1985

ECONOMIC REVIEW

FEDERAL RESERVE BANK
of CLEVELAND

QUARTER II

Economic Review
Federal Reserve Bank of Cleveland

Quarter II 1985

Economic Review is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland, P.O. Box 6387, Cleveland, OH 44101. Telephone: 216/579-2000.

Editor: William G. Murmann. Assistant editor: Meredith Holmes. Design: Jamie Feldman. Typesetting: Lucy Balazek.

ISSN 0013-0281

Vector Autoregressive Forecasts of Recession and Recovery: Is Less More?

by Gordon Schlegel

Gordon Schlegel is a research assistant at the Federal Reserve Bank of Cleveland. *The author would like to thank K.J. Kowalewski, Owen Humpage, Mark Sniderman, James Hoehn, and William Gavin for their helpful comments and suggestions.*

Economic forecasts are valuable tools for decision makers in many different areas. When used with discretion, forecasts can help guide the strategic plans of businesses and corporations. A reasonably sharp picture of the future is also important in the formation of sound fiscal and monetary policy.

Forecasts are particularly important when the economy has just entered a recessionary or expansionary period. Policies that are useful in expansionary periods must often be adjusted before and during contractions, and vice versa. To get an idea of the degree to which policies must change, one needs to forecast the extent of the expansion or contraction to come.

Many economists are turning to the use of vector autoregressive (VAR) models for forecasting. A number of studies have indicated that VARs forecast as well as, if not better than, many large structural models; one such study is that of Lupoletti and Webb (1984). However, the forecast periods used in these studies are not differentiated into expansionary and recessionary periods. An economist using VARs might want to ask the question: "What VAR specification will do the best job in predicting the length and intensity of recessions and recoveries?"

This paper provides a possible answer to this important question. The first section discusses the reasons that VARs are gaining in popularity among forecasters and describes the methodology of VARs. Section II discusses the pros and cons of VARs. Section III describes the various model specifications compared in the study and the measures of forecast accuracy employed in the comparison. Section IV looks at the estimation results for the specified models, while section V considers a more recently developed VAR technique. Finally, section VI sums up the overall results of the study and mentions several cautions concerning the interpretations of the results.

I. VARs: Why and How?

In their never-ending search for the perfect crystal ball, economic forecasters try to obtain high forecast accuracy and, at the same time, use as simple a technique as possible. This is particularly true of business economists who work under significant time and resource constraints which, in turn, limit the degree of sophistication they can apply to their forecasts.

However, the forecasts must still be accurate enough to give a fairly sharp picture of the environment that firms and consumers will be facing in the immediate future. A forecast is, obviously, not useful if it does not predict with an "acceptable" degree of accuracy. However, even if the technique exists to produce a perfect forecast, the method is worthless if it is too complex for a practitioner to apply properly.

VAR techniques have been proposed as a means through which one can have the best of both worlds: simplicity and accuracy.[1] In a VAR system with n lags, each variable being forecast is regressed against its own values in each of the n preceding periods, against the values in each of the n previous periods of all of the other variables being forecast, and against a constant term. For example, a VAR system with three variables, X, Y, and Z, and with two lags would consist of the following equations:

$$X = c_1 + a_{11}X_{-1} + a_{21}X_{-2} + b_{11}Y_{-1} + b_{21}Y_{-2} + c_{11}Z_{-1} + c_{21}Z_{-2} + e_1,$$

$$Y = c_2 + a_{12}X_{-1} + a_{22}X_{-2} + b_{12}Y_{-1} + b_{22}Y_{-2} + c_{12}Z_{-1} + c_{22}Z_{-2} + e_2,$$

$$Z = c_3 + a_{13}X_{-1} + a_{23}X_{-2} + b_{13}Y_{-1} + b_{23}Y_{-2} + c_{13}Z_{-1} + c_{23}Z_{-2} + e_3,$$

1. One cannot, however, perform policy simulations using vector autoregressive models. Lucas (1976) pointed out that under alternative policies, agents will have different views about the way exogenous shocks affect the system. Therefore, one cannot use the same set of parameters for all alternative policies one may wish to examine. This implies that the coefficients obtained through in-sample estimation may not accurately reflect policy changes.

where

X_{-n} = the value of X n periods before the current period,

e_r = the error term of equation r, distributed as a normal random variable with mean 0 and constant variance, and

c_r = the constant term of equation r.

The equations are estimated individually to yield estimates for all parameters and constant terms. One can then calculate the reduced form of the system and predict the values of all variables in the current time period. These values can, in turn, be used as regressors in predicting the next period's values for the variables. The process can be continued indefinitely, enabling one to produce dynamic, out-of-sample forecasts as far into the future as desired, given the information available in the present period.

The regression equations are commonly estimated in one of two ways. With ordinary least squares, the parameters are completely unconstrained and can assume whatever values best fit the data. Bayesian techniques enable a forecaster to explicitly include, in the model, subjective judgment or other objective evidence concerning the values of the parameters, as well as the degree of confidence he has in his judgment. A general discussion of the techniques is given in Todd (1984), while Litterman (1979) approaches the topic from a more technical basis.

In this paper, we first search for the optimal ordinary least squares (OLSQ) specification, where the "optimal" specification is the one that provides the most accurate forecasts, the measures of accuracy being described below.[2] We then compare this specification to one derived through a Bayesian procedure.

II. Advantages and Disadvantages of VAR Models

VARs have a number of characteristics that make them convenient for those who make economic forecasts on a regular basis. Of these characteristics, the following five seem especially worthy of note:

1) It is relatively easy to write a computer program to perform a VAR. A programmer with a moderate amount of skill and a package of standard regression techniques should be able to implement such a program without much trouble.

2) The commands needed to perform an OLSQ VAR can be implemented in virtually any programming language. This would make it unnecessary to buy a specialized package to run VARs and would enable a forecaster to avoid this type of expense. The Bayesian VAR can be implemented with a little more effort, provided that matrix capabilities are available.

3) Since VARs can be programmed fairly easily, it might not be necessary to buy forecasting services from an outside data vendor. Subscriptions to the major econometric forecasting services can cost from $16,000 to $20,000 per year, no bargain if, as Lupoletti and Webb (1984) suggest, the simpler VAR models can perform as well as, or better than, the large models.

4) Because VARs only use a relatively small number of variables, it is easy to update and revise the data series as needed.

5) In their pure form, VARs require no subjective add factors. Large models contain a number of arbitrary constants that a forecaster might be unable to estimate sufficiently well for his purposes, due to a lack of necessary specialized information or expertise. The VAR gets around this problem by avoiding it.

No forecasting technique, however, is without its problems. VAR models have two major disadvantages:

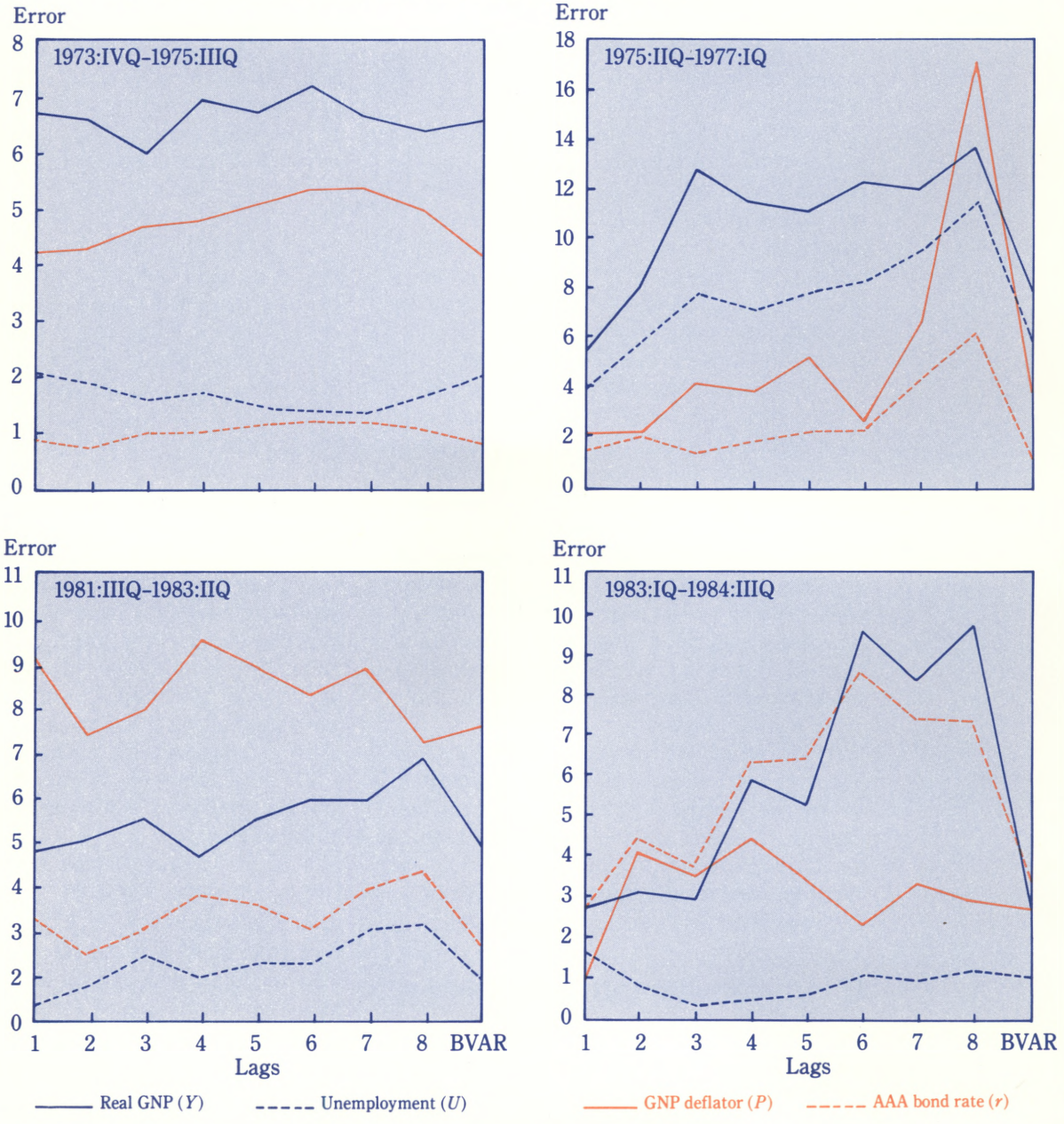
1) Since most aggregate economic time series are highly correlated with their own previous values and with present and past values of other time series, multicollinearity can become a serious problem as more and more series and lagged values of series are added to the model. As the system expands, it can become very difficult to separate the effects of the explanatory variables, and the

2. The in-sample fits of the various specifications are not considered. We only want to predict future values of the variables in the system, not explain their past values.

parameter estimates can become highly sensitive to the combination of variables used in the model.

Also, a high degree of multicollinearity will make it difficult to determine which explanatory variables are significant, since the standard errors of the coefficient estimates will tend to be large. A forecaster considering

Fig. 1 Dynamic Out-of-Sample Root Mean Squared Error



a certain lag structure might want to ask if certain lagged variables can be dropped from the system without sacrificing forecast accuracy. A detailed discussion is found in Intriligator (1978), among others.

As far as the forecasting aspects of multicollinearity are concerned, Christ (1966) points out that if the joint distribution of the regressors changes during a forecasting period, multicollinearity between regressors will affect the accuracy of the forecasts. Given the increasing volatility of aggregate measures of economic activity over the past 10 years, particularly interest rates, it would appear that such changes have taken place. Multicollinearity, therefore, seems to present a problem for VAR forecasting.

2) As the number of variables of a VAR model increases, the number of parameters to be estimated goes up rapidly. If a variable is added to the model, each equation has n more coefficients to be estimated, where n is the number of lags for each variable.

If a lag period is added, each equation has r more parameters, where r is the number of variables in the system. As the number of coefficients increases relative to the amount of available data, random events of the past, as well as systematic relationships, are increasingly reflected in the coefficients. If these coefficients are used in out-of-sample prediction, a set of future random events that differs from the shocks of the past would be expected to result in less accurate forecasts. This problem is discussed in Todd (1984).

III. Model Specification

The model contains four variables: the growth rate of the GNP deflator (P), the growth rate of real GNP (Y), Moody's AAA corporate bond rate (r), and the civilian unemployment rate (U).[3] All variables are expressed as percentages—the growth rates being annualized. We wanted to examine how well the various model specifications estimate the scope of the expansion or recession to come because, as mentioned before, once an expansion or contraction begins, an economist needs an idea of how long the new phase of the business cycle will last.[4]

One-quarter- and eight-quarter-ahead, *ex post*, dynamic, out-of-sample forecasts were produced from two cyclical peaks: the fourth quarter of 1973 and the third quarter of 1981, and from one cyclical trough: the second quarter of 1975. For the period beginning in the first quarter of 1983, a cyclical trough, a seven-quarter-ahead forecast was made rather than one for eight quarters ahead, since revised data for the fourth quarter of 1984 were not available at the time this paper was written.

The first step in our estimation process was to perform a multivariate time series analysis on the four variables for each in-sample period. Using the techniques described in Box and Jenkins (1976) and Tiao and Box (1981),

3. We choose the growth rate of real GNP instead of a measure of the level of this variable. This implies that we are interested in the pattern of GNP growth over our forecast horizon, not just the proportion by which output will have grown seven or eight quarters hence.

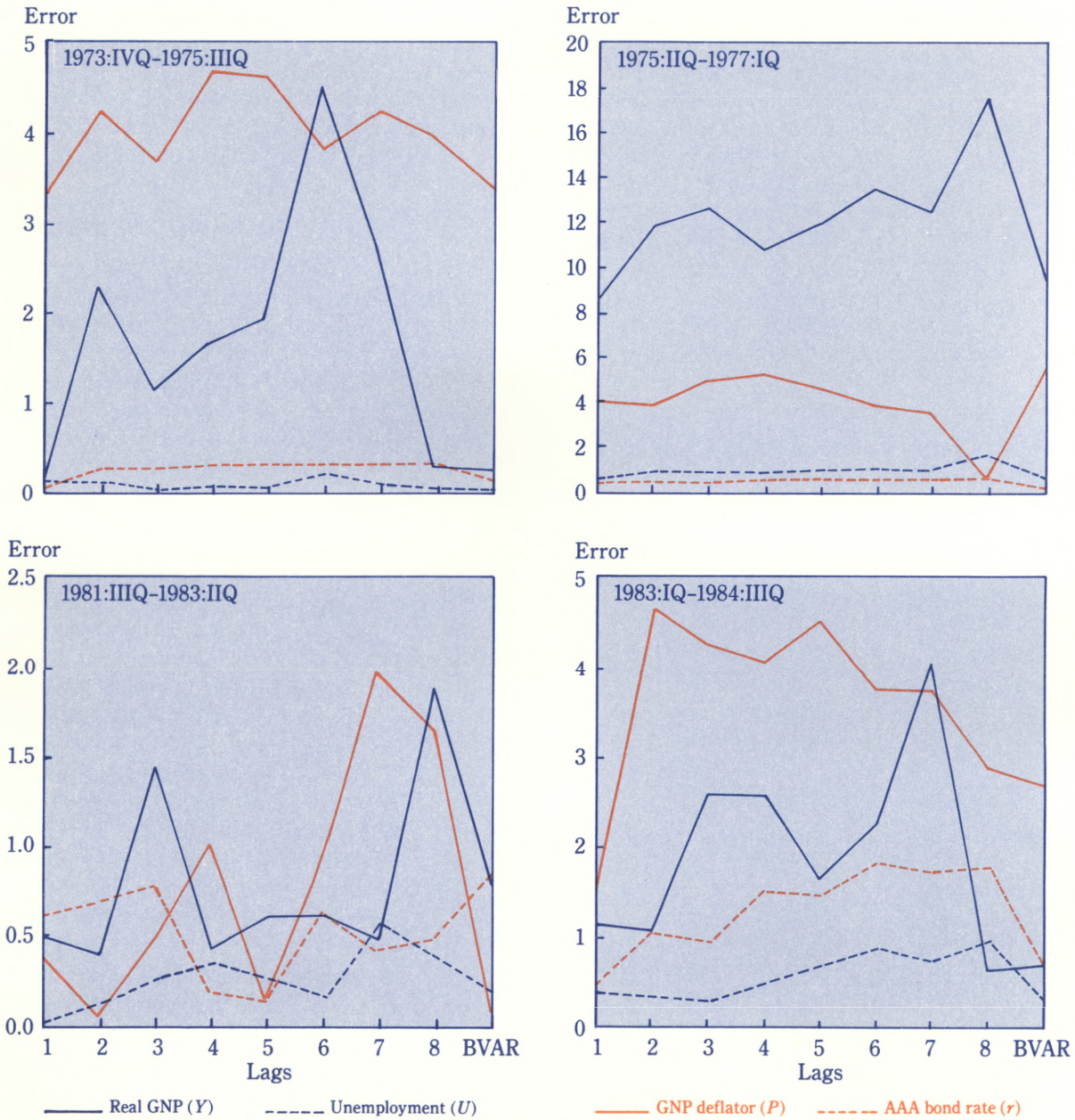
4. Implicit in this methodology is the assumption that turning points are recognized when they occur. In practice, there may be a time lag of several months between the occurrence of a turning point and its recognition by forecasters.

Table 1 Rankings of Root Mean Squared Errors of Dynamic Out-of-Sample Forecasts

Forecast period	Variables	Number of lags							
		1	2	3	4	5	6	7	8
1973:IVQ-	P	1	2	3	4	6	7	8	5
1975:IIIQ	Y	6	4	1	7	5	8	3	2
	r	2	1	3	4	6	8	7	5
	U	8	7	4	6	3	2	1	5
1975:IIQ-	P	1	2	5	4	6	3	7	8
1977:IQ	Y	1	2	7	4	3	6	5	8
	r	2	4	1	3	5	6	7	8
	U	1	2	4	3	5	6	7	8
1981:IIIQ-	P	7	2	3	8	5	4	6	1
1983:IIQ	Y	2	3	5	1	4	6	7	8
	r	4	1	2	6	5	3	7	8
	U	1	2	6	3	5	4	7	8
1983:IQ-	P	1	7	6	8	5	2	4	3
1984:IIIQ	Y	1	3	2	5	4	7	6	8
	r	1	3	2	4	5	8	7	6
	U	8	4	1	2	3	6	5	7
Total ranking		47	49	55	72	75	86	94	98

it was found that, in each period, an AR(1) or, at most, an AR(2) specification provided an adequate in-sample fit.[5] Since these models contain no moving average or lagged error terms, they closely approximate a standard VAR with one or two lags of each explanatory variable.[6] This makes our use of VAR techniques to solve the model under consideration

Fig. 2 One-Step-Ahead Errors (Absolute value)



5. There might, however, be significant moving average terms in the ARIMA specification which provides the best out-of-sample *fit.*

6. Box and Jenkins (1976) show that, for moderate or large samples, the ordinary least squares estimates of the parameters of a VAR equation differ only slightly from those obtained through the Yule-Walker equations used in ARIMA type analyses.

justified by these more general time series analysis procedures.

Each specification of the model consists of four OLSQ regressions. In the equations, each variable at period t is regressed against the values of all four variables at times t-1 through t-n, as well as a constant. For this paper, the lag length n ranged from one to eight. Despite the multicollinearity problems and estimation difficulties mentioned above, OLSQ estimation has been used in such seminal VAR models as that of Sims (1980). Our goal is to compare the different lag specifications to see which size of OLSQ VAR model provides the best out-of-sample forecasts of recession and recovery.

Comparing Forecast Accuracy

There are many measures of forecasting accuracy that one may use to compare different models that propose to explain the same phenomena. For this study, the following techniques were chosen:

1) To compare the one-step-ahead forecasts for each lag specification, we simply compare the absolute values of the one-step-ahead forecast errors. Here, we assume that it is just as undesirable to overestimate the actual values of the variables being forecast as to underestimate them, since either type of error can cause problems. We only want to know the degree to which the forecasts miss the mark.

2) For the seven- or eight-quarter-ahead forecasts, we look at the root mean squared errors of the forecasts for each variable. This seems to be an appropriate procedure, since we are not directly comparing forecasts of different variables.

Also, in business, as forecasts become more inaccurate, the fallout from decisions based on these forecasts increases even faster than the inaccuracy of the forecasts. The more inaccurate a forecast, the more sectors of a business' operation are affected by decisions made on the basis of the incorrect prediction. Thus, we seem justified in using a squared error measure, as opposed to a measure based on the simple difference between the actual and predicted values. Again, this implies that it is equally important to avoid overprediction and underprediction.

3) It would also seem useful to know if the longer-term forecasts consistently overestimate or underestimate the actual values of the variables we are interested in. If forecasts constantly miss the mark in the same direction, the problems caused by the decisions based on the forecasts will be compounded over time, rather than being compensated for by mistakes in the other direction. The measure used here is the bias component of the Theil U decomposition described in Theil (1961). This bias component is calculated as:

$$Bias = (Y - \overline{Y})^2/MSE,$$

Table 2 Rankings of Absolute Values of One-Step-Ahead Forecast Errors

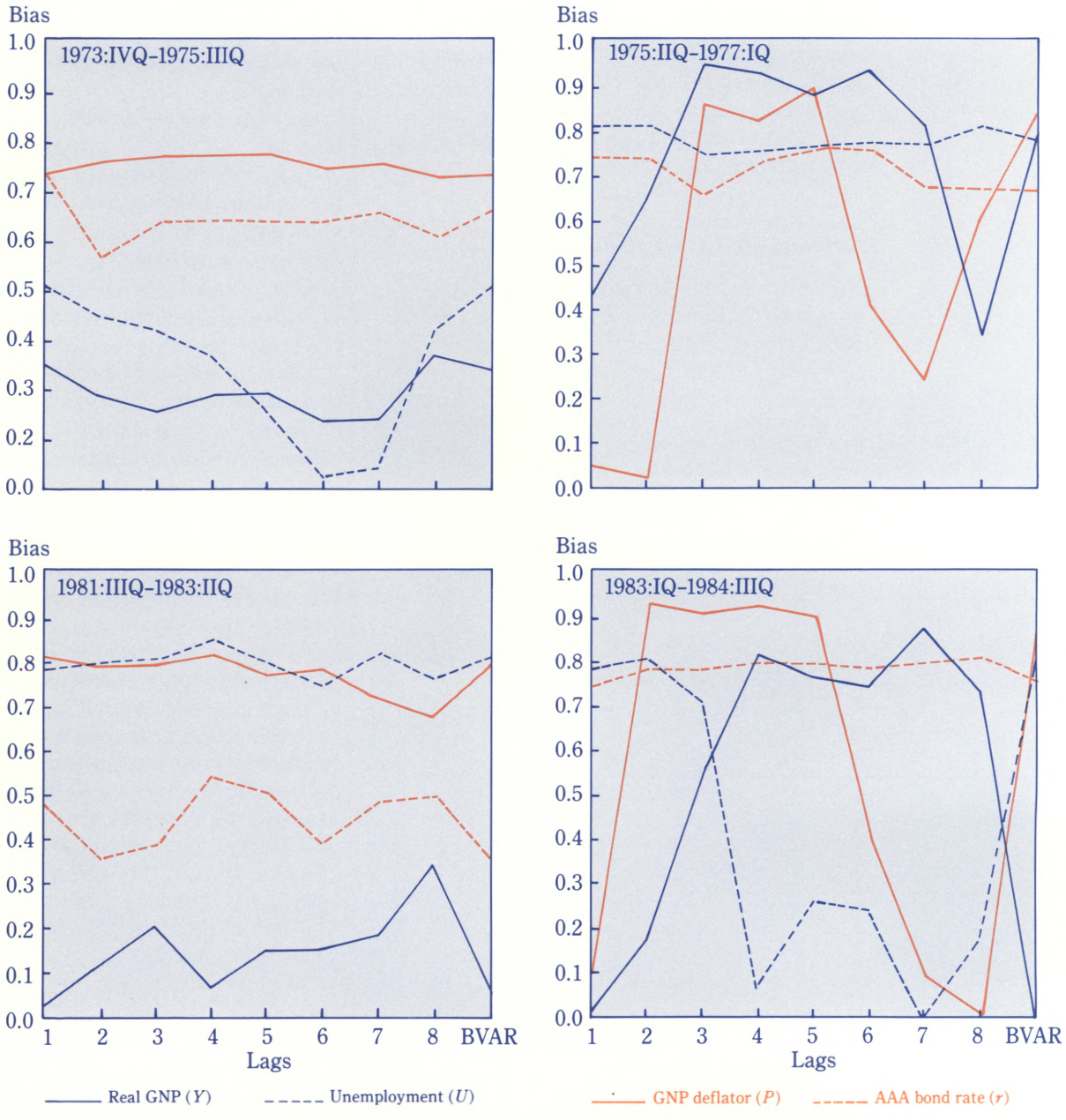
Forecast period	Variables	Number of lags 1	2	3	4	5	6	7	8
1973:IVQ-	P	1	5	2	8	7	3	6	4
1975:IIIQ	Y	1	6	3	4	5	8	7	2
	r	1	2	3	5	7	6	4	8
	U	6	7	1	4	3	8	5	2
1975:IIQ-	P	5	3	7	8	6	4	2	1
1977:IQ	Y	1	3	6	2	4	7	5	8
	r	1	3	2	4	6	5	7	8
	U	1	4	3	2	5	7	6	8
1981:IIIQ-	P	3	1	4	6	2	5	8	7
1983:IIQ	Y	4	1	7	2	5	6	3	8
	r	5	7	8	2	5	6	3	8
	U	1	2	5	6	4	3	8	7
1983:IQ-	P	1	8	6	5	7	4	3	2
1984:IIIQ	Y	3	2	7	6	4	5	8	1
	r	1	3	2	5	4	8	6	7
	U	3	2	1	4	5	7	6	8
Total ranking		38	59	65	73	75	94	87	85

where

Y = mean of the forecast values of Y,

$\overline{Y}$ = mean of the actual values of Y, during the forecast period, and

Fig. 3 Theil U Biases



$\overline{Y}$ = mean of the actual values of Y, during the forecast period, and

MSE = mean squared error of the forecast.

It must be noted that all of these measures of accuracy are subject to McNees's (1975) comments concerning the use of *ex post* forecasts to compare the predictive power of different models. However, McNees's critique does not apply to the VAR models examined here as much as it does to the large models he studies. With VARs, we have no exogenous variables and no subjective adjustments—two factors that McNees feels present a strong case for the use of *ex ante* forecasts when judging the comparative performance of econometric models. For our purposes, the *ex post* forecasts would seem to be appropriate.

To evaluate the rankings of the forecasts, we used the following techniques:

1) For each variable in each forecast period, the smallest error or bias is given a rank of one. The next smallest is given a rank of two and so on, the largest error or bias being assigned a rank of eight. If there is a tie, say, for the third smallest error, the tied errors are each given a rank of three, while the next largest error gets a five ranking. Since there are four forecast periods and four variables involved, we have 16 sets of rankings for each of the three accuracy measures.

2) The 16 sets of rankings for each measure are then added for each of the eight lag lengths. We thus obtain the totals of all the ranks for each lag length, one through eight. The lag length with the smallest total ranking is considered the one that forecasts the best, the length with the second smallest total ranking is considered the one that forecasts second best, and so on.

Several assumptions are implicit in this type of ranking scheme. We assume that all variables and all time periods are equally important. We also assume that the quantitative differences in error measures between forecasts are not important; we only want to know which forecast does better. It must be noted that even if two forecasts have different quantitative error measures, the difference between the measures may not be statistically significant. Ashley, Granger, and Schmalensee (1980) suggest a technique through which one can test the squared errors of forecasts from various models for such significance. However, our methodology generates only four forecasts of a given number of steps ahead for each variable in each model specification. Therefore, we do not have enough forecasts to utilize their method for comparing prediction errors. No test is currently available to examine the Theil U biases of different models for statistical significance.

Table 3 Rankings of Theil U Bias Statistics

Forecast period	Variables	Number of lags 1	2	3	4	5	6	7	8
1973:IVQ–	P	2	5	6	7	8	3	4	1
1975:IIIQ	Y	7	4	3	5	6	1	2	8
	r	8	1	3	6	5	4	7	2
	U	8	7	5	4	3	1	2	6
1975:IIQ–	P	2	1	7	6	8	4	3	5
1977:IQ	Y	2	3	8	6	5	7	4	1
	r	5	6	1	4	7	8	3	2
	U	6	8	1	2	3	5	4	7
1981:IIIQ–	P	7	5	6	8	3	4	2	1
1983:IIQ	Y	1	3	7	2	4	5	6	8
	r	4	1	2	8	7	3	5	6
	U	3	5	6	8	4	1	7	2
1983:IQ–	P	2	8	6	7	5	4	3	1
1984:IIIQ	Y	1	2	3	7	6	5	8	4
	r	1	3	2	6	5	4	7	8
	U	7	8	6	2	5	4	1	3
Total ranking		66	70	72	88	84	63	68	65

IV. Estimation Results

As the lag length increased, the in-sample fits improved. This follows directly from the theory of least squares regression, which states

that as more explanatory variables are added to a model, the in-sample fit should improve or stay the same. However, the graphs and the tables of rankings show that, by the methodology described above, the out-of-sample forecasts worsened as the lag lengths increased. In the case of the seven- or eight-quarter-ahead forecasts, forecast accuracy decreased over the entire range of lag lengths, with one lag giving the best forecasts and eight lags the worst. These results are shown in table 1 and figure 1. In table 2 and figure 2 we see that, in the case of the one-step-ahead forecast errors, the one-period lag gave, by far, the most accurate predictions. The forecasts got uniformly worse, as longer lags were used, until the seven-period lags, when there was a slight improvement. For the Theil U biases, shown in table 3 and figure 3, the rankings deteriorated uniformly from one lag period to four, improved slightly with five-period lags, then returned to a level very close to that of the one-period lag for lag lengths six through eight.

In sum, these results seem to indicate that, in a vector autoregressive system estimated with OLSQ, the best forecasts of recessions and recoveries are obtained by assuming that the value of each variable depends only on the values, in the immediately preceding period, of itself and all other variables in the model. A one-lag model, in essence, restricts the coefficients for all longer lags to zero.

It is possible, however, that a forecaster may have prior information—information not reflected in the data—which indicates that some of the coefficients for variables lagged two or more periods can be nonzero. To explicitly accommodate these "priors" in a statistical model in the hope of obtaining better forecasts, we can use Bayesian vector autoregression.

V. The Bayesian VAR Method

By using the Bayesian vector autoregression (BVAR) techinque, one can include, in the model, subjective estimates of the model's parameters and measures of the forecaster's confidence in his estimates.[7]

Very briefly, the BVAR technique involves the following steps:

1) Choose the lag structure and variables of the model. Here, we use the same variables as before (P, Y, r, and U) and regress each on the past three values of all four variables and a constant term.

2) Make an estimate of the coefficient values and your confidence in the estimates. Here, we have applied what Todd (1984) calls the Minnesota prior. The Minnesota prior assumes that all variables in each equation of the model behave according to a random walk; that is, all coefficients are zero except for the coefficient of the most recent value of the endogenous variable, which is one.

In other words, it is expected that the value of a variable at any given time equals the value of that variable in the preceding period. The Minnesota prior also assumes that one has more confidence in his estimates of the coefficients as the lag lengths get longer; the longer the lag, the more certain the forecaster is that a lagged variable has no effect on the system.[8]

3) Divide the variables of each equation into own and cross variables, where the endogenous variable of any given equation is the own variable for that equation, and all other variables in the equation are cross variables. Once this is done, scale the prior variances of the cross variables to units equivalent to those of the own variable.[9]

4) Multiply all own and cross-variances by hyperparameters H_o and H_c, respectively, to convert the weights determined in steps two and three to estimates of the absolute prior variances. For this estimation, we set H_o at 0.1 and H_c at 0.05 for all cross variables.

5) Perform a mixed estimation simulation using the method described, for example, in Theil (1970). A further discussion of points two, three, and four may be found in Todd (1984).

When we compare the results from the Bayesian VAR with those of the OLSQ estimations, we find that the BVAR performs at a level comparable to that of the non-Bayesian VAR with one lag. The ordinals of the root mean squared errors for the longer term fore-

7. This technique is being used by the Federal Reserve Bank of Minneapolis to model and forecast economic conditions in the Ninth Federal Reserve District. The forecasts are presented in District Economic Conditions, *available free of charge from the Research Department of the Federal Reserve Bank of Minneapolis, Minneapolis, MN 55480.*

8. The Minnesota prior constrains the variance of the coefficient of any n-period lagged variable to be $1/n$ times the variance of the coefficient of that variable when lagged once.

9. This is done by multiplying each relative prior variance of a cross variable by s_o/s_c, where s_o is the standard error of the regression in which the own variable is the endogenous variable, and s_c is the standard error of the equation in which the cross variable is the endogenous variable.

casts show that the BVAR performs slightly better that the one period VAR estimated with OLSQ. For the one-step-ahead forecast errors, the BVAR performs better than all other specifications except for the one-period non-Bayesian VAR, which does a shade better. Finally, the Theil U bias statistics show that the BVAR forecast consistently over- or underestimates the realized values by about the same degree as the one-, six-, seven-, or eight-period, lagged non-Bayesian VAR. The breakdown of the rankings for the BVAR is shown in table 4, while table 5 compares the BVAR performance to that of the OLSQ autoregressions.[10] Figures 1 through 3 chart the BVAR performance against that of OLSQ.

10. The ordinal scores in table 5 for the OLSQ VARs are not strictly comparable to those presented in tables 1 to 3. In table 5, we are comparing nine specifications: eight OLSQ and one Bayesian. The Bayesian model is not ranked in tables 1 to 3.

Table 4 Rankings of Bayesian VAR Model by Variable and Forecast Period

Forecast period	Variables	7-, 8-quarter-ahead RMSE	1-step-ahead forecast error	Theil U bias
1973:IVQ–1975:IIIQ	P	1	2	2
	Y	3	2	7
	r	2	2	8
	U	8	2	8
1975:IIQ–1977:IQ	P	5	9	7
	Y	2	2	4
	r	1	1	2
	U	3	2	6
1981:IIIQ–1983:IIQ	P	3	2	5
	Y	3	1	2
	r	2	9	2
	U	3	4	7
1983:IQ–1984:IIIQ	P	3	2	5
	Y	2	2	1
	r	2	2	2
	U	6	2	8
Total ranking		49	46	76

Table 5 Total Rankings of Bayesian and Ordinary Least Squares Models

Error measure	OLSQ (number of lags) 1	2	3	4	5	6	7	8	BVAR
1-step-ahead	44	71	79	87	89	95	101	98	46
RMSE	55	57	68	84	89	99	108	111	49
Theil U bias	75	79	81	98	93	70	76	72	76

VI. Conclusions and Caveats

The results indicate that, at least when the economy moves from an expansionary period to one of contraction, or vice versa, the forecasting ability of a VAR system deteriorates as longer lags are incorporated into the model. It also seems that a Bayesian estimation procedure does not produce forecasts that are substantially better than those of the non-Bayesian VAR with one lag per variable. Since the Bayesian method is more difficult to implement than the standard OLSQ technique, a forecaster using VAR techniques under these circumstances would probably want to stick with OLSQ.

Three important considerations must be noted, however, concerning these results. First, it may be that the comparative forecasting abilities of VARs with different lag specifications would change if the forecasts were made at points other than those considered here. For example, the one-lag model might not be superior to the others if the forecasts were being made in the middle of a cyclical expansion. Such an investigation might prove to be a useful topic for future work. If the one-lag specification is still the best method at any point of the business cycle, there is no need to use longer lags at any time. If this is not so, then we need a measure of when to change between different VAR specifications in forecasting.

The second issue is that a forecaster usually doesn't know when a recession or recovery has begun until several periods after the fact. Would the one-lag method still be best if applied when a forecaster became aware that the economy had taken a turn, rather than at the turn itself?

Finally, there is no guarantee that the Minnesota prior provides the best Bayesian VAR forecasts at the times we consider. The data may, in fact, be strongly rejecting the imposition of a random walk, producing biased coefficient estimates. A different set of estimates for the values of the coefficients and variances might yield even better predictions.

It must also be noted that, for many economists, it is more important to predict *when* the economy will turn than to forecast the *magnitude* of the turn. How well can VARs forecast the timing of the beginning and end of a recession compared to other small models and large econometric systems? Also, what VAR lag specification calls the timing of the turns most accurately? These questions must be addressed to better evaluate the usefulness of VAR forecasting methods.

As we have seen, VARs, while freeing one from the assumptions underlying a structural economic model, present problems of their own. However, since even the prototype BVARs, for instance, outperformed many commercial forecasters (see, for instance, Doan, Litterman, and Sims [1984]), further research on the models should prove very fruitful in clearing up our crystal balls.

References

Ashley, R., C.W.J. Granger, and R. Schmalensee. "Advertising and Aggregate Consumption: An Analysis of Causality," *Econometrica*, vol. 48, no. 5, pp. 1149–67.

Box, George E.P., and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control*. Revised Edition, San Francisco, CA: Holden-Day Inc., 1976.

Christ, Carl F. *Econometric Models and Methods*. New York, NY: John Wiley and Sons, Inc., 1966.

Doan, Thomas, Robert Litterman, and Christopher Sims. "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, vol. 3, no. 1, pp. 1–100.

Intriligator, Michael D. *Econometric Models, Techniques, and Applications*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1978.

Litterman, Robert B. "Techniques of Forecasting Using Vector Autoregressions," *Working Paper 115*, Federal Reserve Bank of Minneapolis, Revised November 1979.

Lucas, Robert E. "Econometric Policy Evaluation: A Critique," in K. Brunner and A. Meltzer, eds., *The Phillips Curve and Labor Markets, Vol. 1*, Carnegie-Rochester Conference Series on Public Policy. Amsterdam: North-Holland, 1976, pp. 19–46.

Lupoletti, William M., and Roy H. Webb. "Defining and Improving the Accuracy of Economic Forecasts: Contributions From a VAR Model," *Working Paper 84-6*, Federal Reserve Bank of Richmond, October 1984.

McNees, Stephen K. "An Evaluation of Economic Forecasts," *New England Economic Review*, Federal Reserve Bank of Boston, November/December 1975, pp. 3–39.

Sims, Christopher A. "Macroeconomics and Reality," *Econometrica*, vol. 48, no. 1 (January 1980), pp. 1–48.

Theil, Henri. *Economic Forecasts and Policy*, 2nd Revised Edition, Amsterdam: New-Holland Publishing Co., 1961.

_____. *Principles of Econometrics*, New York: John Wiley and Sons, Inc. 1970.

Tiao, G.C., and G.E.P. Box. "Modeling Multiple Time Series with Applications," *Journal of the American Statistical Association*, vol. 76, no. 376 (December 1981), pp. 802–16.

Todd, Richard M. "Improving Economic Forecasting With Bayesian Vector Autoregressions," *Quarterly Review*, Federal Reserve Bank of Minneapolis, Fall 1984, pp. 18–29.

Paul Gary Wyckoff is an economist at the Federal Reserve Bank of Cleveland. *The author would like to thank John Beck and Ronald Fisher for helpful comments on an earlier draft.*

Revenue Sharing and Local Public Expenditure: Old Questions, New Answers

by Paul Gary Wyckoff

During his first four years in office, President Reagan has been an active reformer of the structure of American federalism. In the Omnibus Budget Reconciliation Act of 1981, the President achieved a sweeping reform of the nation's system of categorical grants to state and local governments, consolidating many of these programs into block grants and reducing overall funding levels.

A second major Reagan initiative, a "swap" in which the federal government was to take complete responsibility for Medicaid (which provides medical care for the poor) in exchange for the states' pledge to take over Aid to Families with Dependent Children (AFDC) and food stamps, failed to win the approval of state and local leaders and has been shelved.

Now the Reagan administration proposes to further trim federal assistance to state and local governments by deleting the general revenue sharing program from its latest budget. Even if supporters manage to continue funding for one more year, the program's future is highly uncertain, since its authorizing legislation expires on September 30, 1986.

The evaluation of such a sweeping reform calls for detailed knowledge of the workings of the recipient governments. To answer the questions of the efficiency, equity, and political acceptability of this proposal, a model of local expenditure decision-making is required. Fortunately, there is a rich literature in economics on the effect of lump-sum, general-purpose aid on local spending; the question has become a focal point for the theoretical analysis of local public choice, shaping investigators' viewpoints on larger questions about the nature and efficiency of the local public sector.

The empirical results in this field, however, pose a serious challenge to the generally accepted models of 10 to 20 years ago, and have broken down rather than built consensus among economists. Thus, existing literature offers no unified framework from which to judge the Reagan proposal.

In this paper, I provide some theoretical background to the current public policy discussion on revenue sharing. In section I, the nature of the economists' previous consensus is explored, along with the empirical irregularities that broke down that consensus and invited new approaches to local public choice. Section II reviews the various ways in which economists have tried to amend or replace their previous notions in light of these empirical results. Section III offers a critique of these efforts. A new model to explain these empirical facts is summarized in section IV, along with a description of an empirical test of this model. The concluding section contains a few preliminary comments on the public policy ramifications of this new model.

I. Flypaper Effects

Two approaches have dominated the literature on modeling local public expenditure decisions. The first approach, exemplified in the work of Henderson (1968), Inman (1971), Ehrenberg (1973), Gramlich and Galper (1973), and Deacon (1978), applies standard consumer theory to this sector. Without specifying either the actors in the local decision-making process or their preferences, local governments are assumed to behave as if they are maximizing a well-behaved utility function over public and private goods, subject to a budget constraint that the total income of the community (intergovernmental grants as well as private income) must not exceed the total amount spent on private spending and local public goods.

Although it is seldom made clear in these studies, this approach implicitly assumes that the city's budget is under the control of some individual or party within the city, since a well-behaved utility function for the community will not exist unless this is the case (Arrow 1950). Subject to certain legal limits on the type of taxes collected, this controlling party determines the type and quantity of local public goods produced and the total amounts spent in the public and private sectors of the economy.

Remarkably, even this very unrestrictive approach, in which the identity of the controlling party is left unspecified, carries implications for local expenditure behavior that are inconsistent with the empirical work in this field. Since the controlling party can tax local private income at will, this model acts as if all intergovernmental aid, as well as all private income, were under the control of this anonymous decisionmaker. Just as the choice for a consumer between new furniture or a new car is independent of the composition of income between wages, capital gains, dividends, and interest, so the controlling party's division of resources between private consumption and public goods should be independent of whether the community's money comes from private income or from intergovernmental aid.[1] If all that concerns the city is to maximize some utility function over private consumption and public services, the source of the money used to pay for the city's budget is irrelevant. Therefore, the expenditure effect of a one-dollar increase in revenue sharing ought to be the same as that resulting from a one-dollar increase in aggregate private income in the community.

In his review of the early econometric work on this question, Gramlich (1977) noted that this equivalence was consistently rejected by the data. "Whether half or all the revenue-sharing money goes into higher expenditures, however, at this point all empirical studies indicate long-run responses appreciably greater than would be implied by the response of expenditures to changes in income . . ." (Gramlich [1977], p. 230). This pattern of behavior has come to be known as the flypaper effect: money originally from the public sector (intergovernmental grants) sticks in the public sector and is spent on public goods, while money originally from the private sector (local taxes on private income) sticks in that sector and is spent on private consumption.

The second major approach to modeling local public expenditure decisions retains the

1. Here, I am abstracting from any considerations as to the relative permanence of these different kinds of income. If a wage gain is considered a permanent increase in income, while a capital gain is considered transitory, this will affect the consumer's savings-consumption decision and perhaps may affect the type of durable goods purchases that he will make.

framework of consumer theory but also specifies the identity and preferences of the controlling party. Early writers in the theory of voting (see Hotelling [1929], Bowen [1943], and Black [1948]), showed that whenever binary choice is involved (two political parties, two candidates, or two sides of an issue), a position at the median of the community's preferred spending levels will generate the greatest electoral support. This result ensures that competitive political processes will always produce median outcomes. Drawing on this theoretical foundation, numerous empirical studies have utilized the assumption that local governments behave as if they were maximizing the utility of the median voter in each community (see Bergstrom and Goodman [1973], Borcherding and Deacon [1972], Ladd [1975], Lovell [1977], Perkins [1977], Inman [1978], and Pack and Pack [1978]). Under further assumptions about the demand function for local public goods and the distribution of income and wealth in the community, the income and the tax price facing the median voter can be calculated, and the response of *individuals* to changes in their public and private good budget constraint can be estimated.

Even before this approach was well developed, however, Bradford and Oates (1971) showed that it did not explain flypaper effects. They made their argument with the help of a simple graph, reproduced here as figure 1. The median voter's budget constraint between private goods and public expenditures is displayed, with a slope equal to the negative of the median voter's tax share (here labeled T). A lump-sum, general-purpose grant of amount A (which I will refer to later as simply a lump-sum grant) shifts out the budget constraint in parallel fashion. Since the budget constraint is a straight line, an income increase of amount TA ought to generate the same final budget constraint as under the aid increase, and hence the same equilibrium amounts of private goods and public expenditures. Thus, under the median voter model, an income increase of amount TA is *equivalent* to an aid increase of amount A.

Another way to think about this result is to note that the median voter controls a share of the lump-sum aid equal to TA. Since the median voter is the dominant actor in local politics, he or she can move this bundle of resources in and out of the public sector as desired. If, for example, the median voter decides to use none of the lump-sum aid for public expenditures, the money would be used to lower taxes and the median voter would receive a rebate in the amount TA. Under the median voter model then, the voter's "public income" (TA) can simply be added to his or her private income (Y) to derive the total income (Z):

$$Z = Y + TA.$$

It follows that under the median voter model an increase in the median voter's share of lump-sum aid (TA) ought to have the same expenditure effect as an increase in his or her private income (Y).

Fig. 1 Aid in the Median Voter Model
Equivalence of a Lump-Sum Grant of Amount A to an Income Increase of Amount TA

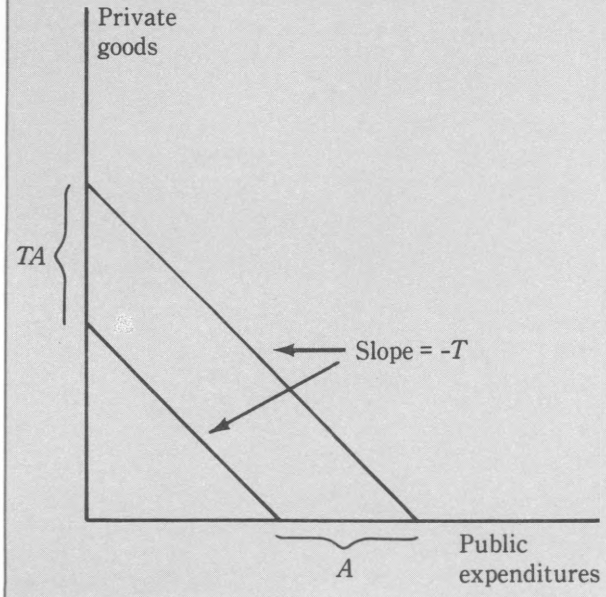


Table 1 (reproduced with permission from Fisher [1982]) shows the results of a recent survey of tests of the flypaper effect in both the median voter model and in the older expenditure-as-utility-maximization literature. For each study, the first column shows the expenditure effect that would be predicted for lump-sum aid if flypaper effects were

Table 1 Estimates of the Flypaper Effect

Study	Predicted by theory	Estimated	Error[a]
Total local government expenditures			
Gramlich-Galper (1973)	$0.03 \leq dE/dA \leq 0.05$	$dE/dA = 0.25$	\$0.20 - \$0.22
	$0.06 \leq dE/dA \leq 0.10$	$dE/dA = 0.43$	0.33 - 0.37
Inman (1971)[b]	$0.02 \leq dE/dA \leq 0.04$	$dE/dA = 1.00$	0.96 - 0.98
Ehrenberg (1973)[b]	$0 \leq \epsilon_{E,A} \leq 0.08$	$\epsilon_{E,A} = 0.22$	0.14 - 0.22
Study	**Predicted by theory**	**Estimated**	**Error[a]**
Education			
Feldstein (1975)[b]	$0 \leq \epsilon_{E,A} \leq 0.05$	$\epsilon_{E,A} = 0.21$	0.16 - 0.21
	$0 \leq \epsilon_{E,A} \leq 0.05$	$\epsilon_{E,A} = 0.06$	0.01 - 0.06
Inman (1971)[b]	$0 \leq \epsilon_{E,A} \leq 0.06$	$\epsilon_{E,A} = 0.71$	0.65 - 0.71
Ladd (1975)	$0 \leq \epsilon_{E,A} \leq 0.05$	$\epsilon_{E,A} = 0.03$	—
Inman (1978)	$0 \leq \epsilon_{E,A} \leq 0.06$ and $0 \leq \epsilon_{E,A} \leq 0.08$	$\epsilon_{E,A} = 0.23$ and $\epsilon_{E,A} = 0.40$	0.15 - 0.34
Olsen (1972)[b]	$0.02 \leq dE/dA \leq 0.04$	$dE/dA = 0.27$	\$0.23 - \$0.25
Weicher (1972)[b]	$0 \leq dE/dA \leq 0.001$	$0.41 \leq dE/dA \leq 0.58$	\$0.41 - 0.58
Gramlich-Galper (1973)	$0.01 \leq dE/dA \leq 0.02$	$dE/dA = 0.10$	\$0.08 - 0.09
Johnson (1979)[b]	$0.004 \leq dE/dA \leq 0.006$	$0.38 \leq dE/dA \leq 1.61$	\$0.37 - 1.60

a. Reported in cents per dollar of grant for studies measuring marginal effects and in points for studies measuring elasticities.
b. These works do not appear in this article's reference list. They can be found in Inman (1979) and Fisher (1982).
SOURCE: Used with permission from Fisher (1982). For references, see Inman (1979) and Fisher (1982).

absent, based on that study's estimate of the expenditure effects of income. The second column displays the actual effect of aid on expenditures, while the last column shows the discrepancy between the actual and predicted effects.

In the case of studies reporting marginal effects, the expenditure effect of lump-sum aid ranged from $0.20 to $1.60 larger than predicted by the theory. For those studies reporting elasticities, the expenditure effects were from zero to 71 percent larger than expected. As table 1 makes clear, although these effects are not ubiquitous (see, for example, Gramlich [1982]), the vast majority of studies support the idea that flypaper effects are significant and in need of explanation. Moreover, flypaper effects results occurred across a wide variety of data sets and empirical methodologies, as discussed below.

II. Previous Explanations of the Flypaper Effect

In examining the theoretical literature on flypaper effects, I begin with six conservative approaches. These six explanations, while modifying the theory briefly outlined above, retain the assumption that local expenditure decisions can be modeled as the choice of a single, rational decisionmaker such as the median voter. These studies blame flypaper effects on misinformation, arguing 1) that previous investigators have missed salient features of the problem in modeling the response of communities to grants-in-aid, or 2) that the median voter himself is mistaken about the effects of grants on his budget constraint.

Chernick (1979) and Fisher (1979) assert that previous analysts have classified much government aid as lump-sum although it does not properly belong in that category. Chernick notes that, if lump-sum aid is construed to include project grants, this money may represent the outcome of utility-maximizing decisions by the bureaucratic agency that administers the program. This creates two problems in estimating the effect of aid on expenditures.

First, the process of awarding grants appears to be influenced by the number and dollar amount of previous grant applications, so that actions of the community influence the amount of grants it receives. If these grant applications are correlated with community expenditures, a simultaneous equations bias exists in which expenditures affect aid and aid affects expenditures.

Second, in a more fundamental argument, Chernick says that grant determination is a complex process that involves the bureaucrat's utility benefit from additional expenditures in that community and the community's willingness to share in the costs of the new project. Therefore, both grant amounts and local expenditures are endogenous variables in the model; they are not related by any consistent function that can be compared to the effect of income on expenditure. Depending upon the level and rates of change of the truly exogenous variables in the model, any combination of grant and local expenditure levels can occur.

Fisher argues that, when lump-sum aid includes revenue sharing, the frequent inclusion of tax-effort factors into the distribution formula for this money creates what amounts to a price effect as well as an income effect on local government spending. A community's tax effort is usually defined as the compound fraction formed by taking the ratio of the community's tax revenue, divided by its tax base, to the tax revenue of the entire nation or state, divided by the tax base of this larger political unit.

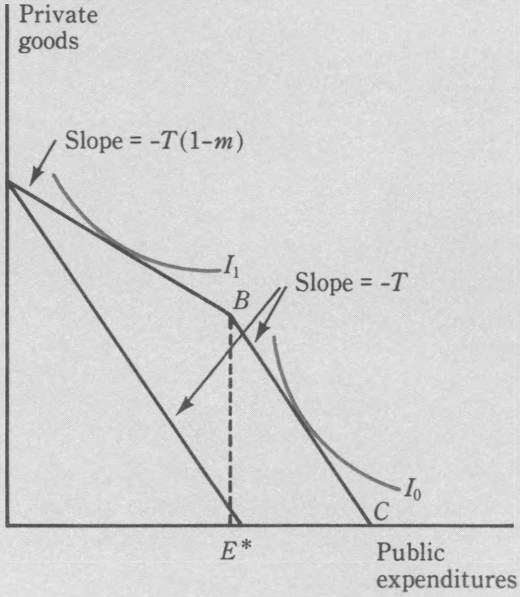
When such a factor is included in a revenue-sharing formula, it creates an incentive for local governments to raise taxes and expenditures in order to raise their tax effort and re-

ceive more aid from higher levels of government. In other words, the price of another unit of expenditure by the community is reduced by the effect of this spending on its tax effort and revenue-sharing collections. Because of this price effect, Fisher argues, we ought not to expect revenue sharing to have the same effect as an equivalent amount of private income.

In a related but more complex argument, Moffitt (1984) examines the role of closed-end matching grants on the budget constraint of the median voter. In many cases, these grants have been considered lump-sum aid on the grounds that, once the program's upper limit has been achieved, the cost of each additional unit of the good is unaffected by the grant.

This effect is shown in figure 2, which depicts the median voter's budget constraint with and without the program. When the community's expenditures are supplemented by the program, the slope of the voter's budget constraint is $-T(1-m)$, where m is the federal government's matching rate, up to some limit E^*. Above that level of expenditures, the grant amount remains unchanged, and the slope reverts to $-T$ (as in figure 1). For any community locating between B and C, the budget constraint is shifted by the program, but its slope remains the same.

Fig. 2 The Case of Closed-End Matching Grants



Moffitt argues that when the budget constraint becomes nonlinear, estimation becomes much more complicated and previous techniques yield biased results. For example, suppose that the functional form used in estimation implies a preference function that includes indifference curve I_0, but that communities have diverse preferences so that median voters in some cities have indifference curve I_1. Then the variation in preferences will be picked up by the error term. Notice, however, that the change in preferences implies a change in the equilibrium price faced by the voter so that the error term and the price variable are correlated. This contemporaneous correlation will lead to bias in the estimated coefficients. Moffitt also presents suggestive evidence (using a more sophisticated estimating technique, but employing an ad hoc demand equation to test for flypaper effects) that, in the case of AFDC grants, flypaper effects disappear when these nonlinearities are accounted for.

Hamilton (1983) believes that previous analysts were fooled because they failed to realize that, in many cases, private income represents both a pool of resources for consumption and a surrogate for certain unobserved factors in the production of local public goods. His case is strongest with respect to local education: not only does increased income in a community make possible increased spending on schools, but educational studies show that children from families with higher income and educational levels tend to learn more rapidly than other children. Thus, as income increases, expenditure increases may be held down by the fact that children from higher-income homes require fewer educational resources to achieve a given level of educational achievement. This effect will again cause lump-sum

aid to have a greater expenditure effect than income increases.

Courant, Gramlich, and Rubinfeld (1979) and Oates (1979) argue that it is the voter, and not the analyst, who is being fooled by the effect of intergovernmental grants.[2] Specifically, since the typical voter has little information about the extent of grants to his community, the voter estimates the unknown marginal cost of public goods using other known variables. By taking the ratio of his tax payments to total expenditures in the community, the voter can determine the average cost of public goods and use this as an approximation for their marginal cost. When lump-sum aid is present, however, the use of this proxy will cause the voter to err in his estimate of marginal cost. If the lump-sum aid is used to finance additional expenditures, total expenditure will increase while the median voter's tax payments will remain unchanged, thus driving down the average price of public goods and leading the voter to mistakenly demand more public goods. Because of this "fiscal illusion," these writers argue, lump-sum aid has a price as well as an income effect and we should not expect the aid to have an expenditure impact that is equivalent to the effect of an income increase.

In contrast to these six arguments, Romer and Rosenthal (1980) and Filimon, Romer, and Rosenthal (1982) insist that a more radical revision of the model is needed to explain flypaper effects. In these papers, the authors remove the median voter from his preeminent position in local decision-making and replace him with a bilateral monopoly model in which both the voter and a budget-maximizing bureaucracy are important actors. Flypaper effects occur, they say, because of the influence of this bureaucracy. This influence springs from the agencies' superior knowledge as compared to that of the median voter and/or the bureaucrats' ability to control the agenda of the decision-making process.

The "asymmetric information" model presented in Filimon, Romer, and Rosenthal is straightforward: the median voter is simply

2. It should be noted that Oates' model includes a budget-maximizing bureaucrat, and in that sense his model replaces rather than reforms the standard median voter model. However, the bureaucrat in this model derives his power solely from the voter's misperception of the marginal cost of local public goods. For that reason, I have included it in this section.

A Primer on Aid Types

Intergovernmental aid can be classified according to two criteria. The first involves restrictions placed on the recipient government about how the money is to be used. The second way of classifying aid is by determining how closely the amount of aid is tied to the recipient's expenditures. Grants are usually identified according to their positions along these two dimensions.

At one end of the spectrum of restrictions placed on recipient governments are *categorical grants*, which can be used only for a single, well-defined purpose. Federal grants for highways are of this type. Many categorical grants are of the *project grants* type, in which money is awarded for a specific undertaking (usually a capital project) at the discretion of the federal agency administering the program. Urban development action grants fit under this category. Somewhat less restrictive are *block grants*, which allow state and local governments to use aid for a broad class of activities. Examples include the federal government's community development block grant, social service block grant, and elementary and secondary education block grant. At the other end of this spectrum lies *general purpose aid*, which can be used for whatever the recipient government wants, including lowering taxes. *Revenue sharing* is an example of general purpose aid.

Along the second dimension, *matching aid* requires that the recipient government spend its own money as well as funds from grants on the aided goods. Typically, as in the aid to families with dependent children (AFDC) program, this takes the form of a cost-sharing arrangement; the federal government pays a percentage of program costs. Matching aid can be closed- or open-ended, depending upon whether the grantor government sets a ceiling upon the amount each recipient can receive (closed-ended), or if aid is available at the matching rate for whatever level of expenditures the recipient chooses (open-ended). At the opposite end of this dimension of grants is *lump-sum aid*, which is entirely independent of the expenditures of the recipient government. Revenue sharing is typically categorized as lump-sum aid, although strictly speaking it has some features of a matching grant if tax effort considerations are used in distributing these funds (see text). In this paper, the term lump-sum aid has also been used as shorthand for the more cumbersome term *lump-sum, general purpose aid.*

unaware of the presence of lump-sum grants in his community (even its impact on the average price of public goods) and the well-informed bureaucrat simply uses all the lump-sum aid for additional expenditures.

The "agenda control" model presented in both Romer and Rosenthal and in Filimon, Romer, and Rosenthal is more complex and more specialized. This model deals only with the case in which voters approve or disapprove local expenditures through a referendum, a situation which is not uncommon in local education. If the school board's request is not approved (and subsequent proposals are also turned down by the voters) the school district's expenditure will be set to a "reversion" level of spending, which is usually mandated by the state. The bureaucrat's power in this situation springs from his ability to determine what proposal, if any, is brought before the voters, who must choose between the board's request and the reversion level. For high reversion levels, the bureaucrat will bring forth no budget at all and will allow the state's reversion level to take effect. For very low (and hence unattractive to voters) reversion levels, the bureaucrat will propose the largest budget which will give the voter the same utility as the reversion level.

The comparative statics of this model are quite complex and depend critically upon the relationship of the reversion level of spending to the median voter's preferred level of spending. Under certain circumstances, however, the model will generate flypaper effects. Suppose for example that the reversion level is very large so that the bureaucrat simply accepts the reversion level. Then increases in income will have no effect on expenditures since it is the exogenous reversion level, not voter preferences, that determines spending. On the other hand, since most states require that aid be included in the reversion level, an increase in lump-sum aid increases spending by the full amount of the grant.

Thus, in this stylized example, a flypaper effect equal to the amount of the grant will occur (based upon the expenditure effect of income, the grant should have no effect on expenditure, but expenditure increases equal to the grant are observed). In other situations, in which the reversion is less than, or in the neighborhood of, the median voter's preferred level, flypaper and even anti-flypaper effects (income generating larger expenditure effects than grants) can occur, depending upon the nature of the voter's preference map.

III. A Critique of Previous Explanations

The explanations outlined above offer only limited descriptions of the flypaper effect that are confined to particular institutional situations, to particular kinds of grants, or to particular government services.

For example, Hamilton develops his argument that income is a proxy for inputs into the production of local public goods in a general way, but is able to offer examples only for local education and police protection. Romer and Rosenthal's "agenda control" model applies only to the case of local direct (not representative) democracy. Chernick's work applies only to project grants, not revenue sharing. Fisher's arguments apply only to revenue sharing that is distributed according to a tax effort formula. Moffitt's model is relevant only for closed-end grants, particularly those with more than one matching rate (such as AFDC) where the applicable rate depends upon the community's expenditures.

In a more subtle way, the fiscal illusion model and the "asymmetric information" model of Filimon, Romer, and Rosenthal are also limited; without further modification, they are confined to the institution of direct democracy. In these models, voters are misinformed about the fiscal situation facing their community and so make incorrect choices. But voters are typically represented by elected officials who know the extent of aid to their communities

(it is a prominent part of each annual budget) and who therefore know that marginal costs are unchanged by lump-sum aid. Moreover, since the decisions made by the voter in the fiscal illusion and asymmetric information models will be suboptimal, elected officials will have a political incentive (in order to maximize their chances of reelection) to both act on this information about the true cost of public goods and to release it to the general public.

For example, if voters would be happier with a smaller public sector and a reduction in local taxes, ambitious politicans have an incentive to give it to them. Thus, in a representative democracy, these models require one of two unpalatable modifications: either elected officials ignore even the most basic elements of their city's financial situation or political competition in the city has completely broken down.

The limited scope of these explanations contrasts sharply with the comprehensive nature of flypaper effects, which appear across a wide range of data sets, local public goods, and empirical methodologies. This means that, for every explanation given above, a study can be found that is beyond the scope of that argument but that still finds evidence of flypaper effects.

For example, Hamilton's hypothesis about income as an input leads to the conclusion that flypaper effects should occur primarily in education and public safety, but Gramlich and Galper (1973) report flypaper effects for social services (health and hospitals, and housing) and urban support (sewers, sanitation, highways, and parks and recreation) as well, while Inman (1971) reports additional flypaper effects for sanitation, sewers, parks and recreation, transportation, libraries, and welfare.

These two studies also carefully separate project grants from their lump-sum aid variable to obviate Chernick's arguments about the exogenous nature of project grants. In a similar way, Wyckoff (1984) removes all categorical grants of any kind from his lump-sum aid variable, thus ensuring that the arguments of Moffitt do not apply.

Since the subject of all the studies in table 1 was representative democracy, none of the arguments that rest on direct democracy (Romer and Rosenthal's agenda control model; Filimon, Romer, and Rosenthal's asymmetric information model; and the fiscal illusion model) are applicable. In addition, Fisher's tax effort considerations are probably not relevant to these results, since those studies took place before the onset of federal general revenue sharing and/or involved independent school districts that do not receive federal revenue sharing money.[3]

It is perfectly possible that flypaper effects are due to a combination of the theories just discussed, with each explanation being more important in a particular place and time. If this were the case, however, we might expect more variation as to the presence or absence of flypaper effects across empirical studies than illustrated in table 1. Without a unifying theory, we are forced to conclude that 10 out of the 11 studies in table 1 happened by chance to choose data sets and empirical techniques that led, through many distinct mechanisms, to flypaper effects.

While this multiple-cause explanation certainly cannot be ruled out, table 1 at least suggests that a more general explanation of flypaper effects might be useful, one which is not tied to a particular public service, institutional situation, or empirical specification. If such a theory existed, it would be easy to explain the consistencies noted in that table. For this reason, the next section summarizes a new attempt to explain flypaper effects, based on institutional features of government that, it is hoped, are more universal than the factors that underlie the explanations given above.

IV. A New Theory of Flypaper Effects

Wyckoff (1985) details a new model of flypaper effects, based upon two basic ideas. First, local public goods are produced by pub-

3. Fisher's point might continue to have some relevance because most states do have a program of revenue sharing or grants for general relief, although these programs are usually small in dollar value. Some of these programs include effort considerations.

lic employees (bureaucrats) whose interests do not always match those of the community. Second, this bureaucracy has influence over city council because it knows more about the true cost of producing public goods than the council does. Because of his or her professional training and day-to-day contact with these matters, the head of each department is assumed to have an advantage over council members in knowing both the production function for public goods (what inputs are needed for a particular level of output) and the minimum cost for these inputs.

To highlight the influence of these two notions, the model uses three simplifying assumptions. Local decision-making is assumed to be a simple two-way struggle between city council and a single, well-informed bureaucrat. Due to political competition, the preferences of city council are taken to accurately reflect those of the median voter in each community. Following Niskanen (1971), the bureaucrat is assumed to be solely interested in increasing the size of his budget, because this budget is systematically related to variables of direct interest to him: salary, fringe benefits, professional prestige, and power over others. Use of this third assumption means that the resulting model is an application and extension of Niskanen's model.

According to the public choice literature on bureaucracy, the bureaucrat's information advantage has an effect on public expenditure, allowing him to expand the city's budget beyond what the median voter would prefer. To increase his budget, the bureaucrat submits the largest request he thinks council will approve. In reviewing this request, city council is hampered by its lack of knowledge of the effects of marginal changes in the budget; since it doesn't know the true cost of public goods, it doesn't know what budget changes will mean in terms of changes in output. A risk-averse city council will therefore tend to avoid making changes in the bureau's budget request.

Moreover, an expansion-oriented bureaucrat will compound the council's timidity in making budget changes by acting strategically. Not only does the bureaucrat have no incentive to reveal correct information about the true cost of public goods, he will try to release distorted information and respond to budget cuts by cutting the most popular programs first ("cutting the meat instead of the fat"). Another budget-increasing tactic is to respond to council's tendency to cut all budget requests by a certain proportion by inflating requests so as to maintain desired spending levels even after allowance is made for token budget-cutting.

By using his information advantage this way, the bureaucrat in this simplified model will push the city council to the point where the median voter is indifferent between the budget that is finally approved and doing without the local public services (and the taxes that go to pay for them) entirely. This is a standard proposition of the Niskanen model. However, the local government case differs fundamentally from the central government case (the subject of Niskanen's study) because city residents have a stronger "exit" option (to use Hirschman's [1970] term) than do citizens of a nation. If he becomes dissatisfied with his community, the voter can always move.

Two standard comparative static results from the Niskanen model carry over to the model in Wyckoff (1985). First, the community's demand function for public goods, as filtered by negotiation with bureaucrats, will always be cost-elastic. Second, a dollar of lump-sum aid to this community will always generate more than a dollar of additional expenditures (for proofs of these two propositions, see Wyckoff [1984]).

Since it is set in the local context, however, the model has additional consequences that explain flypaper effects. The intuition behind these results is that the median voter's bargaining position with respect to the bureaucrat is not the same when he gets lump-sum aid as when he receives an increase in his private income.

When the voter receives an increase in private income, he can use this extra income

both in his present circumstances and in any alternative city he moves to. The increase in the income (and hence the utility) of the voter's next best alternative is of prime importance for the model: this effect leads to greater credibility in the voter's threat to leave if the bureaucrat goes too far. An increase in the value of the voter's alternative helps constrain the bureaucrat's demands and reduces the equilibrium size of the community's budget.

An increase in lump-sum aid, by contrast, improves the voter's current circumstances but cannot be moved to a new location with the voter—it is tied to his current city. Hence there is no corresponding increase in the value of the voter's threat to move in the case of an increase in intergovernmental aid. It is this asymmetry in bargaining position that creates flypaper effects.

The situation facing city council and the bureaucrat is similar to that facing the management of a company and its labor union. During labor negotiations, the wages and working conditions that are eventually agreed upon depend not only on current circumstances, but on each side's alternative situation if an agreement is not reached. For example, if management can creditably assert that it does not really need the plant due to, say, the possibility of filling orders from overseas production, then the perceived value of its next best alternative will be high, and it will be able to more effectively restrain the wage demands of the union.

To continue this analogy, consider management's bargaining position with respect to the union in two situations: 1) an increase in profitability in this one plant due to a reduction in the local price of materials; and 2) an increase in the profitability in the entire company due to a worldwide increase in demand for the product.

The former situation, which parallels the effect of lump-sum aid in the case of local governments, improves management's profit picture in the current situation (with this plant open) but not in any other situation (overseas supply). The latter situation, which is analogous to the effect of private income on local decision-making, increases management's profits in current as well as in alternative production schemes. Because management's threat to move production overseas is more credible in the latter situation than the former, workers will demand higher wage increases when the profit increase is localized to their own plant.

This new model of flypaper effects was tested using 1977 expenditure data from 115 small cities in Michigan. Using a single-equation, double-logarithmic functional form, expenditure was regressed on to population, the median voter's tax share, total income ($Z = Y + TA$), the share of total income from lump-sum aid (TA/Z), non-revenue-sharing aid, and several additional demographic variables.

In testing this bureaucratic model against the standard median voter model, a joint hypothesis test involving two coefficients was employed. First, the coefficient on population was included because of population's role in influencing the cost to the median voter of local public goods. Since the model retains the primacy of the median voter vis-à-vis other citizens in the local decision-making process (so that the preferences of other voters don't matter), if the median voter's tax share is held constant, the only effect of increasing population in a community is crowding of public facilities. If public goods are defined in terms of the resources available to each individual resident (for example, park space per capita), then, *ceteris paribus*, this crowding raises the cost of providing a uniform level of these goods to the median voter.

Second, the coefficient on the share of income from lump-sum aid was also utilized to test for the presence or absence of flypaper effects. If flypaper effects are absent, the composition of the median voter's income between private income and aid should have no effect on expenditures; the coefficient should be

zero. If flypaper effects are present, expenditures should increase with the share of total income coming from lump-sum aid.

Thus, under the bureaucratic model, demand must be cost-elastic and the coefficient of population on total expenditures must be negative. In addition, the coefficient on the share of income from lump-sum grants must be positive, reflecting flypaper effects. By contrast, under the median voter model, there is no restriction at all on the population coefficient, but the coefficient on TA/Z must be zero.

The regressions contained in Wyckoff (1985) show that, when operating expenditures only are the dependent variable, the bureaucratic model is rejected by the data, while the median voter model is not rejected. When capital expenditures are employed, the opposite is true: the median voter model is rejected by the data, but the bureaucratic model is not rejected.[4]

The results suggest that a dichotomy exists with respect to local governments' operating and capital expenditures: the bureaucrat has a great deal of influence on the latter and not much on the former. This is not an implausible result, since in the real world city councils may not be as helpless as portrayed in the simplified model above. Council members can often employ monitoring devices that, although costly in terms of time or money, yield information about bureau performance and the true costs of producing public goods. For example, strict budgeting and expense reporting techniques may be used, cost and output data can be compared with those of other communities, and feedback from citizens and the news media can be cultivated. It is entirely possible that these monitoring devices work well in one context but not in another. The complexity of capital expenditures, along with their ability to be financed by debt, may make it easier for the bureaucrat to press his demands there rather than in operating expenditures.

In addition, as pointed out by Nathan, Manvel, and Caulkins (1975), city councils may be more willing to accede to the bureaucrat's demands in the capital expense area because of a fear that revenue-sharing money might eventually be cut off by the federal government. Rather than using revenue sharing to fund new operating expenditures, which would have to be funded by increased taxes if revenue sharing was discontinued, local governments often chose to channel the revenue sharing money into one-time capital projects such as highway and sewer repairs.[5]

Moreover, the dichotomy of spending patterns between capital and operating expenditures observed in these cities suggests that the bureaucratic model may prove superior to the other explanations of flypaper effects discussed above, although no empirical tests of this hypothesis were undertaken. None of these previously mentioned theories suggest such a dichotomy. In fact, differences between current and capital expenditures are wholly inconsistent with many of these models. For example, if flypaper effects are caused by fiscal illusion, the voter ought to be fooled for both kinds of expenditures. If, on the other hand, fiscal effort provisions in revenue sharing are causing flypaper effects, these effects ought to show up in both capital and operating expenditures. And, finally, if bureaucrats are able to hide grants from voters, this should be registered in both types of spending.

V. Conclusions

The previous discussion ought to establish one important point: any evaluation of proposals to change the current system will be strongly influenced by our model of how the local public sector works. For example, proponents of the Reagan cutbacks have argued that reductions in aid to state and local government will be offset by the increases in state-and-local-government-taxable private income that results when tax and deficit burdens on the economy are reduced. Suppose for the sake of argument that private income does increase just enough so that, in the absence of flypaper effects, local expenditure in each community would be unchanged. If we accept the argu-

4. With regard to the restriction under the Niskanen model that a one dollar increase in lump-sum aid generates more than a one dollar increase in expenditures, this hypothesis applies only to total (current plus capital) expenditures. It may be worth noting, however, that the data appeared to fulfill this restriction of the model. Evaluated at sample medians, a one dollar increase in unrestricted aid generated an extra 56 cents of current expenditures and an increase of 75 cents in capital spending, for a total increase of $1.31.

5. The observations of Nathan, Manvel, and Calkins, however, do not by themselves constitute an explanation of flypaper effects. Although they explain why revenue-sharing money might be used for capital rather than operating expenditures, their arguments fail to show why the money is not used to reduce local taxes—why does the money stick in the public sector? If city councils are in charge of the budget and are responsive to the voters, this should not happen.

ments of Moffitt and Chernick that observed flypaper effects are due to the peculiarities of project grants and closed-end matching grants, the proposed cuts in revenue sharing (which does not share these unique features) will indeed be balanced by an appropriate increase in private income. According to the arguments of Filimon, Romer, and Rosenthal, of Romer and Rosenthal, and of Wyckoff, however, flypaper effects are endemic to the local decision-making process, and it would take very large increases in private income to offset the spending cuts caused by the loss of the revenue-sharing program.

The model of Hamilton, on the other hand, implies a subtle and interesting position on this question. Flypaper effects do occur, he acknowledges, and we ought to expect that the substitution of private income for intergovernmental aid will reduce total state and local government expenditure, but we ought not to conclude from this that the total output of the local public sector has declined. If income enters the local production function for public goods, then, even if purchased inputs (which is what is measured by the local budget) have declined, the increase in income may increase the (unmeasured) output of local public goods in the community.

Perhaps surprisingly, the model in Wyckoff does not have unambiguous public policy implications with regard to economic efficiency. Despite the bureaucrat's expansion of the local budget, the model does not show that the local public sector is either productively or allocatively inefficient in a welfare sense. Because the effective demand function for local public goods is always cost-elastic, the bureaucrat can only maximize his budget by operating at minimum cost, and hence there is no productive inefficiency (see Wyckoff [1984]). And although the budget is larger than the median voter would like, there is no reason to presume that what the median voter desires is allocatively efficient. In fact, two studies have argued that, if the median voter model is operating in the local public sector, the output of that sector is probably suboptimal (see Barlow [1970] and Bergstrom and Goodman [1973]).

The model does have predictions about the likely effects of a repeal of the revenue-sharing program and the political dimensions of such a move. First, as noted above, we ought to expect large cutbacks in state and local expenditures because of this change. Second, the chief opponents of such a cutback would not necessarily be the citizens of each state and local government, since the satisfaction of the median voter in each community is determined not by the amount of aid received by his or her state or local government, but by the utility of the voter's next best alternative community. The aid raises local expenditure levels without increasing his satisfaction with his current community. This result may help explain both the widespread discontent of citizens with state and local governments and the fact that the chief proponents of aid programs are often the employees and managers of these governments.

References

Arrow, Kenneth J. "A Difficulty in the Concept of Social Welfare," *Journal of Political Economy*, vol. 59, no. 4 (August 1950), pp. 328–46.

Barlow, Robin. "Efficiency Aspects of Local School Finance," *Journal of Political Economy*, vol. 78, no. 5 (September/October 1970), pp. 1028–40.

Bergstrom, Theodore C., and Robert P. Goodman. "Private Demands for Public Goods," *American Economic Review*, vol. 63, no. 3 (June 1973), pp. 280–96.

Black, Duncan. "On the Rationale of Group Decision Making," *Journal of Political Economy*, in Kenneth J. Arrow and Tibor Skitovsky, eds., *Readings in Welfare Economics*. Homewood, IL: Richard D. Irwin, Inc., 1969, pp. 133–46.

Borcherding, Thomas E., and Robert T. Deacon. "The Demand for the Services of Non-Federal Governments," *American Economic Review*, vol. 62, no. 5 (December 1972), pp. 891–901.

Bowen, Howard R. "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics*, vol. 58 (November 1943), pp. 27–48, in Kenneth J. Arrow and Tibor Skitovsky, eds., *Readings in Welfare Economics*. Homewood, IL: Richard D. Irwin, Inc., 1969, pp. 115–32.

Bradford, David F., and Wallace E. Oates. "The Analysis of Revenue Sharing in a New Approach to Collective Fiscal Decisions," *Quarterly Journal of Economics*, vol. 85, no. 3 (August 1971), pp. 416–39.

Chernick, Howard A. "An Economic Model of the Distribution of Project Grants," in Peter Mieszkowski and William H. Oakland, eds., *Fiscal Federalism and Grants-in-Aid*. Washington, DC: The Urban Institute, 1979, pp. 81–103.

Courant, Paul N., Edward M. Gramlich, and Daniel L. Rubinfeld. "The Stimulative Effects of Intergovernmental Grants: or Why Money Sticks Where It Hits," in Peter Mieszkowski and William H. Oakland, eds., *Fiscal Federalism and Grants-in-Aid*. Washington, DC: The Urban Institute, 1979, pp. 5–21.

Deacon, Robert T. "A Demand Model for the Local Public Sector," *Review of Economics and Statistics*, vol. 60, no. 2 (May 1978), pp. 184–92.

Ehrenberg, Ronald G. "The Demand for State and Local Government Employees," *American Economic Review*, vol. 63, no. 3 (June 1973), pp. 366–79.

Filimon, Radu, Thomas Romer, and Howard Rosenthal. "Asymmetric Information and Agenda Control: The Bases of Monopoly Power in Public Spending," *Journal of Public Economics*, vol. 17, no. 1 (February 1982), pp. 51–70.

Fisher, Ronald C. "Income and Grant Effects on Local Expenditure: The Flypaper Effect and Other Difficulties," *Journal of Urban Economics*, vol. 12, no. 3 (November 1982), pp. 324–45.

_____. "A Theoretical View of Revenue Sharing Grants," *National Tax Journal*, vol. 32, no. 2 (June 1979), pp. 173–84.

Gramlich, Edward M. "An Econometric Examination of the New Federalism," *Brookings Papers on Economic Activity*, 2:1982, pp. 327–60.

_____. "Intergovernmental Grants: A Review of the Empirical Literature," in Wallace E. Oates, ed., *The Political Economy of Fiscal Federalism*. Lexington, MA: Lexington Press, 1977, pp. 219–39.

_____, and Harvey Galper. "State and Local Fiscal Behavior and Federal Grant Policy," *Brookings Papers on Economic Activity*, 1:1973, pp. 15–58.

Hamilton, Bruce W. "The Flypaper Effect and Other Anomalies," *Journal of Public Economics*, vol. 22, no. 3 (December 1983), pp. 347–61.

Henderson, James M. "Local Government Expenditures: A Social Welfare Analysis," *Review of Economics and Statistics*, vol. 50, no. 2 (May 1968), pp. 156–63.

Hirschman, Albert O. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press, 1970.

Hotelling, Harold. "Stability in Competition," *Economic Journal*, vol. 39 (March 1929), pp. 41–7.

Inman, Robert P. "The Fiscal Performance of Local Governments: An Interpretative Review," in Peter Mieszkowski and Mahlon Straszheim, eds., *Current Issues in Urban Economics*. Baltimore, MD: Johns Hopkins University Press, 1979, pp. 270–321.

_____. "Testing Political Economy's 'As If' Proposition: Is the Median Voter Really Decisive?" *Public Choice*, vol. 33, no. 4 (1978), pp. 45–65.

_____. "Towards an Econometric Model of Local Budgeting," in National Tax Association, *Proceedings of the 64th Annual Conference on Taxation* (1971), Columbus, OH: National Tax Association, 1972, pp. 699–719.

Ladd, Helen F. "Local Education Expenditures, Fiscal Capacity, and the Composition of the Property Tax Base," *National Tax Journal*, vol. 28, no. 2 (June 1975), pp. 145–58.

Lovell, Michael C. "Spending for Education: The Exercise of Public Choice," *Review of Economics and Statistics*, vol. 60, no. 4 (November 1978), pp. 487–95.

Moffitt, Robert A. "The Effects of Grants-in-Aid on State and Local Expenditures: The Case of AFDC," *Journal of Public Economics*, vol. 23, no. 3 (April 1984), pp. 279–305.

Nathan, Richard P., Allen D. Manvel, and Susannah E. Calkins. *Monitoring Revenue Sharing*. Washington, DC: Brookings Institution, 1975.

Niskanen, William A. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton, Inc., 1971.

Oates, Wallace E. "Lump-Sum Intergovernmental Grants Have Price Effects," in Peter Mieszkowski and William H. Oakland, eds., *Fiscal Federalism and Grants-in-Aid*. Washington, DC: The Urban Institute, 1979, pp. 23–30.

Pack, Howard, and Janet Rothenberg Pack. "Metropolitan Fragmentation and Local Public Expenditures," *National Tax Journal*, vol. 31, no. 4 (December 1978), pp. 349–62.

Perkins, George M. "The Demand for Local Public Goods: Elasticities of Demand for Own Price, Cross Prices, and Income," *National Tax Journal*, vol. 30, no.1 (December 1977), pp. 411–22.

Romer, Thomas, and Howard Rosenthal. "An Institutional Theory of the Effect of Intergovernmental Grants," *National Tax Journal*, vol. 33, no. 4 (December 1980), pp. 451–58.

Wyckoff, Paul Gary. "Bureaucracy, Efficiency, and Local Public Choice: An Empirical Test of Some New Implications of Bureaucratic Public Service Provision," unpublished Ph.D. dissertation, University of Michigan, 1984.

_____. "A Bureaucratic Theory of Flypaper Effects," *Working Paper 8501*, Federal Reserve Bank of Cleveland, 1985.

The Federal Reserve Bank of Cleveland publishes an informative research periodical called Economic Commentary. *Following are the titles published since January 1985. If you are interested in receiving this publication, either future or back issues, please contact the Public Information Department, Federal Reserve Bank of Cleveland, P.O. Box 6387, Cleveland, OH 44101, (216) 579-2048.*

Economic Commentary

Ohio's Electricity Prices
Philip R. Israilevich
1/1/85

Will Adjustable Rate Mortgages Survive?
Thomas M. Buynak
1/15/85

Recent Changes in the Consumer Bankruptcy Laws
K.J. Kowalewski
2/1/85

Deposit Rates and Local Markets
Paul R. Watro
2/15/85

The 1985 Humphrey-Hawkins Testimony
Michael R. Pakko
3/1/85

Will Taxing Imports Help?
Michael F. Bryan and
Owen F. Humpage
3/15/85

Federal Reserve's Response to the Problems Experienced by ODGF Thrifts
Karen N. Horn
4/1/85

Imports and Domestic Steel
Amy Kerka
4/15/85

The Debt Burden: What You Don't See
John B. Carlson
5/1/85

CRR and Monetary Control
Michael R. Pakko
5/15/85

The Financial Distress in American Farming
Michael Bryan and Gary Whalen
6/1/85