Vector Autoregressions: Forecasting and Reality

JOHN C. ROBERTSON AND ELLIS W. TALLMAN
The authors are senior economists in the macropolicy section of the Atlanta Fed's research department. *They thank Amanda Nichols for research assistance and Robert Parker and Bruce Grimm of the Bureau of Economic Analysis and Jeff Amato of the Federal Reserve Bank of Kansas City for data revisions.*

CONSTRUCTING FORECASTS OF THE FUTURE PATH FOR ECONOMIC SERIES SUCH AS REAL (INFLATION-ADJUSTED) GROSS DOMESTIC PRODUCT (GDP) GROWTH, INFLATION, AND UNEMPLOYMENT FORMS A LARGE PART OF APPLIED ECONOMIC ANALYSIS FOR BUSINESS AND GOVERNMENT. THERE ARE A VARIETY OF METHODS AVAILABLE FOR GENERATING ECONOMIC FORECASTS. ONE COMMON TYPE OF FORECAST IS A SO-CALLED JUDGMENT-BASED FORECAST. THIS TYPE OF FORECAST IS PREDOMINANTLY THE RESULT OF A PARTICULAR FORECASTER'S SKILL AT READING THE ECONOMIC TEA LEAVES, INTERPRETING ANECDOTAL EVIDENCE, AND HIS OR HER EXPERIENCE AT OBSERVING EMPIRICAL REGULARITIES IN THE ECONOMY.

One difficulty with judgmental forecasts, however, is that it is hard, if not impossible, for an outside observer to trace the source of systematic forecast errors because there is no formal model of how the data were used. Moreover, the accuracy of judgmental forecasts can be evaluated only after a track record is established. In addition, it would not be surprising, given the element of subjectivity in such forecasts, to find that changes in the personnel of the forecasting staff substantially affected the accuracy of judgmental forecasts.

Model-based forecasts provide an alternative. They are easier to replicate and validate by independent researchers than forecasts based on expert opinion alone, and the forecaster can formally investigate the source of systematic errors in the forecasts. An important aspect of a forecast from a model that quantifies future uncertainty is that it allows the forecaster to give not only an estimate of the most likely future outcome but also a probabilistic assessment of a range of alternative outcomes. In this context, to say that GDP growth next year is predicted to be 2 percent conveys somewhat less information about the future than does saying that GDP growth next year is most likely to be 2 percent and the probability of negative growth is less than 10 percent. Another advantage to employing model-based forecasts is that the accuracy of the point forecasts from the model can be statistically evaluated prior to using the forecasts. In other words, a forecast model's performance can be established before it is used by a decision maker. The distinction between judgmental and model-based forecasts cannot be pushed too far, however, because successful model specifications also depend heavily on the skill and ingenuity of particular individuals. No model can be left on automatic pilot for long.[1]

This article describes a particular model-based forecasting approach. The model studied is a vector autoregression (VAR) of six U.S. macroeconomic variables. Although this model is small and highly aggregated, it provides a convenient framework for illustrating several practical forecasting issues. In emphasizing the practical problems of forecasting economic data using a statistical model, the research draws on experience in using such a model at the Federal Reserve Bank of Atlanta. The focus on a simple model is intended to provide potential users with a road map of how one might implement a VAR-based forecasting model more generally.

The article first describes some practical problems in using a VAR model for forecasting purposes. The discussion focuses on a VAR model fitted to monthly data having staggered release dates that uses a distributed monthly estimate of quarterly GDP data. The article then evaluates the relative forecast performance of various alternative specifications of the VAR and offers conclusions based on the study's findings.

Developing a VAR Model for Forecasting

The starting point for any forecasting project should be the question, What is the objective of the forecasting exercise? This question inevitably raises additional questions, such as who will be using the model and for what purpose the forecast will be used. A forecaster will design a model to fit the demands of his or her client. In essence, the end user of the forecast typically determines the variables to be incorporated into the model.

Who Are the Clients and What Data Should Be Forecast? The main client for models designed at the Federal Reserve Bank of Atlanta is the bank president, and his needs determine the model's design. The president serves on the Federal Open Market Committee (FOMC), the voting body of the Fed that determines monetary policy. To support his responsibilities in contributing to policy decisions, a helpful model will be designed to forecast the main economic aggregates of concern to the FOMC, such as measures of inflation, of the employment situation, and of real output.[2] As in Zha (1998), the VAR model described in this article includes real GDP as a measure of real output, the consumer price index (CPI) for urban households as a measure of inflation, and the civilian unemployment rate (UR) as a measure of unemployment. In addition, since it is constructed to help guide monetary policy, two monetary variables are included—the effective federal funds rate (FF), which is a series that the FOMC influences directly, and the M2 money stock, a series that the FOMC influences somewhat less directly. Finally, to allow a role for commodity prices in predicting future inflation, a commodity price series is also included in the variables list.[3]

Handling Mixed-Frequency Data. The VAR model is specified for data at a monthly frequency. But one of the variables in the model, real GDP, is measured only quarterly. Incorporating data of different frequencies into a single-frequency model is a vexing problem. One simple approach is to take all the higher-frequency data and convert them to the frequency of the lowest-frequency data, in this case real GDP. The estimated model would then be a quarterly model in which the other variables would typically be averages of the daily, weekly, or monthly observations. This method is a standard approach in forecasting models involving GDP data.

The focus on a simple model is intended to provide potential users with a road map of how one might implement a VAR-based forecasting model more generally.

There is some evidence, however, that incorporating timely, monthly data can help forecast quarterly data (Miller and Chin 1996; Ingenito and Trehan 1996; Tallman and Peterson 1998).[4] The ability to incorporate high-frequency data into forecasts is perhaps one of the main justifications for using judgmental forecasts. However, measuring the marginal contribution of such procedures in judgmentally adjusted forecasts is difficult because the impacts of using high-frequency data on the forecasts cannot be clearly traced.

The technologies that are available for exploiting monthly information for model-based forecasts of real economic aggregates (like real GDP) take a variety of forms. The approach taken here is to use the distribution technique of Chow and Lin (1971) in constructing a monthly real GDP series. The procedure uses monthly

1. Reifschneider, Stockton, and Wilcox (1997) provide a thorough discussion of the interaction between judgmental and model-based forecasting as practiced at the Board of Governors of the Federal Reserve System.

2. For a monetary policymaker, the Humphrey-Hawkins legislation suggests that the objectives of the Fed (as legislated by Congress) should be consistent with the federal government's goals of full employment and low inflation.

3. Exact definitions of the series used are contained in Appendix 1.

4. Ingenito and Trehan estimate a monthly model of the U.S. economy. Miller and Chin combine forecasts of, say, one month of a monthly series with two actual monthly observations within a quarter, to form preliminary estimates of the quarterly observations. In essence, there are alternative ways to extract information from monthly data and incorporate it into quarterly forecasts. Tallman and Peterson show that incorporating timely monthly information improves the forecast accuracy of a quarterly model for Australian GDP.

data on variables related to GDP (specifically, industrial production, nonagricultural payroll employment, and personal consumption expenditures) to estimate the coefficients of a regression equation for GDP at a quarterly frequency.[5] The regression is then used to construct estimates of monthly real GDP in a way that ensures that the quarterly average of the resulting monthly GDP estimates equals the corresponding quarterly observation of GDP.

Three new values for the monthly GDP index can be constructed with the release of GDP data for a new quarter. GDP and the data on the monthly indicator series are revised on a reasonably regular schedule (see Appendix 1). Hence, the whole index could be reestimated every month as the existing quarterly GDP data and data on the monthly series are revised and when a new GDP observation becomes available. A description of how to implement this procedure is presented in Appendix 2.

An advantage to employing model-based forecasts is that the accuracy of the point forecasts from the model can be statistically evaluated prior to using the forecasts.

Specification and Estimation of the VAR. The idea underlying forecasting with a vector autoregression model is first to summarize the dynamic correlation patterns among observed data series and then use this summary to predict likely future values for each series. Mathematically, a VAR expresses the current value of each of m series as a weighted average of the recent past of all the series plus a term that contains all the other influences on the current values. A VAR can be written compactly as

$$y_t = \nu + B_1 y_{t-1} + \cdots + B_p y_{t-p} + u_t,$$

where y_t denotes the $m \times 1$ vector of variables included in the VAR for month t and where all but the interest rate and unemployment are expressed in natural logarithms.[6] Notice that the $m \times 1$ error vector u_t measures the extent to which y_t cannot be determined exactly as a linear combination of the past values of y with weights given by the constant coefficients ν and B_l, $l = 1, \ldots, p$. Uncertainty about the value of u_t arises because the numbers of lagged observations of y to be included in the VAR, p, along with the values of the coefficients are unknown and hence will have to be estimated from the available data. The uncertainty about u_t is made operational by assuming that u_t is a random vector having a zero mean, the error covariance matrix Σ is positive-definite, and u_t is uncorrelated with lagged values of y_t.

It is not uncommon to find that VAR models freely fitted to data of the type used here have many estimated coefficients whose standard errors are large. Perhaps they are large because the coefficients are actually zero as indicated. Alternatively, the data might not be rich enough to provide sufficiently precise estimates of nonzero coefficients. If the parameters are too imprecise, then the situation is serious because it has been observed that large estimation uncertainty can lead to poor forecasts.[7] Getting imprecise parameter estimates in a VAR is likely to be a common practical problem because the number of parameters is often quite large relative to the available number of observations. For example, in the next section of the article the VAR models are specified with p as large as 13. With six variables in the VAR, a total of seventy-nine coefficients would therefore be estimated in each equation in the VAR. Various solutions to the problem of "overfitting" VAR models have been proposed in the forecasting literature, and these all amount to putting prior constraints on the values of the model's coefficients so as to require less information from the data when determining the coefficient values. These prior restrictions act as nondata information regarding the coefficient values.

One approach to reducing the coefficient uncertainty is to set some coefficients to zero or other preassigned values. These values may or may not have been determined on the basis of prior fitting of models to the data. For example, one might prespecify a maximum lag order p_{max} for the VAR ($p_{max} = 13$ in the empirical analysis) and select the $p \le p_{max}$ that minimizes a specific criterion. This criterion discounts the increase in measured in-sample fit that occurs simply because one is fitting more coefficients to a fixed set of observations. Another way that coefficient restrictions are used in a VAR forecasting model is by predifferencing data series that appear to exhibit trends or quite persistent local levels over time, prior to fitting the VAR. This approach would be mathematically the same as imposing exact restrictions on the coefficients of a VAR in the levels of the data.

An alternative approach is to impose inexact prior restrictions. For example, rather than setting all lags greater than p to zero, the VAR could be estimated in a way that gives more weight to nonzero coefficients on recent observations relative to those on more distant lags but without necessarily setting the coefficients on the more distant lags to zero. Similarly, the degree of uncertainty about the constraints implied by predifferencing the data could be incorporated explicitly into the coefficient estimation strategy.

The methods presented in Litterman (1980, 1986), Doan, Litterman, and Sims (1984), Sims (1992), Kadiyala and Karlsson (1993, 1997), and Sims and Zha (1998) are all essentially ways of imposing inexact prior restrictions on the coefficients of a VAR. How closely the prior restrictions are imposed is usually determined by examining historical out-of-sample forecast performance across various degrees of tightness on the prior restrictions. The specification can be made to resemble an exact restriction if the resulting improvement in forecast performance warrants such a specification, but it has been found that the best performance usually arises by not imposing the restrictions exactly. The implementation of various types of inexact prior restriction is described in some detail in Box 1.

Handling the Staggered Timing of Data Releases. Another data-related problem is that new data are released at irregular intervals. For example, an average interest rate for a month is available at the end of the month being measured. On the other hand, a money stock estimate for a month is not available until the middle of the following month. Moreover, although the distribution of the real GDP series puts the model data on a monthly frequency, the new GDP observations can be estimated only toward the end of the month after the quarter being measured. In order to exploit monthly data that are available on some but not all series in the VAR, the so-called conditional forecasting technique, as described in Doan, Litterman, and Sims (1984) and Litterman (1984), is used. In this framework, at the end of a particular month, say, the value for all data series that are not yet available for that month are forecast "conditional" upon all the variables for which observations are available for the current month. The procedure involves first estimating the VAR model using a sample that contains complete observations on all the variables in the model. At the end of January, then, the VAR would be estimated with data up to the previous December. Then a forecast of all the variables from the VAR for January is made as if no additional data were available. However, the forecasts for the federal funds rate and commodity prices must be exact because their January values are at hand, and this information should allow deducing more accurate forecasts of the other series whose January values are not known. The size of the improvement to the forecast for January will depend on the extent to which knowing the values of January's federal funds rate and commodity prices is useful for predicting the other series' values for that month. This idea can be readily extended to more complicated situations, as occurs at the end of March when the values for the federal funds rate and commodity prices for January, February, and March and values of M2, CPI, and unemployment for January and February are available but there is no value for first-quarter GDP.[8] A simple example that illustrates the implementation of the foregoing conditional forecasting procedure is presented in Box 2.

Measuring Forecast Accuracy. Evaluating the accuracy of forecasts is a form of accountability in the sense that a client would presumably give more weight to a forecasting scheme that can generate relatively accurate forecasts. In addition, forecast evaluation is relevant to the forecaster when deciding on a model specification for subsequent use. The preferences or loss function of the forecast user is key to the selection of the accuracy criterion. In most forecast evaluations the accuracy measures are some form of average error, typically root mean squared error (RMSE) or mean absolute error (MAE), but many other possibilities are available. For example, the proportion of times the direction of a change in a variable is correctly forecast may be relevant for evaluating forecasts if, for example, capturing turning points were viewed as of primary importance. The results reported below use the RMSE as the accuracy criterion, but it is acknowledged that using other forecast accuracy criteria may yield different model rankings.

The idea underlying forecasting with a vector autoregression model is first to summarize the dynamic correlation patterns among observed data series and then use this summary to predict likely future values for each series.

In justifying the final specification to a client, most forecasters would present some evidence regarding the accuracy and reliability of the model over some historical period. This evidence is likely to be a by-product of the model selection process given that a forecaster will probably have spent considerable time conducting historical out-of-sample forecasting experiments in order to tune the model specification. In the empirical application presented below, the period from 1986 to 1997 is used to examine the forecast performance of the various VAR specifications.

5. *These monthly data are the same as those used in constructing the Conference Board's coincident index of economic activity except that this model uses consumption expenditure instead of disposable income and does not use the monthly retail and trade sales series because it has a two-month release lag.*
6. *The selection of the variables included in the model implies a strong exclusion restriction on all other variables that could have been but were not included.*
7. *On this point see Wallis (1989) and Fair and Shiller (1990).*
8. *An algorithm for implementing this procedure is available in the RATS software package. The authors have a more general version programmed using the GAUSS language that allows a wider range of conditioning experiments.*

Inexact Prior Restrictions in VAR Forecasting Models

As described in the text, a VAR model for the $m \times 1$ vector of observations y_t has the form

$$y_t = \nu + B_1 y_{t-1} + \cdots + B_p y_{t-p} + u_t, \qquad \text{(B1)}$$

where the coefficients ν and B_l, $l = 1, \ldots, p$, and the covariance of u_t, Σ, are to be estimated once a value for p is specified. A major problem with using a VAR such as equation (B1) for forecasting based on unrestricted OLS estimates of the coefficients when p is moderately large is that the coefficient values are often not very well determined in a finite set of data. Litterman (1980, 1986) discusses this problem in the context of economic series that exhibit trends or persistent local levels and suggests an alternative, Bayesian, method for estimating the coefficients in these cases. The idea is to treat the coefficients as random quantities around given mean values, with the tightness of the distributions about these prior means determined via a set of hyperparameters. The OLS coefficient estimator is then modified to incorporate the inexact prior information contained in these distributions. The main technical issues involve specifying the form of the prior distributions and determining the form of the estimators.

Litterman's method is often referred to as the Minnesota prior because of its origins at the University of Minnesota and the Federal Reserve Bank of Minneapolis. It is usually implemented as follows.

The prior for the individual elements of each lag coefficient matrix B_l is that they are each independent, normally distributed random variables with the mean of the coefficient matrix on the first lag, B_1, equal to an identity matrix and the mean of the elements of B_l, $l > 1$, equal to zero. Notice that if these restrictions were exact then each variable would be a random walk, possibly with nonzero drift. While the random walk prior might be considered a reasonable specification, there is no need to impose it exactly on the VAR. In particular, the standard deviation of the ij-th element of the l-th lag coefficient matrix B_l can be nonzero, with these being often specified as $\lambda_1 / l^{\lambda_3}$, if $i = j$, and $\sigma_i \lambda_1 \lambda_2 / \sigma_j l^{\lambda_3}$, if $i \neq j$ (see, for example, Sims 1992).

The parameter λ_1 is the prior standard deviation of the ii-th element of B_1, reflecting how closely the random walk approximation is to be imposed. Lowering λ_1 toward zero has the effect of shrinking the diagonal elements of B_1 toward one and all other coefficients to zero.

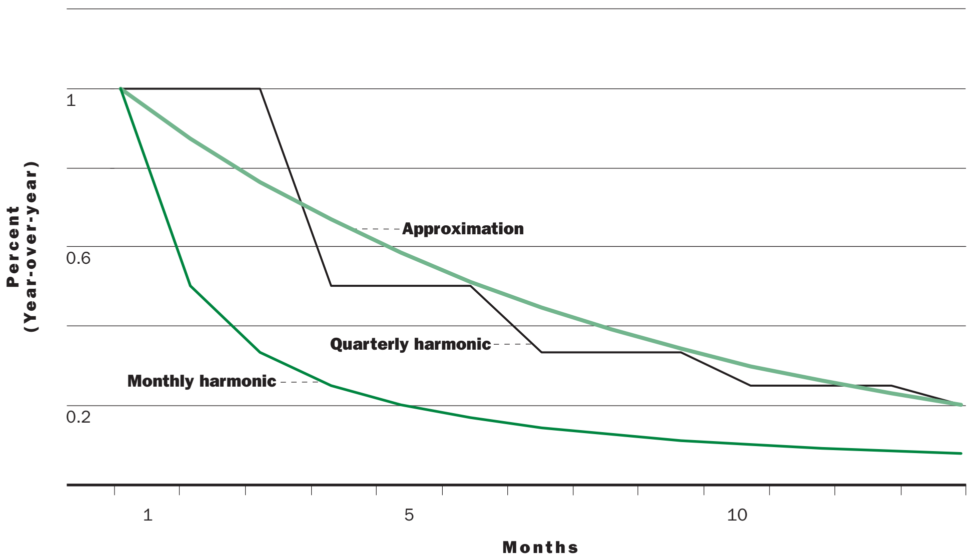
It is reasonable to suppose that most of the variation in each of the dependent variables in the VAR is accounted for by its own lags. Thus, in each row of B_l, coefficients on lags of other variables can be assigned smaller variance in relative terms by choosing a value for λ_2 such that $0 < \lambda_2 \leq 1$. Decreasing λ_2 toward zero has the effect of shrinking the off-diagonal elements of B_l toward zero. Setting λ_2 to unity means that no distinction is made between the lags of the dependent variable and the lags of other variables.

The ratio (σ_i / σ_j) is included in the prior standard deviations to account for the differences in the units of measurement of different variables. If the variability of $y_{i,t}$ is much lower than that of $y_{j,t}$, then the coefficient on $y_{j,t-1}$ in the i-th equation is shrunk toward zero. In practice, the σ_i are usually set equal to the residual standard error from an OLS regression of each dependent variable on p lagged values.

The parameter $\lambda_3 > 0$ is used to determine the extent to which coefficients on lags beyond the first one are likely to be different from zero. As λ_3 increases, the coefficients on high-order lags are being shrunk toward zero more tightly. If λ_3 is set to one, the rate of decay in the weight is harmonic. For a VAR fitted to monthly data one could choose to use a decay rate that approximates a harmonic decay pattern at a quarterly frequency. In particular, in a monthly VAR with $p = 13$ lags, rather than using l^{-1}, the decay could be specified as $\exp(cl - c)$ where $c = -0.13412$. This approximation is depicted in Chart A, which shows the harmonic decay pattern at a quarterly frequency (1, ½, ⅓, ¼, ⅕). The approximately equivalent monthly decay is based on $p = 13$ lags. Notice that this specification ensures that for $l = 13$ the decay rate is exactly ⅕. By contrast, assuming a harmonic decay at the monthly frequency results in a much faster rate of decay, as the chart shows. In fact, under the monthly harmonic decay scheme the weight given to the coefficients on observations from five months ago would be the same as that given to the coefficients attached to observations from thirteen months ago under the quarterly harmonic decay approximation. The choice of the decay pattern will undoubtedly have some influence on the model's forecast performance.

As for the constant term ν, there are numerous ways that an inexact prior restriction could be implemented. Here the prior mean of the constant in the i-th equation is taken to be zero with standard deviation $\sigma_i \lambda_4$. As λ_4 decreases, the constant is shrunk toward zero.

CHART A Lag Decay Rates



To summarize, the Minnesota prior for the coefficients of the i-th equation of the VAR is that the coefficient vector b_i is normally distributed with mean $\bar{b}_i$ and covariance $\bar{G}_i$. For example, if $m = 2$ variables and there are $p = 2$ lags of each variable in the VAR, then $\bar{b}_1 = (0\ 1\ 0\ 0\ 0)'$ and $\bar{b}_2 = (0\ 0\ 1\ 0\ 0)'$. The diagonal prior covariance matrix $\bar{G}_1$ has nonzero elements $(\sigma_1\lambda_4)^2$, λ_1^2, $(\sigma_1\lambda_1\lambda_2/\sigma_2)^2$, $(\lambda_1/2^{\lambda_3})^2$, and $[\sigma_1\lambda_1\lambda_2/(2^{\lambda_3}\sigma_2)]^2$ while $\bar{G}_2$ has nonzero elements $(\sigma_2\lambda_4)^2$, $(\sigma_2\lambda_1\lambda_2/\sigma_1)^2$, λ_1^2, $[\sigma_2\lambda_1\lambda_2/(2^{\lambda_3}\sigma_1)]^2$, and $(\lambda_1/2^{\lambda_3})^2$.

In terms of coefficient estimation, one should note that the usual OLS estimator of the coefficients of the i-th equation of the VAR model in equation (B1) has the form

$$\hat{b}_i^{OLS} = (X'X)^{-1}X'y_i,\ i = 1,\ldots,m,$$

where y_i is a $T \times 1$ vector with T-th element $(y_{i,T})$ and X is a $T \times (mp + 1)$ matrix with T-th row $(1\ y_{1,T-1}\cdots y_{m,T-1}\ y_{1,T-2}\cdots y_{m,T-2}\cdots y_{m,T-p})$. In contrast, the coefficient estimator, or more formally, the mean of the posterior distribution under the Minnesota prior, is

$$\hat{b}_i^{LT} = (\bar{G}_i^{-1} + \sigma_i^{-2}X'X)^{-1}(\bar{G}_i^{-1}\bar{b}_i + \sigma_i^{-2}X'y_i)$$

(see, for example, Lutkepohl 1991).

Under a strict interpretation of the Minnesota prior, the estimator of the error covariance is a diagonal matrix with σ_i along the diagonal and with σ_i determined from the data. The coefficient estimation problem is therefore simplified because it avoids having to specify how the prior distribution Σ is related to the prior distribution for B_l. In practice, the diagonality restriction on Σ is often ignored, and a nondiagonal estimator based on the residual sum-of-squares matrix is used instead.

Over recent years Bayesian VAR techniques have been developed that remove the assumption that the error covariance matrix is fixed and diagonal (see Kadiyala and Karlsson 1993, 1997, and Sims and Zha 1998 for a discussion). For example, one could replace the Minnesota prior with a specific form of a so-called Normal-Wishart prior. Under a Normal-Wishart prior, the prior distribution of coefficients (given Σ) is Normal while the prior distribution of Σ is inverse Wishart (see Drèze and Richard 1983, 539–41). This feature allows the random-walk aspect of the Minnesota prior on the coefficients to be used without having to take independence across the equations of the VAR as an exact restriction.

Under the Normal-Wishart prior, the coefficient estimator (the mean of the posterior distribution) has the form

$$\hat{B} = (\bar{H}^{-1} + X'X)^{-1}(\bar{H}^{-1}\bar{B} + X'y),$$

where $\bar{B}$ is the prior mean of the coefficient matrix $B = (b_1 \ldots b_m)$, and $\bar{H}$ is a diagonal, positive-definite matrix, with elements defined as in Sims and Zha (1998). The corresponding estimator of the error covariance is

$$\hat{\Sigma} = T^{-1}\left(y'y - \hat{B}'(X'X + \bar{H}^{-1})\hat{B} + \bar{B}'\bar{H}^{-1}\bar{B} + \bar{S}\right),$$

where $\bar{S}$ is the diagonal scale matrix in the prior inverse-Wishart distribution for Σ. As a specific example, if $m = 2$ and $p = 2$, then $\text{vec}(\bar{B}) = (0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0)'$; $\bar{S}$ has $(\sigma_i/\lambda_0)^2$ along the diagonal; and $\bar{H}$ has diagonal elements

$(\lambda_0\lambda_4)^2$, $(\lambda_0\lambda_1/\sigma_1)^2$, $(\lambda_0\lambda_1/\sigma_2)^2$, $[\lambda_0\lambda_1/(2^{\lambda_3}\sigma_1)]^2$, and $[\lambda_0\lambda_1/(2^{\lambda_3}\sigma_2)]^2$. The parameter, λ_0, can be thought of as controlling the overall tightness of the prior on Σ.[1]

To see how this setup is related to the Minnesota prior, note that under a Normal-Wishart prior the covariance of coefficients has a form whereby $\overline{H}$ is multiplied by each element of $\overline{S}$ (a Kroneker product operation). Doing this multiplication yields an $m(mp + 1) \times m(mp + 1)$ scale matrix whose elements are exactly the coefficient prior variances under the Minnesota prior but with $\lambda_2 = 1$. This latter restriction is required because the Normal-Wishart prior implies a certain symmetry across the equations of the VAR (apart from scale). In particular, it prohibits the prior from treating lags of the dependent variable differently from lags of other variables in each equation. In a sense the restriction on λ_2 is the price of being able to relax the strong error covariance assumption of the Minnesota prior while still being able to have an estimator that is simple to implement.[2]

Other types of inexact prior information have been introduced as modifications of the Minnesota prior that involve priors on linear combinations of the coefficients in equation (B1). Because this modification introduces nonzero off-diagonal terms into the prior covariance for the individual coefficients, it is usually implemented by mixing a set of dummy observations into the data set rather than directly specifying the prior covariance structure. The magnitude of the weight attached to the dummy observations is used to determine the tightness of the prior restriction.

Two types of initial dummy observations are discussed here. The first is motivated by the frequent practice of specifying a VAR model of data that contain stochastic trends (unit roots) in first differences of the data. This specification corresponds to the restrictions that the sums of coefficients on the lags of the dependent variable in each equation of the VAR equal one while coefficients on lags of other variables sum to zero. Formally, when $\sum_{i=1}^{p} B_i = I$, the VAR can be written as

$$\Delta y_t = \nu + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t,$$

where $\Gamma_l = \sum_{n=l+1}^{p} B_n$ and forecasts of period-to-period changes in the variables will not be influenced at all by the current level of the series. To accommodate this possibility, Doan, Litterman, and Sims (1984) developed the so-called sum of coefficients prior for a VAR specified in the levels of the data. This prior is implemented by adding a set of m initial dummy observations to the data set. For example, if there are $m = 2$ variables then there are two sum-of-coefficients dummy observations for each equation. For the dependent variable in the first equation, these are $(\overline{y}_1^0\ 0)'$ while they are $(0\ \overline{y}_2^0)'$ for the dependent variable in the second equation, and $\overline{y}_i^0$ is the mean of the p presample values for variable y_i. If there are $p = 2$ lags of each of the two variables in the VAR, then the matrix of dummy observations for the regressors in both equations has the form

$$\begin{pmatrix} 0 & \overline{y}_1^0 & 0 & \overline{y}_1^0 & 0 \\ 0 & 0 & \overline{y}_2^0 & 0 & \overline{y}_2^0 \end{pmatrix}.$$

A weight of $\mu_5 \geq 0$ is then attached to these dummy observations, and as $\mu_5 \to \infty$ the estimated VAR will increasingly satisfy the sum of coefficients restriction.[3] Notice also that as $\mu_5 \to \infty$ the forecast growth rates will eventually converge to their sample average values.

If it is supposed that each series contains a stochastic trend, then as $\mu_5 \to \infty$ the sum of coefficients restriction implies that there are as many stochastic trends in the VAR as there are variables in y_t. However, it might be reasonable to suppose that there are fewer than m stochastic trends in the VAR, as would be the case if there were stable long-run relations between the trending series (cointegration). Sims (1992) observed that introducing an additional dummy observation could make some allowance for this possibility. If $m = 2$ and $p = 2$, the Sims initial dummy observation for the dependent variable in the i-th equation is $\left(\overline{y}_i^0\right)$ while $\left(1\ \ \overline{y}_1^0\ \ \overline{y}_2^0\ \ \overline{y}_1^0\ \ \overline{y}_2^0\right)$ is the vector of dummy regressor observations. A weight of $\mu_6 \geq 0$ is then attached to these dummy observations for each equation. If the series individually contain stochastic trends, then as $\mu_6 \to \infty$ increasingly more weight will be put on a VAR that can be written in a form in which all series share a single stochastic trend and the intercept will be close to zero.[4]

1. *For given σ_j, it is only the products $\lambda_0\lambda_1$ and $\lambda_0\lambda_4$ that have a direct influence on the coefficient estimator.*
2. *GAUSS and MATLAB code for implementing this prior is available from the authors upon request.*
3. *It is not necessary that μ_5 be the same in each equation. See, for example, Miller and Roberds (1991).*
4. *A remaining practical question is how to select values for the hyperparameters since the quality of the forecasts will depend on these choices. In practice, values are determined based on examining historical forecast performance across a range of parameter settings.*

Summary of the VAR Forecasting Procedure. Assuming that the forecasts are constructed within a few days of the end of a month, there is a simple pattern for forecast construction across a year. In particular, forecasts formed at the end of January, April, July, and October all have the same structure. Those formed at the end of February, May, August, and November all have the same structure, and those formed at the end of March, June, September, and December have the same structure. As a specific example of the previous discussion, it may be helpful to consider how forecasts would have been formed at the end of January, February, and March 1998.

January. At the end of January 1998 the commodity price index and funds rate for January would be available. The most recent quarterly GDP observation would be the "advanced estimate" for the fourth quarter of 1997. The latest available CPI, unemployment rate, M2 stock, and monthly GDP data (constructed using the Chow-Lin procedure) are all for December 1997. A VAR is then fitted to the latest vintage of monthly data through December 1997. Conditional forecasts of the unemployment rate, CPI, M2, and GDP are made for January, and with these forecasts in hand forecasts of the quarterly average unemployment rate, inflation rate, and rate of GDP growth are constructed, along with corresponding forecasts of the annual averages and growth rates.

February. At the end of February 1998 the commodity price index and funds rate for January and February, the CPI, unemployment rate, and M2 stock for January, and a GDP estimate for the preceding fourth quarter would be in hand. The latest quarterly GDP estimate for the fourth quarter of 1997 is the "preliminary estimate." The Chow-Lin procedure is reapplied to generate a new monthly GDP series, and the VAR is refit using the vintage of data through December 1997, available at the end of February. Conditional forecasts of the unemployment rate, CPI, and M2 are made for February, and a conditional forecast of GDP is made for January and February. Forecasts of the quarterly and annual averages/growth rates are then constructed.

March. At the end of March 1998 the commodity price index and funds rate for January, February, and March, the CPI, unemployment rate, and M2 stock for January and February, and a GDP estimate for the preceding fourth quarter would be available. The latest quarterly GDP estimate for the fourth quarter of 1997 is the "final estimate." The Chow-Lin procedure is reapplied to generate a new monthly GDP series based on the revised data, and the VAR is fitted using the vintage of data through to December 1997 available at the end of March. Conditional forecasts of the unemployment rate, CPI, and M2 are made for March, and a conditional forecast of GDP is made for January, February and March. Forecasts of the quarterly and annual averages/growth rates are then constructed.

The process then repeats itself. For example, at the end of April 1998 the commodity price index and funds rate for April, the CPI, unemployment rate, and M2 stock for March, and a GDP estimate for the first quarter of the current year are available, and so on.

Evaluating the accuracy of forecasts is a form of accountability in the sense that a client would presumably give more weight to a forecasting scheme that can generate relatively accurate forecasts.

Empirical Application

This section reports the results of using various VAR specifications to forecast the unemployment rate, the inflation rate, and the rate of growth in GDP for the current and the next quarter and the current and next calendar year over the period from 1986 to 1997. This comparison is not intended as a formal model evaluation but simply as a demonstration that the nature of prior restrictions on a VAR specification can have important implications for forecast performance. The alternative specifications considered are

- An unrestricted VAR specification in the levels of the data, estimated by OLS, and with $p = 13$ imposed. This specification is denoted as the OLS model.
- A VAR specification estimated by OLS, with the sum-of-coefficients restriction imposed exactly (the data are first-differenced) and with the lag length chosen on the basis of the Akaike Information Criterion (AIC) with p_{max} set at 13 (see Lutkepohl 1991 for a discussion of the AIC). This specification is denoted as the DOLS-AIC model.[9]
- A Litterman VAR as described in Box 1, with the settings for the parameters of the prior standard deviations those suggested by Litterman (1986),

9. *The study also examined the forecasts from a VAR in levels, where the lag length was chosen based on the AIC (denoted as OLS-AIC). The OLS-AIC model's accuracy is better than the unrestricted OLS model but not as good as the DOLS-AIC model. Similarly, an unrestricted VAR in first differences (denoted as DOLS) performs better than the OLS and OLS-AIC specifications but not as well as the DOLS-AIC model. The fact that specifications in which the lag length is selected on the basis of a penalty function outperform those that do not suggests that down-weighting distant lags is advantageous to forecast performance.*

Conditional Forecasting with A VAR

A frequent problem in implementing a VAR model for forecasting is that observations on all the variables in the model for the current month are not available. Rather than using a complete data set, which may involve using relatively old data or postponing the forecast until all new data are available, one can make a forecast of the missing observations using the partially complete data set. If there is significant high-frequency correlation among the observed and yet-to-be-observed data, then this approach should generate better forecasts of the missing observations than if that information were ignored.

As a hypothetical example of the strategy, suppose that there is a VAR with $m = 2$ variables and $p = 1$ lag such that each series is a random walk and the errors have unit variance and a contemporaneous (current-period) correlation of 0.25. Thus, the VAR can be written as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix},$$

where, as at time t, the e_t are expected to be zero with covariance equal to an identity matrix. Now suppose that for some reason, possibly because of lags in the publication of data, at the end of February 1998 there are observations on y_1 for December 1997 and for January and February 1998 while the only observation on y_2 is for December 1997. If T denotes December 1997, then the January and February values are determined according to

$$\begin{pmatrix} y_{1T+1} \\ y_{2T+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1T} \\ y_{2T} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} e_{1T+1} \\ e_{2T+1} \end{pmatrix},$$

and

$$\begin{pmatrix} y_{1T+2} \\ y_{2T+2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1T} \\ y_{2T} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} e_{1T+1} \\ e_{2T+1} \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{pmatrix} e_{1T+2} \\ e_{2T+2} \end{pmatrix}.$$

With only data for December 1997 in hand, the best guess about the future errors is zero, and hence the forecast of January's and February's data values for variable y_i will simply be the corresponding December data values y_{iT}. However, this forecast ignores the fact that the January and February values of y_1: y_{1T+1} and y_{1T+2} are actually already known. The problem is how to use this extra information to refine the forecasts of y_{2T+1} and y_{2T+2}.

Suppose the difference between the December forecasts of y_{1T+1} and y_{1T+2} and their actual values were stacked in a 2×1 vector,

$$r = \begin{bmatrix} y_{1T+1} - y_{1T} \\ y_{1T+2} - y_{1T} \end{bmatrix}.$$

Notice that the first element of r is exactly equal to e_{1T+1} in the example and the second element of r is equal to $e_{1T+1} + e_{1T+2}$. That is, there is only one value for e_{1T+1} that can ensure that the forecast of y_{1T+1} satisfies the constraint, and hence there is only one value for e_{1T+2} that can make the second constraint hold. Denote as e the 4×1 vector of stacked error terms,

$$e = (e_{1T+1}\ e_{2T+1}\ e_{1T+2}\ e_{2T+2})',$$

and let R be a 2×4 matrix,

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix},$$

where the first row of R gives the relationship between the first element of r and e while the second row gives the relationship between the second element of r and e. The result is a system of equations of the form $r = Re$ to solve for e, and in general there will be an infinity of possible solutions.[1]

However, Doan, Litterman, and Sims (1984) showed that a unique vector of forecast errors that both satisfy the constraints and minimize the sum of squared errors $e'e$ is given by $\hat{e} = R'(RR')^{-1}r$. In a least-squares sense this is the most likely set of values for the forecast errors. In the example this solution yields

$$\hat{e} = (\hat{e}_{1T+1}\ \ 0\ \ \hat{e}_{1T+2}\ \ 0)',$$

where $\hat{e}_{1T+1} = y_{1T+1} - y_{1T}$ and $\hat{e}_{1T+2} = y_{1T+2} - y_{1T} - \hat{e}_{1T+1}$. Finally, substituting the elements of $\hat{e}$ back into the VAR yields modified, or conditional, forecasts. In particular, instead of using y_{1T} as the forecast of y_{1T+j} for $j = 1, 2$, their actual known values would be used, and rather than using y_{2T} as the forecast of y_{2T+j}, $y_{2T} + 0.5\hat{e}_{1T+1}$ and $y_{2T} + 0.5(\hat{e}_{1T+1} + \hat{e}_{1T+2})$ would be used, respectively.

Although designed to handle incomplete data matrixes in a VAR model, this procedure can also be extended to allow for conditioning on assumed future paths for variables in the model by simply treating these as known data. For example, in September 1998 one might wish to condition on GDP growth to be an annualized 2 percent per quarter for the current and the next quarter. Doing so would introduce two constraints (one for each quarter) into the VAR to be spread across six months.[2]

Waggoner and Zha (1998) observe that, although $\hat{e}$ are the most likely outcomes for the forecast errors, one can also allow for the fact that over the constraint period (January and February in the example) the values of the nonconstrained errors will not necessarily be zero. In particular, the stacked forecast errors will be randomly distributed with mean $\hat{e}$ and a singular covariance matrix given by $I - R'(RR')^{-1}R$. This result can be used to generate error bands for the conditional forecasts by repeatedly simulating observations from this distribution. The reader is referred to Waggoner and Zha for additional details relating to conditional forecasting in this context and for methods for accounting for uncertainty about the true values of the model's coefficients in these simulations.

1. *In general,* R *is of dimension* n × (mk)*, where* n *is the number of constraints to be satisfied and* k *is the maximum number of months the constraints are imposed.*
2. *This strategy assumes that one does not want to incorporate additional information as to the underlying source of the 2 percent GDP growth. See Litterman (1984) and Waggoner and Zha (1998) for a discussion on this point.*

that is, $\lambda_1 = 0.2$, $\lambda_2 = 0.2$, and $\lambda_3 = 1$, with λ_4 set as 0.3.

- A modified Litterman VAR that uses the base Litterman settings but also incorporates sum-of-coefficients and cointegration dummy observations described in Box 1, with weights $\mu_5 = 5$ and $\mu_6 = 5$, respectively.
- The VAR specification used in Waggoner and Zha (1998) and Zha (1998) and described in Box 1. Following Waggoner and Zha the prior standard deviation parameter values are set at $\lambda_0 = 0.6$, $\lambda_1 = 0.1$, $\lambda_2 = 1$, $\lambda_3 = 1$, $\lambda_4 = 0.1$, $\mu_5 = 5$ and $\mu_6 = 5$. This model is denoted as the ZVAR model.
- A partial ZVAR specification that uses the base parameter settings of the ZVAR model but shuts off the dummy observations by setting $\mu_5 = 0$ and $\mu_6 = 0$.

Each non-OLS specification uses the monthly approximation to a quarterly harmonic lag decay pattern as discussed in Box 1.

Robertson and Tallman (1998) describe how a true real-time forecast experiment would involve using exactly the data series available at the time the forecasts were made, together with a model specification and coefficients determined using these data. The results reported below are obtained by using the real-time vintage of historical data to construct the forecasts and a recent vintage of data (as of July 1998) to evaluate the forecasts. However, because the parameter setting for the ZVAR specification was chosen on the basis of out-of-sample forecast performance, the study does not accurately replicate real-time forecasts. In particular, a forecaster in 1986 would not have been able to use the post-1986 forecast performance to guide the model specification.

For each VAR specification a maximum of $p = 13$ lagged observations of the six variables in y_t are included in each equation. The VAR is first fit to data for the period from February 1960 to December 1985, with the thirteen presample values being those for January 1959 to January 1960. The VAR is reestimated (and lag length reselected in the case of the DOLS-AIC model) every three months through December 1997, and in doing so the coefficient estimates can vary in response to new data and revisions to existing data. As described in the previous section, each time the model is reestimated, (conditional) forecasts of the unemployment rate, inflation, and GDP growth are generated for the current and the two subsequent quarters, as well as forecasts for the current and each of the two subsequent calendar years. Pooling the forecast errors for each period (quarter or year) yields a set of 144 current-quarter forecasts, 141 next-quarter forecasts, 138 subsequent-quarter forecasts, 144 current-year forecasts, 132 forecasts for the next calendar year, and 120 forecasts for the calendar year after next. Each model's forecast accuracy is evaluated on the basis of the RMSE statistic.

Forecast Results. The RMSE results are reported in Table 1. The numbers in parentheses give the ratio of the RMSE of the associated model to the RMSE of the ZVAR forecasts at each horizon. A value greater than one for this ratio means that the RMSE of the given model is larger than for the ZVAR forecasts, indicating that those forecasts are less accurate. The modified Litterman model generally produces the smallest RMSE values across variables and forecast horizons. However, the ZVAR model is only slightly less accurate overall. The DOLS-AIC model also performs very well for the inflation and GDP growth forecasts but generates relatively poor unemployment forecasts. The basic Litterman and partial ZVAR models are ranked next in terms of accuracy, and the unrestricted OLS model is clearly

TABLE 1 RMSE of VAR Forecasts 1986–97[a]

	Current Quarter	First Quarter	Second Quarter	Current Year	First Year	Second Year
			Unemployment			
OLS	0.190 (1.25)	0.367 (1.30)	0.546 (1.33)	0.261 (1.54)	0.920 (1.42)	1.435 (1.46)
DOLS-AIC	0.161 (1.06)	0.324 (1.15)	0.519 (1.26)	0.202 (1.20)	0.879 (1.36)	1.377 (1.40)
Litterman	0.160 (1.05)	0.306 (1.10)	0.449 (1.09)	0.220 (1.30)	0.740 (1.14)	1.089 (1.10)
Modified Litterman	0.151 (0.99)	0.277 (0.98)	0.398 (0.97)	0.166 (0.98)	0.619 (0.96)	0.940 (0.95)
ZVAR	0.152	0.282	0.411	0.169	0.647	0.986
Partial ZVAR	0.159 (1.05)	0.302 (1.07)	0.439 (1.05)	0.216 (1.28)	0.718 (1.10)	1.069 (1.08)
			CPI Inflation			
OLS	1.216 (1.30)	2.105 (1.38)	2.207 (1.39)	0.552 (1.34)	1.719 (1.56)	2.560 (2.32)
DOLS-AIC	1.076 (1.15)	1.701 (1.12)	1.541 (0.97)	0.444 (1.08)	1.074 (0.98)	1.106 (1.00)
Litterman	0.973 (1.04)	1.737 (1.14)	1.787 (1.12)	0.472 (1.15)	1.227 (1.12)	1.451 (1.32)
Modified Litterman	0.920 (0.99)	1.523 (1.00)	1.616 (1.02)	0.400 (0.97)	1.106 (1.00)	1.133 (1.03)
ZVAR	0.933	1.524	1.590	0.411	1.099	1.102
Partial ZVAR	0.999 (1.07)	1.783 (1.17)	1.826 (1.15)	0.495 (1.20)	1.336 (1.21)	1.519 (1.38)
			GDP Growth			
OLS	2.819 (1.32)	3.122 (1.52)	2.982 (1.39)	0.954 (1.35)	2.156 (1.48)	2.509 (1.49)
DOLS-AIC	2.266 (1.06)	2.140 (1.04)	2.322 (1.08)	0.696 (0.99)	1.465 (1.00)	1.641 (0.98)
Litterman	2.340 (1.09)	2.332 (1.13)	2.459 (1.15)	0.800 (1.13)	1.776 (1.22)	1.878 (1.12)
Modified Litterman	2.237 (1.05)	2.036 (0.99)	2.133 (0.99)	0.710 (1.00)	1.430 (0.98)	1.621 (0.96)
ZVAR	2.141	2.058	2.147	0.706	1.456	1.681
Partial ZVAR	2.250 (1.05)	2.351 (1.14)	2.447 (1.14)	0.785 (1.11)	1.775 (1.22)	1.919 (1.14)

[a] The numbers in parentheses give the ratio of the RMSE of the associated model to the RMSE of the ZVAR forecasts at each horizon. A value greater than one means that the RMSE of the given model is larger than for the ZVAR forecasts, indicating that the given model's forecast is less accurate than the ZVAR forecasts.

Sources: Unemployment and CPI, Bureau of Labor Statistics; GDP, Bureau of Economic Analysis

dominated by all the others. The results are described in more detail below.[10]

Unemployment. The modified Litterman model generates the most accurate unemployment forecasts of any of the alternative forecast schemes considered here. However, the ZVAR model performs quite well, having an RMSE no more than 5 percent higher than the modified Litterman model. The Litterman and partial ZVAR models' forecasts are essentially equivalent, but each performed worse than either the ZVAR or modified Litterman models. For example, for the one-year annual forecast, the RMSE from the Litterman specification is 14 percent higher than from the ZVAR model, and the RMSE from the partial ZVAR model is 10 percent higher. The DOLS-AIC is the next best performing specification, but it has an RMSE more than 30 percent higher than the ZVAR for one- and two-year annual forecasts. The OLS model performs very poorly for any variable over any horizon other than the current quarter. For each of the one- and two-year annual forecasts the RMSE is over 40 percent higher than for the ZVAR model.

CPI Inflation. The ZVAR and modified Litterman specifications generate almost equally accurate inflation forecasts at all horizons, with the relative RMSEs differing by no more than 3 percent. The DOLS-AIC specification is marginally more accurate than either the ZVAR or modified Litterman model at the two-quarter and the one-year horizons. But the improvement in RMSE is no more than 3 percent over the ZVAR benchmark. The partial ZVAR generates somewhat worse forecasts than the Litterman model, especially for annual forecasts. More notably, the RMSE for the partial ZVAR is almost 40 percent higher than for the ZVAR model for two-year annual forecasts. The OLS model generates by far the worst-performing inflation forecasts, with an RMSE 132 percent higher than the ZVAR model for two-year annual forecasts.

GDP Growth. The ZVAR and modified Litterman specifications generate almost equally accurate forecasts of GDP growth, with the RMSEs differing by no more than 5 percent. The DOLS-AIC specification also yields almost the same forecast accuracy for the annual forecasts, and it has an RMSE that is no more than 8 percent higher than the ZVAR model for the quarterly forecasts. The Litterman and partial ZVAR models are the next most accurate, but each have RMSEs that are as much as 20 percent higher than the ZVAR model for annual forecasts. Again the OLS model performs the worst, having an RMSE for the two-year annual GDP growth forecast that is almost 50 percent higher than that from the ZVAR model.

The results suggest that using the sum of coefficients restriction, either exactly as in the case of the DOLS-AIC model or slightly more loosely as in the ZVAR and modified Litterman models, can significantly improve forecast performance over specifications that do not use such information. However, the fact that the restriction is not imposed exactly and that the cointegration prior is used might account for why the ZVAR and modified Litterman models decisively outperform the DOLS-AIC model in forecasting the unemployment rate.[11]

Conclusion

This article illustrates in some detail the steps involved in one approach to producing real-time forecasts from a VAR model. It focuses attention on the technical hurdles that must be addressed in a real-time application and methods for overcoming those hurdles. The solutions to technical difficulties include conditional forecasting to handle the staggered release of data and the interpolation of lower-frequency data to match the frequency of monthly data. In addition, the article discusses methods that attempt to improve VAR forecast accuracy by imposing inexact prior restrictions. The goal is to provide a road map for an analyst interested in designing and building a VAR forecasting model using these techniques.

The article then provides some suggestive empirical evidence regarding the performance of various possible specifications of a six-variable VAR in forecasting unemployment, inflation, and output growth. The forecast accuracy results show that using a particular setting for the systemwide inexact prior restrictions of the type described in Sims and Zha (1998) generates more accurate forecasts for unemployment, inflation, and real GDP growth than a VAR that uses the single-equation Litterman (1980) inexact priors. However, this improvement is largely explained by the incorporation of reasonably tight priors on the long-run properties of the VAR. This long-run aspect of the specification appears to matter more for the improvements in accuracy than the systemwide nature of the formulation does. In particular, a modified Litterman model that also incorporates these priors appears to be at least as accurate as the ZVAR model.

VAR models are increasingly being used for forecasting in private business and in policy institutions. It is hoped that the empirical techniques presented in this article will prove useful to those interested in implementing or at least understanding real-time forecasting with a VAR model.

10. *The impact of conducting a real-time forecasting experiment is most noticeable in the short-term GDP forecasts. In particular, the quarterly GDP forecasts can be more accurate one or two quarters ahead than they are for the current quarter. However, when the same experiment is conducted using the July 1998 vintage of historical data throughout, the forecast accuracy uniformly declines as one forecasts beyond the current quarter. Recall that a GDP estimate is revised on at least three occasions, and the size of the revisions to the monthly series used in constructing the monthly GDP data is often quite large. Of the series used to distribute GDP data over a quarter, industrial production in particular is often substantially revised (see Robertson and Tallman 1998 for discussion of this point). These real-time data revision issues affect the near-term (current and next-quarter) forecasting accuracy statistics in a way not captured in a forecasting analysis that uses only the latest available data.*

11. *This article has not searched across parameter settings for the non-OLS VAR specifications that might have improved their performance further.*

Data Definitions

Variables Included in the VAR

Real GDP: The value in real (1992) dollars of the output produced over a given quarter reported at a seasonally adjusted annual rate. Real GDP is measured in chain-weighted dollars to account for changes in relative prices over time. Availability is discussed below. Source: Bureau of Economic Analysis.

Civilian unemployment rate: The percentage of the civilian labor force that is unemployed. Seasonally adjusted. Released on either the first or second Friday of the month following the month measured. Source: Bureau of Labor Statistics.

Price level: Consumer price index (CPI) for all urban consumers. Not seasonally adjusted. Available by about the middle of the month following the measured month. Currently the average CPI value for the years from 1982 to 1984 is set equal to 100. A not seasonally adjusted series is used in the empirical analysis described in the article largely because it was readily available over the forecast period 1986 to 1997. However, there appears to be little seasonality in U.S. CPI data. Source: Bureau of Labor Statistics.

M2 money stock: Seasonally adjusted. Measured in billions of current dollars and available around the middle of the month after the month to which they refer. The aggregate is currently composed of the sum total of coins and paper currency, traveler's checks, demand deposits, other checkable deposits (NOW, share drafts), overnight repurchase agreements, overnight Eurodollars, general purpose and broker/dealer money market funds, money market deposit accounts, savings deposits, and small-denomination time deposits. Source: Board of Governors of the Federal Reserve System.

Effective federal funds rate: The interest charged between banks on loans of reserves held with the Federal Reserve System and measured as the monthly average of federal funds transactions for a group of federal funds brokers who report to the Federal Reserve Bank of New York each day. Source: Board of Governors of the Federal Reserve System.

Commodity prices (CP): Spot raw industrial subindex of thirteen markets for commodity prices. Not seasonally adjusted. Compiled daily and available at a monthly frequency at the end of the current month. The thirteen included prices are burlap, scrap copper, cotton, hides, lead scrap, print cloth, rosin, rubber, steel scrap, tallow, tin, wool tops, and zinc. Source: Commodity Research Bureau.

Monthly Series Used in the Chow-Lin Distribution Procedure

Nonagricultural payroll employment: Available on the first or second Friday following the measured month. Measured in millions of employed. Seasonally adjusted. Source: Bureau of Labor Statistics.

Total industrial production index: Available midmonth following the month measured. Measured as an index of physical output produced in a selection of sectors. 1987 = 100. Seasonally adjusted. Source: Board of Governors of the Federal Reserve System.

Real personal consumption expenditures: Seasonally adjusted. New estimates are usually available by the end of the month after that being measured. Source: Bureau of Economic Analysis.

Data Release Sequence

There is usually a delay of a few weeks between the end of a quarter and the release of the initial estimate of quarterly real GDP (the advanced estimate) for that quarter. Two revised real GDP estimates (preliminary and final) are released in the two subsequent months. Moreover, real GDP data for any particular quarter are subsequently revised, and this process of revision can continue many years after the initial data release (see Robertson and Tallman 1998 for a discussion).

Monthly civilian unemployment rate data are published on the first or second Friday of the month immediately following the month to which they refer, and the CPI data are published around the middle of the month following the month to which the data refer. There is generally little revision to the unemployment and CPI data over time, although major benchmark revisions to the CPI are made approximately every ten years.

The Board of Governors of the Federal Reserve System publishes monthly data on various money aggregates by the middle of the month after the month to which the data refer. These estimates are revised on a continuing basis with the receipt of more accurate source data, and on occasion the historical M2 data have been subject to major redefinition (see Anderson and Kavajecz 1994).

Numerous commodity price indexes are available on a daily basis. The study uses the monthly average of the daily closing CRB/BLS index of spot prices for raw industrial commodities as the commodity price. The federal funds rate data are the monthly average of the daily effective funds rate.

The Chow and Lin (1971) Procedure for Distributing Quarterly GDP Observations

Chow and Lin (1971) derive a procedure for distributing quarterly observations on a flow variable across the months of a quarter. Their algorithm assumes that the T observations in the monthly series y_m to be estimated are related to T observations on n monthly indicator variables X_m via a regression of the form

$$y_m = X_m \beta + u_m, \tag{A1}$$

where y_m is $T \times 1$, X_m is $T \times n$, and the regression error follows a stationary first-order autoregression $u_{m,t} = \rho_m u_{m,t-1} + e_{m,t}$ for $t = 1, \ldots, T$, with the $T \times 1$ vector e_m having zero mean and a covariance matrix $\sigma^2 I_T$. Thus, the covariance matrix of u_m has the standard form $V_m = [\sigma^2/(1-\rho^2_m)]P_m$, where

$$P_m = \begin{bmatrix} 1 & \rho_m & \cdots & \rho_m^{T-1} \\ \rho_m & 1 & & \rho_m^{T-2} \\ \vdots & & \ddots & \vdots \\ \rho_m^{T-1} & & \cdots & 1 \end{bmatrix}$$

(see for example, Hamilton 1994).[1]

The $(T/3)$ quarterly observations are related to the monthly observations via an "averaging" matrix

$$C = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 0 & & & \cdots & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \cdots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & & \cdots & 1 & 1 & 1 \end{bmatrix}.$$

This matrix implies a regression relationship for the quarterly observations of the form

$$y_q = Cy_m = X_q \beta + u_q \tag{A2}$$

and where the covariance matrix of u^q will be

$$V_q = CV_m C'. \tag{A3}$$

Chow and Lin show that the smallest variance linear unbiased estimator of y_m is

$$\hat{y}_m = X_m \hat{\beta} + \hat{P}_m C'(C\hat{P}_m C')^{-1}\hat{u}_q,$$

where $\hat{\beta}$ is the (generalized) least squares estimate of β in equation (A2). To estimate ρ_m notice that the auto-regression coefficient in the quarterly regression, ρ_q, is the ratio of the first to the second elements of the first row of V_q in equation (A3). This ratio reveals that ρ_m can be obtained as the unique solution to the polynomial

$$\rho_q = \frac{\rho_m^5 + 2\rho_m^4 + 3\rho_m^3 + 2\rho_m^2 + \rho_m}{2\rho_m^2 + 4\rho_m + 3}$$

and replacing ρ_q with the (generalized) least squares estimate of ρ_q in equation (A2) provides a consistent estimate of ρ_m.

1. *If* X_m *contains a stochastic trend, then the regression represents a cointegrating relationship. A modification of the Chow and Lin procedure that was suggested by Litterman (1983) would be appropriate in the case that* $\rho_m = 1$.

REFERENCES

ANDERSON, RICHARD G., AND KENNETH A. KAVAJECZ. 1994. "A Historical Perspective on the Federal Reserve's Monetary Aggregates: Definition, Construction, and Targeting." Federal Reserve Bank of St. Louis *Economic Review* 76 (March/April): 1–31.

CHOW, GREGORY C., AND AN-LOH LIN. 1971. "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Time Series." *Review of Economics and Statistics* 53:372–75.

DOAN, THOMAS, ROBERT B. LITTERMAN, AND CHRISTOPHER A. SIMS. 1984. "Forecasting and Conditional Projection Using Realistic Prior Distributions." *Econometric Reviews* 3:1–100.

DRÈZE, JACQUES H., AND JEAN-FRANÇOIS RICHARD. 1983. "Bayesian Analysis of Simultaneous Equation Systems." In *Handbook of Econometrics*, vol. 1, edited by Zvi Griliches and Michael D. Intriligator. Amsterdam: North-Holland.

FAIR, RAY C., AND ROBERT J. SHILLER. 1990. "Comparing Information in Forecasts from Econometric Models." *American Economic Review* 80, no. 3:375–89.

HAMILTON, JAMES. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

INGENITO, ROBERT, AND BHARAT TREHAN. 1996. "Using Monthly Data to Predict Quarterly Output," Federal Reserve Bank of San Francisco *Economic Review* 3:3–12.

KADIYALA, K. RAO, AND SUNE KARLSSON. 1993. "Forecasting with Generalized Bayesian Vector Autoregressions." *Journal of Forecasting* 12, no. 4:365–78.

———. 1997. "Numerical Methods for Estimation and Inference in Bayesian VAR Models." *Journal of Applied Econometrics* 12, no. 2:99–132.

LITTERMAN, ROBERT B. 1980. "A Bayesian Procedure for Forecasting with Vector Autoregression." Massachusetts Institute of Technology, Department of Economics Working Paper.

———. 1983. "A Random Walk, Markov Model for the Distribution of Time Series." *Journal of Business and Economic Statistics* 1 (April): 169–73.

———. 1984. "Forecasting and Policy Analysis with Bayesian Vector Autoregression Models." Federal Reserve Bank of Minneapolis *Quarterly Review* 8, no. 4:30–41.

———. 1986. "Forecasting with Bayesian Vector Autoregressions—Five Years of Experience." *Journal of Business and Economic Statistics* 4 (January): 25–38.

LUTKEPOHL, HELMUT. 1991. *Introduction to Multiple Time Series Analysis*. New York: Springer-Verlag.

MILLER, PRESTON J., AND DAN M. CHIN. 1996. "Using Monthly Data to Improve Quarterly Forecasts." Federal Reserve Bank of Minneapolis *Quarterly Review* 20 (Spring): 16–33.

MILLER, PRESTON J., AND WILLIAM ROBERDS. 1991. "The Quantitative Significance of the Lucas Critique." *Journal of Business and Economic Statistics* 9, no. 4:361–87.

REIFSCHNEIDER, DAVID L., DAVID J. STOCKTON, AND DAVID W. WILCOX. 1997. "Econometric Models and the Monetary Policy Process." *Carnegie Rochester Conference Series on Public Policy*, no. 47:1–37.

ROBERTSON, JOHN C., AND ELLIS W. TALLMAN. 1998. "Data Vintages and Measuring Forecast Model Performance." Federal Reserve Bank of Atlanta *Economic Review* 83 (Fourth Quarter): 4–20.

SIMS, CHRISTOPHER A. 1992. "A Nine-Variable Probabilistic Macroeconomic Forecasting Model." In *Business Cycles, Indicators, and Forecasting*, edited by James H. Stock and Mark W. Watson. Chicago: University of Chicago Press.

SIMS, CHRISTOPHER A., AND TAO A. ZHA. 1998. "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review* 39, no. 4:949–68.

TALLMAN, ELLIS W., AND TIM PETERSON. 1998. "Improving Forecasts of Real Output Growth by Using Timely Data." Unpublished manuscript, March.

WAGGONER, DANIEL, AND TAO A. ZHA. 1998. "Conditional Forecasts in Dynamic Multivariate Models." Federal Reserve Bank of Atlanta Working Paper 98-22, December.

WALLIS, KENNETH. 1989. "Macroeconomic Forecasting: A Survey." *Economic Journal* 99, no. 394:28–61.

ZHA, TAO A. 1998. "A Dynamic Multivariate Model for Use in Formulating Policy." Federal Reserve Bank of Atlanta *Economic Review* 83 (First Quarter): 16–29.

Against the Tide: Malcolm Bryan and the Introduction of Monetary Aggregate Targets

R.W. HAFER

The author is an Atlanta Fed visiting scholar and professor of economics, Southern Illinois University at Edwardsville. *He thanks Stu Allen, Dave Ault, Michael Bischof, Frederick Deming, Bill Dewald, Jerry Dwyer, Milton Friedman, Gail Heyne Hafer, Joe Haslag, Bob Hetzel, Thomas Mayer, Bill Poole, Anna Schwartz, Frank Steindl, Dave Wheelock, and especially Jim Meigs for helpful comments.*

MONETARY POLICY WAS FREED FROM THE STRAIGHTJACKET OF PEGGING U.S. TREASURY INTEREST RATES FOLLOWING THE TREASURY–FEDERAL RESERVE ACCORD IN 1951. THIS NEWFOUND FREEDOM LED TO A GROWING DEBATE INSIDE AND OUTSIDE THE FEDERAL RESERVE SYSTEM ABOUT THE APPROPRIATE MEASURES TO USE AS OPERATING GUIDES. AS THE 1950S PROGRESSED, THE FEDERAL RESERVE FOCUSED ON CONTROLLING MARKET INTEREST RATES TO ACHIEVE ITS POLICY GOALS. THIS POLICY, WHICH REMAINS IN USE TODAY, CAME UNDER FIRE FROM A HANDFUL OF POLICYMAKERS AND POLICY WATCHERS. AN ALTERNATIVE SUGGESTED BY SOME WAS TO PLACE GREATER EMPHASIS ON THE BEHAVIOR OF THE MONETARY AGGREGATES IN SETTING POLICY. THIS ARTICLE EXAMINES THE CONTRIBUTIONS OF MALCOLM BRYAN, PRESIDENT OF THE FEDERAL RESERVE BANK OF ATLANTA FROM 1951 THROUGH 1965, TO THIS DEBATE AND TO THE DEVELOPMENT OF MONETARY POLICY IN THE POSTACCORD ERA.

Bryan parted company with most of his colleagues on the Federal Open Market Committee (FOMC) during the late 1950s and into the 1960s. Bryan tried to steer policy away from focusing on interest rates and money market conditions to placing more weight on the monetary aggregates. A reading of the transcripts of the FOMC meetings during this time reveals that Bryan's alternative policy reflected his desire to prevent the disruptive effects on the economy from short-run fluctuations in money growth and the longer-term effects of expansive Fed actions, namely, inflation.[1] Bryan and a few other committee members considered money market conditions and changes in interest rates to be inadequate indicators of policy actions. Bryan argued that monetary aggregates not only provided better feedback on policy changes than tone and feel but also afforded the FOMC a better gauge for measuring the success of achieving its desired policies.

Bryan was not a lone figure in this debate, but his contributions deserve special mention.[2] For example, in 1959 he became the first member of the FOMC to introduce an explicit quantitative, long-run aggregate target into postwar policy discussions. In 1960 he further separated himself from the majority opinion on the FOMC by developing and introducing short-run, monetary growth targets—growth cones, as they would become known in the 1970s—into the policy debate.[3] Bryan used these quantitative indicators to help interpret and steer Fed policy during the late 1950s and early 1960s. Reviewing the debate over the usefulness of monetary aggregates as operating guides provides an informative case study of U.S. monetary policy, one that offers valuable lessons for monetary policy even today.

To provide a background for Bryan's introduction of aggregate targeting, the article first describes the state of monetary policy and the economy during the period following the 1953–54 recession and through 1958. Of particular interest is how the FOMC interpreted economic developments as the economy emerged from the recession and the inflationary shadow of the Korean War. Against this backdrop, the article discusses Bryan's 1959 introduction of a long-term aggregate target, one based on the postwar trend in the growth of reserves. The next section focuses on the policy debate in 1960 and Bryan's introduction of a short-term, reserve growth cone as an operating guide. The concluding section offers some final observations.

Monetary Policy and the Economy: 1956–58

Prior to the Treasury–Federal Reserve Accord, monetary aggregates essentially were ignored as policy operated within the confines of supporting Treasury security prices. Following the accord, the policy role of monetary aggregates expanded, though the shift was more cosmetic than real (Friedman 1982). The Federal Reserve focused on achieving desired market rates and maintaining orderly markets to hit stated policy goals, such as sustained economic growth and low inflation. Federal Reserve policy during the 1950s and into the 1960s often relied on hitting target levels of free reserves—the difference between banks' excess reserve holdings and reserves borrowed through the discount window—in the banking system to bring about changes in financial markets. Goodfriend (1991), for example, argues that free reserves in the 1950s and 1960s, like nonborrowed reserves in the 1980s, provided a distraction to the Fed's primary policy concern of manipulating market interest rates. Thus, even though free reserves are a "monetary" aggregate, their use in policy was conditioned on activity in the financial markets.

Bryan's aggregate-based approach to monetary policy was a dramatic departure for a member of the FOMC at the time.

Changes in free reserves were brought about as the FOMC and the manager of the Open Market Desk at the Federal Reserve Bank of New York took their cues from events in both the domestic and, later, international financial markets. Monetary policy in the 1950s thus depended heavily on the subjective judgment of the FOMC and the manager of the desk.[4] Their role was to interpret developments in the financial markets—to determine the tone and feel of the markets—and how these would influence and be influenced by policy actions.[5] The combination of free reserves and tone and feel did give the Fed some control over short-term interest rates, essentially the three-month

1. *The FOMC is the policy-making arm of the Federal Reserve System. It is composed of the seven members of the Board of Governors and five of the twelve district bank presidents, four of whom vote on a rotating basis, and the New York Federal Reserve Bank president, who is a permanent voting member. For a recent use of the FOMC transcripts as the basis for analyzing Fed actions, see Edison and Marquez (1998).*
2. *Meigs (1976) chronicles the contributions of D.C. Johns, the president of the Federal Reserve Bank of St. Louis, and Homer Jones, the director of research at St. Louis. Meigs also provides a brief discussion of Bryan's role in the developing debates that would later be centered on the Federal Reserve Bank of St. Louis.*
3. *The term cones comes from their construction. For example, supposing that the base period is the average value for the level of an aggregate in the fourth quarter of a year and assuming that this value is $100, if a 5 percent growth path is the policy objective for the year, then the average value for money in the fourth quarter of the following year would be $105. Allowing for a growth path somewhat higher (for instance, 7 percent) and lower (for instance, 3 percent) would give quarter-average values of $107 and $103, respectively. As shown below, connecting the base period value with the upper and lower ranges creates a cone of possible values.*
4. *Monetary policy is conducted through the Federal Reserve Bank of New York primarily by buying and selling government securities in the open market. This activity takes place through the Open Market Trading Desk, supervised by the manager of the desk.*
5. *Atkinson (1969) shows that the FOMC often switched between free reserves and looking to tone and feel during the 1950s and 1960s. Even though the FOMC officially used free reserves as the operating guide, Atkinson's evidence indicates that doing so did not reduce the variance of interest rates or lead to better control over reserves than proposals that used tone and feel as guidelines. For an early analysis of the problems associated with the use of free reserves, see Brunner and Meltzer (1964).*

Treasury bill rate. But the policy had its costs. Calomiris and Wheelock (1998) assert that the Fed's reliance on free reserves as an operating guide simply recycled the policies that Benjamin Strong had advocated during the 1920s. Operating under the limited constraints of the gold standard, the Fed manipulated free reserves to achieve desired levels of short-term interest rates with one eye toward the domestic economy and the other toward the growing problems of external imbalances. Such policy choices, Wheelock (1997) suggests, help explain the upward drift in money growth and inflation that lasted for the next two decades.

Bryan and a few other FOMC members considered money market conditions and changes in interest rates to be inadequate indicators of policy actions.

Dissension among FOMC members arose over the best course for policy following the 1953–54 recession. The economy in 1955 grew quite rapidly after the recession, with real gross national product (GNP) increasing at an annual rate of more than 6 percent. Brisk real growth and the sharp run-up in prices that followed the Korean War made inflation a primary concern at FOMC meetings throughout 1955. Hetzel notes that inflation was "the primary macroeconomic preoccupation of the political system in the 1950s" (1995, 6). By the end of 1956, however, real growth had slowed considerably, increasing at an annual rate of only 1.4 percent for the year. Even so, the members' inflation fears now seemed justified: the price level increased at an annual rate of more than 3.5 percent during 1956, up from a 2.5 percent annual rate of change in 1955.

The FOMC reacted to the potential of higher rates of inflation with a policy of increased restraint during 1956 and into 1957. William McChesney Martin, the chairman of the Board of Governors, voiced the majority opinion that the Fed should not repeat the mistake it made coming out of the last recession, essentially that of not raising market rates fast enough to curb inflation. The increased policy restraint resulted in money growth (M1) falling from a 2.2 percent annual growth rate in 1955 to a 1.1 percent rate in 1956. This constraint persisted into 1957, with M1 decreasing at an annual rate of 0.5 percent in 1957, bank credit flat, and a three-month Treasury bill rate that rose throughout the year.

This episode intensified the committee's internal debate over the choice of policy guides. Bryan's comments at the January 28, 1957, FOMC meeting are representative of the confusion and uncertainty that using free reserves engendered. He argued that the behavior of free reserves was not "particularly useful at the present time" (FOMC 1957, 13) and that some alternative should be discussed.[6] Board economist Woodlief Thomas noted that using tone and feel, the companion operating guide at the time, often led to changes that unfortunately contradicted the policy desires of the committee. Governor J.L. Robertson, among others, also expressed discontent with tone and feel as an effective operating guide. The FOMC's dilemma was that neither approach seemed to provide very reliable signals about policy actions and their effects on the economy.

Chairman Martin took the position often expressed by Fed chairmen: operating guides such as tone and feel may be less than perfect, but they afforded the FOMC and the manager of the Open Market Desk the needed flexibility to respond to unforeseen changes in financial markets or the economy. Robert Roosa recalled that "the Federal Reserve has had to rely primarily on experimental probing. . . . *[U]tilizing its own qualitative concept of pressure*, it has withheld or released new bank reserves . . . the 'feel' arising from participation in securities markets and broader judgments of current economic trends" (1960, 262; emphasis added).[7] Alfred Hayes, president of the New York Fed and vice chairman of the FOMC, gives another perspective on the process of using tone and feel to guide policy: "[T]he tone of the market is a very difficult thing to describe unless you are actually sitting at this trading desk, which is the nerve center of the bank and the nerve center of the System for keeping in touch with credit and banking and money market developments. But I would say that it [tone] is a compound of all kinds of impressions you get from the volume of trading, the speed of trading, what is happening to prices, what the bank's position is, whether the dealers are hard up for financing or have plenty of financing, whether funds are well distributed throughout the country or not well distributed" (Atkinson 1969, 85).[8] Chairman Martin and others extolled the flexibility that tone and feel offered. Martin also dismissed claims that monetary aggregates could serve as credible operating guides, insisting on more than one occasion that monetary policy would not be constrained by the "dead hand of statistics," something that he associated with reliance on the aggregates. Sentiment for a change in operating guides was not overwhelming.

By the end of 1957, a few members of the FOMC advised that the Fed's restrictive policy was having a deleterious effect on the economy and needed to be reversed. Bryan was one of the more vocal critics of current policy. At the November 12, 1957, meeting he warned that the lack of reserve and money growth was creating "a terrific drag" on the economy. Unless the FOMC moved to quickly reverse this policy, Bryan warned, the Federal Reserve would be "party to producing economic convulsions" (FOMC 1957, 695). This view was not new for Bryan (Bryan 1938, 1948). A decade earlier, in a speech before the Alabama Bankers Association, Bryan suggested that "the central bank must *lean against the breeze* in times of boom and inflation and likewise in times of depression and deflation" (1948; emphasis added). At the December 17, 1957, meeting Bryan reiterated his long-held view on the nature of policy effects, further stating that, "I believe [monetary policy] can play its most effective role in a downturn if monetary ease is injected *during the early stages of a downward movement* rather than after the recession is well underway. . . . [I]n the face of a now clearly perceptible economic downturn, our effective policy, whatever our intentions, has been to allow a reserve base providing for no growth whatever in the economy. I believe it is clear that the continuation of such policy must finally be an important causative factor in promoting a serious recession" (FOMC 1957, 801–2; emphasis added).

Most other members of the FOMC, however, pressed for continued restraint since their primary concern was to avoid "sloppy" financial markets. In general, committee members and the chairman believed that reversing the course of monetary policy would have little impact on curtailing any recessionary momentum that might already be under way in the economy.

Bryan's view of monetary policy and its effect on real economic activity was not ordinary for a member of the FOMC at that time. His approach to judging policy, contrary to his colleagues and many others in the economics profession, employed ideas associated with a small number of monetary economists that were beginning to circulate through the profession. For example, compare Bryan's comments cited above with Milton Friedman's testimony before the Joint Economic Committee in March 1958. Friedman asserted that actions taken by the Federal Reserve were a "causative factor" in explaining past recessions. Friedman's analysis of recent Fed policies concluded that "the tight money policy of 1956 and 1957 which coexisted with rising prices . . . [is] with us in the current recession" (Friedman 1958, 250, quoted in Meigs 1976, 445).

Bryan was aware of and closely followed developments in monetary economics, as a correspondence between Bryan and Friedman indicates. In a letter dated April 7, 1959, Bryan wrote to Friedman asking for a copy of his paper "Some Theoretical and Empirical Aspects of the Supply of Money."[9] From Friedman's response, it is clear that Bryan had sent along a copy of his paper "The Sovereign, the Central Bank, and the Monetary Standard" (1959), which he had delivered several times in speeches. Later Friedman wrote to Bryan, sending him a copy of his recently published *A Program for Monetary Stability*. This correspondence, albeit limited, suggests that Bryan followed the developments and debates in the increasingly active area of monetary economics and sought input from one of its leading theorists and proponents for an aggregates-based policy.[10]

Prior to the Treasury–Federal Reserve Accord, monetary aggregates essentially were ignored as policy operated within the confines of supporting Treasury security prices.

Bryan's policy analysis was not solely the product of others' research, however. Bryan's views on the role and effects of monetary policy, as indicated in a series of speeches, was established as early as 1938. By 1957

6. *Although quotation marks appear, the minutes represent the FOMC Secretary's summary of the discussion and are not necessarily verbatim. Even so, FOMC members had the opportunity to correct the minutes before they entered the permanent record.*
7. *This notion of flexibility can be found throughout FOMC discussions. For example, compare Roosa's and Martin's comments to those of Chairman Paul Volcker at the December 20–21, 1982, meeting of the FOMC: "I think we're left with what could be termed an eclectic, pragmatic approach. It's going to involve some judgment as to which one of these [aggregate] measures we emphasize, or we may shift from time to time. . . . [W]e're going to have to make some judgments as to which one is more significant at any particular point in time against what nominal GNP is or what the goal is or what the real economy is doing and what prices are doing and all the rest. . . . [T]hat's the way the Federal Reserve used to operate, less elaborately, for years when policy by present standards looked pretty good" (FOMC 1982, 41).*
8. *The source of the quote is Hayes's testimony before the Joint Economic Committee in 1961.*
9. *It is likely that Bryan meant "Some Theoretical and Empirical Aspects of the Demand for Money" since there is no reference in the NBER list of publications to the former piece.*
10. *Thanks to Milton Friedman for making this correspondence available. Bryan's views on monetary policy and the effects of money probably reflect the fact that he received postgraduate training at the University of Chicago (see Box 1). This is the school often associated with Friedman and so-called monetarist economics. To get a feel for the opinion that many economists held of such views, the remarks of Richard Davis, an economist at the Federal Reserve Bank of New York in 1969, can be considered: "[T]he view that 'only money matters' or, perhaps more accurately, that 'mainly money matters' was the province of an obscure sect with headquarters in Chicago. For the most part, economists regarded this group—when they regarded it at all—as a mildly amusing, not quite respectable collection of eccentrics" (1969, 119).*

BOX 1

Biography of Malcolm Bryan

Bryan was born in 1902 in Wateska, Illinois, a small town of several thousand a little more than 100 miles south of Chicago. At the age of twenty-two Bryan graduated with a bachelor's degree from the University of Illinois, where he remained for an additional year to earn a master's degree. Presumably the master's degree was in economics since following graduation Bryan took a position in the economics department at the University of Georgia.

During his years at the University of Georgia (1925–36), Bryan served in several positions. In 1929 he was a member of Georgia's Special Tax Commission, and in 1934 he served on the Special Committee on Banking and Taxation under the auspices of the U.S. Treasury. From 1933 to 1937, Bryan served as editor of the *Southern Economic Journal*. Perhaps the most important aspect of this period in Bryan's life is that he spent 1927 and 1928 doing postgraduate work in economics at the University of Chicago, where it is likely that his views on the role of monetary policy and the link between monetary aggregates and the economy were influenced by economists Viner, Knight, Douglas, Mints, and Simon, all faculty members at that time.

Bryan left the University of Georgia in 1936 and began his career in the Federal Reserve System with a two-year stint as an economist at the Board of Governors. Afterward he returned to Georgia as a vice president of the Federal Reserve Bank of Atlanta, a position that he held from 1938 through 1941. In 1941 Bryan was promoted to first vice president of the Atlanta Fed, where he remained until 1946. While Bryan was at the Atlanta Fed, he was elected president of the Southern Economics Association in 1942. He also served on the American Technical Staff, part of the negotiating team at the Bretton Woods Monetary and Financial Conference in 1944.

After the end of World War II, Bryan left the Atlanta Fed to become vice chairman of the Trust Company of Georgia. He remained in this job from 1946 until 1951. During this interval away from the Fed, Bryan served as a member of the Senate Finance Committee's Advisory Committee on Social Security (1947–48) and as part of the Economic and Financial Mission to Peru in 1948. Bryan left the Trust Company of Georgia in 1951, returning to the Atlanta Fed as its president. He remained in this position for the rest of his professional career, retiring in 1965. Bryan died in 1967.

Bryan already was arguing that significant, short-run changes in money growth were likely to influence real economic activity. This date suggests that Bryan's policy stance predates the monetarist position usually associated with the Federal Reserve Bank of St. Louis. For instance, Homer Jones, a leading proponent of using monetary aggregates, did not begin as research director at St. Louis until 1958. And Meigs's observation that "[t]he *new* element in the St. Louis position in 1960 was a recognition that short run changes in the money stock can have adverse effects on income and employment" (1976, 447; emphasis added) was a conviction Bryan used in his policy analysis. The development suggests that the maverick views so often associated with the St. Louis bank were already operating at the Atlanta bank.[11]

Monetary policy abruptly changed direction in 1958. The growth of M1 increased dramatically: after decreasing at an annual rate of 2.1 percent in the first quarter of the year, the growth rate of M1 jumped to more than 6 percent during the next two quarters. At the same time, real output roared out of the recession, increasing at better than a 10 percent annual rate during the second half of the year. In contrast to the positions of his colleagues, Bryan's statements at FOMC meetings throughout 1958 reflect an evolution in his view about the economic effects of short-run fluctuations in money growth. Eschewing the common practice of measuring policy on a meeting-to-meeting basis—often a period of only weeks—Bryan began to compare the level of reserves at one meeting with that of the previous year. While others focused on measuring reserve growth on a meeting-to-meeting basis, Bryan put current policy analysis and discussion into a longer-term perspective in order to understand the current stance of policy actions. This development is reflected in his introduction of a reserve growth target in 1959.

Introducing a Reserve Growth Target: 1959

Bryan's concern about the inflationary effects associated with long-term reserve growth is consistent with his often expressed distress about the dangers of inflation and the Federal Reserve's responsibility to contain it. A popular notion at the time was that a little inflation was good for the economy. Bryan considered this view to be economically naive and morally bankrupt. For example, he publicly argued in a speech that inflation was merely a "transfer of purchasing power from savers in money forms to other classes of society" and a process that erodes the very foundation of a market economy. He believed that "once money is destroyed as a store of value or its function therein seriously impaired . . . the judgment of the consumer, the saver, the businessman and often governments as to their best interests in the presence of inflation as against what their judgment of their best interests would be in the absence of inflation" is negatively affected (1957).

Following the expansionary policies of 1958, the threat of inflation once again pervaded FOMC meetings into 1959. For instance, as early as the January 6, 1959, meeting, the presentation to the FOMC by staff economists characterized the economic situation as a "maturing recovery" with the problems of recession replaced by "problems of sustainable growth" without igniting inflation (FOMC 1959, 5). Money growth continued strong, with M1 increasing at an annual rate of 4 percent during the second half of 1958.

A major problem facing the FOMC was that its desire to reign in potential inflation conflicted with its perceived necessity to accommodate the Treasury's financing needs. In the spring of 1959 a few FOMC members openly expressed dissatisfaction with such policy constraints. D.C. Johns, for example, said at the February 10 meeting that the FOMC should "pay more attention to what was happening to the money supply" in its discussions (FOMC 1959, 92). Governor A.L. Mills suggested at the March 3 meeting that continued use of free reserves would likely repeat the "unhappy experience" of 1958 when Fed actions caused a "supercharged growth in the money supply" (167). Bryan continued in this vein at the April meeting, asserting that recent reserve growth was "sufficient to finance a first-rate inflation and [that] it would easily be possible to get into trouble" (288–89). He repeated this warning in May, noting that policy "had not been particularly restraining" (313).

Bryan also questioned the manner in which policy instructions were communicated to the manager of the desk: the directive. A chronic problem for the committee was that different members often had different interpretations of the directive to which they had agreed. (To provide a perspective on the problem, the directive forwarded to the desk from the May 5, 1959, meeting is reprinted in Box 2.) For example, the consensus at the May 5 meeting called for an even-keel policy, prompted by upcoming Treasury financing needs. Bryan questioned what this meant: was an even-keel policy "measured by net free reserves, net borrowed reserves, the feel of the market, or the intuition of the Account Manager"? (FOMC 1959, 340). The manager, in response, "thought it was a mixture of the things [Bryan] had mentioned" (340). Relying on free reserves to achieve interest rate targets engendered uncertainty over policy directions to the desk.

Operating guides such as tone and feel may be less than perfect, but they afforded the flexibility to respond to unforeseen changes in financial markets or the economy.

Bryan became an increasingly outspoken proponent for changing the directive from a qualitative description of policy desires—firm up the markets, achieve some ease in free reserves, and so forth—to one that established numerical targets for policy. In the summer of 1959, Bryan began to base his policy analysis on the short- and long-run growth of a reserve measure that was developed by the Atlanta bank, called total effective reserves.[12] At the August 18, 1959, meeting Bryan introduced something novel in postwar FOMC deliberations: the idea of gauging policy by tracking total effective reserves relative to their postwar trend, an annual average growth rate of 3.6 percent.[13] Bryan argued that when

11. *In personal discussions, Jim Meigs relates that Bryan and Johns often met outside the FOMC meetings to discuss policy developments. Bryan, the professional economist, is likely to have influenced Johns, a lawyer by training, in matters of monetary policy.*

12. *The total-effective-reserves measure developed at Atlanta is similar to the St. Louis adjusted monetary base series less currency. Total effective reserves are measured by first calculating the average value for the ratio of required reserves to average deposits beginning in May 1958 through December 1959. May 1958 is used since it is the last time reserve requirements were changed. This ratio is 0.1152. For the period prior to May 1958, this ratio is divided by the monthly ratio of required reserves to deposits and the value of this term multiplied by actual reserves—in other words,* $[0.1152/(R_r/D)] \times R$. *From May 1958 onward, actual member bank reserves are used. Both reserve measures are then seasonally adjusted. This computation is outlined in the appendix to the January 26, 1960, FOMC meeting. This measure can be replicated using the data in Appendix D of Meigs (1962). For example, this author's trend estimate is that reserves increase, on average, $42.7 million per month compared with Bryan's reported estimate of $43 million per month.*

13. *Meigs (1976, 445) suggests that Bryan introduced the use of total effective reserves at the November 24 meeting of the FOMC. It was at the November 24 meeting that Bryan introduced the charts upon which his policy discussions actually had been based since August.*

BOX 2

The Directive of May 5, 1959

Thereupon, upon motion duly made and seconded, the Committee voted unanimously to direct the Federal Reserve Bank of New York, until otherwise directed by the Committee:

(1) To make such purchases, sales, or exchanges (including replacement of maturing securities, and allowing maturities to run off without replacement) for the System Open Market Account in the open market or, in the case of maturing securities, by direct exchange with the Treasury, as may be necessary in the light of current and prospective economic conditions and the general credit situation of the country, with a view (a) to relating the supply of funds in the market to the needs of commerce and business, (b) to fostering conditions in the money market conducive to sustainable economic growth and stability, and (c) to the practical administration of the Account; provided that the aggregate amount of securities held in the System Account (including commitments for the purchase or sale of securities for the Account) at the close of this date, other than special short-term certificates of indebtedness purchased from time to time for the temporary accommodation of the Treasury, shall not be increased or decreased by more than $1 billion;

(2) To purchase direct from the Treasury for the account of the Federal Reserve Bank of New York (with discretion, in cases where it seems desirable, to issue participations to one or more Federal Reserve Banks) such amounts of special short-term certificates of indebtedness as may be necessary from time to time for the temporary accommodation of the Treasury; provided that the total amount of such certificates held at any one time by the Federal Reserve Banks shall not exceed in the aggregate $500 million.

Source: FOMC (1959, 341–42)

reserves go below their postwar trend, he "would be inclined to resolve any doubts [on the stance of policy] slightly on the side of ease" (FOMC 1959, 558). Bryan pushed for the adoption of his reserve measure and its trend growth as the operating guide for policy to replace financial market behavior and free reserves. Not surprisingly, little was made of his suggestion at this time.

Monetary policy by late 1959 confronted conflicting economic signals. Domestically, money growth had deteriorated sharply, with M1 decreasing at an annual rate of 3.4 percent in the fourth quarter of 1959. The steel strike that began in mid-July and ended in November disrupted the relation between policy actions and the economy. Research out of the St. Louis bank, for example, suggested that the decline in economic activity due to the strike lowered velocity and made the recent slowing in money growth less restrictive than normal. Uncertainty about the domestic economy and the correct policy response was compounded by a worsening balance-of-payments situation. As Hetzel describes it, "monetary policymakers walked a tightrope requiring them to balance internal and external objectives . . . each requiring conflicting policy responses" (1996, 23). By 1959, the currencies of the European Economic Community countries had become fully convertible into U.S. dollars.[14] A persistent balance-of-payments deficit generated significant gold outflows from the U.S. Treasury. This situation not only caused alarm among policymakers but was also politically unacceptable. With domestic interest rates below European rates, the Fed moved to raise the three-month Treasury bill rate in an attempt to quell the gold outflow. Monetary policy thus was conducted with heightened uncertainty (and disagreement) over which objective—internal or external balance—was more important.

Bryan believed that monetary policy did not cause and could do little to solve external imbalances. At the September 22 meeting he asserted that unless reserves increased fast enough to satisfy seasonal needs plus some positive growth, the "System could easily get itself into the position of bringing about greater tightness this fall than it desired" (FOMC 1959, 649). In other words, the current course of policy would, in all likelihood, cause another recession. In contrast, Chairman Martin said at the November 24 meeting that while he "shared some of the apprehension that had been expressed about the money supply and the relationship of credit to growth . . . he did not believe this was the time to correct it" (828). Martin steadfastly relied on the flexibility given by tone and feel to guide policy decisions even though Robert Rouse, manager of the desk, cautioned at the December 15 meeting that "interpretation of the signals given off by the market was by no means easy" (836). Martin simply

TABLE 1 Compounded Annual Growth Rates fo Effective Reserves[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	1.8	1.0	2.0	2.6	3.2	3.1	3.4	3.3	3.1	2.9	3.3	3.0
1948	x	x	0.2	2.0	2.9	3.5	3.3	3.6	3.5	3.2	3.0	3.4	3.3
1949	x	x	x	3.9	4.3	4.6	4.1	4.3	4.1	3.7	3.4	3.8	3.6
1950	x	x	x	x	4.6	5.0	4.2	4.4	4.1	3.7	3.3	3.8	3.6
1951	x	x	x	x	x	5.4	4.0	4.3	4.0	3.5	3.1	3.7	3.5
1952	x	x	x	x	x	x	2.6	3.8	3.5	3.0	2.6	3.4	3.2
1953	x	x	x	x	x	x	x	5.0	4.0	3.1	2.6	3.6	3.3
1954	x	x	x	x	x	x	x	x	3.0	2.2	1.9	3.3	3.0
1955	x	x	x	x	x	x	x	x	x	1.3	1.3	3.3	3.0
1956	x	x	x	x	x	x	x	x	x	x	2.0	4.3	3.6
1957	x	x	x	x	x	x	x	x	x	x	x	6.6	4.7
1958	x	x	x	x	x	x	x	x	x	x	x	x	2.1
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year. The footnote to the original table reads: "Reserve figures exhibited in [this table] and the chart on effective reserves [Chart 1] are total member bank reserves (monthly averages of daily figures) adjusted for changes in reserve requirements and for seasonal influences. No effort was made to remove the expansion potential of total reserves resulting from shifts in deposits among classes of banks and between types of deposits subject to different requirements.

"Method of computation: For May 1958–November 1959, figures used are actual member bank reserves, adjusted for seasonal influences. Monthly values of effective reserves for January 1947 through April 1958 (when reserve requirements were last changed) have been derived by (1) obtaining the ratio of average required reserves to average deposits subject to legal reserves for May 1958–April 1959; (2) multiplying actual reserves by the percentage the above ratio is of the ratio of required reserves to deposits subject to legal reserves for each specified month; and (3) adjusting the values for seasonal influences."

[b] Eleven months

Source: FOMC (1959, 882)

reiterated his distrust of allowing the behavior of the monetary aggregates to guide policy: "[O]ne should not go overboard on the money supply question unless he was certain that the velocity factor was not playing a part. . . . For this reason he was less wary of restraint" (876). Martin's position, which also was the Committee's consensus view, meant that the FOMC continued to restrain money and credit growth while real economic growth was showing signs of deteriorating.

At this meeting Bryan asserted that the relevant issue had become not whether the Fed should maintain its current policy but how damaging to the economy the Fed's policy of tightening actually would be. Bryan used two analytical tools to drive home his point. One was a set of tables showing the annual growth rates of effective reserves, real GNP, and inflation over the period from 1947 through 1959. These tables are reproduced here as Tables 1–3. The deterioration of reserve growth is evident in Table 1. After increasing at an annual rate of more than 6 percent in 1958, effect reserves were growing only at about a 2 percent rate in 1959. Bryan used the tables to illustrate the connection between fluctuations in reserve growth and economic activity. For example, it is evident from Table 1 that reserve growth in 1956 and 1957—about 1 percent and 2 percent, respectively—was well below the trend rate of growth of about 3 percent. These slow rates of reserve growth precede the downturn in 1958, when real GNP decreased at an annual rate of 2.3 percent (Table 2). Bryan also used the tables to establish the connection between the longer-term

14. The following draws on Hetzel (1996). See also the related discussions in Meltzer (1991), Schwartz (1997), and Wheelock (1997).

TABLE 2 Compounded Annual Growth of the U.S. Economy[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	3.9	1.8	4.1	4.9	4.6	4.6	3.7	4.2	4.0	3.8	3.2	3.5
1948	x	x	–0.2	4.1	5.2	4.8	4.7	3.6	4.3	4.0	3.7	3.1	3.4
1949	x	x	x	8.7	8.1	6.5	6.0	4.4	5.0	4.6	4.3	3.5	3.8
1950	x	x	x	x	7.5	5.4	5.1	3.4	4.3	3.9	3.6	2.9	3.3
1951	x	x	x	x	x	3.4	3.9	2.0	3.5	3.2	3.0	2.2	2.8
1952	x	x	x	x	x	x	4.4	1.3	3.6	3.2	2.9	2.0	2.7
1953	x	x	x	x	x	x	x	–1.7	3.1	2.8	2.6	1.6	2.4
1954	x	x	x	x	x	x	x	x	8.2	5.1	4.0	2.4	3.2
1955	x	x	x	x	x	x	x	x	x	2.1	2.0	0.5	2.0
1956	x	x	x	x	x	x	x	x	x	x	1.8	–0.2	2.0
1957	x	x	x	x	x	x	x	x	x	x	x	–2.3	2.1
1958	x	x	x	x	x	x	x	x	x	x	x	x	6.7
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year, of GNP in 1954 dollars

[b] Three quarters

Source: FOMC (1959, 883)

movements in reserves and inflation, although the focus at this meeting was on the impending downturn.

The other tool was a chart showing the level of total effective reserves plotted around their postwar trend. Chart 1 reproduces Bryan's chart. Bryan used this chart for two purposes. One was to illustrate the relation between effective reserve growth and real economic activity, hence the appearance of National Bureau of Economic Research (NBER)–designated recessions. Bryan argued, on the basis of the growth triangles and the chart, that "a situation appeared to be approaching in which the matter of the growth factor in reserves should have serious consideration" (FOMC 1959, 871).

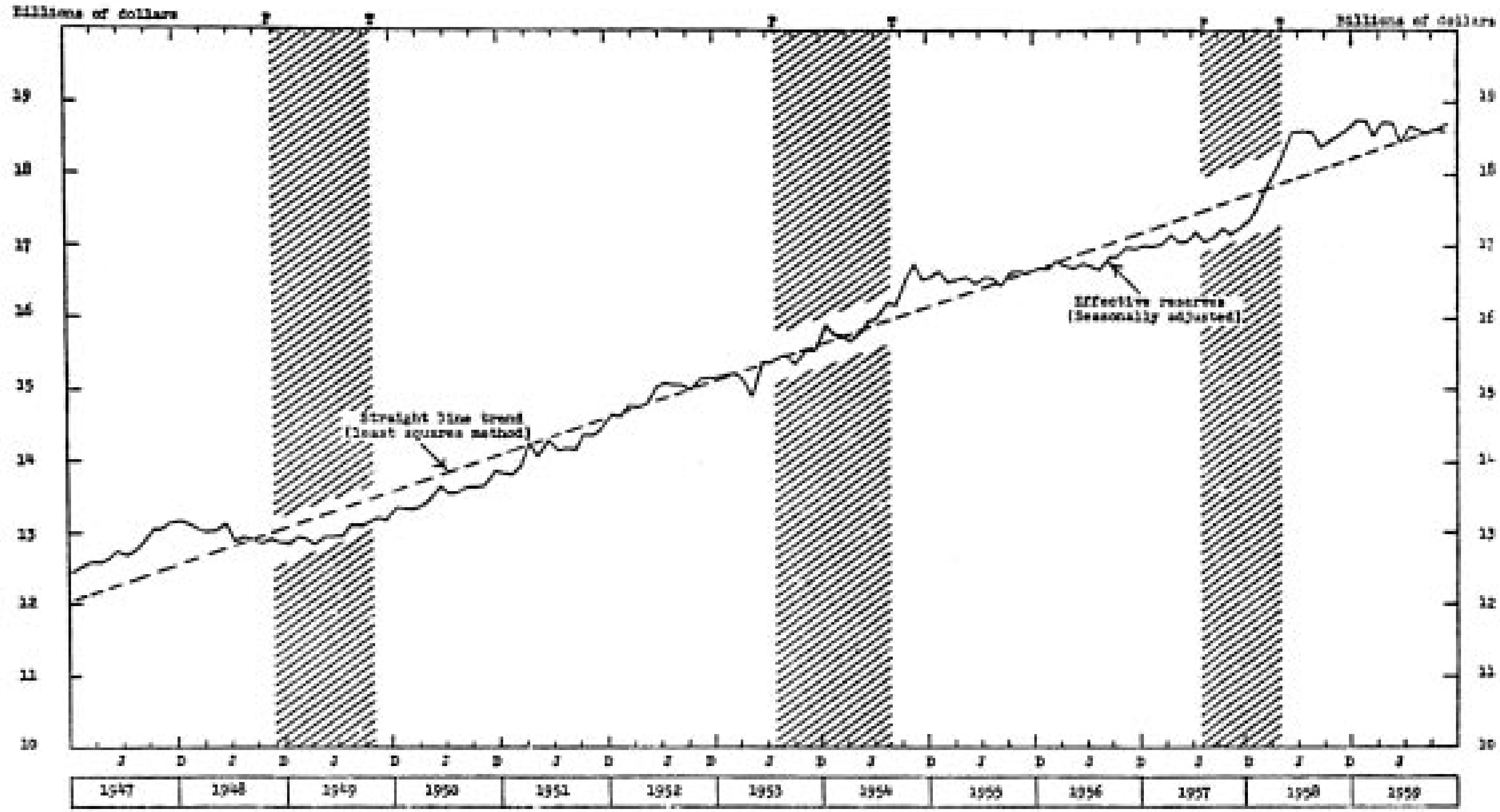
Chart 1 also served another purpose for Bryan. He used it to reiterate his misgivings about the qualitative nature of the directive. Bryan proposed that a quantitative tool like Chart 1 gave the FOMC "a means by which instruction can be given in quantitative rather than qualitative terms" (FOMC 1959, 872). By introducing this simple chart Bryan sought to transform the policy debate from money market conditions to the behavior of the monetary aggregates. This chart not only enabled Bryan to provide an explicit, quantitative analysis of past policies but also to illustrate their effects on the economy.

The next section turns to the events of 1960 as Bryan used his new framework to analyze policy and attempted to convince others of its merits. This appraisal details Bryan's introduction of short-term, aggregate growth cones for setting monetary policy as a means of achieving his longer-term targets and the opposition he faced within the FOMC.

Introducing Aggregate Growth Cones: 1960

The January 12, 1960, FOMC meeting opened with routine reports by staff economists. Guy Noyes, the Board's director of research and statistics, observed that the "customary measures of current economic activity [for example, construction activity, industrial production, and GNP] are all up, and further increases seem as certain for the near term as anything can be" (FOMC 1960, 6). As far as any economist in or out of the Federal Reserve System could predict, this view was correct. (Although not known at the time, real GNP increased at a 7.2 percent rate in the first quarter of 1960.) Robust real output growth seemed assured. Still, some FOMC members were warning that past policies, which left the growth rate of the money supply trending down over the second half of 1959, would exert significant

CHART 1 Trend Growth of Effective Reserves



Note: Shaded areas indicate recessions, according to reference dates of National Bureau of Economic Research. P = Peak; T = Trough.
Last month plotted: November 1959 Trend line exhibits an annual growth of 3.6 percent per year.

Source: Reprinted from FOMC (1959, 881)

TABLE 3 Compounded Annual Growth Rates of Price Inflation[a]

Base Year	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959[b]
1947	x	7.6	3.2	2.5	3.8	3.5	3.1	2.7	2.3	2.2	2.3	2.4	2.2
1948	x	x	–1.0	0.0	2.6	2.5	2.2	2.2	1.6	1.5	1.8	1.9	1.7
1949	x	x	x	1.0	4.4	3.7	3.0	2.4	2.0	1.9	2.1	2.2	2.0
1950	x	x	x	x	8.0	5.1	3.6	2.8	2.2	2.1	2.3	2.3	2.1
1951	x	x	x	x	x	2.3	1.5	1.1	0.8	0.9	1.3	1.5	1.4
1952	x	x	x	x	x	x	0.8	0.6	0.3	0.6	1.2	1.4	1.3
1953	x	x	x	x	x	x	x	0.3	0.0	0.5	1.2	1.5	1.4
1954	x	x	x	x	x	x	x	x	–0.3	0.6	1.5	1.8	1.6
1955	x	x	x	x	x	x	x	x	x	1.5	2.5	2.6	2.1
1956	x	x	x	x	x	x	x	x	x	x	3.4	3.1	2.3
1957	x	x	x	x	x	x	x	x	x	x	x	2.7	1.7
1958	x	x	x	x	x	x	x	x	x	x	x	x	0.7
1959	x	x	x	x	x	x	x	x	x	x	x	x	x

[a] Percentage changes, base year to terminal year, in consumer price index

[b] Ten months

Source: FOMC (1959, 884)

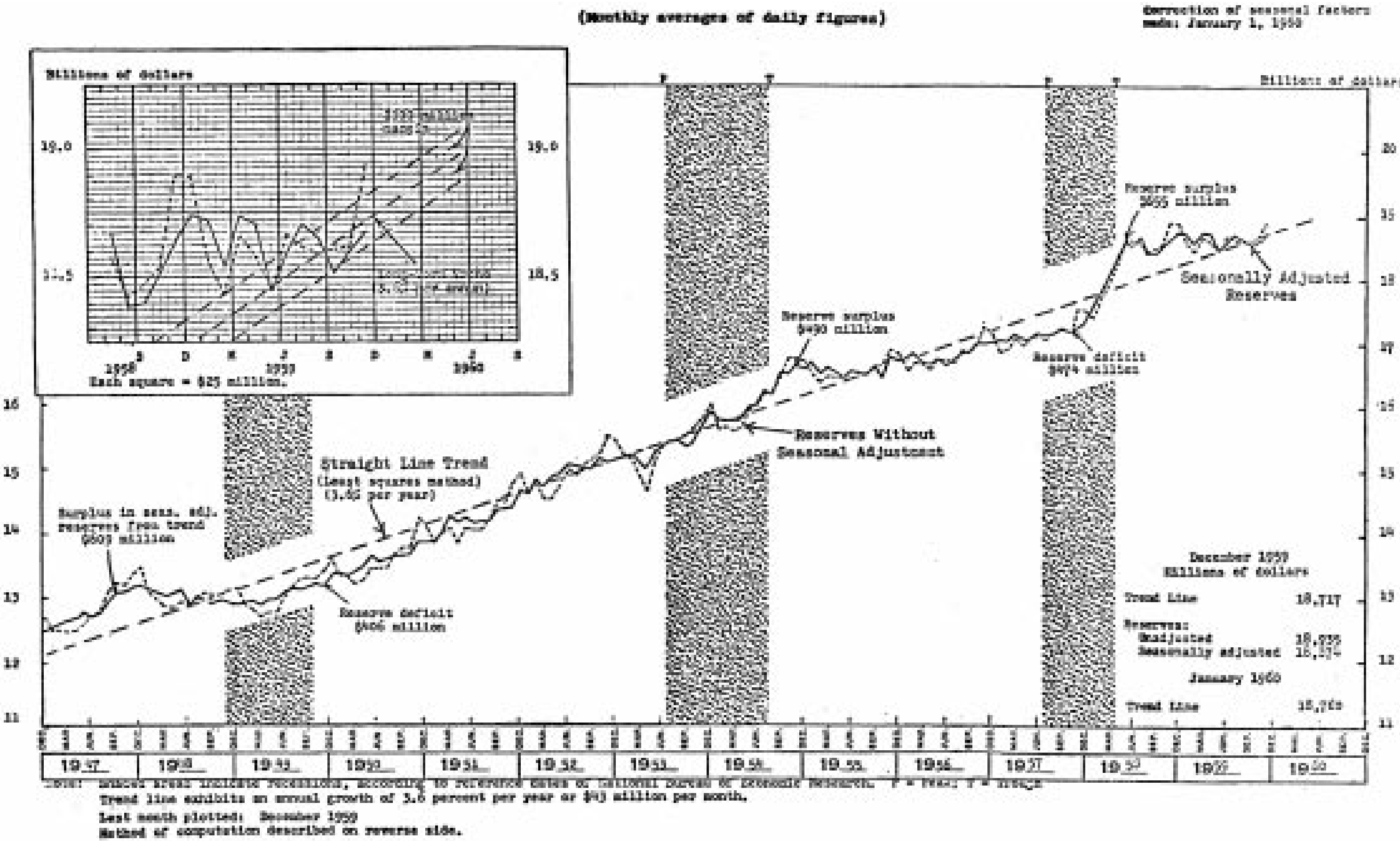
downward pressure on economic growth. If the expansion were to continue, it was argued, action to ease needed to be taken immediately. Even Noyes recognized the potential for a slowing in real output growth as he puzzled over the fact that the "high and growing rate of economic activity [stands] in interesting contrast to wholesale prices and the money supply, both of which are substantially unchanged from year-ago levels" (7).

Discussion at this meeting revealed uncertainty among FOMC members over the economic impact of recent policy actions and what direction policy should take. For example, Governor J.L. Robertson thought that "at this particular stage of the business cycle, it [is] incumbent upon the system to maintain a restrictive policy" (FOMC 1960, 30). Johns of St. Louis agreed that continued restraint was needed "to avoid inflationary developments" (16). New York Federal Reserve Bank President Alfred Hayes offered the view that the current level of restraint was appropriate in the face of the apparent economic expansion.

Bryan's evaluation of the situation separates him from his colleagues, even those who also had pressed for quantitative targets. Bryan presented the FOMC with an updated version of his chart of total effective reserves around its trend, reproduced here as Chart 2. Chart 2 clearly is more complex than Chart 1. In Chart 2, for example, Bryan details the level of reserves relative to the trend as "surplus" or "deficit." Chart 2 makes it clear that Bryan's placement of these notes is not random: deficits appear before recessions (marked off in shaded bars using NBER dating), and surpluses follow. This arrangement likely reflects the impact of the recent findings of other monetary economists on Bryan's thinking. Another item of interest in Chart 2 is that the inset box showing the behavior of effective reserves since September 1958 illustrates the restrictiveness of policy over all of 1959. This point was the basis for Bryan's policy position at this meeting. He judged the restrictive policies of 1958–59 as a necessary "mopping-up" operation to get reserves back on trend and dampen any inflationary pressure that might have built up. But enough was enough: the "justifiable mopping-up operation seems to me to be completed" (FOMC 1960, 97). The FOMC, Bryan suggested, must focus on policies to increase total effective reserves or run the risk of inducing another recession.

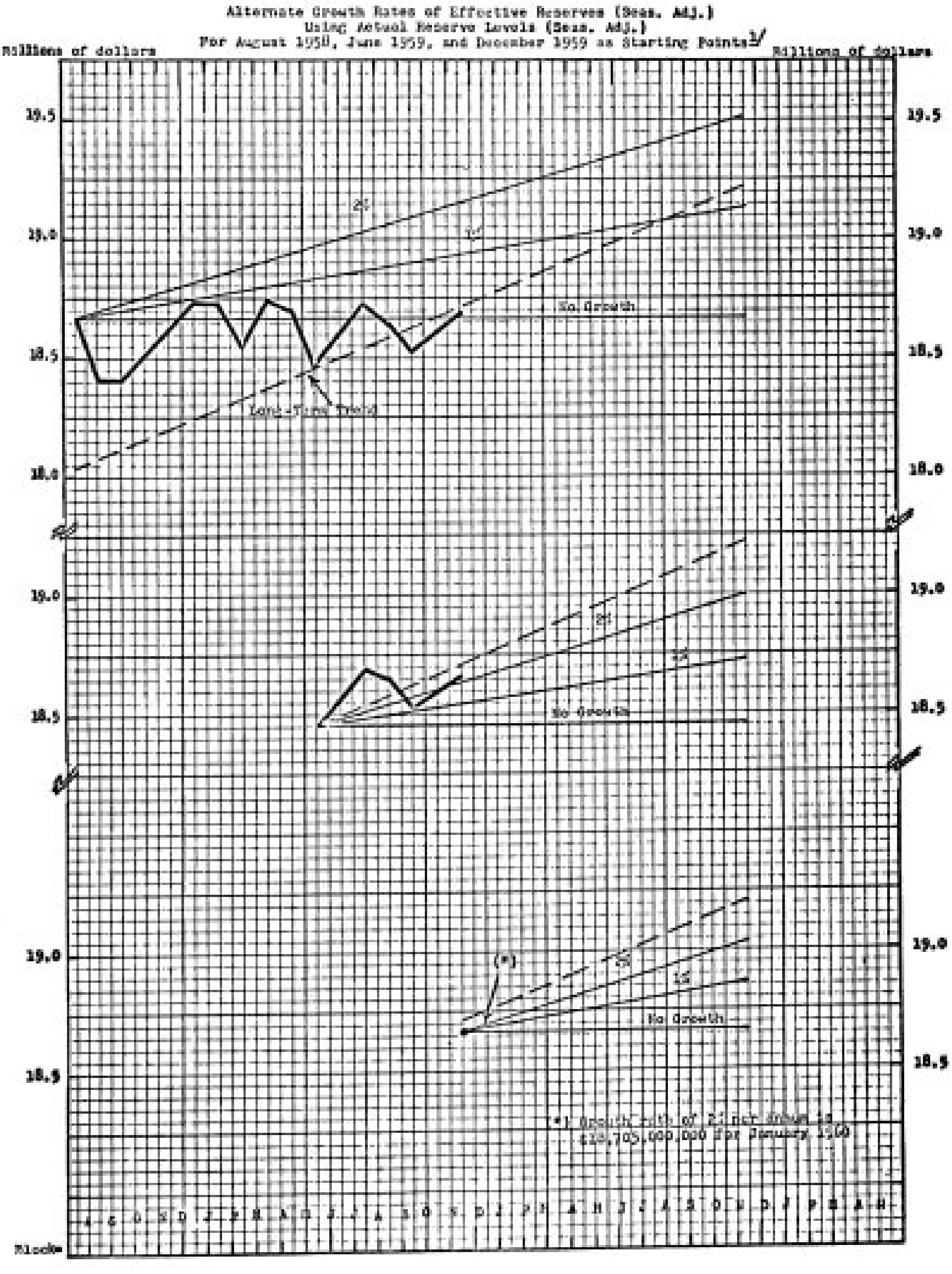
Meigs (1976) argues that even if the FOMC could have agreed on a target rate of money growth, its members would not have known how to accomplish it. Most

CHART 2 Effective Reserves



Source: Reprinted from FOMC (1960, 108)

CHART 3 Short-Run Growth Cones



Source: Reprinted from FOMC (1960, 109)

members of the FOMC believed that the Fed simply did not possess the tools to effectively influence the behavior of the monetary aggregates. At the January 12 meeting Bryan met this challenge by introducing his most novel contribution, short-run growth targets for effective reserves. This innovation squarely addressed the issue of how the FOMC could assess the desk's success in hitting short-run reserve targets. It also marks the first time aggregate growth cones were used in FOMC policy deliberations.

Bryan argued that his reserve growth cones, reproduced as Chart 3, allowed the FOMC to track the behavior of reserves relative to the path stated in the directive. Three different base periods appear in the original version. The choice of August 1958 and June 1959 was based on the fact that in those months the FOMC changed its directive. The August 1958 base period also included the postwar trend growth of reserves as a benchmark. December 1959 simply represented the last full month for which data were available.[15] Bryan recognized that requiring the desk to hit a specific target for reserves would be futile. Thus his proposal was to target total reserves to fall within a range of permissible values, providing the desk with some flexibility "to conditions in the money market as they develop" (FOMC 1960, 97).

These charts and accompanying growth rate tables gave Bryan the quantifiable foundation to argue that the FOMC could in fact assess the success or failure of the desk in meeting its policy objectives as stated in the directive. Bryan's introduction of reserve growth cones served to reinforce the attack on the language and intent of the directive. Writing the directive in terms of a reserve target, the FOMC could easily "avoid qualitative terminology as represented by such indefinable terms as tone, feel, ease, tightness, and so on" (FOMC 1960, 98). The record of this meeting indicates that Chairman Martin responded to Bryan's analysis with little discussion, merely noting that Bryan's suggestion and materials "would be taken under study" (98).

A lengthy discussion about the nature of the directive and operating procedures occurred again at the February 9, 1960, meeting of the FOMC. Bryan used his reserve growth cones to propose a February target range of $18,535 million to $18,635 million for total effective reserves. As shown in Table 4, the midpoint of this range put reserves at a level slightly lower than the postwar trend. Bryan felt that this target level provided a needed increase in reserves while maintaining some degree of restraint: policy would provide some monetary stimulus to the economy and avoid any inflationary effect. Even though the midpoint of Bryan's target indicated some restraint, it was appreciably less than what occurred: actual effective reserves for February was $18,203 million, more than $300 million below the lower bound of Bryan's target range.

In Chairman Martin's summary of the discussion at this meeting, he dismissed the usefulness of Bryan's quantitative targets. Martin maintained that such a strategy was too simplistic, a mechanistic approach that would not be wise given the variable nature of the financial markets that must take precedence in policy deliberations and action. Johns defended Bryan's experiment and hoped that "the Committee would not permit proposals such as those advanced . . . to be laughed out of court by attaching a 'mechanistic' approach label to them [and that] such proposals were worthy of serious study" (FOMC 1960, 167). The minutes reveal little additional support, however.

Bryan's introduction of short-run aggregate growth targets—growth cones—stands out as a significant and innovative development in monetary policy analysis.

Reserve and money growth continued to decline throughout the early part of 1960. A major reason for the decline was the Fed's dilemma in trying to balance internal and external objectives. The problems stemming from the persistent balance-of-payments deficit called for one type of policy (raising interest rates) while mounting evidence of a domestic slowdown in real output growth called for another (easing). In contrast to the ebullient outlook just a month earlier, the staff report on economic conditions at the February 9, 1960, meeting now discounted the expected rebound in real growth following the end of the steel strike in November 1959. Economic indicators suggested a softening in the economy. Chairman Martin, however, believed that an economic downturn was not likely, even though "he put the possibility forward . . . as an intellectual exercise" (FOMC 1960, 149). By March, however, a recession seemed probable to most observers. For example, Arthur Burns, the chairman of the Council of Economic Advisors and a future Federal Reserve Chairman, was advising President Nixon that a recession was likely (Nixon 1962, 124–28, cited in Meigs 1976). Real output, it turned out, declined at an annual rate of 1.6 percent in the second quarter of 1960.

15. The published record does not indicate whether Bryan preferred one base period over another. The fact that Bryan shows little affinity for selecting one base from which to measure changes in reserves causes base drift. For a discussion of this issue and how it influenced monetary policy in the 1970s, see Broaddus and Goodfriend (1984).

TABLE 4 Effective Reserves in 1960: Bryan's Target Level, Trend Level, and Actual[a]

Month	Bryan Target[b]	Trend	Actual
January	$18,700	$18,748	$18,863
February	18,585	18,790	18,203
March	18,653	18,833	18,027
April	NA	18,876	18,101
May	18,240	18,918	18,236
June	18,419	18,961	18,289
July	18,449	19,004	18,515
August	18,658	19,047	18,499
September	NA	19,089	18,566
October	NA	19,132	18,723
November	18,950	19,175	18,973
December	19,450	19,218	19,270

[a] Figures reported are in millions. The "target" values are midpoints of the ranges specified by Bryan. "Trend" values are determined by assuming a constant 3.6 percent increase in effective reserves. "Actual" effective reserves are based on the first reported values for the month listed.

[b] Bryan did not report a target value for April. Lack of values for September and October reflect his absence from the FOMC during the relevant meetings.

Source: Compiled by the author from information in FOMC 1960, various meetings.

There persisted a view among most of the FOMC members that the present course of policy and procedures was appropriate. This opinion is exemplified by statements of New York Bank President and Vice Chairman Hayes. Hayes was an ardent supporter of the status quo and an outspoken opponent of Bryan's, or any other, aggregates-based strategy. He considered the daily conference calls, the system of reports, and the frequency of information about money market conditions to be "so extensive that each member [of the FOMC] has ample opportunity to inform the manager if he sees any deviation from the committee's instructions." Arguing that the FOMC "would be giving up a highly advantageous technique, developed over many years, if we were to attempt to couch the instructions in some very exact mathematical terms" (FOMC 1960, 211), Hayes praised the use of net borrowed reserves in conjunction with money market conditions as the best operating guides for policy. Even with evidence of a slowdown in economic activity mounting, Hayes saw no need to alter policy. As he expressed at the March 1 meeting, there simply was "no evidence to suggest that 1960 will be other than a prosperous year, with an upward trend in the economy through most of the year" (210). Unfortunately, Hayes's outlook was shared by the chairman and a majority of the FOMC members.

Reserves in both February and March fell well below the midpoints of Bryan's target levels and below the postwar trend (see Table 4). This development led Governor A.L. Mills to warn the FOMC at the April 12 meeting that, in keeping with Bryan's views, the decline in reserves already was a serious threat to economic growth which, "if not arrested, would in due course lead to serious financial and economic consequences" (FOMC 1960, 347). Why did the FOMC allow the decline in reserves to persist? Bryan focused blame on the committee's continued reliance on free reserves and money market conditions as operating guides. He believed that these measures gave deceptive signals about the degree of ease or restraint of policy actions. If the FOMC continued down this path, Bryan foresaw "trouble ahead that would be hard to explain" (354).

Martin's consensus view at the April 12 meeting reflects some movement to ease. He stated that the FOMC "should move in the direction of slightly easing the picture as far as reserves were concerned, but with great care on the part of the desk not to do this in an *overt* way" (FOMC 1960, 363; emphasis added). This concession appears as a slightly altered directive from the May 24 meeting: Clause (b) was changed from calling for the desk to "restrain inflationary credit expansion in order to foster sustainable economic growth and expanding employment opportunities" to fostering "sustainable growth in economic activity and employment by providing reserves needed for moderate bank credit expansion" (488). Chairman Martin made it clear, however, that the desk could move to increase reserves only if it did not cause a pronounced change in short-term interest rates. This position again reflects the clash between internal and external policy objectives. While easing reserves would help alleviate the slowing in domestic economic growth, the attendant decline in rates could exacerbate the external imbalances that were becoming politically intolerable.

Table 4 shows that reserves increased slightly following this meeting. Bryan's opinion throughout the summer of 1960 was that policy should aim at increasing the level of effective reserves at a faster pace in order to put them back to the December 1959 level. By June it was clear to all that the economy was in a recession. Bryan urged the FOMC not to "push the panic button" but to undertake immediate actions to reverse the disastrous effects of previous policy. Putting domestic concerns ahead of any external problems, Bryan asserted at the July 6 meeting that it was important to increase effective reserves at a rate that, after seasonal adjustments, would meet the secular needs of the economy without raising inflation. Such a policy "is necessary because of the economic situation and because of the lagged effects of monetary policy" (FOMC 1960, 564). He also noted that his approach would improve economic growth and not "drive short-rates to the ridiculous and obviously unsustainable low levels that have characterized other easing cycles of monetary policy" (565).

Bryan's strategy for getting reserves back on track was based on two key considerations. One was the possible arousal of inflationary expectations, an overarching factor in all of Bryan's policy prescriptions. The other factor was the negative repercussion on the money market if the Fed were to increase reserves immediately and massively.[16] D.C. Johns, however, argued that interest rate movements were of secondary importance. At the July 6 meeting Johns questioned "the System's taking a position of deliberately dampening the downward adjustment of market rates of interest in a period of slack economic growth" (FOMC 1960, 567). Internal and external objectives again gave rise to divergent policy choices.

In light of the deteriorating domestic economy, the FOMC directed the desk to undertake operations that would achieve some ease in reserves by summer's end. Even though the FOMC had decided in favor of easing the reserve position of the banking system, total effective reserves in August and September of 1960 still remained far below their December 1959 level. At the September 13 meeting Bryan pointed out that the August level of effective reserves was lower than in August of the previous year despite the committee's decision to ease. His view was that "the economic situation was deteriorating" and that "[i]n these circumstances, he disagreed with the view of Mr. Hayes that no further monetary ease was required or would be appropriate" (FOMC 1960, 709). The failure of the desk to achieve the desired growth in reserves once again demonstrated the temerity of using free reserves together with money market conditions as operating guides.

Chairman Martin maintained throughout the summer and fall of 1960 that tone and feel was the best policy guide, especially in times of economic uncertainty. At the September 13 meeting he stated that "the System must not let itself be persuaded that if it had expanded the money supply exactly the right amount on a statistical basis, there would not have been any recessions . . . [I]f it got to that point, the only thing necessary would be to keep the levers moving ad infinitum" (FOMC 1960, 736). At the October 4 meeting, tone and feel prevailed over the behavior of any aggregate measure in setting policy. As Robert Rouse, the manager of the open market account, reported, "[A]s the Committee has instructed, we have been operating primarily on the feel of the market rather than on the basis of reserve statistics" (745).

Contrary to the majority opinion of the committee, Bryan advocated a policy of increasing the supply of adjusted reserves in order to provide some monetary stimulation to an economy that was in recession. (Real output decreased at a 0.4 percent rate in the third quarter and at a 3.1 percent rate in the fourth quarter.) As shown in Table 4, Bryan's short-run targets for reserves in November and December 1960 would have put reserves slightly above their postwar trend by the end of 1960. Getting effective reserves back on trend and countering the disastrous policy actions taken during the past year became the principal policy objective for Bryan. The majority of the FOMC took a very different view, however. Chairman Martin, for example, suggested that any overly

16. *The notion that any attempt to make up the shortfall in one action could disrupt the market in undesirable ways was used during the 1979–82 period of monetary aggregate targeting. During that period, intermeeting deviations of the aggregates from targets were reduced gradually in order to prevent undue gyrations in interest rates. For an appraisal of policy actions during the 1979–82 period, see, among others, Hetzel (1982) and Poole (1982).*

expansionary monetary policy, an umbrella under which he placed Bryan's suggestions, would be imprudent. In fact, Martin "could not get very pessimistic about the domestic picture . . . [and] continued to feel that the biggest shadow was cast by the balance-of-payments problem" (FOMC 1960, 834). Policy, Martin averred, "would have to be careful that it did not feed fuel to the fires of pessimism by appearing to embark on a cheap money policy" (834). In contrast to Bryan's warnings, at the December 15 meeting Martin voiced the opinion that there had been too much ease recently. He felt that such an easy money policy would do little to affect domestic economic activity and would only exacerbate the critical problem now confronting monetary policy: the persistent balance-of-payments deficit.[17]

Bryan's attempt to convince the FOMC to adopt his short-run reserve growth targets effectively ended with this meeting. Although he never abandoned his conviction that the behavior of monetary aggregates was vital in setting monetary policy, the charts and approach he favored in 1960 would not reappear during his time left on the FOMC.[18] Bryan's experiment with short-term aggregate growth cones swayed a few members of the FOMC. Statements by Johns suggest that he came to appreciate the economic impact that short-run fluctuations in money growth could have, reflecting Bryan's influence and the research of his own staff. Meigs (1976, 450) suggests that Balderston and Mills also were sympathetic. In the final analysis, however, the majority of the FOMC remained unconvinced, relying instead on the dubious tradition of using free reserves and market conditions as operating guides.

Conclusion

Bryan's pioneering development and use of aggregate growth targets as a policy alternative to free reserves and money market conditions provides an instructive case study in the early development of postwar U.S. monetary policy. The minutes of the FOMC meetings reveal that money market conditions were of uppermost concern for policy. Bryan's proposals to replace money market activity as the policy operating guide faced a hostile reception in meetings of the FOMC, just as proposals to use monetary growth targets would a decade later. His campaign to change policy challenged not only the convention of maintaining orderly domestic financial markets but also the politically pressing charge to maintain external balance.

Bryan's aggregate-based approach to monetary policy was a dramatic departure for a member of the FOMC at the time. He took new and controversial research results coming out of monetary economics and tried to implement them. Bryan's contributions went beyond merely adopting others' ideas, however. His introduction of short-run aggregate growth targets—growth cones—stands out as a significant and innovative development in monetary policy analysis. Even though his targets and procedures were not adopted by the FOMC at the time, his strategy for monetary policy would resurface as the inflation produced by the policies against which Bryan fought became unacceptable.

The role of monetary aggregates in the formation of policy is as limited today as it was forty years ago. Interest rates remain the primary instrument by which the Fed carries out policy. What then is Bryan's legacy? On an individual basis, Bryan's singular contribution is the development and use of monetary aggregate growth targets—the cones. Although recent events have once again pushed the money supply to secondary importance in its discussions, each year the FOMC must state its annual targets for the aggregates. And discussions at the FOMC during the past year indicate that some members, as in Bryan's time, remain concerned about the behavior of the money supply and its potential effects on economic activity and inflation.

Bryan's contribution also should be considered within a larger perspective. Bryan, like some other bank presidents, pursued a research agenda that resulted in policy prescriptions quite different from that of the Board.[19] His willingness to advocate a controversial view within the FOMC promoted an airing of diverse views and concerns over monetary policy. In his own way, Bryan helped to foster an environment in which alternative theories and approaches to economic analysis could be used for improving monetary policy.

17. *The FOMC turned its attention increasingly toward the external balance-of-payments problem as the 1960s unfolded. Bryan believed that monetary policy was not responsible for the problem and could do little to correct it. As he stated at the October 24, 1961, meeting, "for the System to try to correct the balance of payments situation by monetary manipulation [of the Treasury bill rate] struck him as not only absurd but dangerous" (FOMC 1961, 892). For a discussion of how these external events influenced monetary policy during the early 1960s, see, among others, Hetzel (1996), Meltzer (1991), and Schwartz (1997).*

18. *Bryan's distrust of free reserves deepened over time. By April 1963, for instance, he recognized that the "maintenance of a constant level of free reserves would permit indefinite expansion of the money supply and the financing of inflation" (FOMC 1963, 343). In September 1963 Bryan observed that the "free reserve figure might be a rather dangerous one to use for target purposes, since maintaining free reserves at any selected level would mean supplying all of the reserves demanded" (839). By January 7, 1964, he admonished the committee that policy "had been injecting reserves into the banking system at a rate . . . greater than sustainable in the long run without inflation" (FOMC 1964, 46). The inflationary record of the late 1960s proved his warning to be all too correct.*

19. *For a related discussion of this issue, see Wheelock (1998).*

REFERENCES

ATKINSON, THOMAS R. 1969. "Tone and Feel of the Market as a Guide for Monetary Management." In *Targets and Indicators of Monetary Policy*, edited by Karl Brunner. San Francisco: Chandler Publishing.

BROADDUS, ALFRED, AND MARVIN GOODFRIEND. 1984. "Base Drift and the Longer-Run Growth of M1: Experience from a Decade of Monetary Targeting." Federal Reserve Bank of Richmond *Economic Review* (November/December): 3–14.

BRUNNER, KARL AND ALLAN H. MELTZER. 1964. *The Federal Reserve's Attachment to the Free Reserves Concept*. Staff Analysis, House Committee on Banking and Currency, 88th Cong., 2d sess.

BRYAN, MALCOLM. 1938. "Limitations of Monetary Management." Speech to the National Catholic Social Action Committee, May 2, Milwaukee, Wisconsin.

———. 1948. "Some Comments on the Federal Reserve System: Operations and Policy." Speech to the Alabama Bankers Association, August 12, Tuscaloosa, Alabama.

———. 1957. "The Idea of Creeping Inflation." Speech to the Financial Section, Annual Meeting of the American Life Convention, October 11, Chicago, Illinois.

———. 1959. "The Sovereign, the Central Bank, and the Monetary Standard." Paper presented to the Regional Conference on United States Monetary Policy, March 19, cosponsored by Duke University and the American Assembly of Columbia University.

CALOMIRIS, CHARLES W., AND DAVID C. WHEELOCK. 1998. "Was the Great Depression a Watershed for American Monetary Policy?" In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, edited by Michael D. Bordo, Claudia Goldin, and Eugene N. White. Chicago: University of Chicago Press.

DAVIS, RICHARD G. 1969. "How Much Does Money Matter? A Look at Some Recent Evidence." Federal Reserve Bank of New York *Monthly Review* (June): 119–31.

EDISON, HALI J., AND JAIME MARQUEZ. 1998. "U.S. Monetary Policy and Econometric Modeling: Tales from the FOMC Transcripts, 1984–1991." *Economic Modeling* 15:411–28.

FEDERAL OPEN MARKET COMMITTEE. Various years. *Minutes of the Federal Open Market Committee*. (Bound in yearly volumes.) Washington, D.C.: Board of Governors of the Federal Reserve System.

FRIEDMAN, MILTON. 1958. "The Supply of Money and Changes in Prices and Output." In *The Relationship of Prices to Economic Stability and Growth*. Compendium of papers by panelists appearing before the Joint Economic Committee, March 31, 85th Cong., 2d sess. Joint Committee Print.

———. 1982. "Monetary Policy: Theory and Practice." *Journal of Money, Credit, and Banking* (February): 98–118.

GOODFRIEND, MARVIN. 1991. "Interest Rates and the Conduct of Monetary Policy." *Carnegie-Rochester Conference Series on Public Policy* 34 (Spring): 7–30.

HETZEL, ROBERT L. 1982. "The October 1979 Regime of Monetary Control and the Behavior of the Money Supply in 1980." *Journal of Money, Credit, and Banking* (May): 234–51.

———. 1995. "William McChesney Martin and Monetary Policy in the 1960s." Federal Reserve Bank of Richmond. Photocopy, April.

———. 1996. "Sterilized Foreign Exchange Intervention: The Fed Debate in the 1960s." Federal Reserve Bank of Richmond *Economic Quarterly* (Spring): 21–46.

MEIGS, A. JAMES. 1962. *Free Reserves and the Money Supply*. Chicago: University of Chicago Press.

———. 1976. "Campaigning for Monetary Reform: The Federal Reserve Bank of St. Louis in 1959 and 1960." *Journal of Monetary Economics* 2:439–53.

MELTZER, ALLAN H. 1991. "U.S. Policy in the Bretton Woods Era." Federal Reserve Bank of St. Louis *Review* (May/June): 54–83.

NIXON, RICHARD M. 1962. *Six Crises*. Garden City, N.Y.: Doubleday.

POOLE, WILLIAM. 1982. "Federal Reserve Operating Procedures: A Survey and Evaluation of the Historical Record Since October 1979." *Journal of Money, Credit and Banking* (November, part 2): 575–96.

ROOSA, ROBERT V. 1960. "The Change in Money and Credit, 1957–59." *Review of Economics and Statistics* 42 (August): 263–69.

SCHWARTZ, ANNA J. 1997. "From Obscurity to Notoriety: A Biography of the Exchange Stabilization Fund." *Journal of Money, Credit, and Banking* 29 (May): 135–53.

WHEELOCK, DAVID C. 1997. "Monetary Policy in the Great Depression and Beyond: The Sources of the Fed's Inflation Bias." Federal Reserve Bank of St. Louis Working Paper No. 97-011A, April.

———. 1998. "National Monetary Policy by Regional Design: The Evolving Role of the Federal Reserve Banks in Federal Reserve Policy." Paper presented at the Center for European Integration Conference "Common Money, Uncommon Regions," July, Bonn, Germany.

Economics and Crime in the States

ZSOLT BECSI

The author is a senior economist in the regional section of the Atlanta Fed's research department.
He thanks Lucy Ackert, Roberto Chang, and Tom Cunningham for insightful comments and Amanda Swift and Victoria McKenney-Johnson for research assistance.

In poll after poll, crime is the number one public worry. Crime also exacts tremendous costs not factored into official measures of well-being, and it is a favorite subject of campaign promises by politicians. In addition, crime responds to economic conditions and incentives, something economists have known at least since Gary Becker's (1968) path-breaking work on the economics of crime. Despite the substantial work in the area, particularly in the 1970s and early 1980s, the public seems largely unaware of the economics view of crime and its implications. A cynical view would point out that familiarizing the public with the research findings would give people the information to distinguish between political promises about crime control that are merely wishful thinking and promises that might have merit.

This article will first introduce the economics and crime literature by describing a very simple crime model.[1] Models for the economics of crime are easily described in a supply-and-demand framework in which criminals supply crime, the public at large demands protection from crime, and the government provides public protection. The model can be used to show how crime responds to a variety of demographic and economic factors and what results to expect from public policy proposals. Then the article introduces crime data by outlining broad trends of various crime categories in the United States and discussing potential problems associated with the data.

A large section of the article describes broad regional differences and trends in the patterns of crime and their underlying economic determinants using state data from 1971 to 1994. Specifically, the discussion looks at the determinants of total crime and the main categories of crime (property and violent crime) and their most visible and best-measured subcategories (auto theft and murder). It may be surprising that all these categories can be interpreted in an economic framework. While property crimes might be thought of as most responsive to economic conditions, many violent crimes are committed as by-products to crimes for economic gain and thus are also explainable with economics. The description compares individual states' experiences over the years from 1971 to 1994 with the 1990–94 portion of the period used to illustrate recent changes. The article presents state and regional rankings for various quality-of-life indicators as well as a discussion of simple correlations of these variables with crime data. Quality of life is a reflection of various demographic and economic variables such as unemployment rates, expenditures on police protection and police employment, welfare and education, the state population share of prisoners, and population density.[2]

Finally, the data undergo a more in-depth treatment using a panel regression approach that estimates the effects of demographic and economic variables on crime for all states over time. These regressions mirror some of the results found by others but also serve to highlight some serious issues that have been vexing the empirical literature. Generally, the demographic and economic variables explain crime rather well, and estimates for the most part conform with the economic model of crime. One important exception is that police variables are positively associated with crime or are insignificant, a finding that is common in the empirical literature. This result can be easily explained by the fact that estimates capture the response of public crime-reduction efforts in response to increases in crime rather than the independent effect of crime reduction efforts on crime. The concluding section highlights the findings by drawing a few policy implications with the Southeast as an example.

The Supply and Demand of Crime

As with all economic models, the economic model of crime assumes actors who try to make rational economic choices. The three sets of actors usually considered are the criminals, noncriminal households and legitimate businesses, and the government. In the simplest possible framework, criminals determine the supply of crime, the rest of society determines the demand for crime (protection), and the government affects both (directly on demand and indirectly through supply). This section briefly discusses how the interaction of all the actors determines the equilibrium rate of crime and how crime responds to different policies.[3]

The supply of crime is modeled as a choice between legitimate activities and work on the one hand and criminal activities on the other. The choice depends on the net payoff to crime, which is the payoff of the criminal activity itself (or loot) above all other costs associated with the crime. These costs include the forgone wages from legitimate activities, the direct costs of the crime (such as cost of supplies and so on), and the expected future penalties from the crime (including fines, incarceration, and other sanctions). The supply of crime is positively related to the net payoff to criminal activities, meaning that criminals will increase their activities when the net payoff rises (see Chart 1). The supply curve shifts to the right when the crime supplied by criminals for a given net return rises or when the return for a given level of activity falls. Examples of conditions that might cause a rightward shift of the supply curve include demographics (a higher proportion of youth), fewer employment opportunities at a given wage, and reductions in imprisonment. Education and welfare might also be thought to increase the opportunity cost of committing crime by increasing legitimate earnings.

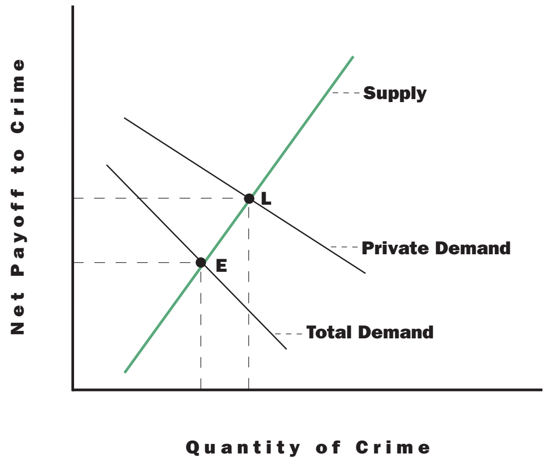
Criminals determine the supply of crime, the rest of society determines the demand for crime (protection), and the government affects both (directly on demand and indirectly through supply).

Though at first the concept of "demand for crime" may sound like an oxymoron, it can be easily explained in terms of two elements. First, there is the direct demand for (the spoils of) crime, whereby the quantity demanded falls as the loot falls, just like any other market good except that the market in this case is part of the shadow economy. Second, there is an indirect demand for crime, which is an inverse demand for protection and insurance and is also negatively related to the payoff of criminal activities. This negative relationship arises because as crime rises individuals step up private efforts at protection (ranging from locking their doors to hiring security personnel and so on), which increase the direct cost of criminal activity and therefore reduce the payoff to crime. The demand curve shifts to the left for any change to household conditions that decreases the payoff to crime for a given rate of crime. Examples include reductions in material well-being or economic growth or an increase in private vigilance.

So far, these ingredients form a market model of crime without a government. The intersection of demand and supply determines the laissez-faire equilibrium rate

1. *Several excellent surveys that capture different aspects of the economics and crime literature are referred to throughout the article. For starters, the reader might consider the articles contained in the Winter 1996* Journal of Economic Perspectives: *DiIulio (1996), Freeman (1996), and Ehrlich (1996).*
2. *Other variables were considered. However, only limited data are available aggregated to the state level with sufficient time variation.*
3. *This is a very simple model along the lines of Ehrlich (1996) and Hellman and Alper (1990). A more in-depth survey of the theoretical literature can be found in, for instance, Eide, Aasness, and Skjerpen (1994).*

CHART 1
Supply and Demand for Crime



of crime and the equilibrium return to crime, denoted point L in Chart 1, a simple graph of demand and supply curves. The laissez-faire crime rate is zero in only the most unusual of circumstances, requiring a high demand for and a low supply of protection and the willingness to pay for it. Note also that even a heinous crime such as murder, when there may be no associated material gains, can have a positive level of activity even though the equilibrium payoff may be negative. For such a crime, the intersection of demand and supply may occur in the lower right-hand quadrant rather than the usual upper right-hand quadrant. Also, the laissez-faire equilibrium will not be optimal from a social welfare perspective. For one thing, crime produces negative spillovers to other parts of the economy that are not reflected in its price and hence will be overproduced. Thus, the laissez-faire equilibrium crime rate will be greater than the socially optimal rate.[4] However, a move toward the socially optimal level of crime requires either some sort of legal or social economywide self-discipline or some sort of market intervention.

The final actor is the government. The government is assumed to be moving the equilibrium toward a (lower) crime rate that has higher social welfare. Economists assume that governments attempt to equate the marginal cost of crime to the marginal social benefit of spending additional moneys on crime prevention for all categories of crime. By this principle, effort and dollars should be targeted to activities most likely to produce results. However, the government does not operate in isolation but under bureaucratic and political constraints that can distort its effectiveness but are not usually considered in simple models. Government anticrime actions can be seen as the public component of the demand curve. The public demand for crime is also negatively sloped because as crime increases the public will respond by stepping up efforts to battle crime, ultimately making crime more costly to the criminal. In particular, the expected future cost of crime (being caught) increases so that the net return to crime falls. Thus, there is an inverse relationship between the net return to crime and the crime rate from the public demand side.

The overall demand curve for crime adds public demand to the private demand. The total demand curve lies below the private demand curve because the combined public and private efforts at crime avoidance mean that there is less crime for a given payoff to crime than if the public acted alone.[5] An exogenous increase in the expected future costs to criminals of public sanctions will shift the total demand curve to the left, as, for example, when tougher laws or harsher sentencing arrangements are enacted.

The equilibrium crime rate (and net return from crime) is determined by the intersection of the demand and supply for crime. This is point E in Chart 1, which is to the left of point L. Exogenous shocks will move the equilibrium, with the strength of the effect determined by the elasticity of supply-and-demand curves. Simply put, the curves tend to be more elastic or flatter as more substitution opportunities occur.[6] The above-mentioned examples of leftward demand shifts (due to falling average incomes or increased vigilance) and supply shifts (because of lower unemployment rates and increased education and welfare) all imply lower equilibrium crime rates. However, the response may be smaller than expected. Consider, for instance, an exogenous increase in conviction rates. Because incarcerating criminals reduces the supply of crime, equilibrium crime will fall. However, convicted criminals will be replaced by new criminals, depending in part on the strength of the deterrence effect of convictions. If the deterrence effect is weak and individuals can easily substitute into the criminal activity, it is possible that crime might not fall significantly. In other words, the more elastic the supply curve, the smaller the response of crime to a given shift in the supply curve, with no change in crime for a perfectly flat supply curve. Thus, analysis of how crime equilibria respond to policy changes must consider the shapes of the supply-and-demand curves, which may differ among crime categories.

Finally, to foreshadow an issue that is important in empirical crime analyses, while it is clear that public efforts to combat crime should reduce the crime rate, this correlation is not always clear. The difficulty is that correlations sometimes fail to distinguish between exogenous shocks and endogenous comovements. For example, more effective police efforts should to some extent reduce crime rates. Thus one would see a negative correlation between crime and police efforts. However, not all police actions are exogenous—that is, independent from crime (or predetermined). Much police activity is in response to perceived changes in criminal activity. Thus, if crime increases for a reason completely unrelated to the crime itself, police activity will increase, too, and the crime and police efforts will be positively related even though exogenous increases in police efforts reduce crime. In empirical work the problem is compounded if many shocks occur at the same time and one does not control for all of the shocks. Then it is possible to estimate correlations that are driven by shocks other than the one being studied. Ideally, to see the effect of police efforts on crime, one would like to tie down the private demand curve and the supply curve for crime so that the only thing moving when police effort changes is the public demand curve for crime. However, in reality there are little data on private efforts at crime control, so the private demand curve is not pinned down and estimated correlation will pick up private and public demand shocks.

Aggregate Trends and Data Problems

Freeman (1996) estimates that the cost of crime in the United States may have been around 4 percent of gross domestic product (GDP) in the early 1990s. He finds that about half of this cost is direct loss from crime (including direct and indirect monetary and nonmonetary losses). The second half is from resources devoted to private and public crime-control activities that could have been put to other uses had there been no crime. To put the cost-of-crime number into perspective, 4 percent of GDP is a greater share than is typically spent on motor vehicles and parts or on clothing and shoes. Thus, without crime, society could afford to have at least twice as much clothing or cars as it does and could feel safer.

One problem with the cost-of-crime estimate is that, strictly speaking, it pertains only to the 1990s and says little about how crime has evolved over time. Occasional press updates on crime trends are not always helpful because they may report information out of context. To understand movements of crime over time, it is helpful to look at crime statistics from the Federal Bureau of Investigation's (FBI's) Uniform Crime Reporting (UCR) program data that have been collected since the 1930s. The FBI collects data on seven index crimes, which can be divided into two major categories—namely, violent crimes and property crimes. Violent crimes are further divided into murder and nonnegligent manslaughter, forcible rape, robbery, and aggravated assault.[7] The category of property crimes contains burglary, larceny-theft, and motor vehicle theft (arson was added in 1978 but will not be considered here). The data are voluntarily provided to the FBI by state or local law enforcement agencies and are measured as a rate per 100,000 residents to adjust for population changes. Chart 2 graphs the overall crime index and its two main categories from 1971 to 1996.

As the chart indicates, overall crime saw dramatic growth (of roughly 50 percent) during the 1970s, peaking in 1980. Since then, overall crime has oscillated within a broad band. Property crime, which accounts for the biggest part of overall crime, has had a similar pattern over this period. However, violent crimes continued to rise over the whole sample period, with a peak in 1990. While all crime rates have fallen off in the 1990s, such movements are not unprecedented, as the experience in the early 1980s shows.

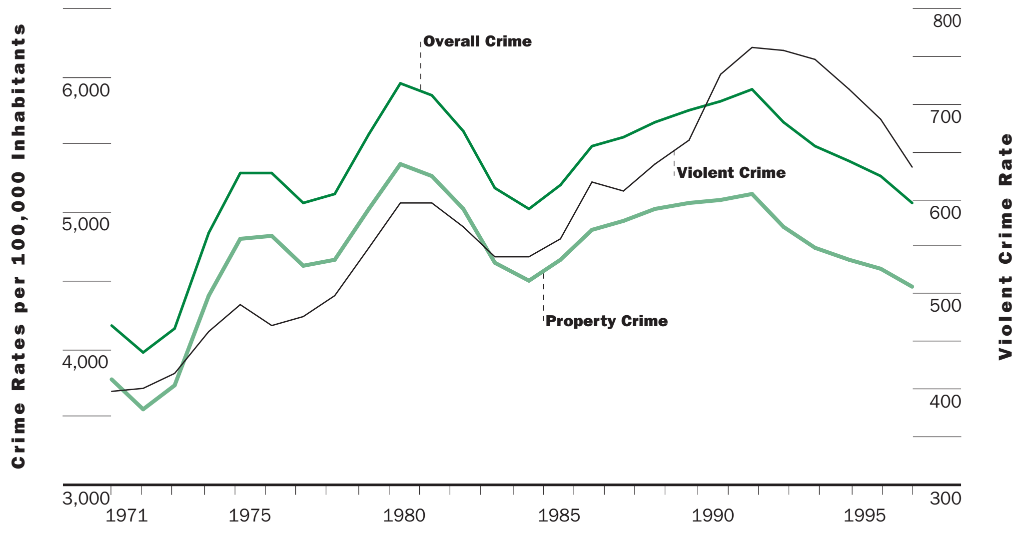
One problem with the cost-of-crime estimate is that, strictly speaking, it pertains only to the 1990s and says little about how crime has evolved over time.

Within the property crime and violent crime categories, developments of the components are shown in Charts 3 and 4. Larceny, by far the largest component of property crime, increased until 1980 and has been fairly steady for the rest of the sample period, making its biggest gains during the 1970s. By contrast, burglary, the second-largest component, has followed a relatively steady downward trend since 1980. However, the smaller auto theft component skyrocketed in the late 1980s and has been falling during the 1990s. For violent crime, the dominant category is aggravated assault, which saw a steady upward trend until the early 1990s and has seen a slight weakening since. Murder, a very visible component, has oscillated within a fairly well-defined range without any clear long-term pattern, although the most recent rates appear at the lower end of their range.

To gauge the relative cost to society, one can rank the different categories according to average cost per

4. *However, even the socially optimal crime rate may not be zero, for achieving it would involve costs that society may not want or cannot afford to pay.*
5. *Public and private efforts may be substitutes and could crowd one another out. For example, if public protection were perfect, then one would not need to install alarms or lock doors, while with private neighborhood watches the necessity for police efforts is reduced.*
6. *Thus, if crime is narrowly defined or the geographic area is small, elasticities will be larger.*
7. *See the data appendix, which defines all variables with citations of data sources.*

CHART 2 Overall Crime Index



Source: See data appendix.

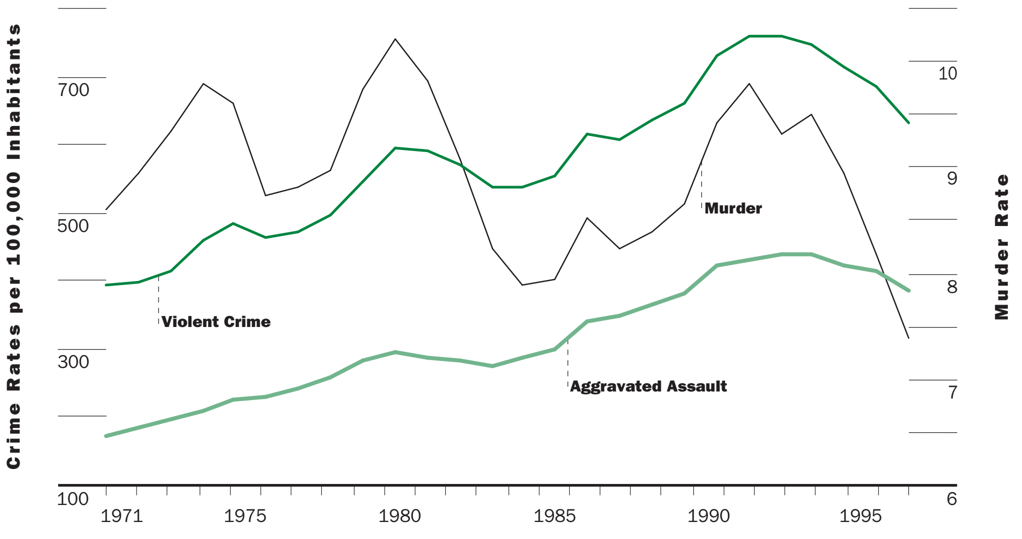
reported crime. Cohen (1988) and Miller, Cohen, and Rossman (1993) provide estimates of the monetary costs of crime (medical bills, property loss, and lost productivity) and, on the basis of jury awards, the quality-of-life reductions caused by pain and suffering. They estimate that the 1992 dollar cost for the average crime was $17,000 for murder, $1,800 for assault, $2,900 for robbery, $1,200 for burglary, $200 for larceny, and $4,000 for auto theft. Estimates of quality-of-life costs were $2.7 million for murder, $10,200 for assault, $14,900 for robbery, $400 for burglary, and $0 for larceny or for auto theft.[8] The estimates give some idea about the relative values associated with the trends in the different crime rate indexes. Thus, while auto theft and murder are the smallest subcategories in terms of crimes per 100,000 individuals, in value terms they are considerably larger. Similarly, although property crimes make up the bulk of the overall crime rate, in terms of quality-of-life costs violent crimes carry much more weight.

The above charts graph those categories most responsive to economic variables, thus not breaking out forcible rape, which is included in the index for all violent crimes. While it would seem that property crime components should be most responsive to economic incentives, it is a misconception that violent crimes are crimes of passion that are impervious to economic factors. Violent crimes are also committed for economic gain, sometimes directly and sometimes as by-products of activities committed for economic gain. Thus, all of the included index crimes are to a greater or lesser extent understandable in the terms of the simple economic model and can be related to some economic variables of interest.

Before analyzing the data, one should note that there are potentially severe data measurement problems with crime data that must be considered when making inferences. As DiIulio (1996) explains, the data problems arise from two sources. First, there is the problem of underreporting by victims. Generally, the problems arise because reporting crimes can be costly to the victims in terms of time, aggravation, and so on. Thus, crime reported to police agencies will be less than the true amount of crimes committed. Second, there is the problem of reporting by local and state law enforcement agencies, which has several aspects. Uniform Crime Reporting data capture only voluntary reports. The fact that the number of agency volunteers has risen over time suggests possible undercounting early in the sample because of underrepresentation. There is also the issue of hierarchical reporting by the FBI, which counts only the most serious crime when several crimes are committed together. Thus, less serious crimes will tend to be undercounted. Finally, there is the issue of undercounting by some local and state departments in order to show that crime has fallen, which could be a temptation around election times. While there have been increasing efforts to improve the quality of the data, particularly in recent times, measurement problems in earlier data will taint any inferences drawn from longer time series.[9]

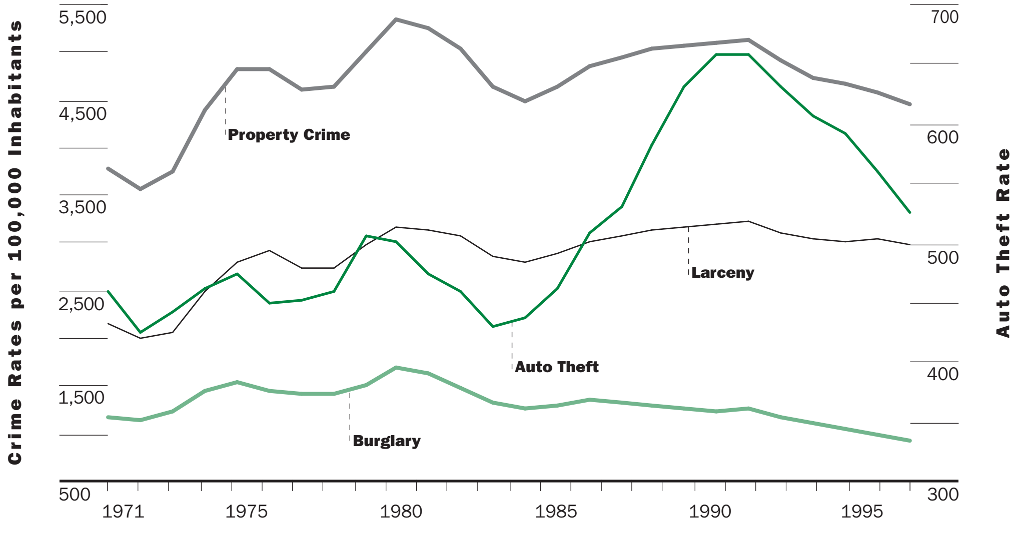
One can see how severe the measurement problem is by examining the results of a study published in 1995 (Department of Justice 1995) that extensively surveyed crime victimization in 1993. The study found that, of the total crimes committed in the United States, at least two-thirds were not reported. Generally, 42 percent of violent

CHART 4 Violent Crime Index



Source: See data appendix.

CHART 3 Property Crime Index



Source: See data appendix.

crimes were reported, compared with only a third of all property crimes. Victims were more likely to report incidents to the police if an injury resulted, forcible entry occurred, or the property loss was high. The two subcategories of murder and auto theft have lesser data measurement problems. For one, murders are generally well reported through the media and documented by funeral home records, and the role of motor vehicle insurance implies that auto thefts will also be well documented. Because of the better documentation for these categories, this study highlights these two series as opposed to the other subcategories.

8. The numbers are taken from Levitt (1996, table 7), which combines numbers from both studies.

9. DiIulio (1996) also compares Uniform Crime Reporting data with other sources of crime data such as the Department of Justice's National Crime Victimization Survey. While this survey may not suffer from the problems of the Uniform Crime Reporting, it has other problems. Note also that prior to the 1980s the two series had different time trends, but now the trends are very similar.

TABLE 1 State Differences in Crime Rates with Rankings

		Crime Rankings: 1971–94										Crime Rankings: 1990–94									
Region	State	Total Crime[a]	Rank	Property Crime[a]	Rank	Violent Crime[a]	Rank	Auto Theft[a]	Rank	Murder[a]	Rank	Total Crime[a]	Rank	Property Crime[a]	Rank	Violent Crime[a]	Rank	Auto Theft[a]	Rank	Murder[a]	Rank
New England	CT	92.2	27	95.3	28	65.1	23	115.0	41	50.4	15	88.4	23	91.6	28	67.4	22	109.7	40	61.1	22
	MA	97.5	32	97.7	32	94.2	35	202.0	50	41.5	13	88.3	22	86.3	19	101.5	38	131.6	46	40.8	13
	ME	68.4	8	73.2	9	28.9	5	37.1	5	27.9	6	61.6	7	68.2	7	17.9	4	24.2	4	19.7	3
	NH	63.4	7	68.4	7	21.6	2	45.6	13	26.1	4	55.9	4	61.8	6	17.0	3	33.5	9	22.3	5
	RI	97.4	31	101.6	34	62.1	22	160.8	49	39.7	11	83.4	20	87.6	21	55.7	17	115.6	43	42.7	15
	VT	71.5	11	77.5	13	22.3	3	38.5	7	27.8	5	67.0	8	74.8	9	15.2	2	23.4	3	23.7	6
Mideast	DE	105.1	36	107.2	38	84.7	30	82.7	31	63.4	21	91.4	26	91.3	27	89.6	32	59.1	22	52.5	18
	MD	108.4	38	104.4	36	141.8	47	105.5	38	113.2	35	108.1	40	104.8	37	130.0	44	115.1	42	127.1	45
	NJ	97.0	30	97.6	31	91.5	34	133.6	46	64.2	22	89.9	25	90.6	24	85.0	30	128.4	45	55.8	20
	NY	110.1	39	103.1	35	168.9	50	146.0	47	128.3	43	102.9	36	96.0	30	148.4	49	143.3	49	140.9	49
	PA	60.6	6	60.6	6	61.4	20	75.6	25	66.1	24	60.1	6	60.4	5	58.1	19	74.5	30	68.1	26
Great Lakes	IL	103.0	35	99.5	33	132.1	46	116.1	42	110.3	34	103.0	37	98.6	31	132.3	46	96.5	38	119.5	42
	IN	81.7	21	84.1	20	61.5	21	76.3	26	75.6	27	82.4	18	84.6	17	67.5	23	70.6	27	79.2	29
	MI	116.9	44	116.2	41	122.6	44	130.0	45	117.8	39	101.4	34	100.7	34	105.8	42	104.6	39	108.0	35
	OH	86.5	23	87.9	23	75.1	27	82.1	29	74.3	26	83.2	19	85.2	18	69.7	25	73.8	28	67.7	25
	WI	76.1	13	81.3	17	32.8	7	52.6	18	36.6	9	75.0	11	80.8	13	36.5	9	63.8	24	48.2	17
Plains	IA	72.9	12	77.5	12	34.1	9	37.6	6	23.1	2	69.7	9	74.1	8	41.1	13	27.8	7	20.3	4
	KS	88.6	24	91.9	26	60.9	19	50.8	16	59.6	20	91.6	27	95.7	29	64.6	21	52.9	18	62.0	23
	MN	80.3	19	84.9	21	42.2	11	62.7	22	28.5	7	79.2	16	84.5	16	44.5	14	55.3	19	33.3	9
	MO	59.2	5	59.6	5	58.0	17	34.6	4	139.7	47	76.9	15	80.0	11	55.9	18	50.4	16	140.9	48
	ND	49.0	2	53.7	2	10.8	1	26.6	2	13.7	1	50.2	2	56.3	2	10.4	1	22.6	2	12.0	1
	NE	71.2	9	74.6	10	43.4	13	44.7	12	36.8	10	76.7	13	81.0	14	48.1	15	42.0	13	36.8	11
	SD	51.3	3	54.3	3	26.9	4	25.1	1	23.1	3	53.4	3	57.5	3	26.4	6	17.9	1	19.1	2
Southeast	AL	77.4	15	76.2	11	86.9	33	55.2	20	133.6	45	89.7	24	87.5	20	104.7	41	55.4	20	122.4	44
	AR	71.5	10	72.1	8	66.7	24	39.9	9	103.6	32	86.5	21	87.8	22	78.0	28	50.9	17	115.8	39
	FL	143.4	50	141.8	49	156.4	49	106.9	39	130.7	44	150.0	50	148.2	50	161.6	50	135.5	47	98.5	33
	GA	96.9	29	97.3	30	94.9	38	80.9	27	146.0	48	112.8	44	115.2	46	97.6	35	95.7	36	121.1	43
	KY	58.9	4	59.3	4	54.4	16	46.0	14	92.9	30	58.8	5	58.4	4	61.2	20	34.1	10	69.5	27
	LA	98.4	33	96.3	29	116.5	42	81.1	28	164.3	50	117.0	47	114.8	45	131.6	45	96.4	37	195.5	50
	MS	91.4	25	90.9	24	94.8	37	85.6	33	104.3	33	92.3	29	91.1	26	100.0	37	84.5	33	109.9	37
	NC	82.0	22	81.7	19	85.7	31	42.3	10	113.9	36	100.9	33	102.7	36	88.9	31	46.5	14	116.9	41
	SC	93.9	28	91.1	25	116.7	43	56.2	21	127.2	42	106.4	39	102.3	35	133.6	47	57.9	21	112.3	38
	TN	79.5	18	78.9	14	84.4	29	82.6	30	114.0	37	91.9	28	90.8	25	98.6	36	92.7	35	109.3	36
	VA	78.1	17	80.6	15	58.5	18	50.2	15	97.1	31	76.1	12	80.2	12	49.3	16	48.8	15	93.5	32
	WV	42.3	1	43.9	1	29.0	6	31.6	3	65.5	23	45.5	1	48.3	1	26.9	7	26.6	6	65.0	24
Southwest	AZ	140.0	49	144.7	50	101.2	40	105.5	37	92.4	29	133.7	49	139.9	49	92.1	34	143.0	48	91.2	31
	NM	114.4	40	114.5	40	112.8	41	69.7	23	117.9	40	114.2	45	113.6	43	117.9	43	63.2	23	100.9	34
	OK	91.6	26	94.0	27	71.7	25	87.7	34	87.9	28	97.7	31	100.0	33	82.1	29	83.5	32	79.1	28
	TX	115.5	41	118.0	42	95.3	39	112.3	40	149.8	49	123.7	48	126.6	48	104.5	40	125.6	44	137.9	47
Rocky Mountains	CO	121.5	46	125.8	46	86.5	32	92.4	36	70.3	25	102.4	35	106.7	39	73.9	27	70.2	26	58.6	21
	ID	76.4	14	80.6	16	41.7	10	39.0	8	40.7	12	71.5	10	76.5	10	38.2	10	28.1	8	30.9	7
	MT	81.2	20	87.1	22	33.2	8	52.4	17	44.6	14	80.2	17	89.0	23	22.2	5	38.8	11	35.5	10
	UT	98.8	34	105.3	37	45.2	14	55.2	19	36.0	8	97.3	30	106.0	38	39.6	11	40.2	12	31.9	8
	WY	77.9	16	81.6	18	47.0	15	42.8	11	58.3	18	76.7	14	82.2	15	40.2	12	24.6	5	39.4	12
Far West	AK	106.7	37	108.2	39	94.3	36	119.9	44	117.6	38	98.3	32	99.5	32	89.8	33	82.3	31	79.9	30
	CA	130.3	47	128.6	47	142.9	48	151.2	48	123.7	41	115.8	46	111.5	41	144.1	48	162.2	50	132.4	46
	HI	115.6	42	124.4	45	43.1	12	91.0	35	58.9	19	110.6	42	122.0	47	35.2	8	66.8	25	41.8	14
	NV	134.2	48	135.2	48	125.4	45	116.8	43	139.7	46	111.5	43	112.6	42	104.2	39	110.0	41	116.0	40
	OR	118.0	45	122.3	44	83.2	28	84.1	32	54.0	17	103.9	38	109.3	40	68.7	24	88.3	34	47.9	16
	WA	116.7	43	122.0	43	72.7	26	73.9	24	53.8	16	108.7	41	114.6	44	69.7	26	74.0	29	52.9	19

[a]Percentage deviation relative to the nation

Source: See data appendix.

TABLE 2 Average Regional Differences in Crime Rankings

Region	1971–94 Total Crime	Property Crime	Violent Crime	Auto Theft	Murder	Comparison of 1990 with Full Sample Total Crime	Property Crime	Violent Crime	Auto Theft	Murder
New England	19	21	15	28	9	–5	–6	–1	–3	2
Mideast	30	29	36	37	29	–3	–5	–1	0	3
Great Lakes	27	27	29	32	27	–3	–4	0	–1	3
Plains	11	11	11	9	13	2	1	2	2	1
Southeast	21	19	31	21	38	7	8	2	4	–1
Southwest	39	40	36	34	37	4	4	0	3	–2
Rocky Mountains	26	28	16	18	15	–5	–3	–3	–6	–4
Far West	44	44	33	38	30	–3	–3	–3	–3	–2

Source: See data appendix.

These points serve as a reminder that any inferences from Uniform Crime Reporting data are inferences about reported crime and not necessarily about true crime, unless it can be shown that there exists some sort of stable relationship between the two. However, Grove, Hughes, and Geerken conclude that for crimes in which both citizens and police agree that a "serious violation of the law" has occurred, such as motor vehicle theft, robbery, burglary, and homicide, Uniform Crime Reporting crime statistics are "reasonably good approximations of the true crime rates" (1985, 489). For the less clear-cut crimes of aggravated assault and rape, they conclude that the evidence that Uniform Crime Reporting data accurately represent serious crime is somewhat weaker, and the larceny rates may overstate the actual crime rate.

A Regional Comparison of Crime and Its Determinants

This section looks at the cross section of states and the cross section's movements across time for the major crime categories. In particular, overall crime and its two major categories, property and violent crime, are examined. Because of the measurement problems discussed above, the discussion also focuses on subcategories that arguably have a smaller data measurement problem: auto theft and murder. For information about the time variation across states without going into the details of year-to-year differences, the study compares the 1990s with averages that span the whole sample. The discussion also looks at the link of the different crime categories with potential explanatory variables. After a brief summary of these quality-of-life variables, simple correlations are discussed.

Table 1 gives a picture of how crime has varied across states for the period 1971 through 1994. The table shows the percentage deviation by state relative to the nation averaged over the full sample period.[10] To make comparisons easier, states' relative crime rates over the sample period are also ranked, with states with the lowest crime rates receiving the highest ranking. Also, states are grouped into eight standard regions defined by the Bureau of Economic Analysis. Finally, to give a rough idea of how the cross section has recently evolved, percentage deviations and rankings for the states are presented for the first half of the 1990s. As an example to help interpret this data, Florida on average had a murder rate that was 30.7 percent above the nation's for the period from 1971 to 1994, earning it a ranking of 44. However, for the first half of the 1990s Florida's murder rate was 1.5 percent below the nation's, and its rank climbed to 33.

To help identify regional characteristics, Table 2 provides unweighted averages of the rankings of all the states in each region for the complete sample period. The table also gives changes in the regional rankings when comparing the first half of the 1990s with the complete sample period. Because unweighted averages give disproportionate weight to small states, use of unweighted averages is meant only to identify patterns in the data, not to summarize the experiences of whole regions. With this in mind, the Plains states tended to have the lowest crime rates across all categories, and Rocky

10. Because of the measurement problems with crime data, one can never be sure that these differences are not an artifact of reporting bias. Ehrlich (1996) suggests that reporting bias may be proportional to crimes reported, but this idea seems untestable.

TABLE 3 Average Regional Rankings of Explanatory Crime Variables

1971–94

Region	Density	15-24-Year Olds	Unemploy-ment Rate	Per Capita Personal Income	Public Welfare	Primary and Secondary Education	Convicts	Police Employment	Police Expenditures
New England	35	30	21	14	12	26	40	26	34
Mideast	46	21	27	14	23	21	20	11	17
Great Lakes	37	28	34	20	17	26	27	25	22
Plains	13	28	10	31	25	20	37	41	42
Southeast	29	21	29	38	31	34	15	31	28
Southwest	14	31	28	35	34	25	16	17	21
Rocky Mountains	7	25	17	31	37	12	36	23	20
Far West	19	24	38	10	24	30	20	17	11

Comparison of 1990 with Full Sample

Region	Density	15-24-Year Olds	Unemploy-ment Rate	Per Capita Personal Income	Public Welfare	Primary and Secondary Education	Convicts	Police Employment	Police Expenditures
New England	0	–6	11	–2	–2	–3	–5	–2	0
Mideast	0	12	1	–4	4	3	–2	0	2
Great Lakes	0	–6	–7	1	2	–2	–2	3	–2
Plains	–1	0	1	1	1	3	2	–2	–2
Southeast	0	8	0	–2	–4	–3	4	–2	2
Southwest	1	0	3	4	–5	–4	–2	–3	–5
Rocky Mountains	0	–15	–1	3	2	5	–1	3	0
Far West	1	–2	–7	2	6	2	1	5	1

Source: See data appendix.

Mountain states had moderately higher crime rates by comparison. On the other hand, states in the Far West had the highest overall and property crime rates while the Southwest and Mideast had relatively high crime rates across all categories.

During the 1990s crime increased dramatically in the Southeast as measured by the overall crime index and the property crime index. The region also saw a slight increase in violent crime from already high levels. Only murders declined slightly from generally high levels. The Southwest also saw important worsening in crime although not as severe as in the Southeast. Violent crime in the region remained unchanged at relatively high rates. All other regions experienced a reduction in overall crime rates. For the Rocky Mountain states and the states in the Far West, this reduction was balanced equally among property crime and violent crime. For the Northeast, Mideast, and Great Lakes regions the reduction in overall crime came mainly from lower property crime, but these regions also saw worsening relative murder rates. Finally, the Plains states saw a marginal worsening from comparatively low rates overall.

Next, the study turns to variables that may help explain crime variation across states. First, two demographic variables are considered—population density and the population share of the young. Population density is thought to be associated with crime primarily because crime is considered an urban phenomenon. Because much crime is committed by young (male) adults, the youth of the population should also be an important explanatory variable. As Freeman notes, the "demographics of the criminal population show that those who commit crimes consist disproportionately of persons with low legitimate earnings prospects—the young, the less educated, persons with low test scores, and so on" (1996, 33). Thus, the analysis looks at the share of the population of fifteen- to twenty-four-year-olds, which is the age group with the highest arrest rates (Marvell and Moody 1991).

Next, economic conditions are captured by two variables, the unemployment rate and per capita personal income. Both of these variables are fairly popular measures in crime analyses (Chiricos 1987; Hsieh and Pugh 1993). The unemployment rate measures reduced legitimate earnings opportunities that are particularly important for the population segment most at risk for engaging in criminal activities. In other words, increases of the unemployment rate imply diminished legitimate earnings expectations and so capture an increase in the net return from crime. By contrast, average personal income can be interpreted as a measure of general material well-being and thus of the potential loot from crime. The average is not really a measure of legitimate earnings for criminals because they tend to be at the lower end of the income distribution.[11]

Finally, five public policy variables are considered, with expenditure variables measured as a share of personal income. The first two variables, welfare expenditures and expenditures on primary and secondary education, measure positive disincentives to crime. Welfare expenditures arguably might be thought of as reducing the pain from unemployment and thus reducing the net return of crime. Alternatively, it can be interpreted as a state's propensity to help disadvantaged population segments. Education expenditures also increase the opportunity costs of crime, first by keeping youths off the streets, so to speak, raising their earnings potential from future legitimate earnings, and giving them tools to evaluate the costs of crime realistically. Finally, three variables are used to measure public disincentives to crime. The first is the state population share of prisoners or convicts, which can be thought of as a summary indicator of the penalties of being caught and the expected cost of crime. Police employment (as a share of state population) and public expenditures on police and protection each measure public efforts to reduce crime and raise the expected cost to criminals.[12]

The unemployment rate measures reduced legitimate earnings opportunities that are particularly important for the population segment most at risk for engaging in criminal activities.

State percentage deviations and rankings of these explanatory variables for the 1990–94 period are depicted in the table provided in the appendix.[13] Table 3 summarizes these data with regional rankings over the full 1971–94 sample period and compares regional rankings for the first half of the 1990s with the full sample period. Some of the regions can be roughly characterized as follows: States in the Northeast had high incomes and welfare expenditures but low prison population and police expenditures. The Far West also tended to have high incomes but relatively high unemployment rates, moderately low education expenditures, and high prison population and police expenditures and employment. The

11. A measure of legitimate earnings for criminals would be the median wage, which tends to be lower than the average but could not be obtained for this study.

12. Per capita police expenditures and police employment were qualitatively very similar.

13. See the data appendix, which defines all variables and cites all data sources.

TABLE 4 Simple Correlations of Crime Variables and Explanatory Variables (Percent)

Variables	1971–94 Total Crime	1971–94 Property Crime	1971–94 Violent Crime	1971–94 Auto Theft	1971–94 Murder	1990–94 Total Crime	1990–94 Property Crime	1990–94 Violent Crime	1990–94 Auto Theft	1990–94 Murder
Population Density	13.8	10.9	23.6	65.4	–8.4	6.5	1.5	22.7	52.5	–1.2
15-24-Year-Olds	–18.8	–18.1	–16.1	–21.8	13.9	–14.8	–14.1	–12.9	–35.3	17.9
Unemployment Rate	24.4	20.1	39.3	30.1	47.9	15.7	8.3	38.0	41.0	43.1
Per Capita Personal Income	40.6	41.5	23.3	61.2	–13.9	13.0	11.4	15.2	47.9	–14.5
Public Welfare	–15.8	–17.8	–1.6	31.0	–16.7	–33.4	–37.4	–9.8	1.9	–5.8
Primary and Secondary Education	–16.7	–14.0	–25.3	–19.6	–23.3	–29.1	–25.7	–33.4	–32.6	–24.7
Convicts	49.3	43.2	66.4	22.7	80.5	58.9	52.2	66.4	48.2	67.5
Police Employment	66.4	62.5	66.7	68.6	35.6	55.8	51.0	57.6	62.8	33.8
Police Expenditures	72.0	69.3	65.0	55.1	42.3	70.0	67.7	59.0	59.6	38.6

Source: See data appendix.

Southeast and Southwest generally had lower incomes, low welfare expenditures, and large prison populations, with the Southwest distinguishing itself by high police expenditures and employment and the Southeast by the lowest expenditures on education. A comparison of the 1990s with the overall sample period reveals that the Southeast and Southwest increased expenditures on welfare and education and increased relative police employment, with the Southeast experiencing rising personal incomes and the Southwest seeing income reductions. At the same time, the Southeast increased its prison population and police expenditures, and the Southwest did the opposite.

Finally, Table 4 ties the variation in crime rates to the set of explanatory variables. The table reports simple correlations of dependent crime variables with potential explanatory variables; that is, each correlation is considered in isolation from other variables, and no effort is made to control for the effects of other explanatory variables. Generally, the findings are that density, age, and personal income are positively correlated with all crimes except murder, for which the correlations with density tend to be relatively weak. Unemployment, the prisoner population share, and both police variables are all positively correlated with all index crime categories. Finally, education and welfare tend to be negatively correlated with all crime categories except that the welfare correlation is positive with auto theft and close to zero for violent crime. Qualitatively, the correlations for the 1990s are much the same as for the overall sample period. The exceptions are that incomes are negatively correlated with murder (but the correlation is small) and that welfare is negatively correlated with violent crime (again, the correlation is small).

Almost all the correlations are consistent with intuition. The demographic variables reinforce the view that crime is committed by youths in urban areas. Correlations of crime with variables for economic conditions support the idea that crime responds positively to the net benefit of crime through unemployment (which measures a reduction in legitimate earnings expectations) and personal income (which measures an increase in the material reward to some crimes). A case can also be made that the crime correlations with welfare and education indicate that these variables can also reduce the net benefit of crime. Even the positive correlations of the police variables and the prisoner share of the population can be explained as a reaction of public efforts to reduce crime rather than as capturing the independent effects of police and imprisonment on crime. This simultaneity bias arises because the public responds to crime and crime responds to public efforts, and simple correlations make no distinction between the two.

However, while the tables are helpful in introducing regional differences and the correlation analysis is suggestive, there are several reasons to view the correlation analysis with skepticism. First, while aggregation across all years might eliminate a great deal of noise in the data, it is also possible that much valuable information is lost and the correlations are purely a result of the time aggregation. Second, there is no way to tell which, if any, of the correlations are statistically significant. Thus, it is possible that the small correlations are statistically different from zero and the large correlations are statistically insignificant. Third, each correlation is treated in isolation without controlling for the effects of other variables.[14] Partial correlations, or correlations that take into account other variables, might be considerably different. All these reasons argue for a multivariate approach that takes time variation into account.

Regression Analysis and Survey of Empirical Studies

This section presents the results of panel regressions that relate the various index crimes to the explanatory variables introduced in the last section. Alternative approaches, such as an aggregate time series analysis (along the lines of the data presented in the previous section) or a simple time-aggregated cross-regression approach (using Table 5 and the appendix table) could have been pursued. However, in either case aggregation of the data might create biases that can easily be prevented through an analysis of panel data.

A panel regression is simply a regression with cross-sectional data that varies over time. The dependent variable (crime) is regressed against all explanatory variables for all states and all years. It is assumed that the coefficients for the explanatory variables are equal for all states, and state-specific variation is allowed through the use of dummy variables.[15] Specifically, the equation to be estimated is of the following form:[16]

$$lnY(it) = a(i) + a * T + b1 * lnX1(it) + b2 * lnX2(it) + \cdots + E(it),$$

where i denotes states and t denotes years; Y is the crime index variable; $X1$, $X2$, and so on are the explanatory variables discussed in the last section; T is a trend; and E is the error term. As in the tables, all variables are defined as ratios of the state variables relative to the corresponding U.S. aggregate for each year, or, for example, $Y(it) = y(it)/y(USt)$. Also, the data used in the regressions have been transformed using a logarithmic transformation because doing so reduces the influence of outliers and allows a simple elasticity interpretation. That is, one can explain the estimated coefficients, $b1$, $b2$, and so on, as the percentage change of Y that is associated with a 1 percent change of any particular explanatory variable.

Table 5 presents the results of the regression for the various crimes, including standard errors and significance values. The regressions have been estimated using White's (1980) formula for correcting for the possibility that the variances of the error term change over the sample. All public expenditure and police variables and the prison population have been lagged one period in order to get an exogenous representation; however, the results do not change much when these variables are not lagged.[17] Also, the estimation results from the initial specification are recorded in the first column for each crime. Overall, the results are intuitive, in some ways more so than the correlation analysis of the last section, but there are also a few surprises. Not surprisingly, property crimes do a better job of conforming with an economic interpretation than do violent crimes and murder and, because property crimes are the largest component of overall crimes, so does the overall crime index.

Generally, density, age, and personal income are positively correlated with all crimes except murder, for which the correlations with density tend to be relatively weak.

In particular, population density is generally insignificant and only significant for auto theft, where it has a positive sign, and for murder, where it has a negative sign.

14. A related point, already hinted at in discussing the police correlations, is that the correlations capture equilibrium comovements and fail to distinguish between exogenous police effort changes and endogenous responses. As the next section makes clear, this problem is only partially solved in a regression framework because, for one thing, there is no way of controlling for all variables that create demand and supply shifts. In particular, there are no (readily available) data on private protection efforts, and thus demand shocks cannot be tied down.

15. The econometric model used is commonly known as the fixed-effect model and is fairly standard in the recent crime literature that deals with multiple time series (see, for instance, Marvell and Moody 1996 and Levitt 1997). One virtue of this model is that the state dummies (or fixed-effect adjustments) help reduce biases that arise because potentially important variables may have been omitted that explain cross-state variation.

16. While earlier discussion of correlations indicates that dividing the sample into smaller time frames such as decades might be interesting, subsample estimates were generally very imprecise and so are not reported. Also, regional dummies were not included, although earlier discussion focused on regional differences, mainly to simplify comparison across states.

17. Using lagged variables is perhaps the simplest way of dealing with the simultaneity biases inherent in empirical crime analysis. One problem with this method is that it may not adequately represent dynamic interrelations in the data and in particular may miss serial correlation effects. Other methods are more structural and require identifying assumptions derived from theory to get at the issue. For instance, Levitt (1997) uses an instrumental variable approach to account for simultaneity biases arising with police variables.

TABLE 5 Panel Regression Results[a]

Variables	Total Crime		Property Crime		Violent Crime		Auto Theft		Murder	
Population Density										
Coefficient estimate	0.018		0.009		0.079		0.632		–0.245	
Standard error	–0.038		0.038		0.068		0.098		0.094	
p-value	0.64		0.81		0.238		0		0.01	
15-19-Year-Olds										
Coefficient estimate	0.64		0.696		0.275		0.522		–0.751	
Standard error	–0.07		0.07		0.119		0.222		0.171	
p–value	0		0		0.02		0.019		0	
20-24-Year-Olds										
Coefficient estimate	0.183		0.114		0.57		1.378		0.682	
Standard error	–0.069		0.071		0.12		0.184		0.21	
p-value	0.008		0.105		0		0		0.001	
Unemployment										
Coefficient estimate	0.121		0.134		–0.048		0.152		–0.08	
Standard error	–0.017		0.018		0.027		0.032		0.037	
p-value	0		0		0.078		0		0.028	
Personal income										
Coefficient estimate	0.371		0.397		0.269		0.559		0.222	
Standard error	–0.076		0.077		0.131		0.164		0.208	
p-value	0		0		0.04		0		0.285	
Welfare										
Coefficient estimate	–0.003		0.0004		0.013		–0.064		–0.105	
Standard error	–0.017		0.018		0.027		0.039		0.047	
p-value	0.874		0.98		0.62		0.097		0.024	
Primary and Secondary Education										
Coefficient estimate	0.016		0.017		–0.024		–0.021		–0.017	
Standard error	–0.028		0.029		0.044		0.056		0.074	
p-value	0.57		0.55		0.58		0.7		0.82	
Convicts										
Coefficient estimate	–0.087		–0.091		–0.046		–0.198		–0.063	
Standard error	–0.015		0.015		0.022		0.032		0.034	
p-value	0		0		0.03		0		0.065	
Police Convicts[b]										
Coefficient estimate		–1.35		–1.475		–0.95		–0.316		0.94
Standard error		0.825		0.884		0.6		0.433		0.757
p-value		0.103		0.095		0.113		0.466		0.211
Residual Convicts[c]										
Coefficient estimate		–0.087		–0.091		–0.046		–0.198		–0.063
Standard error		–0.015		0.015		0.022		0.032		0.034
p-value		0		0		0.03		0		0.065
Police Expenditures										
Coefficient estimate	0.076	0.123	0.08	0.132	0.069	0.1	0.022	0.026	–0.006	–0.044
Standard error	0.024	0.031	0.025	0.033	0.033	0.036	0.025	0.03	0.031	0.039
p-value	0.001	0	0.001	0	0.037	0.004	0.39	0.39	0.83	0.26
Police Employment										
Coefficient estimate	0.126	0	0.138	0	0.09	0	0.012	0	–0.1	0
Standard error	0.083	0	0.088	0	0.06	0	0.043	0	0.06	0
p-value	0.128	0	0.118	0	0.132	0	0.78	0	0.18	0

[a] The initial specification is in the first column of figures for a category, and the alternative specification is in the second column.
[b] Prison population explained by police expenditures and employment
[c] Prison population unaccounted for

Source: See data appendix.

All estimated elasticities are less than one, and for the most part, they are small. Specifically, a 10 percent increase of relative state population density is associated with a 6.3 percent increase in relative auto theft but a 2.45 percent reduction in murder rates. These figures suggest that auto theft is an urban phenomenon but, surprisingly, that murder is not.[18] The share of twenty- to twenty-four-year-olds is always estimated with a positive sign and, except for property crime, is highly significant. Similarly, the share of fifteen- to nineteen-year-olds is always significant at the 98 percent level and positive except for murder. Thus, the estimates reaffirm the notion that crime is associated with youth.[19] However, only for auto theft is the elasticity larger than one, meaning that a baby boom, for example, will have disproportionately large effects for this crime category.[20]

Overall, the estimated effects of the indicators of economic conditions are consistent with the simple economic model of crime, while for the most part education and welfare expenditures are statistically insignificant. The unemployment rate, which is a proxy for the opportunity cost of legitimate work, is generally significant. It is positive for overall crime, property crime, and auto theft but negative for violent crime and murder, perhaps indicating a sort of envy effect on those left out of the legitimate work market.[21] What works consistently for all types of crime (and is consistent with most studies—for example, Eide, Aasness, and Skjerpen 1994) is per capita personal income, which has a positive sign in all the regressions and is significant except for crimes of murder, for which one would expect material gain to be a lesser issue than for other crimes. Public expenditures on primary and secondary education are always insignificant, with a negative sign for most categories of crime except property crime and overall crime. On the other hand, welfare expenditures are weakly significant for auto theft and strongly significant for murder, both with a negative sign, suggesting that welfare spending may play a role in alleviating some crimes.[22]

Finally, the evidence on imprisonment rates strongly suggests that punishment works to reduce crime, with the coefficient on the population share of prisoners always negative and strongly significant except for murder, for which it is weakly significant. Thus, a 10 percent increase in the prison population is estimated to be associated with a 0.5 percent to 1.9 percent reduction in crimes.[23] However, the estimated coefficients on the two police variables are mostly positive (except for murder) and, with a few exceptions, insignificant. There are several possible explanations for this finding. One is that police employment or expenditures may not really matter for crime.[24] Alternatively, it might be that the regressions do not capture the exogenous component of police efforts very well and mostly capture the endogenous response of police activity to changes in crime. In other words, the regression might not be controlling for simultaneity bias

The United States in the 1990s has seen a dramatic fall in crime in almost all categories. However, not all regions have benefited equally.

18. *Partly, this result may be due to the fact that the data were aggregated to the state level. At a more disaggregated level, density usually is positive and significant (Eide, Aasness, and Skjerpen 1994).*
19. *In their surveys Eide, Aasness, and Skjerpen remark that crime is usually positively related to the share of young people but that "considering that most crimes are committed by young people, one would have expected strong results" (1994, 163). They suggest insufficient variability in the data and that "young people are perhaps not different, just poorer" (163) as possible explanations. Marvell and Moody find that only a small share of the ninety studies they reviewed find a significant relationship with the age structure and conclude that crime "forecasts based on demographic trends are not likely to be helpful" (1991, 237).*
20. *The regressions could be used to explain crime variation by states. One simple way is to insert the numbers from the tables of explanatory variables into the regression and calculate predicted crime rates. Thus, for instance, Florida's crime rates in the 1990s are driven by a very high population density and low welfare expenditures.*
21. *Chiricos (1987) and Eide, Aasness, and Skjerpen (1994) survey the mixed evidence on the relationship between unemployment and crime. Chiricos concludes that on balance the evidence favors a positive relationship that frequently is significant. Hsieh and Pugh (1993), in their metanalysis of the literature on poverty, income inequality, and violent crime, conclude that the evidence indicates that resource deprivation is an important determinant of violent crime.*
22. *The results on education and welfare are consistent with studies that use more aggregated data, as in Levitt (1997). However, Zhang (1997) finds a significantly negative relationship to property crime for different welfare measures using 1987 state data. Using a more disaggregated approach, Witte and Tauchen's (1994) study suggests that parochial school education has a significantly negative effect as opposed to other forms of primary and secondary education.*
23. *This estimate is close to those by Marvell and Moody (1994) and Spelman (1994).*
24. *Marvell and Moody (1996) find in a survey of thirty-six studies of the police and crime relationship that there is little evidence of a negative relationship, leading some researchers to dismiss police activity as a source of crime reduction. They argue that overall the studies pay insufficient attention to correcting for simultaneity bias.*

(something that is corroborated by the fact that longer lags produced "better" results for the police variables). It also may be that, since imprisonments are one output of police activities, the coefficient on the share of prisoners already captures the effect of police efforts on crime. If so, this fact would explain why the police variables are largely insignificant in a regression in which both prison and police variables are used as explanatory variables.

To shed more light on these explanations, imprisonment rates were regressed on both police variables. From this regression two variables were extracted: the part explained by the police variables and the residual that was unexplained, which can be thought of as a measure of the exogenous imprisonment rate. These two variables were substituted into the regression in place of the original imprisonment variable. The results of this regression appear in Table 5 in the second column for each dependent variable; the estimates for all other variables were exactly the same as before so that only differences in the estimates are recorded. The exogenous conviction rate (not explained by police variables) is still negative and strongly significant. In fact, the coefficient estimates for the residual are identical to the original conviction rate estimates, reflecting the fact that the police variables did not significantly explain conviction rates in the first place. However, the coefficient on the imprisonment rate due to police efforts is large and negative (except for murder) and can generously be seen as weakly significant. Thus, police efforts can be interpreted as reducing crime through conviction. Interestingly, with this new specification, the coefficient for police employment turns to zero, suggesting that the numbers on police forces affect crime only inasmuch as they affect imprisonment rates. Finally, police expenditures again tend to have a positive coefficient.

The results suggest that adding to the police force or increasing police expenditures will reduce crime only insofar as doing so leads to higher rates of conviction and imprisonment.

In sum, the results of exploratory regressions suggest that crime does fit the economic model, particularly property crime. Generally, the age, unemployment, per capita incomes, and imprisonment rate variables were highly significant and usually of the expected sign (positive for age and income and negative for prisoners) and density was largely insignificant. The exception was that unemployment is positively related to property crimes and negatively related to violent crimes, a result that is consistent with the mixed results reported in the crime literature. The public expenditures analyzed here give a mixed picture. Primary and secondary education expenditures were usually insignificant, the same as when a more inclusive measure of education that included postsecondary expenditures was used. Interestingly, welfare expenditures are found to be significant for auto theft and murder, two of the crime variables that suffer least from measurement error. Studies by Zhang (1997) and Witte and Tauchen (1994) suggest that looking at components of the welfare and education measures might produce stronger results. Finally, police expenditures were usually positive and sometimes highly significant, and police employment had an effect only through imprisonment rates. On balance, the police results suggest that simultaneity bias still persists and that stronger empirical measures are warranted (Marvell and Moody 1994; Levitt 1997).

Conclusion

The United States in the 1990s has seen a dramatic fall in crime in almost all categories. However, not all regions have benefited equally; in fact some areas have seen no improvement. In particular, states in the Southeast have seen dramatic worsening in crime rates relative to other regions. The empirical analysis suggests that for some states increased crime is related to rapid growth in personal incomes, making property crime more lucrative. For other states, high crime rates reflect relatively high unemployment rates and low expected earnings from legitimate work for some population segments. Also, demographic factors may have been at work, with a comparative rise in the share of fifteen- to twenty-four-year-olds, the population share associated most closely with arrest rates.

The empirical analysis suggests several policy conclusions. Analysis of the effects of unemployment on crime indicates that unemployment insurance that alleviates the costs of unemployment might have some effects in ameliorating crime. Other social policies might also be helpful. For instance, the analysis shows that increases in welfare expenditures are associated with reduced auto theft and murder rates. Beyond these types of policies, there is strong empirical support that increases in imprisonment rates will significantly reduce crime. However, for the Southeast, increases in conviction rates may not have offset the crime-increasing effects of welfare reductions combined with other demographic and economic changes. Finally, the results reported here suggest that adding to the police force or increasing police expenditures will reduce crime only insofar as doing so leads to higher rates of conviction and imprisonment. It should be noted, however, that the methods used may have been too weak to capture other effects of police efforts on crime.

Data Definitions and Sources

The data used in this article are annual state-level data for the United States for the years from 1971 to 1994 and until 1996 for the crime data.

Variables

Crime rate: Computed as number of crimes per 100,000 inhabitants (see Special Data Concerns). Data source: FBI (1960–96).

Burglary: "[I]includes any unlawful entry to commit a felony or a theft and includes attempted burglary and burglary followed by larceny." Definition source: Bureau of the Census (1997, 197). Data source: FBI (1960–96).

Larceny: "[I]ncludes theft of property or articles of value without use of force and violence or fraud and excludes embezzlement, 'con games,' forgery, etc." Definition source: Bureau of the Census (1997, 197–98). Data source: FBI (1960–96).

Auto theft: "[I]ncludes all cases where vehicles are driven away and abandoned, but excludes vehicles taken for temporary use and returned by the taker." Definition source: Bureau of the Census (1997, 198). Data source: FBI (1960–96).

Property crime: Includes burglary, larceny, and auto theft. Data source: FBI (1960–96).

Aggravated assault: "[I]ncludes assault with intent to kill." Definition Source: Bureau of the Census (1997, 197). Data source: FBI (1960–96).

Murder and nonnegligent manslaughter: "[B]ased on police investigations, as opposed to the determination of a medical examiner or judicial body, includes willful felonious homicides, and excludes attempts and assaults to kill, suicides, accidental deaths, justifiable homicides, and deaths caused by negligence." Definition source: Bureau of the Census (1997, 197). Data source: FBI (1960–96).

Violent crime: Includes murder, forcible rape, robbery, and aggravated assault. Data source: FBI (1960–96).

Overall crime index: Includes violent and property crime. Data source: FBI (1960–96).

Population: Civilian noninstitutional population, sixteen years or older. Data source: Department of Labor (1970–95).

Population density: Population of state/land area (in square miles) of state or thousands of people per square mile. Data source: Rand McNally (1995).

Population of youth: Youth between the ages of fifteen and nineteen and between twenty and twenty-four as a percentage of total state population. Data source: Bureau of the Census (1998).

Personal Income: Nominal, annual, by state. Data source: DRI/McGraw-Hill.

Unemployment rate: Seasonally unadjusted, by state. Data source: Department of Labor (1970–95).

Police protection employment: Includes all activities concerned with the enforcement of law and order, including "police training academies, coroners, medical examiners, forensic services and crime labs, temporary 'lockups,' police communications and radios services, buildings or other facilities used exclusively for police purposes (including rentals), criminal justice planning, and payments for transporting criminals." Definition source: Bureau of the Census (1992). The employment data reported for police protection employees represent full-time equivalent employment. Data source: Bureau of the Census (1958–95).

Police protection expenditure: All amounts of money paid out by the government for the "preservation of law and order and traffic safety. Includes police patrols and communications, crime prevention activities, detention and custody of persons awaiting trial, traffic safety, and vehicular inspection." Definition source: Bureau of the Census (1996, A-7). Data source: Bureau of the Census (1958–96).

Prison population (convicts): Adult prisoners in state prisons/state population. Data source: Department of Justice (1957–72; 1973–96).

Public welfare expenditure: Includes "support of and assistance to needy persons contingent upon their need. [It] excludes pensions to former employees and other benefits not contingent on need. Expenditures under this heading include: cash assistance paid directly to needy persons under the categorical programs (Old Age Assistance, Aid to Families with Dependent Children, Aid to the Blind, and Aid to the Disabled) and under any other welfare programs; vendor payments made directly to private purveyors for medical care, burials, and other commodities and services provided under welfare programs." Definition source: Bureau of the Census (1996, A-8). Data source: Bureau of the Census (1958–96).

Explanatory Variable Rankings: 1990–94

Region	State	Population Density[a]	Rank	15-24-Year-Olds[a]	Rank	Unemployment Rate[a]	Rank	Per Capita Personal Income[a]	Rank	Welfare[a]	Rank	Primary and Secondary Education[a]	Rank	Convicts[a]	Rank	Police Employment[a]	Rank	Police Expenditures[a]	Rank
New England	CT	934.8	47	107.0	6	95.0	24	135.1	1	94.3	24	102.2	27	115.3	13	99.3	18	79.4	36
	MA	1,066.2	48	100.5	24	110.6	43	118.0	4	129.1	3	86.9	44	51.1	42	103.3	12	86.0	26
	ME	55.3	15	98.8	34	107.1	39	106.0	11	150.5	2	128.4	6	38.5	47	83.4	37	66.6	47
	NH	172.6	33	102.6	17	95.7	27	108.5	8	105.1	18	90.7	42	46.8	46	95.0	24	77.9	38
	RI	1,318.1	49	101.0	23	119.4	47	101.3	17	125.4	6	106.2	20	85.1	25	105.3	9	96.7	15
	VT	85.5	21	96.6	39	85.8	15	93.1	29	118.7	9	131.5	4	64.9	38	76.4	46	72.7	42
Mideast	DE	494.5	44	91.5	47	82.6	13	105.7	12	68.8	46	97.3	35	180.0	1	101.6	13	96.6	16
	MD	691.3	46	99.8	28	86.9	17	100.2	19	74.4	41	91.7	41	125.5	10	107.7	7	93.3	20
	NJ	1,455.6	50	86.6	50	104.2	36	129.0	2	85.6	32	116.4	9	92.4	23	137.7	1	99.3	12
	NY	529.9	45	105.9	10	108.6	42	119.4	3	160.9	1	114.3	11	105.4	19	133.2	2	126.1	5
	PA	370.5	42	99.0	32	101.1	33	102.1	16	109.6	15	102.1	28	65.0	37	83.9	36	72.7	43
Great Lakes	IL	289.3	40	107.2	5	104.2	35	108.1	9	81.1	35	83.8	47	85.3	24	122.4	3	101.5	11
	IN	218.3	35	99.1	31	86.3	16	91.5	31	95.2	23	105.7	22	77.3	29	82.2	39	67.0	46
	MI	230.5	37	105.1	11	117.9	46	98.8	21	106.6	16	110.8	13	125.3	11	82.9	38	99.2	13
	OH	372.1	43	100.0	26	96.1	29	94.6	24	112.8	12	99.7	31	107.2	17	85.3	34	91.8	21
	WI	127.3	27	99.0	33	74.9	9	94.6	23	115.6	11	116.5	8	53.6	41	94.0	27	109.2	10
Plains	IA	69.6	18	106.2	8	65.1	4	90.0	33	96.3	22	110.8	14	50.8	43	78.9	43	77.3	39
	KS	42.7	11	108.8	4	72.5	7	95.2	22	67.1	47	105.5	23	74.2	32	97.7	21	83.1	32
	MN	78.1	20	96.9	38	74.5	8	101.2	18	125.7	5	106.2	19	26.6	49	74.4	47	84.6	29
	MO	76.9	19	92.8	46	114.9	45	69.9	50	102.6	19	106.9	17	111.5	16	84.3	35	75.7	40
	ND	12.8	4	98.4	35	65.7	5	83.3	39	110.5	13	106.1	21	24.4	50	73.4	48	59.3	49
	NE	29.0	9	101.7	22	41.2	1	94.6	25	81.6	34	98.0	34	49.5	44	86.2	32	71.1	44
	SD	13.1	5	95.7	41	54.3	2	86.4	37	78.7	38	98.6	32	65.0	36	77.4	45	62.9	48
Southeast	AL	113.2	26	115.1	2	107.1	40	81.6	41	86.3	31	91.8	40	132.5	8	89.9	29	84.4	30
	AR	63.9	16	97.4	37	101.6	34	76.0	49	105.6	17	101.2	30	105.8	18	79.0	42	68.8	45
	FL	345.9	41	95.9	40	107.2	41	99.6	20	66.5	48	86.1	45	115.0	14	120.6	5	126.1	6
	GA	161.8	30	99.2	30	87.1	18	92.2	30	88.2	30	101.6	29	121.3	12	98.9	20	87.6	25
	KY	131.4	28	89.3	49	97.0	30	80.8	44	123.1	7	105.2	24	84.2	26	71.1	49	73.2	41
	LA	133.0	29	94.1	44	113.7	44	78.9	45	117.0	10	107.4	16	154.7	3	100.4	15	119.5	8
	MS	104.7	24	91.5	48	90.3	20	94.0	26	77.1	39	84.9	46	97.4	21	100.7	14	81.8	35
	NC	194.7	34	101.9	21	76.6	10	88.7	34	80.8	36	98.2	33	93.8	22	94.4	26	88.5	24
	SC	164.8	31	97.5	36	95.2	26	80.9	43	100.7	20	106.7	18	160.6	2	91.7	28	83.7	31
	TN	169.3	32	106.2	9	88.1	19	87.4	35	90.6	28	82.5	48	75.4	31	89.7	30	78.8	37
	VA	223.0	36	104.5	13	80.8	12	104.2	14	58.4	49	92.8	38	105.2	20	88.1	31	84.6	28
	WV	104.0	23	100.2	25	152.2	50	77.0	48	129.0	4	134.5	3	29.2	48	58.3	50	51.5	50
Southwest	AZ	47.6	14	102.2	18	96.0	28	87.4	36	91.2	27	96.0	37	133.7	7	106.0	8	132.0	2
	NM	18.1	8	93.3	45	105.8	38	77.4	46	96.7	21	118.5	7	66.2	35	103.9	11	127.7	4
	OK	64.7	17	95.6	42	92.2	21	81.5	42	92.5	26	102.7	26	142.6	6	96.8	23	83.0	33
	TX	93.8	22	101.9	20	104.8	37	90.9	32	74.2	42	109.9	15	127.3	9	100.2	16	90.2	23
Rocky Mountain	CO	46.5	13	109.3	3	78.1	11	102.5	15	70.3	44	92.6	39	81.0	27	97.7	22	96.5	17
	ID	18.0	7	100.0	27	93.1	23	83.4	38	70.6	43	102.8	25	68.2	34	94.4	25	91.6	22
	MT	7.9	3	102.6	16	95.0	25	81.8	40	94.2	25	130.2	5	58.3	40	85.8	33	82.7	34
	UT	30.9	10	106.6	7	67.1	6	77.4	47	79.5	37	113.9	12	47.1	45	77.7	44	93.7	18
	WY	6.6	2	123.0	1	83.7	14	93.8	27	69.6	45	144.3	2	75.9	30	121.8	4	116.2	9
Far West	AK	1.4	1	102.0	19	123.4	49	110.5	6	121.2	8	165.4	1	152.4	4	99.6	17	144.8	1
	CA	273.2	39	103.2	14	123.3	48	107.0	10	109.7	14	87.9	43	112.0	15	99.3	19	127.8	3
	HI	247.4	38	95.1	43	63.0	3	112.1	5	77.0	40	82.2	49	79.4	28	104.3	10	93.3	19
	NV	16.9	6	105.0	12	93.0	22	109.0	7	49.2	50	77.0	50	143.1	5	120.1	6	123.4	7
	OR	42.9	12	99.5	29	97.1	31	93.1	28	82.3	33	114.5	10	70.3	33	80.6	40	98.3	14
	WA	107.1	25	102.8	15	99.7	32	104.6	13	90.6	29	96.4	36	58.6	39	79.1	41	85.2	27

[a]Percentage deviation relative to the nation

Source: See data appendix

Primary and secondary education expenditures: Public elementary (primary) and secondary education expenditure by the schools themselves (and not state governments on the schools) in thousands of dollars. Includes "payments for instructional, support services, and other activities of local public schools for kindergarten through high school programs . . . the operation of public schools, construction of school buildings, purchase and operation of school buses, and other services ancillary to the provision of public schools." Definition source: Bureau of the Census (1996, A-3) (see Special Data Concerns). Data source: Department of Education (1964–78; 1978–83; 1984–95).

Special Data Concerns

Primary and Secondary Education Expenditures. No data were available for any states for the years 1982 and 1983. In order not to lose all information in these years in the panel analysis of the data, these observations have been estimated for each state in the following ways. For 1982, a weighted average was estimated using the formula (2/3) × (1981 observation) + (1/3) × (1984 observation). The 1983 value was estimated using the weighted average (1/3) × (1981 observation) + (2/3) × (1984 observation). The following list gives the names of the states and how estimates were found for the missing years:

State	Year	Formula
AK, GA, IL	1979	average of 1978 and 1980
NJ	1975	average of 1974 and 1976
WI	1978	(2/3) × (1977 observation) + (1/3) × (1980 observation)
WI	1979	(1/3) × (1977 observation) + (2/3) × (1980 observation)

Crime Rate Variables. The following states did not have complete data for some years, and therefore those crime counts have been estimated by the FBI.

State	Years
KS	1993, 1994
KY	1988
MT	1994
IL	1993, 1994

Illinois data from 1975–84 contain the "Chicago adjustment." Forcible rape data were estimated for Illinois for 1985–94 (forcible rape is included in the violent crime data).

REFERENCES

BECKER, GARY S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 73 (March/April): 169–217.

CHIRICOS, THEODORE G. 1987. "Rates of Crime and Unemployment: An Analysis of Aggregate Research Evidence." *Social Problems* 34:187–212.

COHEN, MARK. 1988. "Pain, Suffering, and Jury Awards: A Study of the Cost of Crime to Victims." *Law and Society Review* 22:537–55.

DIIULIO, JOHN J., JR. 1996. "Help Wanted: Economists, Crime, and Public Policy." *Journal of Economic Perspectives* 10 (Winter): 3–24.

EHRLICH, ISAAC. 1996. "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives* 10 (Winter): 43–67.

EIDE, ERLING, JORGEN AASNESS, AND TERJE SKJERPEN. 1994. *Economics of Crime: Deterrence and the Rational Offender.* Oxford: North-Holland.

FEDERAL BUREAU OF INVESTIGATION. 1960–96. *State Crime Data, 1960–96.* FBI Uniform Crime Reporting Program Data. Available online at <http://www.ojp.usdoj.gov/bjs/datast.htm> [February 24, 1998].

FREEMAN, RICHARD B. 1996. "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?" *Journal of Economic Perspectives* 10 (Winter): 25–42.

GROVE, WALTER R., MICHAEL HUGHES, AND MICHAEL GEERKEN. 1985. "Are Uniform Crime Reports a Valid Indicator of Index Crimes? An Affirmative Answer with Minor Qualifications." *Criminology* 23:451–501.

HELLMAN, DARYL A., AND NEIL O. ALPER. 1990. *Economics of Crime: Theory and Practice.* 2d ed. Needham Heights, Mass.: Simon and Schuster, Ginn Press.

HSIEH, CHING-CHI, AND M.D. PUGH. 1993. "Poverty, Income Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies." *Criminal Justice Review* 18:182–202.

LEVITT, STEVEN D. 1996. "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Crowding Litigation." *Quarterly Journal of Economics* 111 (May): 319–51.

———. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review* 87 (June): 270–90.

MARVELL, THOMAS B., AND CARLISLE E. MOODY. 1991. "Age Structure and Crime Rates." *Journal of Quantitative Criminology* 7, no. 3:237–73.

———. 1994. "Prison Population Growth and Crime Reduction." *Journal of Quantitative Criminology* 10, no. 4:109–40.

———. 1996. "Police Levels, Crime Rates, and Specification Problems." *Criminology* 34 (November): 609–46.

MILLER, TED, MARK COHEN, AND SHELLI ROSSMAN. 1993. "Victim Costs of Violent Crimes and Resulting Injuries." *Health Affairs* 12:186–97.

RAND MCNALLY. 1995. *Road Atlas: United States, Canada, Mexico*.

SPELMAN, WILLIAM. 1994. *Criminal Incapacitation.* New York: Plenum Press.

U.S. BUREAU OF THE CENSUS. 1958–95. *State Distribution of Public Employment*. Washington, D.C.: GPO.

———. 1958–96. *Government Finances*. Washington, D.C.: GPO.

———. 1992. *Governments Finance and Employment Classification Manual*. Available online at <http.://www.census.gov/govs/manual/ch4.txt> [March 16, 1999].

———. 1997. *Statistical Abstract of the United States*. Washington, D.C.: GPO.

———. 1998. *Population Estimates Program*. Washington, D.C.: GPO.

U.S. DEPARTMENT OF EDUCATION. 1964–78. *Statistics of Public Schools*. Washington, D.C.: GPO.

———. 1978–83. *Statistics of Public Schools: Revenues and Expenditures for Public Elementary and Secondary Education*. Washington, D.C.: GPO.

———. 1984–95. *Estimates of School Statistics*. Washington, D.C.: GPO.

U.S. DEPARTMENT OF JUSTICE. Bureau of Justice Statistics. 1973–96. *Prisoners in State and Federal Institutions*. Washington D.C.: GPO.

———. 1995. *Criminal Victimization, 1993*. Washington, D.C.: GPO, May.

———. Federal Bureau of Prisons. 1957–72. *National Prisoner Statistics*. Washington, D.C.: GPO.

U.S. DEPARTMENT OF LABOR. Bureau of Labor Statistics. 1970–95. *CPS Household Survey: Unadjusted Unemployment Rate*. Washington, D.C.: GPO.

WHITE, HALBERT. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817–38.

WITTE, ANN DRYDEN, AND HELEN TAUCHEN. 1994. "Work and Crime: An Exploration Using Panel Data." National Bureau of Economic Research Working Paper 4794, July.

ZHANG, JUNSEN. 1997. "The Effect of Welfare Programs on Criminal Behavior: A Theoretical and Empirical Analysis." *Economic Inquiry* 35 (January): 120–37.