X are the foreign country's imports, I^*, so that the trade balance may be expressed as the difference between foreign and domestic imports, each a function of real national income, or $B = I^*(Y^*) - I(Y)$. Differentiating the trade balance with respect to the subsidy T yields $dB/dT = (dI^*/dY^*)(dY^*/dT) - (dI/dY)(dY/dT)$. Since Wheatley assumes the recipient's and the payer's incomes increase and decrease, respectively, by the exact amount of the subsidy—that is, $dY^*/dT = -dY/dT = 1$—the expression simplifies to $dB/dT = (dI^*/dY^*) + (dI/dY)$, where the right-hand side is the sum of the marginal propensities to import. If this sum is one, as Wheatley assumes, then $dB/dT = 1$, or $dB = dT$, and the trade balance moves into surplus by exactly the amount of the subsidy. In the final analysis, the subsidy is paid in goods. Hence he concludes that "the superiority of our exports above imports must nearly correspond with the amount of our foreign expenditure" (1807, p. 219).

monetary variables or the terms of trade.[9] Here too is his formulation of the Ricardian or strict bullionist doctrine that monetary phenomena are invariant in response to real shocks to the balance of payments.

Wheatley's income-shift theory differed from the dominant gold-flow, price-adjustment theory of his contemporaries (see Fetter [1968], pp. 65–69). They held that real transfers are accomplished through price changes and specie flows prompted by the initial financing of the transfer. Such financing requires the paying country to obtain the recipient country's currency to make the cash payment. The resulting increased demand for foreign exchange bids up the exchange rate. Given national price levels, the rising exchange rate lowers the relative price of goods P/EP^* in the paying country, thus spurring its exports and checking its imports. Net exports get an extra boost when the exchange rate reaches its specie point and the resulting gold drain and monetary contraction deflate the paying country's price level. With deflation and depreciation lowering relative prices, the export surplus expands to effect the transfer in goods.

The view that transfers operate through monetary variables was, of course, anathema to Wheatley, who omits such channels from his model. Transfers were real phenomena. As such, they were entirely independent of monetary phenomena. To dramatize this independence, Wheatley argued that even as the paying country was making massive unilateral transfers abroad, it could still enforce specie inflows to any extent simply by contracting paper issues and deflating prices until the exchange rate fell to its import point.[10] Foreign payments, in other words, had nothing to do with exchange rates and specie flows.

4. POLICY IMPLICATIONS OF WHEATLEY'S WORK

Although Wheatley's analysis was primarily theoretical, it had some practical policy implications.[11] First, exchange rates, gold prices, and specie movements offer infallible indicators of overissue. When they signal monetary excess, it must be occurring since they respond to nothing else. Their invariance to real disturbances means that such disturbances cannot distort their signal and render it ambiguous.

A second implication is that persistent inflation is less likely to occur in convertible than in inconvertible currency regimes. Under convertibility, inflation is self-correcting. It automatically precipitates gold drains and forces

[9] "Foreign payments . . . have no effect to alter the state of our currency, they have no effect to alter the state of the exchange" (1819, p. 29).

[10] "A due compression of our paper circulation would have led to its [gold's] influx at the very moment, that the loan was in payment, and would have glutted the country with specie" (1807, p. 193).

[11] On Wheatley's policy views, see Fetter (1942), pp. 368–74.

banks to contract their note issue to protect their gold reserves. The resulting shrinkage of the money stock ends the inflation. No such corrective drains occur under inconvertibility when gold, if available at all, commands a price that renders specie exports unprofitable. It follows that convertibility offers the stronger safeguard to overissue.

A third implication of Wheatley's work is that price-level stability can be achieved by monetary means. A staunch advocate of such stability, Wheatley stressed the evils of price fluctuations. They arbitrarily redistributed income and wealth among the social classes and provoked social discontent. Avoiding such evils meant removing their monetary causes. To this end, Wheatley advocated (1) ending the suspension of specie payments and restoring convertibility of the British pound, (2) eliminating small notes which he saw as the most unstable component of the money supply, and (3) removing the note-issuing privilege from competing private banks and lodging it with the Bank of England.[12] These reforms, he thought, would prevent or minimize sharp variations in the money stock that constituted the primary obstacle to price stability.

A fourth implication of Wheatley's work is that indexation can immunize real payments from unanticipated movements in nominal ones. In this connection, he proposed price-level indexation of long-term contracts to compensate for fluctuations in the value of money. "Some criterion," he said, "should be assumed for the purpose of providing a graduated scale of the value of money" so that nominal incomes could be adjusted "in conformity to the result." Of the candidate criteria, a general price-index series such as that constructed by Sir George Shuckburgh Evelyn would be "the least objectionable" (1807, pp. 328–29). Earlier writers had advocated stabilizing real incomes by adjusting nominal incomes according to changes in the price of a single commodity such as rye or corn. But Wheatley was the first to recommend a general price-level index number for that purpose. The modern notion of indexation originates with him.

5. WHEATLEY'S PLACE IN CLASSICAL MONETARY THOUGHT

Even at its best, Wheatley's *Essay on the Theory of Money* could hardly match the subtlety and insight of Henry Thornton's *Paper Credit of Great Britain*. Nor could it match the power, brilliance, and lucidity of David Ricardo's *High Price of Bullion.* Still, if originality is any criterion, Wheatley's name belongs in the

[12] Unlike other bullionists, Wheatley was unwilling to exonerate private banks from overissue. By perversely varying their reserve ratios, such banks overissued notes independently of the central bank. He denied that the Bank of England controlled such banks through their reserves or that overissue was prevented by the operation of an interregional price-specie-flow mechanism.

front rank of classical monetary theorists. He formulated the strict bullionist model, which dichotomizes real and monetary sectors and posits neutrality and exogeneity in the short run as well as the long. True, this model looks primitive compared to Thornton's sophisticated schema. But its ultra-simplicity entails certain positive strengths. The model yields clear-cut policy conclusions. And it avoids confusion between real and nominal variables. It emphasizes money's permanent price effects but ignores transitory output and employment effects that might distract the central bank from pursuing its primary goal of price stability. Ricardo, for one, found these properties desirable. He employed a version of the strict bullionist model after Wheatley first presented it.

Nor should Wheatley's other contributions go overlooked. He established, three years before Ricardo, the theoretical underpinnings of the Ricardian definition of excess (see O'Brien [1975], p. 148). This definition says that if (1) the exchange rate is depreciated, (2) gold is selling at a premium, and (3) specie (under convertibility) is leaving the country, then the currency is by definition excessive and must be contracted. Here was the tool strict bullionists needed. With it they could counter antibullionists' and moderate bullionists' claims that such phenomena might well originate in real shocks so that monetary contraction was not required.

Beyond these ideas were his specific contributions to international monetary theory. He presented the clearest and most complete statement of the purchasing power parity doctrine before Gustav Cassel. He was the first to use the rational expectations argument to explain why the terms of trade is always in equilibrium. He originated the distinction between large- and small-open-economy models in which a large country's note issue perceptibly influences world money and prices whereas a small country's issue does not. Likewise, he introduced the notion that all nations expanding in step in a convertible currency regime can do what no single nation can do alone, namely, generate an unlimited rise in money and prices. And his demand-shift, income-expenditure theory of unilateral transfers anticipated the subsequent contributions of Mountifort Longfield, J. E. Cairnes, C. F. Bastable, J. S. Nicholson, and Bertil Ohlin.

Perhaps his most outstanding contribution, however, was his specification of the link between exchange rates, gold prices, and gold flows. In an explanation superior to any before Ricardo, he showed that in metallic and convertible currency regimes gold is not simply a means of discharging international payments. Rather it is a commodity that flows across nations to capture arbitrage profits created when exchange rate movements generate gold price differentials.

A related contribution was his use of the law of one price to show that gold ceases to move across nations when its common-currency value is equalized worldwide such that no arbitrage profit can be realized by shipping it. He showed that money's purchasing power parity over gold, not its purchasing power parity over goods, is what halts gold movements. These contributions, together with his indexation proposal, were the products of an original mind. They identify Wheatley as a creative scientific economist who deserves a prominent rank in the classical pantheon. Above all, his ideas identify him as the most monetarist, or strictly quantity theoretic, of all the classical writers.

Wheatley took three key ideas—(1) the quantity theory of money, (2) the purchasing power parity doctrine, and (3) the notion of gold arbitrage—and endowed them with sharp analytical content. He then combined these concepts into a powerful framework capable of tracing the effects of monetary disturbances produced by England's suspension of convertibility during the Napoleonic Wars. To analyze the effects of real disturbances produced by the wars, he developed a separate demand-shift, income-expenditure mechanism. His achievements merit recognition from economists today.

REFERENCES

Chipman, John S. "Balance of Payments Theory from Locke to Ricardo." Manuscript. University of Minnesota, 1983.

Einzig, Paul. *The History of Foreign Exchange*. London: Macmillan, 1962.

Fetter, Frank W. "The Transfer Problem: Formal Elegance and Historical Realism," in C. R. Whittlesey and J. S. Wilson, eds., *Money and Banking in Honour of R. S. Sayers*. London: Oxford University Press, 1968.

__________. "The Life and Writings of John Wheatley," *Journal of Political Economy*, vol. 50 (June 1942), pp. 357–76.

Frenkel, Jacob A. "Purchasing Power Parity; Doctrinal Perspective and Evidence from the 1920s," *Journal of International Economics*, vol. 8 (1978), pp. 169–191.

Hollander, J. H. "The Development of the Theory of Money from Adam Smith to David Ricardo," *Quarterly Journal of Economics*, vol. 25 (1911), pp. 429–70.

Hume, David. "On Money" and "On the Balance of Trade," in E. Rotwin, ed., *Writings on Economics*. Madison: University of Wisconsin Press, 1955.

Kydland, Finn, and Edward C. Prescott. "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50 (November 1982), pp. 1345–70.

Long, J. B., and C. I. Plosser. "Real Business Cycles," *Journal of Political Economy*, vol. 91 (February 1983), pp. 39–69.

Metzler, Lloyd A. "The Theory of International Trade," in *A Survey of Contemporary Economics*, Vol. 1. Philadelphia: The Blakiston Company, 1948.

Morgan, E. V. *The Theory and Practice of Central Banking, 1797–1913*. London: Cambridge University Press, 1943.

O'Brien, D. P. *The Classical Economists*. London: Oxford University Press, 1975.

Officer, Lawrence. *Purchasing Power Parity and Exchange Rates.* Greenwich, Conn.: JAI Press, 1984.

Ricardo, David. *The High Price of Bullion: A Proof of the Depreciation of Bank Notes*. London: Murray, 1810, reprinted in Piero Sraffa, ed., *The Works and Correspondence of David Ricardo,* Vol. 3, *Pamphlets and Papers, 1809–1811*. Cambridge: Cambridge University Press (for Royal Economic Society), 1951.

Schumpeter, J. A. *History of Economic Analysis*. New York: Oxford University Press, 1954.

Thornton, Henry. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*. 1802. Edited with an introduction by F. A. von Hayek. London: Allen and Unwin, 1939.

Viner, Jacob. *Studies in the Theory of International Trade*. New York: Harper, 1937.

Wheatley, John. *Report on the Reports of the Bank Committees*. Shrewsbury, England: W. Eddowes, 1819.

__________. *An Essay on the Theory of Money and Principles of Commerce,* Vol. 1. London: Cadell and Davies, 1807.

__________. *Remarks on Currency and Commerce.* London: Cadell and Davies, 1803.

Wu, C-Y. *An Outline of International Price Theories*. London: Routledge, 1939.