Vol. 25, No. 2

# ECONOMIC REVIEW

1989 Quarter 2


**Removing the Hazard of Fedwire Daylight Overdrafts** 2
by E.J. Stevens

**Capital Subsidies and the Infrastructure Crisis: Evidence from the Local Mass-Transit Industry** 11
by Brian A. Cromwell

**Employment Distortions Under Sticky Wages and Monetary Policies to Minimize Them** 22
by James G. Hoehn


FEDERAL RESERVE BANK
OF CLEVELAND

ECONOMIC REVIEW

1989 Quarter 2
Vol. 25, No. 2

**Removing the Hazard of Fedwire Daylight Overdrafts** 2
by E.J. Stevens

Free Federal Reserve daylight overdrafts misallocate resources. One reason is the moral hazard of fully insuring a paying bank's access to whatever volume of daylight overdraft credit it needs. This paper contrasts the effects of three recent proposals for pricing daylight overdrafts and demonstrates that reducing moral hazard depends on how, rather than on how much, pricing affects daylight overdrafts. If payment practices and modes of bank financing were unresponsive to pricing, it would suggest that the moral hazard of Federal Reserve daylight overdrafts has been an insidious force behind the rapid growth of interbank lending and securities-market trading in recent decades.

**Capital Subsidies and the Infrastructure Crisis: Evidence from the Local Mass-Transit Industry** 11
by Brian A. Cromwell

Public investment and maintenance decisions are potentially distorted by budget procedures, political pressures, and capital subsidies. Empirical evidence from two recent studies of the mass-transit industry is summarized and suggests that federal capital subsidies have important effects on infrastructure decisions of local governments.

**Employment Distortions Under Sticky Wages and Monetary Policies to Minimize Them** 22
by James G. Hoehn

Sticky nominal wages can result in distortion of employment levels as demand for goods and labor productivity change. This article shows that employment distortions can be minimized by a monetary policy that allows some price deflation when productivity improves. Policies that target nominal income or the price level result in smaller distortions than do policies that target output or money.

*Economic Review* is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. Copies of the *Review* are available through our Public Affairs and Bank Relations Department, 216/579-2157.

Coordinating Economist:
Randall W. Eberts

Editors: William G. Murmann
Robin Ratliff
Design: Michael Galka
Typesetting: Liz Hanna



ISSN 0013-0281

# Removing the Hazard of Fedwire Daylight Overdrafts

by E.J. Stevens

E.J. Stevens is an assistant vice president and economist at the Federal Reserve Bank of Cleveland. The author thanks John Carlson, Charles Carlstrom, Randall Eberts, and William Gavin for their useful comments on earlier drafts of this paper.

## Introduction

The 12 Federal Reserve District Banks extend about $115 billion of credit within a few hours on an average business day, only to take it back again before the close of business. This huge sum reflects banks' daylight overdrafts of their deposit accounts at Federal Reserve Banks when making large-dollar-value payments to other banks using Federal Reserve wire transfer systems.[1] If all goes well, subsequent receipts from other banks extinguish the daylight overdrafts before the end of the day.

Daylight overdrafts via Fedwire are not allocated by any market process and are free, a result of the order in which a bank's payments and receipts occur. The same might seem to be true of checks presented and deposits made to any checking account during a day, but there is a crucial distinction: a Fedwire payment is irrevocable upon receipt, while a check is only a provisional payment. Therefore, the Federal Reserve is the party at risk if a daylight overdraft is not repaid by the end of a day.

Free daylight overdrafts are costly. Of course, the Federal Reserve faces no financing or resource costs in issuing daylight credit because it has the power to create money; failure of a bank to eliminate its daylight overdraft by the end of a day would simply add to Federal Reserve assets (claims on a bank) and liabilities (bank reserve deposits).[2] The costs arise from resource misallocations.

One source of these inefficiencies, and the focus of this paper, is the "moral hazard" involved in providing free daylight overdrafts.[3] Fedwire fully insures a payor bank's access to whatever volume of daylight overdraft credit it needs to make payments that are immediately available

■ 1 These systems include Fedwire, for transfer of reserve balances from one bank to another, and the securities wire, for transfer of book-entry U.S. government securities from one bank to another in return for reserve balances. The term Fedwire will be used here to refer to both systems. A third system, CHIPS (Clearing House Interbank Payment System), is operated by the private New York Clearing House Association; credit extended among participants in this system adds another $45 billion of interbank daylight credit on an average day.

■ 2 Failure to repay might result from a bank's insolvency, perhaps impairing the value of the asset, causing a charge against Federal Reserve income that would reduce Treasury receipts.

■ 3 Stevens (1988) provides a discussion of the probable nature of some resource misallocations resulting from this moral hazard.

and irrevocable. The result is a form of insurance that removes any incentive for payee banks to monitor or manage credit risk in receiving payments that payor banks fund with daylight credit.

Suggestions have been made to price Fedwire daylight overdrafts in an effort to control them. Market sources of funding would replace some or all Fedwire daylight overdrafts in making payments and would require compensation based on credit risk. Market discipline would then provide the now-missing incentive for payor banks to attend to risk, thereby avoiding moral hazard.

This paper suggests that economizing need not bring about the market discipline that would eliminate moral hazard. The first section provides a brief review of Fedwire daylight overdraft history, Federal Reserve payment system risk policy, and the problem of moral hazard. The second part shows how differences among three recently proposed daylight overdraft pricing mechanisms can influence the extent of daylight overdraft reduction and, more important, the way in which banks reduce daylight overdrafts. The final part argues that reducing Fedwire moral hazard does not depend on *how much,* but on *how* banks reduce daylight overdrafts, and that this should be a criterion for choosing among pricing proposals.

## I. Fedwire Daylight Overdrafts and Moral Hazard

**A** bank goes into daylight overdraft when it has made more payments from its account at a Federal Reserve Bank by some point during a day than can be covered by its opening reserve-deposit balance plus payments received by that point in the day. A common example is that of a bank dependent on continuous overnight federal-funds borrowing. Operational convenience leads it to return the borrowed funds each morning, before borrowing replacement funds in the afternoon. The midday period is spent in overdraft, funded by the Federal Reserve.

As recently as 30 years ago, the U.S. large-dollar-value payments system was for the most part a cash-in-advance system. Irrevocable Fedwire payments were riskless both to payees and to the Federal Reserve because they were drawn against positive balances. Since then, Federal Reserve daylight risk exposure has mushroomed, associated with the telecommunications revolution in the payments mechanism, the proliferation of new financial instruments, and the explosion of trading volumes in worldwide money and capital markets.

A simple comparison illustrates the extent of the change. In 1947, reserve-deposit balances represented 700 percent of (seven times) the value of daily debits (Fedwire, checks, etc.) to member-bank reserve accounts. That is, the average bank could make all of its own and its customers' payments for seven successive business days without ever receiving a single offsetting payment, and without exhausting its initial reserve-deposit balance. By 1983, balances were a minuscule 4 percent of daily debits. The average bank could meet demands for payment for only 20 minutes of a single eight-hour business day before it would have to receive some offsetting payments, or go into overdraft.[4]

Initially, the evolution from a cash-in-advance system toward automatic daylight credit seems to have gone undetected, but confronting the growing daylight credit risk problem became unavoidable in the late 1970s under the pressures of technological change and a demand for same-day net settlement service by potentially competing private large-dollar-value payment networks. Originally, starting in 1918, telegraph, telephone, or mail messages to the Federal Reserve were the only mechanisms for transferring ownership of reserve-deposit balances between banks with same-day finality. Related devices were official checks, offering only next-day finality, and interbank messages that simply instructed a bank to use Fedwire to transfer funds.

Introducing computer-to-computer telecommunications technology for payments by Fedwire and by the Clearing House Interbank Payment System (CHIPS), and for interbank message systems, suggested a new possibility in the 1970s. Private payment networks like CHIPS and the then-proposed CashWire network each would be capable of clearing payment messages among its own participants continuously during the day before presenting a single balanced set of net debit and credit positions to the Fed in time to achieve same-day final settlement.

Compared to the next-day systems prevalent then, this would offer the advantage of reducing costly overnight float financing of banks in net debit position by those in net credit position. In addition, it would shorten the length of time during which overnight float exposed banks to credit risk. Operating details of telecommunication devices, accounting-system modifications, backup facilities, and daily time schedules were laid out quickly, but the enterprise foundered on

■ **4** Reduced reserve requirements represent only a small portion of this change. To have maintained the 1947 reserve deposits/debits ratio with the 1983 volume of debits would have involved reserve deposits equal to an impossible two-and-a-third times the total assets of all commercial banks.

the "unpostable debit"—what to do if one of the participants had insufficient funds in its reserve account to cover its private network net debit at settlement hour.

Some found the unpostable debit an operational inconvenience to be ignored: from an operations perspective, it was no problem as long as the accounting system accepted negative numbers. After all, a Federal Reserve Bank did not check to see whether a bank had sufficient funds to cover a Fedwire transfer. Why should a net settlement message be treated any differently? Others found it troubling to design a system in which the central bank automatically would guarantee a private network settlement by accepting an unpostable debit as an offset to irrevocable credits. That issue is not fully resolved even today, but two developments did force some action with respect to daylight overdrafts.[5]

One development was the increasing incidence of *overnight* overdrafts of reserve accounts and adoption of the current Federal Reserve overnight overdraft policy.[6] High interest rates, escalating wire-transfer traffic, and declining reserve requirements were making reserve-deposit accounts a less and less effective buffer stock in banks' daily reserve-balance management. With no formal overnight overdraft policy other than Regulation D (that banks maintain an average required balance over a one- or two-week reserve maintenance period), concern was mounting that banks might abuse the Federal Reserve by running overnight overdrafts when especially profitable opportunities arose.

Developing an overnight overdraft policy led to more widespread realization within the Federal Reserve that daylight overdrafts were a fact of life. Not only was there no mechanism in place to prevent daylight overdrafts, but neither was there a way to know how widespread the practice was. The second development was a carefully constructed survey of the incidence of daylight overdrafts. This provided the factual foundation for debating and developing the initial Federal Reserve payment system risk (PSR) policy: guidelines for determining limits on daylight overdraft positions; continued recording of daylight overdraft positions (in addition to a real-time mechanism to control daylight overdrafts at problem banks and special institutions); and a stated intention to ratchet-down limits over time. Pricing daylight overdrafts now is being suggested as a next step for this policy.

The problem with free Fedwire daylight overdrafts is moral hazard. The term refers to the hazard an insurer faces as a result of the elimination of incentives for an insured party to avoid a risk precisely because any losses arising from that risk are covered by insurance. Fire, life, and casualty insurers protect against moral hazard in a variety of ways. For example, coinsurance in the form of deductibles or copayments gives the insured a stake in preventing loss; inspection and requirements to remove risks give the insurer the ability to manage risk.

Fedwire does have some similar protections. The payor bank's net worth is at stake if it is unable to repay its credit, constituting a form of coinsurance. Regulation, supervision, and examination of banks guard against imprudent banking practices, now extended to include payment practices. However, initial limits on daylight overdraft exposure deliberately have been set high, and do not yet apply to overdrafts from book-entry securities transfers. As a result, Fedwire moral hazard is real, particularly in the short run between bank examinations.

Payee banks have no reason to limit payments received during a day, regardless of the volume of daylight overdrafts per dollar of net worth of the payor bank, because the Federal Reserve is at risk. Payor banks face no external disincentives that would raise the cost of daylight overdraft credit as the volume they use increases and as their credit quality falls. Federal Reserve protections against moral hazard are not yet very strong.

## II. Avoiding Daylight Overdrafts

### Adjustments

Any bank could eliminate daylight overdrafts by holding more overnight reserve deposit balances, by borrowing balances for a few moments or hours during the day, or by modifying its own or its customers' payment practices to prevent a negative balance. Such adjustments might be costly, of course, but would be worthwhile if they cost less per dollar than a daylight overdraft.

■ 5 The most recent effort to resolve the unpostable debit issue is that of the New York Clearing House Association, which has adopted a requirement that CHIPS members participate in a loss-sharing arrangement. It also has proposed federal legislation apparently intended to give legal priority to network payment claims over all others if a network member becomes insolvent. See *American Banker,* April 7, 1989, pp. 1 and 16.

■ 6 Overnight overdrafts are subject to a penalty of the larger of $50, or the larger of 10 percent or a rate 2 percentage points above the federal funds rate prevailing on the day the overdraft is incurred. The penalty charge is in addition to the cost of making up the reserve-deposit deficiency for reserve-requirement purposes.

A cost-minimizing bank might acquire excess reserves in the federal funds market. After meeting its temporary daylight need to cover payments, the bank would then have these extra funds available to hold, or to loan out overnight, if it could. The marginal cost of preventing a daylight overdraft would be the difference between the cost of borrowing and the return on lending.

A private daylight loan market does not now operate, but such a market would provide a second possibility for avoiding Federal Reserve daylight overdrafts.[7] Daylight loans could redistribute existing reserve balances from banks having them and not needing them during the day for payment purposes, but only overnight for reserve-requirement purposes, to banks not having them and needing them during the day, but not overnight. Free Federal Reserve daylight credit preempts such a market now, but if daylight overdrafts were to become costly, and timely delivery were assured, borrowing in a daylight loan market might become an inexpensive way for a bank to prevent overdrawing its reserve account during a day, with repayment before close of business.

Finally, a bank could alter the amounts of debits and credits to its account, or their sequence during the day. It might do this by lengthening the maturity of its liabilities, or by adopting a continuing contract for federal funds borrowing, with daily renegotiation of the rate but no daily repayment and re-receipt of funds. Or, pairs of institutional customers operating in securities markets might be induced to net their transactions obligations during a day, producing a single net obligation for daily payment, again reducing debits that might now precede credits. Or, groups of banks might join in private payment networks, substituting daylight credit on the private networks for Federal Reserve daylight overdrafts. Only net settlement of end-of-day positions would need to be accomplished through Federal Reserve accounts.[8]

Modifying payment practices in these ways would involve some costs, too, such as paying higher rates on longer-term liabilities, or receiving lower prices or revenues for payment services when institutional customers engage in obligation netting, or sharing the cost of a private payment network. Some tactics would be more expensive than others, so the marginal cost of preventing daylight overdrafts in reserve accounts by modifying payment practices would increase with the volume of overdrafts avoided.

In equilibrium, cost-minimizing banks would adopt the unique combination of adjustment mechanisms having marginal costs equal to or less than the marginal cost of a daylight overdraft. Pricing daylight overdrafts would lead banks to adjust from today's zero marginal cost to something higher.

## Three Proposals to Price Daylight Overdrafts

Three specific pricing proposals that have been receiving attention are evaluated in this section.[9] One would treat each daylight overdraft as an automatic overnight discount-window loan, booked at a penalty rate. A second would require a bank to hold additional balances at a Federal Reserve Bank in proportion to its daylight overdrafts. A third would simply impose a slight fee per dollar of daylight overdraft.

*Penalty Rate* The penalty rate proposal comes from Wayne Angell, member of the Board of Governors of the Federal Reserve System. A bank would be required to borrow the amount of any daylight overdraft as a collateralized loan from its Federal Reserve Bank discount window at an above-market penalty rate, but the Federal Reserve Bank would pay an explicit (below-market) rate of return on excess reserves.[10] The combination of the two features means that, under normal circumstances, no bank would run a daylight overdraft intentionally and pay the penalty discount rate, because the maximum alternative cost would be only the interest-rate spread between the cost of financing extra excess reserves, perhaps the federal funds rate, and the earnings rate on excess reserves.

The same spread would become the cost of borrowing daylight funds in the likely event that a private daylight loan market developed. Banks

■ **7** Simmons (1987) contains an extensive discussion of daylight funds market possibilities.

■ **8** Humphrey (1987) and Board of Governors of the Federal Reserve System, Large-Dollar Payments System Advisory Group (1988) contain detailed explanations of a number of these potential modifications of payment practices.

■ **9** These proposals are described in VanHoose (1988), the Angell proposal of a penalty rate; Hamdani and Wenninger (1988), supplemental balances; and Board of Governors of the Federal Reserve System, Large-Dollar Payments System Advisory Group (1988), fees.

■ **10** Penalty-rate borrowing would differ from an overnight overdraft in that a bank would be required to post eligible collateral for the loan associated with a daylight overdraft, but would not involve the cost of making up a reserve-deposit deficiency for reserve-requirement purposes.

would never pay more than this spread for a daylight loan because they could always borrow reserves in the federal funds market and lend at the overnight rate; lenders would never charge less than this spread because they could always sell their reserves at the federal funds rate, of course forgoing the rate earned on excess reserves.

Note, however, that excess reserves and a daylight loan market would be relevant only to the extent that daylight overdrafts were not eliminated by modifications in payment practices that were less costly than the rate spread.[11]

*Supplemental Balances* The supplemental balance proposal has been described by the staff of the Federal Reserve Bank of New York. A bank would be required to hold a special interest-bearing deposit (the supplemental balance) in a current period equal to some fraction (the supplemental balance ratio) of prior-period daylight overdrafts of its combined reserve and supplemental deposit accounts. The maximum cost of a dollar's daylight overdraft today would be the supplemental balance ratio multiplied by the expected next-period spread between the cost of financing a dollar's supplemental balance and the rate earned on the supplemental balance. With both this rate spread and the ratio administratively fixed, the maximum cost of a daylight overdraft would be a simple constant amount per dollar of daylight overdraft.

The cost would set an upper limit on the market rate for daylight loans. And, as in the penalty rate case, supplemental balances and daylight lending would emerge only to the extent that less-expensive modifications in payment practices failed to eliminate daylight overdrafts.

Banks would not use ordinary non-interest-bearing excess reserves to avoid daylight overdrafts, because the cost of financing them at the federal funds rate normally would be greater than the supplemental balance ratio times the rate spread. Unlike the penalty rate proposal, the supplemental balance approach would not necessarily eliminate all daylight overdrafts. Only at a very low earnings rate on supplemental balances (perhaps even a negative rate) would it be certain that banks would find payment-system modifications (or excess reserves) a cheaper way to avoid daylight overdrafts.

*Fees* The fee proposal has been suggested by the Federal Reserve System's Large-Dollar Payments System Advisory Group. It would simply have the Federal Reserve impose a fee for Fedwire overdrafts in excess of a base amount established for each bank. The maximum cost to a bank of a dollar's daylight overdraft would be that fee.

Extra excess reserves would not be used in this case unless the fee were set *higher* than the federal funds rate. A limited daylight loan market could develop, redistributing the required reserves of banks whose need for daylight balances was less than their need for required reserve balances. And, of course, neither daylight overdrafts nor daylight loans might be necessary if sufficient modifications in payment practices were forthcoming at a marginal cost less than the fee.

In brief summary, then, each of the three pricing proposals might be capable of eliminating Federal Reserve daylight overdrafts entirely through inexpensive modifications in payment practices. However, if modifying payment practices and redistributing required reserves through a daylight loan market were not sufficiently responsive to price, the outcome of pricing would differ substantially among the three proposals:

- The penalty rate regime would eliminate *all* remaining daylight overdrafts by expanded holdings of excess reserves and their redistribution in a daylight loan market.
- The supplemental balance regime would eliminate *some* of the remaining daylight overdrafts by expanded holdings of reserves in the form of supplemental balances and their redistribution in a daylight loan market.
- The fee regime would eliminate *none* of the remaining daylight overdrafts, unless the fee became a penalty rate.

## III. Pricing and Moral Hazard

Each of the three pricing proposals could reduce daylight overdrafts, but to what extent would they reduce moral hazard? None of the proposals would directly relate price to a bank's credit quality or to the volume of its daylight overdrafts. Nor would any of them introduce the kind of actuarial relation between price and risk exposure needed to establish an insurance fund.

■ **11** Note also that the penalty rate proposal contains the seeds of a problem for monetary policy. Extra demand for excess reserves would be matched, on average, by extra supply through open market operations, maintaining a policy-desired level of the federal funds rate, on average. However, the variability of the federal funds rate around the average rate might increase, reflecting variations in payment needs for balances within a day, or perhaps day-to-day, unrelated to reserve requirements and monetary growth. A bulge in payment needs that drove up the daylight loan rate during a day would drive up the federal funds rate by the same amount, because the overnight earnings rate on excess reserves is administratively fixed. No creditor would lend federal funds during the day for less than the sum of the daylight loan rate and the overnight rate. As long as policymakers value the federal funds rate as a tool or information variable, adopting the penalty rate proposal might involve some risk of less-precise policy implementation.

Reduced moral hazard would have to come as a by-product of pricing, in some form of enhanced market discipline. This could not be administered by payee banks on Fedwire, for they remain free of any risk in receiving payments. Results, therefore, could come only from the behavior of other creditors, or from eliminating payments requiring daylight funding. Investigating the adjustment mechanisms banks could use in response to pricing, however, reveals an uncertain basis for expecting market discipline to flourish.

### Excess Reserves

Both the penalty rate and the supplemental balance proposals could create a need to finance extra holdings of interest-bearing reserve balances. In both proposals, the earnings rate on those balances would be uniform across all banks, but the rate paid in the market to finance the extra balances might vary with the credit quality of a payor bank. If so, then the marginal cost of avoiding or funding a daylight overdraft would vary with the credit quality of the borrowing bank, injecting market discipline into payments.

Of course, moral hazard in the current deposit-insurance systems tends to dampen the role of credit quality in pricing both deposits and deposit insurance, and in pricing any kind of financing for a bank considered "too big to let fail." However, to the extent that a bank's marginal cost of funds can vary with credit quality, moral hazard would be diminished relative to the current arrangement of free daylight overdrafts.

### Daylight Loans

Similar assertions are made about the market discipline of a daylight loan market: if pricing induced banks needing daylight funds to borrow them from banks having surplus daylight funds, risk premiums would emerge in daylight loan rates, as market scrutiny sorted borrowers by credit quality.

This argument needs two qualifications.[12] One is that neither the supplemental balance nor, more especially, the fee proposal provides much basis for an extensive daylight loan market. Balances available for daylight lending would be limited to those of banks whose need for payment balances was less than their required, or required plus supplemental, reserve balances. This suggests only a limited stock of reserve deposits available for market allocation of daylight loans to replace free daylight overdrafts, at least relative to the penalty rate proposal.

The second qualification recognizes the too-easy presumption that daylight lenders actually would be at risk. The presumption rests on an apparent analogy between unsecured overnight interbank loans in the federal funds market and the envisioned unsecured intraday interbank loans in a daylight loan market. Whatever the similarity between overnight and intraday lending, it does not extend readily to risk of loss.

Federal funds loans are risky even though their dominant maturity is only one day. While deposit insurance and the "too big to let fail" maxim may minimize risk, it is still possible for a bank to be closed, resulting in at least a delay in repayment, if not partial or complete loss of interest and principal to its federal funds market creditors. Even with assurance that a loan is for only one day, banks routinely impose limits on their lending to individual banks as a matter of credit policy, and risk premiums sometimes are required.

Daylight loans would seem to be much closer to a riskless opportunity. Under what circumstances could a borrower fail to repay? One is if regulatory authorities closed the bank *during* a day, rather than following the precedent of closing banks only after close of business.

Closing a bank in the midst of a day's business would seem exceedingly awkward in a financial and legal environment where the timing of competing claims arriving by different means (over the counter, mail, messenger, telephone, day-ahead magnetic tape, off-line telecommunication, on-line telecommunication) is not readily distinguished. In fact, one by-product of pricing daylight overdrafts could be a standard timetable for posting each off-line activity to the daylight balance monitor, and use of that standard for defining priorities among claimants. Such a monitor could make intraday closings easier to arrange, but unless all of this were to become well established, authorities are not likely to close a bank during daylight hours.

Ruling out unexpected daylight closings means that all lending and borrowing banks would have access to Fedwire, and could make irrevocable repayment of daylight loans if they wished to do

■ **12** Another strand of thinking about daylight overdrafts would add a third qualification, also relevant to excess reserves: the "event risk" problem. Creditors might not have a way to assure themselves that the debtor would not borrow additional sums, an event raising the riskiness of their loans after-the-fact. If this were the case, early credit would be underpriced and risk premiums too low. This is a problem for any creditor, and gives rise to restrictive covenants in lending agreements. To be a serious qualification in the daylight loan case, however, would require a demonstration both that the second qualification does not hold, so that private lenders actually are at risk, and that covenants in standard daylight loan agreements combined with innovations in electronics network monitoring, such as already exist in CHIPS, could not deal with the problem. An elaborate treatment of the underpricing/overlending case can be found in Gelfand and Lindsey (1989).

so. Daylight loans could be riskless because, in the normal case, a bank *unexpectedly* in trouble would in no way be prevented from sending Fedwires to repay daylight loans, even though that were to result in a daylight overdraft.

It may seem ludicrous to imagine a bank borrowing in the daylight loan market in order to avoid a daylight overdraft, but then repaying the loan later the same day by going into daylight overdraft—except as part of a tactic calculated to trigger a discount-window loan or an overnight overdraft. Nonetheless, the point is made—that any bank on the ex post monitor could make irrevocable repayment of a daylight loan during banking hours *if it wanted to do so.* Daylight loans would carry the risk of nonpayment only if the borrowing bank preferred to default on the loan rather than overdraw its account at a Reserve Bank. Daylight loans are riskless unless there are good reasons to think that any unexpectedly insolvent bank would prefer default in the market to overdraft at the Federal Reserve and potential closing.

The inexpensive technology of ex post monitoring of daylight overdrafts is perfectly adequate for ex post booking of a penalty rate loan, or ex post calculating of a supplemental balance to be held in the future, or ex post billing of a simple fee. The difficulty with the technology is that it leaves unclear who is at risk, or perhaps makes only too clear who is *not* at risk, in interbank daylight lending. As long as interbank daylight lending is riskless, no market discipline emerges from it. The moral hazard of free Federal Reserve daylight overdrafts would remain the moral hazard of private daylight loans.

### Payment Practices

Modifying payment practices would be expected to reduce moral hazard. For example, as banks replace overnight federal funds with longer-maturity financing, their creditors would accept and demand compensation for additional risk. This risk formerly was accepted by the Federal Reserve, when daylight overdrafts provided an automatic means for an unexpectedly insolvent bank to close without having renewed its overnight credit.

A different example of risk shifting is that of netting the many payments of two customers into a single obligation. This would eliminate moral hazard because self-interest of the parties in the netting process would demand risk evaluation and compensation and would impose limits on any credit-risk exposure they might assume with respect to one another.

As a third example, pricing would encourage the migration of payments from Fedwire to private networks. Moral hazard would diminish as payments shifted to private systems because, with prerequisite credit limits and loss-sharing agreements in place among participants, banks would be expected to ration and/or price network credit on the basis of credit quality.

### How Much Good Would Pricing Do?

One thing certain is that none of the proposals would enlist the self-interest of payee banks directly in monitoring the credit quality of payor banks. As long as Fedwire provides irrevocable ownership of good funds upon receipt, payee banks do not extend credit in the Fedwire payment process, are not at risk, and have no incentive to monitor the credit quality of payor banks.

Market discipline would have to originate from other pressures on payor banks to manage payment risks. That said, the most crucial unknown factor is the rate at which the marginal cost of modifying payment practices rises as the volume of eliminated daylight overdrafts increases. If this marginal cost rises relatively slowly, so that inexpensive modifications effectively will eliminate all Fedwire daylight overdrafts, then moral hazard should disappear, supplanted by the market discipline of risk-sharing agreements in private payment networks, by netting agreements among banks' customers, and by the risk aversion of banks' creditors (and, perhaps in the future, of banks' insurers).

On the other hand, if this marginal cost rises relatively rapidly, the major burden of rationing daylight overdrafts would have to be borne through the direct mechanism of a pricing scheme. In this event, conjecture becomes somewhat more dependable — at least concerning the relative strengths of the three proposals.

The penalty rate proposal, while eliminating daylight overdrafts altogether, is not likely to be effective in removing moral hazard. Ex post daylight overdraft monitoring would leave the Federal Reserve bearing the credit risk of an active interbank daylight loan market, redistributing a much enlarged volume of excess reserves. High-quality banks could borrow excess reserves needed to avoid the penalty rate, not only for their own accounts, but also for riskless lending to lower-quality banks, with repayment assured by irrevocable Fedwire transfers.

The supplemental balance approach would more successfully tie the cost of daylight funding to perceptions of a bank's credit quality in the interday markets (via a risk spread paid for supplemental balances). This seems to be the most effective of the three pricing devices for injecting market discipline into the cost of funding payments.

The simple fee proposal offers little protection against moral hazard to the extent that changes in payment practices fail to eliminate daylight overdrafts. Flat-rate pricing of assured access to daylight credit may discourage its use, but provides no basis for scrutiny of the credit quality of payor banks, and no risk-based market disincentive for payor banks to limit daylight funding of payments.[13]

The higher the proposed price, the more scope there will be for modifications in payment practices to eliminate Fedwire daylight overdrafts. But, in the limit, if sufficient modifications were not forthcoming, a price above the federal funds rate would guarantee elimination of daylight overdrafts, no matter which proposal was adopted, because excess reserves would be the economical way to avoid the price. Charging this high price would transform each proposal into a variant of the penalty rate proposal. However, unless a substantial earnings rate was offered on overnight holdings of excess reserves, daylight overdraft elimination would be quite costly to the banking system. In any case, imposing this net cost on banks and their customers to eliminate daylight overdrafts would not avoid moral hazard to the extent that excess reserves would feed an extensive market in riskless daylight loans.

## IV. Conclusion

Fedwire daylight overdrafts of Federal Reserve deposit accounts create a moral hazard that pricing might reduce. Pricing could have the desired result to the extent that banks would respond by modifying payment practices, or by bringing payments-related credit needs under more effective market discipline based on risk evaluation.

Much of Fedwire payment and daylight overdraft volume can be traced to unsecured interbank lending and to settlement of securities-market trading. Rapid growth of these activities has taken place within the nationwide framework of free Fedwire daylight overdrafts. There is little basis in actual experience, therefore, for predicting the responsiveness to pricing of either Fedwire daylight overdrafts or the financial-market activities they reflect.

The hope is that modifications in payment practices would be sufficiently responsive to price that there would be no need to test the strength of credit-market discipline; that moral hazard could be eliminated at relatively low cost.

The danger is that payment practices would be unresponsive to price and that market discipline would not be engaged because of a large residual element of moral hazard in the form of priced daylight overdrafts or riskless daylight loans. If this were to be the actual outcome, it would suggest that, in addition to efficient allocation of financial resources, an insidious driving force in the rapid growth of interbank lending and securities-market trading in recent decades has been the moral hazard of Fedwire daylight overdrafts.

■ **13** This may overstate the case in one way. Pricing would operate only on daylight overdrafts in excess of a "free" allowance, determined as a percent of capital. Price then depends on credit quality, in that capital influences price. Beyond that first step, however, no discipline from the market or from regulatory credit evaluation would discourage additional borrowing.

## References

**Board of Governors of the Federal Reserve System**, "A Strategic Plan for Managing Risk in the Payments System," Report of the Large-Dollar Payments System Advisory Group to the Payments System Policy Committee of the Federal Reserve System, Washington, D.C., August 1988.

**Gelfand, Matthew D. and Lindsey, David E.**, "The Simple Microanalytics of Payments System Risk," Finance and Economics Discussion Series No. 61, Board of Governors of the Federal Reserve System, March 1989.

**Hamdani, Kausar and Wenninger, John A.**, "The Macroeconomics of Supplemental Balances, Appendix C," *Controlling Risk in the Payments System,* Report of the Task Force on Controlling Payments System Risk to the Payments System Policy Committee of the Federal Reserve System, Board of Governors of the Federal Reserve System, August 1988.

**Humphrey, David B.**, "Payments System Risk, Market Failure, and Public Policy," in Elinor Harris Solomon, ed., *Electronic Funds Transfers and Payments: The Public Policy Issues,* Boston: Kluwer-Nijhoff Publishing, 1987.

**Simmons, Richard D.**, "Would Banks Buy Daytime Fed Funds?" *Economic Perspectives,* Federal Reserve Bank of Chicago, May/June 1987, 36–43.

**Stevens, E.J.**, "Pricing Daylight Overdrafts," *Working Paper* 8816, Federal Reserve Bank of Cleveland, December 1988.

**VanHoose, David**, "The Angell Proposal: An Overview," mimeo, Board of Governors of the Federal Reserve System, June 1988.

# Capital Subsidies and the Infrastructure Crisis: Evidence from the Local Mass-Transit Industry

by Brian A. Cromwell

Brian A. Cromwell is an economist at the Federal Reserve Bank of Cleveland. The author would like to thank Paul Bauer, Michael Bell, John Davis, Randall Eberts, Erica Groshen, James Poterba, and William Wheaton for useful suggestions and discussion. William Lyons and Dottie Nicholas of the Transportation System Center provided invaluable assistance with the data. Financial support from the National Graduate Fellowship Program and the MIT Center for Transportation Studies is gratefully acknowledged.

## Introduction

The condition of the public capital stock—perceived by many to be dilapidated and inadequate—has received considerable attention in political, media, and academic circles in recent years.

Pat Choate and Susan Walter's *America in Ruins* gave striking examples of crumbling infrastructure and suggested that enormous increases in infrastructure investment were needed just to maintain the existing levels of services. The media and political attention given this work was highlighted by tragedies such as the 1983 collapse of the Interstate 95 bridge in Connecticut. More systematic studies by the Urban Institute and the Congressional Budget Office (1983) catalogued the existing state of public infrastructure and projected the need for new public investment.[1]

More recently, the National Council of Public Works Improvement (1988) completed a series of studies examining the state of the nation's public infrastructure, entitled *Fragile Foundations* and concluded that "...the quality of America's infrastructure is barely adequate to fulfill current requirements, and insufficient to meet the demand of future economic growth and development."[2]

Debates and studies of the infrastructure "crisis" involve a wide range of policy issues related to measuring the costs and benefits of public capital. The issue of what level of infrastructure is optimal involves addressing questions of how to measure the current state of and future needs for public capital, how to measure the impact of infrastructure on productivity and regional growth, and how expenditures on public capital should be weighed against other uses of public monies. Questions of financing involve traditional issues of fiscal federalism and public finance, including what level of government should provide infrastructure services, who should pay, and what financing mechanisms raise revenue with the least economic cost.

While most studies argue that increased public investment is needed, a more provocative set of

■ 1 The Urban Institute project included a series of case studies on municipal infrastructure. For example, see Humphrey et al. (1979). For a review of infrastructure needs studies, see Peterson et al. (1986).

■ 2 National Council of Public Works Improvement (1988), p. 1.

questions focuses on how public infrastructure arrived at its present condition and critiques the decision-making process itself. In particular, it is alleged that the structure of infrastructure financing mechanisms, combined with political and budgetary pressures, induce public officials to systematically underfund the maintenance of the existing capital stock, leading to excessive deterioration of public infrastructure. The study of infrastructure maintenance, however, has received little empirical attention due to the lack of data on local maintenance policies and a lack of natural experiments with which to evaluate public-sector maintenance.

This article reviews questions regarding infrastructure policy with a focus on how the costs and benefits of public capital and maintenance decisions are potentially distorted by budget procedures, political pressures, and the structure of federal grant policies. I then describe how the local mass-transit industry provides an opportunity to investigate public-sector investment and maintenance decisions. Empirical evidence from two recent studies of the local mass-transit industry, Cromwell (1988a, 1988b), is then summarized. The results suggest the structure of federal grant policies has important effects on infrastructure decisions of state and local governments.

## I. Infrastructure Policy Incentives

### Budget Processes

Leonard (1986) argues that ignoring depreciation and deferring maintenance are both powerful forms of hidden spending that are not accounted for by local governments. Failure to reinvest or maintain existing infrastructure is, in effect, to live off an inherited bank account. Current taxpayers spend assets provided to them by previous generations. This spending is obscured, however, by the lack of records and comprehensive accounting for fixed-asset investments from year to year.

Current accounting procedures for capital and maintenance by local governments appear to be inadequate for effective management of public infrastructure.[3] The Government Accounting Standards Board, which sets standards for public-sector accounting, requires governments to maintain records of fixed assets recorded at historical cost in a separate account group held apart from operating funds. Recording the value of immovable infrastructure assets—bridges, roads, sewers—is explicitly optional, as is the recording of depreciation. Even if a governmental unit does recognize depreciation, it is shown as an offset to the value of assets, not as an operating cost as in the private sector. When tight funds result in deferred maintenance, there is no notation in capital records of the decline in asset values from the failure to maintain them, making preventive and routine maintenance an attractive target for budget cuts.

In a 1983 survey of city and county officials by the American Planning Association, 29 percent reported having poor information on the current conditions of the city's or county's capital stock and 48 percent felt they had weak methods of evaluating the cost-effectiveness of proposed projects. Hatry et al. (1984, 1986) surveyed over 40 public works agencies and found capital investment decisions to be highly decentralized. In general, agency management determined what analysis should be undertaken and determined priorities. While most agencies had formal procedures for rating and ranking potential projects, these rankings were often based primarily on subjective information. They found few explicit estimates of expected improvement in service levels or expected reductions in future costs from individual proposed projects.

Budgeting procedures for maintenance were found to be even more deficient. The agencies surveyed undertook only a small amount of regular, systematic examination of capital maintenance and repair options and did not regularly and systematically examine trade-offs between preventive maintenance activity (such as painting bridges or cleaning sewers) and other major options, such as rehabilitation or reconstruction. The Hatry study found no examples in which a local government considered the costs of deferred maintenance.

Several proposals for maintenance evaluation procedures have surfaced in recent years for several common forms of public infrastructure. For example, Archuleta (1986) proposed a program for effective preventive maintenance for water and wastewater facilities. Pavement maintenance management systems promoted by the American Public Works Association (1987) enable managers to monitor road pavement conditions and schedule needed repairs. Carlson (1986) of the Federal Highway Administration proposed a similar systematic maintenance review process for bridges. Implementation of such proposals,

■ 3 These arguments were first advanced by Leonard (1986) and are also presented in Blumenfeld (1986) and the National Council of Public Works Improvement (1988).

however, often requires a crisis atmosphere. The state of Connecticut, for example, instituted a comprehensive bridge inspection and repair program that identified and ranked needed bridge reconstruction following the I-95 tragedy. There is no obvious general groundswell of public opinion, however, for the reform of infrastructure accounting procedures.

## Maintenance and Visibility

Many aspects of the infrastructure problem, particularly issues of maintenance and rehabilitation, have low levels of visibility and are not readily apparent to voters and elected officials. The costs of neglected infrastructure accrue over time and are not immediately apparent or measurable. As discussed in Eberts (1988), often they occur in the form of lost productivity and slower regional growth. Even when observed, the long-run benefits of maintenance practices are potentially discounted by elected officials with short time horizons. Cohen and Noll (1984), for example, demonstrate that legislators maximizing the probability of reelection seek to defer such costs.

Elected officials may also derive greater utility from new investment than from maintenance. Possible sources of utility from capital projects for public officials include political support and contributions from direct project beneficiaries. Weingast et al. (1981) present a model of legislative behavior in which the geographic incidence of benefits and costs systematically biases public decisions toward larger-than-efficient projects. Capital projects give benefits directly to a small group, while their costs are widely distributed.

Further political benefits come from being associated with large and visible investment projects that do not accrue from the more mundane activities of maintenance. An assistant secretary for Housing and Urban Development asked, "Have you ever seen a politician presiding over a ribbon-cutting for an old sewer line that was repaired?"[4] Such effects further encourage the substitution of investment for maintenance.

## Capital Financing Policies

The political and budgetary bias against infrastructure maintenance is reinforced by two common features of capital financing: debt-financing of new capital and the traditional emphasis of federal grant policies on capital subsidies.

Local governments often finance new purchases of capital, as well as major reconstruction and rehabilitation, through borrowing. Ordinary maintenance expenditures, however, are counted as operating expenses and are financed through current funds. This treatment of maintenance stems in part from the wide variance of maintenance activities. Certain maintenance activities, such as sweeping sidewalks or patching potholes, have immediate short-term benefits and, according to the benefit principle of public finance (those who benefit from public services should pay), should be paid for by the immediate beneficiaries through current revenues. The benefits of other maintenance activities, such as painting bridges or flushing sewers, accrue over many years. Maintenance of this sort constitutes a form of public investment that according to the benefit principle should be paid over many years through debt-financing.[5]

Treating all maintenance activities as current expenses ineligible for debt-financing ignores their investment component and results in underfinancing when operating budgets are tight. During periods of budget constraints, officials choose between funding preventive maintenance at the expense of cutting back on other programs, or allowing infrastructure to deteriorate until major reconstruction is needed, which can be funded through debt. As the mayor of Lincoln, Nebraska observed, "In the choice between laying off police and maintaining sewers, the sewers always lose."[6]

Federal grant policies for public infrastructure further exacerbate the bias against infrastructure maintenance. Under the rationale that local taxpayers should pay to operate the facilities presented to them, federal grants often heavily subsidize new construction, but provide no assistance for maintenance or other operating expense.

A wide range of federal grant programs provide major assistance for infrastructure at the

■ **4** *Newsweek,* August 2, 1982. Also cited in Leonard (1986).

■ **5** Maintenance is often considered in the operations research and investment literature to be a fixed operating expense. For a standard example, see the optimal equipment replacement model in Jorgenson et al. (1967) and the discussion in Nickell (1978). For good reviews of models of preventive maintenance, see Pierskall and Voelker (1976) and Sherif and Smith (1981). The treatment of maintenance as a form of investment is shown in Bitros (1976). This approach is used in models of housing stock maintenance, in which maintenance expenditures have important effects on rental income and sale price. See Vorst (1987), Arnott et al. (1983), and Sweeney (1974) for examples of such models.

■ **6** *Newsweek,* op. cit.

state and local level. In 1988, $25 billion in federal grants accounted for 26 percent of state and local capital spending. This included $13.7 billion granted by the Federal Highway Administration (FHWA) for the construction and rehabilitation of highways; $2.6 billion from the Environmental Protection Agency for pollution control and abatement; $2.4 billion in capital financing for mass transit administered by the Urban Mass Transit Administration (UMTA); and $3.1 billion granted through the Community Development Block Grant program.[7]

While the structure of grants varies from program to program, most provide capital assistance at a high matching rate, with the state and local government required to meet the matching share. The FWHA provides financing for completion, rehabilitation, and reconstruction of the interstate highway system at a 90 percent matching rate. Discretionary grants from UMTA for major rail and subway systems provide funds up to a 75 percent matching rate. Formula grants from UMTA pay 80 percent of the cost of regular transit vehicle replacement. No corresponding subsidies, however, are provided for maintenance. These subsidies distort the relative prices facing local governments for new investment versus maintenance of existing infrastructure. Even if the federal matching rate is not specified in formula, the expectation of federal aid potentially induces local officials to substitute away from maintenance. The empirical work we now turn to attempts to identify such substitution.

## II. Local Mass Transit: A Natural Experiment on Subsidies and Infrastructure

As discussed in the previous section, several elements of public accounting, political and budget processes, and capital financing potentially lead to underfunding of infrastructure maintenance and result in excessive deterioration of public capital. Empirical research on the relative importance of these issues, however, has been limited by a dearth of data on capital assets and maintenance, and by a lack of obvious natural experiments with which to evaluate public-sector maintenance practices. In two recent studies, Cromwell (1988a) and Cromwell (1988b), however, I examine the impact of capital subsidies on investment and maintenance decisions of local governments, using data on the maintenance policies of both publicly and privately owned local mass-transit providers. While not addressing all issues of infrastructure maintenance, these studies suggest that the structure of federal grants has significant effects on the infrastructure decisions of state and local governments.

The data used were collected under the Section 15 Reporting System administered by the Urban Mass Transportation Administration (UMTA). Section 15 data for fiscal year (FY) 1979 through FY1985 are available for 435 transit systems. The data set contains extensive information on vehicle fleets as well as expenditures and labor hours for vehicle maintenance, providing a consistent measure of public capital and maintenance efforts not previously seen. These data provide an unusually detailed panel of local governments' physical assets. Vehicle inventories for each system are broken down by model, year of manufacture, and mileage.

Data are also available for certain privately owned and operated systems. Their inclusion in the Section 15 data results from contracting with a public recipient of Section 9 funds to provide transit services. As these contracts often provide for the leasing of public vehicles, care was taken to examine maintenance and scrappage decisions only on vehicles owned outright by private operators.

### Federal Transit Policies

The federal government finances a major part of local public mass transportation. The principal federal grant program for entities that only operate bus lines (the focus of these studies) is the Section 9 formula grant program that distributes funds to urbanized areas for use in transit operating and capital expenditures. The Section 9 capital funds are principally used for vehicle replacement and pay up to 80 percent of the cost of a new vehicle. As funds are adequate for normal vehicle replacement, this matching rate represents an enormous marginal subsidy for new capital.

Vehicle maintenance, however, is counted as an operating expense and is ineligible for the capital subsidy. Due to a desire by UMTA to wean local entities away from operating assistance, the Surface Transportation Act of 1982 capped the level of funds available for operating assistance for FY1983 and beyond to some 90 percent of the FY1982 level, or to 50 percent of a property's operating deficit, whichever was

■ 7 See U.S. Office of Management and Budget (1989). For further discussion of federal grants-in-aid, see Delmar and Menendez (1986).

TABLE 1

**Used Transit Vehicle Prices in 1987 and 1988**

| Year of Manufacture | Average Price | Max. Price | Min. Price | Number of Observations |
|---|---|---|---|---|
| Public | | | | |
| 1961-65 | $ 301 | $ 1,000 | $ 100 | 255 |
| 1966-70 | 841 | 3,500 | 400 | 163 |
| 1971-75 | 1,648 | 6,000 | 250 | 239 |
| 1976-80 | 8,863 | 17,000 | 3,300 | 8 |
| Private | | | | |
| 1961-65 | $3,500 | — | — | 11 |
| 1966-70 | 6,590 | — | — | 11 |
| 1971-75 | 7,500 | — | — | 9 |
| 1976-80 | 18,000 | — | — | 1 |

SOURCE: Telephone survey by author.

lower. The overwhelming majority of public-transit properties are constrained by the cap and receive no operating assistance on the margin.

Federal control over maintenance principally consists of setting an upper limit for deterioration of federally purchased equipment. UMTA requires local transit properties to operate buses purchased with federal funds for at least 12 years or 500,000 miles.[8] Failure to do so results in a penalty in federal assistance for new capital purchases. This 12-year limit, however, is below the potential operating life of 15 to 20 years for standard bus models when properly maintained.

The structure of the UMTA grants results in a large distortion in the relative price of maintenance versus new investment for buses over 12 years old. If the capital and maintenance decisions of local government are sensitive to the structure of subsidies, we would expect the following results. First, publicly owned buses should depreciate quickly, with little physical or financial value left after 12 years. Second, we would expect higher average levels of maintenance in the private sector compared to the public sector. Finally, in the public sector we would expect low levels of scrappage before the 13-year point, a marked shift in scrappage at year 13, then high levels of scrappage thereafter. A similar pattern for privately owned vehicles is unlikely, as they are not subject to such a discontinuity in the price of new equipment.[9]

## III. Empirical Evidence on Subsidies and Transit Capital

### Evidence from Used-Bus Prices

Evidence from used-bus prices supports the thesis that public equipment depreciates rapidly. The used-bus market is highly fragmented and ad hoc in nature. The disposition of equipment is not reported in the Section 15 data, and no central data source of used-bus prices or sales exists. UMTA officials report, however, that the used transit bus market is depressed. The supply of public vehicles over 12 years old far exceeds demand—and vehicles are most commonly sold for scrap. Depressed prices, however, are also consistent with systematic undermaintenance of equipment.

To confirm this, I collected transaction prices for some 645 transit vehicles sold in 1987 and 1988 by contacting all properties that solicited bids for used vehicles during this period.[10] The results of this survey are shown in table 1. Prices for publicly owned vehicles manufactured before 1971 ranged from $100 to $3,500, with an

■ **8** See UMTA (June 1985).

■ **9** Previous studies on transit subsidies have used detailed engineering data from specific transit systems to simulate the effects of capital bias in the subsidy structure on scrappage dates. Tye (1969) used data from the Cleveland and Chicago transit systems to simulate the effect of subsidies in the late 1960s that paid for new capital at a 66.6 percent rate, but which provided no assistance for operating expenses. He calculated that the subsidy would lead a cost-minimizing transit firm to replace buses at half the efficient age. For average levels of utilization, this implied scrappage at 8 to 10 years versus an efficient 17 to 20 years, with the resulting waste of resources equaling 27 percent of the subsidy. Similarly, Armour (1980) used data from Seattle Metro and calculated that the 80 percent federal capital subsidy reduced the optimal scrappage point from 20.5 to 26 years to 8.5 to 10 years.

Frankena (1987) is the paper closest in spirit to the empirical work presented here. Using probit estimation with 1961 to 1983 data on scrappage of Canadian buses, this study shows that scrappage increases with age, and that significantly higher average scrappage rates followed the imposition of a capital-biased subsidy program in 1972. He finds no significant change, however, in the scrappage rate when the capital subsidies take effect at age 15 (the critical point in the Canadian subsidy program). In general, the hazard-model estimators used here dominate the probit approach. They allow for variation in the underlying hazard rate over time, and control for bias introduced by vehicles dropping out of the sample when scrapped. The results, as will be seen, show a significant impact on scrappage when subsidies take effect.

■ **10** Used-bus prices were obtained by contacting all agencies soliciting bids in *Passenger Transport* between January 1987 and June 1988. Typically, less than 10 bids were received per auction with a mean of five bids reported by properties that would provide this information. Those bidding included Caribbean nations, church groups, charter-bus operators, people planning to make recreational vehicles, and farmers in need of storage space. If the vehicles were purchased with federal funds, UMTA collected 80 percent of the proceeds with an allowance made for administrative expenses. The costs of soliciting bids or holding an auction, however, often were reported to exceed the remaining local share.

**TABLE 2**

**Vehicle Maintenance Expenses and Labor Hours[a]**

| | Private | Public |
|---|---|---|
| Expenses per mile ($1.00) | 0.77 (0.12) | 0.53 (0.02) |
| Labor hours per 1,000 miles | 37.8 (3.6) | 29.3 (1.4) |
| Percent of fleet > 12 years old | 38.4 | 22.0 |
| Percent mileage on vehicles > 12 years old | 26.7 | 11.2 |
| Number of observations | 22 | 100 |

a. 1984 cross-section sample means (standard errors).
SOURCE: Author's calculations.

average price of $511. Even vehicles reported to be well-maintained typically did not sell for over $3,000. Prices for vehicles manufactured between 1971 and 1975 ranged from $250 for scrapped vehicles to $6,000 for well-maintained vehicles. Prices for newer vehicles manufactured between 1976 and 1980 averaged $8,863.

I was also able to obtain used-vehicle prices for a much smaller sample of privately owned vehicles. These prices, also shown in table 1, suggest that the private vehicles are in better condition and command a higher price, with prices averaging from $3,500 to $7,500 for vehicles manufactured before 1976. Other private companies, however, reported selling their vehicles for scrap at the depressed prices similar to those received by public agencies.

The extremely low prices on used buses suggest that maintenance practices can lead to rapid deterioration of equipment in the public sector. It is important, however, to distinguish between variations in maintenance and depreciation attributable to unavoidable operating conditions, and variations due to capital grant policies or bureaucratic behavior that are potential sources of government inefficiency. The empirical work that follows attempts to identify these separate effects.

### Evidence on Maintenance

The impact of the capital grant structure on average levels of maintenance is examined in Cromwell (1988a). My initial empirical work examines a cross-section of Section 15 data for FY1984 from 122 transit properties. The sample consists of single-mode bus operators—properties that provide only fixed-route bus service as opposed to rail or demand-response service—that operated at least five revenue vehicles. Table 2 reports sample means for maintenance expenses and maintenance employees, scaled by annual vehicle miles. In general, the average levels of both expenses and labor hours follow the predicted patterns. The private systems, on average, spend 45 percent more on maintenance per mile and devote 29 percent more labor hours to maintenance than do the public systems.

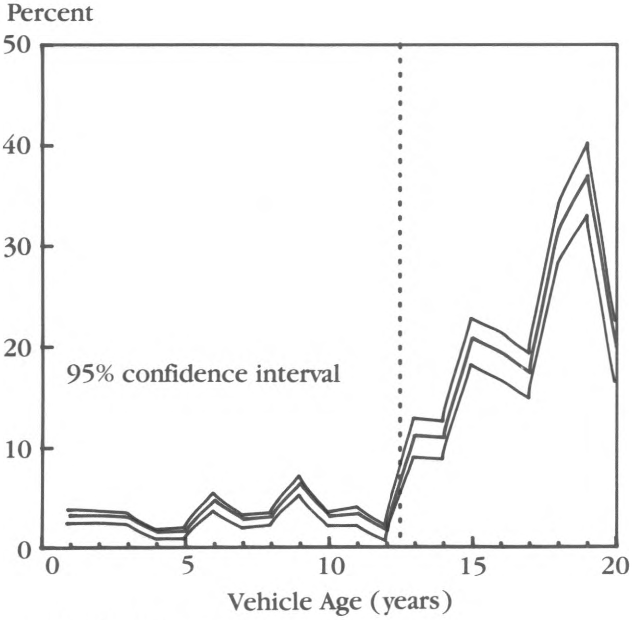
The average age of vehicles in private systems is substantially higher than that for public fleets, with 38.4 percent of the private fleets being more than 12 years old compared to 22.0 percent of the public fleets. The distribution of vehicles weighted by miles is similar, with 26.7 and 11.2 percent of the mileage being run on vehicles older than 12 years for the private and public systems, respectively. The older fleet in the private systems is consistent with privately owned capital deteriorating slower than publicly owned capital as a result of greater maintenance efforts.

The means shown in table 2, while consistent with the predicted results regarding the private versus public operators, do not control for systematic differences due to wages, operating conditions, and fleet composition. For example, many of the private systems operate in the New York metropolitan area, which is noted for its harsh operating conditions. To examine the public/private differential more systematically, I use pooled time-series cross-section regression analysis on a sample of systems between 1982 and 1985. Independent variables include maintenance wage rates, operating conditions, fleet composition, fleet age, and a dummy variable for operation in the New York area. The results show that, controlling for wages, operating conditions, and fleet composition, privately owned transit companies devote some 14 to 17 percent more labor hours to maintenance than do publicly owned and managed transit companies. The analysis then uses this public/private differential, along with cross-state variation in grant policies, to measure the elasticity of maintenance with respect to capital subsidies. The point estimates suggest an elasticity of –0.16, meaning that a 10 percent increase in the subsidy rate for transit capital reduces vehicle maintenance by 1.6 percent.

The estimates are statistically significant and suggest that average maintenance levels are higher in the private sector. They do not necessarily demonstrate, however, that public capital deteriorates at a faster rate than privately owned

**F I G U R E 1**

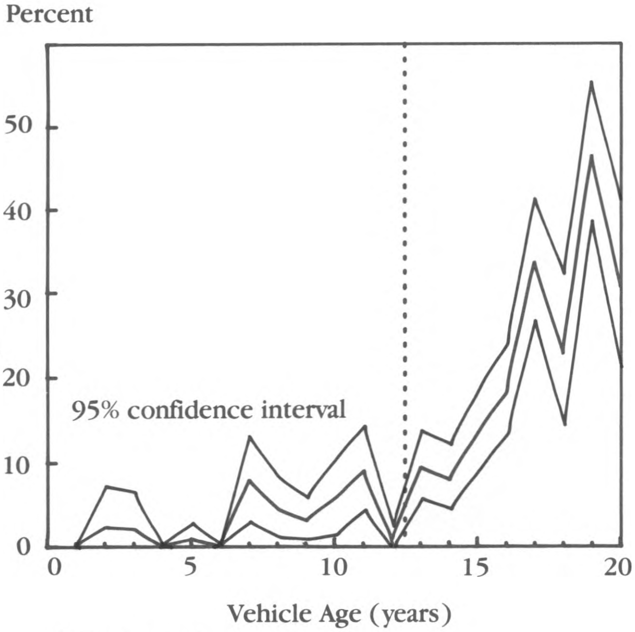
**Scrappage Rate**
**Public Vehicles**




SOURCE: Author's calculations.

**F I G U R E 2**

**Scrappage Rate**
**Private Vehicles**




SOURCE: Author's calculations.

capital. The higher levels of maintenance labor hours could be attributed to less capital-intensive maintenance practices. Furthermore, an implicit assumption that maintenance is qualitatively similar between the two sectors could be false. If one sector fixes equipment upon failure, as opposed to conducting preventive maintenance, differences in overall maintenance levels could result. The companion analysis in Cromwell (1988b), however, directly examines the scrappage and retirement rates of private versus public equipment to determine whether the higher maintenance in the private sector is reflected in longer equipment life.

### Evidence on Scrappage

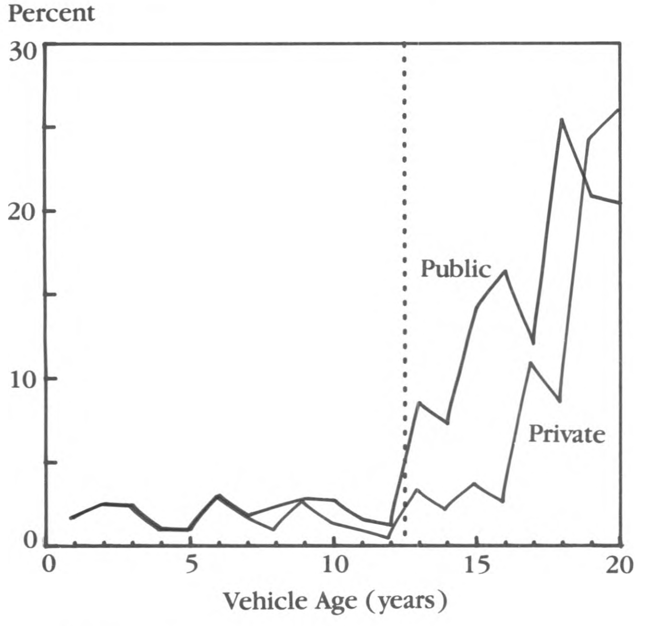
Cromwell (1988b) examines the impact of subsidies on equipment life by tracking vehicles in the UMTA data set from 1982 through 1985. Scrappage decisions were observed for 15,829 vehicles, including 1,005 privately owned vehicles from 11 privately owned companies. Vehicles that changed from active to inactive status or that were dropped from the fleets between report years were counted as scrapped. The results provide strong evidence that federal grant policies have a direct impact on local scrappage decisions.

The probability of scrappage for public and private vehicles of different ages (or empirical hazard) can be estimated directly from the observed scrappage rates and is plotted, with 95 percent confidence intervals, in figures 1 and 2.[11] The estimates in general suggest the importance of federal grant policies for public-sector scrappage. The hazard for public vehicles averages under 4 percent for years prior to age 13, then jumps to over 11 percent at age 13, decreases slightly at age 14, then rises steadily to 37 percent by age 19. Standard errors calculated for these estimates suggest that the hazards for public vehicles are measured with much precision and that the shift at the 13-year point is statistically significant.

■ **11** The empirical scrappage rate presented here is also known as the Kaplan-Meier (1958) hazard estimator, which directly estimates the hazard function from the sample of vehicles. For each time $t$, the number of failures $D(t)$ (that is, the number of vehicles scrapped) is divided by the total number of vehicles at risk at the start of time $t$, $R(t)$. Censored spells (that is, vehicles that are not observed to be scrapped) are included in the risk set previous to their censor time and are dropped thereafter. This treatment of censoring yields a consistent estimate of the true hazard at each time $t$ as long as the censoring mechanism and vehicle age are independent of each other. The standard errors were estimated following a suggestion in Kalbfleisch and Prentice (1980).

FIGURE 3

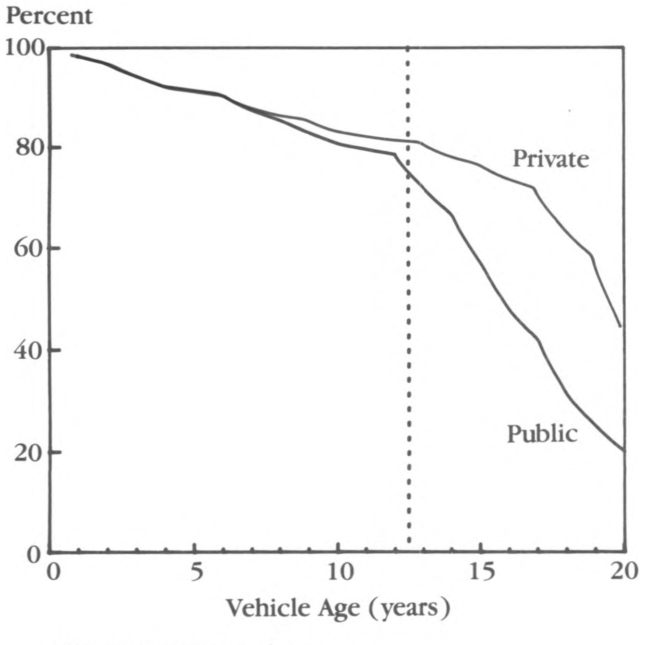
**Scrappage Rate**
**Private vs. Public Vehicles**




SOURCE: Author's calculations.

FIGURE 4

**Survivor Functions**




SOURCE: Author's calculations.

The private-vehicle hazards are estimated with less precision and exhibit more volatility, but in general show a rise in scrappage from near 0 for the 1- to 6-year period to an average 5 percent for the 7- to 10-year period to 9 percent at the 13-year point. Due to only one scrappage out of 143 in the age-12 risk set, however, the estimated hazard at year 12 is quite low, and a shift appears to occur at the 13-year point—contrary to the predicted pattern. This shift can be attributed, however, to the smallness of the sample size and, given the estimated hazards in the surrounding years, the pattern of estimated hazards for private vehicles appears to be markedly different from the public sector.

These empirical hazard rates do not account for heterogeneity across transit systems in prices of maintenance and operating conditions. Given the large number of private vehicles operating in the New York metropolitan area, for example, adverse operating conditions might have a major impact on observed private-sector scrappage. To account for this heterogeneity, I employed a hazard estimator that allows for nonparametric estimation of the baseline scrappage rate, while permitting estimation of the impact of operating conditions, wage rates, and other explanatory variables.[12] The resulting baseline hazards are shown in figure 3. The impact of the grant structure on public-sector scrappage is readily apparent. While the private-sector baseline remains under 5 percent until year 16, and then rises steadily through year 20, the public-sector baseline takes a distinct and significant jump at the 13-year point from 1 percent to over 8 percent, twice that of the private sector. Scrappage then rises to over 14 percent for 15- and 16-year-old vehicles and remains above the private sector until year 19. The distinct difference in scrappage rates can be attributed to the availability of federal grants.

An alternative approach to examining public and private scrappage is to look at the survivor functions for the two sectors. The survivor function is defined as the percentage of vehicles of a given vintage that survive to a given age, as shown in figure 4. The functions further emphasize the difference between public and private

■ **12** The baseline hazard estimates shown here are estimated using the semiparametric hazard estimator shown in Meyer (1988) and first developed in Prentice and Gloeckler (1978). This estimator allows for control of explanatory variables without imposing a specific structural form on the underlying baseline hazard. Cromwell (1988b) also presents estimates using the fully parametric estimator which imposes the commonly used Weibull baseline as shown in Lancaster (1979) and Katz (1986).

scrappage policies. They track closely through year 12, then diverge as public scrappage sharply increases. Again, this shift in the survivor function at the 13-year point can be attributed to the sudden availability of federal subsidies. By age 16, only 47 percent of the public vehicles survive, compared to 73 percent for private vehicles. At age 20, 45 percent of private vehicles are still estimated to be in operation, versus 20 percent for the public sector.

The consistently lower survival rate of publicly owned vehicles after the availability of federal funds is direct evidence that federal capital grants reduce equipment life in the local public sector. It suggests that federal grant policies that subsidize the purchase of new capital, but that ignore the maintenance of existing capital, result in the increased deterioration of public infrastructure. The magnitude of savings for the transit industry from a shift in policies, however, may be small if increased maintenance expenses offset reduced vehicle expenditures. In a simulation of vehicle replacement reported in Cromwell (1988b), this is the case. In spite of increased deterioration of public capital, the net efficiency losses of the federal subsidies appear to be low. There may be unobserved costs, however, in terms of quality of service that result from lower maintenance levels and increased deterioration of equipment.

## IV. Conclusion

Several aspects of public accounting, political and budgetary procedures, and capital financing potentially lead local governments to systematically underfund the maintenance of public infrastructure. The resulting excessive deterioration of public capital has been advanced as a possible source of the "infrastructure crisis" of recent years.

This article summarizes the results of two studies of one aspect of infrastructure maintenance: the impact of large federal capital subsidies for new investment with no corresponding subsidies for maintenance. Using data from the local mass-transit industries, the empirical results suggest federal subsidies for new transit vehicles lower maintenance levels and increase scrappage rates in public transit systems. The extremely low resale value of used vehicles further suggests excessive deterioration. In the case of local mass transit, however, the net cost of the distortion appears to be small. The results suggest that increased purchases of vehicles are offset by lower maintenance costs.

While the efficiency losses of the transit subsidies for new vehicles appear to be small, they still show that local governments respond significantly to incentives in the price of maintenance versus new investment introduced by federal subsidies. Given the several other biases against infrastructure maintenance discussed in section I, this suggests that federal policies should focus more on the maintenance and upkeep of facilities purchased with federal funds. Possible proposals to support maintenance include reducing the distortion in the relative price of maintenance versus new investment facing local authorities through direct federal subsidies of important maintenance activities or through a reduction in the federal subsidy rate for capital projects. Adoption of preventive maintenance programs developed by public works experts could also be a requirement of receiving federal aid. Leonard suggests the development of a maintenance schedule at the time of acquisition of a new capital facility. The financial requirements for maintenance would be a formal liability recorded on a jurisdiction's financial statement. Reforms in this direction would help ensure that existing capital is better preserved and that large projected investments in new infrastructure are not wasted.

Finally, future research in this area could include analysis on how the incentive effects described here for the local mass-transit industry apply to other forms of infrastructure. Using the standard optimal equipment replacement model in Cromwell (1988b), one would expect that the elasticity of optimal equipment life with respect to capital subsidies is larger for capital goods with shorter useful equipment lives, and larger for capital goods whose acquisition costs are large relative to maintenance costs. It would be interesting to examine the difference in magnitude of the distorting effects of federal subsidies for infrastructure with these characteristics.

Furthermore, the distorting effects of capital subsidies are likely to be more severe when the deterioration of infrastructure is less visible—as in the case of sewers, water mains, or the undersides of bridges. Less visibility reduces the ability of voters or federal bureaucrats to monitor the condition of local infrastructure. Such monitoring potentially acts as a check on the incentives to undermaintain that are introduced by capital subsidies.

## References


**American Public Works Association**, "Good Practices in Public Works," prepared for the National Council on Public Works Improvement, August 1987.

**Archuleta, E.**, "Standardized Inventory and Conditions Indexes for Water and Wastewater Facilities," *APWA Reporter,* June 1986.

**Armour, R.**, "An Economic Analysis of Transit Bus Replacement," *Transit Journal,* 1980, *6,* 41–54.

**Arnott, R., Davidson, R., and Pines, D.**, "Housing Quality, Maintenance, and Rehabilitation," *Review of Economic Studies,* 1983, *50,* 467–94.

**Bitros, G.**, "A Statistical Theory of Expenditures in Capital Maintenance and Repairs," *Journal of Political Economy,* October 1976, *84,* 917–36.

**Blumenfeld, Arthur A.**, "Infrastructure and the Public Finance Professional," prepared for the National Council on Public Works Improvement, October 1986.

**Carlson, E.**, "Put Your Bridge Money Where It's Needed," *APWA Reporter,* June 1986.

**Choate, Pat and Walter, Susan**, *America in Ruins: Beyond the Public Works Pork Barrel,* Washington, D.C.: Council of State Planning Agencies, 1981.

**Cohen, L. and Noll, R.**, "The Electoral Connection to Intertemporal Policy Evaluation by a Legislator," Stanford University, Center for Economic Policy Research, Publication No. 36, 1984.

**Congressional Budget Office**, *Public Works Infrastructure: Policy Considerations for the 1980s,* Washington, D.C., 1983.

**Cromwell, Brian**, (1988a) "The Impact of Federal Grants on Capital Maintenance in the Local Public Sector," *Working Paper* 8812, Federal Reserve Bank of Cleveland, November 1988.

________, (1988b) "Federal Grant Policies and Public Sector Scrappage Decisions," *Working Paper* 8811, Federal Reserve Bank of Cleveland, November 1988.

**Delmar, Clare E. and Menendez, Aurelio,** "Infrastructure Financing in the United States: Issues and Mechanisms," prepared for the National Council on Public Works Improvement, September 1986.

**Eberts, Randall W.**, "Some Empirical Evidence on the Linkage Between Public Infrastructure and Local Economic Development," manuscript, Federal Reserve Bank of Cleveland, October 1988.

**Frankena, Mark W.**, "Capital-Biased Subsidies, Bureaucratic Monitoring, and Bus Scrapping," *Journal of Urban Economics,* March 1987, *21,* 180–94.

**Hatry, Harry P., Millar, and Evans,** *Guide to Setting Priorities for Capital Investment,* Washington D.C.: The Urban Institute, 1984.

________, **Neary, Kevin, and Allen, Joan,** "The Capital Investment and Maintenance Decision Process in the Public Sector," prepared for the National Council on Public Works Improvement, Washington D.C.: The Urban Institute, October 1986.

**Humphrey, Nancy, Peterson, G., and Wilson, Peter,** *The Future of Cleveland's Capital Plant,* Washington, D.C.: The Urban Institute, 1979.

**Jorgenson, Dale, McCall, John J., and Radnor, Roy,** *Optimal Replacement Policy,* Chicago: Rand McNally, 1967.

**Kalbfleisch, J. and Prentice, Ross L.**, *The Statistical Analysis of Failure Time Data,* New York: John Wiley & Sons, Inc., 1980.

**Kaplan, E.L. and Meier, Paul,** "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association,* 1958, *53,* 457–81.

**Katz, Lawrence**, "Layoffs, Recall and the Duration of Unemployment," *National Bureau of Economic Research Working Paper* No. 1825, January 1986.

**Lancaster, Tony,** "Econometric Methods for the Duration of Unemployment," *Econometrica,* July 1979, *47,* 939–56.

**Leonard, Herman B.,** *Checks Unbalanced: The Quiet Side of Public Spending,* New York: Basic Books, Inc., 1986.

**Meyer, Bruce,** "Semiparametric Estimation of Hazard Models," mimeo, Northwestern University, 1988.

**National Council of Public Works Improvement,** "Fragile Foundations: A Report on America's Public Works, Final Report to the President and Congress," February 1988.

**Nickell, S.,** *The Investment Decisions of Firms,* Cambridge University Press, 1978.

**Peterson, G., Miller, Ted, Humphrey, Nancy, and Walker, Christopher,** "Infrastructure Needs Studies: A Critique," prepared for the National Council on Public Works Improvement, The Urban Institute, October 1986.

**Pierskall, W. and Voelker, J.,** "A Survey of Maintenance Models: The Control and Surveillance of Deteriorating Systems," *Naval Research Logistics Quarterly,* 1976, *23,* 353–88.

**Prentice, Ross and Gloeckler, L.,** "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics,* 1978, *34,* 57–67.

**Sherif, Y. and Smith, M.,** "Optimal Maintenance Models for Systems Subject to Failure—A Review," *Naval Research Logistics Quarterly,* 1981, *32,* 47–74.

**Sweeney, J.,** "A Commodity Hierarchy Model of the Rental Housing Market," *Journal of Urban Economics,* 1974, *1,* 288–323.

**Tye, William B.,** "The Economic Costs of the Urban Mass Transportation Capital Grant Program," Ph.D. thesis, Department of Economics, Harvard University, Cambridge, Mass., October 1969.

**U.S. Office of Management and Budget,** *Special Analysis D, Budget of the United States Government, Fiscal Year 1990,* Washington, D.C.: U.S. Government Printing Office, January 1989.

**Urban Mass Transportation Administration, U.S. Department of Transportation,** *National Urban Mass Transportation Statistics, Section 15 Annual Report,* Washington, D.C.: U.S. Government Printing Office, various years.

**Urban Mass Transportation Administration, U.S. Department of Transportation,** Circular 9030. 1, June 1985.

**Urban Mass Transportation Administration, U.S. Department of Transportation,** *A Directory of Urban Public Transportation Service,* August 1986.

**Vorst, A.,** "Optimal Housing Maintenance under Uncertainty," *Journal of Urban Economics,* March 1987, *21,* 209–27.

**Weingast, Barry R., Shepsle, Kenneth H., and Johnson, Christopher,** "The Political Economy of Benefits and Costs," *Journal of Political Economy,* August 1981, *89,* 642–64.

# Employment Distortions Under Sticky Wages and Monetary Policies to Minimize Them

by James G. Hoehn

James G. Hoehn is an economist at the Federal Reserve Bank of Cleveland. The author would like to thank Charles Carlstrom, Brian Cromwell, Randall Eberts, William Gavin, Anil Kashyap, and John Roberts for helpful discussions and comments on drafts of this paper.

### Introduction

**A** major problem that monetary authorities must address is that contracts are made in nominal terms. During the contract interval, the terms may become inappropriate and cause misallocations if one of the parties has discretion over activity levels.

The prototype case emphasized by macroeconomists is that of the labor contract, which may run for three years, during which the nominal wage is stuck, despite changes in the marginal productivity and disutility of labor caused by various events. Employers have some discretion over employment levels and can improve profits by adjusting employment in response to changes in the state of the economy. The profit-maximizing employment level will not, generally, be the same as the socially optimal level because the wage is stuck and does not perfectly reflect changes in the disutility of labor. An optimal monetary policy has the effect of tending to make the real wage match the marginal disutility of work in various states of the economy.

This article explores how the money supply can be manipulated by the Federal Reserve to keep the real wage close to the marginal disutility of work in various states of the economy, and thereby minimize social welfare losses associated with the employment distortions arising from sticky wages. The primary contribution of the analysis is to provide a social welfare metric defined in terms of the outcomes of an IS-LM Phillips Curve model. Simulations are run to compare the social loss under various monetary policies, including the one that is optimal in the model, as well as policies that target money, output, nominal income, and the price level. The simulations are not intended to encompass all possible structures of the economy, but instead are meant to suggest how various policies might compare under the assumptions of the model in meeting the social goal of labor-market efficiency.

### I. Employment Distortion Under Nominal Wage Contracts

**A**ccording to the basic neoclassical theory of wage determination, wages tend to be set at a level that reflects both productivity and disutility of work. If the nominal wage is set in advance, it will tend to be set at a level equal to the expected marginal revenue product of labor and the marginal disutility of work. Then, the real wage will be expected, on average, to clear the labor market, and employment will be at optimal levels (leaving aside issues related to monopoly power or other such sources of externalities, which are not essentially monetary problems because there is little the monetary authorities can do to ameliorate them).

Once the nominal wage is set, unanticipated events can render that wage incorrect and cause misallocation. For example, if the demand for commodities rises beyond what was expected at the time contracts were signed, and if monetary policy keeps the money supply constant, the price level will rise, lowering the real wage under contracts. This reduction in the real wage will tend to cause an expansion of employment by profit-maximizing firms. In an extreme case of period-by-period profit-maximization, the expansion of employment would carry to the point at which the marginal product of labor falls to the lower real wage. This expansion of employment is socially inappropriate because the additional employment produces less value of output than the disutility of work it incurs.

To take another example of how predetermined wages can result in inefficiency, consider an autonomous cyclical labor productivity improvement. Further assume, for illustration, that as output supply increases, the price level is kept from falling by monetary expansion. The profit-maximizing firms expand employment in order to take advantage of the higher productivity, but will not face increasing unit labor costs if the contract calls for employees to supply all the labor the firm wants at a predetermined wage. Employment will overexpand because firms are not required to consider the rising disutility of work.

Ideally, real wages should be regulated by policy so that they match the marginal disutility of work. In the case of an autonomous cyclical labor productivity shock, real wages should rise to keep pace with the rise in the disutility of work associated with higher employment. A monetary policy that tended to allow the price level to fall when autonomous increases in labor productivity occur could help real wages match the marginal disutility of work. Then, the employment level would still rise with productivity improvements, but not excessively so. One policy that tends to set up a negative relation between labor productivity shocks and the price level is a nominal income, or GNP, target. In simulations with a model, GNP targets are close to optimal in that people's time tends to be allocated between labor and leisure in an appropriate way.

## II. A Simulation Model

The simulation model combines the notion of sticky wages and the IS-LM demand apparatus with autonomous labor productivity shocks. Elsewhere, I have shown that a simpler (constant-velocity) version of the model can account for stylized facts, such as the natural-rate hypothesis and the mild procyclicity of real wages and productivity (see Hoehn [1988]), so long as forward-looking expectations guide nominal wage contractors. The IS-LM apparatus for representing intuitions about demand is preferred here over simple velocity equations, because the effects of monetary policy can be offset or enhanced by changes in velocity, and because IS-LM allows assessment of the information policymakers can obtain from observations on the nominal interest rate. The model has three shocks: to money demand, to commodity demand, and to the marginal labor productivity schedule. These features provide a model consistent with the stylized facts and containing utilitarian welfare criteria for policy.

Relative to the standard macroeconomic models involving wage stickiness, four changes are offered to make a useful policy model.

(i) Expectations of inflation and productivity are forward-looking (Muthian rational).

(ii) Labor productivity is subject to autonomous cyclical variations (as well as to variations induced by shifts in commodity and labor demand).

(iii) Employment is determined not strictly by demand, but is also influenced by supply.

(iv) The information content of the interest rate is used by goods demanders and the central bank.

To incorporate these features, the following model is offered.

### Supply Sector

Following Fischer (1977), represent multiyear nominal wage bargaining with two-period staggered, or overlapping, contracts. The model economy is composed of two groups of firms, identical in all respects, except for the date at which currently effective labor contracts were signed. Firms having signed wage contracts at the end of last period ($t-1$) are referred to as group one firms, while those that signed wage contracts at the end of the period before last ($t-2$) are referred to as group two firms. The groups are competitive in that they take the commodity price as given, and contract with workers to pay them their expected marginal revenue product.

Economywide aggregates are simulated by taking the average of the two groups' firms.

The main difference between the determination of wages in the model here and that of other sticky-wage models is that contract wages here adjust completely and efficiently to information available at the time of wage bargains. In some other models, such as that of Taylor (1979), wages can take longer than a contract interval to respond completely to events, and are subject to random variations conceived of as wage-setting errors. Taylor's model can be justified as more realistic. However, the model used here is more consistent with microeconomic theory about the determination of wages and is consistent with the natural-rate hypothesis: the average level of employment is invariant with respect to the money supply rule.

As in most sticky-wage models, variations in employment are those for a representative worker. Implicitly, employment variations are variations in hours worked among workers who each have jobs in all states of the economy. The model falls short of accounting for unemployment.

The determination of employment and wages reflects both Keynesian and neoclassical elements. Hall (1980) and Barro (1977) have sought to reconcile the fact of sticky wages with the neoclassical theory of employment determination by arguing that sticky wages need not have any misallocational effects. Efficient contracts, which could be implemented in the absence of transactions or enforcement costs, would involve optimal employment determination as productivity varied, so that sticky wages would have no allocational effects. Here, it is supposed that there are constraints on optimal contracts that prevent workers and firms from effecting optimal contracts. However, the traditional Keynesian assumption that employment is strictly demand-determined is softened. Instead, the employment reflects both the optimal level (the employment level associated with the intersection of demand and notional supply curves) and the demand for labor at prevailing prices and wages. This is simulated by an equation for employment that makes it a weighted average of both the optimal level and the notional demand. The weight attached to the demand can be conceived of as the degree to which sticky wages have misallocational effects or, alternatively, the degree to which the problems of ideal contract enforcement are effective constraints.

In order to derive this employment equation, first the notional labor demand is developed, then the notional labor supply is formulated, and then they are put together. Finally, the employment equation, in conjunction with the production function and stochastic assumptions about productivity disturbances, implies a supply function, or Phillips Curve: a semireduced form equation for output supply as a function of the state of technology and unexpected inflation.

**Notional Labor Demand**

A firm's production function is

$$(1) \quad Y_{it} = U_t N_{it}^{\gamma}, \; 0 < \gamma < 1, \; i = 1, 2,$$

where $Y_{it}$ is the output of a firm in group $i$ in period $t$, $N_{it}$ is the labor input of a firm in group $i$, and $U$ is a global productivity shock. The marginal product of labor is

$$(2) \quad \frac{dY_{it}}{dN_{it}} = U_t \gamma (N_{it})^{-(1-\gamma)}, \; i = 1, 2.$$

In logarithmic form, output is

$$(3) \quad y_{it} = u_t + \gamma n_{it}, \; i = 1, 2,$$

where the lowercase letters $y$, $u$, and $n$ are natural logarithms of their uppercase counterparts. The (log of the) marginal product of labor is

$$(4) \quad ln\left(\frac{dY_{it}}{dN_{it}}\right) = u_t + ln(\gamma) - (1-\gamma) n_{it},$$
$$i = 1, 2.$$

The notional demand for labor by firm $i$ in period $t$, $n_{it}^{d}$, is given by the condition that the real wage equals the marginal product of labor:

$$(5) \quad (w_{it} - p_t) = u_t + ln(\gamma) - (1-\gamma) n_{it}^{d},$$
$$i = 1, 2,$$

or

$$(5') \quad n_{it}^{d} = \frac{1}{1-\gamma}\left[-(w_{it} - p_t) + u_t + ln(\gamma)\right],$$

where $w_{it}$ is the (log of the) wage received by group $i$ firms' workers in period $t$, and $p$ is the (log of the) price level.

**Notional Labor Supply**

The notional supply of labor to a firm is conditioned on the real wage rate:[1]

$$(6) \quad n_{it}^{s} = \beta_0 + \beta_1 (w_{it} - p_t),$$
$$\beta_1 \geq 0, \; i = 1, 2.$$

### Determination of Contract Wage

If the labor market cleared each period, fully reflecting the taste and technology conditions underlying notional labor supply and demand, $n^d_{it} = n^s_{it}$, then the employment level at firm $i$ in period $t$ would be

$$(7) \quad n^*_{it} = [\beta_0 + \beta_1 ln(\gamma)] M_0 + \beta_1 M_0 u_t$$

$$\text{where } M_0 \equiv [1+\beta_1(1-\gamma)]^{-1},$$

with $n^*_{it}$ denoting the market-clearing employment level. If wages were not sticky, but varied to clear the market, they would be

$$(8) \quad w^*_{it} = p_t + [ln(\gamma) - (1-\gamma)\beta_0] M_0 + M_0 u_t .$$

The contractual wage rate is the expectation of the rate that would clear the labor market. The contract wage for group $i$ is found by taking the expectation of (8) conditioned on information available in period $t$-$i$, when the contract was signed.

$$(9) \quad w_{it} = E_{t-i} p_t + [ln(\gamma) - (1-\gamma)\beta_0] M_0 + M_0 E_{t-i} u_t ,$$

where $E_{t-i}$ is the operator that conditions random variables on realizations at $t$-$i$ and earlier. Note that, in this formulation, the nominal wage will generally be different in each of the two periods subject to the contract.

Finally, let $u_t$ be a first-order autoregressive process,

$$(10) \quad u_t = \rho_1 u_{t-1} + \epsilon_t ,$$

$$0 < \rho_1 < 1, \ \epsilon_t \sim N(0, \sigma^2_\epsilon).$$

■ 1 The notional labor supply schedule could be derived from the primitive utility function:

$$c_0 + c_1 Y_t - c_2 N_t^{c_3}, \ c_1 > 0, \ c_2 > 0, \ c_3 > 1,$$

and the budget constraint:

$$Y_t = (W_t / P_t) N_t .$$

The first-order condition on $N$ is:

$$c_1 (W_t / P_t) = c_2 c_3 N_t^{(c_3 - 1)}.$$

Taking the natural logarithm and rearranging it, one obtains the labor supply function:

$$n^s_t = ln\,[c_1 / c_2 c_3] + \frac{w_t - p_t}{c_3 - 1},$$

which is the same as equation (6) of the text for $\beta_0 = ln\,[c_1 / c_2 c_3]$ and $\beta_1 = 1/(c_3 - 1)$. (Thanks to Charles Carlstrom for this argument.)

### Aggregate Commodity Supply

These elements are sufficient to specify the supply sector of the economy, under the assumption that labor input partly reflects the demand, and partly reflects the optimal level:

$$(11) \quad n_{it} = \phi n^d_{it} + (1-\phi) n^*_{it} .$$

The parameter $\phi$ represents the degree to which sticky wages cause misallocations, or employment distortions.

Using (3), (5′), (7), (9), (10), and (11), it can be shown that the (log of the) output of group one is

$$(12) \quad y_{1t} = \gamma A + M_2 \epsilon_t + M_1 \rho_1 \epsilon_{t-1} + M_1 \rho_1^2 u_{t-2} + G_1 (p_t - E_{t-1} p_t),$$

where

$$A \equiv [\beta_0 + \beta_1 ln(\gamma)] M_0$$

$$M_1 \equiv \frac{1 - \gamma M_0}{1 - \gamma} = (1 + \gamma \beta_1 M_0)$$

$$M_2 \equiv \frac{1 - \gamma M_0 (1 - \phi)}{1 - \gamma}$$

$$G_1 \equiv \frac{\gamma \phi}{1 - \gamma},$$

and the output of group two is

$$(13) \quad y_{2t} = \gamma A + M_2 \epsilon_t + M_2 \rho_1 \epsilon_{t-1} + M_1 \rho_1^2 u_{t-2} + G_1 (p_t - E_{t-2} p_t).$$

Total output for the economy is taken as the average of $y_{1t}$ and $y_{2t}$:

$$(14) \quad y_t = \gamma A + M_2 \epsilon_t + M_3 \rho_1 \epsilon_{t-1} + M_1 \rho_1^2 u_{t-2} + G_0 \sum_{i=1}^{2} (p_t - E_{t-i} p_t),$$

where

$$M_3 \equiv \frac{2(1 - \gamma M_0) + \phi \gamma M_0}{2(1 - \gamma)}$$

$$G_0 \equiv \frac{\gamma \phi}{2(1 - \gamma)} .$$

Equation (14) provides a characterization of the supply sector of the economy. It shows that output depends on productivity variations and on unanticipated inflation, both with coefficients that depend uniquely on the elasticity of output with respect to labor input, $\gamma$, the elasticity of notional labor supply, $\beta_1$, and the degree of misallocation, $\phi$. Higher $\beta_1$ values increase the responsiveness of output to productivity variations; the responsiveness of output to unanticipated inflation is proportional to $\phi$.

### Demand Sector

The demand sector of the model is a variant of the familiar IS-LM apparatus, introduced in Hoehn (1987). The main innovation is that goods demanders are allowed to update their inflation expectations in light of the current nominal interest rate and to revise their assessments of the real interest rate accordingly. Much complexity in solutions results from this innovation. The innovation is necessary if the authority's use of the information in the interest rate is to be studied without making the implausible assumption that the authorities know more (specifically, the current interest rate) than do other people. The innovation ensures that any influence monetary policy has over real variables does not arise from superior information.[2]

The commodity demand function, or IS curve, is

$$(15)\quad y_t^d = b_0 - b_1[R_t - (E^+_{t-1}p_{t+1} - p_t)] + x_t,$$
$$b_1 > 0,$$

$$(16)\quad x_t = \rho_2 x_{t-1} + \lambda_t,$$
$$0 < \rho_2 < 1,\ \lambda_t \sim N(0, \sigma^2_\lambda)$$

where

$$E^+_{t-1}p_{t+1} \equiv E[p_{t+1} | \Omega_t],$$

$$\Omega_t \equiv \text{observable state of economy at time } t$$
$$\equiv \{R_t ; S_{t-1}\},$$

and $S \equiv$ state vector (given a specific identity in the next section). The nominal interest rate, $R_t$, is measured as the natural logarithm of unity plus the coupon rate of return. The future price expectation, $E^+_{t-1}p_{t+1}$, is conditioned on the observed state of the economy, $\Omega_t$, an information set that includes the current economywide interest rate, $R_t$, and the lagged state vector, $S_{t-1}$. $E^+_{t-1}p_{t+1}$ can differ from $E_{t-1}p_{t+1}$ because people use the current nominal interest rate to update their inflation expectations. $x_t$ is a stochastic demand shock.

The money-demand function is conventional:

$$(17)\quad m^d_t - p_t = a_0 - a_1 R_t + a_2 y_t + v_t,$$

$$(18)\quad v_t = \rho_3 v_{t-1} + \eta_t,$$
$$0 < \rho_3 < 1,\ \eta_t \sim N(0, \sigma^2_\eta),$$

where $v_t$ is the log of the quantity of money and $v_t$ is a first-order autoregressive random disturbance.

### Policy Sector

Given the model, a policy rule that is adequate for the policy targets and criteria to be considered, is

$$(19)\quad m^s_t = qR_t + \mu_0 + \mu_1 u_{t-1}$$
$$+ \mu_2 v_{t-1} + \mu_3 x_{t-1} + \mu_4 E_{t-2} u_{t-1}.$$

### Harberger Welfare Metric

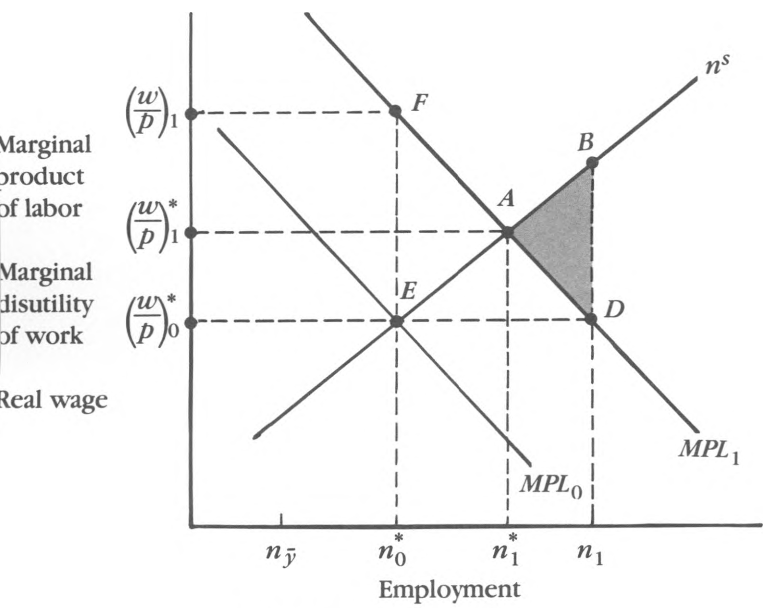
The loss function measures a representative individual's frustration in obtaining an optimal allocation of time between labor and leisure, as productivity and demand conditions change. The method, due to Harberger (1971), of measuring individual frustrations uses the labor supply and demand curves, assuming that they accurately reflect preferences and thereby show how workers and firms would want to adjust output and employment in response to changing productive opportunities. Equilibrium between notional supply and demand is then supposed to be optimal. Equilibrium values of output and employment in this log-linear model are a strict log-linear function of $u_t$, as shown in equation (7). The welfare loss is taken as proportional to the square of the deviation of the actual from the optimal employment level. This welfare-loss metric is proportional to the area of the familiar Harberger welfare-loss triangles, as shown in the figure of the next section.

In the model with two staggered contracting firm groups, an approximate measure of the expected Harberger welfare loss over the span of a contract is

■ 2 The effect of allowing goods demanders to extract information about inflation from the nominal interest rate was analyzed extensively in Hoehn (1987). It can reverse the usual effects of money supply or demand shocks on the price level and output during the temporary period before shocks become fully known to all. For example, output and prices may temporarily rise in response to an increase in money demand. But such cases arise only in cases of extreme policies, such as crude attempts to smooth interest rates by expanding money greatly in response to a rise in the interest rate, or where structural parameters or relative variances of shocks take on extreme values.

FIGURE 1

**Optional Employment Equates the Marginal Product of Labor With the Marginal Disutility of Work**




SOURCE: Author's calculations.

$$(20)\quad \text{Expected Welfare Loss} = E(n_{1t} - n^*_{1t})^2 + E(n_{2t} - n^*_{2t})^2,$$

where the $n_{it}$ are actual employment levels and the $n^*_{it}$ are the market-clearing employment levels of equation (7). This measure is the sum of the variances of employment from optimal for each of the two periods of any contracting firm, during which it will first be a group-one firm, and then a group-two firm.

## III. How Policy Can Minimize Employment Distortions

To understand how a well-chosen policy rule can improve welfare, it is useful to examine the nature of the money-supply responses to various shocks that would fully prevent employment distortions. Such a degree of success is not possible in reality because of policymaker uncertainty about shocks. In the model simulations, it is assumed that the authorities know the structure of the economy, the current interest rate, and the lagged state of the economy; the authorities do not have full information about current shocks. This complicates analysis, motivating a heuristic treatment of the simpler case in which the authorities know the full state and can change the money supply continuously to keep employment for both groups of firms at the ideal level. Readers interested in the final-form solution and the optimal policy rule in the full model may find them available in Hoehn (1989).

The optimal employment level for each group, $n^*_0$, is determined by the intersection of the marginal product of labor schedule, $MPL_0$, and the labor supply or marginal disutility of work schedule, $n^s$, as shown in the figure. This employment level will be chosen by firms only if the real wage is equal to $(w/p)^*_0$. (This statement holds true for any degree of misallocation, $\phi$, except zero, in which case nominal wage stickiness cannot create employment distortion. The case illustrated here is the simple case of pure demand-determination of employment, $\phi = 1$. Of course, the size of employment distortions will be smaller if $\phi$ is a fraction.)

The optimal employment level and the real wage that will induce firms to choose the optimal employment level vary with autonomous labor productivity shocks. For example, a cyclical improvement in labor productivity raises the optimal employment level and the associated real wage. The figure illustrates this with a shift in the marginal product of labor schedule from $MPL_0$ to $MPL_1$, which raises the optimal employment level to $n^*_1$. This optimal level will be chosen by firms if the real wage rises to $(w/p)^*_1$.

The productivity shock case reveals the suboptimality of a price-stabilization policy. Because nominal wages are fixed during the contract interval, stable prices imply that the real wage would remain at the initial level of $(w/p)^*_0$. Firms would choose the employment level $n_1$, at which the marginal product equals the unchanged real wage. The expansion of employment from $n^*_0$ to $n_1$ is an excessive response to the improvement in productivity, because the marginal disutility of work exceeds the marginal product of labor for employment levels above $n^*_1$. The Harberger welfare loss triangle is BAD.

To prevent firms from overexpansion, the monetary authorities should allow the price level to fall by enough to raise the real wage to $(w/p)^*_1$. Somewhat ironically, this policy will involve an expansion in the money supply. If the money stock were unchanged, the price level would fall too much as output rose. For example, if the velocity of money were constant and the quantity of money were constant, then a productivity improvement would raise the marginal product of labor and—via deflation—raise the real wage by the same amount, to $(w/p)_1$, leaving the profit-maximizing level of employment at $n_0^*$. The labor market is then at point $F$ in the figure, with welfare loss triangle EFA. The optimal policy response to the productivity shock is to expand the money supply enough to moderate the deflation, so that real wages rise to $(w/p)_1^*$, but no further.

The shift from point $E$ to point $F$ in response to the productivity improvement will always be obtained under a nominal income target, because that shift lowers the price level and raises the output level by the same proportion, leaving their product unchanged. In the simulations with the IS-LM demand apparatus, the velocity of money falls with favorable productivity shocks. Consequently, the nominal income target will necessarily require increases in money to obtain point $F$. If the increase in money is not forthcoming, as under a constant-money policy, the price level will fall more than one-for-one with the productivity improvement, and the profit-maximizing employment level falls below $n_0^*$. The welfare loss resulting from sticky wages under a productivity shift is greater under a constant money policy than under the nominal income target, once velocity changes are accounted for.

The optimal policy response to a commodity-demand or money-demand shock is easier to understand than the optimal response to a productivity shock. In the model as specified, such shocks do not alter either the marginal product of labor schedule or the marginal disutility of work. Consequently, the optimal level of employment is unchanged. The optimal policy will attempt to prevent the employment level from changing with demand and money shocks. Employment can be insulated from distortions arising from such shocks by a policy that stabilizes the price level. A stable price level prevents the real wage from changing, preventing firms from desiring a change in employment. Money supply should be decreased with increased commodity demand by an amount adequate to prevent inflation. Money supply should be increased one-for-one with increases in the money-demand function.

A policy of output stabilization is unambiguously worse than a policy of price stabilization. Both of these policies give an appropriate response to commodity-demand and money-demand shocks, but the distortion concurrent with a productivity shock is unambiguously larger under the output stabilization policy. As soon as a single-minded output-stabilizing authority observes a productivity improvement, it will deflate the price level by reducing the money supply. The result is deflation sufficient to drive the real wage above $(w/p)_1$, and employment declines below $n_0^*$, say to $n_{\bar{y}}$.

The ability of the authority to stabilize output in this example is limited because recontracting firms can offset the real-wage effects of excessive deflation by lowering nominal wages. As soon as one of the groups recontracts, it will reduce wages to aim at an increased employment level, driving the authorities to further reduce employment in the second group via yet more deflation. The second group cannot protect itself against the negative employment distortions by recontracting for lower nominal wages until one more period passes and the old contract expires. The second group's employment must be reduced, if output is to be stabilized, by enough to offset not only the economywide increase in productivity, but must also offset the increase in employment at the recontracting firms, who will rationally anticipate deflation and reduce wages to allow employment to increase to the optimal employment level. Because the loss function is the sum of squared group employment distortions, the concentration of the employment distortion in the second group of firms leads to a sizeable welfare loss.

## IV. A Numerical Simulation

In order to illustrate how various policy rules influence employment distortions arising from sticky wages, a simulation can be conducted with particular numerical values for structural parameters. The values chosen for this simulation were the following:

$$(21)\quad \beta_1 = 1/2 \quad \gamma = 1/2 \quad \phi = 1$$
$$a_1 = 2 \quad a_2 = 2/3 \quad b_1 = 1$$
$$\sigma_\epsilon^2 = 1 \quad \sigma_\lambda^2 = 2 \quad \sigma_\eta^2 = 5 \quad \rho_i = 4/5,\ i = 1,2,3.$$

The elasticity of labor supply with respect to the cyclical variations in the real wage was set at one-half, an arbitrary but plausible value. The elasticity of output with respect to labor input, $\gamma$, was set at the midpoint of its permissible range, also arbitrary but plausible. The value assigned to the money demand elasticity with respect to the nominal interest rate, $a_1$, implies, for example, that an increase in the rate from 5 to 6 percent would, for given levels of income and prices, lower real money demand by approximately 1.9 percent. The money-demand elasticity with respect to output, $a_2$, was set at somewhat less than unity, as suggested by

## TABLE 1

### Alternative Policy Rules ($\phi = 1$)

| Policy Parameter[a] | Policy Criterion: Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| $q$ | 0.0 | +1.06 | –0.05 | +0.62 | +0.48 |
| $\mu_1$ | 0.0 | –1.97 | +3.35 | +1.73 | +2.06 |
| $\mu_2$ | 0.0 | –2.44 | –1.56 | –2.10 | –1.98 |
| $\mu_3$ | 0.0 | +0.80 | +0.80 | +0.80 | +0.80 |
| $\mu_4$ | 0.0 | b | –0.84 | –0.04 | b |

a. The money supply rule is $m_t = qR_t + \mu_1 u_{t-1} + \mu_2 x_{t-1} + \mu_3 v_{t-1} + \mu_4 E_{t-2} u_{t-1}$, where $u$, $x$, and $v$ are disturbances to goods demand, goods supply, and money demand.
b. The policy parameter $\mu_4$ is irrelevant to the criterion. In simulations, $\mu_4$ is set to zero.
SOURCE: Author's calculations.

## TABLE 2

### Response of Money to Innovations ($\phi = 1$)

| Innovation | Policy Criterion: Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| Productivity | | | | | |
| $t$ | 0.0 | –0.11 | +0.01 | –0.06 | –0.06 |
| $t-1$ | 0.0 | –2.17 | +3.41 | +1.25 | +1.33 |
| $t-2$ | 0.0 | –1.81 | +2.05 | +0.98 | +1.06 |
| Goods Demand | | | | | |
| $t$ | 0.0 | +0.25 | –0.01 | +0.15 | +0.12 |
| $t-1$ | 0.0 | –1.60 | –1.60 | –1.60 | –1.60 |
| $t-2$ | 0.0 | –1.28 | –1.28 | –1.28 | –1.28 |
| Money Demand | | | | | |
| $t$ | 0.0 | +0.31 | –0.02 | +0.19 | +0.15 |
| $t-1$ | 0.0 | +0.80 | +0.80 | +0.80 | +0.80 |
| $t-2$ | 0.0 | +0.64 | +0.64 | +0.64 | +0.64 |

SOURCE: Author's calculations.

abstract analysis of the transactions demand for money. The commodity-demand elasticity with respect to the real interest rate, $b_1$, was set to unity because, of all (equally arbitrary) values, unity is the most straightforward choice. (Econometric evidence currently available does not provide direct knowledge of this elasticity.) The relative sizes of the disturbances give considerable scope to demand-side influences on output and employment, and allow for a relatively unstable money-demand function.

In the basic simulation, firms were assumed to choose employment to equate the marginal product of labor with the real wages, so $\phi = 1$. In a second simulation, $\phi$ was set equal to one-third, in order to see whether the results of the basic simulation were robust with respect to this parameter.

Five different policy rules were simulated, with their response coefficients chosen so as to target (1) money, (2) output, (3) the price level, (4) nominal income, or (5) optimal employment. The last of these is, of course, the only optimal policy by the criterion employed, but it is instructive to compare results of other potential targets.

The policy rules' response coefficients, $q$ and the $\mu_i$, are displayed in table 1. The final-form solution for the money supply is determined by both these coefficients and the solution for the nominal interest rate (because of the $qR_t$ term in the money supply rule), and is shown in table 2

**TABLE 3**

**Welfare Losses Under Alternative Policies ($\phi$ = 1)**

| Loss due to shocks to: | Policy Criterion: Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| Productivity | 2.86 | 14.99 | 1.80 | 0.38 | 0.19 |
| Goods demand | 6.91 | 3.04 | 1.96 | 2.47 | 2.35 |
| Money demand | 4.52 | 0.24 | 1.28 | 0.66 | 0.76 |
| TOTAL LOSS | 14.29 | 18.27 | 5.04 | 3.51 | 3.30 |

SOURCE: Author's calculations.

for each of the five alternative policies. In the immediate period of impact, the monetary authority's response to a shock is equal to $q$, its interest rate response coefficient, times the response of the interest rate to the shock. For example, under a policy of stabilizing output, the money supply is increased 1.06 for each one-point change in the interest rate. A productivity shock in period $t$ reduces the interest rate by –0.10 (not shown in tables) under this policy rule, so the response of money at time $t$ to a productivity shock in period $t$ is 1.06 times –0.10, or about –0.11.

Only after one period has passed can the monetary authority observe all three shocks independently and tailor its response to each one separately. For example, the output-stabilizing policy contracts the money supply by 2.17 at time $t$ for a one-unit innovation to productivity in the previous period, $\epsilon_{t-1}$. This response reflects two channels: first, an indirect channel involving the change in the interest rate, –0.19, times the response coefficient $q$ = 1.06, or about –0.20. To this is added the direct response coefficient on $t-1$ productivity, $\mu_1$ = –1.97. Together, these add to –2.17, the total contraction of the money supply required to prevent period–$t$ output from responding to period $t-1$ productivity innovations. A similar calculation involving direct and indirect effects finds that the output-stabilizing policy contracts the money supply at time $t$ by 1.81 in response to a unit productivity innovation in period $t-2$.

Aside from the constant-money policy, the policies considered are identical in their money-supply responses to goods demand or money demand shocks, once these shocks are observed. In this model, all the activist targets are essentially equivalent in terms of the implied response of the money supply to these demand-side shocks.

The main difference among the active money-supply policies lies in the response of money to productivity shocks. The output-stabilizing policy's response is too restrictive; it contracts money at time $t$ by 2.17 after a unit productivity innovation in period $t-1$, contrasting with an optimal increase of 1.33. The price-stabilization rule responds too expansively; it expands the money supply by 3.41. The nominal income target's response is to expand the money supply by 1.25, very close to optimal. These differences among alternative active policies in their response to productivity shocks account for the relative rankings of their efficiency.

Expected welfare losses under alternative policies, shown in table 3, are the sum of the mean squared deviations of group one and group two employment levels from optimal employment levels. Given the information constraint the authority faces, it can reduce this loss measure to 3.30 using the optimal policy. Most of this loss, 2.35, is attributable to goods-demand shocks occurring in the current period; a small fraction is attributable to productivity shocks occurring in the current period. Distortions due to shocks in period $t-1$ can be completely eliminated by policy responses, while distortions due to $t-2$ or earlier shocks are eliminated by wage recontracting by both groups of firms.

The nominal income targeting policy is close to optimal; its welfare loss is 3.51, only slightly higher than for the optimal policy. The output-stabilizing policy is far worse, with a total expected loss of 18.27, most of which is due to productivity shocks. The constant-money policy is not much better than the output-stabilizing policy; it generates substantial employment distortions in the face of goods-demand and money-demand shocks, which the activist policies make active efforts to prevent. Finally, the price-stabilization policy results in somewhat greater losses than the nominal income policy, but results in much smaller losses than the output or money targeting policies.

TABLE 4

**Response of $(n_{2t} - n^*_{2t})$ to Innovations $(\phi = 1)$**

| Innovation | Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| | Policy Criterion | | | | |
| Productivity | | | | | |
| $t$ | −0.28 | −0.34 | −0.28 | −0.32 | −0.30 |
| $t-1$ | −1.64 | −3.84 | +1.28 | −0.42 | 0.0 |
| $t-2$ | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Goods Demand | | | | | |
| $t$ | +0.70 | +0.88 | +0.70 | +0.78 | +0.76 |
| $t-1$ | +1.58 | 0.0 | 0.0 | 0.0 | 0.0 |
| $t-2$ | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Money Demand | | | | | |
| $t$ | −0.38 | −0.16 | −0.36 | −0.26 | −0.28 |
| $t-1$ | −0.80 | 0.0 | 0.0 | 0.0 | 0.0 |
| $t-2$ | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

SOURCE: Author's calculations.

TABLE 5

**Welfare Losses Under Alternative Policies With $\phi = 1/3$**

| Loss due to shocks to: | Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| | Policy Criterion | | | | |
| Productivity | 0.51 | 1.03 | 0.21 | 0.05 | 0.04 |
| Goods demand | 1.37 | 0.60 | 0.44 | 0.54 | 0.54 |
| Money demand | 0.83 | 0.14 | 0.32 | 0.22 | 0.22 |
| TOTAL LOSS | 2.72 | 1.78 | 0.98 | 0.80 | 0.79 |

SOURCE: Author's calculations.

The deviations of employment from optimal for the two groups can be read from table 4. The table lists the deviations for the second group; the deviations for the first group, $(n_{1t} - n^*_{1t})$, are the same as for the second group for period-$t$ shocks, but recontracting by this group makes the period $-t$ employment distortion equal to zero for $t-1$ or earlier shocks. A one-unit innovation in productivity at time $t$ raises the optimal employment level for both groups by 0.40 in time $t$. Given that the effect of an innovation on the marginal productivity schedule decays at the rate $\rho_1 = .8$, optimal employment increases by 0.32 and by about 0.26 in response to unit productivity innovations in periods $t-1$ and $t-2$.

The gross suboptimality of the output-stabilizing policy reflects the employment distortion in the second, nonrecontracting, group, in response to a productivity innovation in period $t-1$. Because policy responds by contracting the money supply, generating deflation and an excessive rise in the real wage for the nonrecontracting group, employment for that group falls by 3.52, in sharp contrast with the increase of 0.32 in optimal employment. The distortion is then −3.84. In order to keep output fixed, the authorities must reduce employment in the second group, and this reduction must be enough to offset both the economywide productivity improvement and the rise in employment by 0.32 in the first, recontracting, group.

The GNP targeting policy is very close to

**TABLE 6**

**Relation Between Money and Output Under Alternative Policies ($\phi$ = 1)**

| Covariation due to shocks to: | Policy Criterion: Money | Output | Price Level | Nominal Income | Optimal |
|---|---|---|---|---|---|
| Productivity | 0.0 | –1.51 | +8.76 | +3.10 | +3.48 |
| Goods demand | 0.0 | +0.22 | –0.01 | +0.12 | +0.09 |
| Money demand | 0.0 | –1.95 | +0.02 | –0.12 | –0.11 |
| CORRELATION | — | –0.30 | +0.59 | +0.30 | +0.34 |

SOURCE: Author's calculations.

optimal. It handles money-demand and commodity-demand variations appropriately, and generates a mild and nearly optimal deflation in response to productivity improvements. The degree of closeness to optimality depends on various parameters, but is not, it appears, sensitive to the degree to which sticky wages cause misallocations, $\phi$ , at least at the chosen values of the other structural parameters. Table 5 shows the welfare losses in the model for $\phi = 1/3$.

The output targeting policy is generally the worse in terms of employment distortion (except when $\phi = 1/3$, when the constant-money policy is worse). The output targeting policy generates the greatest losses when productivity shocks occur. Output targets handle commodity- and money-demand shocks, however, in an appropriate manner.

The price-stabilization policy results in overemployment when a productivity improvement occurs. The policy is too stimulative; it does not provide for the deflation required to raise the real wage in line with marginal productivity at the new optimal employment level. In the case of commodity- and money-demand shocks, however, a policy of price stabilization provides essentially the same optimal response as does the nominal and real GNP targets.

The constant-money policy accrues losses in the case of all kinds of shocks. The loss attending productivity shocks is less than in the case of the output target, but the money-targeting policy fails to respond appropriately to commodity- or money-demand shocks. In the simulation, the constant-money policy results in less employment distortion than the output-stabilizing policy, unless the degree of misallocation is small, such as $\phi = 1/3$.

## V. Conclusion

**A** monetary policy that seeks to aid wage contractors in avoiding employment distortions due to sticky wages will attempt to keep the real wage equal to the marginal disutility of labor in all states of the economy. Such a policy will require money supply expansion when cyclical improvements in labor productivity occur. To the extent that productivity variations are an important factor in the business cycle, the optimal money supply rule will involve a positive correlation between money and output. (See table 6.) Hence, the belief, common among economists, that sticky-wage models argue for a countercyclical or output-stabilizing policy is not necessarily correct, once productivity shocks are taken account of.

In simulations, it was found that a nominal income target might be reasonably close to the optimal policy. This result is useful because the Federal Reserve may not be able to predict and target optimal employment levels because of uncertainty about the structural parameters and shock variances needed in a welfare analysis, yet can probably predict and target nominal income using its models and judgmental forecasters. After all, the main objective of macroeconometric models has been the prediction and potential control of national income. The analysis of this paper tends to give additional justification to proposals for nominal income targeting, including those by Meade (1978), Tobin (1980), Hall (1983), Gordon (1985), and McCallum (1987).

The relative near-optimality of a nominal income target might not be robust to all conceivable values of the labor market parameters, $\gamma$ and $\beta_1$, however. For example, if the marginal product of labor curve declines steeply ($\gamma$ close to zero), and/or if the notional labor supply curve is nearly horizontal ($\beta_1$ very large), then a price target will do as well or better than a nominal income target. More precisely, if $M_0 = [1 + \beta_1(1 - \gamma)]^{-1}$ is close to unity, then a nominal income target will be close to optimal, but if $M_0$ is close to zero, then a price level target will be close to optimal.[3] In the simulation, $\gamma = 1/2$ and $\beta_1 = 1/2$, so $M_0 = .8$, which is rather close to unity. In order to adequately confirm the relative efficiency of a nominal income target relative to a price target, econometric evidence and a sensitivity analysis are needed to rule out small values of $M_0$. In general, the optimal policy response to a productivity improvement will be one that is less stimulative than that implied by a price target and more stimulative than that implied by a nominal income target.

If the specification of the model were modified to allow for costs of changing commodity prices ("menu costs"), or to allow for some degree of commodity price stickiness, then a price-targeting policy might yet be better than a nominal income target. Many other elements of more detailed macroeconometric models have unknown implications for the welfare analysis. Much more research along these lines is needed for an adequate welfare analysis of monetary policy toward the business cycle.

## Glossary of Variables and Parameters

### Endogenous Variables

- $y$ output
- $y_1$ output of group 1 firms
- $y_2$ output of group 2 firms
- $p$ price level
- $R$ nominal interest rate
- $m$ money stock
- $w$ wage rate
- $w^*$ market-clearing wage rate
- $n$ employment
- $n_1$ employment of group 1 firms
- $n_2$ employment of group 2 firms
- $n^*$ optimal employment level

### Exogenous Variables

- $\epsilon$ innovation to the productivity disturbance, $u$
- $\lambda$ innovation to the commodity-demand disturbance, $x$
- $\eta$ innovation to the money-demand disturbance, $v$

### State Vector

$$S_t \equiv \{\epsilon_t, u_{t-1}, E_{t-2}u_{t-1}; \lambda_t, x_{t-1}, E_{t-2}x_{t-1}; \eta_t, v_{t-1}, E_{t-2}v_{t-1}\}$$

### Information Set, or Observed State

$$\Omega_t \equiv \{R_t; u_{t-1}, E_{t-2}u_{t-1}; x_{t-1}, E_{t-2}x_{t-1}; v_{t-1}, E_{t-2}v_{t-1}\}$$

### Parameters

All nonpolicy parameters are nonnegative.

- $a_1$ = elasticity of money demand with respect to interest rate = $d\ln(M/P)/d\ln(1+R)$
- $a_2$ = elasticity of money demand with respect to output
- $b_1$ = elasticity of aggregate demand with respect to real interest rate
- $\beta_1$ = elasticity of notional labor supply with respect to real wage
- $\gamma$ = elasticity of output with respect to labor input
- $q$ = coefficient of money-supply response to interest rate
- $\mu_i$ = coefficients of money-supply response to lagged state variables (see equation 19 of the text)
- $\sigma_\epsilon^2$ = variance of productivity innovation
- $\sigma_\lambda^2$ = variance of commodity-demand innovation
- $\sigma_\eta^2$ = variance of money-demand innovation

■ 3 Bean (1983) apparently was the first to note this.

### References

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

**Bean, Charles R.**, "Targeting Nominal Income: An Appraisal," *Economic Journal,* December 1983, *93,* 806–19.

**Fischer, Stanley**, "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule," *Journal of Political Economy,* February 1977, *85,* 191–205.

**Gordon, Robert J.**, "The Conduct of Domestic Monetary Policy," in Albert Ando, Hidekazu Eguchi, Roger Farmer, and Yoshio Suzuki, eds., *Monetary Policy in Our Times,* Cambridge, Mass.: The MIT Press, 1985, 45–81.

**Hall, Robert**, "Employment Fluctuations and Wage Rigidity." *Brookings Papers on Economic Activity,* Washington, D.C.: The Brookings Institution, 1980.

_______ , "Macroeconomic Policy under Structural Change," *Industrial Change and Public Policy,* Federal Reserve Bank of Kansas City, 1983, 102–3.

**Harberger, Arthur C.**, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay," *Journal of Economic Literature,* September 1971, *9,* 785–97.

**Hoehn, James G.**, "Monetary Policy Under Rational Expectations With Multiperiod Wage Stickiness and an Economy-Wide Credit Market," *Working Paper* 8716, Federal Reserve Bank of Cleveland, December 1987.

_______ , "Procyclical Real Wages Under Nominal-Wage Contracts with Productivity Variations," *Economic Review,* Federal Reserve Bank of Cleveland, 1988 Quarter 4, *24,* 11–23.

_______ , "A Welfare Analysis of Monetary Policy with Nominal Wage Contracts and Demand and Productivity Shocks," Federal Reserve Bank of Cleveland, unpublished working paper, available from the Bank on request (1989).

**McCallum, Bennett T.**, "The Case for Rules in the Conduct of Monetary Policy," *Weltwirtshaftliches Archiv,* 1987, *123,* 415–29.

**Meade, J.E.**, "The Meaning of Internal Balance," *Economic Journal,* September 1978, *88,* 423–35.

**Taylor, John B.**, "Staggered Wage Setting in a Macro Model," *American Economic Review: Papers and Proceedings,* May 1979, 108–13.

**Tobin, James**, "Stabilization Policy Ten Years After." *Brookings Papers on Economic Activity,* Washington, D.C.: The Brookings Institution, 1980, *1,* 19–72.

# Recent Developments in Macroeconomics

## Conference Proceedings Offered

The papers in this special issue of the *Journal of Money, Credit, and Banking* were presented and discussed at a conference on "Recent Developments in Macroeconomics" held at the Federal Reserve Bank of Cleveland. This conference was organized to discuss the practical aspects of recent developments in macroeconomic research and to discuss the relevance of these developments for economic policy.

## Uses

- Identification of Current Issues
- Classroom Assignments
- Literature Review

## Contents


- **Recent Developments in Macroeconomics: A Very Quick Refresher Course**
  N. Gregory Mankiw
  Comment by Herbert Stein
  Comment by Edmund Phelps
- **Postwar Developments in Business Cycle Theory: A Moderately Classical Perspective**
  Bennett T. McCallum
  Comment by Lawrence Summers
  Comment by Peter K. Clark
- **What Microeconomics Teaches Us About the Dynamic Macro Effects of Fiscal Policy**
  Laurence J. Kotlikoff
  Comment by William Niskanen
  Comment by Preston Miller
- **Some Recent Developments in Labor Economics and Their Implications for Macroeconomics**
  Lawrence F. Katz
  Comment by Robert Topel
  Comment by Thomas Kniesner
- **On the Roles of International Financial Markets and Their Relevance for Economic Policy**
  Alan C. Stockman
  Comment by Patrick Kehoe
  Comment by J. David Germany
- **Financial Structure and Aggregate Economic Activity: An Overview**
  Mark Gertler
  Comment by Marvin Goodfriend
  Comment by Laurence Weiss


For a limited time, you may order a copy of the conference proceedings at the reduced price of $5.00 (regularly $10.00). Consider bulk orders! Please send $5.00 (U.S.) in a check or money order drawn on a U.S. bank. Make checks payable to the Federal Reserve Bank of Cleveland. Complete and detach the form and mail to:

Federal Reserve Bank of Cleveland
Research Department
P.O. Box 6387
Cleveland, Ohio 44101

Please send the Macroeconomics Conference proceedings to:

Name ____________________

Address ____________________

____________________

City ____________________

State ____________________ Zip Code

Country ____________________

# Second Quarter Working Papers

## Working Paper Notice

Current *Working Papers* of the Cleveland Federal Reserve Bank are listed in each quarterly issue of the *Economic Review*. Copies of specific papers may be requested by completing and mailing the attached form below.

Single copies of individual papers will be sent free of charge to those who request them. A mailing list service for personal subscribers, however, is not available.

Institutional subscribers, such as libraries and other organizations, will be placed on a mailing list upon request and will automatically receive *Working Papers* as they are published.

■ **8904**
**Dedicated Taxes and Rent Capture by Public Employees**
by Brian A. Cromwell

■ **8905**
**Modeling Large Commercial-Bank Failures: A Simultaneous-Equation Analysis**
by Aslı Demirgüç-Kunt

■ **8906**
**Do Wage Differences Among Employers Last?**
by Erica L. Groshen

■ **8907**
**The Structure of Supervision and Pay in Hospitals**
by Erica L. Groshen and Alan B. Krueger

Please complete and detach the form below and mail to:

Federal Reserve Bank of Cleveland
Research Department
P.O. Box 6387
Cleveland, Ohio 44101

**Check item(s) requested.**

Please send the following Working Paper(s):

☐ **8904** ☐ **8906**

☐ **8905** ☐ **8907**

**Send to:**

Please print

Name

Address