Dominant Sectors in the US: A Factor Model Analysis of Sectoral Industrial Production

Soroosh Soofi Siavash*

May 13, 2016

Abstract

This paper considers the role of sector-specific production growth shocks as major sources of cross-sector comovement in the business cycle. The dominant sectors are identified according to the number of factors detected in the residuals of a factor model after treating the production growth in some candidate sectors as observed factors. I build on the properties of the principal component method and analytically show this approach can consistently identify the dominant units (sectors) in a large panel data set even if the assumption of strongly influential factors is moderately relaxed. Empirical evidence is provided that growth in a few industrial sectors in the US provide suitable approximations for an unknown common factor. Using data on the intersectoral material input-outputs and the cross-sector capital flows, I find that the dominant sectors have an important role as suppliers of capital products to other sectors.

*Goethe University Frankfurt, ssoofi@hof.uni-frankfurt.de. I am grateful to Mehdi Hosseinkouchack, Pooyan Amir-Ahmadi, Sebastian Kripfganz, Sulkhan Chavleishvili, and the seminar participants at Goethe University Frankfurt for helpful comments and discussions. I would like to acknowledge travel grant from the IAAE to present this paper at the 3rd IAAE conference in Milan.

1 Introduction

Comovement across sectors of an economy is a striking feature of business cycles. This feature appears in large pairwise correlation between the sectoral growth rates in terms of output, value added, and employment. Shea (2002) finds an average correlation of 0.28 and 0.23 using annual growth rate of gross output and value added for 126 industries in the US over the period 1960 to 1986. Even a larger correlation (0.34) is obtained using employment data over 1959 to 1986 according to the same sectoral classification. Furthermore, when it comes to the sources of variability in the growth rates of the macroeconomic aggregates, sectoral covariability appears to have an important role in explaining this aggregate variability. Shea decomposes variance of the aggregate gross output growth in terms of the contribution of diagonal and off-diagonal elements of the sectoral covariance matrix and shows that the off-diagonal elements account for 85% of the aggregate variance. Foerster, Sarte, and Watson (2011) find an average pairwise correlation of 0.19 over 1972 to 2007, and a contribution of 89% for the sectoral covariances using quarterly industrial production data for 117 US sectors. Long and Plosser (1987) and Carvalho (2014), among others, report similar results.

Two main approaches are considered in literature to analyze the sources of intersectoral comovement. One approach relies on factor model analysis, where it is assumed that a few common shocks account for a large portion of comovement across sectors. Comparison of aggregate versus sectoral shocks have been provided by Long and Plosser (1987), Forni and Reichlin (1998), Shea (2002), and Foerster et al. (2011), among others. A common finding is that the idiosyncratic shocks to sectors, next to the aggregate shocks, have important roles in explaining overall movements, where it is shown that around 50% of the overall variability in the US can be attributed to the sector-specific shocks in some periods. Another strand of literature considers propagation of shocks due to intersectoral complementarities (e.g. Long and Plosser (1983), Horvath (1998), and Dupor (1999)). These studies create multisector general equilibrium models, where production technology of a sector (or a firm) is linked to other sectors through input-output interactions. The complementarity creates intersectoral linkages transmitting growth disturbances to downstream and upstream sectors. Sequential shocks propagation might lead to widespread comovement, and in turn to substantial aggregate variability. The common findings among these studies are, in the first place, the input-output linkages are central to propagation of technological progress, and furthermore, the extent of shock propagation depends on the structure of cross-sector linkages. The sectors which have a larger number of trading partners have more important roles in spreading aggregate and idiosyncratic disturbances throughout the model. These studies consider the sectoral versus aggregate shocks with the input-output linkages acting as propagation channels. In this paper, I look at the shocks to individuals sectors as if they potentially act like common macroeconomic shocks. My goal is to identify the "dominant" sectors using factor model analysis and the principal component method, where a sector is defined as a dominant one if its production growth provides a suitable approximation for a common factor.

From the literature on the principal component method, we know that number of factors and the space spanned by them can be consistently estimated when the number of cross-section units (N) and time series observations (T) in a panel data set jointly go to infinity ($N, T \to \infty$).[1] In this paper, the presence of a few unknown common factors is viewed as a starting point for identification of the dominant sectors. Then, the approach from Parker and Sul (2016) (henceforth, PS) is used for the purpose of identifying. In particular, production growth of a candidate sector is treated as an observed common factor. Including the candidate variables into the factor model, we can identify the dominant sectors according to number of factors estimated in the residuals. If a candidate sector is indeed dominant, the production growth in this sector covers the space of an underlying factor, which leads to a reduction in the number of factors in the residuals. In Section 2, I provide the definition of a dominant sector and review the PS approach.

The dominance of a unit (sector) as defined by PS relates to the concept of unknown common factors or diffuse indexes which are usually meant to explain statistical properties of panel data sets. As opposed to a statistical factor model, a model of production network building on a multisector general equilibrium model, incorporates the observed input-output linkages across sectors and uses the network techniques to discuss the distinguishing properties of the sectors or firms. Section 3 builds on this literature and provides a link between dominance of a sector and its network centrality. Starting with the solution of a multisector general equilibrium model (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), and Holly and Petrella (2012)), I obtain a factor model representation as a reduced form of the model which features the production growth in the most

[1] The properties of the principal component method are discussed in detail by Stock and Watson (1998), Bai and Ng (2002), and Bai (2003), among others.

central sectors acting as common factors.

In factor analysis and principal component estimation, it is usually assumed that the factors are common and affect all the cross-sectional units in the data set. This feature (which in our context means that the production growth of the dominant sectors affect all other sectors) together with the assumption that there is limited dependency left among the idiosyncratic errors (i.e. there is a star-like structure with the dominant sectors being neighbors to all others and the rest which all are nondominant have limited dependencies among themselves) ensure that the space of factors can be consistently estimated. In Section 4, I deviate from such a common factor structure and analytically assess consistency of the PS approach under different degrees of factors influence. The cases are considered in which factors are potentially influential on parts of the system. The contribution of this section is motivated by the fact that the principal component method becomes less reliable or even inconsistent as dominance of factors relative to the idiosyncratic error reduces (Boivin and Ng (2006), and Onatski (2012)). To weaken the assumption of commonness of factors, I closely follow Kapetanios and Marcellino (2010) who use an analytically appealing way to incorporate less-than-strong factors to track the implication of different degrees of factors influence on the principal component method.[2] The analysis proceeds in two steps. In the first step, it is shown that the principal component method remains consistent for each time t under a moderate deviation from the case of common factors. Under this setting, the second step provides the conditions required to consistently identify the dominant units (sectors) using the PS approach. A Monte Carlo study is presented in Section 5 to assess small sample properties of the identification approach following the analytical analysis.

Section 6 presents the empirical work, where the dominant sectors in the US are identified using the data set from Foerster et al. (2011). The data set contains disaggregated data on sectoral industrial production in the period covering 1972 to 2007. Two common shocks are detected among production growth rates. The findings are as follows. I provide evidence that one of these common shocks is attributable to the shocks arising in a few sectors, including the heavy machinery industries, while the other one appears external to the model. This data set excludes some service sectors like financial industry. The latter shock could be attributable to the shocks arising in these sectors or another aggregate shock, such as monetary policy affecting the demand for durable goods. In addition, I

[2]See Chudik, Pesaran, and Tosetti (2011) for an elaboration on the strong and weak factors.

find that the dominant sectors mostly have important role as suppliers of capital products to others. Adopting a network perspective, this result is implied from centrality analysis of two tables of input-output linkages; intersectoral material purchases and cross-sector capital flows. Furthermore, I show the growth rate of the sectors whose higher centrality is implied from the table of capital flows tend to have more significant relationships with the other sectoral growth rates in comparison to those implied from the table of intersectoral material input-outputs.

This work builds on the concept of micro shocks acting as sources of macroeconomic variability which is originally established by Jovanovic (1987) and Durlauf (1993). See Carvalho (2014) for a thorough elaboration on the concept. Acemoglu et al. (2012) argue that micro shocks would have nontrivial contributions if they come from the firms having considerably more important roles as suppliers of inputs to others, where the authors provide empirical evidence for the presence of such substantial asymmetries in favor of their argument.[3] In this paper, I use insights obtained from the multisector general equilibrium models of production networks regarding the extreme asymmetries among the sectors of an economy, and combine these insights with factor analysis for the purpose of identifying the dominant sectors. The identification approach used in this paper is less restrictive relative to such structural models (i.e. it relies on the assumptions of a static factor model and does not impose the identifying restrictions implied from the input-output tables). Thus, it would be helpful to use this approach first to detect the dominant sectors based on the statistical properties of the sectoral data. Using these results, one can next assess whether there are common features among the dominant sectors in terms of their importance as suppliers of inputs to others or their relative size in the model.

Furthermore, this paper also uses insights from Chudik and Pesaran (2013) who discuss the role of the variables of the dominant units (the sectoral production growth rates) as common factors in the context of large dimensional VAR models. The discussion provided here elaborates on an approach to identification of such units (sectors) focusing on a static factor model representation.

Next in the paper, Section 7 provides a summary and concluding remarks. The technical appendix includes the proofs of the theorems given in Section 4.

[3] Gabaix (2011) focuses on size of the firms in an economy and develops the hypothesis that idiosyncratic shocks to large firms would have nontrivial contributions to aggregate fluctuations. The author provides empirical evidence that a weighted average of idiosyncratic shocks to the top 100 US firms explain around one-third of fluctuations in the GDP growth.

A brief note on notation. I use 'hat' to present the least squares estimates and 'tilde' to present the principal component estimates. K denotes a generic finite number which is independent of N and T, and $C_{NT} = min\{\sqrt{N}, \sqrt{T}\}$ is an important divergence rate as $N, T \to \infty$. $||A|| = \sqrt{tr(A'A)}$ is the Frobenius norm of the matrix A. Let $\rho_1(A) > \rho_2(A) > \cdots > \rho_r(A)$ denote absolute value of the first r eigenvalues of A in descending order with $\rho_1(A)$ being the spectral radius, and $||A||_2 = \sqrt{\rho_1(A'A)}$ denotes the spectral norm of A. $a_n = O(b_n)$ states that the sequence $\{a_n\}$ is at most of order b_n, and $x_n = O_p(y_n)$ states that the random variable x_n is at most of order y_n in probability. Convergence in probability is denoted by $\xrightarrow{p}$.

2 Factor Model and Review of the PS Approach

To elaborate on the PS approach to identification of dominant sectors, I begin with defining a static approximate factor model for sectoral growth rates. Let x_{it} denote the ith sector's production grow rate, for $i = 1, ..., N$ and $t = 1, ..., T$. I assume the data generating process is a r-factor model

$$x_t = \Lambda f_t + e_t, \tag{1}$$

where $x_t = (x_{1t}, ..., x_{Nt})'$, f_t is a $r \times 1$ vector of unknown factors, Λ is the $N \times r$ matrix of factor loadings, and e_t contains the idiosyncratic errors. The ith row of x_t, Λf_t, and e_t give the process for the ith sector's production growth rate as $x_{it} = \lambda_i' f_t + e_{it}$ with $\lambda_i' f_t$ being the common component of sector i. The model takes the matrix form

$$X = F\Lambda' + e, \tag{2}$$

where $X = (x_1, ..., x_T)'$ is the $T \times N$ matrix of all observations in the panel data set, $F = (f_1, ..., f_T)'$, and $e = (e_1, ..., e_T)'$. Here, f_t is unknown and potentially attributable to growth disturbances of some sectors. The goal of this paper is to identify these sectors.

The model assumptions, which are similar to those made by Bai (2003) are as follows.

Assumption FM1 (Factors): *$E||f_t||^4 \leq K < \infty$, and $F'F/T \xrightarrow{p} \Sigma_f$ where Σ_f is a full rank matrix.*

Assumption FM2 (Loadings): *λ_i is deterministic such that $||\lambda_i|| \leq K < \infty$ for all i, and $\Lambda'\Lambda/N \to \Sigma_\Lambda$ where Σ_Λ is a full rank matrix.*

Assumption FM3 (Weakly dependent idiosyncratic errors): *For all N and T;* (i) *$E(e_{it}) = 0$ and $E|e_{it}|^8 \leq K < \infty$.* (ii) *Let $E(e_{it}e_{js}) = \tau_{ij,ts}$, $|\tau_{ij,ts}| \leq \tau_{ij}$ for all (t,s), and $|\tau_{ij,ts}| \leq \tau_{ts}$ for all (i,j); $\frac{1}{N}\sum_{i,j=1}^{N}\tau_{ij} \leq K$, $\frac{1}{T}\sum_{t,s=1}^{T}\tau_{ts} \leq K$, and $\frac{1}{NT}\sum_{i,j,t,s=1}|\tau_{ij,ts}| \leq K$.* (iii) *$E|N^{-1/2}\sum_{i=1}^{N}(e_{is}e_{it} - E(e_{is}e_{it}))|^4 \leq K$ for all (t,s).*

Assumption FM4 (Weak dependence between factors and idiosyncratic errors, and moment conditions): *(i) $E\left(N^{-1}\sum_{i=1}^{N}||T^{-1/2}\sum_{t=1}^{T}f_t e_{it}||^2\right) \leq K$.*
(ii) *$E||(NT)^{-1/2}\sum_{s=1}^{T}\sum_{i=1}^{N}f_s[e_{is}e_{it} - E(e_{is}e_{it})]||^2 \leq K$ for all t.* (iii) *$E||(NT)^{-1/2} \times \sum_{t=1}^{T}\sum_{i=1}^{N}f_t\lambda_i' e_{it}||^2 \leq K$.*

According to Assumptions FM1-2, there are r common factors. Assumption FM2 ensures that Λ is full column rank and the number of nonzero elements in each of its columns is proportional to N as indicated by convergence of $\Lambda'\Lambda/N$ to a full rank matrix. Assumptions FM3 and FM4 allow for limited dependence in e_{it} across i and t, as well as limited dependence between the factors and the idiosyncratic errors, respectively. These assumptions let the model deviate from a classical factor model which assumes f_t and e_{it} are i.i.d. to the extent that some moment conditions, which hold under a classical factor model, still hold. More precisely, the model (2) has the sample covariance (divided additionally by N)

$$\frac{X'X}{NT} = \frac{\Lambda F'F\Lambda'}{NT} + \frac{\Lambda F'e + e'F\Lambda}{NT} + \frac{e'e}{NT}.$$

Under Assumptions FM1-2, $\Lambda F'F\Lambda'/NT$ has always r nonzero eigenvalues. In contrast, the eigenvalues of $(\Lambda F'e + e'F\Lambda)/NT$ and $e'e/NT$ diminish at a suitable rate as $N, T \to \infty$. These results hold similar to a classical factor model despite allowing for weak dependence in e. Thus, the contribution of the common component in variance of x_t substantially dominates that of idiosyncratic errors when N and T are large. These features let us estimate F and r using simple approaches based on the principal component method. I use here these approaches.

In particular, the principal component method is used to estimate F and Λ which minimizes the average sum of squared errors

$$V(r) = \min_{F,\Lambda} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it} - \lambda_i' f_t \right)^2, \quad (3)$$

subjected to the normalization $F'F/T = I_r$. The solution of (3) yields an estimate for F, $\tilde{F}$, which is $\sqrt{T}$ times the first r eigenvectors of the $T \times T$ matrix XX'/NT, as well as an estimate for Λ which is $\tilde{\Lambda} = X'\tilde{F}(\tilde{F}'\tilde{F})^{-1} = X'\tilde{F}/T$. Furthermore, if we treat $V(k)$ as a function of k for $k = 0, \ldots, r_{max}$ where k is the number of principal components included, r can be estimated by minimizing the following penalized criterion from Bai and Ng (2002)

$$\mathrm{IC2}(k) = ln(V(k)) + k \left(\frac{N+T}{NT} \right) ln \left(C_{NT}^2 \right).$$

Let $\hat{r} = argmin_{0 \leq k \leq r_{max}} \mathrm{IC2}(k)$. Then, $\hat{r}$ equals r with a probability approaching one as $N, T \to \infty$.

To describe the factors which are attributable to the production growth rates of sectors, we need to define a dominant sector precisely. Following PS, an individual unit (sector) is dominant if its variable provides a suitable approximation for a factor, which is defined in terms of an approximate dominant leader (ADL):

Definition of Approximate Dominant Leader (ADL): *Sector i is an ADL for the true factor l, f_{lt}, if and only if* $|f_{lt} - x_{it}| = O(T^{-1/2})$.

The time-t diminishing difference between the variable and the factor ensures that x_i is not asymptotically distinguishable from f_l when $T \to \infty$. In addition, such a diminishing behavior, as indicated by $O(T^{-1/2})$, allows x_i to deviate from the true underlying factor for a limited number of periods, which might lead to the case that there are more than one ADLs attributable to a single factor.[4]

As an example, let $f_t = [f_{1t}', f_{2t}']'$, where f_{1t} is a $r_1 \times 1$ vector which are attributable to ADLs and f_{2t} is a $r_2 \times 1$ vector of the common factors having external sources, such that $r = r_1 + r_2$. Furthermore, suppose that the first sector in the panel data set is the only ADL in the model which corresponds to f_{1t}, such that $|f_{1t} - x_{1t}| = O(T^{-1/2})$ and f_{-1t} is the $(r-1) \times 1$ vector of the other factors. Inserting x_{1t} in (1), we obtain

$$x_t = \Lambda_1 x_{1t} + \Lambda_2 f_{-1t} + u_t, \quad (4)$$

[4] See Section 3 from PS for an elaboration on this issue.

where Λ_1 and Λ_2 are the corresponding matrices of loadings, and $u_t = e_t + O(T^{-1/2})$. PS show that introduction of the approximation error ($O(T^{-1/2})$) into u_t does not affect the distinguishable behavior of the first r eigenvalues of $X'X/NT$, which is now explained by $\{x_{1t}, f_{-1t}\}$, from the remaining eigenvalues for large N and T. This implies that one can assess whether the first sector is an ADL according to the number of factors estimated in u_t, which is asymptotically zero after filtering out the effect of $\{x_{1t}, f_{-1t}\}$.

The key to identification of ADLs is that f_t, properly scaled, is consistently estimated by $\tilde{f}_t$ at each time t as shown by Bai (2003). The estimation error of the principal component method for the unobserved f_t is asymptotically negligible. Nevertheless, f_t is identified from $\tilde{f}_t$ up to scale. This can be seen by noting that f_t and λ_i are not separately identified since $(\lambda_i' M^{-1} M f_t)$ creates an equivalent model for the same observations by any invertible matrix M. To identify the unknown common factors, the PS approach looks into x_is one at a time and includes a single variable, say x_1, together with a vector of $(r-1)$ principal components into the model (4). The following algorithm and the discussions provided in the remaining of this section present the details of this approach.

1. Estimate r and use the principal component method to obtain $\tilde{f}_t$.

2. For a given r, consider a set of m candidate sectors with the variable set $\left\{x_{jt} : j \in \mathscr{I}\right\}$ where $\mathscr{I}$ is the set of the sectors. Do the following steps for each j.

   - Consider r vectors containing the principal components $\tilde{f}_{-kt}$, for $k = 1, ..., r$, where $\tilde{f}_{-kt}$ corresponds to $\tilde{f}_t$ excluding its kth element. For each combination $\{x_{jt}, \tilde{f}_{-kt}\}$, estimate the least squares residuals in the following regression

   $$x_t = \beta_{jk} x_{jt} + \Lambda^{\star}_{jk} \tilde{f}_{-kt} + u^{\star}_{jkt}, \tag{5}$$

   where β_{jk} and $\Lambda^{\star}_{jk}$ are the coefficient matrices.

   - Select x_{jt} as an ADL if the number of factors estimated in the residuals from at least one of these r regressions is zero.

PS argue that the ADLs are consistently identified this way when N and T are large. Note that in Step 2, we need a set of candidate sectors as a starting point.

In the absence of such candidates, we can use a R^2-criterion, where the variables having relatively higher explanatory power for the principal components would be selected as potential dominant sectors. In particular, we estimate R^2 values for every variable in the following regression

$$\tilde{f}_{kt} = a_{kj}x_{jt} + b'_{kj}\tilde{f}_{-kt} + \eta^{\star}_{kjt}, \text{ for } j = 1,...,N \text{ and } k = 1,...,r. \quad (6)$$

If x_{jt} is a good candidate, it would result in a relatively high R^2 value in the regressions for at least one of the principal components, and sector j can then be analyzed as a potential dominant sector.

Clustering of ADLs

The concept of ADLs, as discussed before, allows production growth of a dominant sector to deviate from a true factor for a limited number of periods. This might lead to selection of more than one ADL corresponding to the same factor. To gain a better understanding of how dominant sectors relate to each other and in turn to a true factor, we can use a clustering technique, where we look at all combinations of the ADLs each consisting of r growth series. Then, we check if they can span the space of the factors together. To elaborate on it, consider a two-factor model and a case of four ADLs (this is the setting implied from the analysis of a large number of US sectors in Section 6). Consider all the six pairs of the selected variables. Including these variables as observed factors, we can put the ADLs in one or two groups according to the following regression

$$x_t = C_l \begin{bmatrix} x_{l_1 t} \\ x_{l_2 t} \end{bmatrix} + \zeta_{il}, for\, l = 1,...,6,$$

where C_l is a $N \times 2$ matrix of coefficients, and l_1 and l_2 are indexes of the ADLs in the lth pair. If a factor is detected in the residuals from all six regressions, then all four ADLs belong to a single group since their production growth captures only the effect of one of the underlying factors. But if in some regressions we detect a factor and in others we detect zero factors, then the paired sectors resulting in zero factors belong to two different groups. Using this technique, I conclude in Section 6 that all the ADLs can be put in a single group because a factor remains always in the residuals.

3 Dominant Sectors in Production Networks

Centrality is a key aspect of production networks. Thinking of an economy as a network comprising of N sectors with the network matrix characterized by the corresponding input-output matrix, the central sectors are the most important suppliers of inputs to others and lie at the center of propagation of technology disturbances. In this section, I turn to a stylized structural factor model, which provides a precise definition of the common and idiosyncratic shocks and incorporates the cross-sector input-output linkages. The goal of this section is to provide a link between dominance of a sector as defined before and the concept of centrality from a network perspective. To do so, I start with a structural model of a production network and then describe the network structure under which a factor representation similar to (4) can be obtained.

The key to capturing the production growth of the dominant sectors in the space of the principal components is that the contribution of ADLs in overall variability of X is substantially larger than it from other sectors. This substantial heterogeneity of sectors can be incorporated into a structural factor model as follows. Suppose a structural relationship between x_t exists and be given as

$$x_t = \Gamma_N x_t + z_t, \tag{7}$$

where Γ_N is the $N \times N$ matrix of the cross-sector dependencies containing the exclusion restrictions for the equilibrium values of x_{it}, $i = 1, \ldots, N$, and z_t is a vector of technology innovations. In addition, the ith technology innovation can be written as

$$z_{it} = b'_{i,N} f_{2t} + \varepsilon_{it}, \tag{8}$$

where f_{2t} denotes the $r_2 \times 1$ vector of external common disturbances like before, $b_{i,N}$ is the vector of coefficient, and ε_{it} is the sector-specific technology disturbance which is assumed to be independent across i with $E(\varepsilon_{it}^2) = \sigma_i^2$. I also use the matrix notation $B_N = [b_{1,N}, \cdots, b_{N,N}]'$ from (8).

The model (7)-(8) is a general static relationship which admits the solutions of the multisector general equilibrium models of Acemoglu et al. (2012) and Holly and Petrella (2012), and furthermore, it brings in f_{2t} as common disturbances to productivity of sectors following Foerster et al. (2011).[5] In addition,

[5] Acemoglu et al. (2012) and Holly and Petrella (2012) consider a static variation of the general equilibrium model of Long and Plosser (1983). The solution of a model of this type can be generally

these models characterize Γ_N as

$$\Gamma_N = A_N W_N, \tag{9}$$

where $A_N = diag(a_1, ..., a_N)$ is a diagonal matrix with $a_i \in (0,1)$ for all i, and the $N \times N$ matrix W_N corresponds to the intersectoral input-output matrix.[6] The ijth element of W_N satisfies $w_{ij} \in [0,1]$ for all (i,j) and gives the share of the production of sector j in total intermediate inputs used by sector i.

Sectoral heterogeneity can be assessed based on the matrix of intersectoral dependencies, Γ_N, which specifies how the sector-specific technology disturbances and common productivity components propagate across sectors. Looking at column i of Γ_N, every element amounts to the direct influence of production growth of sector i on another sector. More interestingly for our analysis, the overall influence of an individual sector in the model can be measured by sum of all the elements in the corresponding column in Γ_N. These column-sums correspond to the network outdegree centrality, denoted by $c_i^{out}(N) = \sum_{j=1}^{N} a_j w_{ji}$ for sector i. Another measure to capture centrality in production networks is the Katz-Bonacich centrality.[7] This measure relates to column-sums of R_N in a reduced form of the model (7)

$$x_t = R_N z_t, \tag{10}$$

where $R_N = (I - \Gamma_N)^{-1}$. In the discussion provided here, I focus on the outdegree centrality based on Γ_N, as opposed to R_N which captures propagation of the productivity shocks. The reason is that I can then be consistent with the identification approach which looks at x_t rather than the unobserved z_t.

Introducing a substantial heterogeneity among the sectors such that the out-

cast is an autoregressive-moving-average in form of ARMA(1,1) depending on the timing of material input delivery and whether capital is included as discussed in detail by Foerster et al. (2011).

[6] The production function of sector i is assumed to be Cobb-Douglas with constant returns to scale

$$X_{it} = exp(z_{it})(\prod_{j=1}^{N} X_{ijt}^{w_{ij}})^{a_i} L_{it}^{1-a_i}, \, for \, i = 1, ..., N,$$

where X_{it} is the sector's production, X_{ijt} is the amount of the products of sector j used in the production of sector i, L_{it} is the amount of labor hired, and a_i denotes the share of total intermediate input, such that $a_i \in (0,1)$. The convention $\sum_{j=1}^{N} w_{ij} = 1$ ensures that the function satisfies constant returns to scale. Holly and Petrella (2012) create their model such that their model features common factors, and ε_{it} is serially correlated and independent across i. Acemoglu et al. (2012) focus on a network assessment of the propagation of only ε_t abstracting from external common shocks, and obtain $A = aI_N$ assuming $a = a_1 = ... = a_N$ where I_N is an identity matrix.

[7] See Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015).

degree centrality of some sectors is increasing in N, while the others have limited dependencies, a factor model representation can be obtained from (7)-(8) which includes the most central sectors as the ADLs. Recall that commonness of factors in (4) was defined according to the convergence condition $N^{-1}\Lambda'\Lambda \to \Sigma_\Lambda$ with Σ_Λ being a full rank matrix (Assumption FM2). This condition corresponds to the case that each factor potentially affects all the individual sectors through nonzero loadings. To provide a link between the structural model (7)-(8) and the factor model (4), we need to bring a structure into the coefficient matrices Γ_N and B_N which is similar to Λ. Let $\mathscr{I}$ be the index set of the r_1 most central sectors receiving the largest outdegree centrality (this is the same index set used to indicate the ADLs in the previous section), $\Gamma_{N,\mathscr{I}}$ be the $N \times r_1$ matrix containing the columns of Γ_N corresponding to the sectors $\mathscr{I}$, and $\Gamma_{N,-\mathscr{I}}$ be the $N \times N$ matrix whose columns indicated by $\mathscr{I}$ contain zeros and the other columns are the same as those in Γ_N. The following assumption is made.

Assumption SM (Coefficient matrices Γ_N and B_N): *For all N,* (i) *Let $S_N = [\Gamma_{N,\mathscr{I}}, B_N]$ be the $N \times r$ matrix of deterministic coefficients with bounded elements. S_N satisfies $S_N' S_N / N^{1-2\alpha} \to \Sigma_S$ where Σ_S is a full rank matrix and* $\alpha \in [0, 1/2)$. (ii) *The spectral norm of $\Gamma_{N,-\mathscr{I}}$ satisfies* $||\Gamma_{N,-\mathscr{I}}||_2 < 1$.

Notice that $\alpha \in [0, 1/2)$ is a parameter indicating how influential $\xi_t = [x'_{\mathscr{I}t}, f'_{2t}]'$ is on the sectoral growth rates through the matrix S_N. Initially, consider the case of $\alpha = 0$. Then, S_N has a similar structure to Λ ensuring the contribution of ξ_t has a very dominant role in explaining variance of X. The extent of this dominance becomes smaller and smaller as α increases and approaches 1/2 (the upper bound of the interval). This can be seen by noting that a smaller number of nonzero elements in each column of S_N is needed to satisfy $S_N' S_N / N^{1-2\alpha} \to \Sigma_S$ for a larger value of α. However, the number of nonzero elements remains increasing in N for all $\alpha \in [0, 1/2)$. Assumption SM.ii characterizes the other sectors as if they have limited linkages to others. This feature holds independent of N.

As shortly will be discussed, Assumption SM ensures that the central sectors in the production network correspond to the ADLs in the model (4), and in turn they can be consistently identified using the PS approach for some value of α. To this end, consider the structural model (7)-(8) which is now written as

$$x_t = \Gamma_{N,\mathscr{I}} x_{\mathscr{I},t} + B_N f_{2t} + \Gamma_{N,-\mathscr{I}} x_t + \varepsilon_t.$$

Rearranging the terms and multiplying both sides by $\bar{\Gamma}_N = (I_N - \Gamma_{N,-\mathscr{I}})^{-1}$, we obtain

$$x_t = \bar{\Lambda}_{N,1} x_{\mathscr{I},t} + \bar{\Lambda}_{N,2} f_{2t} + v_t, \tag{11}$$

where $\bar{\Lambda}_{N,1} = \bar{\Gamma}_N \Gamma_{N,\mathscr{I}}$, $\bar{\Lambda}_{N,2} = \bar{\Gamma}_N B_N$, and $v_t = \bar{\Gamma}_N \varepsilon_t$. The factor model representation (11) is obtained as a reduced form having $x_{\mathscr{I},t}$ initially separated from x_t. The error terms v_{it}, for $i = 1,...,N$, whose cross-section dependence is governed by $\bar{\Gamma}_N$, are weakly dependent following Assumption SM.ii.[8] Furthermore, the particular arrangement of nonzero elements in the columns of S_N leads ξ_t to be dominant relative to v_t, with this relative dominance being diminishing in α as follows. The growth rates of the central sectors (or equivalently the ADLs under the setting of Section 2) act as common factors for $\alpha = 0$. As α increases, the relative dominance narrows down but still holds under $\alpha \in [0, 1/2)$. Considering α which is marginally outside of this range ($\alpha \geq 1/2$), the difference between the first r eigenvalues and the rest is not visually distinguishable anymore even for a very large N. See Lemma 1 in Appendix.

Note that centrality, as discussed in this section, is a more general concept than sectoral dominance in factor model analysis because centrality of a sector is defined as a varying function of $\alpha \in [0, 1/2)$, while only the case of $\alpha = 0$ was covered in Section 2. The next section brings the parameter α into the factor model in a consistent way and analyzes consistency of the PS approach in identifying the ADLs under different values of α.

Sections 2-3 elaborate on two aspects of our analysis. First, given that the model (7)-(8) is the true data generating process, it is expected that the domi-

[8] This assumption implies convergence of the power series $\bar{\Gamma}_N = I_N + \Gamma_{N,-\mathscr{I}} + \Gamma^2_{N,-\mathscr{I}} + \cdots$ for all N. Notice $||\Gamma_{N,-\mathscr{I}}||_2 < 1$ is stronger than $\rho_1(\Gamma_{N,-\mathscr{I}}) < 1$ which is made in VAR models for covariance-stationary time series. The assumption is made to ensure boundedness of variance of v_t in the reduced form model (11) as $N \to \infty$. To see this, consider

$$\Sigma_u = E(\bar{\Gamma}_N \varepsilon_t \varepsilon_t' \bar{\Gamma}_N') \leq \bar{\sigma}^2 \bar{\Gamma}_N \bar{\Gamma}_N',$$

where $\bar{\sigma}^2 = max_i\{\sigma_i^2\}$. Then, it is implied that Σ_u has bounded eigenvalues since

$$\rho_1(\bar{\Gamma}_N \bar{\Gamma}_N') = (||\bar{\Gamma}_N||_2)^2 \leq K < \infty,$$

where boundedness of $||\bar{\Gamma}_N||_2$ is implied from $||\Gamma_{N,-\mathscr{I}}||_2 < 1$. One can come up with examples such that $\rho_1(\Gamma_{N,-\mathscr{I}}) < 1$ and $||\Gamma_{N,-\mathscr{I}}||_2 \geq 1$ where variance of v_{it} for different i is increasing in N. See the example in Section 3 from Chudik and Pesaran (2011) in the context of infinite dimensional VARs. This way, separating $\Gamma_{N,\mathscr{I}}$ initially from Γ_N and then obtaining the reduced form of the model, we can incorporate extreme asymmetries in terms of the outdegree centrality which otherwise would not be allowed under the setting from Acemoglu et al. (2012) when $N \to \infty$, where the focus of the authors is on $R_N = (I - \Gamma_N)^{-1}$.

nant sectors, which are identified using the PS approach, match the most central sectors whose greater outdegree centrality is implied from the corresponding sectoral input-output table. Second, this section refers to a stylized general equilibrium framework focusing on the material input-output linkages. There might be other types of intersectoral linkages such as cross-sector capital flows (Foerster et al. (2011) and Bouakez et al. (2014)), or more probably a combination of the intersectoral material and capital input-outputs. The PS approach to identification is based on the statistical factor model which is in general less restrictive than structural model described in this section, in the sense that it does not make any particular assumption on sources of the linkages. Thus, it can be used at first. Next, some measures of relative importance of the sectors in terms of supplying material inputs, supplying capital products, or their size can be used to assess if there are common features among the identified sectors.

4 Identification in Less-Than-Strong Factor Models

The asymptotic properties of the PS approach in a model including common factors ($\alpha = 0$) are discussed by PS, which was briefly reviewed in Section 2. To account for the factors which have different degrees of influence, a more general model is considered in this section by introducing $\alpha \in [0, 1/2)$ into the matrix of factor loadings in (1). In particular, I here answer two questions. Does the principal component method consistently capture the ADLs at each time t when α deviates from 0 and how large α could be in order that the estimator remains time-t consistent? What is the range of α for which the PS approach can consistently identify the ADLs? Theorems 1 and 2 in this section answer these questions, respectively. The discussion provided here sheds light on the limits of the principal component-based methods used in Section 2 to identify the ADLs.

Suppose Λ is the matrix of loading which satisfies Assumption FM2. To deviate from the case of $\alpha = 0$, consider the following data generating process

$$x_t = \Lambda_N f_t + e_t, \tag{12}$$

for $t = 1, ..., T$, with the assumption that

$$\Lambda_N' \Lambda_N / N^{1-2\alpha} \to \Sigma_{\Lambda_N},$$

where Σ_{Λ_N} is a full rank matrix with $\alpha \in [0, 1/2)$. Chudik et al. (2011) define strong and semi-strong factors according to column-sums of $\Lambda_N = (\lambda_{N,il})$. Let $s_l = \sum_{i=1}^{N} |\lambda_{N,il}|$, for $l = 1, ..., r$, where $s_l = O(N^{1-\alpha})$. Then, f_t is a vector of strong factors if $\alpha = 0$, and the factors are semi-strong given that $\alpha \in (0, 1/2)$.[9] Kapetanios and Marcellino (2010) consider a less-than-strong factor model by assuming that Λ_N in (12) satisfies

$$\Lambda_N = N^{-\alpha}\Lambda. \tag{13}$$

This builds on the literature on instrumental variable estimator in the presence of so many weak instruments where an instrument appears to be weak as sample size goes to infinity, and it provides a mathematically tractable way to assess properties of the principal component method for different values of α. I here closely follow Kapetanios and Marcellino (2010) and use (13).

The denominator N^{α} makes the contribution of the common component in $X'X/NT$ a negative function of α. Lemma 1 in the appendix shows

$$\rho_1(X'X/NT) = O_p(N^{-2\alpha}), \tag{14}$$

if $\alpha \in [0, 1/2)$ and $N^{2\alpha}/T \to K < \infty$. The term (14) is driven by $\Lambda_N f_t$ which is dominant relative to e_t when $\alpha \in [0, 1/2)$. From the literature on factor analysis, we know the factor estimates and performance of the factor number estimators are affected as factors become less dominant relative to the idiosyncratic errors. Under this setting, it is not clear if the principal component method remains consistent for all values $\alpha \in [0, 1/2)$. Kapetanios and Marcellino (2010) establish that the time average of the squared difference between the factors estimates and the scaled true factors diminishes if $\alpha \in [0, 1/4)$ and $N^{4\alpha}/T \to 0$. Onatski (2012) shows the principal component method is inconsistent considering the case of $\alpha = 1/2$ under $N/T \to K < \infty$. In what follows, I initially find the sufficient conditions ensuring the consistent estimation of factors at each time t.

Before doing so, I assume the number of factors, r, is known. Though, this assumption seems reasonable under $\alpha = 0$, it is expected that performance of IC2 deteriorates similarly to the principal component method for larger values of α. I confine the assessment of this factor number estimator to the simulation studies provided in Section 5, where it is shown that IC2 performs well in small samples for a given $\alpha \in [0, 1/2)$ for which the principal component method turns

[9] Notice Chudik et al. (2011) define these rates in terms of 1-α which makes for example a strong factor correspond to $1 - \alpha = 1$.

out to be performing well.

The following theorem and corollary establish the consistency of the principal component method for each time t for different values of α.

Theorem 1. *Let $\Lambda_N = N^{-\alpha}\Lambda$. Under Assumptions FM1-4, $\tilde{f}_t$ converges to $H'f_t$ as $N,T \to \infty$ given that $\alpha \in [0,1/4)$ and $N^{4\alpha}/T \to 0$, such that the convergence rate is given as*

$$\tilde{f}_t - H'f_t = O_p(\frac{N^{4\alpha}}{C_{NT}^2}) + O_p(N^{\alpha-1/2}),$$

with $H = (\Lambda_N'\Lambda_N/N)\left(F'\tilde{F}/T\right)V_{NT}^{-1}$ being a rotation matrix, V_{NT} is a diagonal matrix containing the first r eigenvalues of XX'/NT, and $C_{NT}^2 = min\{N,T\}$.

Corollary 1. *Let $\Lambda_N = N^{-\alpha}\Lambda$. Under Assumptions FM1-4,*
(i) if $N^{1/2+3\alpha}/C_{NT}^2 \to 0$, we have

$$N^{1/2-\alpha}\left(\tilde{f}_t - H'f_t\right) = O_p(1);$$

(ii) if $N^{1/2+3\alpha}/C_{NT}^2 \to \tau > 0$, we have

$$\frac{C_{NT}^2}{N^{4\alpha}}\left(\tilde{f}_t - H'f_t\right) = O_p(1).$$

Theorem 1 follows Theorem 1 from Bai (2003). It implies that the estimation error $(\tilde{f}_t - H'f_t)$ is asymptotically negligible given that $N^{4\alpha}/C_{NT}^2 \to 0$. This holds under $\alpha \in [0,1/4)$ and $N^{4\alpha}/T \to 0$. Corollary 1 puts these results in terms of the Bai's theorem focusing only on the convergence rates. Bai shows

$$N^{1/2}\left(\tilde{f}_t - H'f_t\right) = O_p(1),$$

given that $N^{1/2}/T \to 0$, but as $N^{1/2}/T \to \tau > 0$, he shows

$$T\left(\tilde{f}_t - H'f_t\right) = O_p(1).$$

Under $\alpha > 0$, there are slower convergence rates such that a nonzero α leads the rates to decline from $1/N^{1/2}$ to $1/N^{1/2-\alpha}$ in (*i*), and from $1/T$ to $N^{4\alpha}/min\{N,T\}$ in (*ii*). It appears that as factors become less influential, the principal component estimator converges to the scaled f_t at a slower rate and $\alpha \in [0,1/4)$ is a sufficient condition for $\tilde{f}_t$ to remain consistent.

Using Theorem 1, I next analyze the consistency properties of the PS approach for identifying the ADLs. Suppose a general relationship between a candidate variable, x_{jt}, and a true factor, f_{lt}, holds as follows

$$f_{lt} = x_{jt} + o_{ljt}/\sqrt{T} + \delta_{lj}\upsilon_{jt}, \tag{15}$$

where o_{ljt} and υ_{jt} are two random variables with finite variances. This way, x_{jt} is an ADL if $\delta_{lj} = 0$. Dominant sectors are identified following the steps described in Section 2. Recall that for each candidate sector, the regression model (5) is estimated r times for the combinations of the principal components and the candidate sector's production growth, and next, whether the sector's growth rate provides a suitable approximation for a factor is implied according to the number of factors in the least squares residuals of these regressions $\hat{u}^{\star}_{jkt}$, for $k = 1, ..., r$. Let $\widehat{n_f}(\hat{u}^{\star}_{jkt})$ denote the corresponding number of factors estimated. We then have the following result.

Theorem 2. *Let* $\Lambda_N = N^{-\alpha}\Lambda$. *Given that* $\alpha \in [0, 1/6]$ *and* $N/T \to K < \infty$*, and under Assumptions FM1-4*
(i) if $\delta_{lj} = 0$*, we have*

$$\underset{N,T\to\infty}{Prob}\left[\widehat{n_f}(\hat{u}^{\star}_{j1t}) = 0, \text{ or}, \widehat{n_f}(\hat{u}^{\star}_{j2t}) = 0, \text{ or}, ...\widehat{n_f}(\hat{u}^{\star}_{jrt}) = 0\right] = 1;$$

(ii) if $\delta_{lj} \neq 0$*, we have*

$$\underset{N,T\to\infty}{Prob}\left[\widehat{n_f}(\hat{u}^{\star}_{j1t}) = 0, \text{ or}, \widehat{n_f}(\hat{u}^{\star}_{j2t}) = 0, \text{ or}, ...\widehat{n_f}(\hat{u}^{\star}_{jrt}) = 0\right] = 0.$$

Theorem 2 states $\alpha \in [0, 1/6]$ and $N/T \to K < \infty$ as the sufficient conditions for consistent identification of the ADLs and extends the results of PS for the case of α greater than zero. In particular under these conditions, if x_j is a suitable approximation for an underlying factor, then $\widehat{n_f}(\hat{u}^{\star}_{jkt}) = 0$ holds for at least one of $\hat{u}^{\star}_{jkt}$, $k = 1, ..., r$, with a probability approaching one. In the next section, small sample properties of the identification approach is assessed using simulated data and it is shown that under relatively less constraining conditions the approach still performs well in identifying the ADLs.

The discussion about the acceptable values of α is meant to highlight how robust the factor analytic methods are under different degrees of factors influence. An estimator for α has been proposed by Bailey, Kapetanios, and Pesaran

(2015).[10] I use this estimator in Section 6 to first assess whether the cross-sectional properties of the sectoral growth rates in the US accord with the conditions of Theorems 1-2 and the Monte Carlo results of the next section.

5 Monte Carlo Study

Consider a two-factor model

$$x_{it} = \lambda_{N,i1} f_{1t} + \lambda_{N,i2} f_{2t} + e_{it}, \text{ for } i = 1, ..., N, \text{ and } t = 1, ..., T.$$

The factors are correlated as

$$\begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} = U \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix},$$

where U is obtained from the Cholesky decomposition of $\Omega = Var(f_t)$ with $\Omega = [2, 0.2; 0.2, 1]$, and

$$w_{lt} = \rho w_{lt-1} + \varepsilon_{lt}, \text{ for } l = 1, 2,$$

with $\varepsilon_{lt} \overset{i.i.d}{\sim} N(0, 1-\rho^2)$, and $\rho = 0.5$.

Elements of the loading matrix $\Lambda_N = (\lambda_{N,ij})$ are normally distributed

$$\lambda_{N,ij} \overset{i.i.d}{\sim} N(0, \sigma^2_{N,\alpha}),$$

with $\sigma_{N,\alpha} = 1/N^{\alpha}$. With the randomly generated loadings, the model is more general in comparison to the case of deterministic loadings in Assumption FM2. The analytical results would not be affected anyway since $\lambda_{N,ij}$ has finite fourth moment and are independent from the factors and the idiosyncratic errors in our data generating process.

The error term follows

$$e_{it} = \phi e_{it-1} + v_{it} + \beta \sum_{j \neq i, j=-J}^{J} v_{i+j,t}, \tag{16}$$

with $v_{it} \overset{i.i.d}{\sim} N(0, (1-\phi)/(1+2J\beta^2))$. e_{it} is serially correlated and also cross-

[10] The authors propose an estimator for the "exponent of cross-sectional dependence" which formally equals $1-2\alpha$ under (13). To estimate α, I use the estimator from Section 3.2 of their paper using both of the proposed bias correction terms. This bias corrected version of the estimator is denoted by $\overset{0}{\alpha}$.

sectionally dependent on the J-upstream and downstream neighbors. I set $\phi = 0.5$, $J = 4$, and $\beta = 0.1$.

The panel data sets are generated for $\{N, T = 50, 100, 200, 500\}$, such that the data is initially generated for $T + 100$ observations and $N + 20$ units. Next, the first 100 observations are dropped from the beginning of the series and I keep N units from the middle of cross-section units. This way, all the units have $2 \times J$ neighbors for $J \leq 10$. The Monte Carlo results are obtained according to 2000 replications.

I conduct three experiments as follows.

Experiment 1. Number of factors is estimated for different values of α using IC2. This sheds light on the sensitivity of IC2 to different factor structures.

Experiment 2. A case of two ADLs is considered, where the ADLs are generated according to (17) and are replaced by the first two units in the panel data set such that

$$x_{jt} = f_{jt} + \nu_{jt}/\sqrt{T}, \text{ for } j = 1, 2, \tag{17}$$

with $\nu_{jt} \overset{i.i.d}{\sim} N(0,1)$ for all (j,t). In this experiment, the first two sectors are taken as potential dominant sectors and next the PS approach is used to assess how frequently x_1 is correctly identified as an ADL.

Experiment 3. As opposed to the previous experiment that the candidate sectors were known, I first use the R^2 values obtained from regression (6) to select ten candidate sectors. A more general case of four ADLs is considered where two ADLs are incorporated for each factor as

$$x_{jt} = \begin{cases} f_{1t} + \nu_{j1t}/\sqrt{T}, \text{ for } j = 1, 2, \\ f_{2t} + \nu_{j2t}/\sqrt{T}, \text{ for } j = 3, 4, \end{cases} \tag{18}$$

with $\nu_{jlt} \overset{i.i.d}{\sim} N(0,1)$ for all (j,l,t). Furthermore, the idiosyncratic errors of the nondominant units are generated differently from (16). In particular, for $i = 5, \ldots, N$, e_{it} is now generated according to

$$e_{it} \overset{i.i.d}{\sim} N(0, \sigma_i^2), \tag{19}$$

where $\sigma_i^2 = \frac{1}{T}\sum_t C_{it}^2$ with $C_{it} = \lambda_{N,i1} f_{1t} + \lambda_{N,i2} f_{2t}$. This way, a nondominant unit would never act like an ADL since the variability of e_{it} and f_t remain proportional.

This is a general setting which closely follows the simulation study design of PS, and in addition, it incorporates different values of α to adjust the factors influence in the model. I consider $\{\alpha = 0, 0.1, 1/6, 0.2, 0.25, 0.3, 0.4, 0.5\}$, which cover the whole range of strong to relatively weak factors. PS confine their Monte Carlo study to $\{\alpha = 0, 0.5\}$. I extend this analysis considering 1/6 and 0.25, as suggested by Theorems 1-2, and some other intermediate values.

Results

In the first experiment, the performance of IC2 is examined. Table (1) presents the average number of factors estimated and the frequency of correctly detecting two factors. The results show the number of factors is estimated well for $\alpha \leq 0.2$ when $N, T \geq 100$. When $\alpha = 0.25$, this performance drops and a larger sample size such as $N \geq 100$ and $T \geq 200$ is required to restore it. Bai and Ng (2002) consider the case of $\alpha = 0$ and show that IC2 tends to overestimate the number of factors in small samples. Although this overestimation notably appears in the results here, it becomes smaller with an increase in α such that for $\alpha = 1/6$, IC2 appears to perform the best. This improvement with an increase in α is attributable to the tendency of IC2 to provide smaller estimates as the dominance of f_t relative to e_t reduces. This offsets the overestimation of the number of factors and leads to a better performance reflected in the higher frequencies. In addition, we observe a substantial reduction in the frequencies as α becomes larger than 0.3. With the factors becoming less dominant corresponding to larger values of α, we would detect zero factors even when $N, T = 500$. These results accord with the consistency condition $\alpha \in [0, 0.25)$ from Theorem 1 and indicate the number of factors can be estimated well for the values of α under which the principal component method remains consistent.

In the second experiment, I examine the small sample properties of the PS approach for the case of two known ADLs, which are placed as first two units in the panel data set. The upper panel of Table (2) presents the frequency of detecting x_1 as an ADL. The results show a good small sample performance for $\alpha \leq 0.2$ since the ADL is identified with a frequency close to one for almost all combinations of N and T. For $\alpha = 0.25$, the frequencies slightly reduce, but they recover as N and T increase. Furthermore, very similar to the pattern we previously observed for estimation of number of factors, the results show a poor performance of the PS approach in identifying the dominant units when the data is generated using larger values of α.

In most of applications, the potential dominant units are not known and we

Avg. Number of Factors Estimated										
		α	0	0.1	1/6	0.2	0.25	0.3	0.4	0.5
N	T									
50	50		2.51	2.20	2.08	1.98	1.60	1.10	0.25	0.04
50	100		2.80	2.27	2.14	2.09	1.95	1.57	0.51	0.05
50	200		3.03	2.26	2.11	2.07	2.03	1.78	0.74	0.04
50	500		3.24	2.18	2.05	2.02	2.00	1.89	0.85	0.03
100	50		2.41	2.16	2.08	2.03	1.69	1.11	0.18	0.03
100	100		2.04	2.00	2.00	2.00	1.81	1.20	0.10	0.00
100	200		2.05	2.00	2.00	2.00	1.99	1.66	0.37	0.00
100	500		2.07	2.00	2.00	2.00	2.00	1.95	0.76	0.00
200	50		2.21	2.07	2.03	1.99	1.60	0.93	0.05	0.01
200	100		2.01	2.00	2.00	2.00	1.88	1.16	0.03	0.00
200	200		2.00	2.00	2.00	2.00	1.99	1.39	0.03	0.00
200	500		2.00	2.00	2.00	2.00	2.00	1.97	0.44	0.00
500	50		2.03	2.00	1.99	1.87	1.21	0.44	0.00	0.00
500	100		2.00	2.00	2.00	2.00	1.70	0.87	0.00	0.00
500	200		2.00	2.00	2.00	2.00	2.00	1.16	0.00	0.00
500	500		2.00	2.00	2.00	2.00	2.00	1.89	0.00	0.00

Frequency of Estimating Two Factors										
		α	0	0.1	1/6	0.2	0.25	0.3	0.4	0.5
50	50		0.63	0.82	0.85	0.79	0.53	0.21	0.01	0.00
50	100		0.46	0.76	0.86	0.90	0.83	0.51	0.03	0.00
50	200		0.35	0.77	0.89	0.93	0.93	0.72	0.03	0.00
50	500		0.29	0.84	0.96	0.98	0.99	0.88	0.03	0.00
100	50		0.67	0.86	0.91	0.86	0.61	0.22	0.01	0.00
100	100		0.96	1.00	1.00	1.00	0.81	0.23	0.00	0.00
100	200		0.95	1.00	1.00	1.00	0.99	0.66	0.00	0.00
100	500		0.93	1.00	1.00	1.00	1.00	0.95	0.00	0.00
200	50		0.81	0.94	0.96	0.93	0.58	0.12	0.00	0.00
200	100		0.99	1.00	1.00	1.00	0.88	0.18	0.00	0.00
200	200		1.00	1.00	1.00	1.00	0.99	0.39	0.00	0.00
200	500		1.00	1.00	1.00	1.00	1.00	0.97	0.00	0.00
500	50		0.97	1.00	0.99	0.87	0.27	0.01	0.00	0.00
500	100		1.00	1.00	1.00	1.00	0.70	0.02	0.00	0.00
500	200		1.00	1.00	1.00	1.00	1.00	0.16	0.00	0.00
500	500		1.00	1.00	1.00	1.00	1.00	0.89	0.00	0.00

Table 1: The average number of factors estimated and the frequency of correctly estimating two factors are reported. The results are obtained using IC2.

Frequency of Correctly Identifying x_1 (Case of Two Known ADLs)										
		α	0	0.1	1/6	0.2	0.25	0.3	0.4	0.5
N	T									
50	50		0.93	0.96	0.93	0.88	0.72	0.61	0.20	0.02
50	100		0.91	0.94	0.96	0.95	0.87	0.66	0.39	0.03
50	200		0.84	0.94	0.96	0.97	0.95	0.76	0.53	0.03
50	500		0.77	0.94	0.98	0.99	0.99	0.89	0.56	0.03
100	50		0.97	0.98	0.98	0.94	0.80	0.66	0.15	0.02
100	100		0.99	1.00	1.00	1.00	0.86	0.67	0.10	0.00
100	200		0.99	1.00	1.00	1.00	0.99	0.73	0.36	0.00
100	500		0.99	1.00	1.00	1.00	1.00	0.96	0.72	0.00
200	50		0.99	1.00	1.00	0.97	0.85	0.69	0.05	0.01
200	100		1.00	1.00	1.00	1.00	0.92	0.77	0.03	0.00
200	200		1.00	1.00	1.00	1.00	0.99	0.66	0.03	0.00
200	500		1.00	1.00	1.00	1.00	1.00	0.97	0.44	0.00
500	50		1.00	1.00	0.99	0.97	0.82	0.42	0.00	0.00
500	100		1.00	1.00	1.00	1.00	0.89	0.82	0.00	0.00
500	200		1.00	1.00	1.00	1.00	1.00	0.79	0.00	0.00
500	500		1.00	1.00	1.00	1.00	1.00	0.90	0.00	0.00

Frequency of Identifying All Four ADLs (Case of Four Unknown ADLs)										
		α	0	0.1	1/6	0.2	0.25	0.3	0.4	0.5
50	50		0.99	1.00	0.96	0.88	0.57	0.22	0.01	0.00
50	100		0.99	1.00	1.00	0.99	0.89	0.55	0.03	0.00
50	200		1.00	1.00	1.00	1.00	0.98	0.75	0.04	0.00
50	500		1.00	1.00	1.00	1.00	1.00	0.89	0.03	0.00
100	50		1.00	1.00	1.00	0.94	0.65	0.23	0.01	0.00
100	100		1.00	1.00	1.00	1.00	0.81	0.23	0.00	0.00
100	200		1.00	1.00	1.00	1.00	0.99	0.66	0.00	0.00
100	500		1.00	1.00	1.00	1.00	1.00	0.95	0.00	0.00
200	50		1.00	1.00	1.00	0.96	0.59	0.12	0.00	0.00
200	100		1.00	1.00	1.00	1.00	0.88	0.18	0.00	0.00
200	200		1.00	1.00	1.00	1.00	0.99	0.39	0.00	0.00
200	500		1.00	1.00	1.00	1.00	1.00	0.97	0.00	0.00
500	50		1.00	1.00	0.99	0.87	0.27	0.01	0.00	0.00
500	100		1.00	1.00	1.00	1.00	0.70	0.02	0.00	0.00
500	200		1.00	1.00	1.00	1.00	1.00	0.16	0.00	0.00
500	500		1.00	1.00	1.00	1.00	1.00	0.89	0.00	0.00

Table 2: The frequency of correctly identifying the ADLs is reported. The cases of (a) two known ADLs (x_1, x_2); and (b) four unknown ADLs $(x_1, \ldots, x_4)$ are considered corresponding to the second and third experiments. For the latter case, the R^2-criterion is used first to select ten candidate variables. The upper panel of the table reports the frequency of identifying x_1 as an ADL in the case (a), and the lower panel reports the frequency of identifying all four ADLs for the case (b).

		α	0	0.1	1/6	0.2	0.25	0.3
N	T							
50	50		0.41	0.24	0.21	0.25	0.48	0.66
50	100		0.54	0.24	0.14	0.10	0.17	0.48
50	200		0.65	0.23	0.11	0.07	0.07	0.28
50	500		0.71	0.16	0.05	0.02	0.01	0.12
100	50		0.34	0.15	0.10	0.15	0.38	0.66
100	100		0.04	0.00	0.00	0.00	0.19	0.74
100	200		0.05	0.00	0.00	0.00	0.01	0.34
100	500		0.07	0.00	0.00	0.00	0.00	0.05
200	50		0.19	0.06	0.04	0.07	0.42	0.69
200	100		0.01	0.00	0.00	0.00	0.12	0.80
200	200		0.00	0.00	0.00	0.00	0.01	0.61
200	500		0.00	0.00	0.00	0.00	0.00	0.03
500	50		0.03	0.00	0.01	0.13	0.68	0.43
500	100		0.00	0.00	0.00	0.00	0.30	0.84
500	200		0.00	0.00	0.00	0.00	0.00	0.84
500	500		0.00	0.00	0.00	0.00	0.00	0.11

Table 3: The frequency of any nondominant units being falsely identified as ADLs is reported. Note that in this table in contrast to the previous tables, a larger frequency indicates that the PS approach performs worse.

need to select a set of candidate units first. In the third experiment, the simulation study is developed to incorporate four ADLs which are unknown. Ten candidate variables are selected first using the R^2-criterion, and next, the frequency of identifying all four ADLs is computed. The lower panel of Table (2) presents the frequency of identifying all four ADLs. It appears that the results are very similar to those obtained under the previous case of known ADLs. It is worth mentioning a feature in the results corresponding to $\alpha = 0.25$, where the frequencies appear to be very different under the cases of large N and small T and the other way around. For example when N=50 and T=500, the ADLs are detected with a frequency equal to one, but in contrast when N=500 and T=50, the frequency reduces to 0.27. This is in line with the second condition from Theorem 2, which confines divergence rate of N relative to T and suggests that, in a case of small N, a large T would help identify the ADLs as α increases and goes slightly above 1/6.

In the third experiment, I finally elaborate on the frequency that the nondominant units are falsely identified as ADLs. Table (3) reports how frequently any units other than first four units are identified (α equal to 0.4 and 0.5 are excluded because IC2 hardly detects any factor in the first place, that in turn results in smaller frequencies for false identification). It is observed that for a small N and T, the nondominant units are selected frequently, but the performance of the PS approach recovers as N and T increase and leads the frequencies of false identification to reduce to zero.

In sum, the Monte Carlo results indicate good small sample performance of the PS approach under a moderate deviation of α from zero. The results accord with the analytical findings of Theorem 2 that $\alpha \in [0, 1/6]$ and $N = O(T)$ sufficiently ensure consistency of the PS approach. Furthermore, it is shown that the dominant units can be identified well under the values of α, which are slightly outside of this range. In particular when $\alpha = 0.25$, a large temporal dimension in the panel data set such as $T \geq 200$ is required to improve reliability of the PS approach.

6 Dominant Sectors in the US

In this section, I use quarterly growth rates of sectoral industrial production to identify the dominant sectors in the US. The analysis proceeds in two steps using the data set provided by Foerster et al. (2011). In the first step, the PS approach is used to identify the dominant sectors over the period 1972Q2 to 2007Q4. Three levels of sectoral disaggregation are considered (L3-5). This provides $N = 88$, 117, and 138, respectively for L3-5. In the second step, I look into the inter-sectoral material input-output tables, cross-sector capital use tables, and sectoral shares in the aggregate industrial production to elaborate on the potential common features among the dominant sectors identified. The material input-output tables for L3-5 and the capital use tables for L3-4 from 1997 are provided in the data set, which are used in this analysis.[11] Furthermore, let s_i denote the aggregation share of sector i such that the growth rate of aggregate industrial production is computed as $x_t^{agg} = \sum_{i=1}^{N} s_i x_{it}$. I compute s_i as the time average over the sample period, $s_i = \frac{1}{T}\sum_{t=1}^{T} s_{it}$, where s_{it} is the corresponding share at time t.

[11] It is worth mentioning that in generating the input-output tables, Foerster, Sarte and Watson distinguish between the material and capital products according to the period they are used in production. A product is viewed as a capital product if it is used for a period longer than one year.

Level of Sectoral Disaggregation	90% Confidence Band
L3	[0.018,0.042]
L4	[0.030,0.044]
L5	[0.044,0.052]

Table 4: The 90% level confidence bands for the estimate of α are reported for different levels of sectoral disaggregation.

6.1 Identification of the Dominant Sectors

I start with estimation of the number of factors and α which was defined in (13). Using IC2, two factors are detected among the sectoral growth rates for all three levels of disaggregation. This follows the factor model analysis from Foerster et al. (2011), who find two common shocks account for a large portion of variability in the growth rate of aggregate industrial production. When I use the method from Bailey et al. (2015) to estimate α, a large exponent of cross-sectional dependence is found in the data set. Table (4) presents the 90% level confidence bands for the estimates of α, which are in the interval $[0, 1/6]$ (recall that $\alpha \in [0, /16]$ is a condition from Theorem 2). Hence, we can make sure that the PS approach provides a reliable identification approach when both N and T are large.[12]

Taking the presence of two aggregate shocks in the model as a starting point, I turn to identification of the ADLs. To this end, ten candidate sectors are initially selected using the R^2 values obtained from the regression (6). Next, I use the PS approach to analyze whether any of these candidate sectors are dominant. I conduct this analysis on all three levels of disaggregation for the purpose of robustness check. For example, in identifying a sector as dominant according to the fourth level of disaggregation (L4), I accept that sector as an ADL if one of its subsectors using the data for L5 and its supersector in L3 are also identified as dominant sectors.

Using the entire sample, I identify three sectors as ADLs. The sectors are 'commercial and services industrial machinery and other general purpose ma-

[12]Note there is a possibility that the values of α corresponding to the underlying factors are different, and as discussed by Bailey et al. (2015), the estimate of α corresponds to their minimum ($\alpha = min\{\alpha_1, \alpha_2\}$, where α_1 and α_2 assess how influential f_{1t} and f_{2t} are, respectively). Then, an informal assessment of a linear combination of α_1 and α_2 which is strictly greater than their minimum can be done by estimating α again after filtering out the effects of the first principal component. Doing so, I obtain one factor using IC2 for L3-5, with the corresponding 90% level confidence bands for the estimate of α being as [0.156,0.206], [0.181,0.263], and [0.119,0.146], respectively. The latter results show that even after removing the effects of the first component, the degree of cross-sectional dependence remains suitably high for the estimation and identification of unknown factors.

chinery' ('commercial and services machinery' for short), 'metalworking machinery', and 'electrical equipment'. The dominance of the the first two sectors ('commercial and services machinery', and 'metalworking machinery') is implied even after filtering out the first principal component by regressing the rates on the first principal component and repeating the analysis on the retrieved residuals.[13] The latter analysis does not filter out the effects of one true factor, but a combination of both. However, it elaborates on dominance of two heavy machinery industries, which maintain their role as ADLs even after removing the first component.

To analyze robustness of these findings to different sample periods, I look into 97 expanding window, such that the shortest window covers 1972Q2 to 1983Q4, the second window covers 1972Q2 to 1984Q1, and the width of successive windows continues to grow by one quarter each with the widest window covering the entire sample. The choice of the ending date 1983Q4 for the first window relies on the findings of Foerster et al. (2011), who show there is a fall in the importance of the aggregate shocks over 1984 to 2007 relative the previous period, though the variability of sectoral shocks does not change much throughout the entire sample. This leads the contribution of the idiosyncratic shocks in the variability of aggregate industrial production to increase considerably during 1984 to 2007, which potentially affects our analysis in different subsamples. Repeating the analysis for each window, the ADLs are identified. Table (5) presents the dominant sectors and the fraction of windows (out of 97 in total) in which they are dominant using L3-5. Nine sectors are identified out of which only 'commercial and services machinery' and 'metalworking machinery' are selected with a frequency equal to one. Ranking the sectors according to how frequently they are identified using L5, we see 'electrical equipment' and 'fabricated metals: forging and stamping' are next in the list with being dominant in 91% of the expanding periods. Note all the sectors selected are classified under the same name in L3-5 with an exception of 'millwork' which shows up only under L4-5, while it is classified as 'other wood products' under L3. Furthermore, looking at the ending dates of the windows over 1983Q4 to 2007Q4 in which the sectors are dominant, it is found that the sectors ranked 5th and 6th ('textile and fabric finishing, and fabric coating mills' and 'engine, turbine, and power transmission equipment') are always detected as dominant before 2000, while including more recent sectoral growth rates in the model, they are never

[13] Here, three sectors are selected. The additional sector is 'machine shops; turned products; and screws, nuts, and bolts'.

	Sector	L3 ($N = 88$)	L4 ($N = 117$)	L5 ($N = 138$)
1.	Commercial and Service Machinery*	1.00	1.00	1.00
2.	Metalworking Machinery*	1.00	1.00	1.00
3.	Electrical Equipment*	0.64	0.86	0.91
4.	Fabricated Metals: Forging and Stamping	0.13	0.54	0.91
5.	Textile and Fabric Finishing, and Fabric Coating Mills	0.74	0.80	0.73
6.	Engine, Turbine, and Power Transmission Equipment	0.31	0.38	0.72
7.	Other Wood Products (under L3)	0.30		
	Millwork (under L4-5)		0.07	0.24
8.	Machine Shops; Turned Products; and Screws, Nuts, and Bolts	0.05	0.13	0.09
9.	Fiber, Yarn, and Thread Mills	0.73	0.57	0.01

Table 5: The dominant sectors and the frequency that they are selected in different expanding windows are presented according to three levels of sectoral disaggregation (L3-5). There are 97 windows in total. Notice all the sectors are classified under the same name in L3-5 with an exception of 'millwork' which shows up only under L4-5. It is a subsector of 'other wood products' under L3.

selected thereafter.

The three dominant sectors identified using the entire sample might deviate from the true factors in some periods. As discussed in Section 2, this limited deviation is allowed when the ADLs are selected according to the number of factors found in the residuals. In a two-factor model, we can put the dominant sectors in one or two groups using the clustering technique discussed previously. To do so, consider all three possible pairs of the sectors 'commercial and services machinery', 'metalworking machinery', and 'electrical equipment'. The sectors in a pair are put in different groups if they capture the effects of both underlying factors. Using this clustering technique, I find that all three sectors belong to a single group. This suggests one of the factors can be attributed to the idiosyncratic production growth in the identified sectors, while the other factor might have external sources. The latter sources might arise due to monetary policy affecting the demand for durables or the spillover effects from nonindustrial sectors like agriculture, public, financial and service sectors for which I do not have data.

6.2 Common Features Among the Dominant Sectors

I now look into whether there are common features among the dominant sectors in terms of their importance as suppliers of intermediate inputs and capital products to other sectors. To do so, I use the intersectoral material input-output table and the cross-sector capital use table from 1997 (referred to as $\mathrm{W^{IO}}$and $\mathrm{W^{CF}}$, respectively), and use network measures to assess whether the dominant sectors have relatively higher centrality in the data set. To assess the sectoral centrality, I use the outdegree and Katz-Bonacich measures.[14] In addition, I consider size of the sectors, which is computed using the sectoral shares in generating aggregate industrial production (s_i, for $i = 1,...,N$).

Looking at the centrality scores obtained from $\mathrm{W^{IO}}$ and $\mathrm{W^{CF}}$, I find that dominance of the sectors implied from the statistical analysis accords more with the patterns of the intersectoral capital flows. This is reflected in relatively higher ranking of the ADLs according to $\mathrm{W^{CF}}$. I observe that the sectors 'commercial and services machinery' and 'metalworking machinery' are ranked 1st and 2nd according to the outdegree centrality using the cross-sector capital use table (2nd and 1st according to the Katz-Bonacich centrality), while they show very moderate roles in terms of supplying materials to others (based on $\mathrm{W^{IO}}$, 'commercial and services machinery' is ranked 43rd (41st), and 'metalworking machinery' is ranked 51st (42nd) according to the outdegree (Katz-Bonacich) centrality using L4). A similar pattern holds for 'electrical equipment'. Looking at size of these sectors relative to others, it turns out that the largest sector among these three is 'commercial and services machinery' (ranked 11th), and coming after it, there appear 'electrical equipment' (36th), and 'metalworking machinery' (38th) out of 117 sectors. Overall, the results regarding the dominant sectors are more in line with the sectoral centrality implied from $\mathrm{W^{CF}}$, rather than those implied from $\mathrm{W^{IO}}$ and relative size of the sectors.

To further assess the relationship between the network centrality and the effect of a sector's growth rate on the rates of other sectors, I look at the number of significant coefficients in the regression models discussed in the following. In particular, all the sectoral rates are regressed on a constant, the growth rate of a sector whose higher outdegree centrality is implied from $\mathrm{W^{IO}}$ or $\mathrm{W^{CF}}$, and a cross-section average of all the rates which accounts for a common factor in the model. The number of significant coefficients on the sector's growth rate provides a simple benchmark on how significantly its production growth moves with the growth rates of others. Table (6) reports these numbers according to 90% confidence level.[15] Five most central sectors are considered according to

[14]Considering the interaction matrix $W = (w_{ij})$, the Katz-Bonacich centrality for j is measured as $c_j^{KB} = \lambda \sum_{i=1}^{N} w_{ij} c_i^{KB} + \eta$ with η being a base centrality which is assumed to be equal among all nodes, and λ is a parameter. I use 1 and 0.5 for η and λ, respectively.

[15]To avoid the multiple testing problem, the critical values are adjusted based on the Holm-Bonferroni correction. Let $t_{(1)}$,...,$t_{(N-1)}$ be the t-ratios in absolute value in descending order, which are obtained

Rank	Sector	The number of Significant Coefficients
Capital Use Table (W^{CF})		
1	Commercial and Service Machinery ⋆	26
2	Metalworking Machinery ⋆	22
3	Industrial Machinery	9
4	Computer and Peripheral Equipment	3
5	Navigational/Measuring/Electromedical/Control Instruments	5
Material Input-Output Table (W^{IO})		
1	Iron and Steel Products	7
2	Semiconductors and Other Electronic Components	4
3	Plastics Products	8
4	Electric Power Generation, Transmission and Distribution	4
5	Organic Chemicals	0
9 based on W^{CF} 25 based on W^{IO}	Electrical Equipment ⋆	21

Table 6: The central sectors and their corresponding number of significant coefficients are presented. The results are according to the 90% level confidence intervals.

each table, and I report the results only for 117 sectors from L4. The results are similar based on the other disaggregation level and remain to a great extent unchanged using the Katz-Bonacich centrality for ranking of the sectors.

An immediate feature emerging from Table (6) is that the dominant sectors which appear among the most important suppliers of capital products (presented in the upper panel of the table and the bottom row) have the largest number of significant coefficients. The coefficient of the growth rate of 'commercial and service machinery' and 'metalworking machinery' is significant in 26 and 22 of the regressions, respectively. This number is 21 for 'electrical equipment'. Furthermore, the results suggest that growth in production of the major capital products suppliers tend to significantly move with a larger number of other sectors than those having important roles as suppliers of input materials. When I repeat this analysis for 30 most central sectors, I obtain a mean value of 7.37 (5.2) and a median value of 4.5 (3.5) for number of the significant coefficients using W^{CF} (W^{IO}). Foerster et al. (2011) consider a class of the multisector general

from the $N-1$ regressions for a particular central sector. Also let $p^a_{(1)},...,p^a_{(N-1)}$ denote the corresponding adjusted critical values for the significance level a. The critical values are computed as $p^a_{(i)} = \Phi^{-1}(1 - \frac{a}{2[(N-1)-i]})$, where Φ^{-1} is inverse of the normal cumulative distribution function.

equilibrium models whose solution can be written as $x_t = \Theta x_{t-1} + R_0 \varepsilon_t + R_1 \varepsilon_{t-1}$, with Θ, R_0, and R_1 depending on the tables of cross-sector linkages, as well as, other underlying structural parameters of the model. Mostly these models abstract from capital or assume full depreciation of sector-specific capital within a period. The authors incorporate the cross-sector capital flows as another interaction channel. Looking at the cross-sector comovements and aggregate variability implied from the structural models, they show that incorporation of both intersectoral material and capital flows help better explain the key features observed in the data. Though, I do not impose the identifying restrictions related to the input-output tables, the results of this section go in line with their results. In particular, the results highlight the important role of the major capital good producers in the US production network (specially the heavy machinery industries) whose production disturbances broadly spread to others, though they might moderately contribute in the supply chains of material inputs (they are not among the major energy industries nor the major general purpose material producers like the iron, and still producers).

7 Conclusions

This paper looks into a detailed disaggregation of industrial sectors and considers the possibility that the sectoral production disturbances act as major sources of aggregate variability. The production growth rates of the sectors, which are dominant leaders in driving intersectoral comovement, act as macroeconomic factors. To identify such dominant sectors, I use the principal component method and the PS approach using a large N large T panel data set of sectoral growth rates. This identification approach starts with estimation of the number of unknown common factors and analyzes whether any of the variables in the data set spans the factors space. The consistency properties of this approach is analyzed here under different unknown factor structures which deviate from the case of strongly influential factors.

I provide evidence that production growth rates of a few sectors in the US, mainly the heavy machinery industries, provide suitable approximations for an unknown common factor over 1972 to 2007. Using the input-output tables in terms of the intersectoral material purchases and capital flows, I find that the sectors identified as dominant sectors have important roles as suppliers of capital products to the other sectors. I further show that a more central role in terms of supplying capital products explains comovement among sectoral growth rates more significantly that a higher central role as supplier of material inputs. These results shed more light on the empirical relevance of the cross-sector capital flows, next to the material input-outputs, in analyzing intersectoral complementarity.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Networks, shocks, and systemic risk. Technical report, National Bureau of Economic Research.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*(1), 135–171.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bailey, N., G. Kapetanios, and M. H. Pesaran (2015). Exponent of cross-sectional dependence: Estimation and inference. *Journal of Applied Econometrics*, (forthcoming).

Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics 132*(1), 169 – 194.

Bouakez, H., E. Cardia, and F. Ruge-Murcia (2014). Sectoral price rigidity and aggregate dynamics. *European Economic Review 65*, 1 – 22.

Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives 28*(4), 23–48.

Chudik, A. and M. H. Pesaran (2011). Infinite-dimensional VARs and factor models. *Journal of Econometrics 163*(1), 4 – 22.

Chudik, A. and M. H. Pesaran (2013). Econometric analysis of high dimensional VARs featuring a dominant unit. *Econometric Reviews 32*(5-6), 592–649.

Chudik, A., M. H. Pesaran, and E. Tosetti (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal 14*(1), 45–90.

Dupor, B. (1999). Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics 43*(2), 391 – 409.

Durlauf, S. N. (1993). Nonergodic economic growth. *The Review of Economic Studies 60*(2), 349–366.

Foerster, A. T., P.-D. G. Sarte, and M. W. Watson (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy 119*(1), 1–38.

Forni, M. and L. Reichlin (1998). Let's get real: a factor analytical approach to disaggregated business cycle dynamics. *The Review of Economic Studies 65*(3), 453–473.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Holly, S. and I. Petrella (2012). Factor demand linkages, technology shocks, and the business cycle. *Review of Economics and Statistics 94*(4), 948–963.

Horvath, M. (1998). Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics 1*(4), 781 – 808.

Jovanovic, B. (1987). Micro shocks and aggregate risk. *The Quarterly Journal of Economics 102*(2), 395–409.

Kapetanios, G. and M. Marcellino (2010). Factor-GMM estimation with large sets of possibly weak instruments. *Computational Statistics and Data Analysis 54*(11), 2655 – 2675. The Fifth Special Issue on Computational Econometrics.

Long, John B., J. and C. I. Plosser (1983). Real business cycles. *Journal of Political Economy 91*(1), 39–69.

Long, J. B. and C. I. Plosser (1987). Sectoral vs. aggregate shocks in the business cycle. *The American Economic Review 77*(2), 333–336.

Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics 168*(2), 244–258.

Parker, J. and D. Sul (2016). Identification of unknown common factors: Leaders and followers. *Journal of Business & Economic Statistics 34*(2), 227–239.

Shea, J. (2002). Complementarities and comovements. *Journal of Money, Credit and Banking 34*(2), 412–433.

Stock, J. H. and M. W. Watson (1998). Diffusion indexes. Technical report, National Bureau of Economic Research.

Appendix

I first describe the limiting behavior of V_{NT} and the rotation matrix H, which are defined in Theorem 1.

Lemma 1. *Let $\Lambda_N = N^{-\alpha}\Lambda$. Under Assumptions FM1-4, given that $\alpha \in [0,1/2)$ and $N^{2\alpha}/T \to K < \infty$, we have $||V_{NT}|| = O_p(N^{-2\alpha})$ and $||H|| = O_p(1)$.*

Proof: We have

$$||V_{NT}|| = \sqrt{\sum_{i=1}^{r} \rho_i^2(X'X/NT)} \leq \sqrt{r}\rho_1(X'X/NT).$$

Thus, it suffices to look at $\rho_1(X'X/NT)$. Consider the following equation

$$\frac{X'X}{NT} = \frac{\Lambda_N F'F\Lambda_N'}{NT} + \frac{\Lambda_N F'e + e'F\Lambda_N}{NT} + \frac{e'e}{NT}.$$

The Weyl's eigenvalue inequality for Hermitian matrices implies

$$\rho_1(\frac{X'X}{NT}) \leq \rho_1(\frac{\Lambda_N F'F\Lambda_N'}{NT}) + \rho_1(\frac{\Lambda_N F'e + e'F\Lambda_N}{NT}) + \rho_1(\frac{e'e}{NT}).$$

Bai and Ng (2002) show $\rho_1(e'e/NT) = O_p(C_{NT}^{-2})$. Consider the first term

$$\frac{1}{NT}\rho_1(\Lambda_N F'F\Lambda_N') = \frac{1}{NT}\left(||F\Lambda_N'||_2\right)^2 \leq \frac{1}{NT}\left(||F||_2 \cdot ||\Lambda_N||_2\right)^2$$

$$= \rho_1(\frac{F'F}{T})\rho_1(\frac{\Lambda_N'\Lambda_N}{N}).$$

This shows the first term is $O_p(N^{-2\alpha})$ according to Assumptions FM1-2 since $\rho_1(\Lambda_N'\Lambda_N/N) = \rho_1(\Lambda'\Lambda/N^{1-2\alpha}) = O_p(N^{-2\alpha})$.

We now turn to the second term. Using the spectral radius of the square matrix in the second term as a lower bound for the Frobenius norm, we have

$$\begin{aligned}\rho_1(\frac{\Lambda_N F'e + e'F\Lambda_N}{NT}) &\leq ||\frac{\Lambda_N F'e + e'F\Lambda_N}{NT}|| \\ &\leq ||\frac{\Lambda_N F'e}{NT}|| + ||\frac{e'F\Lambda_N}{NT}||.\end{aligned}$$

For $||\Lambda_N F'e/NT||$, we have

$$||\frac{\Lambda_N F'e}{NT}|| = \frac{1}{\sqrt{NT}}\sqrt{tr\left[e'F(\frac{\Lambda_N'\Lambda_N}{N})F'e\right]} = O_p(N^{-\alpha}T^{-\frac{1}{2}})\sqrt{\frac{1}{NT}tr(e'FF'e)},$$

where $tr(e'FF'e)/NT = ||\sum_{t=1}^{T} f_t e_t'||^2/NT = N^{-1}\sum_{i=1}^{N}||T^{-1/2}\sum_{t=1}^{T} f_t e_{it}||^2 = O_p(1)$ according to Assumption FM4-i. Thus, we have $||\Lambda_N F'e/NT|| = O_p(N^{-\alpha}T^{-\frac{1}{2}})$. The same result holds for $||e'F\Lambda_N/NT||$. These results show

$$||V_{NT}|| = O_p(N^{-2\alpha}) + O_p(N^{-\alpha}T^{-\frac{1}{2}}) + O_p(C_{NT}^{-2}).$$

Note the term $O_p(N^{-2\alpha})$ would dominate if $\alpha \in [0, 1/2)$ *and* $N^{2\alpha}/T \to K < \infty$. Furthermore, consider H as follows

$$\begin{aligned}||H|| &= ||(\Lambda_N'\Lambda_N/N)(F'\tilde{F}/T)V_{NT}^{-1}|| \\ &\leq ||(\Lambda_N'\Lambda_N/N)||\cdot||(F'F/T)||^{\frac{1}{2}}\cdot||(\tilde{F}'\tilde{F}/T)||^{\frac{1}{2}}\cdot||V_{NT}^{-1}||.\end{aligned}$$

Assumptions FM1-2, together with $\Lambda_N = \Lambda/N^{\alpha}$ and the normalization $\tilde{F}'\tilde{F}/T = I_r$ imply

$$||H|| = O_p(N^{-2\alpha})||V_{NT}^{-1}|| = O_p(1).$$

This completes the proof. *Q.E.D.*

Lemma 1 shows the contribution of the common component in overall variability of X would dominate that from the idiosyncratic shocks under $\alpha \in [0, 1/2)$ and $N^{2\alpha}/T \to K < \infty$.

Kapetanios and Marcellino (2010) showed consistency of the principal component method for $\hat{F} = X\tilde{\Lambda}_N/N^{1-2\alpha}$, where $\tilde{\Lambda}_N$ is the principal component estimate of loadings. The following lemma deduces the same result for $\tilde{F}$.

Lemma 2. *Let $\Lambda_N = N^{-\alpha}\Lambda$. Under Assumptions FM1-4, we have*

$$\frac{C_{NT}^2}{N^{4\alpha}}\left(\frac{1}{T}\sum_{t=1}^{T}||\tilde{f}_t - H'f_t||^2\right) \quad = \quad O_p(1). \tag{20}$$

given that $\alpha \in [0,1/2)$ and $N^{2\alpha}/T \to K < \infty$.

Proof: Substituting for $\tilde{\Lambda}_N = X'\tilde{F}/T$ in the formula for $\hat{F}$, we obtain $\hat{F} = XX'\tilde{F}/N^{1-2\alpha}T$. Also from the principal component analysis we know $XX'\tilde{F}/NT = \tilde{F}V_{NT}$. This implies $\tilde{F} = N^{-2\alpha}\hat{F}V_{NT}^{-1}$. Thus, we can write

$$\tilde{f}_t - H'f_t = \left(N^{-2\alpha}V_{NT}^{-1}\right)\left[\hat{f}_t - (\frac{\tilde{F}'F}{T})(\frac{\Lambda_N'\Lambda_N}{N^{1-2\alpha}})f_t\right].$$

Lemma 2 follows from Lemma 1, and Theorem 4 from Kapetanios and Marcellino (2010), which says

$$\frac{C_{NT}^2}{N^{4\alpha}}\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{f}_t - (\frac{\tilde{F}'F}{T})(\frac{\Lambda_N'\Lambda_N}{N^{1-2\alpha}})f_t||^2\right) = O_p(1).$$

Q.E.D.

We can now provide the proof for Theorem 1, which builds on Lemma A.2 and Theorem 1 from Bai (2003).

Proof of Theorem 1. We start with the following identity

$$\begin{aligned}\tilde{f}_t - H'f_t \quad &= \quad V_{NT}^{-1}\left[\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_s\gamma_N(s,t) + \frac{1}{T}\sum_{s=1}^{T}\tilde{f}_s\zeta_{st}\right.\\ &+ \quad \left.\frac{1}{T}\sum_{s=1}^{T}\tilde{f}_s\eta_{st} + \frac{1}{T}\sum_{s=1}^{T}\tilde{f}_s\xi_{st}\right]\end{aligned} \tag{21}$$

where
$\gamma_N(s,t) = E(e_s'e_t/N)$,
$\zeta_{st} = e_s'e_t/N - \gamma_N(s,t)$,
$\eta_{st} = f_s'\Lambda_N'e_t/N$,
$\xi_{st} = f_t'\Lambda_N'e_s/N$.

Equation (21) can be written as

$$
\begin{aligned}
\tilde{f}_t - H'f_t &= V_{NT}^{-1}\left[\frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\gamma_N(s,t) + \frac{H'}{T}\sum_{s=1}^{T} f_s\gamma_N(s,t)\right. \\
&+ \frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\zeta_{st} + \frac{H'}{T}\sum_{s=1}^{T} f_s\zeta_{st} \\
&+ \frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\eta_{st} + \frac{H'}{T}\sum_{s=1}^{T} f_s\eta_{st} \\
&+ \left.\frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\xi_{st} + \frac{H'}{T}\sum_{s=1}^{T} f_s\xi_{st}\right] \\
&= V_{NT}^{-1}[I + II + III + IV + V + VI + VII + VIII].
\end{aligned}
$$

Notice the main difference here in comparison to the setting from Bai (2003) is that $1/T\sum_{s=1}^{T}||\tilde{f}_s - H'f_s||^2$ is $O_p(N^{4\alpha}C_{NT}^{-2})$ instead of $O_p(C_{NT}^{-2})$. In addition, η_{st} and ξ_{st} contain $\Lambda_N = \Lambda/N^{\alpha}$ as opposed to Λ. From Lemma 1, we know $||H||= O_p(1)$, which implies the limiting behavior of *II* and *IV* do not change here relative to those presented by Bai (2003). Thus, we have $II = O_p(1/T)$ and $IV = O_p(1/\sqrt{NT})$.

Now, we continue by looking into *I*. We have

$$
||\frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\gamma_N(s,t)|| \leq \left(\frac{1}{T}\sum_{s=1}^{T}||\tilde{f}_s - H'f_s||^2\right)^{\frac{1}{2}}\left[\frac{1}{\sqrt{T}}\left(\sum_{s=1}^{T}\gamma_N^2(s,t)\right)^{\frac{1}{2}}\right].
$$

Assumption FM3-ii implies that $\left(\sum_{s=1}^{T}\gamma_N^2(s,t)\right)^{\frac{1}{2}}$ is bounded from above. Hence, $I = O_p(T^{-1/2}N^{2\alpha}C_{NT}^{-1})$ following Lemma 2. For *III*, in a similar way we can write

$$
||\frac{1}{T}\sum_{s=1}^{T}(\tilde{f}_s - H'f_s)\zeta_{st}|| \leq \left(\frac{1}{T}\sum_{s=1}^{T}||\tilde{f}_s - H'f_s||^2\right)^{\frac{1}{2}}\left(\frac{1}{T}\sum_{s=1}^{T}\zeta_{st}^2\right)^{\frac{1}{2}}
$$

$$= O_p(N^{2\alpha}C_{NT}^{-1}) \left(\frac{1}{NT} \sum_{s=1}^{T} \left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (e_{is}e_{it} - E(e_{is}e_{it})) \right]^2 \right)^{\frac{1}{2}}.$$

We have $N^{-1/2}\sum_{i=1}^{N}(e_{is}e_{it} - E(e_{is}e_{it})) = O_p(1)$ from Assumption FM3-iii, and in turn it is implied that $III = O_p(N^{2\alpha-1/2}C_{NT}^{-1})$.
Looking at VI and $VIII$, we see their limiting behavior changes by a factor $N^{-\alpha}$ in comparison to the case that Λ is the loading matrix. Using the rates obtained by Bai (2003) which are $O_p(N^{-1/2})$ and $O_p(N^{-1/2}T^{-1/2})$ respectively, it follows that $VI = O_p(N^{-\alpha-1/2})$ and $VIII = O_p(N^{-\alpha-1/2}T^{-1/2})$.
We now consider V and VII. Bai (2003) shows they both are $O_p(N^{-1/2}C_{NT}^{-1})$. Following our argument about the changes in the rates due to the different convergence rate for $\tilde{F}$ ($N^{2\alpha}C_{NT}^{-1}$ replaces C_{NT}^{-1}), as well as, use of Λ/N^{α} which brings in a factor of $N^{-\alpha}$, we imply that $V, VII = O_p(N^{-1/2}(N^{2\alpha}C_{NT}^{-1})N^{-\alpha}) = O_P(N^{\alpha-1/2}C_{NT}^{-1})$.

Using $V_{NT}^{-1} = O_p(N^{2\alpha})$ from Lemma 1, we finally obtain

$$\tilde{f}_t - H'f_t = O_p(\frac{N^{4\alpha}}{C_{NT}^2}) + O_p(N^{\alpha-1/2}). \tag{22}$$

which completes the proof. The points (i) and (ii) in Corollary 1 are just implied from (22). *Q.E.D.*

Having the results from Theorem 1, we can finally provide the proof for Theorem 2.

Proof of Theorem 2. Consider the least squares residuals using $\tilde{f}_t$

$$\hat{e}_{it} = e_{it} - \left(\hat{\lambda}'_{N,i} - \lambda'_{N,i}H'^{-1}\right)\tilde{f}_t - \lambda'_{N,i}H'^{-1}\left(\tilde{f}_t - H'f_t\right).$$

From the properties of the least squares method, we know $\hat{\lambda}'_{N,i} - \lambda'_{N,i}H'^{-1} = O_p(T^{-1/2})$. Furthermore, Lemma 1 and Theorem 1 imply

$$\lambda_{N,i}' H'^{-1}\left(\tilde{f}_t - H' f_t\right) = O_P(N^{-\alpha})\left[O_p(\frac{N^{4\alpha}}{C_{NT}^2}) + O_p(N^{\alpha-1/2})\right]$$
$$= O_p(\frac{N^{3\alpha}}{C_{NT}^2}) + O_p(N^{-1/2}).$$

Thus, we obtain

$$\hat{e}_{it} - e_{it} = O_p(T^{-1/2}) + O_p(\frac{N^{3\alpha}}{C_{NT}^2}) + O_p(N^{-1/2}).$$

PS show when $\hat{e}_{it} - e_{it} = O_p(C_{NT}^{-1})$, the probability of detecting zero factors in $\hat{e}_t$ approaches one using the criteria of Bai and Ng (2002) when $N, T \to \infty$. To ensure this, we need $\alpha \in [0, 1/6]$ and $N/T \to K < \infty$.
These results ensure the argument of PS holds even if factors are not strong as long as α, N, and T satisfy the conditions mentioned. Then, the theorem simply follows Theorem 1 of PS, that completes the proof. *Q.E.D.*