Economic Quarterly—Volume 95, Number 4—Fall 2009—Pages 335–355

Heterogeneity in Sectoral Employment and the Business Cycle

Nadezhda Malysheva and Pierre-Daniel G. Sarte

This paper uses a factor analytic framework to assess the degree to which agents working in different sectors of the U.S. economy are affected by common rather than idiosyncratic shocks. Using Bureau of Labor Statistics (BLS) employment data covering 544 sectors from 1990–2008, we first document that, at the aggregate level, employment is well explained by a relatively small number of factors that are common to all sectors. In particular, these factors account for nearly 95 percent of the variation in aggregate employment growth. This finding is robust across different levels of disaggregation and accords well with Quah and Sargent (1993), who perform a similar analysis using 60 sectors over the period 1948–1989 (but whose methodology differs from ours), as well as with Foerster, Sarte, and Watson (2008), who carry out a similar exercise using data on industrial production.[1]

Interestingly, while common shocks represent the leading source of variation in aggregate employment, the analysis also suggests that this is typically not the case at the individual sector level. In particular, our results indicate that across all goods and services, common shocks explain on average only 31 percent of the variation in sectoral employment. In other words, employment at the sectoral level is driven mostly by idiosyncratic shocks, rather than common shocks, to the different sectors. Put another way, it is not the case that "a rising tide lifts all boats." Moreover, it can be easy to overlook the influence of idiosyncratic shocks since these tend to average out in aggregation.

We wish to thank Kartik Athreya, Sam Henly, Andreas Hornstein, and Thomas Lubik for helpful comments. The views expressed in this article do not necessarily represent those of the Federal Reserve Bank of Richmond, the Board of Governors of the Federal Reserve System, or the Federal Reserve System. All errors are our own.

[1] See also Forni and Reichlin (1998) for an analysis of ouput and productivity in the United States between 1958 and 1986.

Despite the general importance of idiosyncractic shocks in explaining movements in sectoral employment, we nevertheless further document substantial differences in the way that sectoral employment is tied to these shocks. Specifically, we identify sectors where up to 85 percent of the variation in employment is driven by the common shocks associated with aggregate employment variations. Employment in these sectors, therefore, is particularly vulnerable to the business cycle with little in the way of idiosyncratic shocks that might be diversified away. These sectors are typically concentrated in construction and include, for example, residential building.

More generally, our empirical analysis indicates that employment in goods-producing industries tends to more tightly reflect changes in aggregate conditions relative to service-providing industries. However, even within the goods-producing industries, substantial heterogeneity exists in the way that sectoral employment responds to common shocks. For instance, the durable goods and construction industries are significantly more influenced by common shocks than the nondurable goods and mining industries. Among the sectors where employment is least related to aggregate conditions are government, transportation, and the information industry.

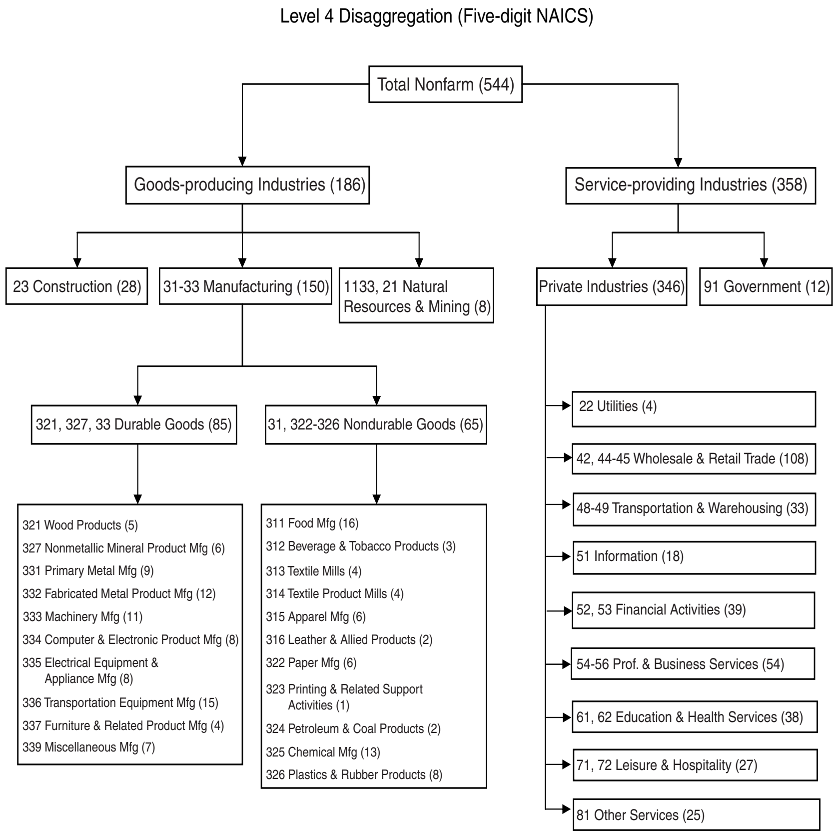
Finally, we present evidence that the factors uncovered in our empirical work play substantially different roles in explaining aggregate and sectoral variations in employment. Although the findings we present are based on a three-factor model, our analysis suggests that one factor is enough to explain roughly 94 percent of the variation in aggregate employment. At the same time, however, that factor appears almost entirely unrelated to employment movements in specific sectors such as in natural resources and mining or education and health services. Interestingly, the reverse is also true in the sense that the analysis identifies factors that significantly help track employment movements in these particular sectors but that play virtually no role in explaining aggregate employment fluctuations.

This article is organized as follows. Section 1 provides an overview of the data. Section 2 describes the factor analysis and discusses key summary statistics used in this article. Section 3 summarizes our findings and Section 4 offers concluding remarks.

1. OVERVIEW OF THE DATA

Our analysis uses data on sectoral employment obtained from the BLS covering the period 1990–2008. The data are available monthly, seasonally adjusted, and disaggregated by sectors according to the North American Industry Classification System (NAICS). Our data cover the period since 1990, the date at which this classification system was introduced. Prior to 1990, BLS employment data were disaggregated using Standard Industry Classification codes, which involve a lower degree of disaggregation and were discontinued

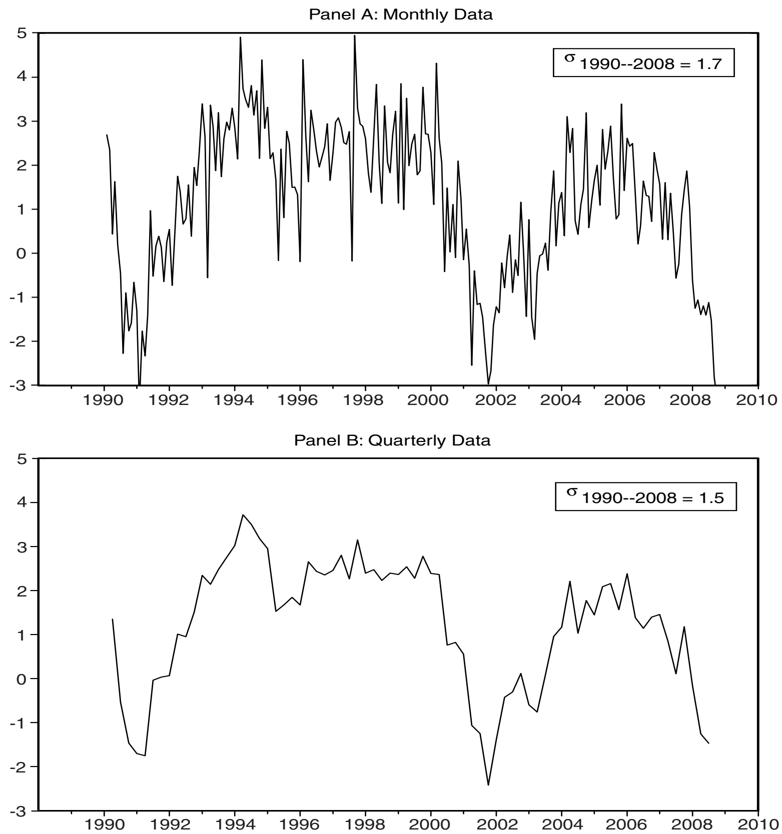
Figure 1 Breakdown of Sectoral Employment Data



as of 2002. For most of the article, we use a five-digit level of disaggregation that corresponds to 544 sectors, although our findings generally apply to other levels of disaggregation as well. The raw data measure the number of employees in different sectors, from which we compute sectoral employment growth rates as well as the relative importance (or shares) of industries in aggregate employment.

When aggregated, the data measure total nonfarm employment. Nonfarm employment is further subdivided into two main groups: goods-producing sectors, comprising 186 sectors at the five-digit level, and service-providing sectors, comprising 358 sectors. The goods-producing sectors are further subdivided into three main categories: construction, with 28 sectors; manufacturing, with 150 sectors; and natural resource and mining, with eight sectors. The manufacturing component of the goods sector contains two main categories: durable goods, comprising 85 sectors, and nondurable goods, with 65 sectors.

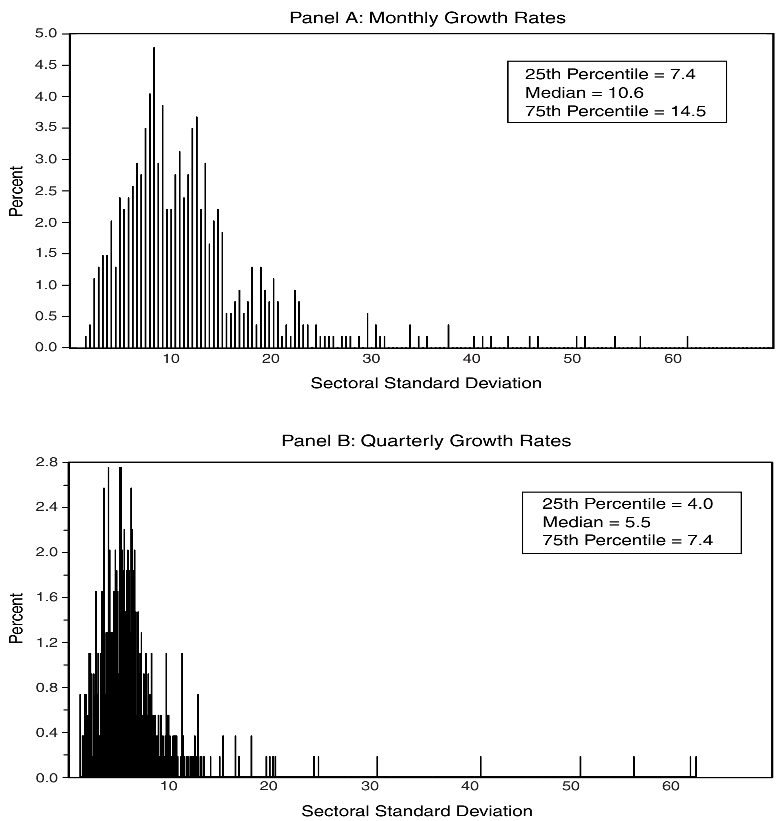
Figure 2 Monthly and Quarterly Employment, All Goods and Services



The service-providing sectors employ more than four times as many workers as the goods-producing sectors. They are made up of two main components: government, with 12 sectors, and a variety of private industries that include 346 sectors spanning wholesale and retail trade, information, financial activities, education and health, as well as many other services. Figure 1 illustrates a breakdown of our sectoral data, along with the number of industries within each broad category of sectors in parenthesis, as well as their corresponding NAICS codes.

Let e_t denote aggregate employment across all goods- and services-producing industries at date t, and let e_{it} denote employment in the i^{th} industry. We construct quarterly values for employment as averages of the months in the quarter. We further denote aggregate employment growth by Δe_t and employment growth in industry, i, by Δe_{it}. At the monthly frequency, we compute Δe_{it} as $1,200 \times \ln(e_{it}/e_{it-1})$ and, at the quarterly frequency, as

Figure 3 Distribution of Standard Deviations of Sectoral Growth Rates (1990–2008)



$400 \times \ln(e_{it}/e_{it-1})$. Aggregate employment growth is computed similarly. Finally, we represent the $N \times 1$ vector of sectoral employment growth rates, where N is the number of sectors under consideration, by $\Delta\mathbf{e}_t$.

Figures 2A and 2B illustrate the behavior of aggregate employment growth at the monthly and quarterly frequencies, respectively, over our sample period. Monthly aggregate employment growth is somewhat more volatile than quarterly employment growth, but in either case the recessions of 1991 and 2001 stand out markedly. At a more disaggregated level, Figures 3A and 3B show the distributions of standard deviations of both monthly and quarterly sectoral employment growth across all 544 sectors. As with aggregate data, quarterly averaging reduces the volatility of sectoral employment. More importantly, it is clear that there exists substantial heterogeneity across sectors in the sense

Table 1 Standard Deviation of Employment Growth Rates

	Monthly Growth Rates	Quarterly Growth Rates
Full Covariance Matrix	1.8	1.5
Diagonal Covariance Matrix	0.7	0.4

Notes: The table reflects percentage points at an annual rate.

that fluctuations in employment are unequivocally more pronounced in some sectors than others.

Let s_i denote the (constant mean) share of sector i's employment in aggregate employment and the corresponding $N \times 1$ vector of sectoral shares be denoted by $\mathbf{s}$. Then, we can express aggregate employment growth as $\Delta e_t = \mathbf{s}'\Delta\mathbf{e}_t$. Furthermore, it follows that the volatility of aggregate employment growth in Figure 2, denoted σ_e^2, is linked to individual sectoral employment growth volatility in Figure 3 through the following equation,

$$\sigma_e^2 = \mathbf{s}'\Sigma_{ee}\mathbf{s}, \tag{1}$$

where Σ_{ee} is the variance-covariance matrix of sectoral employment growth. Thus, we can think of the variation in aggregate employment as driven by two main forces—individual variation in sectoral employment growth (the diagonal elements of Σ_{ee}) and the covariation in employment growth across sectors (the off-diagonal elements of Σ_{ee}).[2]

Table 1 presents the standard deviation of aggregate employment, σ_e^2, computed using the full variance-covariance matrix Σ_{ee} in the first row, and using only its diagonal elements in the second row. As stressed in earlier work involving sectoral data, notably by Shea (2002), it emerges distinctly that the bulk of the variation in aggregate employment is associated with the covariance of sectoral employment growth rates rather than individual sector variations in employment. The average pairwise correlation in sectoral employment is positive at approximately 0.10 in quarterly data and 0.04 in monthly data. Moreover, if one assumed that the co-movement in sectoral employment growth is driven primarily by aggregate shocks, then Table 1 would immediately imply that these shocks represent the principal source of variation in aggregate employment. For example, focusing on quarterly growth rates, the fraction of aggregate employment variability explained by aggregate shocks would amount roughly to $1 - (0.4^2/1.5^2)$ or 0.93. This calculation, of course, represents only an approximation in the sense that the diagonal elements of Σ_{ee} would themselves partly reflect the effects of changes

[2] As in Foerster, Sarte, and Watson (2008), time variation in the shares turns out to be immaterial for the results we discuss in this article.

in aggregate conditions. That said, it does suggest, however, that despite clear differences in employment growth variability at the individual sector level, these differences, for the most part, vanish in aggregation and so become easily overlooked.

2. A FACTOR ANALYSIS OF SECTORAL EMPLOYMENT

As discussed in Stock and Watson (2002), the approximate factor model provides a convenient means by which to capture the covariability of a large number of time series using a relatively few number of factors. In terms of our employment data, this model represents the $N \times 1$ vector of sectoral employment growth rates as

$$\Delta \mathbf{e}_t = \boldsymbol{\lambda} \mathbf{F}_t + \mathbf{u}_t, \tag{2}$$

where $\mathbf{F}_t$ is a $k \times 1$ vector of unobserved factors common to all sectors, $\boldsymbol{\lambda}$ is an $N \times k$ matrix of coefficients referred to as factor loadings, and $\mathbf{u}_t$ is an $N \times 1$ vector of sector-specific idiosyncratic shocks that have mean zero. We denote the number of time series observations in this article by T. Using (1), the variance-covariance matrix of sectoral employment growth is now simply given by

$$\Sigma_{ee} = \boldsymbol{\lambda} \Sigma_{FF} \boldsymbol{\lambda}' + \Sigma_{uu}, \tag{3}$$

where Σ_{FF} and Σ_{uu} are the variance-covariance matrices of $\mathbf{F}_t$ and $\mathbf{u}_t$, respectively.

In classical factor analysis, Σ_{uu} is diagonal so that the idiosyncratic shocks are uncorrelated across sectors. Stock and Watson (2002) weaken this assumption and show that consistent estimation of the factors is robust to weak cross-sectional and temporal dependence in these shocks. Equation (2) can be interpreted as the reduced form solution emerging from a more structural framework (see Foerster, Sarte, and Watson 2008). Given this, features of the economic environment that might cause the "uniquenesses," $\mathbf{u}_t$, to violate the weak cross-sectional dependence assumption include technological considerations, such as input-output (IO) linkages between sectors or production externalities across sectors. In either case, idiosyncratic shocks to one sector may propagate to other sectors via these linkages and give rise to internal co-movement that is ignored in factor analysis. Using sectoral data on U.S. industrial production, Foerster, Sarte, and Watson (2008) show that the internal co-movement stemming from IO linkages in a canonical multisector growth model is, in fact, relatively small. Hence, the factors in that case capture mostly aggregate shocks rather than the propagation of idiosyncratic shocks by way of IO linkages. Thus, for the remainder of this article, we shall interpret $\mathbf{F}_t$ as capturing aggregate sources of variation in sectoral employment.

When N and T are large, as they are in this article, the approximate factor model has proved useful because the factors can simply be estimated by principle components (Stock and Watson 2002). By way of illustration, the Appendix provides a brief description of the principle component problem and its relationship to the approximate factor model (2). Bai and Ng (2002) further show that penalized least-square criteria can be used to consistently estimate the number of factors, and the estimation error in the estimated factors is sufficiently small that it need not be taken into account in carrying out variance decomposition exercises (Stock and Watson 2002).

Key Summary Statistics

Given equation (2), we shall summarize our findings in mainly two ways. First, we compute the fraction of aggregate employment variability explained by aggregate or common shocks, which we denote by $R^2(\mathbf{F})$. In particular, since $\Delta e_t = \mathbf{s}'\Delta\mathbf{e}_t = \mathbf{s}'\boldsymbol{\lambda}\mathbf{F}_t + \mathbf{s}'\mathbf{u}_t$, we have that

$$R^2(\mathbf{F}) = \frac{\mathbf{s}'\boldsymbol{\lambda}\Sigma_{FF}\boldsymbol{\lambda}'\mathbf{s}}{\sigma_e^2}. \tag{4}$$

For the 544 sectors that make up all goods and services at the five-digit level, $R^2(\mathbf{F})$ then captures the degree to which fluctuations in aggregate employment growth are driven by aggregate rather than sector-specific shocks. Second, we also assess the extent to which aggregate shocks explain employment growth variability in individual sectors. More specifically, denoting a typical equation for sector i in (2) by

$$\Delta e_{it} = \lambda_i \mathbf{F}_t + u_{it}, \tag{5}$$

where λ_i represents the $1 \times k$ vector of factor loadings specific to sector i and u_{it} denotes sector i's idiosyncratic shocks, we compute

$$R_i^2(\mathbf{F}) = \frac{\lambda_i'\Sigma_{FF}\lambda_i}{\sigma_{e_i}^2}, \tag{6}$$

where $\sigma_{e_i}^2$ is the variance of employment growth in sector i.

Note that the analysis yields an entire distribution of $R_i^2(\mathbf{F})$ statistics, one for each sector. Consider the degenerate case where $R_i^2(\mathbf{F}) = 1$ for each i. In this case, employment variations in each sector are completely driven by the shocks common to all sectors and idiosyncratic shocks play no role. Put another way, variations in aggregate employment reflect only aggregate shocks and the fate of each sector is completely tied to these shocks. A direct economic implication, therefore, is that the issue of market incompleteness or insurance considerations (at the sectoral level) tend to become irrelevant as there is no scope for diversifying away idiosyncratic shocks. To the extent that factor loadings differ across sectors, aggregate shocks still affect sectoral employment differentially so that there may remain some opportunity to complete markets. However, in the limit where $\lambda_i = \lambda_j \; \forall i, j$, the standard

Table 2 Decomposition of Variance from the Approximate Factor Model

	Monthly Growth Rates			Quarterly Growth Rates		
	1 Factor	2 Factors	3 Factors	1 Factor	2 Factors	3 Factors
Std. Dev. of Δe_t Implied by Factor Model	1.80	1.80	1.80	1.53	1.53	1.53
$R^2(\mathbf{F})$	0.77	0.80	0.80	0.94	0.95	0.95

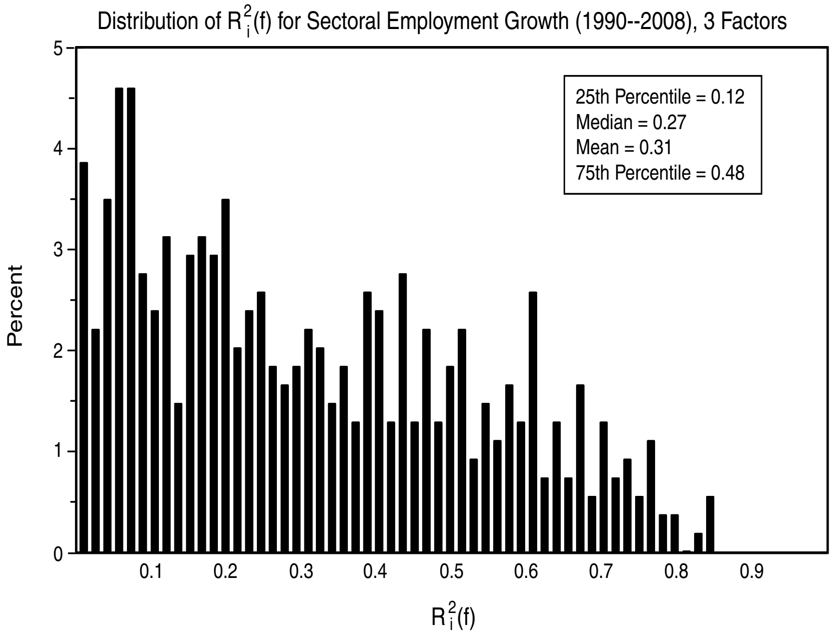
representative agent setup becomes a sufficient framework with which to study business cycles (i.e., without loss of generality). In contrast, when $R_i^2(\mathbf{F}) < 1$ for a subset of sectors, it is no longer true that the fortunes of individual sectors are dictated only by aggregate shocks. Sector-specific shocks help determine sectoral employment outcomes, and the degree of market completeness potentially plays an important part in determining the welfare implications of business cycles.

3. EMPIRICAL FINDINGS

Tables 2 through 4, as well as Figures 4 and 5, summarize the results from computing these key summary statistics using our data on sectoral employment growth rates. We estimated the number of factors using the Bai and Ng (2002) ICP1 and ICP2 estimators, both of which yielded three factors over the full sample period. For robustness, Table 2 shows the factor model's implied standard deviation of aggregate employment (computed using constant shares), as well as the fraction of aggregate employment variability explained by the common factors, $R^2(\mathbf{F})$, using either one, two, or three factors. Most of our discussion will focus on the three-factor model. Two important observations stand out in Table 2. First, the common factors explain essentially all of the variability in quarterly employment growth rates. These common shocks also explain the bulk, or more specifically 80 percent, of fluctuations in monthly growth rates. Second, note that for both monthly and quarterly growth rates, the first factor almost exclusively drives aggregate employment growth, with the second and third factors contributing little additional variability to the aggregate series in relative terms. That is not to say that the absolute variance of the latter factors is small, and we shall see below that these are essential in helping track subsets of the sectors that make up total nonfarm employment.

At a more disaggregated level, Figure 4 illustrates the fraction of quarterly employment growth variability in individual sectors that is attributable to common shocks or, alternatively, the distribution of $R_i^2(\mathbf{F})$. As the figure makes

Figure 4 Contribution of Sector-Specific Shocks to Sectoral Employment



clear, sector-specific shocks play a key role in accounting for employment variations at the sectoral level, with common shocks explaining, on average, only 31 percent of the variability in sectoral employment. In addition, observe that there exists substantial heterogeneity in the way that employment is driven by aggregate and idiosyncratic shocks across sectors. Specifically, the interquartile range suggests $R_i^2(\mathbf{F})$ statistics that are between 0.12 to 0.48, or a 0.36 point gap.

It may seem counterintuitive at first that $R^2(\mathbf{F})$ is close to 1 in Table 2 while the mean or median $R_i^2(\mathbf{F})$ statistic is considerably less than 1 in Figure 4. To see the intuition underlying this result, consider equation (2) when aggregated across sectors:

$$\mathbf{s}'\Delta\mathbf{e}_t = \mathbf{s}'\boldsymbol{\lambda}\mathbf{F}_t + \mathbf{s}'\mathbf{u}_t. \tag{7}$$

When the number of sectors under consideration is large, as in this article, the "uniquenesses" will tend to average out by the law of large numbers. Put another way, since the u_{it}s are weakly correlated across sectors and have mean zero, $\mathbf{s}'\mathbf{u}_t = \sum_{i=1}^{N} s_i u_{it} \rightarrow^p 0$ as N becomes large. This result

Table 3 Fraction of Variability in Sectoral Employment Growth Explained by Common Shocks

Sector	$R_i^2(\mathbf{F})$
Residential Building Construction	0.85
Electrical Equipment Manufacturing	0.85
Wood Kitchen Cabinet and Countertop	0.84
Plumbing and HVAC Contractors	0.84
Printing and Related Support Activities	0.80
Other Building Material Dealers	0.80
Wireless Telecommunications Carriers	0.78
Construction Equipment	0.78
Plywood and Engineered Wood Products	0.77
Semiconductors and Electronic Components	0.77
Management of Companies and Enterprises	0.77
Electrical Contractors	0.77
Lumber and Wood	0.77
Metalworking Machinery Manufacturing	0.76
Electric Appliance and Other Electronic Parts	0.76

holds provided that the distribution of sectoral shares is not too skewed so that a few sectors have very large weights (see Gabaix 2005). In contrast, $\mathbf{s}'\boldsymbol{\lambda}\mathbf{F}_t = \mathbf{F}_t \sum_{i=1}^{N} s_i \lambda_i$ does not necessarily go to zero with N since the λ_is are fixed parameters.[3] Therefore, whatever the importance of idiosyncratic shocks in driving individual sectors (i.e., whatever the distribution of $R_i^2(\mathbf{F})$), $R^2(\mathbf{F})$ will generally tend towards 1 in large panels. The rate at which $R^2(\mathbf{F})$ approaches 1 will depend on the particulars of the data-generating process. In this case, with 544 sectors, we find that $R^2(\mathbf{F})$ is around 0.8 in monthly data and 0.95 in quarterly data.

Interestingly, Figure 4 suggests that at the high end of the cross-sector distribution of $R_i^2(\mathbf{F})$ statistics, there exist individual sectors whose variation in employment growth is almost entirely driven by the common shocks that explain aggregate employment, and, thus, that are particularly vulnerable to the business cycle. Table 3 lists the top 15 sectors in which idiosyncratic shocks play the least role in relative terms. Note that all of the sectors listed in Table 3 are goods-producing sectors. In other words, even though service-providing sectors employ more than four times as many workers as the goods-producing sectors, it turns out that it is the latter sectors that are most informative about the state of aggregate employment. In essence, because employment variations in the sectors listed in Table 3 reflect mainly the effects of common shocks, and because movements in aggregate employment growth are associated with

[3] In Foerster, Sarte, and Watson (2008), the factor loadings correspond to reduced-form parameters that can be explicitly tied to the structural parameters of a canonical multi-sector growth model.

Table 4 Sectoral Information Content of Aggregate Employment

Selected Sectors Ranked by $R_i^2(\mathbf{F})$	Fraction of $\Delta\mathbf{e}_t$ Explained by Selected Sectors
Top 5 Sectors	0.88
Top 10 Sectors	0.92
Top 20 Sectors	0.94
Top 30 Sectors	0.96

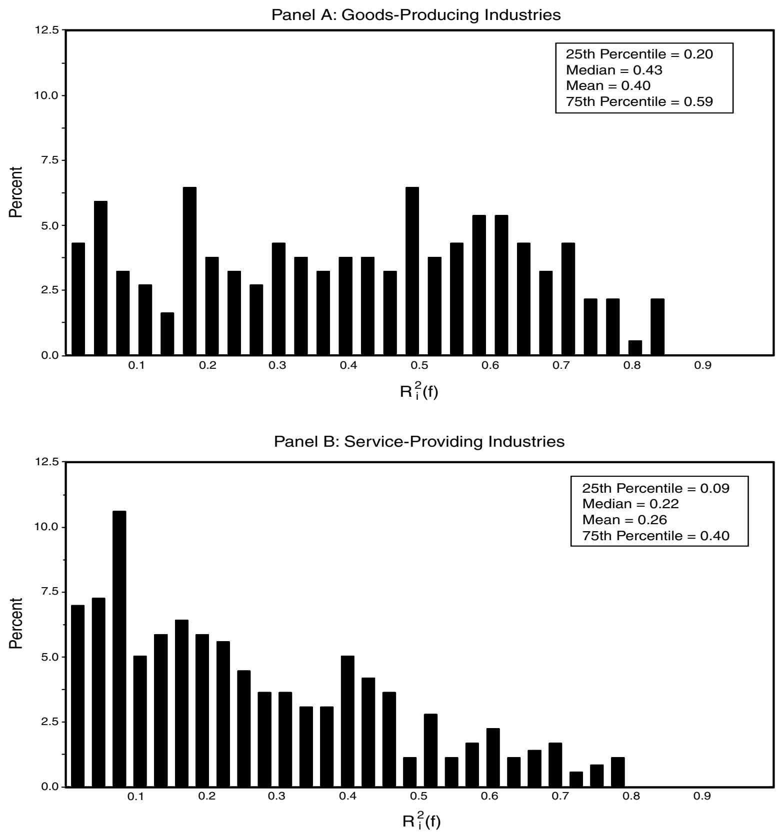
these shocks (Table 2), information regarding aggregate employment tends to be concentrated in these sectors.

This notion of sectoral concentration of information regarding aggregate employment can be formalized further as follows. Consider the problem of tracking movements in aggregate employment using only a subset, M, of the available sectors, say the the five highest ranked sectors in Table 3. This problem pertains, for example, to the design of surveys that are meant to track aggregate employment in real time such as those carried out by the Institute for Supply Management, as well as by various Federal Reserve Banks including the Federal Reserve Bank of Richmond.[4] In particular, the question is: Which sectors are the most informative about the state of aggregate employment and should be included in the surveys? To make some headway toward answering this question, let $\widetilde{\Delta\mathbf{e}}_t$ denote the vector of employment growth rates associated with the M sectors such that $\widetilde{\Delta\mathbf{e}}_t = \mathbf{m}\Delta\mathbf{e}_t$, where $\mathbf{m}$ is an $M \times N$ selection matrix. To help track aggregate employment growth, $\mathbf{s}'\Delta\mathbf{e}_t$, we compute the $M \times 1$ vector of weights, $\mathbf{w}$, attached to the different employment growth series in $\widetilde{\Delta\mathbf{e}}_t$ as the orthogonal projection of $\mathbf{s}'\Delta\mathbf{e}_t$ on $\widetilde{\Delta\mathbf{e}}_t$. That is to say, the weights are optimal in the sense of solving a standard least-square problem, $\mathbf{w} = (\mathbf{m}\Sigma_{ee}\mathbf{m}')^{-1}\mathbf{m}\Sigma_{ee}\mathbf{s}$.

Table 4 reports the fraction of aggregate employment growth explained by the (optimally weighted) employment series related to various sector selections in our data set, $\mathbf{w}'\widetilde{\Delta\mathbf{e}}_t$. Strikingly, using only the sectors associated with the highest five $R_i^2(\mathbf{F})$ statistics in Table 3, this particular filtering already helps us explain 88 percent of the variability in aggregate employment growth. Moreover, virtually all of the variability in aggregate employment growth is accounted for by only considering the 30 highest ranked sectors, according to $R_i^2(\mathbf{F})$, out of 544 sectors. It is apparent, therefore, that information concerning movements in aggregate employment growth is concentrated in a small number of sectors. Contrary to conventional wisdom, these sectors are not necessarily those that have the largest weights in aggregate employment nor the most volatile employment growth series. Because aggregate employment growth

[4] Employment numbers are typically released with a one-month lag and revised up to three months after their initial release. In addition, a revision is carried out annually in March.

Figure 5 Distribution of $R_i^2(F)$ in Goods-Producing and Service-Providing Sectors



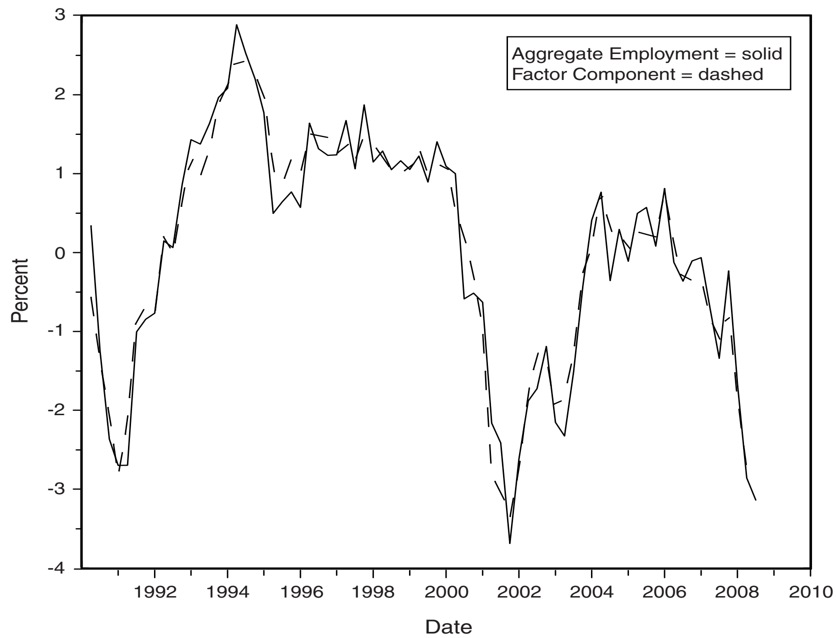
is almost exclusively driven by common shocks, the factor analysis proves useful precisely because it allows us to identify the individual sectors whose employment growth also moves most closely with these shocks.

From the exercise we have just carried out, it should be clear that there is much heterogeneity in the way that individual sector employment growth compares to aggregate employment growth over the business cycle. To underscore this point, Figure 5 depicts the breakdown of $R_i^2(\mathbf{F})$ statistics across the main sectors that make up total goods and services separately. Differentiating between goods-producing and service-providing industries, Figure 5 shows that aggregate shocks play a lesser role in driving employment variations in the service sectors relative to the goods-producing sectors. In particular, both the mean and median $R_i^2(\mathbf{F})$ statistics are notably lower in the service-providing industries than in the goods sectors. That said, it is also the case that there isn't much uniformity within the goods-producing sectors. In particular, we find that employment variations in the durable goods sectors are significantly more subject to common shocks than in the nondurable goods sectors. The median $R_i^2(\mathbf{F})$ statistic is 0.54 in durable goods but only 0.20 in the nondurable goods sectors. In service-providing industries, we find that sector-specific shocks generally play a much greater role in determining employment growth variations. Moreover, the distributions of $R_i^2(\mathbf{F})$ tend to be more similar across service sectors than they are across goods-producing industries. The smallest median $R_i^2(\mathbf{F})$ value across private industries is 0.19, in financial activities, while the largest value is relatively close at 0.29, in the information sector. As indicated above, although employment variations in individual sectors tend to be dominated by sector-specific shocks, these shocks tend to lose their importance in aggregation. To further illustrate this notion, let $\mathbf{s}_j$ denote a vector comprising either the shares corresponding to a particular subsector j of total goods and services, say goods-producing sectors, or zero otherwise. In other words, $\mathbf{s}_j$ effectively selects out employment growth in the different industries making up subsector j. It follows that employment growth in that subsector is given by $\mathbf{s}_j' \Delta \mathbf{e}_t$, and the corresponding factor component in that subsector is $\mathbf{s}_j' \boldsymbol{\lambda} \mathbf{F}_t$. Note that to the degree $\mathbf{s}' \boldsymbol{\lambda} \mathbf{F}_t$ successfully captures the business cycle as it relates to movements in aggregate employment, $\mathbf{s}_j' \boldsymbol{\lambda} \mathbf{F}_t$ captures the analogous concept at a more disaggregated level.

Figures 6 and 7 depict the behavior of $\mathbf{s}_j' \Delta \mathbf{e}_t$ and $\mathbf{s}_j' \boldsymbol{\lambda} \mathbf{F}_t$ for the various sectoral components of our data. Despite the heterogeneity in sectoral employment across sectors as captured by $R_i^2(\mathbf{F})$, the figures suggest that employment growth generally follows movements in the factor component not only at the aggregate level but in subsectors of the economy as well. Of course, at the aggregate level, we have argued that this is to be expected given the results in Table 2 and confirmed in Figure 6. However, we also find that employment growth and the factor component generally move together in goods-producing and service-providing industries separately (Figure 7). In fact, this finding is also true of the main subsectors that make up total goods and services, with the notable exception of government. Perhaps not surprisingly, the latter finding simply reflects the lack of a business cycle component in government services relative to other sectors. Consistent with our earlier findings, our work additionally suggests that employment growth moves less closely with the factor component in service-providing industries than in goods-producing sectors, notably in financial services for instance. On the whole, however, the factor analysis appears to provide a helpful way to track the business cycle as it relates to employment in the broad sectoral components of goods and services.

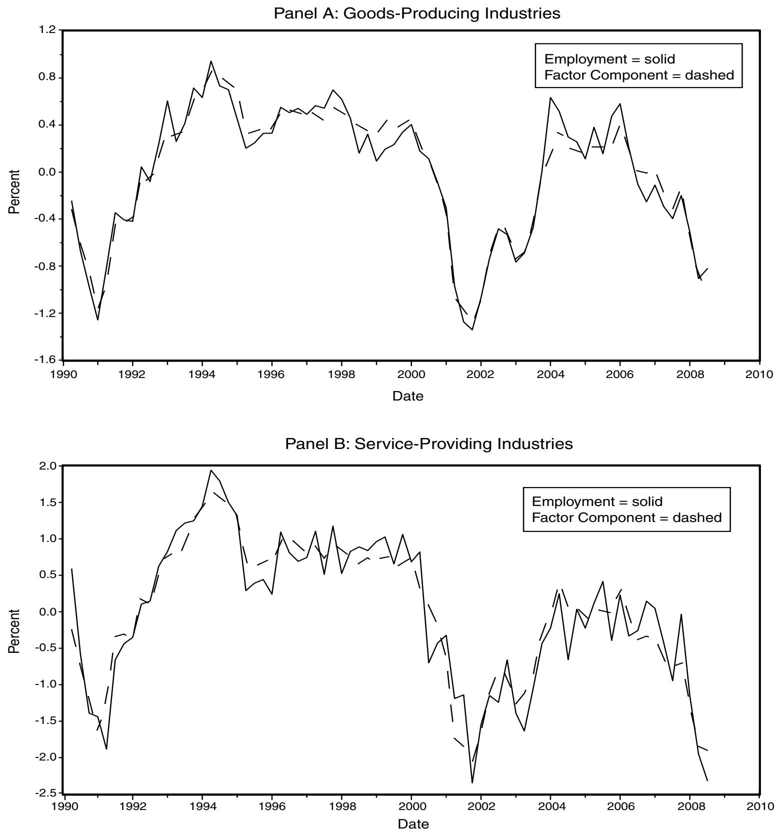
Finally, we note that the factors uncovered in this analysis play substantially different roles in explaining aggregate and sectoral variations in employment. Specifically, even though the first factor alone explains roughly 94

Figure 6 Aggregate Employment Growth and Factor Component



percent of the variation in aggregate employment growth (Table 2), this factor does very little to explain employment growth in particular sectoral components of goods and services. To see this, Figure 8 shows plots of employment growth in natural resources and mining, as well as education and health services, against the factor component using one, two, and three factors. In the first row of Figure 8, we see unambiguously that, despite accounting for the bulk of the variations in aggregate employment, the first factor does very little to capture employment variations in either of the sectors. The correlation between the factor component and employment growth is virtually nil at 0.03 in natural resources and mining and 0.08 in education and health services. In sharp contrast, this correlation jumps to 0.57 in education and health once the second factor is included, and to 0.77 in natural resources and mining once the third factor is included. Note, in particular, that the second factor does little to capture employment growth in natural resources and mining, and it is the third factor alone that helps capture business cycle movements in employment in that sector. In that sense, the Bai and Ng (2002) ICP1 and ICP2 estimators help identify factors that not only explain aggregate employment variations but also account for employment movements at a more disaggregated level.

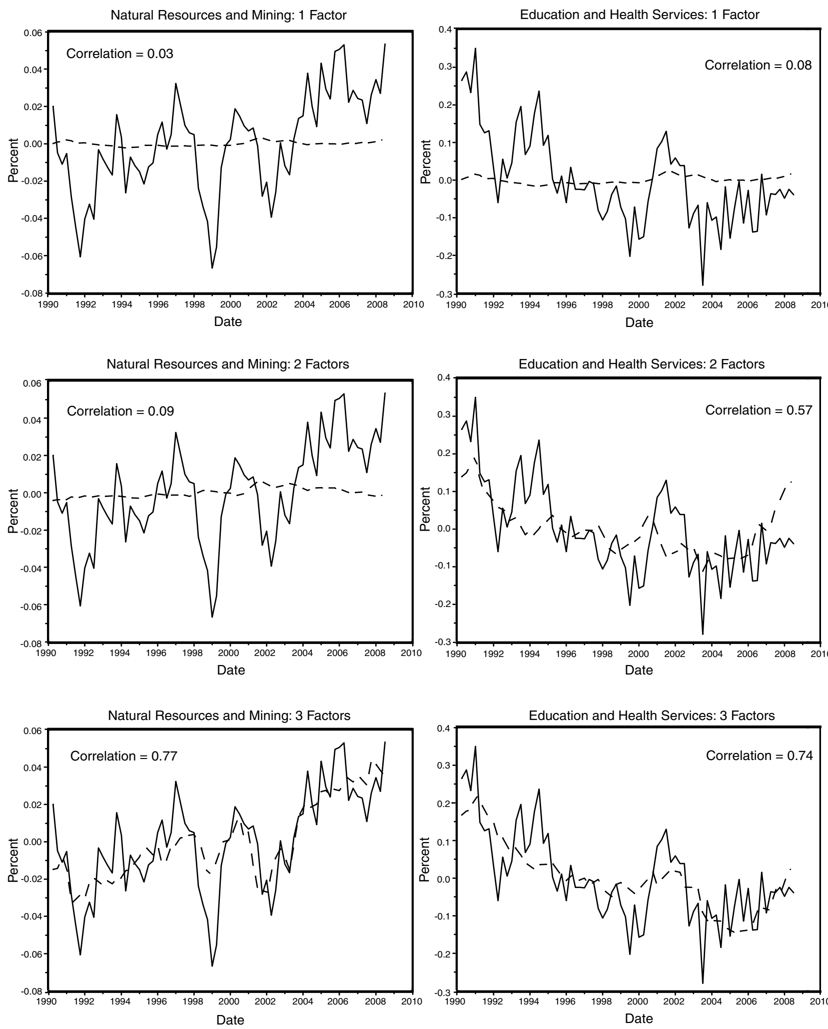
Figure 7 Employment Growth and Factor Component in Goods and Services



4. CONCLUSIONS

In the standard neoclassical one-sector growth model, fluctuations in the representative agent's circumstances are largely determined by shocks to aggregate total factor productivity. This notion is developed, for example, in work going as far back as King, Plosser, and Rebelo (1988). The assumption of a representative agent stands in for a potentially more complicated world populated by heterogenous agents, but where homothetic preferences and complete markets justify focusing on the average agent. Alternatively, we can also think of the representative agent framework as approximating a world in which all agents are essentially identical and affected in the same way by shocks to the economic environment. Under the latter interpretation, a boom in the course of a business cycle characterizes a situation in which "a rising tide lifts all

Figure 8 Contribution of Individual Factors in Explaining Sectoral Employment Growth



boats," and vice versa in the case of a recession. Put another way, idiosyncratic shocks play no role in determining agents' outcomes. More importantly, when individual agents' fortunes are driven mainly by common shocks, the significance of market incompleteness and the importance of insurance considerations tend to vanish since there is no scope for diversifying idiosyncratic shocks away.

Using factor analytic methods, this article documents instead significant differences in employment variations across sectors. In some industries, notably in goods production, variations in employment growth are dominated by aggregate shocks so that these sectors are particularly sensitive to the business cycle. In other industries, in particular some service-providing industries, employment movements are virtually unrelated to aggregate shocks and instead result almost exclusively from sector-specific shocks. The analysis, therefore, suggests that agents working in different sectors of the U.S. economy are affected in very different ways by shocks to the economic environment. Moreover, it underscores the potential importance of market incompleteness and mitigates the usefulness of representative agent models in determining the welfare costs of business cycles.

APPENDIX

This Appendix gives a brief description of the Principle Component (PC) problem based on the discussion in Johnston (1984). See that reference for a more detailed presentation of the problem and its implications.

As described in the main text, suppose we have (demeaned) employment growth observations across N sectors over T time periods summarized in an $N \times T$ matrix, X. In that way, $\Delta \mathbf{e}_t$ in the text is a typical column of X. The nature of the PC problem is to capture the degree of co-movement across these N sectors in a simple and convenient way. To this end, the PC problem transforms the Xs into a new set of variables that will be pairwise uncorrelated and of which the first will have maximum possible variance, the second the maximum possible variance among those uncorrelated with the first, and so on.

Let

$$F_1' = X'\lambda_1$$

denote the first such variable where λ_1 and F_1' are $N \times 1$ and $T \times 1$ vectors, respectively. In other words, F_1' is a linear combination of the elements of X across sectors. The sum of squares of F_1 is

$$F_1 F_1' = \lambda_1' \Sigma_{XX} \lambda_1, \tag{8}$$

where $\Sigma_{XX} = XX'$ represents the variance-covariance matrix (when divided by T) of employment growth rates across sectors. We wish to choose the weights λ_1 to maximize $F_1 F_1'$, but some constraint must evidently be imposed on λ_1 to prevent the sum of squares from being made infinitely large. Thus, a convenient normalization is to set

$$\lambda_1'\lambda_1 = 1.$$

The PC problem may now be stated as

$$\max_{\lambda_1} \lambda_1'\Sigma_{XX}\lambda_1 + \mu_1(1 - \lambda_1'\lambda_1),$$

where μ_1 is a Lagrange multiplier. Using the fact that Σ_{XX} is a symmetric matrix, the first-order condition associated with this problem is

$$2\Sigma_{xx}\lambda_1 - 2\mu_1\lambda_1 = 0.$$

Thus, it follows that

$$\Sigma_{xx}\lambda_1 = \mu_1\lambda_1.$$

In other words, the weights λ_1 are given by an eigenvector of Σ_{xx} with corresponding eigenvalue μ_1. Observe that when λ_1 is chosen in this way, the sum of squares in (8) reduces to

$$\lambda_1'\Sigma_{XX}\lambda_1 = \lambda_1'\mu_1\lambda_1 = \mu_1.$$

Therefore, our choice of λ_1 must be the eigenvector associated with the largest eigenvalue of Σ_{XX}. The first principle component of X is then F_1.

Now, let us define the next principle component of X as $F_2' = X'\lambda_2$. Similar to the choice of λ_1 we have just described, the problem is to choose the weights λ_2 so as to maximize $\lambda_2'\Sigma_{XX}\lambda_2$ subject to $\lambda_2'\lambda_2 = 1$. In addition, however, because we want the second principle component to capture comovement that is not already reflected in the first principle component, we impose the further restriction $\lambda_2'\lambda_1 = 0$. This last restriction ensures that F_2 will be uncorrelated with F_1.

The problem associated with the second principle component may then be stated as

$$\max_{\lambda_2} \lambda_2'\Sigma_{XX}\lambda_2 + \mu_2(1 - \lambda_2'\lambda_2) + \phi\lambda_2'\lambda_1.$$

The corresponding first-order condition is

$$2\Sigma_{xx}\lambda_2 - 2\mu_2\lambda_2 + \phi\lambda_1 = 0.$$

Pre-multiplying this last equation by λ_1' gives

$$2\lambda_1'\Sigma_{XX}\lambda_2 - 2\mu_2\lambda_1'\lambda_2 + \phi\lambda_1'\lambda_1 = 0,$$

or

$$\phi = 0,$$

since $\lambda_1'\lambda_1 = 1$, $\lambda_1'\Sigma_{XX} = \mu_1\lambda_1'$, and $\lambda_1'\lambda_2 = 0$. Therefore, we have that the weights λ_2 must satisfy

$$\Sigma_{xx}\lambda_2 = \mu_2\lambda_2,$$

and, in particular, should be chosen as the eigenvector associated with the second largest eigenvalue of Σ_{XX}.

Proceeding in this way, suppose we find the first k principle components of X. We can arrange the weights $\lambda_1, \lambda_2,...,\lambda_k$ in the $N \times k$ orthogonal matrix

$$\Lambda_k = [\lambda_1, \lambda_2, ..., \lambda_k].$$

Furthermore, the general PC problem may then be described as finding the $T \times k$ matrix of components, $F' = X'\Lambda_k$, such that Λ_k solves

$$\max_{\Lambda_k} \ \Lambda_k' \Sigma_{XX} \Lambda_k \text{ subject to } \Lambda_k' \Lambda_k = I_k. \tag{9}$$

Now, consider the approximate factor model (2) in the text written in matrix form,

$$X = \Lambda_k F + u,$$

where X is $N \times T$, Λ_k is a $N \times k$ matrix of factor loadings, F is a $k \times T$ matrix of latent factors, and u is $N \times T$. One can then show that solving the constrained least-square problem,

$$\min_{\{F_1\}_{t=1}^T,...,\{F_k\}_{t=1}^T, \Lambda_k} \sum_{t=1}^{T} (X_t - \Lambda_k F_t)'(X_t - \Lambda_k F_t) \text{ subject to } \Lambda_k' \Lambda_k = I_k,$$

is equivalent to solving the general principle component problem (9) we have just described (see Stock and Watson 2002).

REFERENCES

Bai, Jushan, and Serena Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70 (January): 191–221.

Foerster, Andrew, Pierre-Daniel Sarte, and Mark W. Watson. 2008. "Sectoral vs. Aggregate Shocks: A Structural Factor Analysis of Industrial Production," Working Paper 14389. Cambridge, Mass.: National Bureau of Research (October).

Forni, Mario, and Lucrezia Reichlin. 1998. "Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics." *Review of Economic Studies* 65 (July): 453–73.

Gabaix, Xavier. 2005. "The Granular Origins of Aggregate Fluctuations." Manuscript, Massachusetts Institute of Technology.

Johnston, Jack. 1984. *Econometric Methods*, third edition. New York: McGraw-Hill Book Company.

King, Robert, Charles Plosser, and Sergio Rebelo. 1988. "Production, Growth and Business Cycles: The Basic Neoclassical Model." *Journal of Monetary Economics* 21: 195–232.

Quah, Danny, and Thomas J. Sargent. 1993. "A Dynamic Index Model for Large Cross Sections." In *Business Cycles, Indicators and Forecasting*, edited by J. H. Stock and M. W. Watson. Chicago: University of Chicago Press for the NBER, 285–310.

Shea, John. 2002. "Complementarities and Comovements." *Journal of Money, Credit, and Banking* 34 (May): 412–34.

Stock, J. H., and M. W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97 (December): 1,167–79.