EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1366 / AUGUST 2011

# STOCK MARKET FIRM-LEVEL INFORMATION AND REAL ECONOMIC ACTIVITY

by Filippo di Mauro,
Fabio Fornari
and Dario Mannucci

EUROPEAN CENTRAL BANK

EUROSYSTEM

# WORKING PAPER SERIES

## NO 1366 / AUGUST 2011

# STOCK MARKET FIRM-LEVEL INFORMATION AND REAL ECONOMIC ACTIVITY[1]

by Filippo di Mauro[2], Fabio Fornari[2]
and Dario Mannucci[3]


In 2011 all ECB publications feature a motif taken from the €100 banknote.

NOTE: This Working Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.



1 Thanks are due to an anonymous referee, Hashem Pesaran, Neil Ericsson, Marco Lombardi, Michael Binder, Til Schuermann, George Kapetanios, Carlo Favero and participants at seminars held at the ECB and the IMF as well as at a conference in Honour of Hashem Pesaran (Cambridge, UK, 1-2 July 2011).
2 European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; e-mails: filippo.di_mauro@ecb.europa.eu and fabio.fornari@ecb.europa.eu
3 Prometeia, Via G. Marconi 43, 40122 Bologna, Italy; e-mail: dario.mannucci@gmail.com

**Address**
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

**Postal address**
Postfach 16 03 19
60066 Frankfurt am Main, Germany

**Telephone**
+49 69 1344 0

**Internet**
http://www.ecb.europa.eu

**Fax**
+49 69 1344 6000



*Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html*

ISSN 1725-2806 (online)

# CONTENTS

**Abstract**

We provide evidence that changes in the equity price and volatility of individual firms (measures that approximate the definition of 'granular shock' given in Gabaix, 2010) are key to improve the predictability of aggregate business cycle fluctuations in a number of countries. Specifically, adding the return and the volatility of firm-level equity prices to aggregate financial information leads to a significant improvement in forecasting business cycle developments in four economic areas, at various horizons. Importantly, not only domestic firms but also foreign firms improve business cycle predictability for a given economic area. This is not immediately visible when one takes an unconditional standpoint (i.e. an average across the sample). However, conditioning on the business cycle position of the domestic economy, the relative importance of the two sets of firms - foreign and domestic - exhibits noticeable swings across time. Analogously, the sectoral classification of the firms that in a given month retain the highest predictive power for future IP changes also varies significantly over time as a function of the business cycle position of the domestic economy. Limited to the United States, predictive ability is found to be related to selected balance sheet items, suggesting that structural features differentiate the firms that can anticipate aggregate fluctuations from those that do not help to this aim. Beyond the purely forecasting application, this finding may enhance our understanding of the underlying origins of aggregate fluctuations. We also propose to use the cross sectional stock market information to macro-prudential aims through an economic Value at Risk.

JEL Classification: C53; C58; F37; G15

Keywords: *Business cycle forecasting*; *granular shock*; *international linkages.*

## Non-technical summary

Real developments, as measured for example by changes in GDP or Industrial Production (IP) Indices over selected horizons, are typically forecast through a combination of macroeconomic variables, financial variables and confidence indicators.

These three sets of variables have been so far typically selected at the aggregate level, i.e. no firm-level information has been regularly employed to forecast business cycle developments. The reason for this is that firm-level shocks should wash out with each other in the aggregate and therefore they should not affect the overall economy.

However, it has been recently shown (Gabaix, 2010) that the cross sectional distribution of firms' size matters a lot for the validity of this assumption. If the distribution of firms' size has fat tails, then firm-level shocks may propagate to the overall economy. Gabaix indeed showed that the idiosyncratic shocks to the rate of growth in the sales of the largest US firms can predict the one-quarter-ahead growth rate of the US GDP.

In this paper we analyse in more detail the implications of Gabaix's theory, taking as well an international perspective that looks at four economic areas. However, we do not restrict ourselves to considering big firms. Rather, we analyse the predictive power stemming from a large cross section of firms'equity prices with the key finding that, in a given month, it is only a small subset of these firms that help improve predictability. Overall, the composition of the set of most predictive firms remains stable for around half a year. It is only after this identification has been made that we investigate which are the firms' characteristics that are associated with an high predictive power for subsequent changes in the IP indices.

Among other results we show that i) idiosyncratic shocks to firm-level equity returns and variances can noticeably improve the prediction of the growth rate in the IP indices especially at horizons between 12 and 24 months; ii) for a given economic area, domestic firms and foreign firms are equally important to improve the forecast and their relative ability to do so changes a lot across the cycle; iii) among the features which make firms helpful in anticipating real growth, size does not seem to be a key factor. Rather the sector in which firms operate as well as other balance sheet items related to the performance of the firms, their investments as well as their international activity seem to be more prominent.

Taken together these findings can help shed more light of the key factors behind aggregate fluctuations.

# 1 Introduction

The recent recession episode that started in the United States in December 2007 stood as another challenge for our ability to anticipate the timing and the amplitude of business cycle fluctuations. Throughout 2007, almost all the forecasts computed by central banks, academics and market participants were not able to detect the approaching sharp decline in real GDP, even when produced around end-2008, right ahead of remarkably negative GDP growth figures. The highly coincident and sharply negative GDP growth rates recorded almost worldwide through the recession, and especially in 2008Q4 and 2009Q1, contribute to make the failure in forecasting even more serious and call, at the very least, for a critical review of the mainstream forecasting methodologies. This paper aims to make some steps in this direction.

So far, economic fluctuations have been predicted almost exclusively through the *aggregate information* conveyed either by i) macro variables (labor market conditions, money, credit, lagged growth), ii) financial indicators (aggregate stock market returns and variances, slope of the yield curve, credit spreads) or iii) confidence (households or business) indicators. Focusing on models including aggregate financial variables, which are also the focus of the present paper, a broad conclusion reached by analyses carried out so far is that their predictive power is broadly unstable over time and also that the set of indicators which are key to improve the forecast of business cycle developments tends to change composition over time.

Fornari and Mele (2009) provide a detailed assessment of the out of sample forecasting ability of univariate linear and non linear models which rely on financial indicators. Overall, their conclusion is that the term spread, together with a time-varying measure of stock market volatility, does a rather good job in anticipating the rates of change in the US post-War industrial production index. However, nearly all of the combinations of variables they look at have their moment of popularity, so that what is eventually judged to be the best model is not the best model consistently across the sample. This finding cannot but confirm that recessions are intrinsically different, both as concerns their roots and the way in which the originating shock propagates across the economy.

But, if recessions are different and shocks transmit both domestically and internationally in a time varying fashion, should not we employ a broader set of regressors - and potentially models - to better track this variability across time? For example, many recent approaches to forecasting consider pooling the individual forecasts stemming from a large number of models, each differing from the other as concerns for example the lag specification, the sample over which estimation is carried out, the number of variables included. This has been the way in which the so-called *uncertain instabilities* have been dealt with in weather forecasting, an approach which has recently spilled over to macroeconomic and financial forecasting (see Amisano and Geweke, 2009; Clark and McCracken, 2006; Jore et al., 2008).

In this paper we come closer to this strategy as we test the hypothesis that a linear combination of selected past idiosyncratic shocks recorded by the equity price of a given firm helps track and forecast aggregate business cycle fluctuations. At this stage we like to anticipate, however, that, somewhat against the benefit achieved by pooling many forecasts suggested by this strand of literature, our conclusions are that pooling individual information does not typically represent a good alternative to a situation in which instead a small number of regressors (i.e. a subset of the full information set whose composition changes over time) are selected according to some real-time criterion of fit. In other words, the largest part of the improvement in predictive ability which is found inside the large cross section of equity prices that we look at comes, at each point in time, from the idiosyncratic equity price movement recorded by a handful of firms out the large number which composes the cross section.

Firm-level information did not receive big attention in macro forecasting so far (see, however, Gilchrist et al., 2009, for an application in which firm-level credit spreads are used for business cycle forecasting)[1] primarily as the idiosyncratic fluctuations of a given equity price should be irrelevant in an aggregate economy characterised by a large number of firms. This assumption, however, heavily depends on the empirical distribution of firms' size having thin tails, i.e. finite variance. However, a fat tailed distribution may be a better proxy of reality, consistently also with the industrial structure of modern economies, in which the weight of large corporations and multinationals has been significantly on the rise over at least the last two decades. It is exactly under the latter conditions that Gabaix (2010) derived his so-called *granular* explanation of aggregate fluctuations.[2] Basically, his empirical evidence shows that the aggregated shock to the rate of growth of the sales made by the 100 largest US firms anticipates the rate of growth of the US GDP over the subsequent quarter and has a power which remains robust to the various controls that he applies. We anticipate, however, that we do not find size (as measured by sales in the empirical evidence in Gabaix) to be the key reason behind the predictive power for aggregate fluctuations that we find in the equity price of specific firms. We also show that the gain in the predictability of business cycle conditions that we find in the cross section of equity prices does not come randomly from any given firm. Rather, it is highly concentrated within a limited subset of these firms whose size, as measured by more than one criterion, is however very scattered. If any, a sector-related explanation has more empirical support than size. In this paper we also consider the international dimension of the granularity hypothesis, i.e. whether the idiosyncratic equity price movement of firms in a given country $i$ matter to explain the aggregate fluctuations in another country, $j$, controlling for some $j$-related pieces of information. As for countries, we look at the United States, the United Kingdom, Japan and a subset of the euro area represented by Germany, France and Italy.

Before moving forward let us also point out that predictive power of firm specific shocks for aggregate fluctuations is also hinted in the *news shocks - animal spirit shocks* interpretation of the innovations to consumer confidence provided in Barsky and Sims (2010). They find that shocks to consumer confidence, while orthogonal to current consumption and growth, give rise to persistent increases in such variables over time. In other words, unexpected developments in confidence seem to be *clean* signals of future rises in productivity. We conjecture that a similar role could be played by the shocks to individual (and aggregate) equity prices. For example, an unexpected decline in the equity price of a firm could stem from the postponement of some of the firms' projects - due for example to lack of demand for its products or tight credit availability - which some market analysts first - and eventually the market as a whole - interprets as a bad signal about the future profitability of the firm. Of course, being firm-specific, this shock will be irrelevant for most of the remaining firms as well as for the aggregate economy in the specific moment in which it is realized. Nonetheless, it may be capturing the first signs of of macroeconomic or financial shocks that later on will eventually spread through the whole economy. The fact that our regressions evidence that the predictability of the changes in the industrial production indices peaks at longer horizons rather than at very short ones would suggest that also shocks to equity price are almost orthogonal to current growth, while anticipating future developments in business cycle conditions over more distant horizons.[3]

[1]This paper points out that not any corporate bond spread helps forecast business cycle developments. Rather the forecasting power of corporate bonds with too high or too low rating is poorer than for bonds with a 'average' rating.

[2]Similarly to what Gabaix proposes, Carvalho, 2009, shows that network effects among sectors generate significant propagation effects. There is also an established literature exploring the impact of microeconomic shocks on aggregate fluctuations, as Jovanovic, 1987; Durlauf, 1993; Horvath, 1998, 2000; Conley and Dupor, 2003.

[3]Always with reference to equity price shocks, Beaudry et al., (2010) analyze the international spillover of news

The paper is organized as follows. In Section 2 we describe the data and the econometric methodology. In section 3 we report some unconditional evidence of the relationships between real activity, aggregate information and firm level variables, in the countries that we consider. This evidence is intended to give a preliminary flavour of the results presented in the remainder of the paper. Section 4 investigates the domestic and the international dimension of the granularity hypothesis through an out of sample econometric exercise. Section 5 looks at the sector-wise composition of the predictive distribution of the firms as well as - limited to the United States - it analyzes whether characteristics of the firms, as captured by key balance sheet items, are related to their predictive power for business cycle developments. Section 6 looks at some robustness issues while Section 7 evidences how the cross sectional equity market information could be used from a macro-financial stability perspective.

# 2 Methodology

The hypothesis that we want to test is whether real economic activity - proxied by industrial production - can be better anticipated when one looks at firm level information[4] in addition to aggregate information. Beyond lagged industrial production, our aggregate variables include the term spread ($Term$) and the return and the variance of the stock market index ($MktRet$ and $MktVar$). Admittedly we do not consider too large a set of macroeconomic indicators and there are two main reasons to do so. First of all, financial variables have been typically found to quickly embody information releases about a broad set of macroeconomic variables. In this respect, we expect financial variables to be good substitutes for macroeconomic information at the monthly frequency we adopt in the paper. In addition, a large body of literature has evidenced that financial variables do a good job in fitting and anticipate business cycle phases (Estrella, 2005). Second, as we take a real-time standpoint in performing the predictive regressions, the different release dates of macroeconomic variables should be properly handled and would need to be examined within a setup similar to Aruoba et al. (2009), leading to a much more complex framework than the simple linear regressions we employ.[5] Although Stock and Watson (2003) are frequently reported as evidence against the existence of predictive power in financial variables, we rely on them especially as the results in Espinoza et al. (2011) point to financial information i) being not useless when one takes an out of sample standpoint and ii) being more important in improving the forecasts in periods characterized by financial turbulence.

We forecast developments in the growth rate of the Industrial Production index in a given country/economic area over $h$ months through the following simple univariate regression:

shocks and conclude that a news shock in a large country can create national business cycles and international business cycles, thereby providing motivation for our research, although in their analysis the spillover of the news shock is related to a concept of geographical proximity.

[4]The firm level variables that we use are the return and the variance of selected equity prices ($Ret^i$, $Var^i$), which match the aggregate information we look at.

[5]See also Giannone et al., 2008, for a related approach.

$$
\begin{aligned}
\Delta_h \ln(ip)_t = \alpha + & \\
& + \sum_{j=1}^{m} \beta_{1,j} \Delta_h \ln(ip)_{t-f(h)} + \sum_{j=1}^{m} \beta_{2,j} Term_{t-f(h)} \\
& + \sum_{j=1}^{m} \beta_{3,j} MktRet_{t-f(h)} + \sum_{j=1}^{m} \beta_{4,j} MktVar_{t-f(h)} \\
& + \sum_{j=1}^{m} \gamma_{1,j}^{i} Ret_{t-f(h)}^{i} + \sum_{j=1}^{m} \gamma_{2,j}^{i} Var_{t-f(h)}^{i} + \\
& + \varepsilon_t
\end{aligned} \qquad (1)
$$

with size(f(h))=m and where $h$, the forecast horizon, is equal to 6, 12, 18 or 24 months and $f(h) = (h+6, h+12, h+18)$ represents the lag structure chosen for the regressors; $ip_t$ is the Industrial Production index, $Term$ and $MktRet$ and $MktVar$ are - respectively - the term spread and the return and the variance of the overall stock market index. As said, $Ret^i$ and $Var^i$ are the return and the variance of the equity price of selected firms. Given the overlapping nature of the data, regressions are always corrected via a heteroskedasticity consistent Newey and West estimator based on a window of data which is a function of $h$. The choice for $f(h)$ made above is of course arbitrary in our regressions. We tried different combinations and always reached the conclusion that long lags are needed to significantly improve forecasts (see also, concerning this choice, the impulse responses of GDP to uncertainty shocks in Bloom, 2009). As said, the inclusion of the term spread and the stock market return and volatility in the regressors is motivated by the remarkable success of these variables reported in the literature (see Estrella, 2005; Fornari and Mele, 2009). Overall, they convey information about financial risk, economic risk premiums and monetary policy. During expansions, market participants exhibit increasing risk appetite and the the risk premiums for long term investments declines. For this reason, and also because monetary policy is typically counter-cyclical, the term spread is expected to be negatively correlated to the economic activity. Stock market volatility, on the other hand, conveys information about the riskiness of financial markets and, more generally, of the overall macroeconomic environment. A riskier environment typically leads firms to under-invest and under-hire (Bloom, 2009), ultimately affecting economic activity, so that higher stock market volatility is expected to lead to lower economic growth. Households are also typically found to postpone spending decisions at times of heightened uncertainty. The aggregate stock market return is included in the regressions mainly to filter out the part of a firm equity return that stems from its systematic co-movement with the market. In fact, what we look at are the idiosyncratic movements of the firms' returns and returns volatilities, relative to the market index. Basically, rather than pre-filtering firms' returns with the market return and firms' variances with the market variance, and having therefore to deal with the problems induced by generated regressors, we directly insert the aggregate market return and variance in the above equation (more on this aspect is in the Robustness section). We insert in the regressions information about one firm at a time, so to assess the significance and extent of every marginal piece of information added by individual firms. We carry out the analysis both at a purely domestic level, i.e. considering how business cycles in the four economic areas we analyze are anticipated by the set of domestic firms only, and at the international level, i.e. consider cross-country interactions. In this way we can ascertain the extent in which global information dominates/is dominated by domestic information. To anticipate, we find that foreign firms can anticipate domestic real developments but on average they can do no better than domestic firms. However, we also show that the relative

| Sector | US | UK | JP | EA | Tot |
|---|---|---|---|---|---|
| 1) Oil & Gas | 25 | 3 | 5 | 2 | 35 |
| 2) Basic Materials | 20 | 6 | 71 | 8 | 105 |
| 3) Industrials | 68 | 47 | 127 | 29 | 271 |
| 4) Consumer Goods | 42 | 17 | 91 | 19 | 169 |
| 5) Health Care | 17 | 2 | 19 | 1 | 39 |
| 6) Consumer Services | 31 | 25 | 29 | 12 | 97 |
| 7) Telecommunications | 4 | - | - | 1 | 5 |
| 8) Utilities | 42 | - | 14 | 5 | 61 |
| 9) Financials | 40 | 64 | 29 | 23 | 156 |
| 10) Technology | 17 | 2 | 12 | 1 | 32 |
| Total | 306 | 166 | 397 | 101 | 970 |

Table 1: Firms distribution across sectors and countries.

importance of domestic and foreign firms in affecting real developments relates to the business cycle position of the domestic country.

# 3 Data and In Sample Evidence

The firm level information that we consider comes from the equity prices of a large set of firms sharing the following characteristics: i) they are based in the United States, the United Kingdom, Japan, France, Germany or Italy; ii) they have been continuously listed in the respective stock exchanges since 1973, the first year for which Thomson Reuters Datastream provides historical data. This results in a set of $n = 970$ firms. They belong to all the industrial sectors of the respective economies, i.e. we do not rule out any sector, a priori, so to maximise our chances of detecting firms with high forecasting power. We consider the so-called level-3 Industry Classification Benchmark - the standard company classification system developed by Dow Jones and FTSE - i.e. a 10-sector classification. Table 1 provides a brief description of the sectoral structure in the dataset. There is a large cross country heterogeneity as for relative sectoral weights, with the industrial and the financial sectors standing out as the most represented. For each firm we collect the daily stock prices and build the end-month realized returns and realized volatilities over various horizons (6, 12, 18 and 24 months) between January 1973 and December 2009. The use of realized volatilities builds on the large literature initiated by Andersen et al. (2003), and basically uses sums of daily absolute equity returns computed within each calendar month. It is important to highlight here that the firms we look at certainly suffer from a survivorship bias. However, considering just the pure forecasting exercise, we could only improve upon the results we present in this paper by considering additional firms. On the other side, we could miss some factors when trying to provide a structural explanation to our forecasting results. For example, we could miss the fact that predictability increases or decreases when default risk reaches critical values, a thing which most likely occurs for the firms which are likely to be about to leave the aggregate index we look at.

As said before, the industrial production index (IP) is our measure of real activity in the selected countries. We also collect daily composite stock market indices, from which we compute end-month realized returns and realized variances (in the same way as for individual firms), as well as the term spread (the difference between the ten-year government bond yield and the three-month T-bills or eurodeposit rate). For convenience we aggregate French, German and Italian series into corresponding *pseudo* euro-area series via weighted averages, with fixed weights based on 1999 GDPs, thus ending up with four main economic areas.

To give a preliminary flavour of what kind of results we will get, we present here the unconditional relationship between real activity, aggregate information and firm level variables. In practice, we estimate model (1) throughout the whole sample and look at the firms' performance as summarized by the regressions' corrected $R^2$ in Figure 1.[6] For each economic area, the horizontal line represents the corrected $R^2$ from the regression of the year-on-year growth rate of the Industrial Production on aggregate information only (lagged IP growth, term spread, aggregate stock market return and variance). The other (downwards sloping) lines in the figure depict the (sorted) corrected $R^2$ after the inclusion of firm level returns and variances, one firm at a time, on top of aggregate information. A quite remarkable feature is that nearly every firm can increase the predictability of real developments relative to aggregate variables. In general, returns seem to be slightly more powerful than variances and when the two variables are jointly included in the model the $R^2$ is, on average, some 50% higher than what provided by aggregate variables only (unreported results confirm the consistency of such findings across different forecast horizons). These unconditional results somewhat anticipate the extent in which firm-level information can improve the predictability of business cycle developments, although important information as the changing role of the firms across the cycle as well as a proper consideration of the data mining issue requires these findings to be confirmed by an out-of-sample exercise, which we tackle in the next section.

# 4 Firm Level Information and Business Cycle Predictability

## 4.1 Concentration in Predictive Power

We measure the amount of business cycle predictability that is associated to firm-level information, for each of the four economic areas that we consider, via out-of-sample predictions of the growth rate of the Industrial Production index through equation (1). For each month between June 1985 and December 2009 (the in sample regression goes back to January 1973) we estimate model (1) over ten-year windows (always using one equation for each forecasting horizon, i.e. a direct forecasting approach rather than iterated forecasting) and make predictions for the IP growth rate over the subsequent 6, 12, 18 or 24 months. We run these regressions for all the n=970 firms but we keep results for domestic and foreign firms separated.

Important for understanding our results, when we report the results of these regressions we switch from the firm-level standpoint (how the predictive power of a given firm evolves across time) to what we call a model-level standpoint, i.e. we aggregate firm-level results that are relatively close each other, over time, into a *model*. To do this we need a criterion to rank the *local* performances of the n=970 firms in each month. Standing for example in month $t-h$ we rank the firms according to their

[6] We report corrected $R^2$ coefficients but the difference in regressors between the specification with aggregate information only and with aggregate and firm-specific information is not particularly large as only one firm at a time is considered and the additional variables are only two with three lags each, quite a minor difference with more than 400 observations.

forecasting performance, for the specific horizon $h$, recorded over the previous 6 months, as measured by the RMSE.[7] Before ranking the models in this way we need to make sure that this *backwards looking* RMSE (as said computed between $t-h-6$ and $t-h$ when standing in $t$) is strongly correlated with the actual predictive ability of the firms in $t$, for any given forecast horizon $h$ (of course, the forecasting performance of the firms in month $t$ will be only known only ex-post). We do not report these results in order to save space but we indeed find an almost one-to-one relationship between the backward looking RMSE and the subsequent actual predictive ability for almost all the firms. The presence of short-term persistence in the predictive power for future IP developments at the firm level is therefore key in allowing us to identify the firms which in a given point in time are more likely to have high predictive power over the subsequent few months. We stress also that the computation of the 6-month backwards RMSE is obviously irrelevant to the aim of producing the actual forecasts. It only has the role of providing a criterion to aggregate, in each given month, the many firm-level based forecasts of future IP growth rates into *models*.

Based on this backwards looking measure of RMSE and abstracting for the moment from the actual forecasting power exhibited by the firms, Figure 2 (black solid line) shows how many times each of the domestic firms shows up in the first decile of the predictive distribution for the respective country's economic activity. It also compares the actual occurrences to a confidence interval for those that we would see if firms were instead randomly selected through uniform odds, both cross sectionally and across time (i.e. we compare the actual number of times a given firm shows up in the top decile with the number that could be expected if all firms had the same chance to be extracted in any given month).[8] We find that two small sets of firms are significantly different from all the others. The first set includes firms that show up too few times relative to a random selection, while the second set comprises firms that show up too many times relative to this benchmark. We may less formally find the same result browsing through the names of the firms which ranked in the top ten positions of the predictive distribution, as firms' names tend to remain stable, on average, for a relatively large number of months. A snapshot for the period between September 2009 and August 2010 is reported in Table 2. The existence of short-term persistence in the relative predictive power supports the view that some firm are different from others and that randomness does not represent the main driver of our results (see Section 6 for additional support). It can be noticed from the Table that once a firm begins to exhibit high predictive power for developments in economic activity (i.e. it belongs to the top 10 firms in terms of predictability), it continues to do so for around six months, before beginning to lose importance. Overall, these *special* firms are around one tenth of the population of domestic firms in each country. Such findings would support a *granular* interpretation of aggregate fluctuations, with aggregate economic activity dynamics being embedded in the *grains* represented by the small set of highly predictive firms that we have identified.

## 4.2 Domestic Predictability and Spillovers

Figure 3 shows the actual RMSE split across domestic and foreign firms. Each RMSE value refers to a given *model*, as explained in the previous sub-section. For example, model number $w$ is the

[7]The choice of 6 months is of course arbitrary. Ideally we would like to have a measure of instantaneous fit and this is the main reason to choose 6 months. In this way we can reduce the complexity which we would encounter in choosing one forecast or a subset of forecasts out of the large number of forecasts that we produce (more than 900 for each month in each economic area when we look at the full set of domestic and foreign firms). In each month, the RMSE computed over a small number of previous months could be employed also to produce a weighted pooling of the individual forecasts, similarly to the log-score criterion used in, among others, Amisano and Geweke (2009).

[8]Details available on request.

| $1^{st}$ | $2^{nd}$ | $3^{rd}$ | $4^{th}$ | $5^{th}$ | $6^{th}$ | $7^{th}$ | $8^{th}$ | $9^{th}$ | $10^{th}$ |
|---|---|---|---|---|---|---|---|---|---|
| SUNOCO | DOMTAR | PREC.CAST | APPLIED MATS. | NATIONAL | ARCHER | INTEL | ASHLAND | RPM INTL. | MDU |
| SUNOCO | DOMTAR | ARCHER | PREC.CAST | APPLIED MATS. | MDU | NATIONAL | SCHLUMBERGER | SPX | RPM INTL. |
| SUNOCO | DOMTAR | MDU | ARCHER | SCHLUMBERGER | APPLIED MATS. | NATIONAL | JOHNSON | PREC.CAST | AQUA |
| MDU | SUNOCO | ARCHER | DOMTAR | SCHLUMBERGER | EXXON | EATON | AQUA | ALCOA | NATIONAL |
| MDU | SUNOCO | ARCHER | DOMTAR | SCHLUMBERGER | EXXON | AQUA | PERKINELMER | ALCOA | EATON |
| MDU | PERKINELMER | SUNOCO | HESS | AQUA AMERICA | JACOBS | SPX | SCHLUMBERGER | DOMTAR | ALCOA |
| MDU | PERKINELMER | HESS | MOTOROLA | SUNOCO | AQUA | VALMONT | SPX | SCHLUMBERGER | VULCAN |
| MDU | PERKINELMER | VALMONT | MOTOROLA | HESS | VULCAN | AQUA | SUNOCO | SPX | SCHLUMBERGER |
| MDU | VALMONT | PERKINELMER | VULCAN | MOTOROLA | RADIOSHACK | ASHLAND | AQUA | HESS | HELMERICH |
| VALMONT | VULCAN | MDU | PERKINELMER | RADIOSHACK | MOTOROLA | HELMERICH | ASHLAND | AQUA | HESS |
| VALMONT | PERKINELMER | VULCAN | MDU | RADIOSHACK | ASHLAND | MOTOROLA | HELMERICH | HESS | SKYWORKS |
| MDU | VALMONT | PERKINELMER | ASHLAND | HELMERICH | RADIOSHACK | SKYWORKS | SCHLUMBERGER | AQUA | SPX |
| BSS | HANSA TRUST 'A' | LAND SECURITIES | HUNTING | GREENE KING | SPIRAX-SARCO | BROWN (N) | CHEMRING | PZ CUSSONS | FINDEL |
| BSS | HANSA | GREENE KING | CHEMRING | HUNTING | SPIRAX-SARCO | LAND SECURITIES | PZ CUSSONS | HELICAL BAR | CALEDONIA |
| BSS | HANSA | CALEDONIA | MUCKLOW | BARR (AG) | GREENE KING | CHEMRING | HUNTING | HELICAL BAR | SPIRAX-SARCO |
| HELICAL BAR | BARR (AG) | BSS | CHEMRING | CALEDONIA | HANSA | GREENE KING | PZ CUSSONS | SPIRAX-SARCO | LAND SECURITIES |
| HELICAL BAR | BARR (AG) | CALEDONIA | BSS | JPMORGAN M.C.I | CHEMRING | GREENE KING | HANSA | SPIRAX-SARCO | GREAT PORTLAND |
| HELICAL BAR | BARR (AG) | CALEDONIA | JPMORGAN M.C.I | GREAT PORTLAND | MENZIES (JOHN) | BSS | CHEMRING | HANSA | RESTAURANT |
| HELICAL BAR | BARR (AG) | MENZIES (JOHN) | BSS | CALEDONIA | HANSA | GREAT PORTLAND | MUCKLOW | RESTAURANT | CHEMRING |
| HELICAL BAR | BARR (AG) | MENZIES (JOHN) | BSS | GLOBAL SMALLER. | HANSA | MUCKLOW | BRITISH EMPIRE | CHEMRING | CALEDONIA |
| HELICAL BAR | BARR (AG) | MENZIES (JOHN) | BSS | GLOBAL SMALLER. | HANSA | MUCKLOW | BRITISH EMPIRE | CHEMRING | CALEDONIA |
| HELICAL BAR | BARR (AG) | MENZIES (JOHN) | BSS | GLOBAL SMALLER. | MUCKLOW | HANSA | BRITISH EMPIRE | DAEJAN | CHEMRING |
| HELICAL BAR | BARR (AG) | DAEJAN | MENZIES (JOHN) | GLOBAL SMALLER. | BSS | MUCKLOW | HANSA | BRITISH EMPIRE | CHEMRING |
| HELICAL BAR | DAEJAN | BARR (AG) | BSS | MENZIES (JOHN) | GLOBAL SMALLER. | MUCKLOW | HANSA | BRITISH EMPIRE | CHEMRING |
| WUESTENROT | INTESA SANPAOLO | THYSSENKRUPP | DYCKERHOFF | PIRELLI | DYCKERHOFF PREF. | K + S | SCA HYGIENE | BANCA FINNAT | UNICREDIT |
| WUESTENROT | INTESA SANPAOLO | PIRELLI | DYCKERHOFF | THYSSENKRUPP | DYCKERHOFF PREF. | K + S | UNICREDIT | SCA HYGIENE | EURAZEO |
| DYCKERHOFF | PIRELLI | GILDEMEISTER | WUESTENROT | INTESA SANPAOLO | DYCKERHOFF PREF. | LECHWERKE | K + S | THYSSENKRUPP | UNICREDIT |
| CICCOLELLA | OLDENBURGISCHE | DYCKERHOFF | PIRELLI | K + S | WUESTENROT | THYSSENKRUPP | DYCKERHOFF PREF. | INTESA SANPAOLO | BASF |
| CICCOLELLA | OLDENBURGISCHE | DYCKERHOFF | PIRELLI | BASF | UNI LAND | THYSSENKRUPP | GENERALI | INTESA SANPAOLO | WUESTENROT |
| CICCOLELLA | OLDENBURGISCHE | BASF | PIRELLI | PPR | THYSSENKRUPP | GENERALI | DYCKERHOFF | UNI LAND | INTESA SANPAOLO |
| OLDENBURGISCHE | CICCOLELLA | PPR | PIRELLI | UNI LAND | BASF | THYSSENKRUPP | CLUB MEDITERRANEE | GENERALI | DASSAULT |
| OLDENBURGISCHE | CICCOLELLA | PPR | BASF | UNI LAND | GENERALI | DASSAULT | CLUB MEDITERRANEE | THYSSENKRUPP | PIRELLI |
| OLDENBURGISCHE | CICCOLELLA | PPR | UNI LAND | BASF | DASSAULT | GENERALI | CUSTODIA | THYSSENKRUPP | PIRELLI |
| OLDENBURGISCHE | CICCOLELLA | UNI LAND | CUSTODIA | PPR | DASSAULT | THYSSENKRUPP | GEA | BASF | GENERALI |
| OLDENBURGISCHE | CICCOLELLA | UNI LAND | CUSTODIA | PPR | GEA | PILKINGTON | DASSAULT | BASF | GILDEMEISTER |
| OLDENBURGISCHE | CICCOLELLA | CUSTODIA | UNI LAND | PILKINGTON | PPR | GEA | DASSAULT | BASF | GILDEMEISTER |
| TOHO | TOBU | SOMPO | HOKUETSU KISHU | KAO | SUMITOMO | KEIHIN | OJI PAPER | NIPPON EXPRESS | NEC |
| KEIHIN | TOBU | NEC | SOMPO | NOMURA | MITSUI | HOKUETSU | SANKYO-TATEYAMA | NAGOYA | SUMITOMO |
| MAKINO | MITSUI | NOMURA | NEC | HITACHI CHEMICAL | SUMITOMO | SUMITOMO | SOMPO | TOKYO ELECT | TOBU |
| MAKINO | SUMITOMO | MITSUI | NEC | SOMPO | NOMURA | HITACHI | SANKYO-TATEYAMA | MITSUI SUGAR | DAIDO STEEL |
| MARUI | IBIDEN | SOMPO | KINDEN | ORIX | MITSUI SUGAR | OLYMPUS | DAIICHI SANKYO | SUMITOMO M. | FUJITA KANKO |
| FUJITA KANKO | MARUI | NIKKISO | HITACHI | MITSUI SUGAR | IBIDEN | ORIX | KINDEN | SOMPO | NAGOYA |
| HITACHI | FUJITA KANKO | NIKKISO | NAGOYA | MITSUI SUGAR | MARUI | DAIDO STEEL | HANKYU HANSHIN | FUJIKURA | KINDEN |
| HITACHI | NIKKISO | FUJITA KANKO | MITSUI SUGAR | TOKYOTOKEIBA | DAIWABO | TOHO ZINC | FUJIKURA | DAIDO STEEL | NAGOYA |
| HITACHI | NIKKISO | DAIWABO | TOHO ZINC | TOKYOTOKEIBA | MITSUI SUGAR | FUJITA KANKO | FUJIKURA | DAIDO STEEL | KEISEI ELEC. |
| HITACHI | NIKKISO | DAIWABO | TOHO ZINC | MITSUI SUGAR | TOKYOTOKEIBA | FUJITA KANKO | DAIDO STEEL | FUJIKURA | KEISEI ELEC. |
| HITACHI | NIKKISO | DAIWABO | TOHO ZINC | MITSUI SUGAR | TOKYOTOKEIBA | DAIDO STEEL | KEISEI ELEC. | FUJIKURA | FUJITA KANKO |
| HITACHI | DAIWABO | NIKKISO | TOHO ZINC | TOKYOTOKEIBA | MITSUI SUGAR | DAIDO STEEL | FUJIKURA | KEISEI ELEC. | AIRLINES |

Table 2: Top ten firms in the ranking of business cycle predictability. US (top block), UK, Euro Area and Japan (bottom block), from September 2009 to August 2010, 24-month ahead forecasts.

model that considers the IP forecasts originating from those firms that in each month ranked $w-th$ in the distribution of the *backwards looking* RMSE. Each of the models is therefore a collection of potentially a large number of firms over the long sample we examine. The regressions run for the foreign firms are identical to equation (1) but the firms we look at are only those that do not belong to the country under examination. More specifically, we assess the predictive power for the rates of change in the IP index in country $i$ coming from the equity prices and variances of firms in country $j$. Figure 3 shows that based on the RMSEs a large number of models induce a remarkable improvement in the forecasting performance relative to models looking at aggregate information only (the horizontal straight lines in the Figure). This is true especially of the euro area, although spillovers seem to be important in all the four economies, including the United States. While adding foreign firms can significantly improve the prediction of domestic business cycles over and above aggregate information, the best foreign model (i.e. the collection of the best foreign firms over time) has approximately the same predictive power as the best domestic model (i.e. the collection of the best domestic firms over time). Figure 4 reports a more formal assessment of the relative performance of the firms against aggregate information only and is based on the Diebold and Mariano (1995) test for equal predictive ability. The test is computed as the t-ratio of the constant in a regression of the difference between the absolute values of the error series produced by the competing models on a constant.[9]

Comparing Figure 2 in sub-section 4.1 to Figures 3 (RMSE) and 4 (Diebold and Mariano test) seems to suggest that the number of firms that outperform aggregate information is by and large overestimated in these latter two Figures.[10] While we deal with this data-snooping bias more formally in Section 6, we also notice here that one thing that may additionally bias the findings for foreign firms is that in constructing the RMSEs for these firms we do not consider that a part of their predictive power could derive from information common also to domestic firms. To control for the domestic component of the predictive power exhibited by foreign firms we should run a large number of regressions (around half a million with 970 firms) and therefore we explore a simpler but possibly less effective, alternative. In each month, and for each country, we compute through the backwards looking RMSE-based ranking of the domestic firms (introduced in section 4.1) two domestic factors (one return and one volatility factor), as simple averages of the return and the volatility of the first 10 firms in this ranking. Being built in real time, these factors should maximize predictability and therefore, as just explained, reduce significantly the information content of the foreign firms, should it be overlapping with domestic information. The specification of these regressions is analogous to eq. (1), i.e.:

[9]The Diebold and Mariano (DM) test is not suited for the comparison of nested models, as we have in this paper. In this case in fact the properties of the statistic collapse, as numerator and denominator are asymptotically the same. Clark and McCracken (2001) examine empirically the properties of the DM test when dealing with nested models and conclude that when the out-of-sample estimation is based on rolling windows, as we do, the test if still reliable, although modestly inferior to the test they propose, in which critical values have to be bootstrapped from the specific predictive regression employed. Based on this evidence we chose to continue to use the Diebold and Mariano test and the normal critical values, as computation times for the bootstrap would be extremely large with the sample size and the cross sectional dimension that we employ.

[10]The RMSE of the model including only the lagged change in the IP index is not far from the RMSE obtained including also aggregate financial information. A random walk forecast does instead much worse, as we consider rather long forecasting horizons. We do not report results for the random walk in the text but for example for the United States it is only at the 12 month horizon that the random walk has approximately the same RMSE of the model that uses only lagged changes in the IP. At the same time its RMSE remains much higher than the for the model that employs the slope of the yield curve and the variance of the stock market. Its performance at all three remaining forecast horizons is much worse than the other models that use only aggregate information and therefore also of the best models that consider firm-specific variables.

$$
\begin{aligned}
\Delta_h \ln(ip)_t = \alpha &+ \sum_{j=1}^{m} \beta_{1,j} \Delta_h \ln(ip)_{t-f(h)} \\
&+ \sum_{j=1}^{m} \beta_{2,j} Term_{t-f(h)} + \sum_{j=1}^{m} \beta_{3,j} MktRet_{t-f(h)} \\
&+ \sum_{j=1}^{m} \beta_{4,j} MktVar_{t-f(h)} \\
&+ \sum_{j=1}^{m} \gamma^{i}_{1,j} Ret^{i}_{t-f(h)} + \sum_{j=1}^{m} \gamma^{i}_{2,j} Var^{i}_{t-f(h)} + \\
&+ \sum_{j=1}^{m} \phi_j DRet_{t-f(h)} + \sum_{j=1}^{m} {}_j DVar_{t-f(h)} + \varepsilon_t
\end{aligned}
\tag{2}
$$

where $DRet$ and $DVar$ are the domestic return and variance factors built as just described and the other variables are the same as defined in equation (1). We overall find (results are just described to save space) that the RMSEs produced by foreign firms and reported in Figure 3 do not change much when domestic factors are also considered, therefore supporting the view that foreign firms can provide significant information beyond domestic aggregate information. At the same time, however, as already mentioned, Figures 3 and 4 also evidence that foreign firms hardly provide information in addition to the best domestic firms (i.e. the blue line - foreign firms - is never below the black line - domestic firms - for the firms providing the lowest RMSEs, those in the left hand side of the Figure). This consideration however applies to the average month in the sample. In fact, as we show in the next subsection, the monthly ranking of the firms evidences significant changes over time in the relative importance of domestic and foreign firms.

## 4.3 Country patterns

Figure 5 shows the country breakdown for the firms that rank in the first decile of the predictive distribution for each of the four economic areas. These breakdowns are generated through model (2), i.e. controlling foreign firms for the information already embodied in domestic factors[11]. Unlike what sample averages suggest, the relative weights of the foreign countries - for each given domestic country - recorded significant variations through time. The swings in the foreign weights have at times been common (so that a global factor seems to have been influencing real developments in all the areas analyzed) while in other periods some specific country has tended to gain relatively more weight. To mention a few interesting cases it is worthwhile looking at Japan, where domestic firms were key to explain the recession recorded around the end of the 90s, but almost irrelevant to anticipate the recent episode, with euro area and UK firms having instead a much more relevant role. The 2001 US recessions could have been anticipated almost equally likely by looking at US, UK or euro area firms; however, standing in December 2006, the US recession that would have started twelve months later could have been anticipated more by UK firms than by US firms. Identifying the reasons behind the observed changes in the relative weights of the countries goes beyond the aim of this paper (but

[11] As explained in the previous sub-section, these factors are the average return and its variance computed for the ten domestic firms that in a given month have exhibited the highest predictive ability

see section 5.2 for some attempts on this respect with reference for the domestic firms in the United States only).

With the aim of analyzing the different information about future economic developments that are conveyed by different firms, Figures 6 and 7 present range forecasts for the US and euro area 12-month growth rate in the IP index that come from different types of firms, i.e. those in the top 20 percent and bottom 20 percent of the predictive distribution (i.e. from top and bottom quintiles). Forecast ranges for IP are built - for each quintile - by taking the forecasts provided by the top and bottom 10 percent of firms, respectively, within the quintile. We present the forecasts ranges split by domestic and foreign firms, each of them being also computed as just described. Results are presented for the 12-month horizon only, but they are broadly similar for the remaining three forecast horizons and are not reported to save space. Two things are worth evidencing. First, there exist large differences in the fit provided by firms in the first and fifth quintiles, for both the United States and the euro area. While we are not surprised by the different performances of the two sets of firms, as they have been selected precisely according to a forecasting accuracy criterion, still it seems remarkable that such different forecasting performances can be achieved through firms belonging to the same stock market index. Second, looking at the top quintile (top panel in both Figures) the confidence intervals for the IP forecast (at the 12-month horizon) are nearly always including the actual rate of growth in the IP index and are rather tight around it, although with some time variation reflecting changes in the volatility of the IP growth rates. Overall there are no big differences across the confidence intervals provided by domestic and foreign firms in the top quintile of the predictive distribution, while predictive power is slightly higher for foreign firms in the bottom quintile, although this is largely irrelevant to the aim of selecting good forecasts. Visual inspection of the forecast ranges associated to firms in the bottom quintile evidences instead large errors throughout the majority of the sample as well as a bigger width of the confidence intervals relative to those produced by the firms in the top quintile.

One last thing to be pointed out is that focusing on the comparison of the IP forecasts across domestic and foreign firms or between firms in the top and the bottom quintile hides the gain in predictive power that the firm-level information provides relative to aggregate information. Figure 8 shows the actual values of the 12-month rate of change in the US IP index over the out-of-sample period (red line), along with its forecast based on aggregate information (black line) and the range forecast (as said above the 10-th and 90-th percentile inside the top quintile) coming from the top quintile of the firms (blue lines). Focusing for example on the 1990/1991 recession, one can see that aggregate information was broadly irrelevant to anticipate the coming slowdown in activity. At the same time, a significant number of models (i.e. those in the top decile of (the top quintile) of firms) could anticipate it rather well. Similar episodes can be detected also in phases of positive economic growth as well as in the other US two recession episodes included in the out-of-sample period.

# 5 Characteristics of the Firms and Predictability

## 5.1 Sectoral Patterns

Having established that the returns and the variances of domestic and foreign equity prices boost the predictability of aggregate fluctuations at various horizons, we look more in detail at how the predictive power is split across sectors within a given country. As the message conveyed by the joint observation of Figures 2-4 is that only a relatively small number of firms (domestic and foreign) can sizeably improve business cycle forecasting, asking ourselves whether these firms are special due to the particular sector in which they operate is quite a straightforward question. The bottom line here is that the conditional standpoint is again key to unveil the existence of sectoral patterns.

Figure 9 reports the weight of each of the ten sectors in the first and the last deciles of the predictive distribution for future IP changes averaged over time, along with the corresponding actual weight of the sector in the sample. From this unconditional perspective no sector shows up in the top and bottom deciles with a different proportion than it has in the sample. In other words, each of the ten sectors can successfully forecast business cycle fluctuations proportionally to its weight in the sample. As anticipated, instead, once a conditional standpoint is taken, the sector to which firms belong emerges as an important feature of their forecasting performance, with sectors having firms with returns and variances that tend to anticipate recessions but not expansions while other sectors have firms that display more power in anticipating expansions.

Figure 10 provides a description of the sectoral patterns evidenced within the (12-month ahead) predictive distribution of the firms, focusing for brevity on the largest sectors in the sample only. Financial firms seem to have some success in predicting recessions especially in the United States and in Japan and remarkably so for the last episode. The same sectoral pattern, however, is not evidenced around the burst of the dot-com bubble around late 1999, when especially firms within the Consumer Goods and Consumer Services sector were more predominant in anticipating real developments. As for other sectors a broad finding is that industrial firms seem to be good predictors of economic expansions while Consumer Goods are not strongly associated to the observed movements in the IP indices. [12]

## 5.2 Balance Sheet Items

The industrial sector, a proxy for the core business of the firms, is only one of the many variables which capture their characteristics, other choices being for instance the value of their assets, sales, revenues or debt, their size as measured for example by the number of employees and so on. To shed light on the importance of other key characteristics of the firms on their ability to anticipate business cycle developments we collect, from Worldscope, yearly data for a number of key balance sheet items over the longest available sample for each of the firms included in the regressions presented so far. These data have been retrieved for US firms only as the corresponding information is richer, between 1985 and 2009. We aim to identify a relationship between balance sheet items and economic activity by regressing the h-month rate of change in the US Industrial Production index (where, as before, h=6,12,18 and 24 months) on the same set of aggregate variables as in eq. (1) and on a single balance sheet item in turn, i.e.:

[12] As forecasts are made 12 months before the start of a recession, the returns and the volatilities recorded within some sectors relative to others seem to be able to capture signs of forthcoming changes in business cycle conditions.

$$
\begin{aligned}
\Delta_h \ln(ip)_t = & \\
& \alpha + \sum_{j=1}^{m} \beta_{1,j} \Delta_h \ln(ip)_{t-f(h)} \\
& + \sum_{j=1}^{m} \beta_{2,j} Term_{t-f(h)} \\
& + \sum_{j=1}^{m} \beta_{3,j} MktRet_{t-f(h)} \\
& + \sum_{j=1}^{m} \beta_{4,j} MktVar_{t-f(h)} \\
& + \gamma_1^i BS_{t-h}^{i,top} + \gamma_2^i BS_{t-h}^{i,bottom} + \varepsilon_t
\end{aligned} \tag{3}
$$

with i=1,...,34 balance sheet indicators. In this regression the terms BS measure the difference between the average value of the $i-th$ balance sheet item recorded by, respectively, the ten top and the ten bottom firms in the predictive distribution for future IP growth, as of the end of the previous calendar year (i.e when considering the IP growth recorded in March 2007, the difference in a given balance sheet item refers to 31 December 2005, if the 6-month horizon is examined). Although balance sheet items are available only yearly, the ranking of the firms as a function of their predictive power changes potentially in each month (see Table 2 for a snapshot referred to the United States) and therefore the balance sheet indicator records nonetheless a noticeable monthly variation.[13] Notice that the balance sheet items in these regressions are dated $t-h$ and as such they are lagged by as much as the forecast horizon, consistently also with the fact that it is based on the ranking of the firms made in $t-h-1$, when forecasting at the h-month ahead horizon. However, one could also suppose that some given firms are more able than others to capture future developments in business cycle conditions because of their balance sheet characteristics at time $t$, rather than at $t-h$. On this respect we also ran regressions (3) placing the balance sheet information at time $t$, i.e. as $BS_t$. Despite some changes in the size of the coefficients there did not seem to be particular variations in the significance pattern in Table 3 below.

Basically regression (3) allows us to verify whether gaps in balance sheet items across US firms, given the ranking of these firms in the predictive distribution for future US IP changes, can account for the difference in their predictive power. Results are reported in Table 3, where we grouped the significant balance sheet items into a few categories (Performance, Liquidity, Size, Investments, Foreign Activity, Inventories). The items displayed in the Table are only those (out of 34 selected items) for which either $\gamma_1$ or $\gamma_2$ (see equation (3)) were significant. It seems to be especially cross sectional divergences in items capturing Performance and Investments to be connected to subsequent real developments. Differences in Inventories and in the International Activity of the firms seems to be also able to anticipate business cycle developments. Measures of firm liquidity and indebtedness are also significant as well as some measures of size, i.e. employees and capitalization.

[13]Balance sheet data are at times missing for some firms or availability starts later than the beginning of our sample. Therefore there may not be a complete match between the IP predictions and the features of the top firms that have generated it.

| | | 6-month | 12-month | 18-month | 24-month |
|---|---|---|---|---|---|
| Performance | Book value per share | -0.44 | -4.46 | -11.73 | -10.57 |
| | Cash per share | -0.77 | -8.91 | -26.54 | -17.86 |
| | Sales per share | -0.61 | -3.71 | -3.95 | -2.16 |
| | Cost of goods to sales | -2.51 | 1.55 | 8.11 | 16.16 |
| | Assets per employee | 0.13 | 0.11 | 0.04 | 0.01 |
| Liquidity | Current ratio | -75.11 | -215.51 | -277.83 | -253.66 |
| | Quick ratio | -162.02 | -295.33 | -292.86 | -287.85 |
| Size | Employees | -0.025 | -0.022 | -0.0002 | 0.0015 |
| | Capitalization | 5.3*$10^{-6}$ | 1.15*$10^{-5}$ | 1.05*$10^{-5}$ | 8.65*$10^{-6}$ |
| Investments | Cap expenditures/sales | -19.88 | -35.77 | -14.66 | 7.94 |
| | Capl Expenditures / Total Assets | -29.31 | -37.28 | -13.05 | 33.42 |
| | Reinvstment Rate | - | - | - | 20.11 |
| | R&D | -0.71 | -75.24 | -117.19 | -111.49 |
| | Assets Turnover | -38.13 | 160.04 | 410.39 | 729.43 |
| | Other Invstments | 3.68*$10^{-5}$ | 4.64*$10^{-5}$ | 2.19*$10^{-5}$ | 4.14*$10^{-5}$ |
| Foreign Acitivty | Foreign/Total Assets | 5.71 | 7.81 | -0.85 | 11.78 |
| | Foreign/Total Income | -0.49 | 0.49 | 0.97 | 0.89 |
| | Foreign/Total Sales | 6.65 | 2.27 | -9.03 | -0.85 |
| Inventories | Inventory Turnover | 2.41 | -1.41 | 16.72 | 8.24 |
| | Inventory, days in | 2.26 | 1.03 | -2.64 | -3.38 |
| | Inventory/Total Assets | 10.49 | 25.08 | 26.15 | 27.64 |
| | Inventories | 3.68*$10^{-5}$ | -1.66*$10^{-4}$ | -2.0*$10^{-4}$ | -2.44*$10^{-4}$ |

Table 3: The Table reports the difference $\gamma_1^i$ - $\gamma_2^i$ as defined in equation (3), i.e the coefficients that come from a regression of the rate of growth of the US Industrial Production index, over $h$ months, where $h = 6, 12, 18$ and 24, on lagged values of the change in the IP index, lagged slope of the US term structure, lagged US aggregate stock market return and variance and the lagged gap in a given balance sheet item computed for the top 15 firms and bottom 15 firms. The differences in the $\gamma$ parameters have been reported only when at least one of the two coefficients was statistically different from zero. Regressions have been run for the period January 1985 - December 2009.

# 6 Robustness of results

In this section we try to i) to better understand the origin of the higher predictive power found for some firms and ii) to make sure that the advantages in looking at firm level information beyond aggregate series are indeed statistically significant and not driven by randomness. As for the first issue, having included both the aggregate stock market return and variance in regression (1) we capture firm-level information that reflects only the idiosyncratic movement of their equity prices. However, one could wonder about how firms that help predict future business cycle developments behave, relative to the market, when compared to firms that do not help in forecasting aggregate real developments. To shed light on this, for each given month in the out of sample period, i.e. 1985 - 2009, and with reference only to the top and bottom 30 firms in the predictive distribution for future IP changes, we perform the following exercise. We cast the equity return of each firm in turn and the aggregate stock market return into a bivariate garch(1,1) model, which is estimated via DCC (Cappiello et al., 2006)[14] over fixed-length windows of 60 months. In this bivariate model, the mean equations are specified so that each of the endogenous variables (the firm's return and the market return) include the first own lag as well as the first lag of the other variable. Once the DCC is estimated we store the firm's idiosyncratic variance, its correlation with the market return and its beta coefficient always relative to the market.[15] For each rolling sample, these three measures are stored in relation to the three lags employed in regression (1), i.e. the lags specified by the f(h) functions. The three measures are then aggregated over, respectively, the top and bottom 30 firms (identified in the previous sections through the *backwards looking* RMSE) so to have their averages across firms with high or low predictive power for future industrial production. Looking at the 12-month forecasting horizon (the remaining three horizons broadly provide the same picture), the differences in the three measures across the two groups of firms are significantly related to GDP developments and we evidence this via the following regression, where only the first lag for each of the three measures (12 months, as we are focusing on the 12-month predictive horizon) has been employed, as they were rather highly autocorrelated:

$$log\frac{IP_t}{IP_{t-12}} = \alpha_0 + \Sigma_{i=1,3}\alpha_j X_{j,t-12} + \varepsilon_t \tag{4}$$

where the vector X collects the three variables computed above (variances, correlations, betas) across the two groups of firms (top 30 and bottom 30).

$$log\frac{IP_t^{US}}{IP_{t-12}^{US}} = 0.018^{**} - 0.44^{**}DiffVar_{t-12} + 0.063^{**}DiffCorr_{t-12} + 0.004DiffBet_{t-12}$$

$$log\frac{IP_t^{UK}}{IP_{t-12}^{UK}} = 0.004^{**} - 0.37^{**}DiffVar_{t-12} - 0.035DiffCorr_{t-12} + 0.002DiffBet_{t-12}$$

$$log\frac{IP_t^{EA}}{IP_{t-12}^{EA}} = 0.007^{**} - 0.42^{*}DiffVar_{t-12} + 0.238^{**}DiffCorr_{t-12} - 0.110^{**}DiffBet_{t-12}$$

$$log\frac{IP_t^{JP}}{IP_{t-12}^{JP}} = 0.010^{**} - 0.82^{*}DiffVar_{t-12} + 0.023DiffCorr_{t-12} - 0.022DiffBet_{t-12}$$

In these regressions DiffVar is the difference in average idiosyncratic variances among the top 30 and the bottom 30 firms and DiffCor and DiffBet are the corresponding differences among the

[14]DCC stands for Dynamic Conditional Correlation and is a convenient and quick way to estimate a multivariate conditionally heteroskedastic model.

[15]The beta is typically used in finance to measure the sensitivity of an asset relative to the market.

correlations and the beta coefficients of the firms' idiosyncratic return with respect to the market return. The equations show that indeed aggregate business cycle developments between time $t-12$ and time $t$ are related to differences in the conditional variances of the top 30 and bottom 30 firms as measured at $t-12$, with a negative relationship prevailing in the four economic areas (i.e an highest variance of the top 30 firms relative to the bottom 30 firms tends to anticipate recessions). Differences in correlations are significant in the United States and in the euro area only, and differences in betas are much less significant. In a nutshell, these regressions suggest that it is not the existence of differences in the relationships with the market return, the betas or the correlations, to drive firms' predictive power. Rather, the latter seem to depend on the cross sectional gap among the firm-level idiosyncratic stock return volatilities.

As for the second point, the so-called data-snooping problem may seriously undermine the significance of the results presented in the previous sections. This issue was first raised by White (2000) and further addressed by Hansen (2005) via a test for *superior predictive ability* ($SPA$).[15] Hansen's (2005) improvement to the White (2000) reality check test has to do with the fact that the latter was shown to be negatively affected when a large number of models representing poor and irrelevant alternatives were added to the comparisons. In fact, adding useless models, $\frac{\alpha}{m}$, which is used as a significance threshold, where $\alpha$ is the chosen significance level, can be arbitrarily pushed towards zero.

The test for superior predictive ability is based on the relative performance of two models, defined as

$$d_{k,t} = L(\xi_t, \delta_{0,t-h}) - L(\xi_t, \delta_{k,t-h}) \tag{5}$$

where $k = 1, ..., m$, so that $d_{k,t}$ measures the performance of model $k$ relative to the benchmark at time $t$. When the pairwise comparisons between the $m$ models and the benchmark are collected into the vector $\mathbf{d}$, the null hypothesis can be cast as $H_0 : \mathbf{d} \leq 0$. As the derivation of the test assumes asymptotic normality for $\mathbf{d}$, then a quadratic form of the test could be employed but this is difficult to implement for large $m$. Therefore, only the diagonal elements of the covariance matrix $\Omega$ are considered and, due to this nuisance, a bootstrapped derivation of the test statistics must be adopted. In a nutshell the test statistics proposed is

$$T_n^{SPA} = \max k = 1, .., m[\frac{n^{0.5}\overline{d}_k}{\hat{w}_k}, 0] \tag{6}$$

where $\hat{w^2}$ is a consistent estimator of $var(n^{-0.5}\overline{d}_k)$. The null distribution is based on a mean computed as

$$\hat{\mu}^c = \frac{\overline{d}_k \mathbf{1} n^{-0.5}\overline{d}_k}{\hat{w}_k} \leq \sqrt{2\log\log(n)} \tag{7}$$

for $k = 1, 2, ...m$.

For our case (results are just discussed to save space) the p-values of the SPA test are ranging between p=0.01 and p=0.15 (depending on which of the three versions of the tests proposed in Hansen, 2005, is used). In any case, the test suggests that it is very likely that some firms indeed

[15]In principle, the Diebold and Mariano test (1995) (see Figure 4) should be able to tell whether a model provides or not the same predictive ability of another model for a given variable of interest. However, the DM test is derived under the null of equal predictive ability (EPA) while testing for superior predictive ability (SPA) is more complex. In fact EPA involves a simple null hypothesis while SPA leads to composite hypotheses and is known to involve asymptotic distributions which are affected by nuisance parameters (and a a result the null hypothesis is not unique).

have additional predictive power for future IP changes relative to aggregate information as well as to other firms and predictability does not seem to be driven by the randomness in the data.

# 7 Implications for Macroeconomic and Financial Stability

Beyond improving our ability to anticipate business cycle developments, could the same firm level information used so far be useful also to financial stability aims? In other words, can we derive macro-financial stability considerations from the cross sectional dispersion of the monthly IP forecasts for a given horizon $h$?

As explained in the previous sections, in each month and for each of the four economic areas, we generate almost 1000 IP forecasts, for each of the four horizons considered (h=6, 12, 18 and 24 months), based on aggregate financial information and firm level equity returns and volatilities. Once eliminated some outliers in the predicted IP growth rates, typically associated to exceptional declines or gains recorded by a few equity prices (for example trimming the data above and below pre-assigned thresholds, with plus and minus 70 percent monthly equity price changes having been used in the paper), we can for example i) compute the probability density function of the IP forecasts at horizon $h$ via a kernel estimator, ii) construct range forecasts associated to selected percentiles of this distribution and last iii) compute a monthly Value at Risk ($VaR_p$), with $p$ being a pre-specified p-value, for the IP growth rate. [16] Last, the returns and the variances of the firms that form a given model could be used[17] to $filter$ the time series of a number of unobserved leading indicators[18] for the business cycle at the selected horizon.

What is worth noticing in the construction of this economic VaR is that it does not come, as traditional VaR measures, from a simulation carried out via an single equation or a multivariate model describing the dynamic behavior of the IP index (see Manganelli and Engle, 2001, for the typical approaches to measuring Value at Risk, and De Nicolo' and Lucchetta, 2010, for a related model). Rather it exploits the cross sectional information of the equity returns, so that the uncertainty about future IP rates of change does not relate to the density of the past forecast errors in predicting such IP changes but rather from the current configuration of the cross sectional equity returns and variances. In principle a mixed approach could also be employed whereby, as $VaR$ is typically computed, future paths of the IP rate of change would be bootstrapped out of each firm-level regression and then these predictive densities would be averaged, either using equal weights or by means of the RMSE criterion that we have been using so far. Whether one strategy dominates over the other is an empirical issue that we do not tackle, to save space, in this paper.[19]

For illustrative purposes, Figure 11 reports the 5 percent Value at Risk (the negative rate of change in the IP index that can be observed with a 5 percent probability over the next 12 months) for the IP growth at the 12-month horizon for the four areas, while Figure 12 reports, for the United States only, the 12-month changes in the IP index as well as the 1- and 5-percent VaR. Also, Figure 13 reports, for the US only, four predictive density functions for IP growth over the subsequent 12

[16] As each individual IP forecast is associated to a RMSE criterion, the forecast ranges associated to each, say, quintile can be assigned a probability value. These can be used to assess the distance in the expected fit of the models.

[17] See Table 2 to recall how firms are related to models.

[18] Potentially there are as many as the number of firms in the cross section.

[19] Although some of the firms we identify provide the highest *local* predictability to subsequent developments in the IP indices, we cannot conclude that they are also systemically important firms, able to spread a given shock to the remaining firms in the economy (see for example the VAR for VaR described in White et al., 2011). Basically, we use individual equity prices just as filtering devices for a set of unobservable shocks, with no implications for the causal relationships among the set of firms.

months referred to dates spanning the last three recession episodes. For the 1990-1991 and the 2001 recessions the density correctly had a negative median as well as negative skewness for the central months of the recession episodes. For the last recession, the economic slowdown is instead captured with some delay (see also Figure 6), as the median of the density becomes negative only in December 2008, i.e. one year after the official start of the recession, as computed by the NBER. It has to be said, on this respect, that sharply negative GDP growth figures were recorded exactly in 2008Q4 and in 2009Q1 so that at least from the nowcasting standpoint the VaR predicted by the model would have nonetheless been useful.

# 8 Conclusions

This paper has shown, with reference to four economic areas and using a sample that for the out of sample exercise starts in 1985, that the idiosyncratic returns and variances of individual equity prices contain the seeds of future real developments. Importantly, the forecasting ability of a given firm is found to be persistent, averaging around six months, which leads to exclude that randomness is a driver of the forecasting power, as also confirmed by ad-hoc tests. Domestic firms as well foreign firms are successful in predicting domestic real developments but the relative weight of the two sets of firms depends significantly on the cyclical position of the domestic country. We do not find size, as Gabaix (2010) proposes, to be the key factor behind predictability. We unveil that some sectoral patterns are related to predictability as well as that some balance sheet items are related to the observed predictive power of the firms. However, we feel that additional efforts must be undertaken to shed more light on this challenging and promising issues.

# References

[1] G. Amisano and J. Geweke. Optimal prediction pools. *ECB Working Paper, No. 1017*, 2009.

[2] T. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Modelling and forecasting realized volatility. *Econometrica*, 71:529–626, 2003.

[3] S. B. Aruoba, F. X. Diebold, and C. Scotti. Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, 27:417–427, 2009.

[4] R. B. Barsky and E. Sims. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review*, forthcoming, 2010.

[5] P. Beaudry, M. Dupaigne, and F. Portier. Propagation of international news shocks. *Review of Economic Dynamics*, 14, 2011.

[6] N. Bloom. The impact of uncertainty shocks. *Econometrica*, 77:623–685, 2009.

[7] R. Cappiello, L., R.F. Engle, and K. Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4:537–572, 2006.

[8] V. Carvalho. Aggregate fluctuations and the network structure of intersectoral trade. *Working Paper CREI, Universitat Pompeu Fabra, Barcelona*, 2009.

[9] T. E. Clark and M. W. McCracken. Tests of equal forecast accuracy and ecompassing for nested models. *Journal of Econometrics*, 105:85–110, 2001.

[10] T. E. Clark and M. W. McCracken. Averaging forecasts from VARs with uncertain instabilities. *Federal Reserve Bank of Kansas City, Working Paper 06-12*, 2006.

[11] T. Conley and B. Dupor. A spatial analysis of sectoral complementarity. *Journal of Political Economy*, 111:311–352, 2003.

[12] G. De Nicolo' and M. Lucchetta. Systemic risks and the macroeconomy. *International Monetary Fund, Working Paper, No. 10/29*, 2010.

[13] F. X. Diebold and R. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–265, 2005.

[14] R. Espinoza, F. Fornari, and M. J. Lombardi. The role of financial variables in predicting economic activity. *Journal of Forecasting*, 2011.

[15] A. Estrella. Why does the yield curve predict output and inflation? *The Economic JOurnal*, 115:722–744, 2005.

[16] F. Fornari and A. Mele. Financial volatility and economic activity. *Working Paper*, 2009.

[17] X. Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79:733–772, 2010.

[18] D. Giannone, L. Reichlin, and D. Small. Nowcasting: The real time informational content of macroeconomic data releases. *Journal of Monetary Economics*, 55:665–676, 2008.

[19] S. Gilchrist, V. Yankov, and E. Zakrajsek. Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets. *Journal of Monetary Economics*, 56:471–493, 2009.

[20] White H. A reality check for data snooping. *Econometrica*, 68:1097–1127, 2000.

[21] White H., T-H. Kim, and S. Manganelli. A VAR for VaR: Measuring systemic risk using multivariate regression quantiles. *ECB Working Paper Series*, forthcoming, 2011.

[22] P. R. Hansen. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23:365–380, 2005.

[23] M. Horvath. Cyclicality and sectoral linkages: Aggregate fluctuations from sectoral stocks. *Review of Economic Dynamics*, 1:781–808, 1998.

[24] M. Horvath. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45:69–106, 2000.

[25] A-S. Jore, J. Mitchell, and S. P. Vaheyc. Combinign forecast densities from VARs with uncertain instabilities. *Norges Central Bank, Working Paper 08-01*, 2008.

[26] B. Jovanovic. Micro shocks and aggregate risk. *Quarterly Journal of Economics*, 102:395–409, 1987.

[27] S. Manganelli and R. F. Engle. Value at risk models in finance. *European Central Bank, Working Paper No.75*, 2001.

[28] J. Stock and M. W. Watson. Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature*, 41:788–829, 2003.

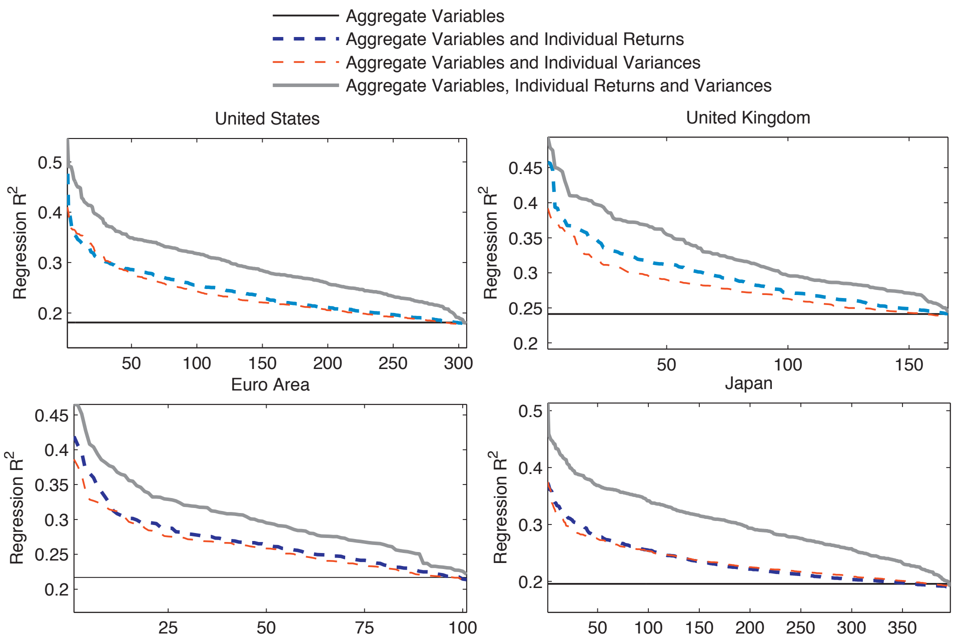



Figure 1: Ranked $R^2$ from the regression of the Industrial Production year on year growth rate on different combinations of lagged aggregate and individual variables, see equation (1). The forecasting horizon is 12 months. Data are monthly from January 1973 to December 2009. Aggregate variables include the lagged rate of change in the IP index, the term spread and the time varying volatility of the composite stock market index. The x-axis reports the number of firms in a given country.

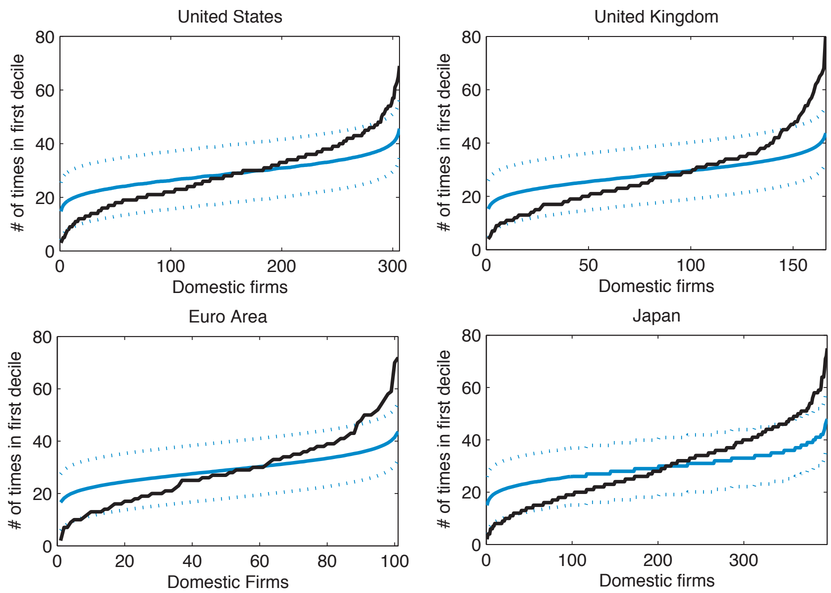



Figure 2: Actual ranked occurrences of each firm in the four economic areas in the first decile of the respective predictive distribution for subsequent changes in the IP index (black line) compared to a case in which all the firms have the same chances of having predictive (solid blue line, dotted lines are 95 percent confidence intervals).

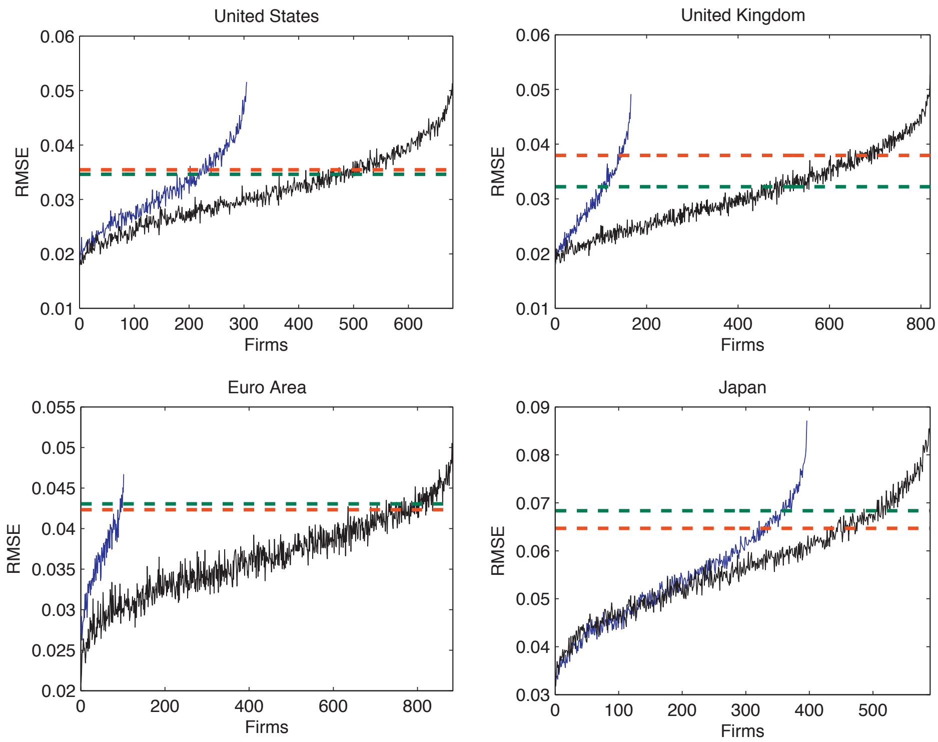



Figure 3: RMSEs of models conditioned upon domestic (blue) and non-domestic (black) firm level information. Horizontal lines refer to RMSE of models with only aggregate variables (lagged IP (red), lagged IP and term spread (green)). The forecast horizon is 12-month and the out of sample analysis refers to the period June 1985 - December 2009 and is based on rolling windows of equal length (10 years).

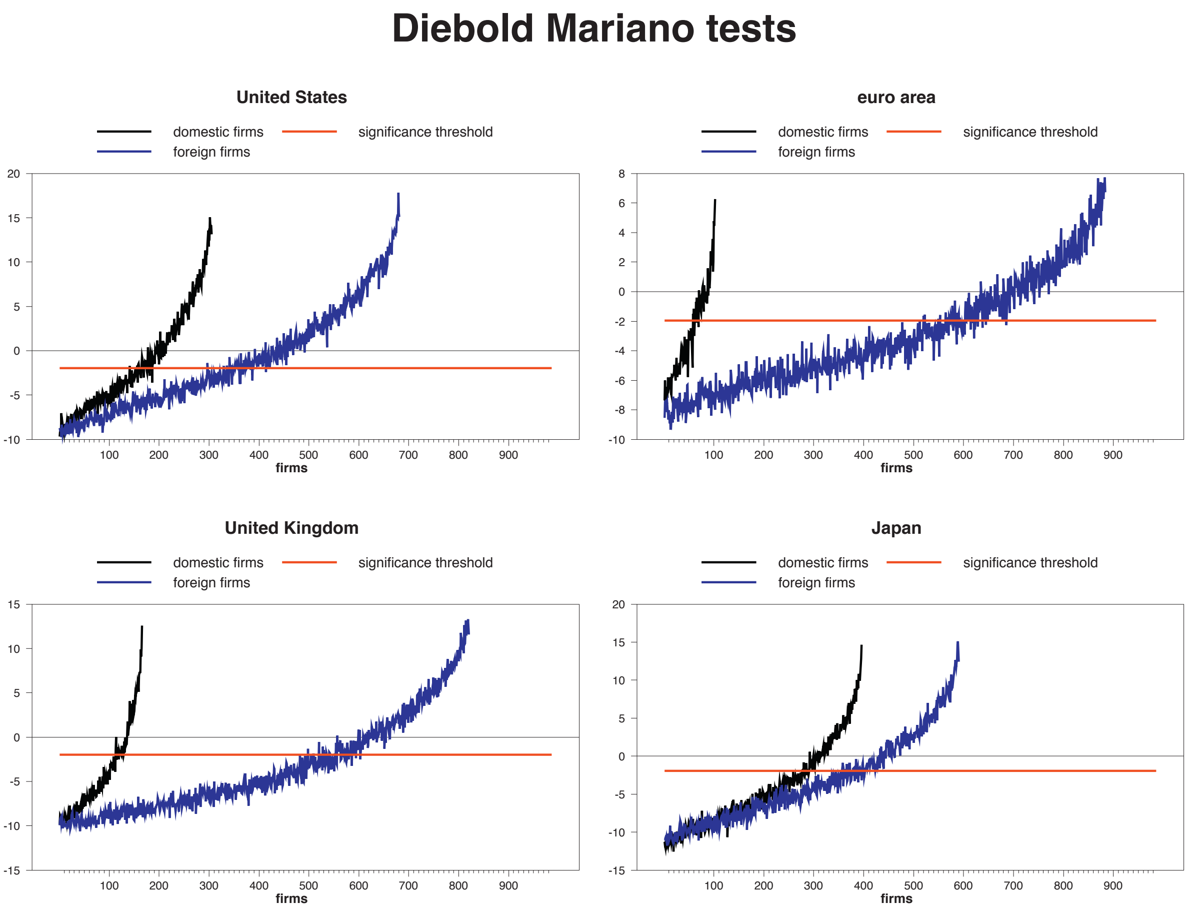



Figure 4: Diebold and Mariano tests of predictive ability of models including i) domestic (black) and ii) non-domestic (blue) firm level information, relative to aggregate domestic information only. The horizontal (red) line is the 5 percent significance threshold for the null of equal predictive ability. The out of sample forecasts are referred to the 12-month maturity and estimation is based on fixed-length windows of 10 years between June 1985 and December 2009.

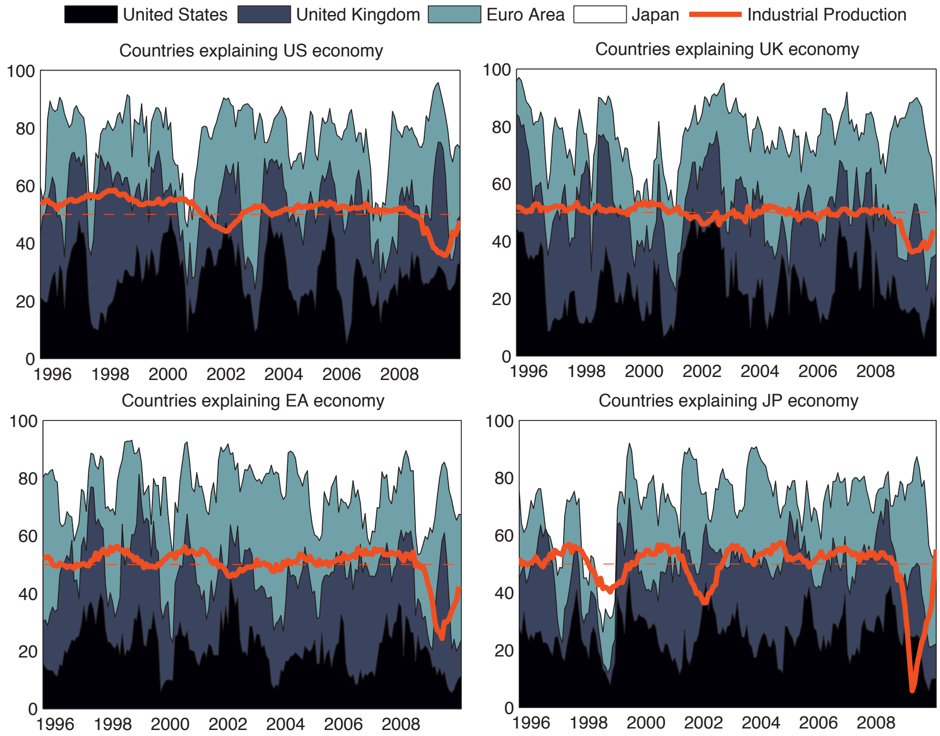
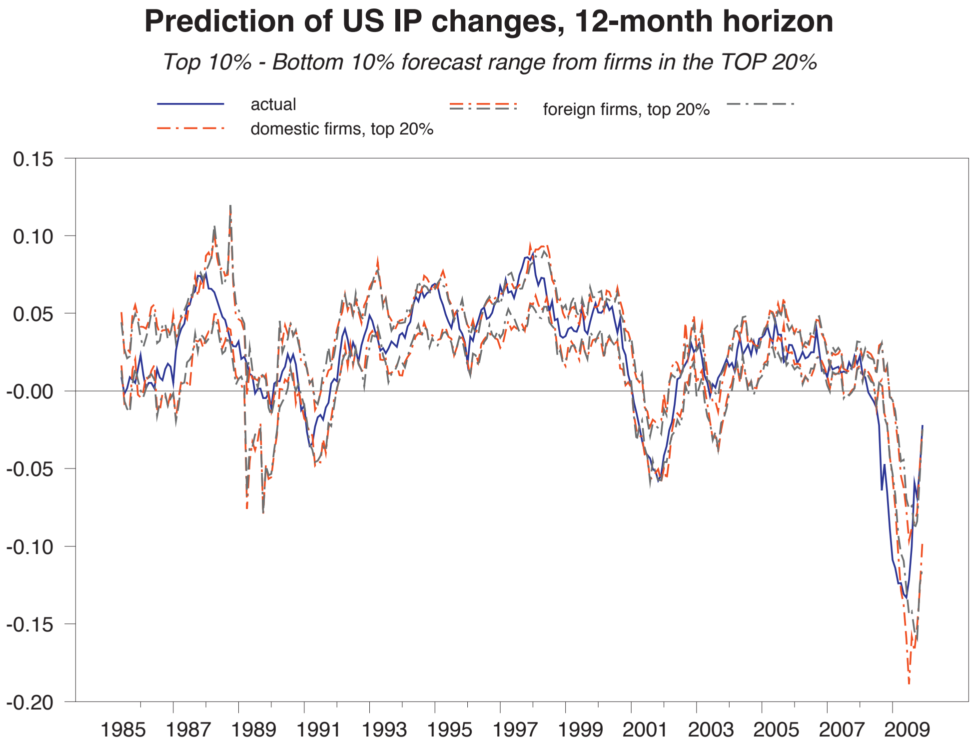



Figure 5: Composition by country of the firms in the first decile of the predictive distribution for 12-month ahead rates of change in the Industrial Production index. Data are monthly from July 1995 to December 2009.



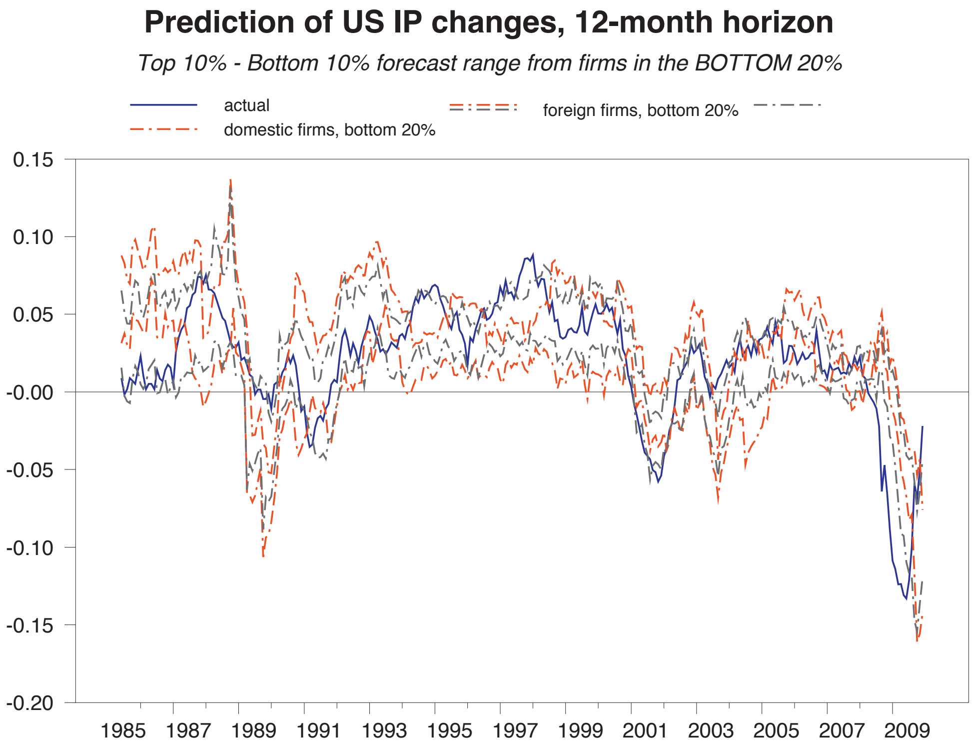
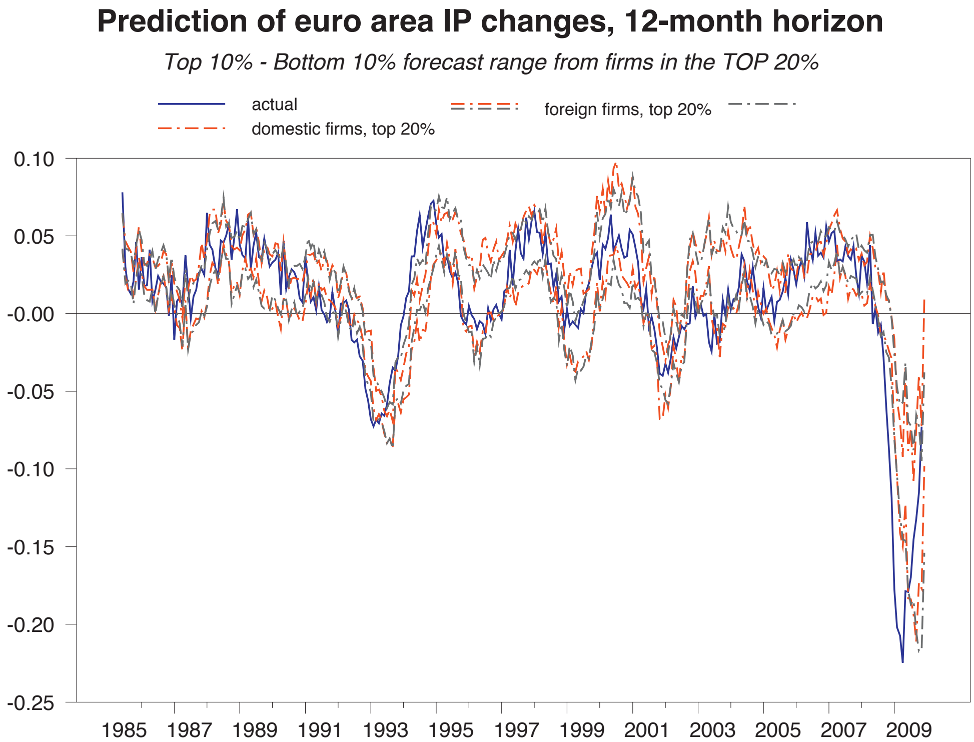




Figure 6: Range of the predictions for the 12-month ahead rate of change in the US IP index, using domestic firms (red) and foreign firms (grey) in the top and bottom 20 percent of the predictive distribution. Monthly data between June 1985 and December 2009.



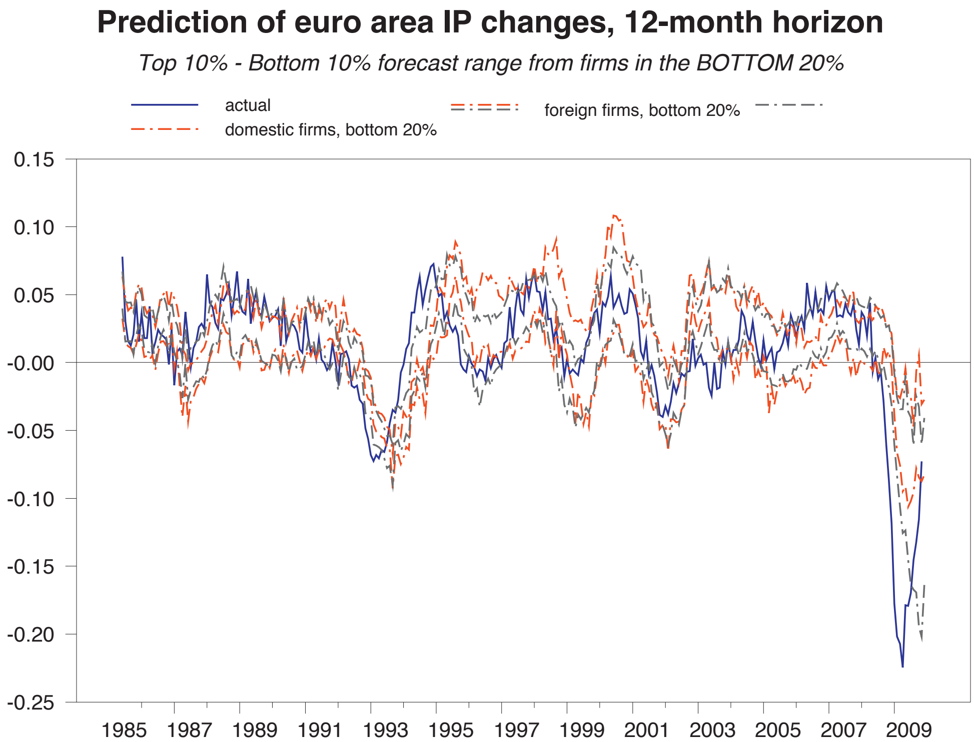




Figure 7: Range of the predictions for the 12-month ahead rate of change in the euro area IP index, using domestic firms (red) and foreign firms (grey) in the top and bottom 20 percent of the predictive distribution. Monthly data between June 1985 and December 2009.

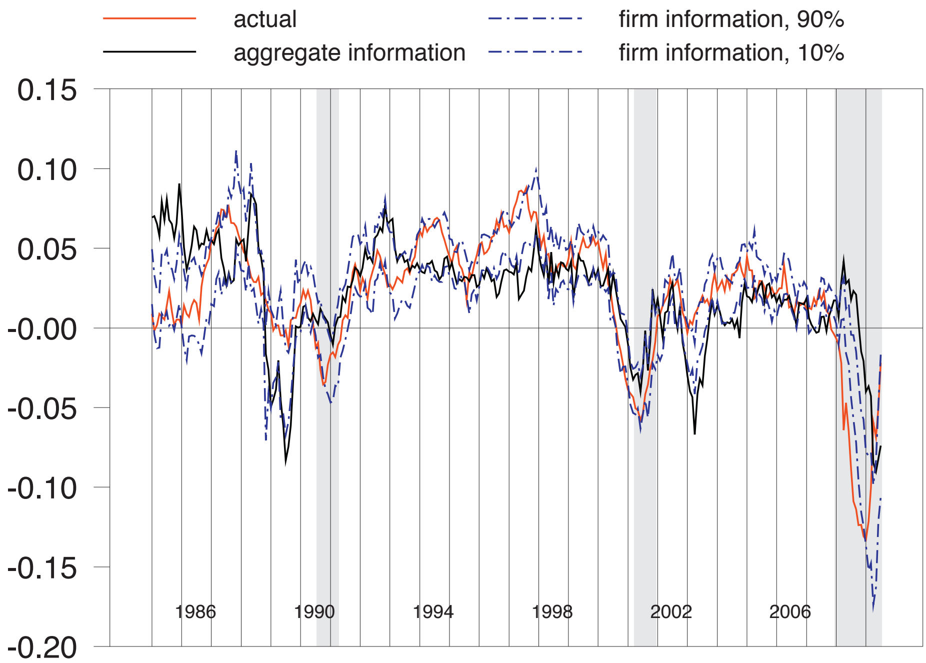



Figure 8: Actual values of the 12-month rate of change in the US IP index and alternative forecasts, Out of sample analysis between June 1985 to December 2009.

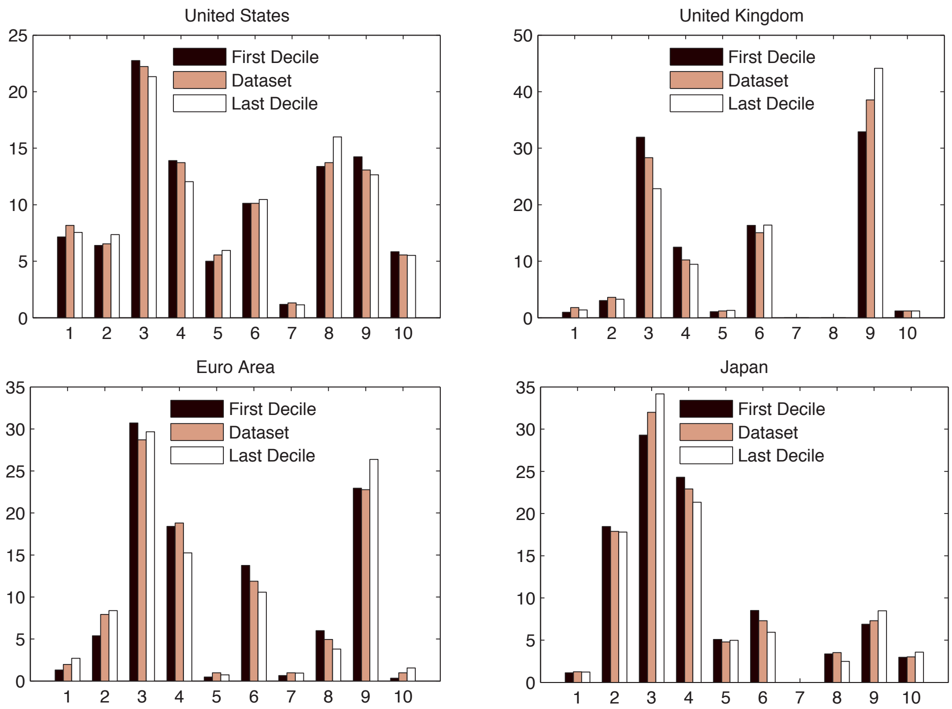



Figure 9: Average sectoral composition of the firms in the first and last deciles of the predictive distribution of 12-month ahead rates of change in the IP indices and corresponding frequencies of the same sectors in the full sample. Averages refer to the out-of-sample period June 1985 - December 2009.

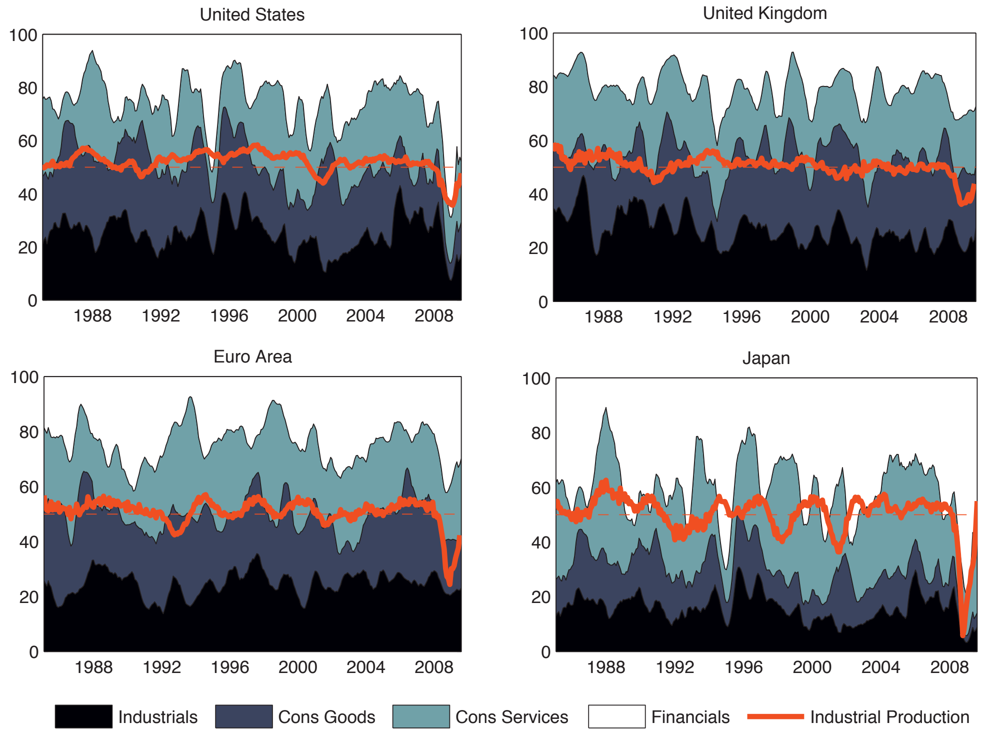



Figure 10: Relative frequency of selected sectors in the first decile of the 12-month ahead predictive distribution for the Industrial Production indices. Based on out of sample analysis between June 1985 and December 2009.

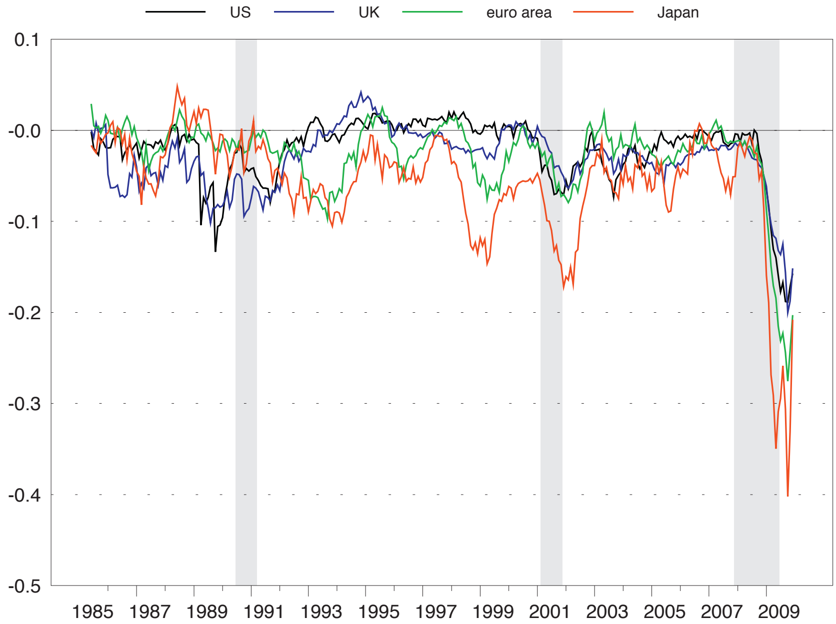



Figure 11: Five percent Value at Risk (VaR) for the year-on-year IP growth rates. Monthly data between June 1985 and December 2009.

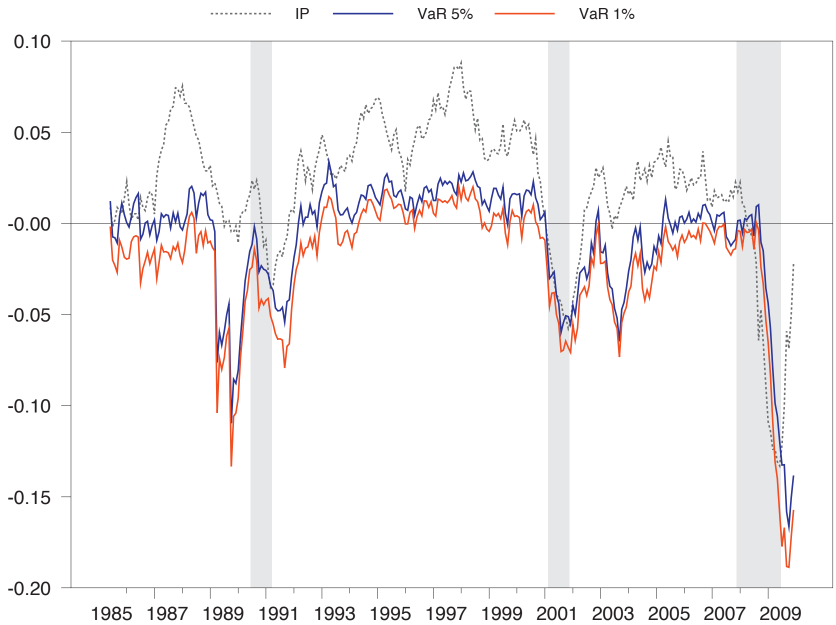



Figure 12: Actual year-on-year growth rate in the US IP index and corresponding 1- and 5-percent Value at Risk (VaR). Monthly data between June 1985 and December 2009. In both panels, the VaR reported for a given month is 'predicted' 12 months before that month. Shaded areas are US NBER-based recessions.

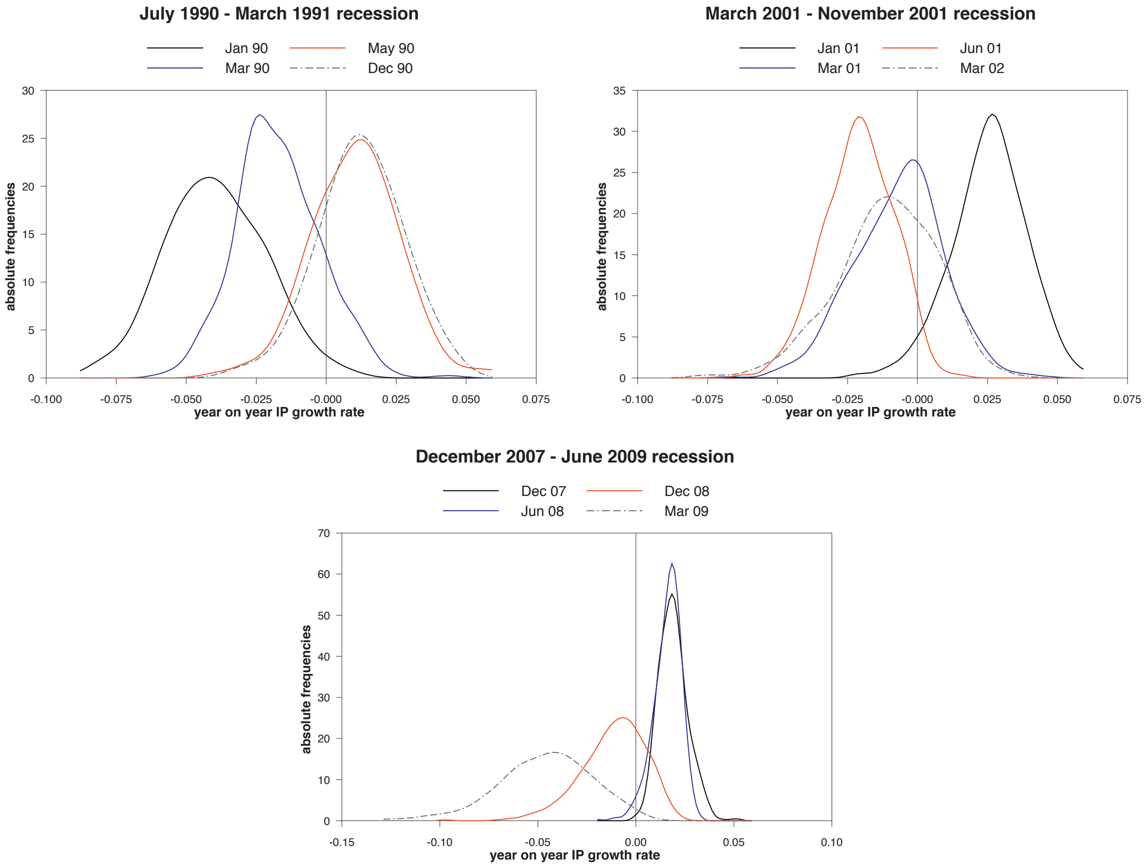



Figure 13: Distribution of the expected year-on-year growth of the US IP index at selected dates around the last three recession episodes.