Discussion Paper 77

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
Research Department
250 Marquette Avenue
Minneapolis, Minnesota 55480-0291

November 1992

A DYNAMIC INDEX MODEL FOR LARGE CROSS SECTIONS

Danny Quah*

London School of Economics

Thomas J. Sargent*

Stanford University
and University of Chicago

ABSTRACT

This paper shows how standard methods can be used to formulate and estimate a dynamic index model for random fields—stochastic processes indexed by time and cross section where the time-series and cross-section dimensions are comparable in magnitude. We use these to study dynamic comovements of sectoral employment in the U.S. economy. The dynamics of employment in sixty sectors is well explained using only two unobservable factors; those factors are also strongly correlated with GNP growth.

*Conversations with Gary Chamberlain and Paul Ruud have clarified our thinking on some of the issues here. John Geweke, Bruce Hansen, Mark Watson, James Stock, Jeff Wooldridge, and an anonymous referee have also provided helpful comments, not all of which we have been able to incorporate. Finally, we gratefully acknowledge the hospitality of the Institute for Empirical Macroeconomics where portions of this paper were written.

1. Introduction

In studying macroeconomic business cycles, it is convenient to have some low-dimensional characterization of economy-wide fluctuations. If an aggregate measure of economic activity—such as GNP or total industrial production—could be widely agreed on as accurately assessing the state of the economy, then all attention could focus on that measure alone. Econometric analysis and forecasting would then be, in principle, straightforward.

But any one given aggregate measure is likely affected by many different disturbances: understanding economic fluctuations then involves disentangling the dynamic effects of those different disturbances. For instance, much has been written on decomposing aggregate measures of output into different interpretable components. Some of this work uses multivariate time series methods to analyze such decompositions. Little empirical work has been done, however, proceeding in the opposite direction, i.e., using information from a wide range of cross-section data to shed light on aggregate fluctuations. In other words, in the business cycle analysis triad of *depth*, *duration*, and *dispersion*, the third has typically been ignored.

We know that professional forecasters look at a variety of different indicators to predict aggregate activity. Interest rates, measures of consumer sentiment, the money supply, and the spectrum of asset prices are all candidates for helping to predict economy-wide movements. There again, to understand aggregate fluctuations, one needs to go beyond examining a single time series in isolation.

This paper takes the natural next step on these observations: it provides a framework for analyzing commonalities—aggregate comovements—in dynamic models and data structures where the cross section dimension is potentially as large, and thus potentially as informative, as the time series dimension. Why might this be useful for analyzing macroeconomic business cycles?

When the aggregate state of the economy affects and in turn is affected by many different sectors, then it stands to reason that *all* those sectors should contain useful information for estimating and forecasting that aggregate state. It should then simply be good econometric practice to exploit such cross section information jointly with the dynamic behavior to uncover the aggregate state of the economy.

Standard time series analysis, however, is not well suited to analyzing largescale cross sections; rather, it is geared toward the study of fixed-dimensional vectors evolving over time. Cross section and panel data analysis, on the other hand, are typically not concerned with the dynamic forecasting exercises that are of interest in macroeconomics. It is this gap in macroeconometric practice between cross section and time series analyses that this paper addresses. While there are many interesting issues to be investigated for such data structures, we focus here on formulating and estimating dynamic index models for them.

The plan for the remainder of this paper is as follows: Section 2 relates our analysis to previous work on index structures in time series. Section 3 shows how (slight extensions of) standard econometric techniques for index models can be adapted to handle random fields. Section 4 then applies these ideas to study sectoral employment fluctuations in the US economy. Appendices contain more careful data description as well as certain technical information on our application of the dynamic index model to random fields.

2. Dynamic Index Models for Time Series

This paper has technical antecedents in two distinct literatures: the first includes the dynamic index models of Geweke (1977), Sargent and Sims (1977), Geweke and Singleton (1980), and Watson and Kraft (1984). The second literature concerns primarily the analysis of common trends in cointegrated models, as in, e.g., Engle and Granger (1987), King, Plosser, Stock, and Watson (1991), and Stock and Watson (1988, 1990). Because of the formal similarity of these two models—the dynamic index model and the common trends model—it is sometimes thought that one model contains the other. This turns out to be false in significant ways, and it is important to emphasize that these models represent different views on disturbances to the economy. We clarify this in the current section in order to motivate the index structure that we will end up using in our own study.

Notice first that neither of these standard models is well suited to analyzing structures having a large number of individual time series. When the number of individual series is large relative to the number of time series observations, the matrix covariogram cannot be precisely estimated, much less reliably factored—as would be required for standard dynamic index analysis.

Such a situation calls for a different kind of probability structure from that used in standard time series analysis. In particular, the natural probability model is no longer that of vector-valued singly-indexed time series but instead that of multiply-indexed stochastic processes, i.e., *random fields.* In developing an index model for such processes, this paper provides an explicit parametrization of dynamic and cross-section dependence in random fields. The analysis here thus gives an alternative to that research which attempts robust econometric analysis without such explicit modelling.

Begin by recalling the dynamic index model of Geweke (1977) and Sargent and Sims (1977). An $N \times 1$ vector time series X is said to be generated by a **K-index model** ($K < N$) if there exists a triple (U, Y, a) with U $K \times 1$ and Y $N \times 1$ vectors of stochastic processes having all entries pairwise uncorrelated, and a an $N \times K$ matrix of lag distributions, such that:

$$X(t) = a * U(t) + Y(t), \tag{2.1}$$

with $*$ denoting convolution. Although jointly orthogonal, U and Y are allowed to be serially correlated; thus (2.1) restricts X only to the extent that K is (much) smaller than N. Such a model captures the notion that individual elements of X correlate with each other only through the K-vector U. In fact, this pattern of low-dimensional cross-section dependence *is* the defining feature of the dynamic index model.

Such a pattern of dependence is conspicuously absent from the common trends model for cointegrated time series. To see this, notice that when X is individually integrated, but jointly cointegrated with cointegrating rank $N - K$, then its common trends representation is:

$$X(t) = aU(t) + Y(t), \tag{2.2}$$

where U is a $K \times 1$ vector of pairwise orthogonal random walks, Y is covariance stationary, and a is a matrix of numbers (Engle and Granger (1987) or Stock and Watson (1988, 1990)). As calculated in the proof of the existence of such a representation, U has increments that are perfectly correlated with (some linear combination of) Y. It is not hard to see that there need be no transformation of

(2.2) that leaves the stationary residuals Y pairwise uncorrelated and uncorrelated with the increments in U. Thus, unlike the original index model, representation (2.2) is ill suited for analyzing cross-sectional dependence in general. In our view, it is not useful to call Y idiosyncratic or sector-specific when Y turns out to be perfectly correlated with the common components U.

3. Dynamic Index Structures for Random Fields

While model (2.1) as described above makes no assumptions about the stationarity of X, its implementation in e.g. Sargent and Sims (1977) relied on being able to estimate a spectral density matrix for X. In trending data, with permanent components that are potentially stochastic and with different elements that are potentially cointegrated, such an estimation procedure is unavailable.

Alternatively, Stock and Watson (1990) have used first-differenced data in applying the index model to their study of business cycle coincident indicators—but in doing so they restrict the levels of their series to be eventually arbitrarily far from each series's respective index component. Stock and Watson chose such a model after pretesting the data for cointegration and rejecting that characterization. Thus, their non-cointegration restriction might well be right to impose, but again, the entire analysis is thereafter conditioned on the variables of interest having a particular cointegration structure.

When the cross section dimension is potentially large, standard cointegration tests cannot be used, and thus the researcher cannot condition on a particular assumed pattern of stochastic trends to analyze dynamic index structure. [For instance, if N exceeds T, the number of time series observations, no (sample) cointegrating regression can be computed.] Our approach explicitly refrains from specifying in advance any particular pattern of permanent and common components; instead, we use only the orthogonality properties of the different components to characterize and estimate an index model for the data. It turns out that by combining ingredients of (2.1) and (2.2) one obtains a tractable model having three desirable features: (1) the cross-section dependence is described by some (fixed) low-dimensional parametrization; (2) the data can have differing, unknown orders of integration and cointegration; the model structure should not depend on knowing those patterns; and finally, (3) the key parameters of the model are consistently estimable even when N and T have comparable magnitudes.

Let $\{X_j(t),\ j = 1,\ 2,\ \ldots,\ N,\ t = 1,\ 2,\ \ldots,\ T\}$ be an observed segment of a random field. We are concerned with data samples where the cross section dimension N and the time series dimension T take on the same order of magnitude. We hypothesize that the dependence in X across j and t can be represented as:

$$X_j(t) = a_j * U(t) + Y_j(t), \tag{3.1}$$

where U is a $K \times 1$ vector of orthogonal random walks; Y_j is zero mean, stationary, and has its entries uncorrelated across all j as well as with the increments in U at all leads and lags; and, finally, a_j is a $1 \times K$ vector of lag distributions. Model (3.1) is identical to (2.1) except that (3.1) specifies U to be orthogonal random walks. A moment's thought, however, shows that the random walk restriction is without loss of generality: an entry in a_j can certainly sum to zero and thereby allow the corresponding element in U to have only transitory effects on X_j.

As is well known, model (3.1) restricts the correlation properties of the data: in particular, the number of parameters in X's second moments grows only linearly with N, rather than quadratically, as would be usual.[1] Further, it is useful to repeat that (3.1) differs from the common trends representation (2.2) in having the j-specific disturbances uncorrelated both with each other and with U; neither of these orthogonality properties holds for (2.2). Thus, (3.1) shares with the original index model (2.1) the property that all cross correlation is mediated only through the lower dimensional U—independent of N and T. Such a property fails for the common trends representation.

Since Y_j is orthogonal to U, (3.1) is a projection representation and thus (relevant linear combinations of) the coefficients a_j can be consistently estimated by a least squares calculation: this holds regardless of whether X_j is stationary, regardless of the serial correlation properties of Y_j, and regardless of the distributions of the different variables. Notice further that for this estimation, one does not require

[1] If necessary such statements concerning moments can be read as *conditional on initial values*. For brevity, however, such qualification is understood, and omitted in the text.

observations on U, but only consistent estimates of the conditional expectations of certain cross products of U and X_j. Finally, since Y_j is uncorrelated across j's, and the regressors U are common, equation by equation estimation is equivalent to joint estimation of the entire system. Many of these properties were already pointed out by Watson and Engle (1983); here we add that if the parameters of (3.1) are consistently estimated as $T \to \infty$ independent of the value of N, then they remain consistently estimated even when N has the same order of magnitude as T.

The estimation problem is made subtle, however, by the U's being not observable. Nevertheless, using the ideas underlying the EM algorithm (see e.g., Watson and Engle (1983), Wu (1983), or Ruud (1991)), model (3.1) can be easily estimated. The key insight has already been mentioned above, and that is that to calculate the least squares projection, one needs to have available only sample moments, not the actual values of U. When, in turn, these sample moments cannot be directly calculated, one is sometimes justified in substituting in their place the conditional expectation of these sample moments, conditional on the observed data X. Note, however, that these conditional expectations necessarily depend on all the unknown parameters. Thus using them in place of the true sample moments, and then attempting to solve the first order conditions characterizing the projection—even equation by equation—could again be intractable, for large N.

Instead, treat the projections in (3.1) as solving a quasi-maximum likelihood problem, where we use the normal distribution as the quasi-likelihood. Then the reasoning underlying the EM algorithm implies that under regularity conditions the algorithm correctly calculates the projection in (3.1).

This procedure has two features worthy of note: first, estimation immediately involves only least squares calculations, equation by equation, and thus can certainly be performed with no increased complexity for large N. To appreciate the other characteristic, one should next ask, Given this first property, how has the cross-section information been useful? The answer lies of course in the second feature of this procedure—the cross-section information is used in calculating the conditional expectation of the relevant sample moments.

To make this discussion concrete, we now briefly sketch the steps in the estimation. These calculations are already available in the literature; our contribution here has been just to provide a projection interpretation for the EM algorithm for estimating models with unobservables—thus, to distance the EM algorithm from any tightly-specified distributional framework—and to note its applicability to the case when N is large.

Write out (3.1) explicitly as:

$$X_j(t) = \sum_{k=1}^{K} \alpha_{jk}(L)U_k(t) + Y_j(t); \tag{3.2}$$

in words, each X_j is affected by K common factors $U = (U_1, U_2, \ldots, U_K)'$ and an idiosyncratic disturbance Y_j. This gives a strong-form decomposition for each observed X_j into unobserved components: the first, in U, common across all j, and the second Y_j specific to each j, and orthogonal to all other X's. It will be convenient sometimes to add a term $d_jW(t)$ to (3.2), where W comprises purely exogenous factors, such as a constant or time trend.

The common factors in (3.2) have dynamic effects on X_j given by:

$$\alpha_{jk}(L) = \sum_{m=0}^{M_a} a_{jk}(m)L^m, \qquad M_a \text{ finite.}$$

Take U_k's to be integrated processes, with their first differences pairwise orthogonal and having finite autoregressive representation:

$$\Gamma(L)\Delta U(t) = \eta_U(t), \tag{3.3}$$

where $\Gamma(L)$ is diagonal, with k-th entry given by:

$$1 - g_k(1)L - g_k(2)L^2 - \cdots - g_k(M_g)L^{M_g}; \qquad M_g \text{ finite,}$$

and η_U is K-dimensional white noise having mean 0 and the identity covariance matrix. As suggested above, the normalizations following (3.3) are without loss of generality because U affects observed data only after convolution with a. If X_j

were stationary, the sequences a_j.—provided that $M_a \geq 1$—can contain a first-differencing operation to remove the integration built into (3.3).

Recall that the idiosyncratic or sector-specific disturbances Y_j are taken to be pairwise orthogonal as well as orthogonal to all ΔU. We now assume further that each Y_j has a finite autoregressive representation:

$$\beta_j(L)Y_j(t) = \epsilon_j(t) \qquad \text{for all } j \tag{3.4}$$

with

$$\beta_j(L) = 1 - b_j(1)L - b_j(2)L^2 - \cdots - b_j(M_b)L^{M_b}, \qquad M_b \text{ finite.}$$

Combining (3.2) and (3.3) and defining $\phi_{jk}(L) = \beta_j(L)\alpha_{jk}(L)$, we get:

$$\beta_j(L)X_j(t) = \sum_{k=1}^{K} \phi_{jk}(L)U_k(t) + \epsilon_j(t).$$

Notice that because α and β are individually unrestricted, we can consider the model to be parametrized in β and ϕ, being careful to take the ratio ϕ/β whenever we seek the dynamic effects of U on X. When W—the set of exogenous variables—is lag-invariant, as would be true for a time trend and constant, then $\beta_j(L)\,(d_j W(t))$ spans the same space as just $W(t)$. Thus, without loss of generality, we can write the model as simply:

$$\beta_j(L)X_j(t) = \sum_{k=1}^{K} \phi_{jk}(L)U_k(t) + \epsilon_j(t) + d_j W(t) \tag{3.5}$$

(possibly redefining d_j). Subsequent manipulations will exploit this representation (3.5) of the original strong-form decomposition of X given in (3.2).

Directly translating the model above into state space form—we do this in the Technical Appendix below—gives:

$$\begin{aligned} X(t) &= AZ(t) + dW(t) + Y(t); \\ Z(t+1) &= cZ(t) + \eta(t+1). \end{aligned} \tag{3.6}$$

The state vector Z, unfortunately, has dimension $O(N)$: this is computationally intractable for data structures with large N and T. Conceptually more important, however, such a state space representation (3.6) simply describes N time series in terms of an $O(N)$-dimensional vector process—certainly neither useful nor insightful. The reason for this is clear: from (3.5), lagged X_j's are part of the state description for each X_j due to the β_j's being nontrivial, i.e., due to the serial correlation in each idiosyncratic disturbance.[2]

The solution to this is to notice that a large part of the state vector happens to be directly observable; the unobservable part turns out to have dimension only $O(K)$, independent of N. Exploiting this structure gives Kalman smoother projections and moment matrix calculations that are, effectively, independent of N.

The assertions of the previous two paragraphs are detailed in the Technical Appendix.

The intuition then is that increasing the cross-section dimension N can only help to estimate more precisely (the conditional expectations of) the (unobserved part of the) state and its cross moments. This must imply the same property for estimates of the parameters of interest. Notice further that deleting entries of X leaves invariant the orthogonality properties on an appropriate reduced version of (3.6). Thus, if the model is correctly specified, estimators that exploit the orthogonality conditions in (3.6) remain consistent independent of N; at the same time, smaller N-systems must imply less efficient estimators. This observation is useful for building up estimates for a large N system from (easier to estimate) smaller systems; additionally, it suggests a Wu-Hausman-type specification test for these index structures.[3]

Next, when the unknown distribution of (X, U, Y, W) generates a conditional

---

[2] This serial correlation pervades our model in particular and business cycle data in general. It is this feature that distinguishes our work from many EM algorithm/unobservable common factor applications in finance, e.g., Lehmann and Modest (1988), where the data are either only a large cross-section or, alternatively, taken to be serially independent.

[3] We do not follow up this testing suggestion below, but leave it for future work.

expectation $E(U \mid X, W)$ that equals the linear projection of U on X and W, then standard Kalman smoother calculations yield the conditional expectations $E\big(Z(t)Z(t)' \mid X, W\big)$ and $E\big(Z(t) \mid X, W\big)$, taking as given a particular setting for the unknown parameters. We will assume that the underlying unknown distribution does in fact fall within such a class.[4]

Iteration on this scheme is of course simply the EM algorithm and, under weak regularity conditions, is guaranteed to converge to a point that solves the first order (projection) conditions. Estimation of the dynamic index model for random fields with large N and T is thus seen to be feasible.

4. An Application: Sectoral Employment

This section gives a sectoral, employment-disaggregated description of US economic fluctuations as interpreted by our index-model structure. We consider the behavior of annual full-time equivalent (FTE) employment across 60 industries.

The National Income and Product Accounts report on an annual basis the number of full time equivalent workers in different industries (NIPA, Table 6.7b). Excluding the government sector and US workers employed outside this country, there are 60 industrial sectors for which these data are available from 1948 through 1989. The industries range from Farming, Metal Mining, and Coal Mining through Motion Pictures, Legal Services, and Private Household Services. (The entire set of 60 industries—complete with Citibase codes—is reported in the Data Appendix.) Figure 4.1 gives a 3-dimensional depiction of these FTE employment data; the *Sectors* axis in this graph arrays the different industries in the order given in NIPA Table 6.7b and the Data Appendix. The vertical axis is the log (times 10) of

---

[4] Elliptically symmetrical distributions would be one such class. Some researchers, such as Hamilton (1989), prefer to work with state vectors that have a discrete distribution—this of course need not invalidate the hypothesized equality of conditional expectations and linear projections, although it does make that coincidence less likely. In principle, one can directly calculate those conditional expectations using only laws of conditional probability. Thus, any simplification—whether by using linear projections or discrete distributions—is not conceptual but only computational.

employment in each sector. It is possible to give a more traditional time series plot of these 60 series. Doing so, however, reveals very little information: the different time series lines quickly merge and the page is simply awash in black.

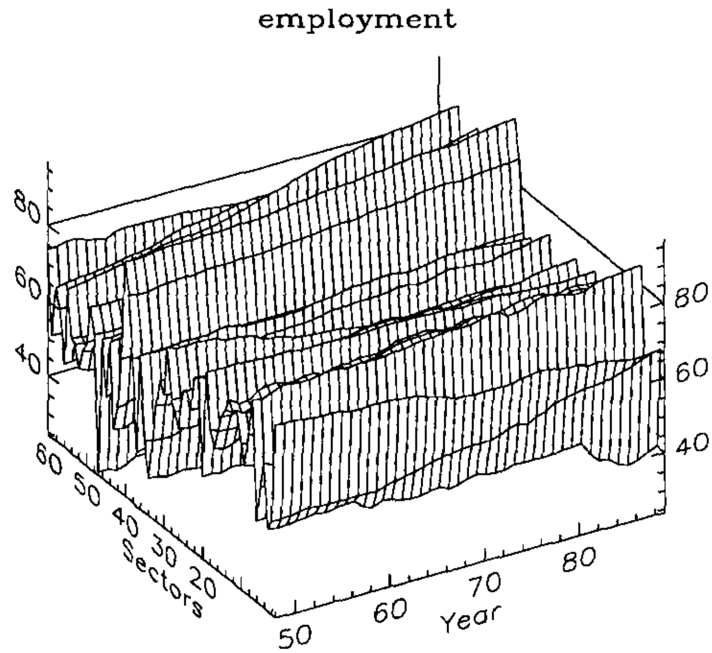

Figure 4.1: *Log (×10) of employment, across sector and time*

Since these employment figures are only available annually from 1948 through

1989, our data are larger in the cross section than in the time series dimension ($60 = N > T = 42$). No full rank (60×60) covariance matrix estimator is available; consequently no full-rank spectral density estimator can be formed. From Figure 4.1 it is also clear that the data are trending, with potentially differing orders of stochastic and deterministic permanent components. Again by $N > T$, no cointegrating regression could be calculated; no cointegration tests could be performed.

Figures 4.2 through 4.4 explore the extent to which the cross-correlation across sectors can be captured by two observable measures typically used by empirical researchers: first, (the log of annual) real GNP, and second, (the log times 10 of) total—equivalently, average—employment across the sixty sectors. When we refer to total employment subsequently we mean this second series and not total US employment. These figures plot residual sample standard deviations from second-order autoregressions, fitted independently across individual sectors, over 1951–1987, and always including a constant and time trend. (In a study with as high dimensionality as this one, presenting alternative specifications—varying lag lengths, for instance—quickly becomes awkward; the signal-noise ratio in presentation falls rapidly and dramatically. Unless stated otherwise, the main conclusions hereafter should be taken as robust across small lag length increases.)

Figure 4.2 graphs on its vertical axis the residual sample standard deviation when GNP at lags -2 through 2 are included as additional regressors, versus that without on the horizontal axis.[5] By least-squares algebra, no point in Figure 4.2 can lie above the 45° line. The further, however, that points fall *below* the 45° line, the more successfully does GNP—common to all sectors—explain employment in each sector. From this graph we conclude that aggregate GNP does appear to be an important common component in sectoral employment fluctuations.

Figure 4.3 is the same as 4.2, but replaces GNP with total employment across

[5] Notice that these regressions include past, present, and future values of GNP. Below, we shall compare the residual variances from these regressions with residual variances from regressions on estimates of our indexes. Those estimates are projections of our indexes on past, present, and future values of all the employment series.

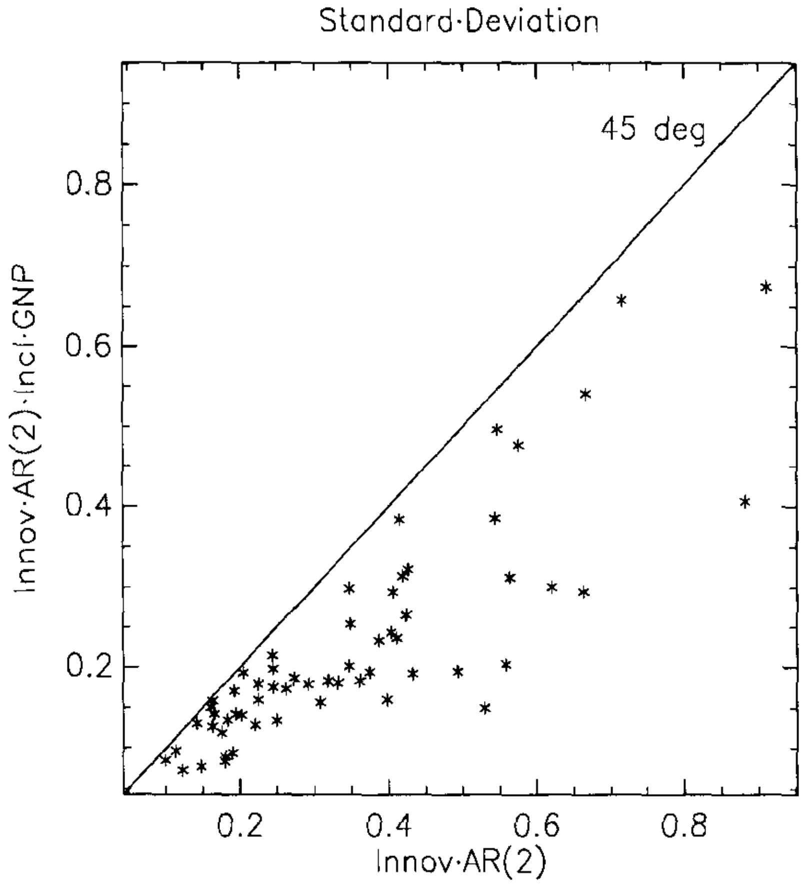

Figure 4.2: Residual standard deviations projecting on own two lags plus GNP, against those projecting on just own two lags

the sixty sectors. The message remains much the same. There appear to be common comovements in sectoral employment, and those comovements are related to aggregate GNP and total employment movements. We emphasize that in the regressions above, both lagged and lead aggregate measures enter as right-hand-side variables. The sample standard deviations increase significantly when lead

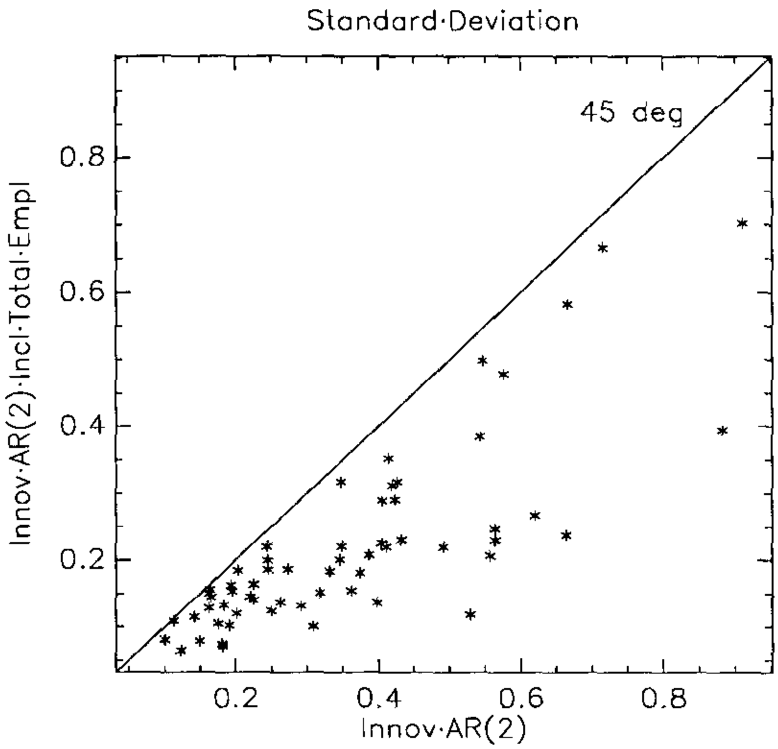

Figure 4.3: Residual standard deviations projecting on own two lags plus total employment, against those projecting on just own two lags

measures are excluded.

Figure 4.4 compares these two measures of common comovements by plotting against each other the sample standard deviations from the vertical axes of Figures 4.2 and 4.3. We conclude from the graph here that both GNP and total employment give similar descriptions of the underlying comovements in sectoral

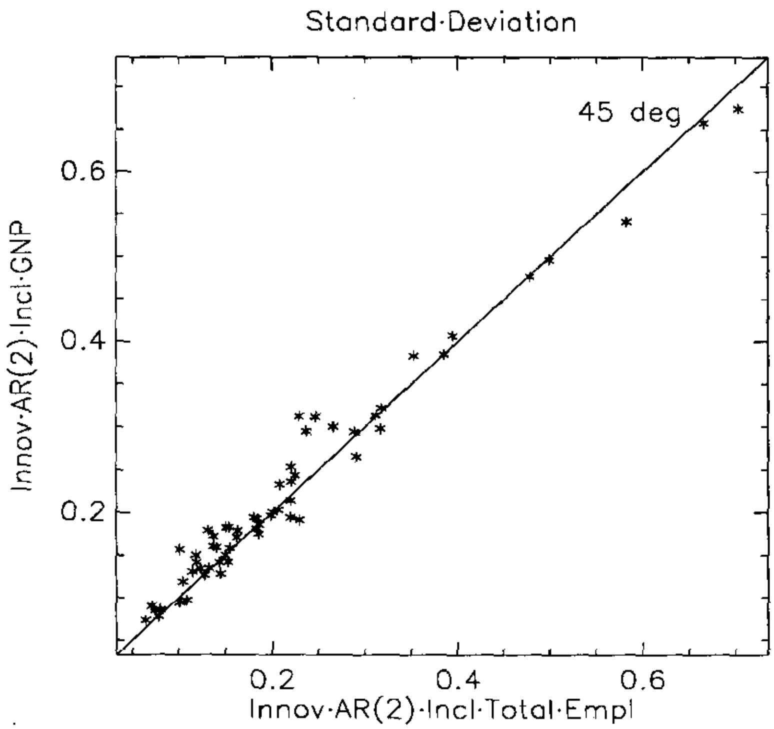

Figure 4.4: Residual standard deviations projecting on own two lags plus GNP, against those projecting on own two lags plus total employment

employment.

We have estimated, for the period 1948 through 1989, one- and two-index representations for sectoral employment.[6] (In the notation of the previous section we take $M_a = 1$, $M_g = 1$, and $M_b = 2$; again, small increases in lag lengths do not affect our conclusions materially.) Figure 4.5 plots standard deviations of the innovations in the idiosyncratic disturbances $\epsilon_j(t)$, under the two-index representation, against the residuals in sector-by-sector projections including GNP (i.e., the vertical axis of Figure 4.2). In other words, the vertical axis describes the innovations upon removing two common unobservable indexes, and imposing extensive orthogonality conditions; the horizontal axis describes the innovations upon removing the single index that is GNP, and without requiring the resulting innovations to be orthogonal across sectors.[7] Since the models are not nested there is no necessity for the points to lie in any particular region relative to the 45° line. Again, however, to the extent that these points fall below that line, we can conclude that the comovements are better described by the model represented on the vertical axis than that on the horizontal.

In this Figure, twelve sectors lie marginally above the 45° line, six marginally below, and the remainder quite a bit below. Overall we conclude that the two unobservable indexes provide a better description of underlying commonalities in sectoral employment than does aggregate GNP.

Figure 4.6 is the same as Figure 4.5, but it replaces GNP by total employment. We draw much the same conclusions from this as the previous graph.

Figure 4.7 replaces the horizontal axes of Figures 4.5 and 4.6 with the standard deviation of idiosyncratic innovations from a single-index representation. Notice that the improvement in fit of the additional index is about the same order of magnitude as that of the two-index representation over either of the single observable aggregates.

We do not present here the calculations we have performed comparing the

---

[6] All index-model and regression calculations and all graphs in this paper were executed using the authors' time-series, random-fields econometrics shell **tsrf**.

[7] Thus, the vertical dimension of Figure 4.5 contains enough information to compute a normal quasi-likelihood value for the unobservable index model; for the horizontal dimension, however, the cross-correlations are nonzero and cannot, as a whole, be consistently estimated.

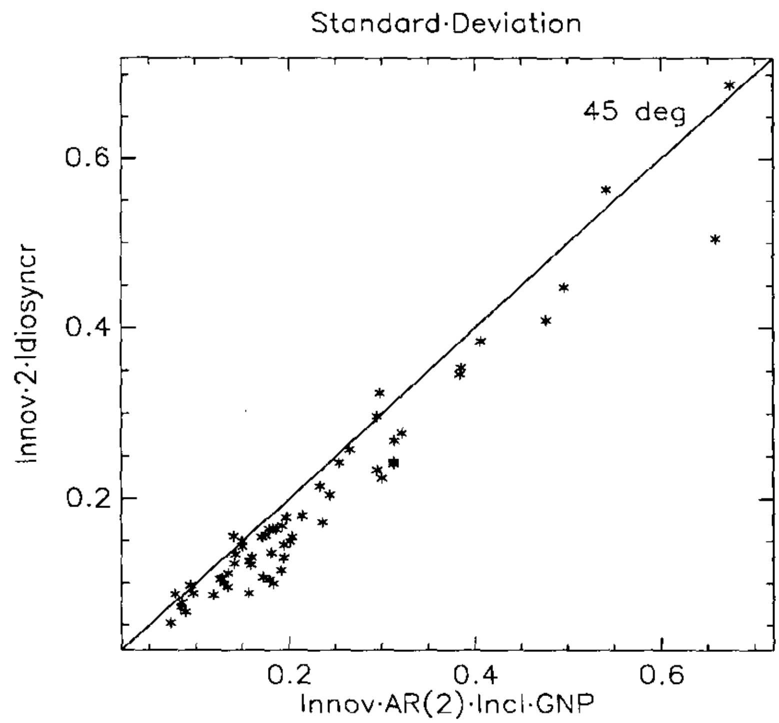

Figure 4.5: Standard deviations of idiosyncratic disturbances in two-index model, against those of the residuals in the projection on own two lags plus GNP

single unobservable index with the two observable aggregates. The calculations show much what one would expect from the analysis thus far. Aggregate GNP and average employment are about as good descriptions of sectoral comovements

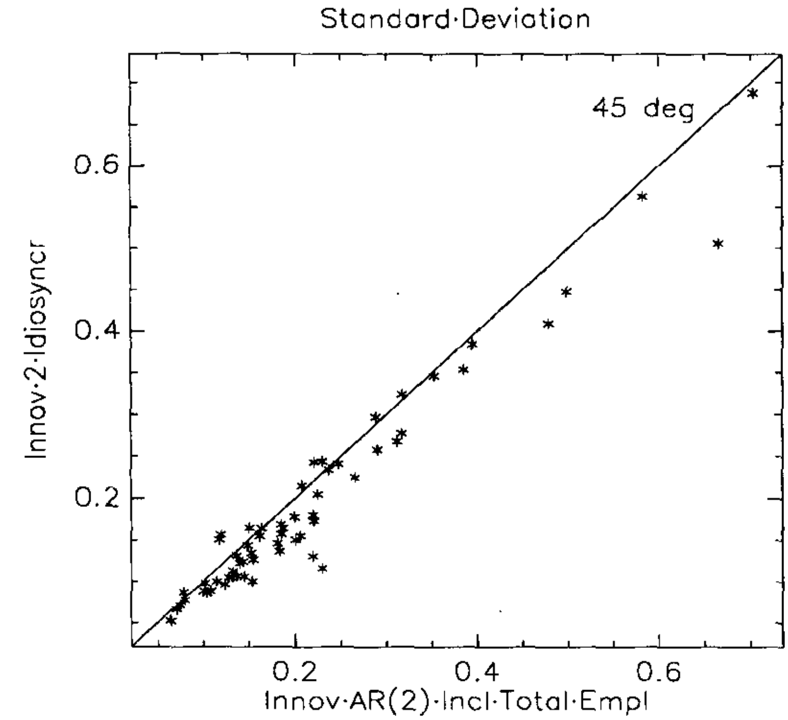

Figure 4.6: Standard deviations of idiosyncratic disturbances in two-index model, against those of the residuals in the projection on own two lags plus total employment

as is the single unobservable index model.[8]

[8] This finding is related to patterns detected in previous applications of unobservable index models to aggregate U.S. time series. For example, Sargent and

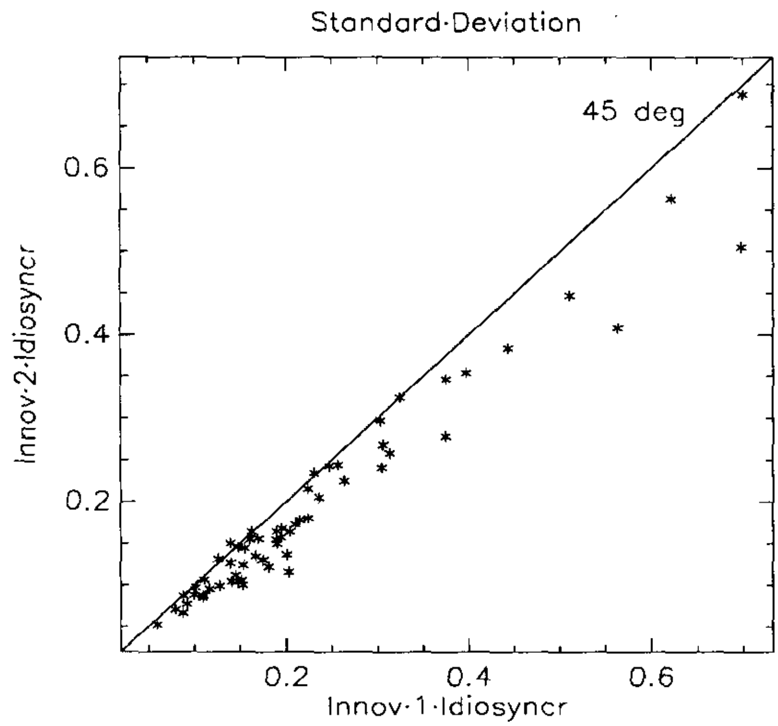

Figure 4.7: Standard deviations of idiosyncratic disturbances in two-index model, against those in one-index model

So what have we learned up until now? If one were concerned only about goodness-of-fit in describing the commonalities in sectoral employment fluctua-

Sims (1977) encountered so-called Heywood solutions at low frequencies: in those solutions the coherence of GNP with one of the indices is unity.

tions, then one might well simply use just total GNP or average employment.[9] In our view, index models should be motivated by a Burns-Mitchell kind of 'pre-GNP accounting,' dimensionality-restricted modelling. The idea is to allow a one-dimensional measure of 'business activity' to emerge from analyzing long lists of different price and quantity time series. We have implemented one version of this vision, and arrived at projection estimators of one- and two-index models. In computing these estimated indexes, we have used only employment series. Now we wish to see how, in particular dimensions, our estimated indexes compare with GNP, the most popular single 'index' of real macroeconomic activity. GNP is, of course, constructed using an accounting method very different in spirit from that used by us and other workers in the Burns-Mitchell tradition.

Thus, we turn to some projections designed to explore this difference. A regression (over 1952 through 1989) of GNP *growth rates* on a constant and first differences of our fitted two indexes, lags 0 through 3, gives an R^2 of 83%. A similar regression using just the single index (estimated from the one-index model), again lags 0 through 3, gives an R^2 of 72%. GNP growth rates are thus highly correlated with the estimated index process. This correlation captures the sense in which a purely mechanical low-dimensional index construction yields an aggregate that closely tracks GNP growth.[10]

We have also experimented with vector autoregressions in GNP growth, average employment growth, and first differences of the indexes. Exclusion tests of lag blocks of the indexes here, however, *cannot* be viewed as Granger causality tests because the estimated indexes are themselves two-sided distributed lags of employment. It would be possible to use the estimated parameters of our index

---

[9] We should emphasize again that this is true only when one proxies for those common comovements using both leads and lags of these indicators. As already stated above, the projection residuals get considerably larger when future aggregates are excluded.

[10] Note that sample versions of these indexes are estimated by a Kalman smoothing procedure and therefore use observations on sectoral employment over the entire sample, past and future. This is also why in Figures 4.2 on, we always used future and past observable aggregates to make the comparison fairer.

representations to construct one-sided-on-the-past index projections. These one-sided projections could then be used to conduct Granger causality tests. We have not done this here, but think that it would be a useful exercise.

In concluding this empirical section, we judge that the unobservable index application to the large cross-section here has been, in the main, successful. The empirical results here encourage us to be optimistic about continued use of large cross-sections for dynamic analysis. There are, however, dimensions along which relative failure might be argued. Most notably, the refinement in the description of commonalities given by the unobservable index model was not spectacular relative to that given by observable measures, such as aggregate GNP or total employment. While the two-index representation is a better description of sectoral employment fluctuations, relative to a single-index representation, not both indexes turn out to be equally important for predicting GNP. For this exercise, the tighter parametrization implicit in a single index appears to dominate the marginal increase in information from a two-index representation. These failures should be contrasted with our two principal successes: (i) the tractability of our extension of standard index model analysis—simultaneously to encompass differing nonstationarities, large cross-section dimensions, and extensive orthogonality restrictions; and (ii) the strong informational, predictive content for GNP of our estimated common indexes.

Unlike in research on interpreting unobservable disturbances—such as in VAR studies—we do not attempt here to name the two unobservable factors. Rather, the goal of the present paper has been to carry out the methodological extensions in (i) of the previous paragraph, and to examine the forecasting properties of the resulting common factors in (ii), also of the previous paragraph. Future work could, in principle, attempt the same exercise for the unobservable index models here as have been performed for VAR representations elsewhere.

5. Conclusion

We have provided a framework for analyzing comovements—aggregate dynamics—in random fields, i.e., data where the number of cross-section time series is comparable in magnitude to the time length. We have shown how, upon reinterpretation, standard techniques can be used to estimate such index models.

In applying the model to estimate aggregate dynamics in employment across different sectors, we discovered that a model with two common factors turns out to fit those data surprisingly well. Put differently, much of the observed fluctuations in employment in those many diverse industries are well explained by disturbances that are perfectly correlated across all the sectors.

The econometric structure we have used here seems to us quite rich and potentially capable of dealing with many different interesting questions. Among others, this includes issues of (a) the relative importance of different kinds of disturbances, e.g., Long and Plosser (1983) and Prescott (1986); (b) convergence across economic regions, e.g., Barro and Sala-i-Martin (1991) and Blanchard and Katz (1992), (c) aggregate and sectoral co-movements, e.g., Abraham and Katz (1986), Lilien (1982), and Rogerson (1988), and (d) location dynamics, conditioning on exogenous policy variables, e.g., Papke (1989). Previous work, however, have used measurement and econometric techniques that differ substantially from that which we propose here; clearly, we think our procedure is closer in spirit to the relevant economic ideas. Questions on the appropriate definition and measurement of inflation, comovement in consumption across economic regions, the joint behavior of asset prices observed for many assets and over long periods of time, can all be coherently dealt with in our framework. In future research we intend to apply our techniques to these and similar, related issues.

6. Technical Appendix

This appendix constructs a state-space representation for the model of Section 3; this is needed to compute the Kalman-smoothed projections that are in turn used in applying the EM algorithm to our estimation problem.

Recall that $\phi_{jk} = \beta_j \alpha_{jk}$ and that the maximum lags on β_j and α_{jk} are M_b and M_a respectively. Define $M_f = M_b + M_a$; this is the maximum lag on ϕ_{jk}. From M_g, the maximum lag on Γ in (3.3), define $M_h = \max(M_f, M_g)$, and let:

$$\tilde{U}(t) = \left(U(t)', U(t-1)', \ldots, U(t-M_h)'\right)'$$

and conformably

$$\tilde{\eta}_U(t) = \left(\eta_U(t)', 0', \ldots, 0'\right)'.$$

Then write (3.3) in first order form as:

$$\tilde{U}(t) = C\tilde{U}(t-1) + \tilde{\eta}_U(t), \tag{6.1}$$

where C is $(1 + M_h) \times K$ square, with the following structure: Call a collection of K successive rows (or columns) a K-row (or column) block. The last M_h K-row blocks comprise simply zeroes and ones, in the usual way, forming identities. Consider then the k-th row ($k = 1, 2, \ldots, K$) in the first K-row block. The first K-column block of this row vanishes everywhere except in the k-th position where it contains $g_k(1)+1$; the $(1+M_g)$-th K-column block of this row vanishes everywhere except in the k-th position where it contains $-g_k(M_g)$. For K-column block m ($m = 2, 3, \ldots, M_g$), the entries again vanish everywhere except in the k-th position where they equal $g_k(m) - g_k(m-1)$. This pattern of coefficents comes from (3.3)'s being in first differences whereas (6.1) is in levels.

Turning to the observables, write

$$X(t) = \left(X_1(t), X_2(t), \ldots, X_N(t)\right)';$$

since this is observed at time t, we denote

$$\mathit{info}(t) = \left\{X(t), X(t-1), X(t-2), \ldots\right\}.$$

Let

$$\tilde{X}(t) = \left(X(t)', X(t-1)', \ldots, X(t-M_b+1)'\right)'.$$

Also, write out $\phi_{jk}(L)$ explicitly as

$$\phi_{jk}(L) = \sum_{m=0}^{M_f} f_{jk}(m) L^m.$$

We can then rewrite equation (3. 5) as

$$X(t) = A \begin{pmatrix} \tilde{U}(t) \\ \tilde{X}(t-1) \end{pmatrix} + dW(t) + \epsilon(t) = aZ(t) + dW(t) + \epsilon(t), \qquad (6.2)$$

where d has rows formed from d_j, and A has the following structure. Its j-th row has the first $(1+M_f)K$ entries given by:

$$\left(f_{j1}(0), \cdots, f_{jK}(0), f_{j1}(1), \cdots, f_{jK}(M_f)\right);$$

the remaining entries in this row differ from zero only in M_b places. After the first $(1+M_h)K$ entries (not first $(1+M_f)K$) there are M_b N-row blocks. Each such block vanishes except in the j-th entry which equals $b_j(m)$, for $m = 1, 2, \ldots, M_b$.

Now augment the transition equation (6.1) with $\tilde{X}$ and W, i.e., write:

$$Z(t+1) = cZ(t) + \delta W(t) + \eta(t+1) \qquad (6.3)$$

with

$$c = \begin{pmatrix} C & 0 \\ c_{21} & c_{22} \end{pmatrix}$$

where the first N rows of $\begin{pmatrix} c_{21} & c_{22} \end{pmatrix}$ contain a and the remaining rows comprise just 0's and 1's; the matrix δ vanishes everywhere except in the N rows after the first $(1+M_h)K$—in those N rows it is comprised of d; and finally,

$$\eta(t+1) = \begin{pmatrix} \tilde{\eta}_U(t+1) \\ \epsilon(t) \\ 0 \end{pmatrix}.$$

Notice that (6.3) contains within it (6.2).

Since all but the first $(1+M_h)K$ entries of $Z(t+1)$ are observed at time t, it will be natural to let $P\big(Z(t+1) \mid info(t)\big)$ equal $\tilde{X}(t)$ except in its first $(1+M_h)K$ entries. Similarly, we will choose the conditional mean square $\mathrm{Var}(Z(t+1) \mid info(t))$ to be zero everywhere except in the leading $(1+M_h)K$ square diagonal block.

In summary our state space representation is:

$$Z(t+1) = cZ(t) + \delta W(t) + \eta(t+1);$$
$$X(t) = AZ(t) + dW(t) + \epsilon(t)$$

with $\mathrm{Var}(Z(1) \mid info(0))$ initialized to vanish everywhere except in the leading $(1+M_h)K$ diagonal block. The disturbance vector $(\eta(t+1)', \epsilon(t)')'$ is serially uncorrelated and has covariance matrix:

$$\Omega = \begin{pmatrix} \Omega_\eta & \Omega_{\eta\epsilon} \\ \cdot & \Omega_\epsilon \end{pmatrix},$$

where Ω_η is singular, $\Omega_{\eta\epsilon}$ contains Ω_ϵ, and Ω_ϵ is $N \times N$ diagonal. We write $\big(\Omega_{\eta\epsilon}\big)_1$ to denote the first $(1+M_h)K$ rows of $\Omega_{\eta\epsilon}$; and Ω_{η_1} to denote the leading $(1+M_h)K$ diagonal block of Ω_η.

Partition $Z(t)$ as $\big(Z_1(t)', Z_2(t)'\big)'$ where

$$Z_1(t) = \tilde{U}(t) \qquad (1+M_h)K \times 1$$

and

$$c = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \qquad \text{with } c_{11} = C \ (1+M_h)K \text{ square and } c_{12} = 0.$$

Thus $\eta(t)$ is also partitioned into $\big(\eta_1(t)', \eta_2(t)'\big)'$ with $\eta_1 = \tilde{\eta}_U$ $((1+M_h)K \times 1)$. We can now write the measurement equation

$$X(t) = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} Z_1(t) \\ Z_2(t) \end{pmatrix} + \epsilon(t)$$

where a_1 comprises the first $(1+M_h)K$ columns of a. Note that $\mathrm{Var}(Z_1(t)|info(s))$ is thus only $(1+M_h)K$ square, and has dimensions that are invariant to N.

References

[1] Abraham, Katherine G. and Katz, Lawrence F. (1986), "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances", *Journal of Political Economy*, 94(3):507–522, June.

[2] Barro, Robert J. and Sala-i-Martin, Xavier. (1992), "Convergence", *Journal of Political Economy*, 100(2):223–251, April.

[3] Blanchard, Olivier Jean and Katz, Lawrence F. (1992), "Regional Evolutions", *Brookings Papers on Economic Activity*, (1):1–75, April.

[4] Engle, Robert F. and Granger, Clive W. J. (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55(2):251–276, March.

[5] Geweke, John. (1977), "Labor Turnover and Employment Dynamics in US Manufacturing Industry", in Sims, Christopher A., (ed.), *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis, Minneapolis MN 55480.

[6] Geweke, John and Singleton, Kenneth J. (1981), "Maximum Likelihood 'Confirmatory' Factor Analysis of Economic Time Series", *International Economic Review*, 22(1):37–54, February.

[7] Hamilton, James D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57(2):357–384, March.

[8] King, Robert G., Plosser, Charles I., Stock, James H., and Watson, Mark W. (1991), "Stochastic Trends and Economic Fluctuations", *American Economic Review*, 81(4):819–840, September.

[9] Lehmann, Bruce N. and Modest, David M. (1988), "The Empirical Foundations of The Arbitrage Asset Pricing Theory", *Journal of Financial Economics*, 21(2):213–254, sep.

[10] Lilien, David M. (1982), "Sectoral Shifts and Cyclical Unemployment", *Journal of Political Economy*, 90(4):777–793, August.

[11] Long, John B. and Plosser, Charles I. (1983), "Real Business Cycles", *Journal of Political Economy*, 91(1):39–69, February.

[12] Papke, Leslie E. (1989), "Interstate Business Tax Differentials and New Firm Location: Evidence from Panel Data", Working Paper 3184, NBER, Cambridge MA, November.

[13] Prescott, Edward C. (1986), "Theory Ahead of Business Cycle Measurement", *Federal Reserve Bank of Minneapolis Quarterly Review*, pages 9–22, Fall.

[14] Rogerson, Richard. (1987), "An Equilibrium Model of Sectoral Reallocation", *Journal of Political Economy*, 95(4):824–834, August.

[15] Ruud, Paul A. (1991), "Extensions of Estimation Methods Using The EM Algorithm", *Journal of Econometrics*, 49:305–341.

[16] Sargent, Thomas J. and Sims, Christopher A. (1977), "Business Cycle Modelling without Pretending to Have Too Much A Priori Economic Theory", in Sims, Christopher A., (ed.), *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis, Minneapolis MN 55480.

[17] Stock, James H. and Watson, Mark W. (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, 83(404):1097–1107, December.

[18] Stock, James H. and Watson, Mark W. (1990), "Business Cycle Properties of Selected U.S. Economic Time Series", Working Paper 3376, NBER, Cambridge MA, June.

[19] Watson, Mark W. and Engle, Robert F. (1983), "Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC, and Varying Coefficient Regression Models", *Journal of Econometrics*, 23:385–400.

[20] Watson, Mark W. and Kraft, Dennis F. (1984), "Testing the Interpretation of Indices in a Macroeconomic Index Model", *Journal of Monetary Economics*, 13(2):165–181, March.

Data Appendix

Real GNP in the text refers to the annual version of **GNP82** obtained from Citibase July 1991. This series is normalized to constant 1982 dollars, and is no longer available—the new constant-dollar GNP series in Citibase is normalized to constant 1987 dollars, and is available only from 1959. Since the employment data we use goes back through 1948 we decided to use the older **GNP82**.

Next, the sectoral description we study is from:

National Income and Product Accounts: Employment

1. FTE EMPLOYEES BY INDUSTRY (ANNUAL)

Private Industries

1. Farms (**GAFAF1**)
2. Agrc., Forestry, and Fisheries (**GAFAF7**)
3. Metal Mining (**GAFM10**)
4. Coal Mining (**GAFM12**)
5. Oil and Gas Extraction (**GAFM13**)
6. Nonmetallic Minerals, Except Fuels (**GAFM14**)
7. Construction (**GAFCC**)

Manufacturing

Durable Goods

8. Lumber & Wood Products (**GAFD24**)
9. Furniture & Fixtures (**GAFD25**)
10. Stone, Clay, and Glass Products (**GAFD32**)
11. Primary Metal Industries (**GAFD33**)
12. Fabricated Metal Products (**GAFD34**)
13. Machinery, exc. Electr. (**GAFD35**)
14. Electric & Electronic Eq. (**GAFD36**)
15. Motor Vehicles & Equipment (**GAF371**)
16. Other Transportation (**GAFD37**)
17. Instruments and Related Products (**GAFD38**)

18. Misc. Mfg. Industries (**GAFM39**)
Nondurable Goods
19. Food & Kindred Products (**GAFN20**)
20. Tobacco Manu. (**GAFN21**)
21. Textile Mill Products (**GAFN22**)
22. Apparel & Other Textile Prod. (**GAFN23**)
23. Paper & Allied Prod. (**GAFN26**)
24. Printing & Publ. (**GAFN27**)
25. Chemicals & Allied Prod. (**GAFN28**)
26. Petroleum & Coal Prod. (**GAFN29**)
27. Rubber & Misc. Plastic Prod. (**GAFN30**)
28. Leather & Leather Prod. (**GAFN31**)
Transportation & Public Utilities
29. Railroad Transportation (**GAFT40**)
30. Local & Interurban Passenger Transit (**GAFT41**)
31. Trucking & Warehousing (**GAFT42**)
32. Water Transportation (**GAFT44**)
33. Transportation by Air (**GAFT45**)
34. Pipelines, except Gas (**GAFT46**)
35. Transportation Services (**GAFT47**)
Communication
36. Telephone & Telegraph (**GAF481**)
37. Radio & TV Broadcastings (**GAF483**)
38. Electric, Gas, & Sanitary Svcs. (**GAFUT**)

39. Wholesale Trade (**GAFW**)
40. Retail Trade (**GAFR**)
Finance, Insurance, and Real Estate
41. Banking (**GAFF60**)
42. Credit Agencies other than Banks (**GAFF61**)
43. Security & Commodity Brokers (**GAFF62**)
44. Insurance Carriers (**GAFF63**)
45. Insurance Agents and Brokers (**GAFF64**)

46. Real Estate (**GAFF65**)
47. Holding & Other Investment Co. (**GAF67F**)

Services

48. Hotels and Other Lodging Places (**GAFS70**)
49. Personal Services (**GAFS72**)
50. Business Services (**GAFS73**)
51. Auto Repair Svcs., & Garages (**GAFS75**)
52. Misc. Repair Svcs. (**GAFS76**)
53. Motion Pictures (**GAFS78**)
54. Amusement & Recr. Svc. (**GAFS79**)
55. Health Services (**GAFS80**)
56. Legal Services (**GAFS81**)
57. Educational Services (**GAFS82**)
58. Social Svcs. & Membership Org. (**GAFS86**)
59. Misc. Professional Svcs. (**GAFS89**)
60. Private Households (**GAFS88**)

List of Figure Legends