# FORECASTING REGIONAL SHIFT USING RANDOM WALK MODELS

*Paul Kochanowski**

## Introduction

Time series modeling relies on the assumption that a time series inherently contains information that, once uncovered, will allow predictions of future values of that series. Such models depend on how present values of the series mirror past values (autoregressive models) and on how random components of the series behave over time (moving average models).

The simplest of these time series models is the random walk process. Under random walk, the changes in a time series (that is, the steps it takes from one period to the next) follow random and unpredictable directions; hence, the term *random walk*. This inherent unpredictability implies, in turn, a simple forecasting principle. If the expected value of the random movements equals zero, then the most efficient forecast of the value of a random walk series in t + 1 is its value in t. For example, if weather patterns follow a random walk, and if it rains today, then tomorrow's forecast based on the random walk model will call for rain. Of course, it may not rain tomorrow. But if weather patterns follow a random walk, forecasts made in this manner will predict, on average, the actual weather outcome more often than forecasts generated by any other method.

For many years models of random walk were confined to analyzing highly unstable processes such as stock and commodity prices. In a 1982 study Nelson and Plosser, however, found evidence that GNP and other macroeconomic time series also behave like random walks. These findings brought numerous attempts to see whether other economic processes might behave similarly. (Pindyck and Rubinfeld, 1991). But forecasting models based on random walk had surfaced in the regional forecasting literature even before the Nelson and Plosser study. Brown (1969), for instance, tested regional forecasting models that assumed the rate of change in a region's share followed a random process. Others challenged Brown's findings, offering alternative models. Stevens and Moore (1980) review many of these forecasting models.

The popularity of shift-share models based on random walk was brief, however. First, as Stevens and Moore conclude: "none of the models investigated [was] sufficiently accurate and dependable for

---

*Paul Kochanowski is a professor of economics, Indiana University South Bend. The author would like to thank Stanley Keil and two anonymous referees for many helpful suggestions.

policy and planning purposes."[1] (1980, p. 433) Second, new, presumably more accurate, econometric models came on the regional forecasting scene. These new models assumed a stable equilibrium relationship between regional and national economic variables. For more than a decade they dominated the scene, seemingly closing the door on random walk as a regional forecasting approach. A paper by Brown, Coulson, and Engle (1991), however, reopened the debate.

Brown, Coulson, and Engle found that state shares of national employment, payroll, or output in such sectors as manufacturing, retail, and services did not exhibit tendencies to approach any long-run equilibrium values. This implied that state shares were nonstationary and that once a state's share was altered, it remained permanently altered. Brown, Coulson, and Engle's results further indicated that because most regional econometric models implicitly assumed such equilibria existed, these models would provide inconsistent estimates and poor forecasts. As a consequence, Brown, Coulson, and Engle retested disequilibrium models. Their results showed conclusively that in postsample forecasts, random walk models forecasted more accurately than did multiparameter equilibrium models.

This study attempts to build on the Brown, Coulson, and Engle analysis in two ways. First, it applies the Brown, Coulson, and Engle shift-share disequilibrium models to 11 small MSAs in Indiana ranging in employment size from about 50,000 to 250,000 to determine whether Brown, Coulson, and Engle's findings for states generalize to smaller areas. Second, the study recasts the econometric and mathematical controversy raised by Brown, Coulson, and Engle into a debate over two economic views—the views of the neoclassicists and the views of those termed the neoskeptics. Neoclassicists, whose models dominated the regional forecasting literature during the 1970s and 1980s, maintain that over time the growth rate for employment in each sector of a region approaches the growth rate of that sector for the nation. For this hypothesis to be valid, a region's long-run share of output, employment, and payroll must remain constant. Neoskeptics, on the other hand, point out that the equilibrium mechanisms required for the neoclassical result often fail to materialize, that regions often grow for extended periods at rates far different from those of the nation, and

[1] In part the inaccuracy probably arose from the long forecast intervals over which the models were tested. Brown, for example, using SMSA data forecasted for intervals of 1947-1954, 1954-1958, and 1958-1963. Intervals of this length place a random walk model at a decided disadvantage. A forecast that $y_t = y_{t-1}$ will be more accurate over short intervals such as quarters than over longer intervals such as several years.

that shocks to a region's economy often permanently alter that region's share of national output or employment.

The results of this study tend to confirm the neoskeptic view that stable equilibrium relationships between regional employment and national employment in the same sector do not exist. Brown, Coulson, and Engle's conclusion that state shares of national industries do not exhibit tendencies to approach long-run equilibrium values appears equally valid for smaller regions such as the MSAs in Indiana. Furthermore, the results indicate that for small Indiana MSAs, random walk models of regional shift generally outperform other more complex models.

## Background

Regional forecasting models have developed along two fronts, fronts that are indicative of two different views of reality. The neoclassicists, whose models dominate the forecasting scene, believe that external shocks to a region's economy cannot alter a region's relative position permanently. This belief stems from two propositions. First, the long-run relative position of a region depends on locationally fixed inputs of the region in contrast to other regions (i.e., supply side factors). And second, mechanisms exist to insure the stability of the long-run equilibrium. Demand and supply shocks may alter a region's share of output temporarily, but mechanisms exist to bring that region's share back to its long-run value. Given this view of reality, neoclassicists favor self-correcting autoregressive models. Such models take the general form:

(1) $y_t = \beta_0 + \beta_1 y_{t-1} + \Sigma\beta_i x_{it}$

where:

x's = Exogenous variables.

If $\beta_1 < 0$, then its long-run equilibrium value remains unaltered regardless of what happens to $y_t$ in the short run. This, Brown, Coulson, and Engle write, is the standard analysis which

> presumes that factor price adjustment and factor migration ... equalize ... prices across regions [thereby] forcing an interregional equilibrium wherein all regions would grow at the same national rate in the long run. A key by-product of such a model's steady state is that the share of national output, employment and earnings generated in the region will remain constant over time. (Brown, Coulson, and Engle, 1991, p. 2)

The alternative view held by many practicing regional forecasters rejects the neoclassical position. These neoskeptics note, for instance, that in the 1960s and 1970s shocks to the manufacturing sector of the upper midwest region appear to have lowered that region's share of manufacturing permanently. Moreover, many of these forecasters (relying on the shift-share model and their own forecasting experience) believe that the best forecast of a region's future share of output, employment, or earnings is its current share. A model based on the proposition that the best forecast of tomorrow is today obviously economizes on data, but it also provides optimal forecasts if a region's share does follow a random walk.

To see the elements of this disagreement more clearly, consider the following model of regional employment. Let the production function for the jth industry in region i be given as:

$$(2)\ Q_{it} = A_t F_i L_{it}^{\beta}$$

where:

A = A productivity shift factor independent of region;
$F_i$ = The locationally fixed factor;
$L_i$ = Employment; and
$\beta$ = The output elasticity of employment.

Converting this production function to logs yields:

$$(3)\ q_{it} = a_t + f_i + \beta l_{it}.$$

Moreover, the first order condition for profit maximization with respect to the variable factor labor is given as:

$$(4)\ w_{it} - P_{it} = \log\beta + qit - l_{it}$$

where:

$w_{it}$ = The log of the wage rate; and
$P_{it}$ = The log of product j's price.

In a small region such as an MSA, the most observable quantity is labor. Combining equations (3) and (4) gives the log of labor as:

$$(5)\ l_{it} = \frac{1}{1-\beta}[\log\beta - w_{it} + p_{it} + f_i + a_t].$$

Equation 5 indicates that labor's growth rate in industry j in region i depends on the growth rate of wages, product price, fixed factors, and productivity. Now define the ith region's share of labor in the jth industry as $S = L_i/L_k$ and use the result given by equation (5) for both regions i and k to determine the log share of region i relative to region k as:

$$(6)\ l_{it} - l_{kt} = \frac{1}{1-\beta}\{(w_{kt} - w_{it}) + (p_{it} - p_{kt}) + (f_i - f_k)\}.$$

Thus, in the short run region i's share of employment in industry j depends on relative wages and prices and differences in the level of the fixed input.

The neoclassicists argue that although it is possible for the price of a good to differ from one region to another in the short run, these prices should be equal in the long run. Similarly, through labor migration or the factor price equalization theorem, wage rates should be equalized. Moreover, as noted by Brown, Coulson, and Engle, the neoclassicist conclusions are robust to several variations. For instance, the inclusion of other variable factors of production such as capital or more disaggregated types of labor or the inclusion of cost of living differences into the equation do not alter the basic conclusion that the long ratio of labor in the two regions will be a constant.[2] Consequently, in the long run, equation (6) reduces to:

$$l_{it} - l_{kt} = \frac{1}{1-\beta}\{(f_i - f_k)$$

thereby indicating a constant long-run share dependent entirely on locationally fixed input differences.

Those who see the neoclassicist view as flawed (such as the neoskeptics), point out, however, that wages and other input prices may not have long-term tendencies to equalize across regions and may be nonstationary, in effect causing a region's log share to be nonstationary. Furthermore, they note that even if such prices and factor costs do equalize, a shock to region i's economy may alter the productivity of its fixed input permanently, thereby changing $(f_i - f_k)$. For example, the value of a fixed advantage such as access may be altered by changing migration patterns, transport technology, and perhaps types of goods consumers purchase. The neoskeptics finally point out that the output elasticity of labor may not be constant across regions and potentially moves in a nonstationary fashion.

[2]Where long-run cost of living differences do persist, the long-run share, while constant, will depend on both the relative amount of fixed factors and the relative cost of living differences.

As so often happens in economics, such disputes cannot be settled at the theoretical level. The remainder of this paper describes the use of modern time series methods to test the constant share hypothesis and to use the results of those test to design regional forecasting models.

**Statistical Test of Long-Run Equilibrium Shares**

The basic point of contention is the following: does a region's share of the jth industry tend to revert to some long-run equilibrium following a shock, or does it follow a random walk and remain permanently altered? The neoclassicists insist that it reverts. Their detractors are not sure. The statistical test for a long-run equilibrium share relies on the regression:

(7) $l_{it} = \hat{\beta}_0 + \hat{\beta}_1 l_{kt} + e_t$.

If a long-run equilibrium does exist, the slope coefficient should be one, with the intercept giving an estimate of the log share. For the estimates in this regression to be statistically consistent, however, requires that the linear combination

$l_{it} = \hat{\beta}_0 + \hat{\beta}_1 l_{kt}$

generates a stationary error term, $e_t$. The equation for this error term

(8) $e_t = \rho e_{t-1} + \varepsilon_t$

implies, in turn, that stationarity and a long-run equilibrium share will exist only if $\rho < 1$, because only then will the impact of any shock on $l_{it}$ become smaller with time. Engle and Granger term a regression that provides consistent estimates of a long-run equilibrium (if one exists), a *cointegrating regression.* Dickey and Fuller provide tests of the hypothesis that $e_t$ is stationary. (Pindyck and Rubinfeld, 1991, pp. 465-467) Hence, verification of the hypothesis that $l_{it}$ and $l_{kt}$ are cointegrated thus supports the neoclassical view.

To test for cointegration, quarterly employment data were gathered for 11 MSAs in Indiana, ranging in employment size from 50,000 to 250,000, on major employment categories such as total; manufacturing; transportation, communication, and utilities; trade (wholesale and retail); and services. These databases represent all of the non-Indianapolis MSAs for which data were available (that is to say, the smaller Indiana MSAs such as Anderson (1981.1), Bloomington (1984.1), Elkhart-Goshen (1981.1), Evansville (1970.1), Fort Wayne (1960.1), Gary-Hammond (1969.1), Lafayette (1981.1), Muncie

(1960.1), New Albany (1960.1), South Bend-Mishawaka (1960.1), and Terre Haute (1960.1).[3] The date in parentheses indicates the beginning quarter available for each MSA. All models were estimated through the end of 1988, with forecasts then made from 1989.1 to 1990.4.

The results given in Table 1 strongly confirm the findings of Brown, Coulson, and Engle. They indicate that MSA employment sectors are not cointegrated with national series.[4] Lack of cointegration seems as prevalent, if not more so, for smaller regional economies as it is for larger ones.

The lack of evidence of cointegration invalidates the neoclassical view of a long-run steady state where all regions grow at the same rate as the nation and have constant shares of employment. It furthermore suggests that shocks to a region's share of employment in a sector are likely to be permanent. This has important implications. Regional policies such as tax or subsidy policies designed to impact particular industries or local shocks such as environmental regulations or infrastructure investment will have persistent effects (Brown, Coulson, and Engle, 1991, p. 11). Regional forecasting models that can be designed to take into account the random, nonstationary characteristics of regional shares will forecast more accurately.

**Shift-Share Forecasting Models**

The cointegration versus noncointegration controversy fits readily into the context of the general shift-share model. Letting upper case letters be actual values and lower case letters be their logs, employment in region i in industry j is given as:

(9) $L_{it} = L_{US,t} S_t$

where:

$S_t$ = Share the ith region has of U.S. employment in industry j.

It follows then that

---

[3]Analyses and findings similar to those presented in this paper have been found by the Indiana University Center for Econometric Model Research for the Indianapolis MSA.

[4]Before a test of cointegration (in the Engle-Granger sense) can proceed, the involved variables must be integrated of order one (i.e., I(1)). Preliminary investigation of the data indicated that (with two exceptions: retail employment for Anderson and manufacturing employment for Lafayette) the hypothesis of a unit root could not be rejected for any of the MSA employment variables. Nor could the hypothesis of a unit root be rejected for any of the national employment series.

(10) $l_{it} = l_{us,t} + s_t$

(11) $(l_{it} - l_{it-1}) = (l_{us,t} - l_{us,t-1}) + (s_t - s_{t-1})$

(12) $\Delta l_{it} = \Delta l_{us,t} + \Delta s_t$.

Equation 12 indicates that the growth rate of employment in the ith region in industry j equals the growth rate of employment in the jth sector in the U.S. (the share effect) plus the change in the log share (the shift effect). The neoclassicists argue that over longer periods $\Delta s_t$ equals zero, so that each region's employment in industry j grows at the rate of the national industry. The cointegration tests given above show, however, that $s_t$ does not approach an equilibrium value, implying that over the long run $\Delta s_t$ is not equal to zero. Thus, a region's growth may be persistently higher or lower than that of the nation.

Those who traditionally have ascribed to the nonequilibrium view of regional growth often have forecasted regional growth by forecasting $\Delta s_t$. Several such models of regional shift exist in the literature and are presented by Brown, Coulson, and Engle.[5] The nine models given below make use of the following definitions:

- The log share of region i for industry j: $s_{it} = l_{it} - l_{us,t}$;
- The regional shift for region i and industry j: $RS_t = \Delta s_t$;
- The error in predicting the log share: $\varepsilon_t = s_t - E_{t-1}s_t$;
- The growth rate of industry j in the U.S.: $IG_t = \Delta l_{us,t}$

Model 1: The Martingale share model implies that the share follows a random walk.

$E_{t-1}[RS_t] = 0$

$\Delta s_t = \varepsilon_t$

Model 2: The Martingale shift model implies that the shift follows a random walk.

$E_{t-1}[RS_t] = RS_{t-1}$

$\Delta^2 s_t = \varepsilon_t$

[5]The nine models Brown, Coulson, and Engle test are general forms of models found in the regional science literature. Brown (1969) and Hellman (1978) provide extensive tests of model 2, the Martingale shift. In addition, the substate models found in Shapiro and Fulton (1985), although taking into account other factors ignored by Brown, Coulson, and Engle, can be written in the general forms of models 5 and 6.

Model 3: The constant shift model implies that the share is a random walk with a drift.

$E_{t-1}[RS_t] = \alpha_0$

$\Delta s_t = \alpha_0 + \varepsilon_t$

Model 4: The autoregressive shift model implies that the expected regional shift in time period t is a function of the regional shift in t - 1. Model 4 includes models 1, 2, and 3 as special cases.

$E_{t-1}[RS_t] = \beta_0 + \beta_1 RS_{t-1}$

$\Delta s_t = \beta_0 + \beta_1 \Delta s_{t-1} + \varepsilon_t$

Model 5: The autoregressive employment model implies that the regional economy is independent of the national economy.

$E_{t-1}[RS_t] = \zeta_0 + \zeta_1 l_{it-1} - lG_t$

$l_{it} = \zeta_0 + (\zeta_1 + 1) l_{it-1} + \varepsilon_t$

Model 6: The constant short-run elasticity model implies a non-unit impact elasticity of national employment on regional employment.

$E_{t-1}[RS_t] = \eta_0 + \eta_1 lG_t$

$\Delta s_t = \eta_0 + \eta_1 \Delta l_{us,t} + \varepsilon_t$

Model 7: The autoregressive share model implies a long-run equilibrium share.

$E_{t-1}[RS_t] = \theta_0 + \theta_1 s_{t-1}$

$\Delta s_t = \theta_0 + \theta_1 s_{t-1} + \varepsilon_t$

Model 8: The unit elastic error correction model combines features of models 6 and 7 and implies a correction factor in the form of $s_{t-1}$.

$E_{t-1}[RS_t] = \kappa_0 + \kappa_1 lG_t + \kappa_2 s_{t-1}$

$\Delta s_t = \kappa_0 + \kappa_1 \Delta l_{us,t} + \kappa_2 s_{t-1} + \varepsilon_t$

Model 9: The error correction model generalizes the unit elasticity error correction model given in model 8 to allow non-unit long-run elasticity and implies increasing or decreasing comparative advantage.

$E_{t-1}[RS_t] = \pi_0 + \pi_1 l_{it-1} + \pi_2 lG_t + \pi_3 s_{t-1}$

$$\Delta s_t = \pi_0 + (\pi_1 + \pi_3) \{l_{it-1} - \frac{\pi_3}{\pi_1 + \pi_3} l_{us,t-1}\} + \pi_2 \Delta l_{us,t} + \varepsilon_t.$$

The nine models fall into two categories—noncointegrating versus cointegrating—depending on whether they permit equilibrium solutions for $s_t$. Except for model 5, which makes the local economy independent of the nation, thereby precluding cointegration, and models 2 and 4, which are second order difference equations, the models take the general first order difference equation form:

$$s_t + as_{t-1} = c$$

with the time path for $s_t$ given as:

$$s_t = (s_0 - \frac{c}{1+a})(-a)^t + \frac{c}{1+a}.$$

Since $c/(1 + a)$ equals the steady state equilibrium share, $s^*$, the time path for s can then be rewritten as:

$$s_t = (s_0 - s^*)(-a)^t + s^*.$$

Based on this equation, a region's share will converge to the equilibrium value $s^*$ if and only if $|a|$ is less than 1. By checking the coefficient on $s_{t-1}$ in each model, the model's convergence to an equilibrium share, as well as whether local and national employment series are cointegrated, can be ascertained. Because the coefficient on $s_{t-1}$ is equal to 1 in models 1, 3, and 6, these models have no long-run equilibrium shares and hence are disequilibrium, noncointegrative models. Similarly, Brown, Coulson, and Engle assume that the absolute value of the coefficients on $s_{t-1}$ in models 7, 8, and 9 is less than 1, thereby implying long-run equilibrium values of $s_t$ and the cointegration of local and national employment series. Model 5 precludes any equilibrium share because it does not contain national employment. Nonetheless, if the absolute value of the coefficient on $l_{it-1}$ is less than 1, the local area attains a long-run equilibrium value of industry employment.[6]

---

[6]Because models 2 and 4 are second order difference equations whose dynamic properties depend on the characteristic roots of a quadratic equation, they are more complex to analyze. In general, however, the time path will converge to an intertemporal equilibrium if and only if the absolute value of every root is less than 1. Model 2 has only one characteristic root equal to one and hence has no equilibrium for $s_t$. Model 4 has two characteristic roots given as:

$$b_1, b_2 = \frac{(1+\beta) \pm \sqrt{(1+\beta)^2 - 4\beta}}{2}$$

which are dependent on the value of beta.

Cointegration models 7, 8, and 9 are included for comparison purposes. Brown, Coulson, and Engle show that cointegrative models tend to perform better than noncointegrative models within the sample period but perform worse in postsample forecasts. Moreover, largely because of the dominance of the neoclassicists' view, the specifications implied by models 7, 8, and 9 are found in many regional forecasting models. For example, the regional models designed by Shapiro and Fulton (1985) for MSAs in Michigan are of the cointegrative type. Thus, comparisons such as those made by Brown, Coulson, and Engle and those made here, which demonstrate such models are inferior forecasting tools, can help convince regional forecasters to look more closely at the disequilibrium view of regional economic development.

To test their in-sample and out-of-sample forecasting ability, each of the nine models given above is estimated for the MSA-industry combinations. The in-sample estimation results are summarized in Table 2. In each case, the log of employment in each sector for each MSA is first predicted from the equation

(13) $l_{it} = l_{it-1} + lG_t + RS_t$,

and the antilog of the predicted value of $l_{it}$ is compared to the actual employment level. From these values, root mean squared percentage errors are computed.

The in-sample results presented in Table 2 are similar to those found by Brown, Coulson, and Engle. Larger models, such as models 8 and 9, with three and four parameters, respectively, are somewhat better than the smaller disequilibrium models. Close behind are models 4 and 6, which are noncointegrative models with two parameters. Models 1, 3, and 5 are in the next group with zero, one, and two parameters, respectively. Models 2 and 5 have the lowest in-sample accuracy.

To test the models' postsample forecasting accuracy, the nine models are used to forecast employment in each major sector of each MSA. Forecasts are for eight quarters and are of two types. First, eight postsample one-step-ahead forecasts are made using all available information at each step. Second, one through eight step postsample forecasts are made using all available information through the estimation period and forecasted values whenever needed during the forecasts. Both forecasting methods use actual values of $lG_t$. For example, the forecasting equation implied by model 1 (the Martingale share) is:

(14) $\hat{l}_{it} = l_{it-1} + lG_t$.

In the one-period-ahead model, actual values of $l_{it-1}$ are available for each of the eight forecasts. In the eight-step-ahead forecast only the first quarter's forecast uses the actual value of $l_{it-1}$, with the remaining seven quarters using forecasted values of $l_{it-1}$. Actual values of $IG_t$ are used in both cases. Forecasting errors in the postsample period are calculated in the same manner as for the in-sample forecasts: the antilog of predicted employment is compared to actual employment, and root mean squared percentage errors are computed.

The postsample forecasting results are found in Tables 3 and 4. While the in-sample forecasting accuracy lends support to the neoclassical cointegration models 7, 8, and 9, the postsample results reinforce the disequilibrium view. With one exception, cointegration models 7, 8, and 9 have larger postsample forecasting errors than the fewer parameter disequilibrium models. The Martingale share random walk model with zero estimated parameters outperforms any of the other models in virtually every case.[7] Moreover, disequilibrium model 4, which includes models 1, 2, and 3 as special cases, generally outperforms the cointegration models.[8]

The results in Tables 3 and 4 not only demonstrate that simple random walk models are superior but also show that such models generate reasonably accurate forecasts. The errors given in Tables 3 and 4 are conditioned on the accuracy of the national industry forecasts. Nonetheless, given the difficulty of forecasting for a small region, conditional forecasting errors in the 2 percent to 3 percent range for total employment and for manufacturing seem acceptable. Moreover, the relatively higher forecasting errors for the service sector probably imply that shift-share models are inappropriate for this sector. Services (retail also) perhaps should be modeled as a function of MSA total employment rather than as a function of national growth rates and regional-national shifts. Treating such sectors as endogenous to the local region thereby may improve forecasting accuracy.

## Conclusions

This study has attempted to evaluate the accuracy of two different views of regional economic change as these views pertain to the accuracy of regional forecasting models. The findings contradict the view

---

[7]Model 9 outperforms the Martingale share model for the multiperiod forecasts of transportation, communications, and utilities. Possibly, the nature of this industry, which includes regulated industries, accounts for this finding.

[8]Of the nine models, the Martingale shift model has the largest forecasting errors. A region's share may be altered permanently and follow a random walk, but it is not the case that the rate of change of that share has the same properties.

held by the neoclassicists that over the long run a region's industries will grow at the rate of those industries nationally and that a region's long run share of a national industry will be a constant. Overwhelmingly, the statistical tests performed on 11 Indiana MSAs support the alternative view that the process of regional change is inherently unstable and more likely will result in a permanent disequilibrium than in any long-run steady state. The findings here have important implications for forecasting small regional economies. In the highly unstable regional economic environment, highly parameterized models often will fail to forecast well. This study finds that starting from the shift-share model

(14) $\Delta l_{it} = \Delta l_{us,t} + RS_t$

the most accurate forecast of the growth rate of labor in sector j of region i results from forecasting $RS_t$ as a random walk process. Thus, the findings in this study reinforce the conclusion of Brown, Coulson, and Engle that because of "its simplicity, ease of computation, and its ability to forecast, [the Martingale share model] becomes highly recommended, both as a forecasting device ... and as a description of the temporal process of share" (p. 22).

## References

1. Brown, S.J., N.E. Coulson, and R.F. Engle, "Non-Cointegration and Econometric Evaluation of Models of Regional Shift and Share," National Bureau of Economic Research, Inc., working paper no. 3291 (1991).

2. Brown, H., "Shift Share Projections of Regional Growth," *Journal of Regional Science*, 9 (1969), pp. 1-18.

3. Engle, R., and C. Granger, "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55 (1987), pp. 251-276.

4. Hellman, D., "Shift-Share Models as Predictive Tools," *Growth and Change*, 7 (1978), pp. 3-8.

5. James, F., and J. Hughes, "A Test of Shift and Share Analysis as a Predictive Device," *Journal of Regional Science* (1973), pp. 223-231.

6. Pindyck, R.S., and D.L. Rubinfeld, *Econometric Models & Economic Forecasts*, 3rd edition (New York: McGraw-Hill, Inc., 1991).

7. Shapiro, H.T., and G.A. Fulton, *A Regional Econometric Forecasting System: Major Economic Areas of Michigan* (Ann Arbor: The University of Michigan Press, 1985).

8. Stevens, B., and C. Moore, "A Critical Analysis of the Literature on Shift-Share as a Forecasting Technique," *Journal of Regional Science*, 20 (1980), pp. 419-438.

**Table 1—Cointegration Test of MSA and National Employment**

| Model | Total | Mfg. | T.C.U. | Trade | Service |
|---|---|---|---|---|---|
| AN | 2.1/3.5 | 0.8/3.5 | 1.8/3.5 | **4.2/3.5** | 3.0/3.5 |
| BL | **3.8/3.6** | 0.8/3.6 | **4.4/3.6** | 1.7/3.6 | 0.7/3.6 |
| EL | 2.1/3.5 | 1.5/3.5 | 1.9/3.5 | 3.1/3.5 | 2.2/3.5 |
| EV | 2.0/3.4 | 2.5/3.4 | 2.0/3.4 | 1.7/3.4 | 2.4/3.4 |
| FW | 1.6/3.4 | 2.2/3.4 | 1.8/3.4 | 1.2/3.4 | 1.7/3.4 |
| GH | 1.7/3.4 | 0.8/3.4 | **4.2/3.4** | 1.9/3.4 | 2.1/3.4 |
| LA | **5.6/3.5** | 0.2/3.5 | 1.8/3.5 | 2.7/3.5 | 0.1/3.5 |
| MU | 2.0/3.4 | 1.0/3.4 | 0.6/3.4 | 1.6/3.4 | 3.1/3.4 |
| NA | 1.9/3.4 | 1.6/3.4 | 2.0/3.4 | 3.2/3.4 | 1.2/3.4 |
| SB | 1.4/3.4 | 1.4/3.4 | 1.9/3.4 | 1.6/3.4 | 2.1/3.4 |
| TH | 1.6/3.4 | 0.7/3.4 | 1.7/3.4 | 1.8/3.4 | 1.6/3.4 |

All of the tests reported above are performed with the Micro TSP package's time series test UROOT. The first number in each pair of numbers represents the Dickey-Fuller t-statistic (absolute value) on the coefficient of the lagged residual from the regression of the MSA employment against national employment (both in logs). The second number represents the absolute value of the MacKinnon critical value at the 5 percent level. Under the hypothesis that the series are not cointegrated and that there exists a unit root in the residuals, the expected value of the Dickey-Fuller t-statistic is zero. Cases where cointegration exist are shown in bold type (i.e., cases where the Dickey-Fuller t is larger than the critical value)

AN = Anderson;
BL = Bloomington;
EL = Elkhart;
EV = Evansville;
FW = Fort Wayne;
GH = Gary-Hammond;
LA = Lafayette;
MU = Muncie;
NA = New Albany;
SB = South Bend;
TH = Terre Haute.

**Table 2—In-Sample Forecasting Errors Across MSAs and Sectors**

| Model | Total | Mfg. | T.C.U. | Trade | Service |
|---|---|---|---|---|---|
| 1 | 2.7 | 3.8 | 3.3 | 2.2 | 2.1 |
| 2 | 4.2 | 5.5 | 4.9 | 3.0 | 2.9 |
| 3 | 2.6 | 3.7 | 3.3 | 2.1 | 2.0 |
| 4 | 2.5 | 3.7 | 3.3 | 2.1 | 2.0 |
| 5 | 2.7 | 4.1 | 3.5 | 3.2 | 2.4 |
| 6 | 2.6 | 3.7 | 3.2 | 2.0 | 1.9 |
| 7 | 2.5 | 3.6 | 3.2 | 2.1 | 2.0 |
| 8 | 2.5 | 3.6 | 3.1 | 2.0 | 1.8 |
| 9 | 2.4 | 3.6 | 3.1 | 1.9 | 1.8 |

**Table 3—Postsample Forecasting Errors Across MSAs and Sectors, One Step Ahead for Eight Quarters**

| Model | Total | Mfg. | T.C.U. | Trade | Service |
|---|---|---|---|---|---|
| 1 | 2.18 | 1.55 | 2.89 | 1.63 | 1.92 |
| 2 | 3.89 | 2.12 | 3.84 | 2.37 | 2.34 |
| 3 | 2.22 | 1.79 | 2.86 | 1.67 | 1.90 |
| 4 | 2.06 | 1.68 | 2.91 | 1.72 | 1.99 |
| 5 | 2.14 | 1.91 | 3.30 | 2.57 | 2.24 |
| 6 | 2.26 | 1.82 | 3.09 | 1.58 | 1.93 |
| 7 | 2.41 | 1.91 | 3.14 | 1.99 | 2.10 |
| 8 | 2.61 | 1.99 | 3.18 | 1.91 | 2.08 |
| 9 | 2.38 | 1.92 | 3.57 | 2.12 | 2.47 |

**Table 4—Postsample Forecasting Errors Across MSAs and Sectors, Eight Step Ahead for Eight Quarters**

| Model | Total | Mfg. | T.C.U. | Trade | Service |
|---|---|---|---|---|---|
| 1 | 2.43 | 2.84 | 7.39 | 2.54 | 4.65 |
| 2 | 28.94 | 12.46 | 18.08 | 14.70 | 20.06 |
| 3 | 2.96 | 4.34 | 7.49 | 2.88 | 4.62 |
| 4 | 2.69 | 3.69 | 7.58 | 2.93 | 4.64 |
| 5 | 2.99 | 4.44 | 7.57 | 3.07 | 4.58 |
| 6 | 2.79 | 4.10 | 7.23 | 2.60 | 4.55 |
| 7 | 2.61 | 3.97 | 7.76 | 2.81 | 4.88 |
| 8 | 2.99 | 4.42 | 7.77 | 2.88 | 4.84 |
| 9 | 2.56 | 3.90 | 6.40 | 3.17 | 5.46 |