An Error-Correction Model of U.S. M2 Demand

Yash P. Mehra

Much applied research in monetary economics has been devoted to the specification of the money demand function. Money demand specification has important policy implications. A poorly specified money demand function could yield, for example, spurious inferences on the underlying stability of money demand—a consideration of central importance in the formulation of monetary policy.

This paper is concerned with one aspect of money demand specification, namely, the choice of the form in which variables enter the money demand function. It is common to specify the money demand function either in log-level form or in log-difference form. The log-level form, popularized by Goldfeld's (1974) work, has often been criticized on the ground that the levels of many economic variables included in money demand functions are nonstationary. Therefore, the regression equations that relate such variables could be subject to "the spurious regression phenomenon" first described in Granger and Newbold (1974). This phenomenon, later formalized in Phillips (1986), refers to the possibility that ordinary least-squares parameter estimates in such regressions do not converge to constants and that the usual t- and F-ratio test statistics do not have even the limiting distributions. Their use in that case generates spurious inferences. In view of these considerations, many analysts now routinely specify the money demand functions in first-difference form.

Quite recently, the appropriateness of even the first-difference specification has been questioned. In particular, if the levels of the nonstationary variables included in money demand functions are cointegrated as discussed in Engle and Granger (1987),[1] then such regressions should not be estimated in first-difference form. This is because level regressions which relate the cointegrated variables can be consistently estimated by ordinary least-squares without being subject to the spurious regression phenomenon described above.[2] One implication of this work is that money demand functions estimated in first-difference form may be misspecified because such regressions ignore relationships that exist among the levels of the variables.

Since there are potential problems with money demand functions specified either in level or in first-difference form, some analysts have recently begun to integrate these two specifications using the theories of error-correction and cointegration. In this approach, a long-run equilibrium money demand model (cointegrating regression) is first fit to the levels of the variables, and the calculated residuals from that model are used in an error-correction model which specifies the system's short-run dynamics.[3] Such an approach permits both the levels and first-differences of the nonstationary variables to enter the money demand function. This approach also makes it easier to distinguish between the short- and long-run money demand functions. Thus, some variables that are included in the short-run part of the model might not be included in the long-run part and vice versa, thereby permitting considerable flexibility in the specification of the money demand function.

This paper illustrates the use of the above approach by presenting and estimating an error-correction model of U.S. demand for money (M2) in the postwar period. The money demand function presented here exhibits parameter stability. Money growth forecasts generated by this function are

---

[1] Let X_{1t}, X_{2t}, and X_{3t} be three time series. Assume that the levels of these time series are nonstationary but first differences are not. Then these series are said to be cointegrated if there exists a vector of constants (α_1, α_2, α_3) such that $Z_t = \alpha_1 X_{1t} + \alpha_2 X_{2t} + \alpha_3 X_{3t}$ is stationary. The intuition behind this definition is that even if each time series is nonstationary, there might exist linear combinations of such time series that are stationary. In that case, multiple time series are said to be cointegrated and share some common stochastic trends. We can interpret the presence of cointegration to imply that long-run movements in these multiple time series are related to each other.

[2] The usual t- and F-ratio statistics can be used provided some other conditions are satisfied and other adjustments are made. See Phillips (1986) and West (1988).

[3] This approach, popularized by Hendry and Richard (1982) and Hendry, Pagan and Sargan (1983), has been applied to study U.K. money demand behavior by Hendry and Ericsson (1990) and U.S. money demand behavior by Small and Porter (1989) and Baum and Furno (1990).

consistent with the actual behavior of M2 growth during the last two decades or so. A key feature of the results presented here is that consumer spending is found to be a better short-run scale variable than real GNP, even though it is the latter that enters the long-run part of the model.[4]

The plan of this paper is as follows. Section 1 presents the error-correction model and discusses the issues that arise in the estimation of such models. Section 2 presents the empirical results. The summary observations are stated in Section 3.

I. An Error-Correction Money Demand Model: Specification and Estimation

Specification of an M2 Demand Model

The error-correction money demand model has two parts. The first is a long-run equilibrium money demand function

$$rM2_t = a_0 + a_1\, rY_t - a_2\, (R\text{-}RM2)_t + U_t \quad (1)$$

where all variables are expressed in their natural logarithms and where rM2 is real M2 balances; rY, real GNP; R, a short-term nominal rate of interest; RM2, the own rate of return on M2; and U, the long-run random disturbance term. Equation 1 says that the public's demand for real M2 balances depends upon a scale variable measured by real GNP and an opportunity cost variable measured as the differential between the nominal rate of interest and the own rate of return on M2. The parameters a_1 and a_2 measure respectively the long-run income and opportunity cost elasticities. A key aspect of the specification used here is that the own rate of return on M2 is relevant in determining M2 demand (Small and Porter, 1989, and Hetzel and Mehra, 1989). The conventional specification usually omits this variable (see, for example, Baum and Furno, 1990).

The second part of the model is a dynamic error-correction equation of the form

$$\Delta rM2_t = b_0 + \sum_{s=1}^{n1} b_{1s}\, \Delta rM2_{t-s} + \sum_{s=0}^{n2} b_{2s}\, \Delta rY_{t-s} - \sum_{s=0}^{n3} b_{3s}\, \Delta(R - RM2)_{t-s} + \lambda\, U_{t-1} + \epsilon_t \quad (2)$$

where all variables are as defined above and where ϵ_t is the short-run random disturbance term; Δ, the first difference operator; $n_i (i = 1,2,3)$, the number of lags; and U_{t-1}, the lagged value of the long-run random disturbance term. Equation 2 gives the short-run determinants of M2 demand, which include, among others, current and past changes in the scale and opportunity cost variables and the lagged value of the residual from the long-run money demand function. The parameter λ that appears on U_{t-1} in (2) is the error-correction coefficient. At a more intuitive level, the presence of U_{t-1} in (2) reflects the presumption that actual M2 balances do not always equal what the public wishes to hold on the basis of the long-run factors specified in (1). Therefore, in the short run, the public adjusts its money balances to correct any disequilibrium in its long-run money holdings. The parameter λ in (2) measures the role such disequilibria plays in explaining the short-run movements in money balances.[5]

---

[4] The results presented here are in line with those given in Small and Porter (1989) but differ from those given in Baum and Furno (1990). The error-correction model of M2 demand reported in Baum and Furno does not exhibit parameter stability. One possible reason for this is the use of inappropriate scale and/or opportunity cost variables. The money demand function reported in Baum and Furno measures the opportunity cost variable by a short-term market rate of interest, thereby implicitly assuming that the own rate of return on M2 is zero. Furthermore, real GNP is used in the long-run as well as in the short-run part of the model.

[5] It should, however, be pointed out that the size of the coefficient on the error correction variable in (2) is influenced in part by the nature of serial correlation in the random disturbance term of the long-run money demand model and is not necessarily indicative of the speed of adjustment of money demand to its long-run level. To explain it further, for illustrative purposes assume the restricted simple money demand model of the form

$$m^*_t = a_0 + a_1\, y_t + U_t \quad (a)$$

where changes in money balances follow the partial-adjustment model

$$m_t - m_{t-1} = \delta\, (m^*_t - m_{t-1}),\ 0 < \delta \leq 1 \quad (b)$$

The parameter δ measures the speed of adjustment. m^* is the long-run desired level of real money balances, and other variables are as defined before. Assume now that the random disturbance term U_t in (a) is stationary and follows a simple AR(1) process of the form

$$U_t = \rho\, U_{t-1} + \epsilon_t,\ 0 \leq \rho < 1 \quad (c)$$

The parameter ρ is determined by the nature of shocks to money demand.

Note that the empirical work in the text relies on a long-run demand specification like (a), but allows for more general dynamics than embedded in (b). Equations (a), (b) and (c) imply the following reduced form equation for changes in money balances

$$m_t - m_{t-1} = \delta a_1\, \Delta y_t - \delta(1-\rho)\, U_{t-1} + \delta\, \epsilon_t. \quad (d)$$

Equation (d) resembles the error-correction model of the form (2) given in the text. As can be seen, the size of the coefficient on the lagged level of U_t depends upon two parameters δ and ρ. If ρ is close to unity, then the error-correction parameter will be small even if δ is large.

An important assumption implicit in the above discussion is that the random disturbance term U_t is stationary. Intuitively, this assumption means that actual M2 balances do not permanently drift away from what is determined by long-run factors specified in (1). If this assumption is incorrect so that U_t is in fact nonstationary, then the regression equation (1) if estimated is subject to the spurious regression phenomenon. Furthermore, the coefficient λ in (2) is likely to be zero. To see this, first-difference the equation (1) as in (3)

$$\Delta rM2_t = a_1 \, \Delta rY_t - a_2 \, \Delta(R-RM2)_t + U_t - U_{t-1} \qquad (3)$$

Assume now that U_t follows a first-order autoregressive process of the form

$$U_t = \rho \, U_{t-1} + \epsilon_t$$

where ϵ_t is a pure white noise process. Then we can rewrite (3) as in (4)

$$\Delta rM2_t = a_1 \, \Delta rY_t - a_2 \, \Delta(R-RM2)_t + (\rho - 1) \, U_{t-1} + \epsilon_t \qquad (4)$$

Equation (4) is similar in spirit to equation (2). If ρ is less than unity so that U_t is stationary, then $\rho - 1$ [which equals λ in (2)] is different from zero. If $\rho = 1$ so that U_t is nonstationary, then $\rho - 1$ [and λ in (2)] is zero. Hence, the dynamic error-correction specification (2) exists if U_t is a stationary variable.

It can now be easily seen that if U_t is nonstationary, then the money demand regression estimated in first-difference form is appropriate [as λ in (2) is in fact zero]. On the other hand, if U_t is stationary, then the first-difference regression is misspecified because it omits the relevant variable U_{t-1} [as λ in (2) is in fact nonzero].

Estimation of the Error-Correction Model

If the random disturbance term Ut is stationary, then the money demand regression (2) can be estimated in two alternative ways. The first is a two-step procedure. In the first step, the long-run equilibrium M2 demand model (1) is estimated using a consistent estimation procedure, and the residuals are calculated. In the second step, the short-run money demand regression (2) is estimated with U_{t-1} replaced by residuals estimated in step one (see, for example, Hendry and Ericsson, 1991, and Baum and Furno, 1990). The money demand regression estimated in the first step of this procedure generates estimates[6] of the long-term income and opportunity cost elasticities (a_1 and a_2). The short-run money demand parameter estimates are generated in the second step.

The alternative procedure is to replace U_{t-1} in (2) by the lagged levels of the variables and estimate the short-run and long-run parameters jointly. To explain it further, substitute (1) into (2) to obtain a combined equation

$$\begin{aligned}\Delta rM2_t = {} & d_0 + \sum_{s=1}^{n1} b_{1s} \, \Delta rM2_{t-s} \\ & + \sum_{s=0}^{n2} b_{2s} \, \Delta rY_{t-s} \\ & - \sum_{s=0}^{n3} b_{3s} \, \Delta(R-RM2)_{t-s} \\ & + d_1 \, rM2_{t-1} + d_2 \, rY_{t-1} \\ & + d_3 \, (R-RM2)_{t-1} + \epsilon_t \qquad (5)\end{aligned}$$

where $d_0 = (b_0 - a_0 \lambda)$

$d_1 = \lambda$

$d_2 = -\lambda \, a_1$

$d_3 = \lambda \, a_2$

Equation 5 can be estimated using a consistent estimation procedure and all parameters of (1) and (2) can be recovered from those of (5). For example, the error-correction coefficient λ is d_1; the long-term income elasticity (a_1) is d_2 divided by d_1; and the long-term opportunity cost elasticity (a_2) is d_3 divided by d_1 (see, for example, Small and Porter, 1989).

If one wants to test hypotheses about the long-run parameters of the money demand function (1), it is easier to do so under the second framework than

---

[6] It should be pointed out that if all of the variables included in (1) are nonstationary, then ordinary least squares estimates of (1) are consistent. However, the usual t- and F-ratio statistics have nonstandard limiting distributions because U_t in (1) is generally serially correlated and/or heteroscedastic. This means one can not carry out tests of hypotheses about the long-run parameters in the standard fashion. Furthermore, if even a single variable in (1) is stationary, then ordinary least squares estimates are inconsistent. West (1988) in that case suggests using an instrumental variables procedure.

under the two-step procedure.[7] The reason is that the residuals in the equilibrium model estimated in step one of the first procedure are likely to be serially correlated and possibly heteroscedastic. Hence, the usual t- and F-ratio test statistics are invalid unless further adjustments are made. In contrast, the residuals in the money demand regression (5) are likely to be well behaved, validating the use of the standard test statistics in conducting inference. In view of these considerations, the error-correction money demand model is estimated using the second procedure, i.e., the money demand function (5).

As noted above, the long-term income elasticity can be recovered from the long-run part of the model (5), i.e., a_1 is d_2 divided by d_1. It may however be noted that the short-run part of the model (5) may yield another estimate of the long-term scale elasticity, i.e., $\sum_{s=0}^{n2} b_{2s}/(1 - \sum_{s=1}^{n1} b_{1s})$. If the same scale variable appears in the long- and short-run parts of the model, then a "convergence condition" might be imposed to ensure that one gets the same point-estimate of the long-term scale elasticity. To explain further, assume that real income appears in the long- and short-run parts of the model and that the long-term income elasticity is unity, i.e., $a_1 = 1$ in (1). This restriction implies that coefficients that appear on rY_{t-1} and $rM2_{t-1}$ in (5) sum to zero. This restriction pertains to the long-run part of the model and is expressed as in (6.1)

$$d_1 + d_2 = 0 \tag{6.1}$$

Furthermore, if the long-term income elasticity computed from the short-run part of the model is unity, then it also implies the following

$$\sum_{s=0}^{n2} b_{2s}/(1 - \sum_{s=1}^{n1} b_{1s}) = 1. \tag{6.2}$$

Equivalently, (6.2) can be expressed as

$$\sum_{s=0}^{n2} b_{2s} + \sum_{s=1}^{n1} b_{1s} = 1.$$

[7] It should be pointed out that these remarks apply to the case in which the equilibrium model (1) is estimated by ordinary least squares, as suggested by Engle and Granger (1987). However, if the equilibrium money demand model is estimated using the procedure given in Johansen and Juselius (1989), then one can conduct various tests of hypotheses of the long-run parameters. The approach advanced in Johansen and Juselius is, however, quite complicated.

In general, if different scale variables appear in the short- and long-run parts of the model, then these restrictions may or may not be imposed on the model.

Tests for Cointegration

An assumption that is necessary to yield reliable estimates of the money demand parameters is that U_t in (1) should be stationary. Since the levels of the variables included (1) are generally nonstationary, the stationarity of U_t requires that these nonstationary variables be cointegrated as discussed in Engle and Granger (1987). Hence, one must first test for the existence of a long-run equilibrium relationship among the levels of the nonstationary variables in (1).

Several tests for cointegration have been proposed in the literature (see, for example, Engle and Granger, 1987, and Johansen and Juselius, 1989). The test for cointegration used here is the one proposed in Engle and Granger (1987) and consists of two steps. The first tests whether each variable in (1) is nonstationary. One does this by performing a unit root test on the variables. The second step tests for the presence of a unit root in the residuals of the levels regressions estimated using the nonstationary variables. If the residuals do not have a unit root, then the nonstationary variables are cointegrated. For the case in hand, if U_t in (1) does not have a unit root, then the nonstationary variables in (1) are said to be cointegrated.

Data and the Definition of Scale Variables

The money demand regression (5) is estimated using the quarterly data that spans the period 1953Q1 to 1990Q4. rM2 is measured as nominal M2 deflated by the implicit GNP price deflator; rY by real GNP; R by the four- to six-month commercial paper rate and; RM2 by the weighted average of the explicit rates paid on the components of M2.

The theoretical analysis presented in McCallum and Goodfriend (1987) implies that the scale variable that appears in a typical household's money demand relationship is real consumption expenditure. Mankiw and Summers (1986) have presented empirical evidence that in aggregate money demand regressions consumer expenditure is a better scale variable than GNP. Their reasoning is based on the observation that some components of GNP, such as business fixed investment and changes in inventories, do not generate as much increase in money balances as does consumer expenditure. The money demand regressions estimated by Mankiw and Summers are in level form and use distributed lags on the scale and interest rate variables. Their empirical work implies that consumer expenditure is a better scale variable than GNP in the short run as well as in the long run. In contrast, Small and Porter (1989) used consumer spending as the short-run scale variable, and GNP as the long-run scale variable. Here I formally test which scale variable is appropriate in the short and long run.[8]

II. EMPIRICAL RESULTS

Unit Root Test Results

The money demand regression (5) includes the levels and first-differences of money, income and opportunity cost variables $rM2_t$, $\Delta rM2_t$, rY_t, ΔrY_t, $(R-RM2)_t$, and $\Delta(R-RM2)_t$. The alternative scale variable considered is real consumer expenditure: rC_t and ΔrC_t. The Augmented Dickey Fuller test[9] is used to test the presence of unit roots in these variables. The test results are reported in Table 1.

[8] All the data (with the exception of RM2 and M2) is taken from the Citibank database. M2 for the pre-1959 period and RM2 are constructed as described in Hetzel (1989).

[9] The unit root test procedure used here is described in Mehra (1990).

These results suggest the presence of a single unit root in $rM2_t$, rY_t and rC_t, implying that the levels of these variables are nonstationary but the first-differences are not. The financial market opportunity cost variable $(R-RM2)_t$ does not have a unit root and is thus stationary.[10]

Cointegration Test Results

The unit root test results presented above imply that except for $rM2_t$ and rY_t all other variables included in the money demand regression (5) are stationary. If $rM2_t$ and rY_t are cointegrated, then (5) can be estimated by ordinary least squares and the resulting parameter estimates are not subject to the spurious regression phenomenon.

The results of testing for cointegration[11] between $rM2_t$ and rY_t are presented in Table 2. As can be seen, the residuals from a regression of $rM2_t$ on rY_t

[10] Schwert (1987) has shown that usual unit root tests may be invalid if time series are generated by moving as well as autoregressive components. In order to check for this potential bias, unit root tests were repeated using longer lags on first-differences of time series. In particular, the parameter n in Table 1 was set at 8 and 12. Those unit root test results (not reported) yielded similar inferences.

[11] For a simple description of this cointegration test see Mehra (1989).

Table 1

Unit Root Test Results, 1953Q1–1990Q4

Z_t	$\rho(t:\rho-1=0)$	$\beta(t:\beta=0)$	Φ_3 $(\rho=1, \beta=0)$	n	$\chi^2(1)$	$\chi^2(2)$
	First Unit Root					
$rM2_t$	.97 (−2.2)	.20 (2.0)	2.67	1	.76	1.59
rY_t	.95 (−2.5)	.39 (2.5)	2.50	2	1.50	1.72
rC_t	.94 (−2.5)	.46 (2.5)	3.13	2	.96	1.03
$(R-RM2)_t$	.80 (−4.2)*	.57 (1.2)	9.07*	4	.37	.42
	Second Unit Root					
$\Delta rM2_t$	.59 (−5.3)*			1	.28	.39
ΔrY_t	.31 (−6.5)*			2	.62	1.28
ΔrC_t	.29 (−5.3)*			4	.45	.55
$\Delta(R-RM2)_t$	.09 (−7.0)*			2	.10	.68

Notes: Regressions are of the form $Z_t = \alpha + \sum_{s=1}^{n} d_s \Delta Z_{t-s} + \rho Z_{t-1} + \beta T + \epsilon_t$. All variables are in their natural logs; rM2, real balances; rY, real GNP; rC, real consumer spending; R-RM2, the differential between the four- to six-month commercial paper rate (R) and the own rate on M2 (RM2); T, a time trend; and Δ, the first-difference operator. The coefficient reported on trend is to be multiplied by 1000. The parameter n was chosen by the "final prediction error criterion" due to Akaike (1969). The coefficients ρ and β (t statistics in parentheses) are reported. Φ_3 tests the hypothesis $(\rho, \beta) = (1,0)$. $\chi^2(1)$ and $\chi^2(2)$ are Chi square statistics (Godfrey, 1978) that test for the presence of first- and second-order serial correlation in the residuals of the regression.

An "*" indicates significance at the 5 percent level. The 5 percent critical value for $t: \rho-1=0$ is 3.45 (Fuller, 1976, Table 8.5.2) and that for $\Phi_3:(\rho=1, \beta=0)$ is 6.49 (Dickey and Fuller, 1981, Table VI).

Table 2

Cointegrating Regressions, 1953Q1–1990Q4

$X1_t$	$X2_t$	b	d (t:d=0)	n	$\chi^2(1)$	$\chi^2(2)$
rM2	rY	1.01	−.10 (−3.5)*	1	1.1	1.1
rY	rM2	.98	−.10 (−3.5)*	1	1.1	1.1
rM2	rC	.91	−.05 (−2.3)	1	1.6	2.1
rC	rM2	1.08	−.05 (−2.3)	1	1.6	2.1

Notes: Each row reports coefficients from two regressions. The first regression is the cointegrating regression of the form $X1_t = a + b\,X2_t + U_t$, where U_t is the residual. The second regression tests for a unit root in the residual of the relevant cointegrating regression and is of the form

$$\Delta U_t = d\,U_{t-1} + \sum_{s=1}^{n} f_s\,\Delta U_{t-s}.$$

The coefficient reported from the first regression is b and the coefficient d is from the other regression. n is the number of lags chosen by Akaike's final prediction error criterion. $\chi^2(1)$ and $\chi^2(2)$ are Godfrey (1978) statistics that test for the presence of first- and second-order serial correlation in the residuals of the second regression.

An "*" indicates significance at the 5 percent level. The 5 percent critical value is 3.21 (Engle and Yoo, 1987, Table 3).

(or of rY_t on $rM2_t$) do not possess a unit root, implying that these two variables are cointegrated. Table 2 also presents test results for cointegration between $rM2_t$ and rC_t; those results suggest that $rM2_t$ and rC_t are not cointegrated.

Estimated M2 Demand Regressions

The cointegration test results above imply that the appropriate scale variable that enters the long-run part of the money demand model (5) is real GNP, not real consumer spending.[12] It is, however, still plausible that real consumer spending is a better short-run scale variable than real GNP. In order to examine this issue, (5) is also estimated using ΔrC_t in the short-run part of the model.

The results of estimating (5) are reported in Table 3. The regressions A and B in Table 3 use real GNP and real consumer spending respectively as the short-run scale variable. The long-run part of the model still uses real GNP as the scale variable. The regressions are estimated without imposing the restrictions (6.1) and (6.2). The regressions also included zero-one dummies to control for the transitory effects of credit controls and the introduction of MMDAs and Super-NOWs. As can be seen, both regressions appear to provide reasonable point-estimates of the long-run and short-run parameters. The long-run real GNP elasticity computed from the long-run parts of the models is unity, and the point-estimate of the long-run financial market opportunity cost elasticity ranges between −.10 to −.12. The short-run coefficients that appear on the scale and opportunity cost variables are generally of the correct signs and are statistically significant. The residuals from these regressions do not indicate the presence of any serial correlation (see Chi square and Q statistics reported in Table 3).

The cointegrating regressions between $rM2_t$ and rY_t reported in Table 2 suggest that the estimated long-term real GNP elasticity is not economically different from unity ($\hat{a}_1$ = 1.0 or .98; Table 2). If this hypothesis is true, then it implies that the restriction (6.1) is also true. F1 in Table 3 is the F statistic that tests whether (6.1) is true. F1 is .026 for regression A and .44 for regression B. Both values are small and thus imply that the long-run real GNP elasticity is not different from unity.

Evaluating the Demand Regressions

The money demand regressions reported in Table 3 are further evaluated by examining their

[12] The long-run money demand functions are assumed to be of the form

$$rM2_t = a_0 + a_1\,rY_t - a_2\,(R - RM2)_t. \quad (1)$$

A key feature of this specification is that the opportunity cost of holding M2 depends upon the differential between a market rate of interest (R_t) and the own rate of return on M2 ($RM2_t$). This specification thus implies that coefficients that appear on R_t and $RM2_t$ in (1) are of opposite signs but equal absolute sizes. The unit root test results presented in the text implies that $(R-RM2)_t$ is stationary, whereas $rM2_t$ and rY_t are not. The cointegration test results presented in the text implies that $rM2_t$ and rY_t are cointegrated. These results together then imply the presence of a single cointegrating vector among the variables postulated in (1). See Goodfriend (1990).

Table 3

The Error-Correction M2 Demand Regressions, 1953Q1–1990Q4

A. *Real GNP in the Short- and Long-Run Parts of the Model*

$$\underset{}{\Delta rM2_t} = \underset{(1.5)}{-.19} + \underset{(4.3)}{.33}\, \Delta rM2_{t-1} + \underset{(1.5)}{.11}\, \Delta rM2_{t-2} + \underset{(1.7)}{.09}\, \Delta rY_t + \underset{(2.0)}{.12}\, \Delta rY_{t-1} - \underset{(5.8)}{.01}\, \Delta(R-RM2)_t - \underset{(4.5)}{.01}\, \Delta(R-RM2)_{t-1}$$

$$\underset{(1.7)}{-.04}\, rM2_{t-1} + \underset{(1.6)}{.04}\, rY_{t-1} \underset{(2.8)}{-.005}\, (R-RM2)_{t-1} \underset{(2.3)}{-.014}\, CC1 + \underset{(1.7)}{.010}\, CC2 + \underset{(4.7)}{.026}\, D83Q1$$

SER = .0055 $\quad \chi^2(1) = .24 \quad \chi^2(2) = 2.3 \quad$ Q(36) = 23.3

$N_{rY} = 1.0 \quad N_{R-RM2} = -.12 \quad$ F1(1,139) = .026

B. *Real Consumer Spending in the Short-Run Part and Real GNP in the Long-Run Part*

$$\Delta rM2_t = \underset{(1.9)}{-.24} + \underset{(4.4)}{.31}\, \Delta rM2_{t-1} + \underset{(1.6)}{.12}\, \Delta rM2_{t-2} + \underset{(2.3)}{.17}\, \Delta rC_t + \underset{(2.0)}{.15}\, \Delta rC_{t-1} - \underset{(5.5)}{.01}\, \Delta(R-RM2)_t - \underset{(4.3)}{.001}\, \Delta(R-RM2)_{t-1}$$

$$\underset{(2.0)}{-.05}\, rM2_{t-1} + \underset{(2.0)}{.05}\, rY_{t-1} \underset{(3.1)}{-.005}\, (R-RM2)_{t-1} \underset{(1.7)}{-.009}\, CC1 + \underset{(1.5)}{.009}\, CC2 + \underset{(4.5)}{.025}\, D83Q1$$

SER = .0055 $\quad \chi^2(1) = .001 \quad \chi^2(2) = 1.4 \quad$ Q(36) = 20.7

$N_{ry} = 1.0 \quad N_{R-RM2} = -.10 \quad$ F1(1,139) = .44

Notes: The regressions are estimated by ordinary least squares. All variables are defined as in Table 1. CC1, CC2, and D83Q1 are, respectively, 1 in 1980Q2, 1980Q3 and 1983Q1 and zero otherwise. SER is the standard error of the regression; $\chi^2(1)$ and $\chi^2(2)$ are Godfrey statistics for the presence of first- and second-order correlation in the residuals, respectively; Q the Ljung-Box Q statistic; N_{ry} the long-term income elasticity; and N_{R-RM2} the long-term financial market opportunity cost elasticity. The long-term income elasticity is given by the estimated coefficient on rY_{t-1} divided by the estimated coefficient on $rM2_{t-1}$ and the long-term opportunity cost elasticity is given by the estimated coefficient on $(R\text{-}RM2)_{t-1}$ divided by the estimated coefficient on $rM2_{t-1}$. F1 is the F statistic that tests whether coefficients on $rM2_{t-1}$ and rY_{t-1} sum to zero. F1 is distributed with F(1,139) degrees of freedom.

structural stability and out-of-sample forecast performance.

Table 4 presents results of the Chow test of structural stability over the period 1953Q1 to 1990Q4. The Chow test is implemented using the dummy variable approach and potential breakpoints covering the period 1970Q4 to 1980Q4 are considered. (The start date is near the midpoint of the whole sample period and the end date near the introduction of NOWs in 1981.) The slope dummies are considered for the long-run as well as for the short-run coefficients.[13] F-S in Table 4 is the F statistic that tests whether slope dummies for the short-run coefficients are zero. F-L tests such slope dummies for the long-run coefficients. F-SL tests all of the slope dummies including the one on the constant term. As can be seen in Table 4, these F statistics generally are not statistically significant and thus imply that the regressions reported in Table 3 do not depict the parameter instability.

[13] The results reported in Table 3 suggest that the restriction $a_1 = 1$ is not inconsistent with the data. This constraint was imposed on the long-run part of the model while implementing the test of stability.

The out-of-sample forecast performance is evaluated by generating the rolling-horizon forecasts of the rate of growth of M2 as in Hallman, Porter and Small (1989).[14] The relative forecast performance of the two competing money demand models is compared over the period 1971 to 1990.[15]

Table 5 reports summary statistics for the errors that occur in predicting M2 growth over one-year-,

[14] The forecasts and errors were generated as follows. Each money demand model was first estimated over an initial estimation period 1953Q1 to 1970Q4 and then simulated out-of-sample over one to three years in the future. For each of the competing models and each of the forecast horizons, the difference between the actual and predicted growth was computed, thus generating one observation on the forecast error. The end of the initial estimation period was then advanced four quarters and the money demand equations were reestimated, forecasts generated, and errors calculated as above. This procedure was repeated until it used the available data through the end of 1990.

[15] The money demand models that underlie this simulation exercise are from Table 3. The predicted values are, however, generated under the constraint that the long-term scale variable elasticity is unity whether computed from the long-run part or from the short-run part of the model. The out-of-sample prediction errors from the error-correction money demand models estimated with this constraint are generally smaller than those from models estimated without the constraint.

Table 4

Stability Tests, 1953Q1–1990Q4

Break Point	Equation A			Equation B		
	F-S	F-L	F-SL	F-S	F-L	F-SL
1970Q4	.6	1.1	1.0	.3	.8	.5
1971Q4	.6	.5	.7	.8	.6	.7
1972Q4	1.9	.1	1.6	1.7	.3	1.2
1973Q4	1.4	.7	1.2	1.4	2.7	1.4
1974Q4	1.9	1.1	1.6	1.6	1.9	1.4
1975Q4	.9	.4	.7	.8	.0	.7
1976Q4	1.2	.4	1.0	1.0	1.2	.9
1977Q4	2.1*	.6	1.6	1.7	.2	1.1
1978Q4	1.8	1.2	1.3	1.6	.9	1.2
1979Q4	1.7	.6	1.2	1.4	.3	1.0
1980Q4	1.3	1.6	1.2	1.3	1.1	1.1

Notes: The reported values are the F statistics that test whether slope dummies when added to equations A and B (reported in Table 3) are jointly significant. The breakpoint refers to the point at which the sample is split in order to define the dummies. The dummies take values 1 for observations greater than the breakpoint and zero otherwise. F-S tests whether slope dummies for the short-run coefficients are zero and are distributed F(6,131) degrees of freedom. F-L tests whether slope dummies for the long-run coefficients are zero and are distributed F(2,131) degrees of freedom. F-SL tests whether all of slope dummies including the one on the constant term are zero and are distributed F(9,131) degrees of freedom.

An "*" indicates significance at the 5 percent level.

two-year-, and three-year-ahead periods. Statistics for regression A are shown within brackets. The period-by-period errors are reported only for the M2 demand regression with real consumer spending as the short-run scale variable. These results suggest two observations. The first is that the regression with real consumer spending provides more accurate forecasts of M2 than does the regression with real GNP. For all forecast horizons the root mean squared errors from regression B are smaller than those from regression A (see Table 5). The second is that the error-correction model with real consumer spending as a short-run scale variable does reasonably well in predicting the rate of growth of M2. The bias is small and the root mean squared error (RMSE) is 1.0 percentage points for the one-year horizon. Moreover, the prediction error declines as the forecast horizon lengthens.

The out-of-sample M2 forecasts are further evaluated in Table 6, which presents regressions of the form

$$A_{t+s} = a + b\ P_{t+s},\ s = 1,\ 2,\ 3, \qquad (7)$$

where A and P are the actual and predicted values of M2 growth. If these forecasts are unbiased, then a = 0 and b = 1. F statistics reported in Table 6 test the hypothesis (a,b) = (0,1). As can be seen, these F values are consistent with the hypothesis that the forecasts of M2 growth are unbiased.

III.
Summary Remarks

The money demand equations have typically been estimated either in log-level form or in log-difference form. The recent advances in time series analysis have highlighted potential problems with each of these specifications. As a result, several analysts have begun to integrate these two specifications using the theories of error-correction and cointegration. In this approach, a long-run money demand model is first fit to the levels of the variables, and the calculated residuals from that model are used in an error-correction model which specifies the system's short-run dynamics. Such an approach thus allows both the levels and first-differences of the relevant variables to enter the money demand regression.

Using the above approach, this paper presents an error-correction model of M2 demand in the postwar period. It is shown here that real GNP, not real consumer spending, should enter the long-run part of the model. The point-estimate of the long-run real GNP elasticity is not different from unity. Real consumer spending however appears more appropriate in the short-run part of the model. The error-correction model with real consumer spending as a short-run scale variable provides more accurate out-of-sample forecasts of M2 growth than does the model with real GNP. However, both of these models are stable by the conventional Chow test over the sample period 1953Q1 to 1990Q4.

The out-of-sample forecasts presented here suggest that M2 growth in the 1980s is well predicted by the error-correction model that uses real consumer spending as a short-run scale variable. The rate of growth in real consumer spending, which averaged 3.97 percent in the 1983 to 1988 period, decelerated to 1.2 percent in 1989 and .2 percent in 1990. The rate of growth in M2 has also decelerated over the past two years. The money demand model presented here implies that part of the recently observed deceleration in M2 growth reflects deceleration in real consumer spending and is not necessarily indicative of any instability in M2 demand behavior.

Table 5

Rolling-Horizon M2 Growth Forecasts, 1971–1990

Year	1 Year Ahead Actual	1 Year Ahead Predicted	1 Year Ahead Error	2 Years Ahead Actual	2 Years Ahead Predicted	2 Years Ahead Error	3 Years Ahead Actual	3 Years Ahead Predicted	3 Years Ahead Error
1971	12.6	12.4	.2	–	–	–	–	–	–
1972	12.0	10.9	1.1	12.3	11.4	.9	–	–	–
1973	6.9	8.9	–1.9	9.5	9.3	.2	10.5	10.1	.5
1974	5.7	5.9	–.2	6.3	7.7	–1.3	8.2	8.4	–.2
1975	11.4	10.1	1.3	8.6	8.4	.2	8.0	8.5	–.5
1976	12.5	12.4	.1	11.9	11.4	.5	9.9	9.7	.2
1977	10.6	11.1	–.5	11.6	11.8	–.2	11.5	11.5	–.0
1978	7.7	8.6	–.9	9.1	9.6	–.4	10.3	10.5	–.3
1979	7.8	8.7	–.9	7.7	8.5	–.8	8.7	9.1	–.4
1980	8.6	8.7	–.1	8.2	8.9	–.7	8.0	8.6	–.6
1981	8.9	8.4	.5	8.7	8.7	.0	8.4	8.8	–.4
1982	8.7	7.9	.8	8.8	7.9	.8	8.7	8.4	.4
1983	11.5	9.6	1.9	10.1	8.7	1.4	9.7	8.5	1.2
1984	7.7	6.4	1.3	9.5	8.6	.9	9.3	8.4	.9
1985	8.3	8.5	–.2	8.0	7.5	.6	9.2	8.7	.5
1986	8.8	7.4	1.4	8.7	8.1	.6	8.3	7.6	.7
1987	4.2	3.1	1.1	6.5	5.1	1.4	7.2	6.3	.9
1988	5.0	5.8	–.8	4.6	4.1	.5	6.0	5.3	.7
1989	4.5	4.4	.1	4.8	5.2	–.4	4.6	4.4	.2
1990	3.8	5.1	–1.3	4.2	4.7	–.5	4.5	5.1	–.6
Mean Error			.16[–.16][a]			.19[.0][a]			.17[.0][a]
Mean Absolute Error			.85[1.19][a]			.66[.94][a]			.49[.64][a]
Root Mean Squared Error			1.02[1.43][a]			.77[1.14][a]			.57[.77][a]

Notes: Actual and predicted values are annualized rates of growth of M2 over 4Q-to-4Q periods ending in the years shown. The predicted values are generated using the money demand equation B of Table 3. (See footnote 14 in the text for a description of the forecast procedure used.) The predicted values are generated under the constraint that the long-term scale variable elasticity is unity whether computed from the long-run part or from the short-run part of the model.

[a] The values in brackets are the summary error statistics generated using the money demand regression A of Table 3.

Table 6

Error-Correction M2 Demand Models: Out-of-Sample Forecast Performance, 1971–1990

Short-Run Scale Variable	1 Year Ahead a	1 Year Ahead b	1 Year Ahead F3(2,18)	2 Years Ahead a	2 Years Ahead b	2 Years Ahead F3(2,17)	3 Years Ahead a	3 Years Ahead b	3 Years Ahead F3(2,16)
Real Consumer Spending	.0	1.01	.2	.1	1.0	.6	.5	.96	.9
	(.8)	(.09)		(.6)	(.08)		(.6)	(.08)	
Real GNP	.4	.93	.2	.6	.91	.3	.8	.89	.7
	(1.2)	(.13)		(1.1)	(.12)		(.8)	(.09)	

Notes: The table reports coefficients (standard errors in parentheses) from regressions of the form $A_{t+s} = a + b\,P_{t+s}$, where A is actual M2 growth; P predicted M2 growth; and s (= 1,2,3) number of years in the forecast horizon. The values used for A and P are from Table 5. F3 is the F statistic that tests the null hypothesis (a,b) = (0,1), and are distributed F with degrees of freedom given in parentheses following F3.

REFERENCES

Akaike, H. "Fitting Autoregressive Models for Prediction." *Annals of International Statistics and Mathematics* 21 (1969): 243-47.

Baum, Christopher F. and Marilena Furno. "Analyzing the Stability of Demand for Money Equations via Bounded-Influence Estimation Techniques." *Journal of Money, Credit and Banking* (November 1990): 465-71.

Dickey, David A. and Fuller, Wayne A. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49 (July 1981): 1057-72.

Engle, Robert F. and Byund Sam Yoo. "Forecasting and Testing in a Cointegrated System." *Journal of Econometrics* 35 (1987): 143-59.

Engle, Robert F. and C. W. J. Granger. "Cointegration and Error-Correction Representation, Estimation, and Testing." *Econometrica* (March 1987): 251-76.

Fuller, W. A. Introduction to Statistical Time Series, 1976, Wiley, New York.

Godfrey, L. G. "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables." *Econometrica* 46 (November 1978): 1303-10.

Goodfriend, Marvin. "Comments on Money Demand, Expectations and the Forward-Looking Model." *Journal of Policy Modeling* 12(2), 1990.

Granger, C. W. J., and P. Newbold. "Spurious Regressions in Econometrics." *Journal of Econometrics* (July 1974): 111-20.

Hallman, Jeffrey, J., Richard D. Porter, and David H. Small. "M2 Per Unit of Potential GNP as an Anchor for the Price Level." Staff Study #157, Board of Governors of the Federal Reserve System (April 1989).

Hendry, D. F. and N. R. Ericsson. "An Econometric Analysis of UK Money Demand in Monetary Trends in the United States and the United Kingdom," by Milton Friedman and Anna J. Schwartz. *American Economic Review* (March 1991): 8-38.

Hendry, David F., Adrian R. Pagan, and J. Denis Sargan. "Dynamic Specification." In *Handbook of Econometrics*, edited by Z. Griliches and M. Intriligator. Amsterdam: North Holland, 1983.

Hendry, David F., and Jean-Francois Richard. "On the Formulation of Empirical Models in Dynamic Econometrics." *Journal of Econometrics* 20 (1982): 3-33.

Hetzel, Robert L. "M2 and Monetary Policy." Federal Reserve Bank of Richmond, *Economic Review* 75 (September/October 1989): 14-29.

Hetzel, Robert L., and Yash P. Mehra. "The Behavior of Money Demand in the 1980s." *Journal of Money, Credit and Banking*, (November 1989): 455-63.

Johansen, Soren and Katrina Juselius. "The Full Information Maximum Likelihood Procedures for Inference on Cointegrating-with Application." Institute of Mathematical Statistics, University of Copenhagen, Reprint No. 4, January 1989.

Mankiw, N. Gregory and Lawrence H. Summers. "Money Demand and the Effects of Fiscal Policies." *Journal of Money, Credit and Banking* (November 1986): 415-29.

McCallum, Bennett T. and Marvin S. Goodfriend. "Demand for Money: Theoretical Studies." In *The New Palgrave*. A Dictionary of Economics edited by John Eatwell, Munay Milgate, and Peter Newman. Vol. 1, 1987, 775-81.

Mehra, Yash P. "Real Output and Unit Labor Costs as Predictors of Inflation." Federal Reserve Bank of Richmond, *Economic Review* 76 (July/August 1990): 31-39.

__________. "Some Further Results on the Source of Shift in M1 Demand in the 1980s." Federal Reserve Bank of Richmond, *Economic Review* 75 (September/October 1989): 3-13.

Phillips, P. C. B. "Understanding Spurious Regressions in Econometrics." *Journal of Econometrics* 33 (1986): 311-40.

Schwert, G. W. "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data." *Journal of Monetary Economics* 20 (1987): 73-103.

Small, David H. and R. D. Porter. "Understanding the Behavior of M2 and V2." *Federal Reserve Bulletin* (April 1989): 244-54.

A Total Production Index for Washington, D.C.

Dan M. Bechter, Zoltan Kenessey, Fred Siegmund, and Ray D. Whitman *

Introduction

This article describes the methods and procedures used in computing a new total production index for the District of Columbia.[1] The new index accounts for changes in services production as well as goods production.[2] The index made its public debut in a release issued March 15, 1991, by the Center of Economic and Business Statistics of the University of the District of Columbia. That same day, *The Washington Post* featured the new index on the first page of its business section. In subsequent months, the Center has issued updates of the index under the release's name, *D.C. Economy*.

At the national level, a monthly production index provides a timely measure of cyclical changes in economic output between calendar quarters. Quarterly figures for gross national product provide the most comprehensive measures of production; between quarters, the monthly index of industrial production compiled by the Federal Reserve Board has proved to be an important and carefully watched economic indicator.

At the regional level, timely measures of output are valuable to business and government officials because economic activity in any region can differ significantly from the national average,and because gross state product figures are available only after a long delay. Output indexes compiled by the Federal Reserve Banks have helped meet the demand for regional economic information used in analyzing economic growth and business cycles, and in economic policy formulation. The attention given to the Federal Reserve's *Beige Book* is one example of the interest of policymakers and the public in reports on economic activity around the country.

The District of Columbia economy is different in composition from the economies of surrounding states and the nation as a whole. For example, the government-based D.C. economy has a relatively small manufacturing sector. Manufacturing indexes for Maryland and Virginia therefore provide little guidance about the current state of economic activity in the nation's capital. The D.C. economy also behaves differently, although it is not always as insulated from the national business cycle as is commonly believed. For example, although employment remained relatively flat in the District of Columbia during the U.S. recession of 1974-75, it declined by a larger percentage than U.S. employment over the two recessions of the early 1980s. Also, the booms associated with the D.C. metropolitan area have been much less evident in the city itself. In the past 20 years, for example, employment in the District of Columbia has grown only 20 percent, in contrast to the 89 percent increase in the entire Washington metro area. It is clear, therefore, that although economic activity in the District of Columbia is usually less volatile than in the nation as a whole, it does change in intensity and, sometimes, direction.

Many individual economic indicators are used to help track the D.C. economy. *The Washington Post*, for example, regularly features charts and data for several different economic sectors. It is difficult to extract from them, however, a clear sense of the general condition and direction of the economy of the District of Columbia. That is, no single indicator fits the pieces of the Washington economy together in a coherent fashion. A timely monthly index of total production does that.

* Dan Bechter is a vice president at the Federal Reserve Bank of Richmond, Zoltan Kenessey is a senior economist at the Board of Governors of the Federal Reserve System, Fred Siegmund is a professor at the University of the District of Columbia, and Ray Whitman is a professor and the director of the Center of Economic and Business Statistics at the University of the District of Columbia. The views expressed are those of the authors and should not be attributed to the Federal Reserve or to the University of the District of Columbia.

[1] The "Total Production Index for the District of Columbia" is computed by the Center for Business and Economic Statistics of the University of the District of Columbia and was developed by the Center in cooperation with the Federal Reserve Bank of Richmond and from the experience and advice from staff at the Federal Reserve Board in Washington. The work has been facilitated by an advisory panel consisting of economists from each of these organizations, who met from time to time to review progress on the project. The authors thank Tapas Ghosh for valuable research assistance.

[2] The measurement of services production draws heavily on earlier work by Zoltan Kenessey. See, in particular, Kenessey (May 1988; November 1988).

Background on Production Indexes

The definitive history of production indexes has yet to be written. More than 60 years ago, however, Arthur Burns referred to the European production indexes by Neumann-Spallart of 1887 and Armand Julin of 1911, and to William Leonard's 1913 index dealing with extractive industries in America (Burns, 1930).

The Federal Reserve System has a long history of involvement in the measurement and analysis of monthly production developments. From its first issue in 1915, the *Federal Reserve Bulletin* contained business conditions data, including some on production. After January 1919, the *Bulletin* reported, in more extended form, monthly data on the "physical volume of trade" (including production). The Federal Reserve Board introduced indexes of production in the *Bulletin* in the spring of 1922, and in more refined form in the winter of the same year (Federal Reserve Board, March 1922 and December 1922).

Work on indexes of production was also underway outside of the Federal Reserve. Wesley C. Mitchell published an annual index number of production in 1919 (Mitchell, 1919). Mitchell and others at the National Bureau of Economic Research (NBER, incorporated in 1920) played a continuing role in U.S. macroeconomic measurement throughout the 1920s and 1930s and greatly influenced the development of production estimates in general. At Harvard University, Edmund Day produced the Harvard-Census index, also called the Day-Thomas index, by using quinquennial Census of Manufacturers data to adjust annual production indexes (Day, 1920). Walter Steward, who earlier worked at the War Production Board (led by Mitchell) and who became the director of research at the Federal Reserve Board in 1922, was among those who published articles about production index numbers in those days. During the 1920s, the U.S. Department of Commerce also issued various physical volume data and indexes of output, similar to those in the *Federal Reserve Bulletin*, which it published in the *Survey of Current Business*.

In 1927 the Federal Reserve Board introduced a new index of industrial production (back to 1919), which can be deemed the beginning of the more elaborate and advanced work on the subject in the United States.[3] The index was extensively revised in 1940, 1953, 1959, 1971, 1976, 1985, and most recently in 1989. Over the decades, the Federal Reserve established its preeminent role in monthly industrial output indexes (Federal Reserve Board, 1986). Meanwhile, important research efforts were made elsewhere, notably at the NBER. The work by Arthur Burns, Frederick Mills, Solomon Fabricant and others influenced not only the way industrial production was estimated, but also how all other components of the gross national product were measured.

The basic conceptual issues on production indexes developed by the Federal Reserve have been applied with some adaptations for use in specific regional economies. Regional production indexes compiled by the Federal Reserve Banks, principally for manufacturing, go back to the 1950s, with the earliest attempts undertaken at Atlanta, San Francisco, and Dallas. Today the Midwest Manufacturing Index of the Federal Reserve Bank of Chicago, the Mid-Atlantic Manufacturing Index of the Federal Reserve Bank of Philadelphia, the Fifth District Manufacturing Index of the Federal Reserve Bank of Richmond, and the Texas Industrial Production Index of the Federal Reserve Bank of Dallas command interest at the regional level among circles in business, government and academia (Kenessey, 1990).

Nationally, economic policymakers want information about developments in the various sectors and parts of the country. The uneven behavior of regional economies in recent economic expansions and contractions has heightened interest in this kind of information. State and local officials, many of whom are currently faced with budgetary shortfalls, clearly need better information about trends in their area economies. Consumers (and workers) also care a great deal about economic conditions affecting them; the popularity of state and metropolitan business magazines, business journals, and newspaper business sections attests to the public interest in local economic news. To help supplement the supply of state and regional economic information, the Federal Reserve Bank of Richmond calculates and publishes indexes of manufacturing output for each of the five states in the Fifth District (Bechter, et al., 1988).[4] Now, the total output index for the District of Columbia, reviewed here, is available.

Production indexes are coincident, not anticipatory, indicators of economic activity. Nationally, the

[3] *Industrial Production, With a Description of the Methodology*, Board of Governors of the Federal Reserve System, 1986, pp. 17-162.

[4] The Fifth Federal Reserve District comprises Maryland, North Carolina, South Carolina, Virginia, most of West Virginia, and the District of Columbia. The manufacturing index for Maryland incorporates the estimate for the District of Columbia.

index of industrial production is one of the four key coincident economic indicators used in identifying peaks and troughs of business cycles. Regionally, production indexes can be used similarly to provide important confirmatory evidence about the current status of output developments in particular economic areas. Regional production indexes are typically used for comparing the performance of a state or area economy with the national total and with other regions. Such analysis, whether it focuses on performance over time or across areas, usually highlights the movements observed for the most recent periods (months or quarters) in a region's economic activity. Importantly, improved regional measures of output may provide new leading indicators for swings in U.S. economic activity, as some regions may lead (and others lag) national business cycle developments.

The Concept of a Production Index

A production index is an index of the *quantity* of output, free of any influence of month-to-month changes in prices.[5] The focus on quantity precludes an index that compares current with past dollar values of production, as such an index would measure changes in prices and production together, not just changes in production. One alternative would be to measure production in constant dollars. Such an approach, however, would require a monthly set of price deflators for each product or product group. It seems useful, therefore, to adopt a methodology that relies mainly on physical measures of production such as tons of coal or taxi miles. Such physical measures of output do not require deflation to eliminate the effect of price changes. Along with the application of proper weights for aggregation, a production index covering several products can be estimated for each month in a timely fashion.

Most production indexes are of the Laspeyres (base-weighted) type. A Laspeyres quantity index can be expressed as:

$$I_t = \frac{\sum_{i=1}^{N} q_{it}p_{io}}{\sum_{i=1}^{N} q_{io}p_{io}} = \sum_{i=1}^{N} q_{it}\left(\frac{p_{io}}{\sum_{i=1}^{N} q_{io}p_{io}}\right)$$

$$= \sum_{i=1}^{N} \frac{q_{it}}{q_{io}}\left(\frac{q_{io}p_{io}}{\sum_{i=1}^{N} q_{io}p_{io}}\right) = \sum_{i=1}^{N} \frac{q_{it}}{q_{io}} w_{io}$$

where the summation is over the N individual goods and services included in the index, q denotes the quantities produced of these items, p denotes a term—usually price—used in weighting items in the index, t refers to the current period and o refers to the base period. The weight, w_{jo}, assigned to the j^{th} item and term q_{jt}/q_{jo} in the right side of the formula, is that item's share of the value of total output in the base period, or $q_{jo}p_{jo}/\Sigma q_{io}p_{io}$. The weights are held constant over a period of several years until changes in the relative importance of the various items of production have become so extensive that a revision of weights is warranted. Given its constant weights, the production index changes over time, as it should, only with changes in the output of goods and services.

As the right side of the formula shows, a quantity index covering several items can be expressed as a weighted average of the production indexes for individual items. The item weights, or product shares of the base period value of output, add up to 1. The individual and overall quantity indexes are usually expressed as percentages, with 100 the value for the base year.

Application to the District of Columbia

To formulate a production index for the District of Columbia, it was first necessary to decide how much productive activity to include. An index of manufacturing output alone was not likely to be very informative; in the District of Columbia, manufacturing consists largely of printing and publishing and is a small share of total employment, personal income, or production. In the District of Columbia, therefore, where the services-producing sectors dominate economic activity much more than in most of the rest of the country, it was appropriate to design an index of *total* production to include all significant segments of the economy: communications, construction, manufacturing, public utilities, public

[5] It is usually fairly easy to measure the change in output of a single homogenous agricultural or industrial commodity, such as bushels of #1 grade durum wheat, or tons of low sulphur bituminous coal. It is quite another matter, however, even when good data are available, to arrive at a "correct" measure of change in overall production when several commodities or grades of commodities are involved. The problem of adding apples and oranges is usually addressed in an economically appealing fashion by using constant prices along with the dollar values of output in some reference period. But because the reference period is usually fixed for a time, measures of change in overall production are plagued by index number problems—for example, the sensitivity of all index numbers to the choice of weights used in the weighted average. The problem in measuring production or prices intertemporally is further complicated by changes in the types and qualities of items in the "market basket" over time. These problems are addressed elsewhere in the literature on index numbers.

administration, services, trade, transportation, finance and real estate.[6]

Ideally, a total production index for the District of Columbia (referred to hereafter as the DC index) would draw on a broad range of physical output measures that fit neatly into the categories of the Standard Industrial Classification (SIC). In practice, ideal data series are not available. In the District of Columbia, several different agencies compile data for monthly use, and while many of these data do fit into the SIC categories, others do not. Moreover, as there are tens of thousands of different goods and services being produced, it was not practical to try to include all of them explicitly in the index. Instead, selected items of production were chosen to represent the monthly changes in output in various sectors. In selecting representative indicators, an effort was made to include one or more series for each major field of production.

Unfortunately, data on physical units of production were available for only one-sixth of total production, as measured by gross product in the base year. Fortunately, the theory of production suggests an alternative way to estimate physical output in the absence of these data. According to production theory, which has ample empirical support in the literature (e.g., the Cobb-Douglas production function), physical units of output can be expressed as a function of physical units of inputs. Moreover, over relatively short periods of time, a production function can be assumed stable, and the inputs of capital and land can be assumed fixed, with production varying with changes in labor input. Together with benchmark information on output provided by gross product data, therefore, labor input data provide a method to interpolate and extrapolate monthly estimates of production.[7]

Thus, to help construct the DC index where product series were deficient, employment data were used, alone as proxies for quantity series, or as supplements to incomplete quantity series. For example, to measure the production of construction in progress, construction-worker hours are used along with building permits to capture the ongoing nature of the work. Fortunately, labor data are available for all significant productive activities, so employment or production-worker hours by industry can be used as input proxies for production.

When the use of labor is applied as a proxy for production, some account must be made for changes in labor productivity over time. To adjust for the rise in productivity, past increases in average productivity are extrapolated from changes calculated between the most recent years reported by the Bureau of Economic Analysis in its gross product figures for the District of Columbia. For example, if between 1980 and 1986 the change in output of a certain good was 10 percent higher than the change in its labor input, then the average annual increase in labor productivity in the years since was about 1.6 percent, and the monthly increase was therefore assumed constant at about 0.13 percent.

In view of the federal government's very large share (36 percent in 1986) of productive activity in the economy of the District of Columbia, productivity movements of government workers are of particular interest for estimating output changes in the area. Fortunately, an extensive effort by the U.S. Bureau of Labor Statistics (BLS) within the framework of the Federal Productivity Measurement System (FPMS) produced quantitative results relevant to this topic (BLS, 1990). For fiscal year 1988, for example, FPMS covered 342 organizations within 61 federal agencies representing 2.1 million persons, 69 percent of the executive branch civilian work force. About 3,000 different products and services were measured in the system. The majority of the 28 major governmental functions, which compose total governmental activity reviewed, were services-producing areas. Yet, BLS was able to find representative product measures for these areas just as for goods-producing activities.

The BLS study found that output per employee increased at an average annual rate of 1.4 percent in the 1967-88 period and 0.7 percent between 1983 and 1988. This finding suggests that the usual assumption of unchanged productivity of federal employees in estimating government output is untenable. In the context of the DC index, the BLS results provide the productivity factors necessary for estimating output changes in an important segment of the economy. Moreover, future refinements of the DC index may draw on the FPMS experience. The various government product series that the FPMS identified could be utilized to estimate monthly production directly on the basis of output data rather

[6] Quantifying the output of services can present problems, but is often easier than it might seem at first blush. Haircuts are an obvious measure of barber production, for example, and court cases might be used to index the output of lawyer services.

[7] The manufacturing output indexes created by the Federal Reserve Bank of Richmond use two inputs, employment and electrical power usage, to estimate changes in output. While both are input measures, they are accounted for in physical units, just as is the production series, rather than in monetary terms.

than indirectly via labor proxies related to inputs. Thus, the large percentage share of labor-based series could be reduced and the number of product series increased in the DC index.

In several instances, production indexes are currently represented in the DC index both by an output series and by an input (employment) series. Rail transportation production, for example, is represented both by the number of AMTRAK passengers and by hours worked by railroad employees; telephone production is represented both by the number of business calls and by communications employment; and so on. When an activity is represented by two series, the SIC weight is split on the basis of their relative significance or in 50-50 proportion between the output series and the labor series, respectively.

The DC index is adjusted for workdays and seasonal variations. Workdays within any month vary from year to year, and seasonal variations occur as well. In making the workday adjustments, it was necessary to establish a normal workweek for each production category. Hotels, for example, do not normally close on weekends, while many retail or banking establishments close on Sundays or perhaps on both Saturday and Sunday.

The DC index is a base-period weighted, Laspeyres-type index. Individual production indicators were assigned weights based on their shares of total value added in 1986, the most recent year for which gross state product data are available. The value-added weights are derived from the 1986 Gross State Product figures published by the Bureau of Economic Analysis of the U.S. Department of Commerce.

Results

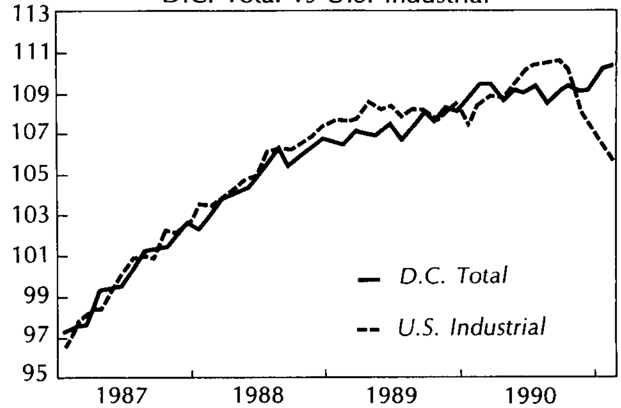
The seasonally adjusted values for the DC index over the past four years are charted here along with the seasonally adjusted values of the U.S. Index of Industrial Production (Chart 1). The DC index shows the behavior of total production in the District of Columbia since early 1989 to have been quite different from the behavior of U.S. industrial production.

Total production in the District of Columbia grew (on a December-to-December basis) at a rate of 4.0 percent in 1988, 1.2 percent in 1989, and 1.1 percent in 1990 despite its essentially flat path over most of that year. Chart 1 indicates that, according to the DC index, the economy of the District of Columbia

Chart 1

INDEXES OF PRODUCTION

D.C. Total vs U.S. Industrial



slowed in 1989, peaked in January 1990, showed no clear trend through December 1990, and rose in early 1991. It should be noted that the index reflects increases in labor productivity assumed in connection with measuring some output components by using labor data proxies.

The DC index covers all goods- and service-producing industries included in the Standard Industrial Classification. Normally, production is classified into four major areas: primary production (agriculture and mining), secondary production (manufacturing and construction), tertiary production (transportation, communications, utilities, retail and wholesale trade), and quaternary production (finance, insurance, and real estate, services and public administration). In the DC index, however, production is classified in three areas—goods, tertiary services, and quaternary services—because primary productive activity (agriculture and mining) is virtually nonexistent in the District of Columbia. Separate tabulations are made also for a total services index which combines tertiary and quaternary production, a private sector index which includes everything but government (federal and local) production, and a public sector index that includes only federal and local government activity. The appendix to this article tabulates the monthly values for all of these indexes from January 1987 through early 1991 (Table 5).

In the DC index, goods production accounts for about 11 percent of total production. This 11 percent is divided mainly between construction (7.1 percent) and printing and publishing (2.6 percent).

Goods production has been volatile in recent years and has exhibited some weakness since late 1988. Chart 2 compares the behaviors of the DC index goods component with the Maryland-D.C. index of manufacturing compiled by the Federal Reserve Bank of Richmond. Construction activity in the District of Columbia dominates the DC goods index, while manufacturing activity in Maryland dominates the MD/DC manufacturing index. It is understandable, therefore, why the two indexes tell different stories about cyclical swings in these respective activities in the vicinity of the nation's capital. In particular, the severity of the recent recession in the D.C. construction sector is clearly evident.

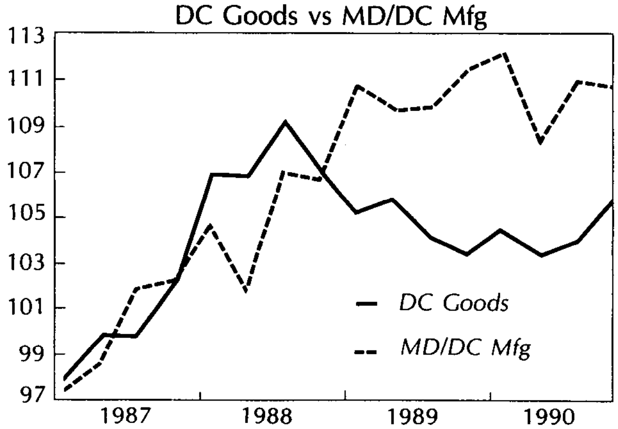
Services production accounts for about 89 percent of total production in the District of Columbia (vs. 68 percent nationally). The growth in D.C. services production slowed to 1.8 percent in 1989 from an annual rate of 4.4 percent in 1988 (December/December). The DC services-production index peaked in January 1990, stayed at or below that peak through the year, and then rose above it in early 1991. By way of comparison, the national services index[8] grew less rapidly in 1988, its growth did not slow in 1989, and it did not stop growing until late in 1990. D.C. services production did decline briefly after July 1990, the month marking the beginning of the recent national recession.

[8] An experimental index developed by Zoltan Kenessey, circulated by the Coalition of Service Industries.

Chart 2

INDEXES OF PRODUCTION

DC Goods vs MD/DC Mfg



Private production in the District of Columbia was more volatile than government production, as one would expect, partly because private production includes goods production (all government production is by definition services production). The growth in private production was a vigorous 5.1 percent in 1988, then declined to 1.9 percent in 1989 and 1.2 percent in 1990. Not all of the greater volatility in private production was due to goods production; private services production was also somewhat more volatile than government (services) production. Private services production grew an estimated 5.9 percent in 1988 (compared to a 2.5 percent increase in government production), then slowed to 3.1 percent in 1989 and to 0.9 percent in 1990 (compared to growth of 0.1 percent and 0.9 percent in 1989 and 1990, respectively, in government production).

Tertiary services production (wholesale and retail trade, transportation, communication and utilities) and quaternary services production (finance, insurance, real estate, business and personal services, and government) behaved similarly in the District of Columbia over the period studied. The growth in tertiary production was less even than the growth in quarternary production, however, as was exemplified by the sharp decline in tertiary production in late 1990.

The first results for the Total Production Index for the District of Columbia indicate that output in the nation's capital peaked in March of 1990, but stayed roughly flat through the year, even when the U.S. economy went into recession. Components of the DC index generally confirm the stabilizing role played by the high proportion of services production in the District of Columbia. D.C. goods production, which is heavily concentrated in construction, peaked in August 1988 and has remained well below that peak through early 1991. The DC index figures are just estimates, of course; the index likely understates the magnitude of the downturn in economic activity in the District of Columbia because labor productivity in recessions usually declines rather than rises as has been assumed for the entire period.

References

Bechter, Dan M., Christine Chmura, and Richard Ko. "Fifth District Indexes of Manufacturing Output," Federal Reserve Bank of Richmond, *Economic Review* 74 (May/June 1988).

Burns, Arthur F. "The Measurement of the Physical Volume of Production," *Quarterly Journal of Economics* 44 (February 1930): 243.

Day, Edmund E. "An Index of the Physical Volume of Production," *Review of Economic Statistics* 2 (1920): 246-59, 309-37, 361-67.

Federal Reserve Board. "Indexes of Trade and Production," *Federal Reserve Bulletin* 8 (March 1922): 292-96.

__________. "Index of Production in Selected Basic Industries," *Federal Reserve Bulletin* 8 (December 1922): 1414-21.

__________. *Industrial Production*. 1986 Edition. With a Description of the Methodology.

Kenessey, Zoltan. "Experimental Indexes of Services Production," 50th Anniversary Conference on Research in Income and Wealth, Washington, D.C., May 1988.

__________. "A New Index of Service Production: A Companion to the Index of Industrial Production," American Economic Association, Annual Meeting, New York, N.Y., December 1988.

__________. "Monthly Total Production Indexes for Regions," Paper presented at the 37th North American Meeting of the Regional Science Association International, Boston, Massachusetts, November 9-11, 1990.

Mitchell, Wesley C. *History of Prices During the War*, War Industries Board Price Bulletin, No. 1 (1919): 44-46.

U.S. Bureau of Labor Statistics. Productivity Statistics for Federal Government Functions, Fiscal Years 1967-88 (February 1990).

Appendix

A Tabular Walk-Through the Calculation of the Total Production Index for the District of Columbia

Table 1

Menu for Calculating a Product Component in the Total Production Index when Physical Units are Used to Measure Output

(1) data:	(2) data:	(3) data:	(4) calculate:	(5) calculate:	(6) data:	(7) calculate:	(8) calculate:
Year & Month (indicated by a "t" subscript)	Physical units of output of product "i" in month "t"	Workdays in month "t" for this product (these change from year to year)	Daily average output of product "i" in month "t" c2/c3, or	Value of unadjusted output index for product "i" in month "t"[1]	(from earlier calculation) Seasonal factor for the month for this product[2]	Value of seasonally adjusted daily output index for product "i" in month "t" c5/c6, or	Component value for product "i" to be included in the Total Production Index[3]
	Q_{it}	A_{it}	$Q_{it}/A_{it} = q_{it}$	$100 \times q_{it}/q_{io} = I^u_{it}$	S_{it}	$I^u_{it}/S_{it} = I^a_{it}$	$w_{io} \times I^a_{it} = TPI^a_{it}$

[1] q_{io} = daily average output of this product in the base year (1987).

[2] The seasonal factor for a month (e.g., March) is the same from year to year, and the same for every day in the month. The seasonal factors were computed using the ratio-to-centered-moving-average method: each month's index was calculated as the ratio of its average value over a four-year period, 1987-90, to the average value of the index during the six months before and the six months after this month in this period. The steps in table columns (6) and (7) can be skipped if the Total Production Index is to be unadjusted for seasonal variations.

[3] The constant weight w_{io} is equal to this product's share of the value of total production in the base year. The value of the seasonally adjusted Total Production Index is the sum of all of its components, or $TPI^a_t = \Sigma TPI^a_{it}$.

Table 2

An Example of a Total Production Index Component—Railroad Transportation—Calculated from Physical Units of Production (Number of Amtrak Passengers)

(1) Year & Month	(2) Number of Amtrak passengers	(3) Workdays in this month	(4) Daily average number of passengers for this month (7822.05 in 1987)	(5) Unadjusted index for this activity	(6) Seasonal factor	(7) Seasonally adjusted index for this activity	(8) The DC index component value for this activity (weight = 0.0022)
1988 10	274,036	31	8839.9 = 274036/31	113.01 = 100 x 8839.9/ 7822.05	1.0125	111.61 = 113.01/ 1.0125	0.25 = 0.0022 x 111.61
1988 11	283,698	30	9456.6	120.90	1.0748	112.48	0.25
1988 12	267,107	31	8616.4	110.15	0.9445	116.63	0.26
1989 01	255,402	31	8238.8	105.33	0.8502	123.89	0.27
1989 02	237,194	28	8471.2	108.30	0.8814	122.87	0.27

Table 3

Menu for Calculating a Product Component in the Total Production Index when Employment Units are Used as a Proxy for Output

(1) data:	(2) data:	(3) data:	(4) calculate:	(5) calculate:	(6) data:	(7) calculate:	(8) calculate:
Year & Month (indicated by a "t" subscript)	Employment units used to produce output of product "i" in month "t"	Production factor coefficient (accounts for the estimated constant monthly change in productivity for this product)	Adjusted employment units used to produce output of product "i" in month "t" c2 x c3, or	Value of unadjusted output index for product "i" in month "t"[4]	(from calculations made prior to table construction) Seasonal factor for this product for this month	Value of seasonally adjusted index for product "i" in month "t" c5/c6, or	Component value for product "i" to be included in the Total Production Index
	E_{it}	F_{it}	$E_{it} \times F_{it} = L_{it}$	$100 \times L_{it}/L_{io} = I^u_{it}$	S_{it}	$I^u_{it}/S_{it} = I^a_{it}$	$w_{io} \times I^a_{it} = TPI^a_{it}$

[4] L_{io} = average monthly labor input used to produce this product in the base year (1987).

Table 4

An Example of a Total Production Index Component—Construction—Calculated from Employment Units of Production (Construction Worker Hours)

(1) Year & Month	(2) Number of construction worker hours (in thousands)	(3) Construction labor production factor coefficient this month (increases 0.35%/mo.)	(4) Adjusted number of construction worker hours (in thousands for month) (average = 15.032 in 1987)	(5) Unadjusted index for this activity	(6) Seasonal factor	(7) Seasonally adjusted index for this activity	(8) The DC index component value for this activity (weight = 0.0626)
1988 10	14.3	1.081	15.46 = 14.3 x 1.081	102.85 = 100 x 15.46/ 15.032	1.0026	102.58 = 102.8468/ 1.0026	3.66 = 0.0626 x 102.58
1988 11	14.5	1.085	15.73	104.66	1.0116	103.45	3.69
1988 12	14.2	1.089	15.46	102.85	0.9913	103.76	3.70
1989 01	13.8	1.093	15.08	100.31	0.9621	104.26	3.72
1989 02	13.5	1.097	14.80	98.48	0.9701	101.51	3.62

Table 5

Total Production Index for the District of Columbia

Seasonally Adjusted; 1987 = 100

Type/Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Q1	Q2	Q3	Q4	Year
TOTAL																	
1987	97.3	97.6	97.7	99.2	99.4	99.6	100.2	101.3	101.3	101.5	101.9	102.7	97.6	99.4	100.9	102.0	100.0
1988	102.3	103.1	103.8	104.1	104.3	104.8	105.6	106.3	105.5	106.0	106.4	106.8	103.1	104.4	105.8	106.4	104.9
1989	106.6	106.5	107.2	107.0	107.0	107.6	106.6	107.4	108.2	107.7	108.3	108.1	106.7	107.2	107.4	108.0	107.3
1990	109.4	109.4	109.5	108.7	109.2	109.1	109.4	108.4	109.1	109.4	109.1	109.3	109.4	109.0	109.0	109.2	109.2
1991	110.3	110.4															
GOODS																	
1987	98.0	97.7	98.0	99.8	100.2	99.7	99.9	98.4	101.2	100.9	100.1	105.3	97.9	99.9	99.8	102.1	99.9
1988	104.9	106.4	109.7	107.1	105.5	108.3	108.7	111.3	107.8	105.9	108.9	106.6	107.0	106.9	109.3	107.1	107.6
1989	106.7	104.8	103.7	105.6	106.9	105.0	104.5	104.1	104.0	103.7	103.5	103.2	105.1	105.8	104.2	103.4	104.6
1990	104.1	105.0	104.2	103.2	103.4	103.4	104.1	103.8	103.9	106.2	105.9	105.7	104.5	103.4	104.0	105.9	104.4
1991	105.9	105.7															
SERVICES																	
1987	97.3	97.6	97.7	99.1	99.3	99.6	100.3	101.7	101.3	101.6	102.1	102.3	97.5	99.3	101.1	102.0	100.0
1988	102.0	102.7	103.1	103.7	104.1	104.4	105.2	105.7	105.3	106.1	106.1	106.8	102.6	104.1	105.4	106.3	104.6
1989	106.5	106.7	107.6	107.2	107.0	107.9	106.8	107.8	108.7	108.2	108.9	108.7	106.9	107.4	107.8	108.6	107.7
1990	110.1	109.9	110.1	109.3	109.9	109.8	110.1	108.9	109.7	109.8	109.4	109.7	110.0	109.7	109.6	109.6	109.7
1991	110.9	111.0															
TERTIARY																	
1987	97.2	98.5	98.6	98.5	99.6	98.0	99.5	102.5	100.6	102.4	103.1	102.5	98.1	98.7	100.9	102.7	100.1
1988	105.3	104.0	105.9	105.5	106.6	106.5	107.4	107.2	105.9	109.2	107.0	107.6	105.1	106.2	106.9	107.9	106.5
1989	107.0	107.0	107.0	107.4	107.1	108.6	108.1	105.3	109.4	106.3	110.1	108.9	107.0	107.7	107.6	108.5	107.7
1990	108.5	109.3	108.1	109.4	108.0	109.5	108.5	108.9	109.2	107.9	106.4	106.7	108.6	108.9	108.9	107.0	108.4
1991	109.2	107.7															
QUATERNARY																	
1987	97.3	97.4	97.5	99.2	99.2	99.9	100.4	101.5	101.4	101.4	102.0	102.3	97.4	99.5	101.1	101.9	100.0
1988	101.4	102.5	102.6	103.4	103.7	104.0	104.8	105.4	105.1	105.5	106.0	106.7	102.2	103.7	105.1	106.1	104.3
1989	106.5	106.6	107.7	107.1	107.0	107.8	106.6	108.3	108.6	108.5	108.6	108.7	106.9	107.3	107.8	108.6	107.7
1990	110.4	110.0	110.5	109.3	110.2	109.9	110.4	108.9	109.8	110.1	110.0	110.2	110.3	109.8	109.7	110.1	110.0
1991	111.2	111.6															
PRIVATE																	
1987	96.7	97.2	97.2	99.2	99.4	99.9	99.6	100.9	101.4	102.1	102.5	103.6	97.1	99.5	100.6	102.8	100.0
1988	103.0	104.1	105.2	105.3	105.7	106.3	107.6	108.5	107.7	107.9	108.3	108.9	104.1	105.8	107.9	108.4	106.5
1989	108.5	108.3	109.4	109.1	109.4	110.3	109.0	110.0	111.1	110.3	111.4	111.0	108.7	109.6	110.1	110.9	109.8
1990	112.6	112.4	112.6	111.6	112.2	111.9	112.9	112.0	112.1	112.5	112.0	112.3	112.5	111.9	112.3	112.3	112.3
1991	113.6	113.8															
PUBLIC																	
1987	98.3	98.2	98.5	99.0	99.4	99.3	101.2	102.0	101.1	100.6	101.0	101.1	98.3	99.3	101.4	100.9	100.0
1988	101.2	101.6	101.7	102.2	102.2	102.5	102.5	102.9	102.3	103.2	103.5	103.6	101.5	102.3	102.5	103.4	102.4
1989	103.7	103.7	103.7	103.7	103.5	103.5	102.7	103.4	103.6	103.7	103.6	103.7	103.7	103.6	103.3	103.6	103.5
1990	104.6	104.7	104.7	104.2	104.6	104.9	104.2	102.8	104.6	104.6	104.5	104.6	104.7	104.6	103.8	104.6	104.4
1991	105.3	105.3															

The Stealth Budget: Unfunded Liabilities of the Federal Government

*Roy H. Webb**

The federal budget is important. It is the basis for planning government programs, it is a significant element in plans of individuals in the private sector, and it is the starting point for assessing the federal government's current impact on macroeconomic conditions. Past budgets are used to study significant economic questions, such as the extent to which federal fiscal actions affect aggregate output, prices, and interest rates.

The traditional statement of the federal budget provides important information about current receipts and expenditures, but is nevertheless incomplete. Actions have been taken that will require spending in the future; provision for that future spending does not, however, appear in the budget accounts. As a result, stated federal spending does not reveal the total resource demands placed on the private economy and stated federal debt does not reveal the full tax burden that taxpayers will face in the future. In other words, a stealth budget that is unseen by most observers will generate future taxing and spending.[1]

The stealth budget is not trivial. The programs discussed in this paper had unfunded liabilities in 1989 in excess of $4 trillion. To put that number in perspective, the conventionally stated gross federal debt in that year was less than $3 trillion.

Although the conventional federal budget omits important information when unfunded liabilities are present, there is a straightforward alternative that would produce a more revealing budget: explicitly state the present value of expected future spending when a program is created. In addition, each future budget could restate that amount due to either the passage of time or legislative revisions.

The next section of this paper will discuss some of the federal programs that have created unfunded liabilities. The focus will be on only those programs (1) that promise specific benefits to specific persons and thus resemble private contracts,[2] or (2) for which current or past actions make future action unavoidable. Deposit insurance, for example, promises an exact benefit to particular deposit holders; and the creation of nuclear waste as a byproduct of weapons production makes disposal or treatment essential. Other federal spending programs that predictably pay benefits but are not embodied in current legislation will not be considered. For example, if a drought reduces crop yields it is virtually certain that Congress will enact a payment scheme; the exact payments to particular individuals, however, are impossible to guess.

SPECIFIC PROGRAMS

Many programs that create unfunded liabilities will be discussed in this section. Each will be briefly

* The author gratefully acknowledges helpful comments from William E. Cullison, Robert Hetzel, Thomas M. Humphrey, Anatoli Kuprianov, and Marvin M. Phaup, Jr., and valuable research assistance from Craig Carlock. A version of this paper was presented to the Western Economic Association International Conference, July 1990. The views and opinions expressed in this article are solely those of the author and do not necessarily represent those of any other person or of the Federal Reserve Bank of Richmond.

[1] The traditional source for fiscal information is *The Budget of the United States* that is prepared by the Office of Management and Budget (OMB) for each fiscal year. Its presentation of future liabilities has improved in recent years. The 1991 and 1992 Budget each contain a section that is analogous to the footnotes in a corporate annual report; that section discusses many, but not all, of the unfunded liabilities discussed in this paper. The content of that section has also changed between the two years, and has changed from similar information presented in the *Special Analyses* book in the set of budget documents for prior years. There is no summary table that has been consistently presented over time that would facilitate discussion of the future resource demands that the federal government has committed to placing on persons and firms in the future.

[2] Legislated promises are of course not exactly equivalent to private contracts. An individual may not voluntarily agree to participate in a program such as Social Security but may still be compelled to participate. Also, if the government later reneges on its promises, there is often no legal recourse for the individual.

described. In cases where the present value of unfunded liabilities can be at least roughly estimated, an estimate will appear in Table 1 and the methodology will be briefly explained. Each entry will be a present value of expected future real payments by the government, net of expected future real receipts, as of the end of the government's 1989 fiscal year.[3] In other cases the source of possible taxpayer liability will be mentioned in Table 2.

To understand most of the programs listed it is important to distinguish between fiscal actions and financial intermediation. Any program that is in essence a combination of taxing and spending is a fiscal program. Many federal fiscal programs are obscured by being described in the language of insurance or banking. For example, a bona fide insurance company will attempt to set premiums on an actuarial basis and will hold sufficient reserves to pay expected future claims. A fiscal program masquerading as an insurance program will set low premiums that have little relation to risk and are insufficient to cover the expected value of future payments. Similarly, a commercial lender will attempt to charge sufficient interest or other fees to compensate for any credit risk; a disguised fiscal program will lend at low rates to poor risks.

Why is the language so obscure? The Appendix to this paper presents some elements of political economy that help explain the incentives for elected officials to use language that fails to reveal the full cost of many programs.

[3] *Present value* is the value of a future stream of cash flows adjusted for the time value of money. For example, a single payment P received in N years over which the market rate of interest is R would have a present value $P(1+R)^{-N}$. For a series of payments the individual items can simply be added together. To adjust for inflation it is often helpful to express the cash flow in constant dollars, or in *real* terms; a series of real cash flows is properly adjusted by using a *real interest rate*, which is the difference between a market rate of interest and expected inflation. In this paper a real rate of 4 percent is used in several calculations, reflecting a market rate of 8 percent and expected inflation of 4 percent. Those values are approximately correct for September 1989, the particular point in time that is used for the calculations. Uncertainty is addressed by looking at *expected* cash flows. An expected value is the product of the value if some event occurs times the probability of that event occurring; those products are then calculated and added over all possible events. For example, if you receive a dollar if a coin flip is heads and a dime if it is tails, the expected value of a coin flip is 55 cents.

By using these definitions, one can compute values that make sense when they are added together. The entries in Table 1 are all present values of expected real cash flows.

Towe (1990) has a good discussion that relates present values of expected cash flows to government budgets, particularly his section on the "actuarial balance" of particular programs.

Deposit Insurance

Deposit insurance has become a well-known example of the type of program that can create future liabilities. It was first offered by the federal government in the 1930s and is now raising the level of federal spending. In some years the insurance system was labled "off-budget" and therefore was not included in spending and deficit calculations. In other years cash payments and expenditures were included in the budget, but no mention was made of rapidly growing future taxpayer liabilities for deposits in insolvent institutions. When major changes in the law raised the expected value of future payments to insured depositors, such as the 1980 increase in the amount of deposits covered, those higher payments did not raise stated spending or debt. Even today the budgeted liability understates the likely total taxpayer expenditure.

Deposits up to $100,000 at banks, savings and loan associations, and credit unions are explicitly insured by federal agencies. In addition, the Federal Deposit Insurance Corporation (FDIC) has treated large banks as "too big to fail" and has extended de facto insurance to uninsured depositors and other creditors.[4] Prior to 1989 legislation (the Financial Institutions Reform, Recovery, and Enforcement Act, or FIRREA) depositors at savings and loan associations were insured by the Federal Savings and Loan Insurance Corporation (FSLIC); they are now insured by the FDIC's new Savings Association Insurance Fund. Bank depositors who were insured by the FDIC are now covered by the FDIC's Bank Insurance Fund. Credit union depositors are insured by the National Credit Union Association's Share Insurance Fund.

Savings and Loan Associations The FIRREA acknowledged a liability of $115 billion over three years, to be paid by taxpayers and by higher insurance fees. Many assumptions behind that number were too optimistic, however. The Secretary of the Treasury (Brady 1990) has estimated that costs will be between $90 billion and $130 billion, in addition to funds already spent.

The way that such a large liability was accrued is instructive and will briefly be described below.[5] FSLIC insurance was established in 1934; it allowed

[4] Todd and Thompson (1990) describe the logic and evolution of the idea that some banks are too big to fail.

[5] For more complete discussions, see Benston and Kaufman (1990) or Dotsey and Kuprianov (1990).

savings and loan associations to substantially increase their leverage, thereby increasing their returns but also increasing the risk that they would not be able to make promised payments to depositors and other creditors. Holding short-term liabilities and long-term assets with fixed returns, the industry was especially vulnerable to interest rate risk. In the 1970s nominal interest rates rose sharply and reduced asset values and the net worth of the industry; the market value of many associations became negative. The political system responded perversely. First, the problem was denied—accounting procedures were altered to obscure the losses that had already occurred. Second, the problem was worsened—the Monetary Control Act in 1980 raised the amount of insurance coverage from $40,000 per account to $100,000, thereby making it easier for insolvent institutions to raise funds. By 1982 much of the savings and loan industry was economically insolvent.[6] A policy of regulatory forbearance kept insolvent institutions from being closed. They were instead allowed to make risky loans funded by insured deposits. Many of the risky loans failed and thus further raised the taxpayer burden that is now being recognized.

Table 1

Unfunded Liabilities of the Federal Government

Billions of 1989 Dollars

Program	Estimated Amount
Savings and loan deposit insurance	130
Social Security	
Retirement and disability benefits	1,052
Health benefits	1,412
Railroad retirement	30
Federal employee retirement and disability benefits	
Civil service	643
Military	513
Medical benefits	279
Pension Benefit Guarantee Fund	16
Crop insurance	25
Flood insurance	5
Defense nuclear waste disposal	68
Loans and loan guarantees by government agencies	77
Total	4,250

Note: The sources of the estimates are described in the text. Each estimate is the present value at the end of the government's 1989 fiscal year of expected real future spending net of any offsetting receipts.

The entry in Table 1 for unfunded savings and loan insurance is $130 billion. It represents the upper bound of the Treasury Secretary's admitted range, which was stated in 1989 dollars. The upper bound is used since all previous official estimates have substantially understated the cost of deposit insurance for savings and loan associations. That estimate is consistent with others prepared by independent analysts; one range was given as $86.5 billion to $136.4 billion (Brumbaugh, Carron, and Litan, 1989). Confusing the issue are competing estimates that add in future nominal interest costs that would result from borrowing the funds to be spent. Those estimates are difficult to interpret and are ignored in this paper.

The official estimates may still be conservative. The perverse incentives created by deposit insurance still exist. Also, the solvency of existing thrift institutions is often overstated by conventional accounting procedures. Until those factors change it is likely that some thrifts will create additional liabilities for Savings Association Insurance Fund and the taxpayer. In addition, the official estimates assume that the assets of failed associations will be sold in a prompt and efficient manner. Kane (1991a), however, estimates that the disposal agency, the Resolution Trust Corporation, cost taxpayers $40 billion in its first year of operation by mismanaging the assets of failed savings and loan associations, with additional costs likely in the future.

Banks The banking industry shares some important similarities with the savings and loan industry several years ago.

(1) Deposit insurance has given banks the incentive to lower their holdings of capital.

(2) Poorly capitalized banks are allowed to stay in business. One study found 30 banks without any capital on a risk-adjusted basis in mid-1989, and another 31 with capital below 3 percent of deposits (Brumbaugh and Litan, 1990). That study was based on conventional accounting data.

---

[6] *Economically insolvent* means that the market value of liabilities, including deposits, is greater than the market value of assets such as loans. It is possible for an institution to be solvent according to an accounting system, but to be economically insolvent. This could occur if loans are assigned higher values than realistic estimates of future cash flows, or if assets such as goodwill are given positive values on the balance sheet but not in the market. According to Benston and Kaufman (1990), "By 1982 some two-thirds of the [savings-and-loan] industry was economically insolvent, with aggregate negative net worth of about $100 billion."

(3) Banks state assets and liabilities at book value rather than market value. Many banks have thereby overstated asset values. Loans to impoverished third-world governments, for example, are routinely traded in private markets at lower values than are recognized by some large banks.

(4) Barriers to branching result in loan portfolios that are not regionally diversified and are therefore vulnerable to localized shocks to the economy. Just as banks and savings and loans in Texas in the mid-1980s were vulnerable to the weak regional economy, banks in the Northeast are now feeling effects of a regional economic downturn.

(5) The FDIC is paying more to close insolvent banks than it is receiving in premiums. In 1990 the bank insurance fund lost $3.5 billion, in 1989 it lost $2.0 million, and in 1988 it lost $4.2 billion.[7]

The parallels with the thrift industry are not exact. Many observers (for example, *The Economist* [1991] and analysts quoted in Rehm [1991b]) believe that on average banks are more profitable, better capitalized, better managed, and better regulated than were savings and loan associations in the 1970s and 1980s.

Without detailed knowledge of the market value of individual banks' assets and liabilities, it is impossible to estimate losses the FDIC will incur. It is therefore impossible to estimate the expected loss to taxpayers due to insurance of bank deposits and other liabilities. One estimate, Kane (1991b), puts the cost to taxpayers at roughly $40 billion. A more optimistic view has been stated by the head of the FDIC, in essence that the present value of future bank premiums for deposit insurance is large enough to close insolvent banks, pay liability holders, and rebuild the Bank Insurance Fund. This view is also held by Ely (quoted in Kleege [1991]) who stated "Losses of this amount [$20 to $40 billion to close insolvent banks in the near future] . . . can be fully paid by the banking industry."

No estimate of taxpayer liability is therefore made. Instead, the face value of insurance provided banks is entered in Table 2, consisting of the deposits of the banking system at the end of 1989.[8]

Credit Unions Credit unions also offer insured deposits. According to one study,[9] although 86 insolvent credit unions are being allowed to remain open, another 122 have very low capital, and another 294 have substandard capital, their insurance fund is unlikely to require taxpayer assistance. Table 1 therefore contains no entry for credit unions. Their total deposits are listed in Table 2 as an insured liability of the government.

Social Security

In 1935 the Social Security system was founded as a mandatory old-age pension plan with benefits loosely based on prior taxable earnings. Benefits, the tax base, and tax rates have been substantially increased over time. The most notable increase in benefits occurred when health insurance was introduced in 1965. The system has always had unfunded liabilities. At times the payroll tax collections were so far below benefit levels that the necessity for major change was obvious. The last such occurrence was in 1983, when Congress cut projected future benefits and substantially raised taxes. The system is now enjoying record annual surpluses of cash receipts over expenditures.

Despite its apparent prosperity, many estimates show substantial future liabilities for the system. The trust fund for hospital insurance is projected to be exhausted by 2006.[10] At that point, current taxes will not pay current benefits and there will be no cushion to draw on. And as the baby boom generation begins to receive retirement benefits, the retirement and disability funds will also decline and eventually become exhausted.

The 1992 Budget contains a range of estimates for the present value of future liabilities for the Old Age and Survivors Insurance and the Disability Insurance Funds. Using a midrange set of actuarial assumptions,

[7] These historical figures describing the Bank Insurance Fund are from the Budget for 1991 and 1992 (1991, Section Two, p. 1115, and 1992, Part Four, p. 1105).

[8] On the one hand, deposits over $100,000 in banks that are not too big to fail are incorrectly included in that entry. On the other hand, some nondeposit debt of banks that are too big to fail is implicitly insured and is incorrectly excluded from that entry. The entry in Table 2 is therefore at best an approximation.

[9] The study by James R. Barth and R. Dan Brumbaugh is discussed in Rehm (1991a).

[10] The source for this estimate and most others in this section is the 1992 Budget, Part II, Chapter VIIIb. A fuller explanation of the programs is given by Aaron, Bosworth, and Burtless (1989).

the funds will become insolvent in the year 2043. Over the next 75 years the present value of that deficit is $1,174 billion. The entry in Table 1, $1,052 billion, is that value augmented for losses more than 75 years out, restated as a present value in 1989.

The Federal Hospital Insurance Trust Fund pays certain medical expenses of elderly Americans. Despite increases in the payroll tax rate and the tax base, spending is growing faster than revenues due to a growing elderly population and rapid growth in the cost of providing medical care. One Treasury projection put the expected future deficit for this program at $312 billion in 1989. Another medical care program, Supplemental Medical Insurance, is funded primarily by general revenues. Spending for that program was $33 billion in 1990 and has been growing rapidly. Assuming that spending growth for that program is only one percent higher than inflation, the present value of spending for Supplemental Medical Insurance is $1.1 trillion. The combined amount for health insurance is $1,412 billion and is entered in Table 1.

Another unfunded liability is a retirement pension program for railroad employees. With three retirees receiving benefits for every employee currently paying taxes, benefit payments are much larger than receipts. The Railroad Retirement Board has received congressional assistance five times in the last 16 years. The 1992 Budget contains an estimate of the unfunded liability of $34 billion. That value, restated for 1989, is listed in Table 1.

Estimates of future Social Security taxes and spending are very sensitive to economic and demographic assumptions such as population and productivity growth, health-care expenses, interest rates, and life expectancy. Any estimated liabilities are thus extremely imprecise. Perhaps more important is the possibility of major changes in the programs. If the economic assumptions are not terribly inaccurate, the growing size of future deficits may lead to substantial changes in taxes, benefits, and even the structure of the medical care industry.

Federal Employee Retirement Benefits

Federal employees are promised retirement and disability benefits, as are many private sector workers. Unlike private firms, the government does not fully accrue reserves to pay those benefits for workers hired before 1985. Also, in some ways the benefits are more generous than those of most private firms. For example, many federal pensions are fully indexed for inflation. The effect is that the cost of federal programs is understated as the full personnel costs are not recognized.

Table 1 contains an entry of $643 billion for civilian employee retirement and disability benefits, which is taken from the 1992 Budget. That amount represents the excess of the present value of expected plan benefits over net assets available for benefits. The funding of retirement benefits for military personnel differs in several details from the civilian program. The 1992 Budget, nonetheless, estimates an unfunded deficit of $513 billion for pre-1985 service. That value is also listed in Table 1.

Federal retirees also receive subsidized health insurance. Agencies' budgets include payments for persons who have already retired but make no provision for future payments for current employees. An admittedly rough estimate of the present value of that amount is $155 billion, the midpoint of a range given in the 1992 Budget. No estimate is made in that document for health benefits for retired military personnel, which include essentially free care in many cases at military facilities. Table 1 presents a rough estimate that the unfunded liability for health care for military retirees has the same proportion to unfunded civilian health care as the unfunded military retirement program has to the civilian retirement program.

Insurance of Private Pensions

The Pension Benefit Guaranty Corporation (PBGC) insures defined benefit pension payments promised by private firms to their workers. In 1989 almost 40 million persons were insured, with promised benefits near $750 billion. Although most defined benefit plans were clearly solvent, some were obviously underfunded.

Before legislation passed in 1987 took effect, a flat premium per covered worker was charged. Premiums now vary according to book values of plan assets and liabilities, but are not completely set on an actuarial basis. Based on plans already terminated the PBGC has a deficit of more than $1 billion; the effect of future pension plan terminations has been projected by many observers to greatly exceed future premium payments at current levels.

Hirtle and Estrella (1990) have simulated pension plan behavior by using Compustat data for 1,512 firms that employ almost 20 million workers. They estimated that plans of those firms would generate future liabilities for the PBGC over the next hundred years with a present value of $27 billion; future premiums, however, would have a present value of $12 billion. Future plan terminations, therefore, have a present value of $15 billion.

That estimate may be conservative. First, it does not cover all insured workers. Hirtle and Estrella point out that as many as 31 million workers may be covered. Second, their simulations do not allow for formation of new firms with defined benefit pension plans that may become insolvent in the future. Third, their dynamic models do not allow for strategic behavior in response to incentives. For example, a firm near insolvency has the incentive to undertake risky behavior. If the risks pay off, managers and equity owners will receive a large return. If the risks fail, creditors, including the PBGC, will bear most of the loss. All three effects would make the PBGC's unfunded liability even greater.

Another possibility is raised by the voluntary termination of defined benefit pension plans, with accrued benefits replaced with annuities issued by insurance companies that may have low quality assets. Although the PBGC does not recognize an obligation to insure such benefits, others believe that a legal or political obligation does exist; in that case the PBGC has stated that such an obligation would add "tens of billions" to the liabilities already insured (Rose and Wessel, 1990). That amount is not included in the tables.

The total unfunded liability of the PBGC for single-employer defined benefit pension plans can therefore be estimated as $16 billion. The largest part is the estimate of Hirtle and Estrella for the unfunded cost of future plan terminations, $15 billion. Adding $1 billion for the deficit from past terminations yields a $16 billion estimate.

Other Insurance Programs[11]

The government has several other programs that are described in the language of insurance. Each promises payments if certain events occur, collects periodic receipts, and may subject taxpayers to future payments if receipts fail to cover expenditures. Some of the programs include flood insurance for owners of buildings in flood-prone areas, crop insurance, war-risk insurance for airplane and ship owners, political-risk insurance for certain foreign investment projects owned by U.S. corporations, and eight life insurance programs for military veterans.

The actuarial soundness of the programs can be hard to assess. Crop insurance has recently been subsidized at the rate of roughly one billion dollars per year. The program's managers are attempting to lower the federal subsidy as a fraction of receipts but are also attempting to raise the fraction of crops that are insured. The two changes would tend to have offsetting effects on total federal spending. The estimate in Table 1 therefore ignores those changes and is simply the present value of current average subsidy payments.

The entry in Table 1 also contains an amount for flood insurance. That estimate was prepared by the agency running the program, and is the amount that would be needed to satisfy policyholder claims in nine out of ten decades. For the other insurance programs mentioned above there is no estimate in Table 1. Instead the face value of the programs is included in Table 2.

Nuclear Waste from Weapons Production[12]

The Department of Energy is responsible for 280 facilities in the nuclear weapons production program. Many of the facilities were built in the 1940s or 1950s and are obsolete. Unavoidable future costs have thus been created; some examples follow. Two facilities have nearly 100 million gallons of high-level wastes in "temporary" storage containers awaiting permanent storage. Leaks in those containers have been a continuing problem, making the necessity for a permanent storage method clear. In addition to leaks of high-level wastes, low-level waste has been put directly into the ground. Substantial soil and groundwater contamination has thus occurred at several sites and needs to be cleaned up. Also, an older nuclear reactor has been taken out of service to avoid substantial safety expenditures; its dismantling is another unfunded liability.

It is not clear what disposal and cleanup methods will eventually be used. As the Secretary of Energy put it, "Today's technology is not sufficiently mature or cost-effective to assure meeting either the Department's goals or the efficient use of public resources" (Department of Energy, 1989). As a result, any estimated cost is highly uncertain. In 1988 congressional testimony, one Energy Department employee

[11] This section is based primarily on the 1992 Budget, Part Two, Chapter VIIIa.

[12] This section is based on Alvarez and Makhijani (1988), United States General Accounting Office (1988), and United States Department of Energy (1989).

put the cost at $100 billion. The General Accounting Office later gave a range of $100-$130 billion. Apparently, neither is a present value, but instead represents spending over a lengthy period. To state the numbers in the same form as the rest of the paper, it is assumed that outlays of $5 billion per year (1989 dollars) for 20 years will dispose of existing nuclear waste and put abandoned production sites in conformity with civilian environmental standards. The present value is $68 billion. It should be emphasized that it is a very imprecise estimate.

Loans and Loan Guarantees[13]

Many government agencies have made loans to individuals and firms; the outstanding volume in 1989 was $207 billion. Programs with more than $10 billion of outstanding debt include foreign military sales, agricultural credit insurance, rural housing insurance, agricultural export credit, and rural electric and telephone utilities. There are also a host of smaller loan programs.

The outstanding volume of direct loans has been declining, but has been more than replaced by loan guarantees. Federal agencies guaranteed $588 billion of primary credit (that is, net of secondary loan pools) at the end of 1989. Programs generating more than $10 billion of loan guarantees include student loans, loans to small businesses, and housing loans from the Federal Housing Administration (FHA), the Government National Mortgage Association, and the Veterans Administration (VA).

Government loans and loan guarantees enable recipients to obtain credit on better terms than would be available in private markets. Some favored parties include poor credit risks and other borrowers who commit less collateral for government credit than would be required by private creditors. Government lending to such parties creates an obvious credit risk for taxpayers. The failure to provide adequate loan loss reserves for outstanding loans certainly creates an unfunded liability.

An example of a lending agency creating an unfunded liability is the Farmers Home Administration (FmHA). The agency lends to farmers unable to obtain credit from normal commercial lenders. According to one report,[14] many of the borrowers lose money due to poor farming practices, such as inadequate care of livestock and crops, or planting on poor land. After defaulting on an FmHA loan, such a borrower is then able to obtain new loans from the same agency. According to the 1991 Budget, the FmHA credit fund had therefore reached a negative net worth of $28 billion.

The 1992 Budget contains estimates for the value of expected losses on loans and loan guarantees made in 1990 and before. For direct loans the expected loss rate is 23.4 percent of the amount of outstanding loans. For loan guarantees the expected loss rate is 4.8 percent. Each loss rate is then applied to the volume of outstanding loans at the end of 1989 and the figure entered in Table 1.

Those figures do not include many activities of government-sponsored enterprises (GSEs), which had lent $763 billion through 1989.[15] GSEs are organizations that have federal charters and some degree of private ownership mixed with some degree of government control. Prominent GSEs include the Federal National Mortgage Association, the Federal Home Loan Banks, the Federal Home Loan Mortgage Corporation, the Student Loan Marketing Association, the Farm Credit Banks, and the Federal Agricultural Mortgage Corporation. Debt issued by a GSE does not have explicit backing by the government but is widely believed to have an implicit guarantee. Evidence of this implicit guarantee can be seen in credit markets, where GSE debt carries a higher interest rate than comparable Treasury debt, but a lower rate than the safest corporate debt.

As with any financial intermediary, a GSE is subject to credit and interest rate risk. The 1992 Budget judges those risks to the taxpayers from current operations to be small. No attempt is therefore made to estimate any taxpayer liability that might occur due to GSE activity; the amount of their lending is listed in Table 2.

There remains the risk that a GSE could change its management strategy in ways that increase risks to the taxpayer. That potential has led to proposals to lessen or eliminate that risk. They include full privatization, increased capital requirements, or the mandatory issuance by GSEs of subordinated debt that is explicitly not guaranteed.

---

[13] This section and the next two sections are primarily based on the *Special Analyses* documents (1989 and 1990), General Accounting Office (1989), and the 1992 Budget.

[14] The General Accounting Office report is cited in Bovard (1988).

[15] A good explanation of the structure of GSEs and the evaluation of their financial risk is given by the Congressional Budget Office (1991).

Table 2

Sources of Possible Liabilities of the Federal Government

Billions of 1989 Dollars

Program	Insured Amount
Insurance of bank deposits	2,175
Insurance of credit union deposits	164
War-risk insurance	239
Veterans life insurance	27
Political-risk insurance of direct investments abroad	9
Lending of government-sponsored enterprises	763
Total	3,377

Note: Each insured amount is a value subject to implicit or explicit government insurance at the end of the 1989 fiscal year. No estimate of expected taxpayer liability is calculated.

CONCLUSION

The stealth budget is enormous. As indicated in Table 1, estimates of unfunded liabilities in a few areas of the federal budget exceeded $4 trillion. Such disparate areas as civil service retirement benefits, deposit insurance for thrift accounts, and disposal of defense-related nuclear waste will contribute to future spending. To put that number in perspective, total federal spending in 1989 was $1.1 trillion and gross federal debt at the end of the 1989 fiscal year was $2.9 trillion.

The $4 trillion estimate is most likely to err on the low side. Several federal insurance programs may produce losses, but the amount is difficult to quantify. The face value of that insurance approached $3.4 trillion.

The stealth budget should concern macroeconomists. The extent to which federal debt affects consumer spending has been the focus of many empirical papers, with conflicting evidence produced.[16] The existence of $4 trillion of unfunded liabilities suggests substantial measurement error in conventional time series of federal spending, debt, and deficits. In general, any conventional measurement of the wealth or income effect of fiscal actions is likely to be misspecified.

The stealth budget should also concern supporters of balanced-budget or other spending-limit legislation. Current examples of such proposals would not constrain unfunded liabilities. As a result, attempts to limit stated spending may simply change the form of spending. For example, a loan guarantee to an insolvent borrower could easily replace a direct subsidy.

Finally, the stealth budget should concern anyone who believes that better information leads to better public policy choices. The magnitude of unfunded liabilities suggests that many decisions by voters and by their elected representatives have been made without a full understanding of either the government's current fiscal position or of the full costs of programs under consideration.

While the estimates in this paper show that substantial unfunded liabilities do exist, the numerical total should be recognized as crude at best. Better estimates for many programs could be produced by the agencies themselves. Their specialists with full knowledge of the programs and with informed access to relevant data, subject to comprehensive review by interested persons outside the agencies, could reveal a wealth of information. Those estimates could then be presented in a consistent format over time to allow easy access to the estimates by non-specialists. Unfortunately, as the Appendix suggests, the very incentives to create unfunded liabilities are also incentives to obscure their costs.

---

[16] A survey of some recent papers is Barth et al. (1991).

APPENDIX

The Political Economy of Unfunded Liabilities

Why does the government have unfunded liabilities? An observer with little information might guess that simple historical accident could explain their existence. Another guess might be that poor management of basically sound programs has allowed some unfunded liabilities to emerge. In either case, a little tinkering would fix the system, eliminate unfunded liabilities, and make the budget more transparent.

The point of this section is to argue that the existence of unfunded liabilities is not accidental. Instead, the American political system has characteristics that produce incentives for politicians—that is, elected officials and their senior-appointed subordinates—to deliberately fail to fund or to fully reveal liabilities that result from current programs. To motivate this interpretation, some key features of a model of political activity will be briefly described below. A fuller discussion of most of these elements can be found in Downs (1957).

Rationally Ignorant Voters

Voters acquire information as long as the marginal benefit of doing so exceeds the marginal cost. A major benefit of voting could occur if a particular voter happened to cast the deciding ballot in an election. The expected value of voting for that reason, however, is very low since the probability that a national election would be decided by a single vote is extremely low. Other benefits of an individual vote, such as expressing an opinion or promoting good citizenship, can also be small. As a result, the marginal benefit of acquiring information is typically very small and voters accordingly acquire little information on candidates and issues.

Vote-Maximizing Politicians

If a politician does not maximize the number of votes received, he or she can be replaced by one who does. It is therefore assumed that all politicians are vote maximizers. A corollary is that politicians are primarily motivated by the prospect of holding office, rather than by ideology.

Interest Groups

Interest groups can lower voter costs of acquiring information on a small subset of issues, can inform politicians on voter attitudes, and can acquire and distribute resources in political campaigns. Interest groups are often formed around issues that affect voter incomes and wealth, although other types of interest groups are also possible.

A political system that contains the above elements can be expected to behave in a predictable manner. A few predictions are given below.

Politicians Respond to Interest Groups

A small tax on all taxpayers may not affect many votes. If all the funds are distributed to a small number of voters represented by a single interest group, however, voting behavior of that group's members may well be changed. If the presence or absence of that program makes a large difference to the wealth of the interest group's members, many (who are rationally ignorant on other issues) will choose to vote for the candidate most strongly supporting that program.

Hidden Costs

A politician can gain support by transferring wealth to members of interest groups. To the extent that the resulting costs can be hidden from any voters who pay them, the politician can benefit from a spending program without suffering adverse consequences from the resulting taxes.

Optimal Ambiguity by Politicians

In order to appeal to a wide range of voters, vote-maximizing politicians will often "becloud their policies in a fog of ambiguity" (Downs, p. 136). By not stating positions clearly, a politician can attempt to appeal to a large fraction of the electorate. In contrast, a clear statement on a controversial issue can often alienate a group of voters.

Public Interest Rhetoric

Voters observing a politician funding interest groups may conclude that his or her actions are likely to be costly. Politicians will therefore attempt to justify their actions as pursuing the public interest whether or not that interpretation is valid. Separating the actual effects from stated purposes of complex programs can be so difficult that many rationally ignorant voters will not bother to try.

Logrolling

Suppose that a local spending program enriches only one interest group in a single congressional district. The representative of that district may support similar programs in other districts in exchange for additional support for the local program. Although the support of other programs will raise taxes for constituents, the support of the local interest group may still provide more votes than are lost by the tax increase. A result is that a program benefiting only a few can obtain broad legislative support.

Summary

These elements can explain the workings of a political system, with the explanation emphasizing the incentives that lead voters and politicians to choose specific actions. Are these predicted actions actually observed? While it is beyond the scope of this article to survey a vast literature, it is appropriate to note that many writers have produced empirical evidence that supports key predictions of the theory sketched above. Representative articles include Peltzman (1984), Snyder (1990), and Grier and Munger (1991). Although the model is not a complete description of the political system in its full complexity, it is sufficient to reveal important incentives for politicians to create unfunded liabilities.

Deposit insurance is perhaps the best known example of a program that creates unfunded liabilities. It lowers the funding cost of insured financial intermediaries by reducing the risk of loss to a depositor below that of alternatives lacking federal insurance such as money market funds. To the extent that premiums paid by a depository institution fail to cover expected future losses, that institution receives a subsidy. Since calculating expected future losses from such a complex program is difficult, politicians have been able to give valuable benefits to customers and owners of many financial institutions without losing votes for increasing either taxes or the reported federal debt. Other programs that generate unfunded liabilities similarly hide the full costs to current and future taxpayers.

REFERENCES

Aaron, Henry J., Barry P. Bosworth, and Gary Burtless. *Can America Afford to Grow Old? Paying for Social Security.* The Brookings Institution, Washington, D.C., 1989.

Alvarez, Robert and Arjun Makhijani. "Radioactive Waste: Hidden Legacy of the Arms Race." *Technology Review* (August/September 1988): 42-51.

Barth, James R., George Iden, Frank R. Russek, and Mark Wohar. "The Effects of Federal Budget Deficits on Interest Rates and the Composition of Domestic Output," in Rudolph G. Penner, ed., *The Great Fiscal Experiment.* The Urban Institute Press, Washington, D.C., 1991, 69-141.

Benston, George J. and George G. Kaufman. "Understanding the Savings and Loan Debacle." *The Public Interest* No. 99 (Spring 1990): 79-95.

Bovard, James. "Jubilee Time at Farmers Home Administration." *The Wall Street Journal*, April 26, 1988.

Brady, Nicholas F. Statement before the Senate Committee on Banking, Housing, and Urban Affairs. May 23, 1990.

Brumbaugh, R. Dan, Andrew S. Carron, and Robert E. Litan. "Cleaning Up the Depository Institutions Mess." *Brookings Papers on Economic Activity* (1989:1): 243-97.

Brumbaugh, R. Dan and Robert E. Litan. "The Banks Are Worse Off Than You Think." *Challenge* (January/February 1990): 4-12.

Dotsey, Michael and Anatoli Kuprianov. "Reforming Deposit Insurance: Lessons from the Savings and Loan Crisis." Federal Reserve Bank of Richmond *Economic Review* (March/April 1990): 3-28.

Downs, Anthony. *An Economic Theory of Democracy.* New York: Harper and Row, 1957.

The Economist, March 16, 1991. "Congress Fiddles While the Financial System Burns," 75-76.

Grier, Kevin B. and Michael C. Munger. "Committee Assignments, Constituent Preferences, and Campaign Contributions." *Economic Inquiry* (January 1991): 24-43.

Hirtle, Beverly and Arturo Estrella. "Alternatives for Correcting the Funding Gap of the Pension Benefit Guaranty Board." Typescript, Federal Reserve Bank of New York, May 1990.

Kane, Edward J. "The S&L Insurance Mess." *Contemporary Issues* Series 41, February 1991, Washington University, St. Louis.

__________. "The Spreading Deposit Insurance Mess: Déjà Vu All Over Again." Speech delivered at the Federal Reserve Bank of Richmond, May 23, 1991. Summary reprinted in "As Fund for Banks Shrinks, Woes Grow." *Richmond Times Dispatch*, May 26, 1991, Section G, p. 1, 4.

Kleege, Stephen. "Ely Finds Banks in Fair Shape." *The American Banker* (May 24, 1991): 5.

Pelzman, Sam. "Constituent Interest and Congressional Voting." *Journal of Law and Economics* (April 1984): 181-210.

Pension Benefit Guaranty Corporation. *1989 Annual Report.*

Rehm, Barbara A. "Agency Faulted in Handling of the Industry's Wounded." *The American Banker*, March 19, 1991.

__________. "Keefe Bruyette Says Bank Woes Are Exaggerated." *The American Banker*, April 11, 1991.

Rose, Frederick and David Wessel. "Junk-Bond Woes Put Retirement Benefits in Danger for Many." *The Wall Street Journal*, February 12, 1990.

Snyder, Jr., James M. "Campaign Contributions as Investments: The U.S. House of Representatives, 1980-86." *Journal of Political Economy* (December 1990): 1195-1227.

Todd, Walker F. and James B. Thompson. "An Insider's View of the Political Economy of the Too Big to Let Fail Doctrine." Manuscript, Federal Reserve Bank of Cleveland, June 1990, presented to the Western Economic Association International Conference, San Diego, July 1990.

Towe, Christopher M. "Government Contingent Liabilities and the Measurement of Fiscal Impact." Working Paper 90/57, International Monetary Fund, Washington, D.C., June 1990.

United States Congressional Budget Office. *Controlling the Risks of Government-Sponsored Enterprises.* Washington, D.C., April 1991.

United States Department of Energy. *Environmental Restoration and Waste Management: Five-Year Plan.* Document DOE/S-0070, Washington D.C., 1989.

United States General Accounting Office. "Nuclear Health and Safety: Dealing with Problems in the Nuclear Defense Complex Expected to Cost Over $100 Billion." Document GAO/RCED-88-197BR, Washington, D.C., July 1988.

__________. "Managing the Cost of Government: Proposals for Reforming Federal Budgeting Practices." Document GAO/AFMD-90-1, Washington, D.C., October 1989, p. 18.

United States Office of Management and Budget. *Special Analyses, Budget of the United States Government.* Fiscal Year 1989 and Fiscal Year 1990. Part 2, Chapter F.

__________. *Budget of the United States Government.* Fiscal Year 1991 and Fiscal Year 1992.