



FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 97-28

REGIONAL EMPLOYMENT DYNAMICS

Keith Sill
Research Department
Federal Reserve Bank of Philadelphia

December 1997

WORKING PAPER 97-28
Regional Employment Dynamics

Keith Sill¹
Research Department
Federal Reserve Bank of Philadelphia
Ten Independence Mall
Philadelphia, PA 19106
email: ksill@frbphil.org

December 1997

¹The views expressed in this paper are the author's and not necessarily those of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System.

Abstract

There is a widespread belief that different geographic regions of the U.S. respond differently to economic shocks, perhaps because of factors such as differences in the composition of regional output, adjustment costs, or other frictions. We investigate the comovement of regional employment series using a common features framework. Little evidence is found to suggest that regions move synchronously; rather, it takes about three months before regions respond in a similar fashion to a common shock. We identify leading and lagging regions. None of the regional employment series appears to share a common, synchronous cycle with aggregate U.S. employment.

1 Introduction

There is a widespread belief that different geographic regions of the U.S. respond differently to economic shocks, perhaps because of factors such as differences in the composition of regional output, adjustment costs, or other frictions. Indeed, economists have long recognized that there are relative swings in the fortunes of regions (Hall [8]). More recently, there has been an increased interest in developing the facts about regional cycles.¹

The interest in regional dynamics can be motivated in several ways. First, regional shocks can have redistribution consequences. As a result, regional reallocations may lead to aggregate fluctuations. Thus, regional shocks might have predictive content for aggregate fluctuations. Similarly, regional shocks may trigger different adjustments of labor, capital, and output than do aggregate shocks. Second, to the extent that regions are specialized in the production of goods and services, national markets are somewhat arbitrary constructs from an economic point of view. The study of regional dynamics would then be the more natural approach to take. Finally, if regional dynamics are significantly different, policymakers may want to identify regional differences and/or imbalances in formulating policies and assessing their consequences.

This paper contributes to the stylized facts on regional employment dynamics using the common features framework first introduced by Engle and Kozicki [5] and then extended by Vahid and Engle [16]. The analysis identifies the presence of codependent cycles in the regional employment series and investigates the extent to which these cycles are synchronized. In short, we investigate whether there are common factors driving short-run movements in the regional series and the timing of the impact of the factors on the employment data.

Common features were first introduced as a way of investigating contemporaneous comovement in a set of series. A set of variables has a serial correlation common feature when a linear combination of them is unpredictable based on the history of the system, even when the series individually display serial correlation. An implication of this structure is that all of the persistence in the series can be embodied in a factor that is common to all the variables in the system. We might think of this factor as a common business cycle.

¹See, for example, Barro and Sala-i-Martin [1], Blanchard and Katz [2], Davis, Lounghani, and Mahidhara [4], Quah [13], and Carlino and DeFina [3]

While the magnitude of a response to the factor may differ across variables in the system, the timing of the response will be synchronized exactly.

The restriction implied by the serial correlation common feature is a strong one. We can imagine that while regional series may share a common driving factor, the timing of the response to shocks to the factor may differ for reasons alluded to above. Thus, if there is a shock to the common factor today, region 1 may respond today while region 2 may respond one month from today. Vahid and Engle [16] show how to test for, and estimate, this type of structure in a set of variables.

We use these methods to investigate the comovement of regional employment and to identify regions that tend to lead or lag aggregate U.S. employment. There is little evidence that regional cycles are synchronous: in pairwise comparisons of regions, only two sets of regions showed common, contemporaneous comovement. For the most part, it takes about three months before regions are responding in a similar fashion to economic shocks. This suggests that there are significant economic differences between geographic regions or that regional shocks spread out, with a lag, to other regions. We are also able to identify a group of regions that lead aggregate U.S. employment (New England, Great Lakes, Plains, and Southeast) and a group that lags aggregate employment (Mideast, Southwest, and Far West). None of the regions appears to move contemporaneously with aggregate U.S. employment. On balance, the evidence is consistent with the view that recessions and expansions "roll" across the U.S.

The plan of the paper is as follows. Section 2 lays out the empirical framework and discusses one possible source of differing regional dynamics: regional industry mix. This section follows closely the discussion and layout in Vahid and Engle [16]. Section 3 contains the empirical analysis for regional employment, investigating comovement in the context of bivariate and system-wide models. Section 4 investigates leading and lagging relationships between regional and aggregate employment using a generalized method of moments methodology. Section 5 concludes.

2 The Empirical Framework

The empirical model is motivated by an interest in quantifying the extent to which the U.S. regional employment series move together in both the short run and the long run. Over the long run, in a well-integrated economy,

regional employment growth rates might eventually converge in the face of diminishing returns to factors of production (if long-run regional productivity growth rates are similar). However, the short-run dynamics of employment growth across geographic regions of the U.S. may be quite different if regions have significant differences in their economic characteristics. The most obvious candidate for such differences is industry mix. Different industries within the broad economy may be affected by the same economic shock in different ways and with some variation in the timing of a response to shocks. If regional employment compositions by industry differ in an economically meaningful way, there could be dissimilar responses of regions to the same economic shock. In addition, economic shocks that originate within a specific region may propagate with a lag to neighboring regions. Regions with weak economic ties to the shock-originating region may show different dynamics in the response of employment to the shock.

How much does employment composition by industry vary across regions of the U.S.? Table 1 provides evidence on industry mix, measured as number of employed workers in an industry relative to total regional employment. The numbers represent employment shares averaged over the period 1955-1995. The definitions of the BEA regions are given in Appendix A. The industry classifications are single-digit SIC categories, so the entries in the table may mask significant differences in industry mix that would be apparent if we were to go to a finer level of disaggregation. However, even at this coarse level of disaggregation, there are significant differences in industry-level employment across regions.

Manufacturing shows the most variability across regions, from a low of 14 percent of employment in the Rocky Mountain region to a high of 31 percent in the Great Lakes region. There is also a fair amount of variability in employment shares for the services and government sectors. Services employment ranges from a low of 13 percent in New England to a high of 20 percent in the Far West. For the government employment share, New England has a low of 14 percent compared to the high of 22 percent in the Rocky Mountain region. The other industrial classifications show little variation in employment but generally account for about 40 percent of employment in a region.

The fact that a substantial share of industry employment varies significantly across regions leads us to suspect that the dynamics of regional responses to economic shocks will differ. Further, the largest volatility in employment shares within regions (measured by standard deviation) occurs in

manufacturing and services, two of the industries that show the most variability across regions.

There are several ways to quantify the intuitive notion that a set of variables move together over the short run. We might investigate correlation coefficients between pairs of series or impulse response functions in a vector autoregression of all series jointly. A recent alternative measure of comovement is due to Engle and Kozicki [5], who introduced the idea of a serial correlation common feature (SCCF). Simply put, a set of variables exhibits an SCCF when each of the series is individually serially correlated but a linear combination of them is not. This is a strong measure of comovement, since it implies that a common shock that has a persistent effect on the individual series will have no such effect on a specific linear combination of the series. Thus, the presence of an SCCF among a set of variables suggests that they share a common cycle that moves the series contemporaneously.

The implications of an SCCF in a set of variables can be illustrated in a simple bivariate example. Suppose we have two scalar time series y_{1t} and y_{2t} that have the following representation in terms of the common factor f_t :

$$\begin{aligned} y_{1t} &= \delta_1 f_t + e_{1t} \\ y_{2t} &= \delta_2 f_t + e_{2t} \\ f_t &= \beta f_{t-1} + w_t \end{aligned}$$

where e_{1t} and e_{2t} are innovations. While y_{1t} and y_{2t} are individually serially correlated through the common factor f_t , the linear combination $z_t = \delta_2 y_{1t} - \delta_1 y_{2t}$ is white noise. In this case, y_{1t} and y_{2t} share an SCCF. Note that while the factor f_t affects y_{1t} and y_{2t} differently via the coefficients δ_1 and δ_2 , the common effect is synchronous: both series are affected at the same time by the movement in f_t .

We could allow for spillover effects by setting up the model a little differently:

$$\begin{aligned} y_{1t} &= \delta_1 f_t + e_{1t} \\ y_{2t} &= \delta_2 f_t + e_{2t} \\ f_t &= \beta f_{t-1} + \gamma_1 e_{1t-1} + \gamma_2 e_{2t-1}. \end{aligned}$$

This model also possesses an SCCF, and thus the series have a common synchronous cycle, since the linear combination $\delta_2 y_{1t} - \delta_1 y_{2t}$ is serially uncorrelated. Note then that the presence of an SCCF implies that all dependence

on the past can be attributed to a factor common to all the variables in the system.

Vahid and Engle [16], building on the work of Gourieroux and Peaucelle [6, 7] and Tiao and Tsay [14, 15], generalized the SCCF notion to allow for common cycles that are nonsynchronous. Returning to the simple factor model example, the model structure might be given by:

$$\begin{aligned} y_{1t} &= \delta_1 f_t + e_{1t} \\ y_{2t} &= \delta_2 f_{t-1} + e_{2t} \\ f_t &= \beta f_{t-1} + w_t. \end{aligned}$$

The variable y_{2t} depends on f_{t-1} while y_{1t} depends on f_t . Note, though, that the linear combination $z_t = \delta_2 y_{1t} - \delta_1 \beta y_{2t}$ is a process with only first-order serial correlation even though y_{1t} and y_{2t} are individually serially correlated of a much higher order. The series share the common factor f_t , but now the cycles that are propagated through f_t are nonsynchronous – y_{2t} responds with a 1-period lag. Similarly, we can imagine that series share a nonsynchronous cycle at j lags.

Nonsynchronized common cycles satisfy the three axioms of a testable feature as described in Engle and Kozicki [5]. Following Vahid and Engle [16], let the *feature* A be defined by “predictable from the joint history of the n -vector series Z prior to and including $t-j$.” Then X has A according to:

$$\begin{aligned} X \text{ has } A &\text{ if } E(X|F_{t-j}) \neq 0 \\ &\text{where } F_{t-j} \text{ is the } \sigma\text{-field generated by } (Z_s, s \leq t-j) \end{aligned}$$

The feature A is common to the series if they all have A but a linear combination of them does not have A .

The common feature framework is a natural way in which to examine the common dynamics of the regional employment series. This methodology can be used to investigate whether regions move together synchronously or nonsynchronously and to get an idea of how nonsynchronous the movements might be.

Testing for the presence of common features is relatively straightforward using the scalar components model (SCM) of Tiao and Tsay [15]. The test uses canonical correlation analysis to determine the order of vector ARMA models. Details are given in Tiao and Tsay. We first repeat their definition of an SCM process:

Definition 1 (Tiao and Tsay (1989, pg 164)): For a k -dimensional process z_t , a non-zero linear combination $y_t = v_0^T z_t$ is said to follow an SCM(p_1, q_1) structure if there exist p_1 k -dimensional vectors $[v_1, \dots, v_{p_1}]$ such that:

1. v_{p_1} is non-zero when $p_1 > 0$ and
2. the linear combination of z_t, \dots, z_{t-p_1} , $u_t = v_0^T z_t + \sum_{j=1}^{p_1} v_j^T z_{t-j}$ satisfies:

$$E(a_{t-m} u_t) \begin{cases} \neq 0 & \text{if } m = q_1 \\ = 0 & \text{if } m > q_1 \end{cases}$$

where a_t is the vector of innovations for z_t .

Consider then a vector process z_t that follows the vector ARIMA model:

$$\Phi(L)z_t = \Theta(L)a_t$$

where $\Phi(L)$ and $\Theta(L)$ are polynomials in the lag operator L . If z_t has an SCM(0,j) representation, it can be written as:

$$v_0^T z_t = v_0 a_t + \sum_{i=1}^j h_i^T a_{t-i}.$$

If the process is SCM(0,j), linear combinations of the variables have an MA(j) representation and thus have a short memory relative to that of the process z_t . These are the types of SCM process investigated in this paper. If the vector process z_t has an SCM(0,j) representation, there is linear dependence in the impulse responses of z_t after j periods.

Tiao and Tsay proposed a test for SCM(0,j) in a vector y_t that amounts to testing whether some of the canonical correlations between y_t and $Z_{h,t-j-1} = \{y_{t-j-1}, \dots, y_{t-j-1-h}\}$, $h \geq 0$ are zero. Let $\hat{\lambda}_i(j)$ be the i th smallest squared canonical correlation (eigenvalue) between y_t and $Z_{h,t-j-1}$. The proposed test statistic is given by:

$$c(j, s) = (T - h - j) \sum_{i=1}^s \log\left(1 - \frac{\hat{\lambda}_i(j)}{d_i(j)}\right)$$

with $d_i(j)$:

$$d_i(j) = 1 + 2 \sum_{v=1}^j \hat{\rho}_v(\hat{\alpha}^T y_t) \hat{\rho}_v(\hat{\gamma}^T Z_{ht})$$

and $\rho_v(x_t)$ the lag- v sample autocorrelation of the process x_t and $\hat{\alpha}$ and $\hat{\gamma}$ the canonical variates associated with the squared canonical correlations $\hat{\lambda}_i(j)$. Under the null, the statistic is distributed as χ^2 with degrees of freedom equal to $s \times \{h \times k + s\}$ where k is the number of elements in the vector y_t .

There is an intuition behind searching for zero-canonical correlations between y_t and its lagged values that is easiest to express in terms of the SCM(0,0) test, which is a test for SCCF. Suppose we have two sets of variables, $y_t = \{y_{1,t}, \dots, y_{k,t}\}$ and a conditioning set $z_t = \{y_{t-1}, \dots, y_{t-m}\}$. Let $\mu_{it} = \alpha_i y_t$ and $v_{it} = \beta_i z_t$ be specific linear combinations of y_t and z_t . Canonical correlations are the set of orthogonal α_i 's and β_i 's that deliver maximal correlations between μ_{it} and v_{it} . Thus, the statistically zero canonical correlations are linear combinations of the y_t 's that are uncorrelated with all linear combinations of the z_t 's, since they are uncorrelated with the linear combination that gives the maximum correlation between μ_{it} and v_{it} .

Testing for the presence of common cycles in a series, whether synchronous or nonsynchronous, is then straightforward. We simply check for statistically zero canonical correlations between the series y_t and its relevant past z_{t-j} beginning with $j = 0$. If the model is rejected, we test for $j = 1$ and continue until we find the appropriate SCM(0,j) structure. In the case where $j = 0$, the eigenvectors associated with the zero canonical correlations between the dependent variables and instruments are limited information maximum likelihood estimates of the cofeature vector. When $j > 0$, the eigenvectors associated with the zero canonical correlations give non-optimal generalized method of moments (GMM) estimates of the cofeature vectors, as shown by Vahid and Engle [16]. They also show how to construct optimal GMM estimates of the cofeature vectors, which we exploit below in examining leading and lagging relationships between the regional employment series and aggregate U.S. employment.

Some care must be taken in choosing the instrument set used to test for and estimate the cofeature vectors. The description of the procedure above assumed that all of the variables under analysis were stationary. In our case, as we are looking at regional employment, the series are nonstationary and may be cointegrated in certain combinations. In such a situation, testing for synchronous and nonsynchronous cycles must be modified slightly.

Suppose then that the vector y_t is nonstationary in levels but stationary in first differences and that the elements of y_t are cointegrated. We can model

y_t using a vector error-correction model (VECM):

$$\Delta y_t = \alpha + \beta y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \Delta y_{t-h} + \epsilon_t.$$

In this case, we will search for cointegration in the first differences of the elements of y_t . The appropriate instrument set used in testing and estimation is composed of lagged differences in y_t as well as the lagged value of the error correction term (βy_{t-1}) from the VECM estimation. In the case where the system is characterized by r cointegrating vectors, the degrees of freedom for the SCM(0,j) statistic are $s \times \{h \times k + r + s\}$. If a linear combination of nonstationary $I(1)$ variables is SCM(0,j), it implies that their cycles in the Beveridge-Nelson decomposition representation have an SCM(0,j-1) structure. Further details are in Vahid and Engle [16].

3 Common Cycles in Employment Across Regions

We consider the evolution of regional employment in the post-World War II period. The data are state-level employment aggregated into eight BEA-defined regions. The data are quarterly and run from 1955 to 1995Q2 for a total of 162 observations. A plot of the log-level of regional employment data is given in Figure 1. The figure shows that, over the long run, the series generally move together, but it also suggests that there are significant differences in average employment growth across regions for the sample period. Table 2 gives some sample statistics for the series.

The Mideast region has the lowest average level of employment growth over the sample period, less than half the average growth of the fastest growing regions, the Southwest and Rocky Mountain. The Great Lakes region is most volatile, in part because of the jump in the level of the series early in the sample period (see Figure 1). However, the Great Lakes region still shows the most volatility when the sample statistics are calculated over 1956:1 to 1995:2. The table indicates that there is a considerable disparity in growth rates across regions. The slow-growing regions are New England and the Mideast; the moderate growth regions are the Great Lakes and Plains; and the fast growth regions are the Southeast, Southwest, Rocky Mountain, and Far West.

3.1 Bivariate Analysis

Consider first the extent to which the regions move together when examined in pairwise combinations. Thus, we initially test for SCCF (or, what is the same thing, SCM(0,0)) in bivariate models of regional employment. Each of the regional series was tested individually for unit roots in its log levels, and in each case, the null of a unit root in the series could not be rejected. When first differenced, each of the series appears stationary, according to standard unit root tests, and, furthermore, displays significant serial correlation.

We next examined whether regional employment is cointegrated in pairwise combinations, since the presence of cointegration has implications for how the SCM tests are set up. Bivariate models were estimated for each of the regions using Johansen's [10, 11] FIML procedure. Again, and for all the discussion that follows, the regional data are in logs. The trace test statistics for the null of no cointegration are presented in Table 3. The critical value for the statistics at the 90 percent confidence level is 13.33 according to the tables in Osterwald-Lenum [12]. Perhaps surprisingly, there is no strong evidence of cointegration in any pair of regional employment series. In other words, when taken in pairwise combinations, the regional employment series do not seem to share a common stochastic trend. This implies that the dynamics of the bivariate models can be adequately captured by a VAR in first differences.

It is an interesting finding that the regions do not seem to share a common long-run trend, especially given their degree of economic integration. We will see below that when a model is estimated with all regions jointly, there is evidence of common stochastic trends among the eight regions. The lack of such a finding in the bivariate relationships is consistent with a complicated stochastic structure for the joint variables. There may be several stochastic trends driving the long-run behavior of the regional series, such that no linear combination of any two series is sufficient to cancel out the long-run trends.

We next turn to whether the series have common short-run movements as characterized by the presence of a serial correlation common feature. This amounts to testing whether the pairwise combinations of regional employment have an SCM(0,0) structure. The tests look for statistically zero canonical correlations between the sets of variables $v_{ik,t} = \{\Delta y_{i,t}, \Delta y_{k,t}\}$ and $z_{ik,t} = \{v_{ik,t-1}, \dots, v_{ik,t-1-h}\}$ where i, k index regions and $\Delta y_{i,t}$ is employment growth in region i at time t . The test statistics are presented in Table 4 for $h = 1$. The test statistic $C(1)$ is asymptotically chi-squared with 3 de-

degrees of freedom under the null of a single white-noise component and $C(2)$ is asymptotically chi-squared with 8 degrees of freedom under the null of two white noise components. The first entry in each cell of the table is the computed $C(1)$ statistic and the second is the $C(2)$ statistic.

The tests indicate that, for the most part, the regions do not share a common cycle when taken in pairs. Only in the case of the Great Lakes and Plains and the Great Lakes and Southeast pairs is there evidence that the regions share a synchronous common cycle. The evidence from the $C(2)$ test statistics indicates as well that there is significant serial correlation in the growth rates of regional employment in all of the pairwise combinations.

The result on the lack of synchronous cycles in employment is not that surprising, given the finding of significant differences in industry mix across regions (Table 1). The regions of the U.S. are well integrated economically, but their economic structures are not identical. However, it is only a conjecture that differences in industry mix lead to a lack of synchronized cycles in the regional employment series. We have no formal statistical tests of this conjecture.

We next step through the $SCM(0,j)$ tests for increasing j to examine the regional data for the presence of nonsynchronous cycles. There is no evidence of codependence at 1 lag, but within two lags, many of the regions appear to show evidence of codependence. The results of the testing are in Table 5. The first row in each cell gives the lag at which evidence of common cycles was found. The second row gives the value of the $C(1)$ test statistic (in no cases were the $C(2)$ statistics not rejected), and the third row gives the eigenvector corresponding to the smallest eigenvalue. Recall that the eigenvector is a consistent (but not efficient) estimate of the cofeature vector. Thus, the eigenvector gives the linear combination of the regions uncorrelated after j lags.

All of the regional pairwise combinations showed evidence of an SCM structure, and thus common nonsynchronous cycles, at some lags between 0 and 4 except for the Great Lakes-New England combination. About half (12) of the bivariate combinations showed evidence of codependence at 2 lags or fewer with the remaining showing codependence at lags 3 and 4. New England shows the least variability in lags for finding codependence across regions, generally showing nonsynchronous cycles at lag 3. The Far West region also shows fairly small variation in lags when finding codependence with other regions, bouncing between 2 and 3 lags.

It appears then that the U.S. regions share a common driving factor that

generates serial correlation in employment growth rates, but that the cyclical comovement is not contemporaneous. Within roughly one year, however, the impulse response functions of regional employment growth are moving together. Thus, there is evidence that economic shocks "roll" through the U.S. regions in contrast to a scenario that has all of the regions responding contemporaneously to an economywide shock.

3.2 System Estimation

The bivariate analysis suggests that economic shocks have different effects, in terms of timing, on regional employment. Additional, and more efficient, information on synchronous/nonsynchronous cycles in regional employment can be gained if we consider all of the U.S. regions jointly, in a multivariate analysis.

The extension of the above-outlined technique to a multi-region system is straightforward. We first undertake a specification analysis in order to build a model that adequately captures the dynamics of the system composed of the eight regional employment series. A vector error correction model (VECM) was formed, both with and without an unrestricted linear time trend. This model was then tested for lag-length using the AIC and SIC criteria. Let y_t denote the stacked 8x1 vector of regional employment levels in logs. The VECM model has the form:

$$\Delta y_t = \alpha + \delta_0 y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_k \Delta y_{t-k} + \mu t + \epsilon_t$$

and was tested for lag length k . The results of the AIC and SIC tests are given in Table 6. The criterion functions used are, for the AIC criterion, $T |\Sigma_k| + 2n^2k$ and for the SIC criterion $T |\Sigma_k| + nTk \log(T)$ where T is the number of observations, n is the number of variables, and k is the number of lags. Tests were conducted under two cases: one where μ is unrestricted and one where $\mu = 0$.

The results for both the AIC and SIC test statistics indicate that a single lag is sufficient to describe the dynamics of the system, whether or not an unrestricted time trend is included in the VECM.

Conditional on a single lag in the VECM representation of the system, the test statistics for the number of cointegrating relationships among the eight regional employment series are presented in Table 7. The statistics were constructed for the two cases in question: with and without an unrestricted time trend in the model. For the model with a time trend, the

maximum-eigenvalue test suggests two cointegrating vectors at the 5 percent level of significance, while the trace statistic suggests three cointegrating vectors characterize the system. A similar result obtains for the model without a time trend: either two or three cointegrating vectors, depending on the test statistic used.

In contrast to the bivariate analysis above, we find that the regional employment series, when considered jointly, do share common stochastic trends. Indeed, with three cointegrating relationships, there are five common stochastic trends among the eight regions. In light of this finding, it is perhaps not that surprising to find no evidence of cointegration in pairwise comparisons of regional employment. More than two regions need to be considered in order to capture all of the stochastic trends driving the long-run behavior of the system.

We next tested whether an unrestricted time trend belonged in the VECM representation. Likelihood ratio tests were conducted under the restriction of $h = 1, 2, \dots, 8$ cointegrating vectors characterizing the system. The test statistics, distributed as $\chi^2(8)$, rejected the null, suggesting that a time trend cannot be safely excluded from the model. In sum, the specification tests suggest that the regional employment data can be adequately characterized by a VECM with a time trend, one lag and two or three cointegrating relationships.

Consider now the short-run dynamics of the series as characterized by the presence of common features. Let $\Delta y_t = \{\Delta y_{1,t}, \dots, \Delta y_{nt}\}$ denote the vector of regional employment growth rates and $z_t = \{\Delta y_{t-1}, \beta y_{t-1}, t, 1\}$ denote the conditioning variables from the VECM. We first test for SCCF among the regional employment series using the SCM(0,0) test under the assumption that the system is characterized by three cointegrating vectors. Thus, we check for the number of statistically zero canonical correlations between y_t and z_t . The results of the tests are reported in the first column of Table 8. The statistics indicate that the eight regional series do not share a serial correlation common cycle. If they do share common cycles, the cycles are nonsynchronous.

Do the regional employment series share cycles that are nonsynchronous? As mentioned in Section 2, this amounts to searching for statistically zero canonical correlations between y_t and z_{t-j} for increasing j using the SCM(0,j) test. These zero correlations represent linear combinations of y_t that are unpredictable based on time $t - j$ information.

The results of the test for an SCM(0,j) structure for $j = 1, 2, 3, 4$ are

presented in Table 8. The tests again assume three cointegrating vectors in the construction of the error correction term. Each column of the table gives the $C(s)$ test statistics for a particular value of j . The final column of the table gives the critical values of the chi-squared distribution at 95 percent.

The table suggests that a single codependence relationship is present among the eight employment series at lags 1 and higher. Note that if a model is SCM(0,j), it should pass the test for an SCM(0,j+1) structure. This holds true for statistics under investigation. At 1 lag, the test statistics indicate that there are two codependence relationships among the regional series. At 4 lags there are 5 codependence relationships.

The results from the system estimation are broadly in accord with those of the bivariate analysis. Regional employment does share a common cycle, but that cycle is nonsynchronous: some regions lag others in response to a shock. The analysis indicates that, by and large, the regional employment series are moving together after about three months from the time of a shock.

4 Leading and Lagging Regions

The analysis above indicates that regional employment in the U.S. does not move synchronously over the short run. Since the regions do not move contemporaneously, it suggests that some regions may be leading regions and some may be lagging regions. That is, some regions may respond to common shocks consistently more quickly than do other regions. To investigate this possibility in a tractable way, we investigate whether regions lead or lag total employment for the U.S. (exclusive of employment for the region under investigation). Thus, we look at bivariate lead-lag relationships between regional employment and aggregate employment.

We investigate lead-lag relationships in a generalized method of moments (GMM) framework as follows. Suppose we have a bivariate model of the form:

$$\begin{aligned} y_{1t} &= \delta_1 f_t + e_{1t} \\ y_{2t} &= \delta_2 f_{t-1} + e_{2t} \\ f_t &= \beta f_{t-1} + w_t \end{aligned}$$

where $e_{1,t}, e_{2,t}, w_t$ are serially uncorrelated innovations that may be contemporaneously correlated. Thus, this model follows an SCM(0,1) structure

where y_{2t} lags y_{1t} . Consider the vector $y_t = \{y_{1,t}, y_{2,t+1}\}$. There is a linear combination of these variables $(\delta_2 y_{1,t} - \delta_1 y_{2,t+1})$ that is uncorrelated with time $t - 1$ information. Conversely, there is no such linear combination of the elements of $y_t^* = \{y_{1,t+1}, y_{2,t}\}$ that will be uncorrelated with time $t - 1$ information. Let $z_{t-1} = \{y_{1,t-1}, y_{2,t-1}\}$. We consider whether a vector β can be estimated that satisfies:

$$E(\beta' \tilde{y}_t \otimes z_{t-1}) = 0 \quad (1)$$

for $\tilde{y}_t = y_t, y_t^*$. These orthogonality conditions can be exploited in a GMM estimation of the parameter vector β .

The details of GMM estimation are well known (Hansen [9]) so here we give only a brief overview geared to our particular application. Let g_T be the sample analog to the set of orthogonality conditions:

$$g_T(\beta; Y, Z) = (1/T) \sum_{t=1}^T (\beta' \tilde{y}_t \otimes z_{t-1})$$

It is assumed that $\beta' y_t$ and z_{t-1} are stationary stochastic processes. The GMM estimator is the value of β that minimizes the quadratic form:

$$g_T(\beta^*; Y, Z)' W_T g_T(\beta^*; Y, Z)$$

for W_T a positive-definite weighting matrix that can depend on sample information. In the case under investigation, we can write the quadratic form, for a given value of W_T as:

$$\frac{\beta' Y' Z}{T} W_T \frac{Z' Y \beta}{T}$$

Normalizing the first element of β to 1 and denoting the rest of the elements by δ , the GMM estimator of δ is

$$\hat{\delta} = (Y_2' Z W_T Z' Y_2)^{-1} Y_2' Z W_T Z' Y_1.$$

Hansen [9] shows that the optimal weighting matrix W_T is an estimate of:

$$S^{-1} = \left[\lim_{T \rightarrow \infty} (1/T) \sum_{t=1}^T \sum_{v=-\infty}^{\infty} E(\beta' \tilde{y}_t \otimes z_{t-1})(\beta' \tilde{y}_{t-v} \otimes z_{t-1-v}) \right]^{-1}$$

the inverse of the asymptotic variance matrix of the sample mean of the orthogonality conditions.

As indicated in Vahid and Engle [16], we can use:

$$\hat{W}_T = \left[\hat{\sigma}^2 \left((1/T) \sum_{t=1}^T z_t z_t' \right) + \sum_{i=1}^q \hat{\gamma}_i \left((1/T) \sum_{t=i+1}^T (z_t z_{t-i}' + z_{t-i} z_t') \right) \right]^{-1}$$

as an optimal estimate of W_T where $\hat{\sigma}^2$ is an estimate of the variance of $\beta' y_t$ and $\hat{\gamma}_i$ is an estimate of the i th autocovariance of $\beta' y_t$.

In the first stage of the GMM procedure the values of $\hat{\sigma}$ and $\hat{\gamma}_i$ are unknown, so the quadratic form is minimized with the weighting matrix $W_T = (Z'Z)^{-1}$. Estimates of $\hat{\sigma}^2$ and $\hat{\gamma}_i$ are then obtained. In the second stage, these estimates are used to construct an optimal weighting matrix and new parameter estimates are derived.

When the number of orthogonality conditions exceeds the number of parameters to be estimated, the model is overidentified. Hansen [9] suggested a model specification test based on the orthogonality conditions being close to zero. He shows that, if the orthogonality conditions are true, T times the minimized value of the objective function is distributed chi-squared with degrees of freedom equal to the number of overidentifying restrictions. This is the test to be used in assessing leading and lagging relationships between the regions and the U.S. employment series.

Our procedure is as follows. We undertake a GMM estimation of a bivariate model consisting of a region's employment and aggregate employment for the rest of the nation. The aggregate employment number is the sum of employments for the seven other regions. We first lead the regional employment series by j periods and estimate a codependence vector under the restrictions implied by equation (1). Hansen's J-statistic is examined to see if the model is rejected. We then repeat the procedure but instead lead the aggregate employment series relative to the regional series by j periods. The model is re-estimated and the J-statistic is again examined. When the null that the orthogonality conditions are close to zero is not rejected, it is interpreted as evidence that we have identified a leading/lagging relationship between the series in question.

First, the eight measures of aggregate employment (U.S. employment less region j employment) were tested for unit roots. We could not reject the hypothesis of a unit root in levels but can reject in first differences. The eight bivariate models were then examined for cointegration: in no case could the null of no cointegration be rejected. Thus, each of the bivariate models of regional and national employment can be expressed solely in terms of first

differences. Finally, as described in the bivariate SCM(0,j) tests described above, the models were tested for an SCM(0,j) structure. In each case, evidence of an SCM(0,j) structure was found for some $j = 1, 2, 3, 4$. So, there does appear to be a common cycle in regional and national employment. In no case was an SCM(0,0) not rejected, so the cycles appear to be nonsynchronous.

We then proceeded with the GMM estimation, stepping through $j = 1, 2, \dots$ in search of models that were not rejected. The instrument set consisted of two lags each of the regional series and the aggregate series. One of the coefficients in the codependence vector was normalized to one, so that a single parameter was estimated for each of the eight models using four orthogonality conditions. Thus, the J-statistics are distributed as chi-squared with 3 degrees of freedom under the null that the orthogonality conditions are close to zero.

The results of the estimations are given in Table 9. The first entry in each cell is the J-statistic for a model where the region leads the nation and the number in parentheses is the J-statistic for a model where the nation leads the region. The critical value of the test statistic at the 5 percent level is 7.81. Numbers with an asterisk are therefore not significantly different from zero at the 5 percent level and thus indicate that the overidentifying restrictions are not rejected by the data.

The tests suggest that the following regions have employment series that lead the nation by one quarter: New England, Plains, and Southeast. None of the regions appears to lag the U.S. aggregate by one quarter. The Great Lakes appears to lead the aggregate by two quarters, while the Mideast, Southwest, and Far West lag U.S. employment by two quarters. The Rocky Mountain region appears to lead the aggregate by four quarters. However, this Rocky region is small relative to the others so we are a bit skeptical that a robust leading relationship has been found.

Thus, the leading regions, relative to the U.S., appear to be New England, Great Lakes, Plains, and Southeast, perhaps with the Rocky Mountain region included. The lagging regions relative to the U.S. are the Mideast, Southwest, and Far West. The leading regions tend to have a somewhat higher fraction of employment in manufacturing and a lower fraction in services and government relative to the lagging regions (see Table 1). However, the Mideast (a lagging region) has an industry mix almost identical to the Southeast (a leading region), indicating that industry mix is probably not the whole story in accounting for leading and lagging relationships between

aggregate and regional employment. Once again, the GMM tests indicate that economic shocks do not affect all regions of the country contemporaneously but rather that the effects of the shocks propagate through the economy geographically.

5 Conclusion

Evidence has been presented supporting the view that economic shocks "roll" across geographic regions of the U.S. economy. That is, though regional employment series do share a common factor that drives short-run persistence, shocks to the common factor do not affect regions contemporaneously. However, evidence from scalar components model tests indicate that regional employment does largely move in a similar fashion after about three months from the time of an economic shock.

We also investigated whether regions tend to lead or lag the rest of the nation. GMM test statistics indicate that the employment series for New England, Great Lakes, Plains, and Southeast tend to lead the nation while the Southeast, Southwest, and Far West tend to lag the nation. For the most part, regions that lead the U.S. tend to have more employment in manufacturing and less in services and government than do the lagging regions.

Appendix
Definitions of Regions

New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont.

Mideast: Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania.

Great Lakes: Illinois, Indiana, Michigan, Ohio, Wisconsin.

Plains: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota.

Southeast: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia.

Southwest: Arizona, New Mexico, Oklahoma, Texas.

Rocky Mountain: Colorado, Idaho, Montana, Utah, Wyoming.

Far West: California, Nevada, Oregon, Washington.

References

- [1] R.J. Barro and X. Sala i Martin. Convergence. *Journal of Political Economy*, 100(2):223–251, 1992.
- [2] O.J. Blanchard and L.F. Katz. Regional evolutions. *Brookings Papers on Economic Activity*, pages 1–61, 1992.
- [3] G. Carlini and R. DeFina. Regional income dynamics. *Journal of Urban Economics*, 37:88–106, 1995.
- [4] S.J. Davis, P. Loungani, and R. Mahidhara. Regional unemployment cycles. Unpublished manuscript, 1996.
- [5] R.F. Engle and S. Kozicki. Testing for common features. *Journal of Economic and Business Statistics*, 11:369–395, 1993.
- [6] C. Gourieroux and I. Peaucelle. Detecting a long run relationship (with an application to the ppp hypothesis). CEPREMAP Working paper 8902, 1988.
- [7] C. Gourieroux and I. Peaucelle. Series codependantes: application a l’hypothèse de parité du pouvoir d’achat. *Revue d’Analyse Economique*, 68:283–304, 1992.
- [8] R.E. Hall. Turnover in the labor force. *Brookings Papers on Economic Activity*, pages 709–764, 1972.
- [9] L.P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.
- [10] S. Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–254, 1988.
- [11] S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59:1551–1580, 1991.
- [12] M. Osterwald-Lenum. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, 54(3):461–471, 1992.

- [13] D. Quah. One business cycle and one trend from (many,) many disaggregates. *European Economic Review*, 38:605–613, 1994.
- [14] G.C. Tiao and R.S. Tsay. A canonical correlation approach to modeling multivariate time series. In *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, pages 112–120, 1985.
- [15] G.C. Tiao and R.S. Tsay. Model specification in multivariate time series. *Journal of Royal Statistical Society, Series B*, 51:157–213, 1989.
- [16] F. Vahid and R.F. Engle. Codependent cycles. *Journal of Econometrics*, 80:199–221, 1997.

Table 1: Regional employment shares by major industry

Region	Mining	Const	Mfg	Trans	Trade	FIRE	Svcs	Govt
NE	.00	.04	.30	.05	.21	.06	.13	.14
ME	.00	.04	.25	.06	.21	.07	.19	.17
GL	.00	.04	.31	.06	.22	.05	.16	.15
PL	.01	.05	.21	.07	.24	.05	.17	.18
SE	.01	.06	.25	.06	.22	.05	.17	.18
SW	.04	.06	.16	.07	.24	.05	.18	.19
RM	.03	.06	.14	.07	.24	.05	.16	.22
FW	.00	.05	.21	.06	.23	.06	.20	.19

The regions are: New England (NE), Mideast (ME), Great Lakes (GL), Plains (PL), Southeast (SE), Southwest (SW), Rocky Mountain (RM), and Far West (FW).

The industry classifications are: Mining; Construction; Manufacturing; Transportation and Public Utilities; Wholesale and Retail Trade; Finance, Insurance, and Real Estate; Services; Total Government.

Table 2: Sample Statistics for Regional Employment Growth 1955:1 to 1995:2, quarterly growth rates

Region	Mean	s.t.d	min	max
New England	.0036	.0076	-.024	.024
Mideast	.0027	.0057	-.016	.018
Great Lakes	.0052	.0209	-.041	.241
Plains	.0051	.0059	-.015	.017
Southeast	.0075	.0067	-.022	.025
Southwest	.0081	.0064	-.012	.023
Rocky Mnt	.0080	.0066	-.011	.029
Far West	.0072	.0069	-.012	.025

Table 3: Trace Test Statistics for Cointegration in Bivariate Models of Regional Employment

	NE	ME	GL	PL	SE	SW	RM
ME	3.75						
GL	2.78	3.62					
PL	1.82	2.95	11.44				
SE	2.61	4.46	12.26	6.15			
SW	4.85	4.18	8.86	5.98	4.65		
RM	4.14	3.82	10.04	7.54	4.55	4.60	
FW	9.43	7.93	11.52	7.12	5.31	5.55	4.45

Critical value for null of no cointegration is 15.41 at the 95 percent significance level and 13.33 at the 10 percent level. See Osterwald-Lenum [12] for test statistic tables.

Table 4: SCM(0,0) tests for bivariate case, h=1

	NE	ME	GL	PL	SE	SW	RM
ME	9.4 166.2						
GL	29.6 186.5	28.2 120.8					
PL	62.1 211.3	41.7 149.0	4.3* 107.1				
SE	72.5 235.9	37.5 166.3	5.6* 142.4	21.5 165.7			
SW	113.2 281.5	89.4 233.0	62.9 206.8	76.2 219.1	115.4 266.5		
RM	69.9 217.7	78.6 185.7	57.1 139.3	57.9 150.8	57.9 188.5	28.5 174.5	
FW	73.2 240.9	29.3 174.9	20.7 161.9	27.1 179.6	65.0 215.6	80.7 258.1	65.9 202.2

* denotes significant at 5 percent level

Table 5: Summary of SCM(0,j) tests for bivariate case, h=1

	NE	ME	GL	PL	SE	SW	RM
ME	j=3 5.4 (-.41,.91)						
GL	****	j=2 7.2 (.95,.30)					
PL	j=3 2.6 (.43,.90)	j=3 3.9 (.98,.17)	j=0 4.3 (.80,-.59)				
SE	j=2 5.2 (-.19,.98)	j=3 7.4 (.99,-.03)	j=0 5.6 (-.59,.80)	j=2 3.7 (-.51,.86)			
SW	j=3 3.2 (.14,.99)	j=2 6.1 (.79,.62)	j=4 2.7 (-.15,.99)	j=4 2.1 (-.26,.97)	j=4 5.2 (-.19,.98)		
RM	j=3 1.7 (.78,.62)	j=2 4.9 (.91,.41)	j=2 5.0 (.89,.49)	j=2 7.7 (.86,.51)	j=2 4.9 (.95,.30)	j=4 4.7 (.85,-.52)	
FW	j=3 6.4 (-.34,.94)	j=2 7.3 (-.37,.93)	j=3 7.2 (.44,.89)	j=2 4.7 (-.54,.84)	j=3 2.9 (.89,.46)	j=2 6.6 (.18,.98)	j=3 2.7 (.14,.99)

Table 6: Lag Length Test Statistics

lags	<i>AIC</i>		<i>SIC</i>	
	trend	notrend	trend	notrend
0	-13839.8	-13758.2	-13839.8	-13758.2
1	-14050.9	-14005.7	-13857.8	-13812.6
2	-14041.1	-14000.4	-13654.9	-13614.2
3	-14041.9	-13996.1	-13462.5	-13416.8
4	-14045.9	-14011.4	-13273.5	-13238.9
5	-14036.2	-13988.6	-13070.7	-13023.1
6	-14046.7	-13991.6	-12888.0	-12832.9

Table 7: Test Statistics for Cointegration Among the 8 Regional Employment Series

p-r	<i>Trend</i>				<i>No Trend</i>			
	$\lambda - max$	c.v.	trace	c.v.	$\lambda - max$	c.v.	trace	c.v.
8	102.31	54.25	279.06	170.80	102.08	51.42	246.99	156.00
7	60.77	48.45	176.75	136.61	51.77	45.28	144.91	124.24
6	39.94	42.48	115.99	104.94	32.12	39.37	93.14	94.15
5	31.58	36.41	76.04	77.74	23.19	33.46	61.02	68.52
4	21.76	30.33	44.47	54.64	19.42	27.07	37.83	47.21
3	14.75	23.78	22.71	34.55	8.87	20.97	18.42	29.68
2	7.92	16.87	7.96	18.17	6.37	14.07	9.54	15.41
1	0.04	3.74	0.04	3.74	3.17	3.76	3.17	3.76

Table 8: SCM(0,j) tests statistics for system estimation

C(s)	j=0	j=1	j=2	j=3	j=4	c.v. 95
C(1)	15.04	2.69	6.36	1.69	1.59	9.5
C(2)	34.07	17.88	14.30	5.09	5.09	18.3
C(3)	55.39	30.83	25.13	15.48	10.17	28.9
C(4)	100.76	62.69	41.49	28.23	20.62	41.3
C(5)	194.31	133.64	65.75	56.65	50.78	55.8
C(6)	339.98	199.57	126.59	94.21	100.63	72.2
C(7)	530.05	268.28	170.75	129.91	132.97	90.5
C(8)	764.03	331.83	217.64	163.88	166.63	110.9

Table 9: J-statistic tests for leading and lagging regions relative to aggregate employment

Region	j=1	j=2	j=3	j=4
New England	2.62* (44.32)			
Mideast	11.59 (11.32)	12.31 (6.30*)		
Great Lakes	15.52 (24.66)	5.34* (15.23)		
Plains	0.45* (16.43)			
Southeast	3.55* (15.63)			
Southwest	86.92 (37.44)	98.98 (7.66*)		
Rock Mntn	42.43 (31.26)	27.39 (24.19)	32.70 (15.78)	2.34* (11.79)
Far West	16.39 (19.89)	28.08 (4.52*)		