A Dynamic Factor Analysis of Firm Growth

Lucia Alessi*
Matteo Barigozzi†
Marco Capasso‡

Laboratory of Economics and Management
Sant'Anna School for Advanced Studies, Pisa

Preliminary version: July 2006

Abstract

We use the Generalized Dynamic Factor Model proposed by Forni et al. [2000] in order to study the dynamics of the rate of growth of sales and investment. By using quarterly firm level data relative to hundreds of US firms from 1985 to 2005, we disentangle the component of industrial dynamics which is due to economy-wide shocks from the component which relates to sectoral or even firm-specific phenomena. The analysis of the dynamics of these components gives insights on the nature of cyclical fluctuations, on the evolution of phenomena in the short, medium and long run, as well as on their features at business cycle frequencies. Finally, we study the patterns of comovement of the common component of the series.

Keywords: dynamic factor analysis, business cycle, comovements, technology
JEL-classification: C51, E32, O30

* *E-mail*: l.alessi@sssup.it
† *E-mail*: m.barigozzi@sssup.it
‡ *E-mail*: m.capasso@sssup.it

We thank Giovanni Dosi and Marco Lippi for helpful comments and suggestions. We are also grateful to Marc Hallin and Roman Liska for providing us with their codes. Of course, the responsibility for any error is entirely our own.

1 Introduction

A variety of important issues on industrial dynamics can be addressed by means of factor analysis. Because of their own nature, factor models focus on the few driving forces underlying the evolution of industries, sectors and firms. Indeed, each economic series of interest is represented as the sum of a common component, driven by the non observable economy-wide factors, and an idiosyncratic component, which is peculiar to each series and vanishes via aggregation. Moreover, this kind of model allows using large cross-sections of time series - in principle, an infinite number of series could be used - thereby taking into account as much information as possible. Finally, the whole dataset can be broken up into smaller subsamples - say, sectoral or country-specific - in order to carry on a factor analysis on each of the subsamples and compare the results across different subsets. This work gives a contribution in this sense and, more generally, goes in the direction of highlighting "which statistical properties and relations are invariant across industries, and, conversely, which ones depend on the technological and market characteristics of particular sectors" (Dosi [2005]).
Whatever the focus of the analysis, it is fundamental to understand at which level phenomena occur, i.e. whether they are common to all firms belonging to the same country or the same region, or they are sectoral or even market- or firm-specific. The following brief review of the literature is intended to give some examples of issues for which the distinction between a common feature or shock and an idiosyncratic component is crucial. Cooper and Haltiwanger [1990] analyze the role of inventories in the comovements of employment and output across sectors. Davis and Haltiwanger [1992] and Davis and Haltiwanger [1999] study job reallocation at the plant level and draw conclusions on the cyclical behavior of the labor market. In investigating the driving forces of employment fluctuations they distinguish between allocative and aggregate shocks. Haltiwanger [1997] measures the importance of shocks at the 4-digit industry level as less than 10% of the whole variation across sectors of output, employment, capital equipment and capital structures. Foster et al. [1998] analyze the within-plant contribution to the overall productivity growth for manufacturing and their sensitivity to the business cycle. Recently, Scarpetta et al. [2004] compare the process of creative destruction across twentyfour countries and two-digit industries and investigate the effect of resource reallocation on productivity. Bartelsman et al. [2005b] compare ten OECD countries and identify idiosyncratic country effect in the firm size distribution and in the entry-exit dynamic. Finally, Bartelsman et al. [2005a] underline how the presence of idiosyncratic components in cross-country panels may pose severe problems when standard techniques are applied.

The novelty of our contribution stands in the use of the Generalized Dynamic Factor Model (GDFM) by Forni et al. [2000] to investigate industrial dynamic issues. Its power stands in the fact that it allows a truly dynamic analysis of a large number of series: by reducing the dimension of the problem it combines the virtues of panel data techniques and VAR models. On this, Reichlin [2002] states:

> "Modern macroeconomic theory is based on the representative agent assumption, but macroeconomic empirics is mostly based on aggregate data. What is the cost of simplicity, i.e. are we losing valuable information by working with econometrics models containing few aggregate variables? How detailed do our models have to be to have a chance to provide the essential information on the macroeconomy? To try to answer these questions there is a need to develop econometric models which (a) are able to handle the analysis of many time series by reducing the number

of the essential parameters to estimate; (b) can provide an answer on what is the relevant stochastic dimension of a large economy, i.e. on how many aggregate shocks are needed to study the macroeconomy which emerges from the behavior of many agents; (c) can help us identifying these (possibly few) shocks and studying the propagation mechanism through agents or through geographical space. This is what will help to bridge the gap between purely time series studies and the cross-sectional approach."

Indeed, the dynamic factor approach presents basically three big advantages with respect to the traditional VAR models (see Forni et al. [2006a]):

- it allows dealing with much larger datasets, virtually infinite, i.e. a wider information set, virtually the whole economy;
- it doesn't need to impose any restriction in order to disentangle the common part and the idiosyncratic part of each series in the sample;
- it needs only a few identification restrictions for the structural shocks, being their number very small and equal to the number of underlying common factors, while in VAR models the number of structural shocks equals the number of series in the dataset.

Ultimately, dynamic factor models allow for an easier economic interpretation. Moreover, they don't suffer from two shortcomings of the VAR approach, i.e. the sensitivity to the choice of the variables to include in the analysis and the substantial arbitrarity of the identification restrictions. For these reasons we do believe that the dynamic factor approach better than others answers to the need of an empirically grounded economic research.
This paper is in the spirit of the work by Forni and Reichlin [1998], where they apply a dynamic factor model *à la* Sargent and Sims [1977] for studying the dynamics of output and productivity in US manufacturing. As in this latter paper, we basically investigate the following issues. Firstly, we want to assess the stochastic dimension of firm growth. In other words, we want to determine the number of fundamental forces driving this economic phenomenon, i.e. the number of non observable factors common to the whole economy which lead the process of firm growth and output fluctuations, from a microeconomic and macroeconomic point of view respectively. Secondly, we measure the relative importance of the common component with respect to the idiosyncratic component by means of a straightforward variance decomposition. By computing variance ratios both for the entire sample and for sectoral subsamples we are able to identify three levels of analysis, i.e. economy-wide, sectoral- and firm-specific.
However, unlike Forni and Reichlin [1998], we do not limit our study to the manufacturing sector, being our wider samples composed by 660 and 355 series, for sales and investments respectively, belonging to all sectors of the US economy. Moreover, we have the use of about 20 years of quarterly data, from 1984 to 2005. Besides, the most important difference is in the model itself: we choose to apply the GDFM, which allows for a limited amount of cross-sectional correlation across idiosyncratic components. Indeed, we believe that the assumption of mutual orthogonality of idiosyncratic components might be too restrictive when dealing with balance-sheet data relative to sales and investments. Finally, we improve on the heuristic procedure for determining the number of common factors by implementing the information criterion recently proposed by Hallin and Liska [2005], which indicates the minimum number of common dynamic factors to include in the model.
The dynamic factor approach allows to tackle a variety of issues all at once. For example, we obtain information with regards to the features of positive and negative comovements in

the economy. This kind of analysis, when run at the sectoral level, may give interesting contribution to the debate on reallocation effects, checking which sectors comove positively and which negatively, and at which frequencies. Indeed, the theoretical model by Mortensen and Pissarides [1999] and the empirical studies by Davis et al. [1998] and Davis and Haltiwanger [1999] narrowly link the issues of job reallocation and nature of cyclical economic fluctuations. Telling whether sectoral shocks may reflect into structural fluctuations is crucial, see for example the sectoral shifts hypothesis by Lilien [1982], the subsequent critique by Abraham and Katz [1986] and the stream of literature which builds on Lilien, e.g. the works by Loungani and Rogerson [1989], Rissman [1997], and Figura [2002]. Alternative models in which sectoral shocks may generate aggregate fluctuations have been proposed by Long and Plosser [1983] in their classic real business cycle model and by Cooper and Haltiwanger [1990], among others. In this paper, we address the problem of whether sectoral shocks may produce fluctuations at business cycle frequencies for the whole economy, and analogously whether firm-specific shocks may produce business-cycle effects at the sectoral level, by studying separately the spectra of the common component and the idiosyncratic component of the series.
The paper is structured as follows. In the following section we outline the GDFM and the estimation procedure. In section 3 we present the COMPUSTAT data and the deflators we use for building the dataset for the empirical analysis. For determining the number of factors to include in the model, in section 4 we apply on all our samples both the criterion by Hallin and Liska and a heuristic procedure. In section 5 we retrieve the common and the idiosyncratic component of each series and study their spectral profile and their cross-spectra. Section 6 concludes and proposes complementary analyses and empirical applications.

2 The Model

We denote as $x_t = (x_{1t} \dots x_{Nt})'$ an N-dimensional vector process. Each of the series is stationary and second order moments $\gamma_{ik} = E[x_{it}x'_{it-k}]$ exist finite for all i and k. In the Generalized Dynamic Factor Model (GDFM), as proposed by Forni et al. [2006b], it is assumed that each series x_{it} can be written as the sum of two mutually orthogonal unobservable components, the *common component* χ_{it} and the *idiosyncratic component* ξ_{it}. The common component is driven by a small number q of dynamic common factors u_{jt} with $j = 1, \dots, q$, which are loaded with possibly different coefficients and lags. Formally:

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \dots + b_{iq}(L)u_{qt} + \xi_{it} \qquad i = 1, \dots, N \tag{1}$$

The q-dimensional vector process $u_t = (u_{1t} \dots u_{qt})'$ is an orthonormal white noise. The N-dimensional vector process $\xi_t = (\xi_{1t} \dots \xi_{Nt})'$ has zero mean and is stationary. Moreover, ξ_{it} is orthogonal to u_{jt-k} for all k, i and j. The polynomials in the lag operator $b_{i1}(L) \dots b_{iq}(L)$ are square-summable, two-sided filters in principle of infinte order.
In order to move to the frequency domain and to have consistent estimates of the spectral densities, we need to assume that the process x_t admits a Wold representation $x_t = \sum_{k=0}^{+\infty} C_k w_{t-k}$ where innovations have finite fourth order moment and the entries of the matrices C_k satisfy $\sum_{k=0}^{+\infty} |C_{ij,k}|k^{1/2}$. We denote the spectral density matrices of the common part and the idiosyncratic part respectively as $\Sigma^{\chi}(\theta)$ and $\Sigma^{\xi}(\theta)$, with $\theta \in [-\pi, \pi]$. Since we are interested in analyzing only business cycle frequencies, we evaluate the spectral densities only for $\theta \in [-\pi/2, \pi/2]$ and impose that they are zero outside this interval, thereby discarding frequencies corresponding to periods shorter than one year. Finally, we assume that the q largest eigenvalues of $\Sigma^{\chi}(\theta)$ diverge almost everywhere as the number of series goes to infinity,

while all the eigenvalues of $\Sigma^{\xi}(\theta)$ are bounded. This last condition, in other words, relaxes the assumption of mutual orthogonality of idiosyncratic components by allowing for a limited amount of cross-sectional correlation.

The estimation of the model follows the procedure proposed in Forni et al. [2006b]. Firstly, the spectral density matrix of x_t, $\Sigma^x(\theta)$, is estimated by applying the Fourier transform to the sample covariance matrices $\hat{\Gamma}_k$. Then the dynamic principal component decomposition is applied, thereby selecting the first q largest eigenvalues of $\hat{\Sigma}^x(\theta)$ and the corresponding eigenvectors $(p_1(\theta), \ldots, p_q(\theta))$. From such eigenvectors we can build the corresponding filters using the inverse Fourier transform:

$$\tilde{p}_j(L) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \left[\int_{-\pi}^{\pi} p_j(\theta) e^{i\theta k} \mathrm{d}\theta \right] L^k \tag{2}$$

The factor space is the minimal closed subspace that contains the q largest dynamic principal components: $U = \overline{\text{span}} \{\tilde{p}_j(L)x_t \ , \ j = 1 \ldots q\}$; and we know that the common part is just the projection of x_t onto the factor space. Given the orthonormalithy of the dynamic eigenvectors the following holds:

$$x_t = \left[\sum_{j=1}^{n} \tilde{p}_j^{\dagger}(L) \tilde{p}_j(L) \right] x_t$$

where with $p^{\dagger}$ we mean the transposed and conjugate of p. Finally, since by assumption the common and the idiosyncratic part are orthogonal we have:

$$\hat{\chi}_t = \left[\sum_{j=1}^{q} \tilde{p}_j^{\dagger}(L) \tilde{p}_j(L) \right] x_t = K(L) x_t \tag{3}$$

We the obtain the idiosyncratic component simply as difference between the original series x_t and $\hat{\chi}_t$.
Note that the filter $K(L)$ is two-sided and in principle is also of infinte order, but for $t \leq 0$ and $t \geq T$ we don't have observations for x_t, hence we need to truncate the filter. Such operation will cause the loss of part of the variance of χ_t even for $N, T \to \infty$ so we must concentrate only to the central part of x_t. In practice this implies the choice of a truncation lag M_T such that consistency of all the estimates is preserved[1]; this is ensured provided that $M_T \to \infty$ and $M_T/T \to 0$ as $T \to \infty$ (for instance, we choose $M_T \sim [\sqrt{T}/2]$).
Thus the estimated filter that in practice we use is:

$$\hat{K}(L) = \frac{1}{2(2M_T + 1)} \sum_{k=-M_T}^{M_T} \left[\sum_{h=-M_T}^{M_T} \left(\sum_{j=1}^{q} p_j^{\dagger}(\theta_h) p_j(\theta_h) \right) e^{ik\theta_h} \right] L^K \tag{4}$$

where $\theta_h = \pi h / 2(2M_T + 1)$ are the $2M_T + 1$ points for which the spectral density is estimated[2].

[1]See Forni et al. [2000] for a detailed discussion of the problem. Moreover, estimations of the common component at the beginning and at the end of the sample are not reliable, thus such method is useless for prediction while it is good for structural analysis as it is in our case.

[2]This in turn comes from the estimated covariances $\hat{\Gamma}_k$ for $k = -M_T \ldots M_T$. Note that if we take the whole interval $[-\pi, \pi]$ instead of keeping just a smaller frequency band, the fractional term in (4) becomes $\frac{1}{2M_T+1}$ and the expression for θ_h becomes $\theta_h = 2\pi h/(2M_T + 1)$. In our specific case, half of the $2(2M_T + 1)$ points are out of the relevant interval $[-\pi/2, \pi/2]$.

A number of studies compare the GDFM with the dynamic factor model *à la* Stock and Watson [2002][3]. The key difference between the GDFM dynamic approach and the static principal component method used by Stock and Watson is that the latter gives a static representation of the dynamic model. In other words, while the first exploits the information contained in lagged covariance matrices, the latter makes use of contemporaneous covariances only. While this fact may turn helpful in those contexts where computational simplicity plays a role, when we come to structural analysis it becomes crucial to tell what the real dynamic factors are, rather than what their static representation is. In a nutshell, structural analysis would be pointless if we would treat one shock and its lags as many different shocks, as it is done in the estimation via static principal components.

3 The Data

For the empirical analysis we use COMPUSTAT quarterly data relative to sales and investments of US firms, by means of which we intend to proxy firm size. The whole COMPUSTAT database contains data on about 10,000 actively traded U.S. companies, standardized by specific data item definition and by financial statement in order to allow for intertemporal and inter-firm comparison. Given the large amount of data at our disposal, we can exclude from the analysis all those series presenting problems concerning outliers or missing values in the period under analysis, so that the database we use is not affected by any arbitrary manipulation. What follows is the COMPUSTAT definition of sales and investments:

- Sales: this item represents gross sales (the amount of actual billings to customers for regular sales completed during the period) reduced by cash discounts, trade discounts, and returned sales and allowances for which credit is given to customers. The result is the amount of money received from the normal operations of the business (i.e., those expected to generate revenue for the life of the company).

- Capital expenditures: this item represents cash outflow or funds used for additions to the company's property, plant, and equipment, excluding amounts arising from acquisitions.

We don't consider those firms which have been affected at any point in the considered time span by any mergers or acquisitions or any kind of accounting changes, as well as those data which present discrepancies with respect to the standard definition (see Appendix A for a detailed description of how the data are built from balance-sheet items and a list of possible anomalies). The sales and investments time series surviving the above selection have been grouped into two distinct datasets: the dataset for sales goes from the second quarter 1985 to the first quarter 2005 - 80 observations in total - and includes 57 firms; the dataset for investments goes from 1984 first quarter to 2005 first quarter - 85 observations in total - and includes 355 firms. The difference in the number of firms in the two samples is due to the fact that often while the investments time series is long enough, the sales time series is too short for the same firm, so we had to drop those firms from the sales sample. However, since the cross dimension plays a crucial role for the consistency results of the dynamic factor model we use, we run the analysis on a wider sales dataset as well, including 660 series. In this latter dataset those firms experiencing significant mergers or acquisitions, whereby the effects on the prior year's sales constitute 50% or more of the reported sales for that year, are not included.

[3]See for example Boivin and Ng [2005] and the works by Marcellino, although in a forecasting context.

Running the analysis on two sales dataset, i.e. the cleanest possible and the widest possible, ensures the accuracy of the results because both the quality and the quantity of the data are taken into account.
We are interested in inter-sectoral comparisons too. Therefore, we break up the two big samples into sectoral subsamples and apply the GDFM to the more numerous subsets. For the definition of sectors we follow the North American Industrial Classification System (NAICS), introduced in 1997, which is more disaggregated than the Standard Industrial Classification (SIC) and better designed. Moreover, it introduces detail on key sectors belonging to services and IT[4]. We focus on the following three sectors, identified on the basis of the three-digit NAICS code:

- Computer and electronic product manufacturing (NAICS code: 334—): 89 firms in the sales subsample, investments subsample too small
- Chemical manufacturing (NAICS code: 325—): 81 firms in the sales subsample, 37 firms in the investments subsample
- Machinery manufacturing (NAICS code: 333—): 61 firms in the sales subsample, investment subsample too small

These raw data have been deflated by means of the deflator series published by the U.S. Department of Labor - Bureau of Labor Statistics (BLS). For deflating capital expenditure series we used the Producer Price Index Finished Goods - Capital Equipment with base year 1982 (figure 1).

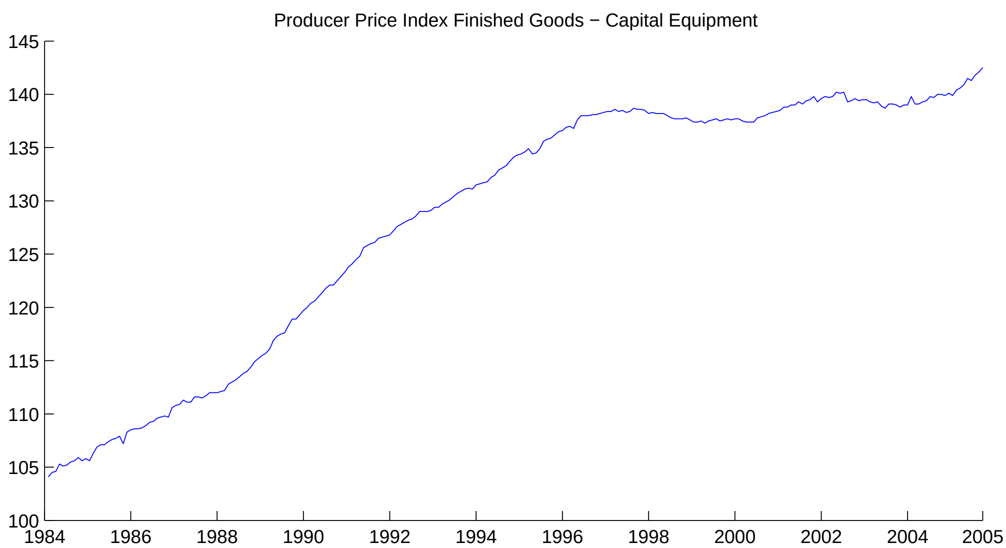


Figure 1: Producer Price Index Finished Goods - Capital Equipment.

For deflating sales series, we use the Producer Price Index Revision-Current Series, which are NAICS-based sectoral deflators. According to the definition, they "reflect price movements for the net output of producers [...]. To the extent possible, prices used in constructing the indexes

[4]Those firms belonging to some very recent niches, as the "Dot-com" enterprises in the computer manufacturing sector, are absent in our dataset.

are the actual revenue or net transaction prices producers receive for sales of their outputs [...]. The PPI is meant to measure changes in prices received by domestic producers, import products are not priced in the survey.". We have been able to match firms and deflators with a good deal of accuracy, being the great majority of the codes at the six-digit level. In some cases, for instance in the case of big enterprises active in more than one sector or when the six-digit NAICS code deflator was not available, we took deflators relative to wider industry definitions, keeping the three-digit level as the maximum limit for aggregation.
Series have been seasonally adjusted by a simple moving average method, again aiming at manipulating the raw data as less as possible (see Appendix B for details).
Finally, series have been differenced in order to get stationarity, de-meaned and standardized, so that we work with rates of growth ultimately.

4 Determining the Number of Factors

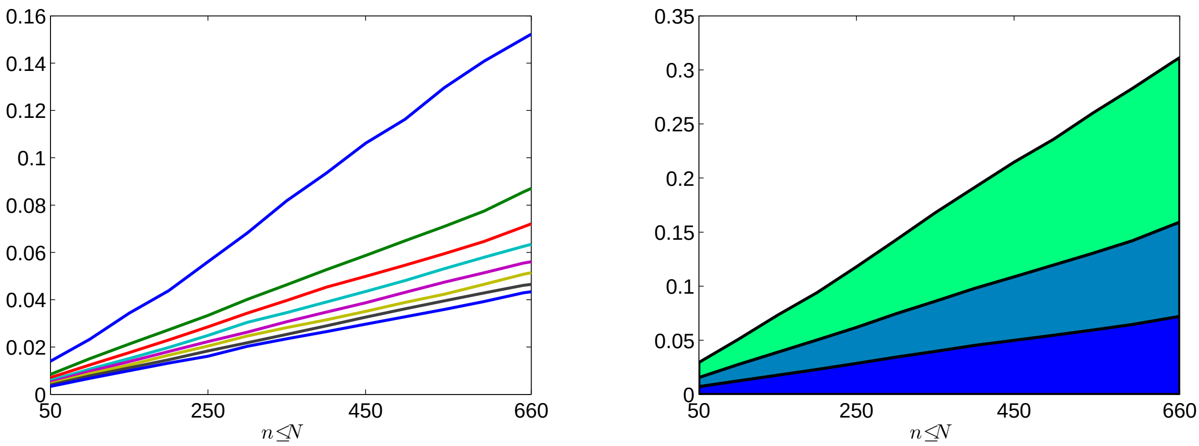
Firstly, we verify that our dataset does fulfill GDFM assumptions on the eigenvalues $\lambda_i(\theta)$ of the spectral density matrix of x_t. According to Brillinger [1981], we define the variance explained by the i^{th} factor as:

$$EV_i = \frac{\int_{-\pi/2}^{\pi/2} \lambda_i(\theta) d\theta}{\sum_{j=1}^{N} \int_{-\pi/2}^{\pi/2} \lambda_j(\theta) d\theta} \tag{5}$$

We require that, as $N \longrightarrow \infty$:

$$\begin{cases} EV_i \longrightarrow \infty & \text{for} \quad i = 1, \ldots, q \\ EV_i \leq M & \text{for} \quad M \in \mathbb{R}^+ \quad i = 1, \ldots, q \end{cases} \tag{6}$$

Indeed, as shown figure (2) for the 660 series sales sample, this is the case. The same figure shows the cumulated explained variance relative to the first eigenvalues for the same sample.

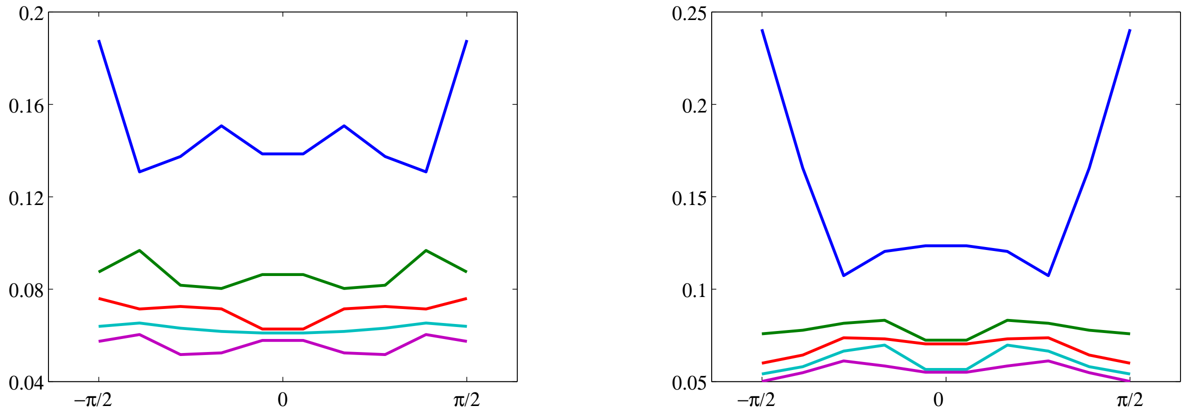


(a) Explained variance

(b) Cumulated explained variance

Figure 2: Variance explained by each dynamic common factor and cumulated explained variance for the 660 series sales sample.

We verify that the first eigenvalue is actually the greatest at all frequencies, although it would be enough the first eigenvalue to be the highest almost everywhere, and analogously that the other eigenvalues up to the q^{th} are at all frequencies respectively the second-highest, the third-highest, and so on. The q eigenvalues have been computed in the $[-\pi/2, +\pi/2]$ frequency band. Indeed, as shown in figure (3) relative to the two big samples for sales and investments, the first eigenvalue takes the higher value at all frequencies, the second eigenvalue takes the second-highest value at all frequencies, and so on. In other words, eigenvalues never cross each other. It is also interesting to note that the first eigenvalue is actually much larger than all the others, this suggesting the presence of just one common dynamic factor.



(a) Sales, 660 series

(b) Investments, 355 series

Figure 3: Highest eigenvalues (frequencies on the horizontal axis).

However, we do not rely only on the intuition coming from the graphs for determining the number of factors to include in the model. We implement two complementary procedures: we apply on the one hand the Hallin and Liska [2005] information criterion for determinig the minimum number of common factors, and on the other hand a heuristic procedure still based on the variance explained by the eigenvalues.
The criterion by Hallin and Liska exploits the relation in the GDFM between the number of common factors and the number of diverging eigenvalues of the spectral density matrix of the observations. We choose the logarithmic form[5] of the covariogram-smoothing version of the criterion. For given N and T, it consists in choosing the number of factors $\hat{q}_N^T$ so to minimize the following

$$IC_N^T(q) = log\left[\frac{1}{N}\sum_{i=q+1}^{N}\frac{1}{2M_T+1}\sum_{h=-M_T}^{M_T}\lambda_{Ni}^T(\theta_h)\right] + qp(N,T), \qquad 0 \le q \le q_{max} < \infty \tag{7}$$

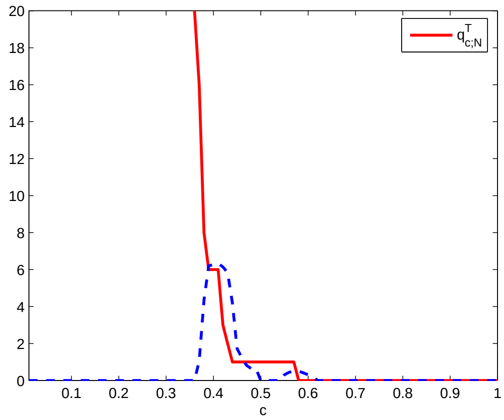
where θ_h and M_T are defined as in (4) and $p(N)$ is a penalty function satisfying

$$\lim_{N\to\infty} p(N) = 0 \quad \text{and} \quad \lim_{N\to\infty} Np(N) = \infty. \tag{8}$$

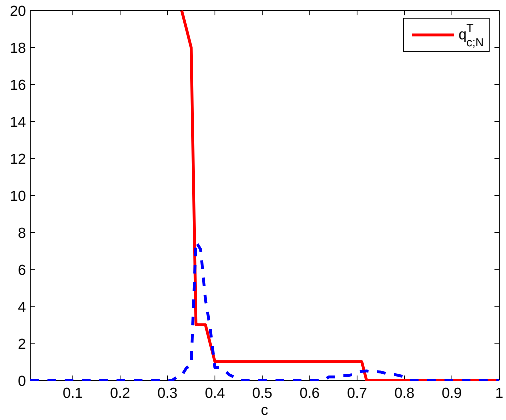
In principle, the maximum number of factors allowed q_{max} is the number of series in the dataset. Therefore, the penalty function should be large enough to avoid overestimation of $\hat{q}_N^T$, but at the same time it should not overpenalize. Multiplying the penalty function by a constant c is a way to tune the penalizing power of $p(N)$[6].

[5]Hallin and Liska suggest that this form has better finite sample performance. However, results obtained by using the non-logarithmic form of the criterion do not significantly differ.

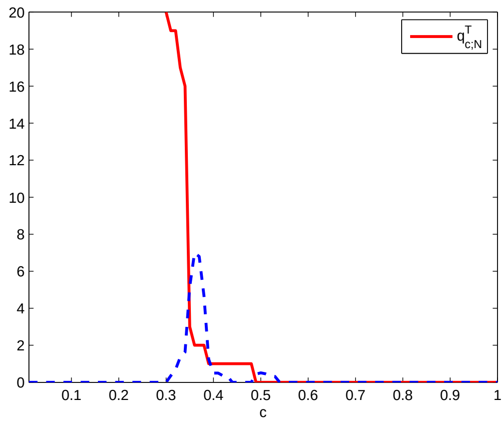
[6]We use $p(N,T) = (M_T^{-2} + M_T^{1/2}T^{-1/2} + N^{-1})log(min[N, M_T^2, M_T^{-1/2}T^{1/2}])$.



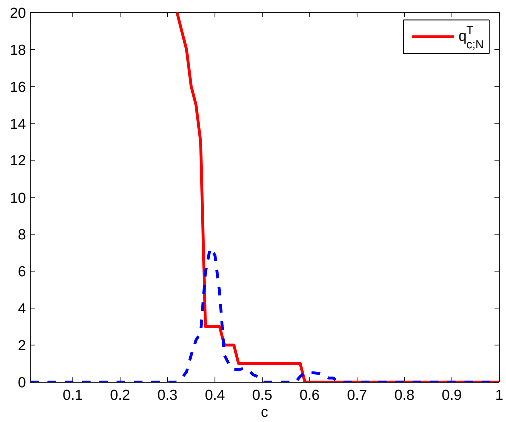
(a) Whole small sample, 57 series



(b) Electronics subsample, 89 series



(c) Chemicals subsample, 81 series



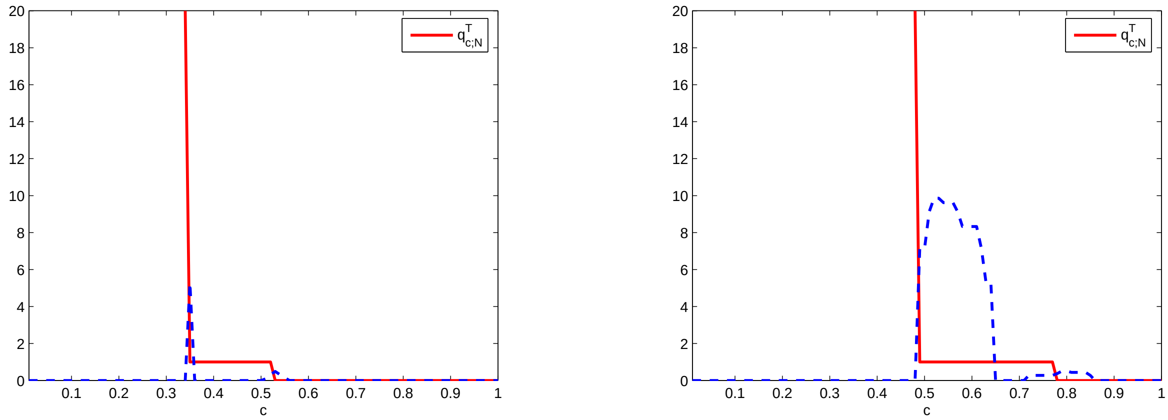
(d) Machinery subsample, 61 series

Figure 4: Sales samples. Hallin-Liska criterion plots.

Hallin and Liska propose an automatic procedure for selecting $\hat{q}_N^T$ which basically explores the behavior of the variance of the selected q_N^T for the whole region of values of the constant c for N and T going to infinity. What we seek is the first stability region compatible with $\hat{q}_N^T < N$. In the plots below, relative to the samples we use in the analysis, the solid line indicates the value of q_N^T while the dashed line represents the variance of q_N^T, thus detecting stability intervals. In other words, $\hat{q}_N^T$ corresponds to the second plateau of the solid line associated with a flat zero-level dashed line. In all the cases, the Hallin-Liska criterion indicates the existence of at least one common factor.

At the same time, we implement a complementary heuristic procedure which takes into account the amount of explained variance relative to the first q eigenvalues. In line with the literature, we aim at explaining at least 30% of the variance[7], hence we keep as many dynamic common factors as it is necessary in order to fulfill the objective. Table (1) summarizes the number of dynamic common factors chosen for each sample (the number of lags has been kept fixed at 4, i.e. one year) and the amount of explained variance for each sample. Moreover, we compute the average variance of each series' common part over the total variance, giving also its standard

[7]We do not consider a higher threshold because our data are pretty disaggregated.



(a) Whole sample, 355 series

(b) Chemicals subsample, 37 series

Figure 5: Investments samples. Hallin-Liska criterion plots.

deviation and its maximum and minimum values: this latter statistic in particular answers the question of why on average the variance of $\hat{\chi}_{it}$ is lower than the variance explained by the corresponding dynamic common factors. Indeed, we might interpret the variance explained by the first q dynamic common factors as a potential value reachable at the cost of including more and more lags to each dynamic common factor, trading off model parsimony and explanatory power.

Sample	No. of series	Hallin-Liska criterion: q	Variance explained by the first q eigenvalues	Average variance of $\hat{\chi}_i t$ over total: av	std	max	min	Heuristic criterion: q	Variance explained by the first q eigenvalues	Average variance of $\hat{\chi}_i t$ over total: av	std	max	min
Sales	57	1	17%	8%	10%	47%	0.3%	3	38%	19%	13%	67%	1%
	660	1	15%	7%	7%	54%	0.1%	3	31%	15%	11%	83%	2%
Sales Chemicals	81	1	15%	6%	5%	25%	0.2%	3	34%	15%	9%	47%	2%
Sales Electronics	89	1	21%	13%	17%	80%	0.3%	2	31%	19%	19%	87%	0.6%
Sales Machinery	61	1	18%	9%	10%	54%	0.3%	3	38%	20%	15%	85%	2%
Investments	355	1	16%	7%	4%	22%	0.5%	3	31%	14%	6%	43%	3%
Investments Chemicals	37	1	23%	8%	4%	17%	0.8%	2	33%	14%	5%	27%	5%

Table 1: Factor decomposition.

5 Common and Idiosyncratic Components

Typically, the common component and the idiosyncratic component of the series present different spectra: low frequencies are more important in the common part, higher frequencies in the idiosyncratic part. Long run movements of the economy as a whole reflect into that part of firm growth which by definition is linked to the deep driving forces of the economy, while short run phenomena in the life of the firm are likely to be idiosyncratic to the firm itself. Finally, factors affecting firm growth may be classified at a third intermediate level, the sectoral level, when they influence more than one firm but their effects do not overcome the borders of a specific industry.

Several studies have focused on the effects of macroeconomic factors on firm growth. In their analysis of the ultimate drivers of Fed monetary policy, Giannone et al. [2004] point out that the stochastic dimension of the US economy is two, i.e. there are two fundamental shocks, real and nominal. Basically, the real shock drives output while the nominal shock drives inflation. However, bridging our analysis between macroeconomic and microeconomic perspectives, studying both firm output and investment growth, we cannot leave out monetary issues. Indeed, the study of the main mechanisms of transmission of monetary policy all the way down to firms and, in turn, the relations between the financial sphere of the firm and the performance of the firm itself, have recently aroused great interest. For example, in their 2006 works Bottazzi et al. study the nexus between firms' financial condition and the properties of their size and growth dynamics, as well as the linkages between the financial structure of the firm and its performance in terms of productivity and profitability. Among their major findings, on the one hand, the presence of small or slowly growing firms among those classified as more risky, due to the presence of a general credit rationing mechanism influencing firm size and growth. On the other hand, the presence of risky big and fast growing firms, which in turn suggests that these liquidity constraints do not affect all the firms in the same way. Moreover, risky firms typically attain the worst performance, with some interesting exceptions at the top productivity and profitability levels. Although financial decisions are predominantly idiosyncratic to the firm itself, we may still regard as a macroeconomic common factor the general financial structure of the economy, including banking system features (e.g. the role of the lender of last resort) and laws regulating credit concession and transfer (e.g. Basel II international accords), as well as equity market patterns and exchange rate dynamics. Finally, as any macroeconomic common factor, the monetary factor does not affect all firms in exactly the same manner: some sectors are likely to be more exposed to shocks on credit while some others rely prevailingly on equity financing, some mainly address the domestic market while some others trade abroad, and so on. For example, after the Dot-com bubble burst investments have dropped more substantially in ICT firms relatively to non-ICT firms. This sensitivity is a consequence of the typical patterns of external finance raising of ICT firms, for which the cost of debt is higher than the average and the main source of funding is equity. As a second example, let us take rates on bank loans: as explained in Sundell [2002], in case of a monetary easing farm loan rates typically fall less and less quickly than non-farm rates. This happens basically because rural banks prefer keeping their small business loan rates as stable as possible, and also because for their loans they depend heavily on consumer deposits, whose rates typically lag open market rates.

As for the real shock, primary driver of output and ultimate cause of business cycle, it may take various forms. For instance, it can be a movement in the terms of trade, possibly due to the introduction of tariffs. It can be a shock on the demand side as a change in the quantity of government purchases, or in consumers' preferences between saving and spending, or in business investment decisions, as well as a fiscal policy. Finally, it can be a shock on the supply side, as a change in the labor supply or a productivity shock. This latter case, in turn, may be due to a variety of causes: the adoption of a new technology enhancing efficiency, the diffusion of new management techniques, an improvement in the quality of human capital as a consequence of better education policies, changes in the regulation affecting production either directly or indirectly (e.g. patent regulation and appropriability institutions in general), unusual metereological conditions, natural disasters, wars, et cetera. Among all the causes of a change in the production function, technological change is the most important, interesting and controversial.

Firstly, there is no consensus on one theoretical explanation of technological change. On the one hand, the straight neoclassical interpretation models technology by means of isoquants and assumes that an infinite number of input combinations is possible, being the chosen one the cheapest given input relative prices. In the neoclassical theory technology is a pool of knowledge that is generally applicable and easy to reproduce. On the other hand, Dosi [1982] proposes the notion of technological paradigm:

> "A "model" and a "pattern" of solution of *selected* technological problems, based on *selected* principles derived from natural sciences and on *selected* material technologies."

In this framework, the pursuit of innovation requires occasional shifts of paradigm, prescripting the direction of R&D. Innovation happens in two major ways: it may consist in finding out new attributes for a given object or it may be the discovery that another object carries some desired attributes. Whatever the type of innovation, however, it usually happens along some inference rule or pattern, which corresponds to a paradigm. A technological paradigm indicates in the first place what has been chosen as a solution for a generic technological task, thereby excluding alternative possibilities from the research field. Moreover, a paradigm indicates the valuable attributes or properties of this solution, which will be the object of further enhancement. The fulfillment of a paradigm is the technological trajectory, defined as "the pattern of "normal" problem solving activity on the ground of a technological paradigm" (ibidem). A technological trajectory identifies the dynamic of improvement of the artifacts along the economic and technological dimensions - or trade-offs - defined by the paradigm, it indicates the direction toward which innovation moves. To make a concrete example, in the case of the nanoscale paradigm technical change is represented by a trajectory of improvement in nanocircuits properties like information transmission and information storage. The dynamic theory of production, built on the notion of technological paradigm, assumes as basic drivers of change the processes of learning (in several forms, but basically innovation and imitation), selection (change in the weight of individual entities plus death of firms) and entry of new firms. A big issue concerns the process through which a technological path is chosen among all the possible paths: economic forces together with institutional and social factors do select the problems to tackle on the basis of profitability, feasibility, experience. The influence relation between technology on one side and economic and social environment on the other runs both ways.

A second source of debate on technological progress concerns its role in explaining business cycle. On the one hand, Real Business Cycle (RBC) theory studies business cycles with the assumption that they are basically driven by technology shocks[8]. In RBC models even the sum

[8]As pointed out in Rebelo [2005], Prescott [1986] estimates that technology shocks account for more than 50% of business cycle fluctuations, with peaks near 75%. However, the use he makes of total factor productivity is likely to everestimate the importance of technology shocks.

of many small technological shocks may generate fluctuations at business cycle frequencies, while nominal variables do not explain economic fluctuations. On the other hand, neither new-Keynesian economics, which imputes business cycle fluctuations mainly to demand shocks, nor the Monetarists, for which business cycle results from changes in money supply, recognize any fundamental function to technology shocks at business cycle frequencies. Indeed, the literature on the empirical investigation of the nature of business cycles, deeply linked to the issue of the decmposition of output into trend and cycle, is farly wide, going from the influential papers by Beveridge and Nelson [1981] and Nelson and Plosser [1982] to the updated unobserved component approach by Watson [1986], to the use of Structural VARs as in Blanchard and Quah [1989] and Gali [1996], to some recent works yielding opposite results on the contribution of technology shocks to the cycle (among the many, see Khan and Tsoukalas [2005] and Dupaigne et al. [2005]).
To conclude this overview on macroeconomic common factors, it is intersting to recall the study by Baum et al. [2002] on the effects of macroeconomic uncertainty. Indeed, they show how not only the sign and the magnitude of the shocks, whatever they are, do impact on firms' behavior, but also the variability itself of the macroeconomic conjuncture plays a role. The work focuses on firm's demand for liquidity, pointing out how a tranquil macroeconomic environment will lead managers to behave more idiosyncratically, adjusting decisions on the basis of firm-specific needs and relatively more reliable predictions. On the other hand, in high volatile periods firms will tend to adopt more homogeneous policies, being unable to tailor them precisely given the uncertainty on the macroeconomic situation in general and, in turn, on their own possible scenarios in particular. To proxy the state of the macroeconomy they use four proxies based on GDP, Industrial Production, CPI and S&P500: by so doing they capture different important aspects of the picture. However, we believe that also those macroeconomic factors left out from the analysis are likely to have the same influence on firms' attitude.

Although factor models, by their own nature, focus on the common forces driving economic phenomena, those aspects peculiar to the single units of analysis deserve attention too. Being our atomic agent the firm, it would be impossible to give a comprehensive picture of all the elements influencing what we have called the *idiosyncratic component* of each output or investment time series. The firm is a complex entity and any attempt to build an exhaustive list of all the potentially idiosyncratic items would necessarily leave out something. However, management science, business studies and theories of the firm give us interesting outlooks. For example, the organizational capability view of the firm focuses on those intangible "particular forms of organizational knowledge that account for the organization's ability to perform and extend its characteristic "output" actions - particularly the creation of a tangible product or the provision of a service, and the development of new products and services" (Dosi et al. [2000]). Empirical research in this field builds upon a quite flexible conceptual framework centered on the notions of organizational capabilities, organizational routines, individual skills, dynamic capabilities and core competences: all these "building blocks" of the firm typically are not codified and thus quite difficult to identify. They vary substantially from one firm to another and are levers for performance improvement, being among the factors that account for firms' success and decline. Moreover, the actions of a firm are interdependent: the final outcome is the result of the interaction among pieces of the system, but the functional relations among interrelated components are only partly understood. The deep mechanisms that lead to the rise of interdependence among parts of the firm and ultimately shape the behavior of the organization in relation to the environment are idiosyncratic to the firm itself.

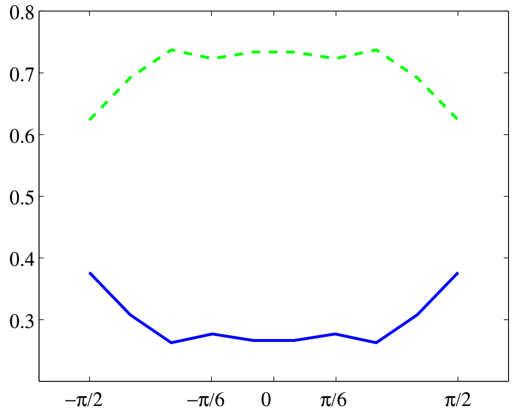
More specifically, Rivkin and Siggelkow [2003], addressing the issue of how interdependencies arise in organizations, identify three key organizational elements - vertical hierarchy, incentive system, decomposition into diverse departments - plus two contextual variables relative to the underlying pattern of interaction among firm's decisions and to the ability of managers to process information. Firms differ in their design in the sense that the characteristics of these five elements are different across firms.

The results of our empirical analysis are consistent with the idea that macroeconomic common factors, leaving their specific nature, relative importance and interpretation out of consideration, do affect firm's output and investment mainly at business cycle frequencies. On the contrary, what of firm's performance is due to idiosyncratic elements, i.e the idiosyncratic part of sales and investment series, turns to be important at higher frequencies. This is evident from the spectra of the common and the idiosyncratic components of the series (retrieved on the basis of the heuristic criterion for the choice of the number of factors). For each sample and sectoral subsample we have averaged the spectra of the common component and of the idiosyncratic component of each single series and plotted this two synthetic spectra in figure (6) and figure (7). These pictures show how, on average, the spectrum of the common component of output and investment series (solid line) has a peak at the $\pi/6$ frequency, corresponding to a 3 year period. As for the peak at $\pi/2$, we might explain it as some residual seasonality of the common component itself that has not been washed away by the deseasonalization operated on the original data. The spectral density of x_t being normalized at 1 at each frequency, the idiosyncratic component (dashed line) presents a mirror-like spectrum: at business cycle frequencies its importance is low, while it increases more and more going towards higher frequencies.
Samples relative to the whole economy and sectoral samples look pretty much alike, although the idiosyncratic component weights slightly uniformly more in these latter cases than in the whole samples. Note the presence of more pronounced peaks at $\pi/6$ in the common component spectrum in the sectoral electronics and machinery sales samples: it suggests that also sectoral-wide shocks, not only economy-wide shocks, play a role at business cycle frequencies, i.e. may generate fluctuations for the whole economy. The chemical sector, however, seems to play a more important role at frequencies corresponding to shorter periods rather than at business cycle frequencies: especially in the case of investments, the peak at $\pi/6$ disappears and gets spread on the entire frequency band going from 0 to about $\pi/4$, corresponding to a period of two years, while the bulk of the variance corresponds to periods from one to two years.

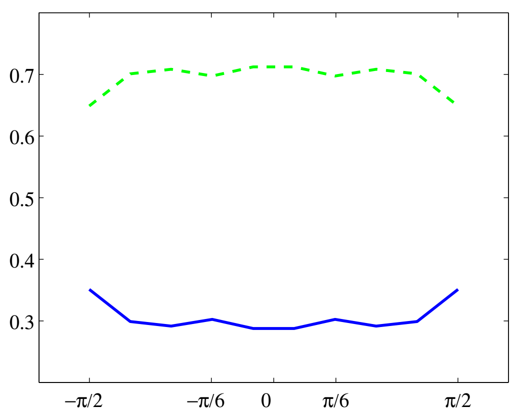
Besides the analysis of the spectrum of the common component and of the idiosyncratic component, it is interesting to check how common components comove with each other. Focusing on comovements involving only the common component of the series adds value to the mere analysis of comovements among the series of sales and investments as such. Indeed, by taking into account only on the part of the series which is linked to the deep structure of the economy, we don't run the risk of interpreting as structural those comovements which are just due to peculiarities of some firms.
The cospectrum relative to the common parts of series i and series j is defined as:

$$s_{ij}(\theta) = \sum_{k=-\infty}^{\infty} c_{ij}^{k} cos\theta k \tag{9}$$

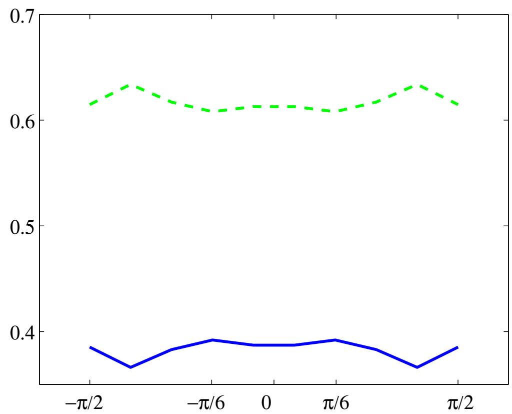
where c_{ij}^{k} is the correlation coefficient between the two common parts at lag k. We follow



(a) Investment sample - 355 series, average spectra

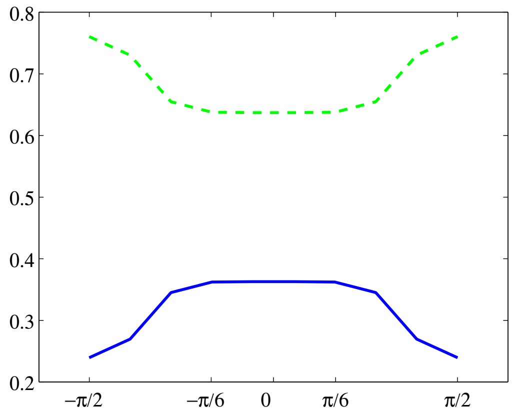


(b) Sales sample - 660 series, average spectra

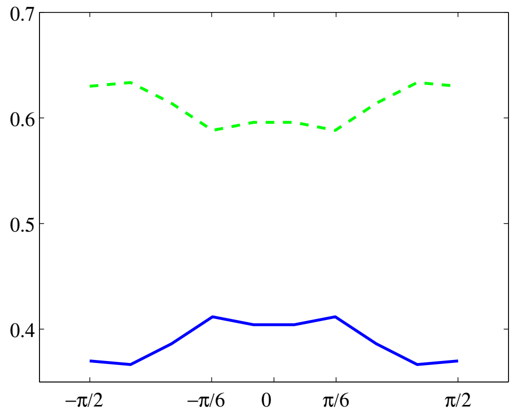


(c) Sales sample - 57 series, average spectra

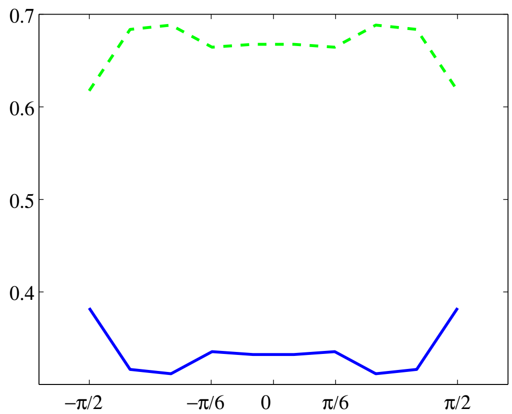
Figure 6: Average spectral densities for whole samples. Solid line:χ_t, dashed line: ξ_t.



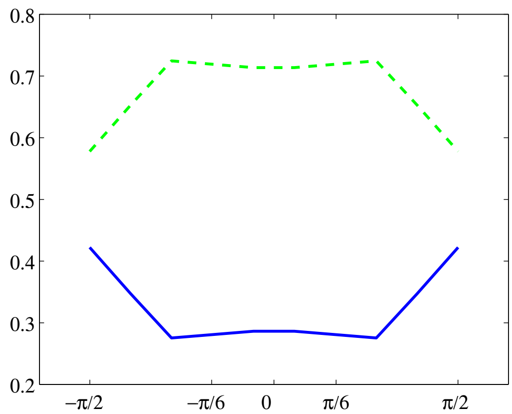
(a) Electronics sample - sales - 81 series, average spectra



(b) Machinery sample - sales - 61 series, average spectra



(c) Chemicals sample - sales - 81 series, average spectra



(d) Chemicals sample - investments - 37 series, average spectra

Figure 7: Average spectral densities for sectoral samples. Solid line:χ_t, dashed line: ξ_t.

the approach proposed by Forni and Reichlin [1998] for measuring positive and negative comovements. Firstly, the cospectrum is decomposed into the sum of a positive cospectrum and a negative cospectrum, defined respectively as:

$$s_{ij}(\theta)_+ = [|s_{ij}(\theta)| + s_{ij}(\theta)]/2 \tag{10}$$

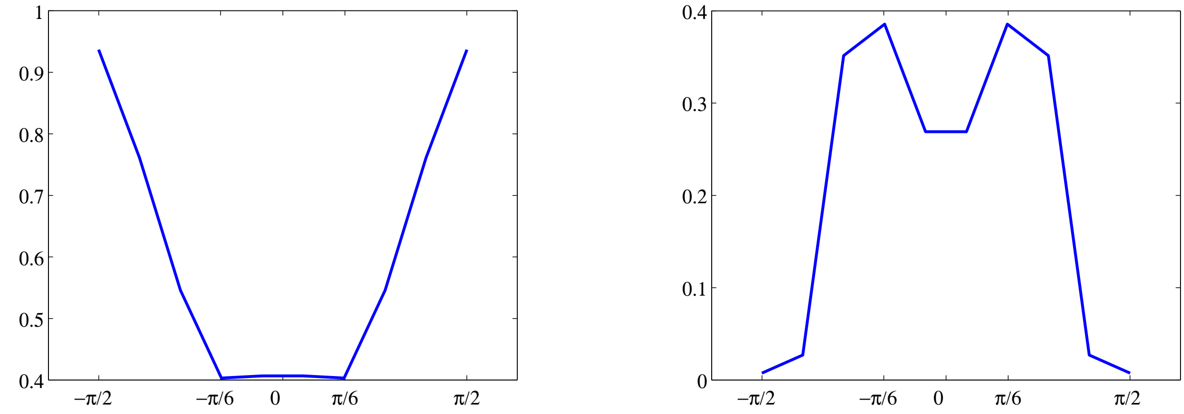
and

$$s_{ij}(\theta)_- = [s_{ij}(\theta) - |s_{ij}(\theta)|]/2. \tag{11}$$

Secondly, a synthetic measure is computed:

$$S(\theta) = -\frac{\sum_{ij} s_{ij}(\theta)_-}{\sum_{ij} s_{ij}(\theta)_+}. \tag{12}$$

A value of $S(\theta)$ close to 1 at a given frequency indicates strong negative comovements, while a value close to 0 indicates that positive comovements are important. As shown in figure 8, where the quantity $S(\theta)$ has been plotted for the biggest samples of sales and investments, the pattern of comovements is pretty different between sales and investment series. Firms' sales definitely positive comove in the long run,i.e. for periods longer than three years. In the medium and short run, however, they begin to comove negatively, and the shorter the period the more negatively they comove. This is reasonable: while in the long run every firm will be influence by the economic conjuncture, thereby homogenizing its performance to the general boom or slowdown of the economy, in the medium run and especially in the short run the mechanism of competition plays the major role. In the short run, the struggle for market shares by definition causes negative comovements on sales. The plot of $S(\theta)$ for investment series, instead, looks pretty much the opposite. Firms' investment behavior is extremely homogeneous in the short run, while it becomes more and more heterogeneous the longer the period. The peak of negative comovements in investment series corresponds to a 3 year period. A tentative interpretation of this finding relies in the monetary policy transmission mechanism, which is effective in bringing all the way down to firms each change in interest rates, which ultimately affects investments. In the short run, every firm will be affected in the same way by monetary policy. On the other hand, investment decisions in the long run seem to be inherent to the firm itself, depending on a number of factors as the attitude of the owner, in the case of a small firm, or the strategic decisions of the management in the case of a big company. Plots relative to sectoral subsamples do not diverge substantially from the bigger samples.



(a) Sales sample - 660 series

(b) Investments sample - 355 series

Figure 8: Behavior of $S(\theta)$.

6 Conclusions and further research

This work is an application of the Generalized Dynamic Factor Model to industrial dynamics. In particular, we have analyzed firm growth from a dynamic factor perspective, thereby aiming at disentangling that part of the phenomenon which is due to economy-wide shocks from that component which is idiosyncratic to each firm. We have proxied firm size by means of sales and investments quarterly series belonging to the COMPUSTAT database and covering the last 20 years. Thanks to the characteristics of the GDFM, we have been able to use extremely wide cross-sections, up to 660 series, and decompose them in sectoral subsamples, still about ten times wider than those used for traditional VAR analysis.
We have applied different criteria for the choice of the number of dynamic common factors to include in the model. Consistently with the preliminary graphical analysis of the eigenvalues, the Hallin-Liska criterion indicates just one dynamic common factor, explaining from 15% to 23% of the variance depending on the sample. However, it is necessary to include up to three dynamic common factors if we aim at explaining at least 30% of the variance.
The spectra of the common and of the idiosyncratic component are consistent with the idea that macroeconomic dynamic common factors - being they nominal, as monetary policy interventions or exchange rate movements, or real, as technological advances - do drive firm growth at business cycle frequencies. On the other hand, the importance of the idiosyncratic component - indeed a fairly complex entity to interpret - increases at higher frequencies. The average spectrum of the common component of the series belonging to sectoral samples suggest that shocks in the electronics or in the machinery industry may generate fluctuations at business cycle frequencies, while the chemical sector plays a role in periods from one to two years. Finally, we analyzed comovements across common components, detecting two opposite patterns in sales and investments. While sales comove negatively in the short run and positively in the long run, investments comove negatively in the long run and positively in the short run.

The main direction for further research is the identification of the dynamic common factors, going beyond the distinction between common and idiosyncratic components and studying the building blocks of the common component itself. However, since the identification of each shock would require a number of restrictions, we would rather focus on the technological common factor. Indeed, in the light of the overview on macroeconomic common factors in the previous section, it seems fair to expect the technological common factor to play a fundamental role and it would be interesting to understand the nature and the characteristics of the this force driving output and investments. Note however that the model representation that we used in this work is not fundamental. As explained in Lippi and Reichlin [1994], nonfundamentalness poses a serious problem concerning the identification of economically sensible impulse-response functions associated to each shock u_{jt}. Thus, before any identification strategy is proposed, the problem of nonfundamentalness has to be solved.

References

K. G. Abraham and L. F. Katz. Cyclical unemployment: Sectoral shifts or aggregate disturbances? *Journal of Political Economy*, 94(3):507–22, June 1986.

E. Bartelsman, J. Haltiwanger, and S. Scarpetta. Measuring and analyzing cross country differences in firm dynamics, March 2005a. Paper prepared for NBER Conference on Research in Income and Wealth.

E. Bartelsman, S. Scarpetta, and F. Schivardi. Comparative analysis of firm demographics and survival: Micro-level evidence for the oecd countries. *Industrial and Corporate Change*, 14(3): 365–391, April 2005b.

C. F. Baum, M. Caglayan, N. Ozkan, and O. Talavera. The impact of macroeconomic uncertainty on non-financial firms' demand for liquidity. Boston college working papers in economics, Boston College Department of Economics, Dec. 2002.

S. Beveridge and C. Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7:151–174, 1981.

O. J. Blanchard and D. Quah. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79:655–673, 1989.

J. Boivin and S. Ng. Understanding and comparing factor-based forecasts. NBER Working Papers 11285, National Bureau of Economic Research, Inc, May 2005.

G. Bottazzi, A. Secchi, and F. Tamagni. Financial fragility and growth dynamics of italian business firms. Working Papers 2006/07, LEM, Sant'Anna School for Advanced Studies, March 2006a.

G. Bottazzi, A. Secchi, and F. Tamagni. Productivity, profitability and financial fragility: Empirical evidence from italian business firms. Working Papers 2006/08, LEM, Sant'Anna School for Advanced Studies, March 2006b.

D. Brillinger. *Time Series: Data Analysis and Theory.* Holden-Day, Inc., San Francisco, CA, 1981.

R. Cooper and J. C. Haltiwanger. Inventories and the propagation of sectoral shocks. NBER Working Papers 2425, National Bureau of Economic Research, Inc, July 1990.

S. J. Davis and J. Haltiwanger. On the driving forces behind cyclical movements in employment and job reallocation. *American Economic Review*, 89(5):1234–1258, December 1999.

S. J. Davis and J. C. Haltiwanger. Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3):819–63, August 1992.

S. J. Davis, J. Haltiwanger, and Schuh. *Job creation and destruction.* MIT Press, 1998.

G. Dosi. Technological paradigms and technological trajectories. a suggested interpretation of the determinants and directions of technical change. *Research Policy*, 1982.

G. Dosi. Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature*, 26(3):1120–71, September 1988.

G. Dosi. Statistical regularities in the evolution of industries. a guide through some evidence and challenges for the theory. LEM Papers Series 2005/17, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy, July 2005.

G. Dosi, R. Nelson, and S. Winter. *The Nature and Dynamics of Organizational Capabilities*, chapter Introduction. Oxford University Press, 2000.

M. Dupaigne, P. Fève, and J. Mathéron. Technology shocks and employment: Do we really need dsge models with a fall in hours? IDEI Working Papers 349, Institut d'Économie Industrielle (IDEI), Toulouse, Apr 2005.

A. Figura. Is reallocation related to the cycle? a look at permanent and temporary job flows. Finance and Economics Discussion Series 2002-16, Board of governors of the Federal Reserve System, 2002.

M. Forni and L. Reichlin. Let's get real: A factor analytical approach to disaggregated business cycle dynamics. *Review of Economic Studies*, 65(3):453–73, July 1998.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: identification and estimation. *The Review of Economics and Statistics*, 82:540–554, 2000.

M. Forni, D. Giannone, M. Lippi, and L. Reichlin. Opening the black box: Structural factor models with large cross-sections. Technical report, Interuniversity Attraction Pole, 2006a.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, 2006b. Forthcoming.

L. Foster, J. Haltiwanger, and C. Krizan. Aggregate productivity growth: Lessons from microeconomic evidence. NBER Working Papers 6803, National Bureau of Economic Research, Inc, Nov. 1998.

J. Gali. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations. NBER Working Papers 5721, National Bureau of Economic Research, Inc, Aug. 1996.

D. Giannone, L. Reichlin, and L. Sala. Monetary policy in real time. Working Papers 284, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University, 2004.

M. Hallin and R. Liska. The generalized dynamic factor model: determining the number of factors. Technical report, Interuniversity Attraction Pole, 2005.

J. C. Haltiwanger. Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence. Technical report, Federal Reserve bank of St.Louis Economic Review, May 1997.

G. Kapetanios and M. Marcellino. A comparison of estimation methods for dynamic factor models of large dimensions. Working Papers 489, Queen Mary, University of London, Department of Economics, Apr. 2003.

H. Khan and J. Tsoukalas. Technology shocks and uk business cycles. Macroeconomics 0512006, EconWPA, Dec. 2005.

D. Lilien. Sectoral shifts and cyclical unemployment. *Journal of Political Economy*, 90:777–793, 1982.

M. Lippi and L. Reichlin. VAR analysis, non-fundamental representation, Blaschke matrices. *Journal of Econometrics*, 63:307–325, 1994.

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

P. Loungani and R. Rogerson. Cyclical fluctuations and sectoral reallocation: Evidence from the PSID. *Journal of Monetary Economics*, 23(2):259–273, March 1989.

N. G. Mankiw. Real business cycles: A new keynesian perspective. *Journal of Economic Perspectives*, 3(3):79–90, Summer 1989.

D. Mortensen and C. Pissarides. Job reallocation, employment fluctuations and unemployment. CEP Discussion Papers 0421, Centre for Economic Performance, LSE, Apr. 1999.

C. Nelson and C. Plosser. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10:139–162, 1982.

E. C. Prescott. Theory ahead of business cycle measurement. Staff Report 102, Federal Reserve Bank of Minneapolis, 1986.

S. T. Rebelo. Real business cycle models: Past, present, and future. NBER Working Papers W11401, National Bureau of Economic Research, Inc, June 2005.

L. Reichlin. Factor models in large cross-sections of time series. CEPR Discussion Papers 3285, C.E.P.R. Discussion Papers, Mar. 2002.

E. R. Rissman. Measuring labor market turbulence, 1997.

J. Rivkin and N. Siggelkow. Balancing search and stability: Interdependencies among elements of organizational design. *Management Science*, 49:290–321, 2003.

T. Sargent and C. Sims. Business cycle modeling without pretending to have too much a-priori economic theory. In C. Sims et al., eds., New Methods in Business Cycle Research, 1977.

S. Scarpetta, J. Haltiwanger, and E. Bartelsman. Microeconomic evidence of creative destruction in industrial and developing countries. Policy Research Working Paper Series 3464, The World Bank, Dec. 2004.

J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20:147–162, 2002.

P. Sundell. Macroeconomic factors behind the fall in farm interest rates. Electronic Outlook Report from the Economic Research Service AIS-78-01, United States Department of Agriculture, March 2002.

M. W. Watson. Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18:49–75, 1986.

Appendix A

What follows is drawn from the COMPUSTAT User's Guide.

Sales (*data2*)
This item includes:

1. Advertising companies' net sales and commissions earned
2. Airline companies' transportation related revenues including passenger, cargo, mail, charter, and other
3. Any revenue source expected to continue for the life of the company
4. Banks and savings and loans' interest income and fee revenue
5. Banks' total current operating revenue and net pretax profit or loss on securities sold or redeemed
6. Broadcasting companies' net of agency commissions
7. Commissions
8. Equity in earnings/losses even if negative for real estate investment trusts and investors
9. Finance companies' earned insurance premiums and interest income. Finance companies' sales are counted only after net losses on factored receivables purchased are deducted from Sales - Net
10. Franchise operations' franchise and license fees and sales
11. Hospitals' sales net of provision for contractual allowances (sometimes includes doubtful accounts)
12. Income derived from equipment rental considered part of operating revenue
13. Installment sales
14. Leasing companies' rental or leased income
15. Life insurance and property and casualty companies' net sales in total income
16. Management fees
17. Net sales of tobacco, oil, rubber, and liquor companies' after deducting excise taxes
18. Oil and extractive companies' mineral royalty income
19. Operative builders' interest income from mortgage banking subsidiaries
20. Other operating revenue
21. Reimbursements for out of pocket expenses reported on the Income Statement
22. Rental income, if included by the company in Sales

23. Research and development companies' equity income from research and development joint ventures (when reported as operating income) and government grant income
24. Retail companies' finance charge revenues
25. Retail companies' sales of leased departments when corresponding costs are not available but are included in expenses
26. Royalty income and/or management fees when considered as part of operating income (such as, oil companies, extractive industries, publishing companies)
27. Security brokers' other income
28. Shipping companies' operating differential subsidies and income on reserve fund securities shown separately
29. Utilities' net sales total current operating revenue

This item excludes:

1. Broadcasting companies' agency commissions
2. Casinos' promotional allowances
3. Cost of delivery expenses for paper mills
4. Discontinued operations
5. Equity in earnings of unconsolidated subsidiaries
6. Excise taxes excluded from Sales (Net)
7. Gain on sale of securities or fixed assets
8. Interest income
9. Nonoperating income
10. Other income
11. Provision for contractual allowances for hospitals
12. Rental income

Capital Expenditures (*data90*)
This item includes:

1. Any items included in property, plant and equipment on the Balance Sheet
2. Expenditures for capital leases
3. Increase in funds for construction
4. Increase in Leaseback Transactions

5. Logging roads and timber
6. Reclassification of inventory to property, plant, and equipment

This item excludes:

1. Capital expenditures of discontinued operations
2. Changes in property, plant, and equipment resulting from foreign currency fluctuations when listed separately
3. Decreases in funds for construction presented as a use of funds
4. Deposits on property, plant and equipment
5. Net assets of businesses acquired
6. Property, plant, and equipment of acquired companies
7. Property, plant and equipment for real estate investment trusts
8. Software costs (unless included in property, plant and equipment on the Balance Sheet)

This item is not available for banks.
Data reflects year-to-date figures for each quarter.

What follows is the list of data anomalies causing a firm to be dropped from the databases used for the analysis:

- Data reflects an acquisition (purchase and/or pooling)
- Data reflects an accounting change
- Reflects fresh-start accounting upon emerging from Chapter 11 bankruptcy
- Excludes discontinued operations
- Includes excise taxes
- Includes other income/excludes some operating revenue
- Includes sales of leased departments
- Some or all data is not available because of a fiscal year change
- Some or all data is not available because the company has been in operation less than one year or presents more than or less than 12 months of data in their statements
- Includes six months of a merger or acquisition
- Includes nine months of a merger or acquisition
- Includes 12 months of a merger or acquisition
- Excludes six months of discontinued operations
- Excludes nine months of discontinued operations
- Excludes 12 months of discontinued operations

Appendix B

To seasonally adjust our dataset we used a multiplicative method based on the ratio between the original series y_t and its centered moving average, defined as:

$$z_t = \frac{1}{} \quad - \qquad\qquad\qquad\qquad -$$

$$\overline{(i_1 i_2 i_3 i_4)^{1/4}}$$

The interpretation is that the original series is s_j percent higher in quarter j relative to the adjusted series, which is finally computed as the ratio between y_t and the seasonal factors.