EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1455 / AUGUST 2012

SURVEY-BASED NOWCASTING OF US GROWTH

A REAL-TIME FORECAST COMPARISON OVER MORE THAN 40 YEARS

by Antonello D'Agostino and Bernd Schnatz

In 2012 all ECB publications feature a motif taken from the €50 banknote.

Acknowledgements
The views expressed in this paper are those of the author and do not necessarily reflect those of the Central Bank of Ireland, European Financial Stability Facility or ECB. It includes many insightful comments by Gabe de Bondt, Michael Ehrmann, Moritz Karber, Giulio Nicoletti, Caroline Willeke and participants at an internal ECB presentation.

Antonello D'Agostino
at European Financial Stability Facility (EFSF), 43 avenue John F. Kennedy, L-1855 Luxembourg and the Central Bank and Financial Services Authority of Ireland;

Bernd Schnatz
at European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; e-mail: bernd.schnatz@ecb.europa.eu

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

ISSN 1725-2806 (online)

Abstract

Reliable and timely information about current economic conditions is crucial for policy makers and expectations formation. This paper demonstrates the efficacy of the Survey of Professional Forecasters (SPF) and the Purchasing Manager Indices (PMI) in anticipating US real economic activity. We conduct a fully-fledged real-time out-of-sample forecasting exercise linking these surveys to US GDP and industrial production growth over a long sample period. We find that both indicators convey valuable information for assessing current economic conditions. The SPF clearly outperforms the PMI in forecasting GDP growth, while it performs quite poorly in anticipating industrial production growth. Combining the information included in both surveys further improves the accuracy of both, the PMI and the SPF-based forecast.

JEL classification: E37, E47, C22, C53.

Keywords: US, Business Cycle, PMI, Forecasting, Real Time Data

Non-technical summary

The considerable delay in the publication of national accounts data undermines the policy makers' need for reliable and timely information about current economic conditions. In the United States, for instance, the first (advance) data release of GDP growth for the current quarter is published only at the end of the first month of the next quarter. Therefore, many policy makers and financial institutions devote significant resources to exploit alternative sources of information in order to gauge the continuously evolving state of the real economy.

Policy makers and market participants regularly attach great importance to survey evidence to measure current economic conditions. In this context, the Survey of Professional Forecasters (SPF) and the Purchasing Manager Indices (PMI) have become very influential yardsticks. Against this background, this paper mainly addresses two issues: Firstly, we assess which survey (if any) outperforms in terms of anticipating current economic conditions prior to their release. Secondly, we ask, if and how the information from the SPF and the PMI surveys can be combined to get an even more accurate picture of the current state of the US economy, rather than using just one of these surveys.

To our knowledge, this is the first paper that provides a systematic, rigorous and comparative analysis of the performance of two the most prominent US surveys with a long history in a fully-fledged real-time out-of-sample comparison exercise over a long sample period of more than forty years. In the estimation, we link these surveys to real-time data vintages on US GDP and industrial production available on the SPF website of the Federal Reserve Bank of Philadelphia. This ensures that no information is taken into account that was not available at the time of actual forecasting. In order to account for unpredictable data revisions, which are a common feature for these output measures, we compare our forecasts to the figures published after the next two subsequent quarters; however, we report the results also for the last data vintage available.

Overall, we find robust evidence that both indicators convey valuable information for assessing current economic conditions (compared with naïve univariate benchmarks). The SPF clearly outperforms the PMI in forecasting GDP growth, while it performs quite poorly in anticipating industrial production growth. However, quite strikingly, combining the information included in both surveys further improves the accuracy of both, the PMI and the SPF-based forecast.

1. Motivation

Reliable and timely information about current economic conditions is crucial for policy makers to take decisions in real time and for steering agents' expectations formation about the state of the economy. Such assessments are, however, thwarted by the considerable delay in the publication of national accounts data. More specifically, in the United States, the first (advance) data release of GDP growth for the current quarter is published only at the end of the first month of the next quarter. Therefore, significant resources need to be devoted to exploit alternative sources of information in order to gauge the continuously evolving state of the real economy. Many policy makers and market participants take recourse to survey evidence to measure current economic conditions. This is widely evidenced by monetary policy communications, which frequently point to survey evidence when describing the current macroeconomic situation.

For economic activity in the United States, two prominent surveys with a long history stand out. Firstly, the Federal Reserve Bank of Philadelphia Survey of Professional Forecasters (SPF) releases direct forecasts of US economic activity indicators – such as GDP or industrial production – in the middle of each quarter. Secondly, the (manufacturing) Purchasing Managers' Index (PMI) – released by the Institute for Supply Management (ISM) – has become a very influential yardstick for applied economists and the financial press as it is even timelier and available at a monthly basis.

On the PMI, so far, most of the applied literature has studied the usefulness of the PMI indicators per se. For instance, Harris (1991) attributes significant explanatory power to the PMI in anticipating US economic activity. Harris et al. (2004) also present evidence that the US manufacturing PMI provides a good gauge of US economic activity. This is consistent with Koenig (2002), who concludes that the PMI is a valuable tool for tracking the health of the US manufacturing sector.[2]

However, it is crucial to move beyond simple univariate benchmark models and judge the performance of these survey indicators against an appropriate competitor. In this tradition, Lahiri and Monokroussos (2011) compare PMI-based models to the forecasts of the dynamic factor model

2 De Bondt (2011) provides affirmative evidence for the nowcasting power of the PMI for the euro area.

of Giannone et al. (2008, 2010).[3] They also find evidence that the PMIs can improve on the forecasts of US GDP growth based on the factor model. However, from a technical point of view, a dynamic factor model including more than hundred macro variables prevents employing a fully fledged out-of-sample forecast comparison exercise based on real time data over a long sample span. Accordingly, the authors employ a pseudo real-time dataset which is based on a single data vintage and focus their out-of-sample analysis on the recent crisis episode. One noteworthy exception is Liebermann (2011), who constructs a novel real-time database for a panel of US variables and compares the performance of a factor model to that of the SPF over a ten-year period from 2000 to 2010. She finds that the SPF does not carry additional information with respect to the best factor model, implying that the often cited superiority of the SPF is rather weak in her sample.

This paper aims at simulating, for the US, the true real-time situation of a forecaster at each point in time over more than forty years. Therefore, we need a compromise benchmark. On the one hand, this benchmark should be more sophisticated than a naïve univariate model, but on the other less data demanding than an all-inclusive dynamic factor model. In fact, the SPF is a sensible candidate in this regard. It includes the views of a large number of professional forecasters, who, in turn, base their assessment on a large variety of macro data available at the time. In fact, the SPF has been shown to encompass a number of convenient properties: Firstly, from a more theory-related perspective, this survey has been used to test the rationality of agents (forecasters). Secondly, evidence has been provided that the SPF improves and complements the forecasts of traditional macro models (see Campbell (2007), D'Agostino and Whelan (2007) and D'Agostino, Mc Quinn and Whelan (2011)). Thirdly, D'Agostino et al. (2006) show that a good forecasting performance (relative to that of a simple benchmark model) is mainly achieved for short horizons (nowcast) and that the forecast accuracy of such surveys has reclined remarkably after the "Great Moderation".

[3] In terms of terminology, we use the term "forecasting" throughout the paper although our concern is about coincident economic conditions. Other papers used the term "nowcasting" in such contexts. For instance, Banbura et al. (2010) use dynamic factor models to produce a sequence of nowcasts for euro area activity. For the global economy, Jakaitiene and Dees (2009) proposed a number of factor model-based approaches to forecast short-term changes in selected world economic variables. See also Aruoba et al. (2009) for a prominent application for the US. Using their indicator would be interesting but is unfeasible in the present real-time analysis, because data vintages are available since 2008 only.

Finally, using the median SPF is consistent with the widespread empirical finding that simple forecast averaging methods provide stable and good results (see Stock and Watson, 2004).[4]

In this paper, we mainly address two issues: Firstly, we assess which survey (if any) outperforms in terms of anticipating current economic conditions prior to their release. Secondly, we ask, if and how the information from the SPF and the PMI surveys can be combined to get an even more accurate picture of the current state of the US economy, rather than using just one of these surveys.

This is, to our knowledge, the first paper providing a systematic, rigorous and comparative analysis on the performance of two most prominent US surveys with a long history in a fully-fledged real-time out-of-sample comparison exercise. In our empirical work, we use the median forecasts for GDP and for industrial production as our SPF activity variables. We include industrial production in the analysis, because the manufacturing PMI might be more closely aligned with industrial production than with broader definitions of economic activity. At the same time, industrial production is much more volatile and therefore more difficult to project by professional forecasters. In order to account for unpredictable data revisions, which are a common feature for these output measures, we compare the forecasts with the figures published after the two subsequent quarters (see Romer and Romer 2000); however, we report the results also for the last data vintage available. The choice between these two vintages is non-trivial: The last data vintage characterises best the "true" state of the economy at that point in time. Correspondingly, it could be considered the most appropriate benchmark. However, the forecaster makes the projection based on (unrevised) data available at that time. Therefore, he cannot anticipate benchmark revision, which makes using the last data vintage perhaps overly ambitious.

The paper is organised as follows. Section two shows some stylised facts and briefly recalls the construction of the PMI indices, their merits and limitations. Section three describes the forecasting exercise and section four summarises the results. It demonstrates in the out-of-sample forecast comparison exercise the efficacy of PMI-based models and the SPF relative to a naïve benchmark for projecting growth in US GDP and industrial production. For GDP growth, the SPF seems to

[4] Capistrán and Timmermann (2009) showed in a pseudo real-time forecasting exercise, that using the simple equal-weighted average method for combining individual forecasts performs best for most variables. For the euro area, Genre et al. (2010) show that alternative combinations of the survey of professional forecasters deliver only small quantitative improvements to the equal weighted combination for GDP growth.

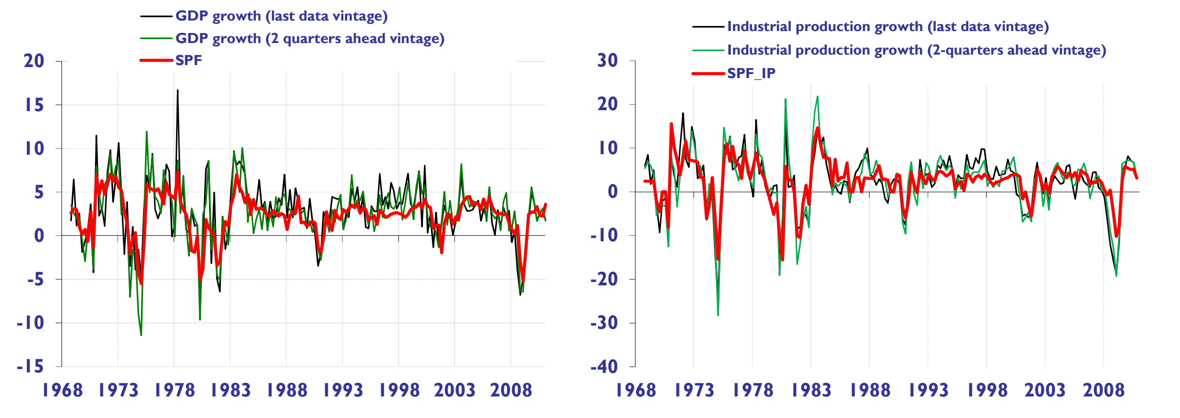
outperform the PMI, while it is vice versa for industrial production growth. Section four shows that combining the PMI and SPF forecasts indeed further improves the forecast accuracy. Section five concludes.

2. Stylised facts

2.1. Survey of Professional Forecasters (SPF)

The oldest quarterly survey of macroeconomic forecasts in the United States is the SPF.[5] Respondents include Wall Street financial firms, banks, consulting groups, and forecasters at large corporations. It appears reasonable to assume that it summarises economic news available in the public domain, although the methods these forecasters use to create their forecast are commonly not revealed. The survey is conducted early in the second month of each quarter and released few days later.[6] By that time, the first (advance) release of GDP growth of the previous quarter is available.

Chart 1: Projections of the SPF and realisations for GDP and industrial productions growth, last data vintage and two-quarter ahead data vintage
in % annualised rates



Source: Federal Reserve Bank of Philadelphia (SPF).

Chart 1 suggests that the SPF tracks GDP and industrial production growth rather well, no matter if compared with the latest vintage of data or if measured against the data available two quarters after the respective quarter, showing correlation coefficients between 0.72 and 0.85. This underscores the

5 When it began in 1968, it was conducted by the American Statistical Association and the National Bureau of Economic Research. In 1990, the Federal Reserve Bank of Philadelphia took over the survey.

6 Since 2005, the SPF has been commonly published at around the 10th of the second months of each quarter. Before that, the SPF was published roughly in the middle of the months, sometimes only around the 20th day.

benchmark role of the SPF in the literature. Indeed, it has proven challenging to systematically outperform the SPF in forecasting US economic activity. Only in a few periods, the SPF seems to have underestimated the strength of the US growth momentum. This is somewhat apparent in the mid-1980s and in the late-1990s, but also the strength of the recovery after the most recent global crisis was initially somewhat stronger. Over the entire period, however, the SPF seems to provide good forecasts. Between 1968 and 2011, the median SPF growth projection of the US economy was 2.6% (in quarterly annualised terms), which is slightly below the actual growth data that was available two quarters later at 2.7% (see Table 1). Only in the latest data vintage, the median growth rate stood somewhat higher at 3.0%.[7] As one would expect, the standard deviation of the SPF is also much smaller than the actual data, but still substantial.

Table 1: Descriptive statistics of the GDP and SPF data

	GDP^{final}	GDP^{2q}	GDP^{SPF}	IP^{final}	IP^{2q}	IP^{SPF}
Mean	2.86	2.61	2.33	2.37	2.38	2.43
Standard deviation	3.49	3.49	2.54	6.80	6.33	4.78
Equality test (p-value)	*0.11*	*0.40*		*0.94*	*0.93*	
Median	3.00	2.65	2.56	3.00	3.42	3.06
Equality test (p-value)	*0.09*	*0.28*		*0.46*	*0.78*	
Correlation with SPF	0.72	0.78		0.79	0.85	

For the equality of mean test, a standard t-test is applied, for the equality of medians, the Wilcoxon/Mann-Whitney test is applied. Final refers to the final data vintage, 2q refers to the data vintage two quarters ahead. SPF refers to the published forecasts by the Survey of Professional Forecasters.

2.2. ISM/PMI indices

The PMI is a natural competitor (or complement) to the SPF projections, which can be verified over a long time span. The US PMI data from the Institute of Supply Management (ISM) is also designed to provide a snapshot of the health of the economy. We employ the ISM manufacturing production indicator, because these data range back to 1948, while an index including non-manufacturing activities is available only since 1998, which is too short to be analysed systematically.[8] The data is based on a monthly survey of more than 300 purchasing and supply

7 Formal tests of forecast efficiency clearly confirm the unbiasedness hypothesis for industrial production, while for GDP growth, there is some evidence for a bias if the final data vintage is used (see Timmermann, 2006). This is in line with Patton and Timmermann (2010), who use Green Book data for US GDP growth.

8 De Bondt and Schiaffi (2011) provide an analysis of the composite indicator for a shorter time span, but their objective is more to assess whether consumer confidence indicators have additional explanatory power in a regression-based rather than real-time out-of-sample exercise. They show some robustness checks also for the manufacturing PMI over long periods.

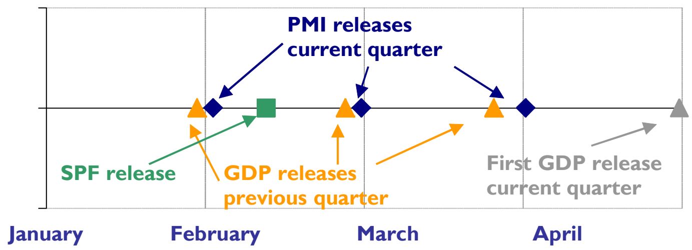
executives from across the country. Survey respondents are asked whether their output has risen, fallen or remained unchanged on that of one month ago. The unweighted net balance of survey responses is converted into a (seasonally adjusted) diffusion index – with a level of 50 being the threshold value between contraction and expansion.

$$PMI_t^{\text{var}} = 100\frac{I + 0.5N}{I + N + D}$$

where "I" is the number of respondents reporting increases, "N" is the number of respondents reporting no change and "D" is the number of respondents reporting decreases. A reading above 50 in the diffusion index implies that more firms report expanding activity than contracting activity. In practise, the index constitutes a hybrid indicator based on subjective responses which encompasses both actual data elements and a confidence element.

One of the most attractive features of the PMI is its timeliness. The PMI for the manufacturing sector for a certain month is released on the first business day of the following month. In the regular quarterly data dissemination cycle, this implies that first information on economic activity in the *current* quarter is available very shortly after the advance estimate of US GDP growth for the *previous* quarter and more than two weeks before SPF forecasts for the current quarter will become available. Chart 2, which provides the stylised release calendar in a typical quarter, also illustrates that this information is available also almost three months before the first release of US GDP growth in the present quarter.

Chart 2: Stylised representation of data releases over the quarter



The PMI has also the convenient feature that it is not subject to revisions. This implies that issues of the "real-time data vintage" of the explanatory variables can be ignored. The most important limitation of the PMI index is its construction as a diffusion index. A higher PMI reading simply means that more respondents are reporting improving (rather than deteriorating) conditions compared to the month before. As pointed out by Vermeulen (2012), the indicator does neither control for the intensity of the change in business condition, nor does it weigh the responses according to the size of the firm. However, he also shows that using alternative distributional assumptions to map the PMI survey results into growth forecasts for US industrial production yields overall very similar estimates.

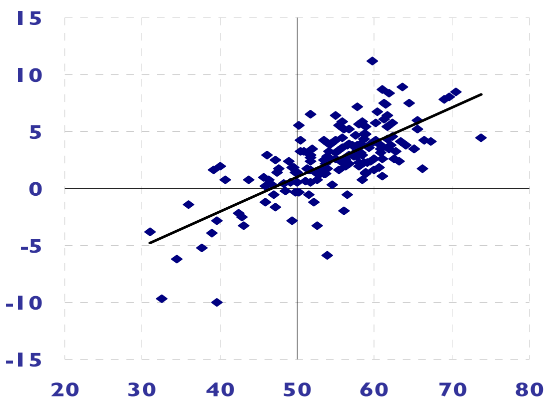
The scatter plot below (Chart 3) demonstrates the close positive relationship between the PMI (manufacturing) output index and growth in US real GDP and industrial production. It illustrates that the link is closer for GDP data available two quarters after the forecast than for the finally revised data. Interestingly, the scatters also suggest that the actual threshold between expansion and contraction is below 50 for GDP growth, but above 50 for industrial production growth.

Chart 3: PMI and US real GDP growth

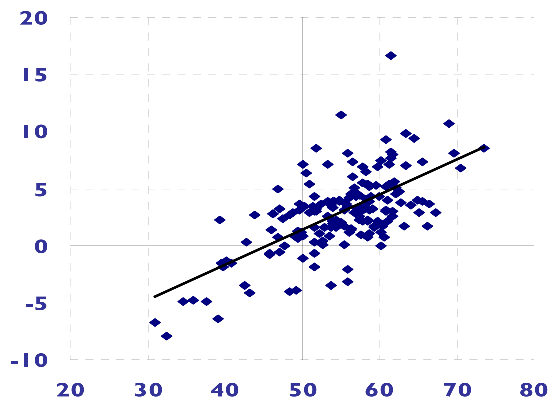
Sample period: 1969Q1-2010Q4

Horizontal axis: PMI: diffusion index, vertical axis: GDP/industrial production growth in %

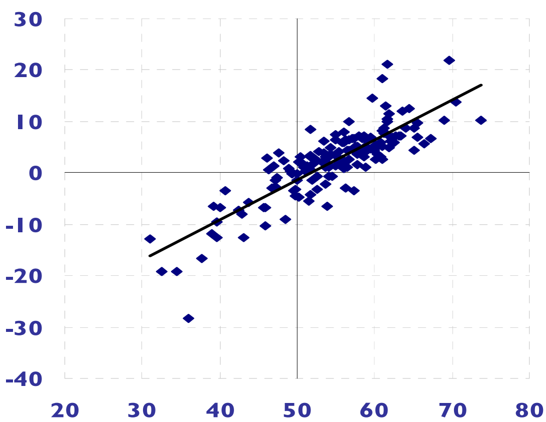
GDP data vintage:
Two-quarters ahead



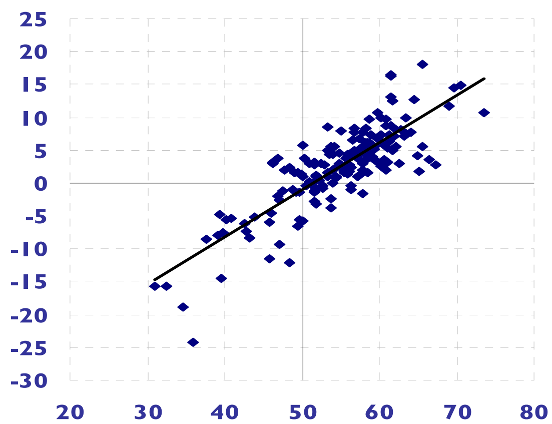
GDP data vintage:
Latest data available



Industrial production data vintage:
Two-quarters ahead



Industrial production data vintage:
Latest data available



Source: Markit, Federal Reserve Bank of Philadelphia.

Following Koenig (2002), this can be more formally verified by running the following regression:[9]

$$\Delta y_t = \beta_1 (pmi_t - \beta_o) + \varepsilon_t$$

In this regression, the constant term β_0 represents the level at which the regression line intersects the horizontal axis in Chart 3, consistent with zero growth. The estimation – performed over the sample period from 1968Q4-2010Q4 – confirms the highly significant relationship between the PMI and growth in GDP and industrial production. A one-unit decline in the PMI index is consistent with 0.3 pp lower GDP growth and a 0.8 pp decline in industrial production growth. The stronger response of industrial production is consistent with the higher variance of this series. A Wald-test for the intercept term shows, that the actual no-growth threshold for GDP is significantly below 50, but significantly above 50 for industrial production (see Table 2).

Table 2: OLS regression results

Dependent variable	GDP^{final}	GDP^{2q}	IP^{final}	IP^{2q}
Slope β_1	0.31	0.34	0.72	0.78
(t-value)	*(11.5)*	*(9.1)*	*(10.7)*	*(13.0)*
Constant β_0	45.65	47.14	51.51	51.76
(t-value)	*(42.8)*	*(57.0)*	*(75.82)*	*(102.5)*
Wald-test H_0: β_0=50	16.6	5.7	12.1	12.1
(p-value)	*(0.00)*	*(0.02)*	*(0.03)*	*(0.00)*
No. of obs.	169	168	168	168
R^2 (adj.)	0.43	0.51	0.69	0.71

Newey-West HAC Standard errors and covariance (lag truncation=4).

3. Forecast comparison exercise

3.1. Forecasting models and evaluation

This section sets up a fully-fledged real-time out-of-sample assessment of the link between US economic activity growth and survey evidence, thereby going well beyond analysing the in-sample properties of these data in the previous section. Out-of-sample procedures are crucial to assess the forecasting performance of indicators. As emphasised in Carriero and Marcellino (2007), it is always possible to explain the behaviour of a specific variable reasonably well when a set of

9 Koenig (2002) also includes the change in the PMI as another exogenous variable. This does not change the results materially.

parameters is carefully chosen, but that there is no reason to expect that such equation are also good forecasting tools. To address this critical issue, we use real-time data vintages on US GDP (quarterly frequency with quarterly vintages) and industrial production (monthly frequency with quarterly vintages) available on the SPF website of the Federal Reserve Bank of Philadelphia. This ensures that no information is taken into account that was not available at the time of actual forecasting.

In the first step of the following empirical analysis, we ask whether the survey evidence is any better than a naïve (constant growth model) benchmark and which indicator performs best. In the second step, we analyse, whether a forecast combination can further improve on models including just one survey indicator. Finally, we conduct a robustness test over the Great Moderation episode to underpin the strength of our results.

In order to deal with the multi-frequency of the data, we follow Parigi and Schlitzer (1995) and Hahn and Skudelny (2008) and use bridge equations. The bridge equation maps developments in the PMI data, which is published at a monthly frequency, to quarterly growth rates. We use the following mapping procedure: Once the PMI for the first month of the quarter is released (z = 1), simple autoregressive models ($s \leq 4$, consistent with the Bayesian Information Criterion) project the PMI over an horizon (h=2) of the next two months. After two monthly releases of PMI data, just the third month is projected (h=1), using the same method. More formally, this implies:

$$(1) \qquad p\hat{m}i^{m}_{t+h|z} = \hat{\mu} + \sum_{s=0}^{p} \hat{\gamma}_s \, pmi^{m}_{t-s}$$

The series is then converted to the quarterly frequency. Note that at the very beginning of the next quarter, all three monthly PMI observations for the current quarter are available, which implies that the quarterly PMI in the last line is not based on any estimates:

$$(2) \qquad \begin{aligned} p\hat{m}i^{q}_{t|1} &= \tfrac{1}{3}(p\hat{m}i^{m=3}_{t|1} + p\hat{m}i^{m=2}_{t|1} + pmi^{m=1}_{t|1}) \\ p\hat{m}i^{q}_{t|2} &= \tfrac{1}{3}(p\hat{m}i^{m=3}_{t|1} + pmi^{m=2}_{t|1} + pmi^{m=1}_{t|1})\,. \\ pmi^{q}_{t|3} &= \tfrac{1}{3}(pmi^{m=3}_{t|1} + pmi^{m=2}_{t|1} + pmi^{m=1}_{t|1}) \end{aligned}$$

Overall, we need to define three PMI series depending on how much actual PMI information has been available in each quarter. More formally, the PMI-based model simply uses static linear regressions (OLS) between the quarterly $pmi_{t|z}$ available at time t (ignoring the q suffix in the following) and the respective available data vintage for US real activity growth y_t. z = 1, 2, 3, depending on the number of available PMI releases in a certain quarter.[10] In equation (3), $\hat{y}_{t|z}^{pmi}$ denotes the real-time out-of-sample PMI-based forecast of the US activity variable y_t computed at time t conditional on z-months of available PMI data.

(3) $$\hat{y}_{t|z}^{pmi} = \hat{\alpha} + \hat{\beta}\, p\hat{m}i_{t|z}\,.$$

The second forecast is readily available from the median of SPF.

(4) $$\hat{y}_t^{spf} = y_t^{spf}$$

The out-of-sample real-time forecasts of these survey-based models are compared to a naïve benchmark model, which is simply the average of past US activity growth rates over the estimation periods (random walk in levels), including m observations:[11]

(5) $$\hat{y}_t^{nve} = \frac{1}{m}\sum_{i=1}^{m} y_{t-i}^{v}$$

where y_t is the growth rate in US GDP or industrial production for the data vintage v available at time t.

The empirical exercise is divided in two parts. In the first part we use real-time data back to 1948 to produce recursive out-of-sample PMI-based forecasts over the full sample ranging from the fourth quarter of 1968 to the second quarter of 2011. We compare the performance of these forecasts with that of the SPF and the naïve benchmark. In the second part we test, if a forecast combination, based on PMI and SFP predictions, can improve on the single variable forecasts. Following Granger and Ramanathan (1984) and Timmermann (2006), the combination is built in the following way. First, we recursively estimate the regression coefficient of this simple model, which includes a constant term and does not impose the constraint that the parameters add up to one in order to allow for the possibility that the underlying forecasts are biased:

(6) $$y_t = c + \alpha_1 \hat{y}_{t|z}^{pmi} + \alpha_2 \hat{y}_t^{spf}$$

---

10 The results are robust to using a dynamic specification, which may include up to four lags of the activity and PMI variables.

11 For industrial production, the construction of the naïve forecast is somewhat more complicated. In each quarter, we assume that industrial production data for the first month is available. Then, we compute the average monthly growth rate of industrial production over the respective data vintage, which is used to extend the series in levels for two months to the end of the quarter. Finally, we compute the growth rate of the current quarter relative to the previous quarter as the naïve benchmark.

Second, we use the estimated coefficients to combine the forecasts at time t as follows:

(7) $$\hat{y}_{t|z}^{com} = \hat{c} + \hat{\alpha}_1 \hat{y}_{t|z}^{pmi} + \hat{\alpha}_2 \hat{y}_t^{spf}$$

The forecast accuracy of all models is evaluated through the Mean Square Forecast Error (MSFE) statistic, so that the forecasts minimise a symmetric quadratic loss function. However, to facilitate the comparison, the accuracy of each model is compared (ratio) with that obtained by the naïve model, used as the benchmark. We also report the statistic proposed by Clark and West (2007) to test, if the forecast produced by the various models can be considered statistically different form the naïve benchmark.

4. Empirical results

4.1. Performance of individual surveys

Table 3 shows that both the PMI-based model and the SPF contain valuable information for forecasting US real GDP growth. In both models, the MSE ratio is clearly below one, which implies that these simple models outperform the naïve benchmark model, irrespective of whether the two-period ahead or the last data vintage for real GDP growth are used.[12] Already with PMI data availability of just one month, the error of the PMI-based model is more than 30% smaller when using the last data vintage and more than 40% smaller when using the 2-quarter-ahead GDP data. Furthermore, as expected, the accuracy of the PMI forecasts improves over the quarter as more PMI information becomes available. This is evidenced by the decreasing relative MSFE as the number of months is increasing. However, the SPF is not only improving over the naïve benchmark, but it is also clearly better than the PMI-based model, even when considering the availability of PMI data for the full quarter. For all models, the Clark-West statistics suggest that the SPF and the PMI-based forecasts are significantly different from those of the naïve model.

For industrial production, the results confirm the usefulness of the PMI for forecasting, particularly if the data two quarters ahead is used as a yardstick. The relative MSE is still clearly below 1, but higher than for GDP growth. This suggests the PMI-based model to be better suited for projecting

[12] The magnitudes of the outperformance are consistent with findings by Liebermann (2011) over a shorter evaluation period. She also finds that the MSFE of the naïve benchmark model is nearly twice that of her factor model and the SPF.

GDP than industrial production, although it is based on the manufacturing survey. Quite strikingly, the SPF performs very poorly in forecasting industrial production growth.

Table 3: Relative MSFEs of various models

GDP growth	2-quarter ahead vintage	Last Vintage
PMI/Naïve (month=1)	0.58**	0.68***
PMI/Naïve (month=2)	0.48**	0.57***
PMI/Naïve (month=3)	0.45**	0.59***
SPF/Naïve	0.38**	0.48***
Memo item: MSFE Naïve (MSE)	*13.08*	*12.65*
Industrial production		
PMI/Naïve (month=1)	0.80**	0.98*
PMI/Naïve (month=2)	0.62***	0.88**
PMI/Naïve (month=3)	0.60***	0.88**
SPF/Naïve	1.23	0.97**
Memo item: MSFE Naïve (MSE)	*11.21*	*11.21*

//*** denotes significance of the Clark-West-Statistics at the 10%/5%/1% level. Clark and West statistic for nested models is the standard Diebold Mariano test adjusted for a negative term which measures the mean squared difference between predictions done under the two alternative models. We use Newey-West standard errors.*

4.2. Performance of forecast combination

In this paragraph we show that using forecast combination methods further improves the accuracy of the forecasts. In this step, the first estimation is performed over the sample 1968:Q4 – 1972:Q4 and it is iterated until the end of the available sample. Altogether, this provides 152 quarterly forecasts, which we compare to the benchmark models.

Table 4 shows the MSFE of the combined forecast relative to the PMI-based forecasts (for each month) and relative to the SPF forecast. Again, the analysis is conducted for growth in US GDP and in industrial production. As before, the forecasts are compared to the two-quarter-ahead data vintage and to the final data vintage.

Combining the forecasts from the PMI and the SPF significantly improves the forecast for both, US GDP and for industrial production. For GDP, this result is not unexpected when relating the combined forecast to the PMI-based forecast, as the previous section showed that the SPF outperforms the PMI model. However, adding the PMI-based forecast to the SPF reduces the error by almost 10% already at a time, when only the PMI for the first month of the quarter is available.

As more PMI information is released over the quarter, the advantage of the combined forecast continues to increase. This result is robust to the use of different data vintages for GDP growth.

Table 4: Relative MSFE of the combined models

GDP growth	2-quarter ahead vintage		Last Vintage	
	MSFE of combined forecast relative to PMI forecast	MSFE of combined forecast relative to SPF forecast	MSFE of combined forecast relative to PMI forecast	MSFE of combined forecast relative to SPF forecast
PMI month=1	0.61	0.91**	0.67***	0.90**
PMI month=2	0.70*	0.87**	0.76***	0.86**
PMI month=3	0.72**	0.85**	0.75***	0.86**
Industrial production	2-quarter ahead vintage		Last Vintage	
PMI month=1	0.82***	0.60**	0.74**	0.82**
PMI month=2	0.85**	0.47**	0.75**	0.73***
PMI month=3	0.83*	0.45**	0.73*	0.71***

//*** denotes significance of the Diebold-Mariano statistics at the 10%/5%/1% level.*

For industrial production, the earlier analysis showed that the PMI-based model outperforms the SPF based forecast, the latter being even outperformed by a naïve forecast (for the two-quarter ahead vintage). Against this background, it is remarkable that adding the SPF to the PMI-based model clearly improves the overall forecast for US industrial production growth.

4.3. Robustness of results

In order to assess the robustness of the results over time, we computed the relative MSFE of the combined forecast model for GDP and industrial production also for the "Great Moderation" episode, i.e. over the period 1985 to 2007. While earlier research suggested that the forecasting performance deteriorates over this period, the results are broadly stable in this exercise (see D'Agostino et al. (2006)). The performance is very robust for GDP growth and even improves relative to the PMI-based forecast. For the industrial production growth, the gain of the combined forecast relative to the SPF seems to be smaller for both data vintages (see Table 5).

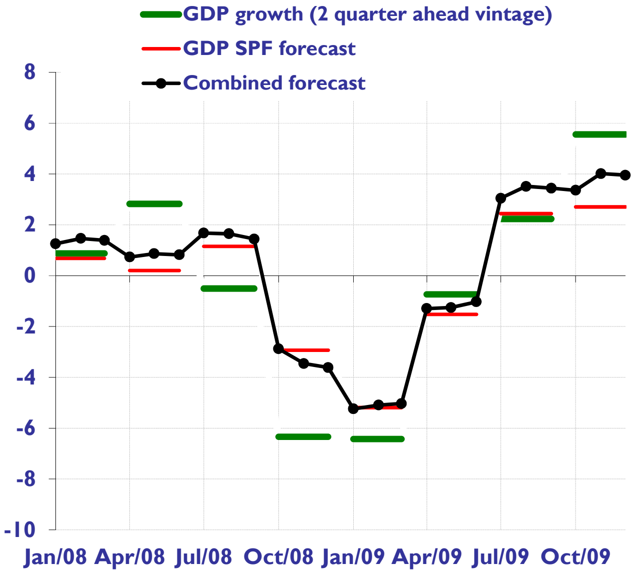
Table 5: Relative MSFE of the combined models in the Great Moderation

GDP growth	2-quarter ahead vintage		Last Vintage	
	MSFE of combined forecast relative to PMI forecast	MSFE of combined forecast relative to SPF forecast	MSFE of combined forecast relative to PMI forecast	MSFE of combined forecast relative to SPF forecast
PMI month=1	0.64**	0.92***	0.57***	0.88***
PMI month=2	0.67**	0.88***	0.64***	0.86***
PMI month=3	0.65**	0.86***	0.61***	0.87***
Industrial production	2-quarter ahead vintage		Last Vintage	
PMI month=1	0.75**	0.69***	0.66*	0.97***
PMI month=2	0.77***	0.59**	0.68**	0.91***
PMI month=3	0.74***	0.58**	0.67**	0.92***

//*** denotes significance of the Diebold-Mariano statistics at the 10%/5%/1% level.*

As regards the model performance in the crisis, Chart 4 shows the evolution of GDP growth (based on data available two quarters later), the GDP projections of the best-performing survey, i.e. the SPF, and the combined SPF/PMI forecast. It shows that in the middle of the third quarter of 2008, i.e. some weeks before the failure of Lehman Brothers, the surveys had not yet priced in the sharp decline in the GDP growth in that quarter.

Chart 4: US GDP growth during the crisis, SPF and combined forecasts
in % annualised rates



This is not very surprising: when the survey was conducted in August 2008, the advance estimate showed a GDP growth rate of almost 2% in the first quarter (annualised). Whereas this number was revised down to 1% in late-August, the preliminary estimate released at the same time suggested buoyant US GDP growth at 3.3% in the second quarter (annualised). Accordingly, professional forecasters assumed that the positive growth momentum would evolve into the third quarter, and also the SPF/PMI-based model projected a robust positive growth rate. As the PMI declined sharply in the survey released at the beginning of October, the overall decline compared to the previous quarter was muted given the rather strong PMI-readings for July (and August).

In the fourth quarter of 2008, survey respondents quickly adjusted their outlook, albeit not fully anticipating the magnitude of the downturn. Already at the beginning of November (based on PMI-data for October), the combined SPF/PMI-based model would have suggested a sharp decline of the US economy. It suggested a drop in US activity by around 3% (annualised), consistent with the SPF results released later in the same month. Over that quarter, incoming PMI data suggested a further deterioration of economic conditions, revising the forecast towards -4% in annualised terms, thereby providing strong indications of a sharp recession of the US economy. In the end, the downturn was even sharper as GDP declined by more than 6% in the fourth quarter of 2008. The US economy contracted by roughly the same magnitude in the first quarter of 2009. In this quarter, both the combined SPF/PMI and the SPF continued to correctly anticipate a further sharp decline of the US economy. Also quite strikingly, both surveys predicted the stabilisation of US growth in the second quarter of 2009 and the rebound thereafter.

5. Conclusions

This paper has shown that prominent survey indicators for the US economy – the SPF and the ISM PMI indices – are very powerful in anticipating US real economic activity in the present quarter. Such “nowcasts” of economic activity are crucial for policy makers, who need timely information about business cycle conditions. We employed a fully-fledged real-time out-of-sample exercise, simulating the situation of a forecaster each month over the past around thirty years. For real GDP, the paper demonstrates that the SPF portrays growth conditions more accurately than the PMI, while for industrial production, the PMI seems to outperform the SPF. Overall, the precision of the PMI-based forecasts improve as more information about the current quarter is released. Strikingly, however, combining the PMI-based forecasts and the SPF projections further improves the forecast accuracy.

Looking ahead, we consider several potentially fruitful extensions of our basic theme: Firstly, one could use different specifications. For instance, Vermeulen (2012) suggested that a non-linear specification of PMI-models provide slightly better forecast for economic activity. Regime-switching dynamics or smooth-transition models (see de Bondt and Schiaffi, 2011) provide other avenues to enrich the simple linear approach followed in this paper. Secondly, addressing the issue whether the assumption of a symmetric loss function may be indeed optimal has been beyond the scope of this paper. However, if the "costs" of over- and underpredicting economic activity were asymmetric, it might also be optimal to bias the forecast accordingly (see Elliott, Komunjer, and Timmermann (2004)).

References

Aruoba A. B., F. X. Diebold and C. Scotti (2009), "Real-Time Measurement of Business Conditions", *Journal of Business and Economic Statistics*, 27, 4, 417 – 428.

Bańbura M., D. Giannone and L. Reichlin (2010), "Nowcasting", ECB Working Paper 1275, Frankfurt am Main.

Bondt, G.J. de (2011), "Nowcasting: trust the Purchasing Managers' Index or wait for the flash GDP estimate?", Paper presented at Athens Institute for Education and Research (ATINER) 5th Annual International City-Break Conference: Business and Society in a Global Economy, 19-22 December, Athens and forthcoming in Conference Proceedings.

Bondt, G. de and S. Schiaffi (2011), "Confidence matters for nowcasting GDP: euro area and US evidence from a PMI-based model", Social Science Research Network (SSRN) eLibrary Paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1951869&download=yes.

Campbell, Sean, D., (2007), "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters", *Journal of Business and Economic Statistics*, 25, pp. 191-200.

Capistrán C. and A. Timmermann (2009), "Forecast Combination With Entry and Exit of Experts", *Journal of Business and Economic Statistics*, 27, 4, 428-440.

Carriero, A., and M. Marcellino (2007), A Comparison of Methods for the Construction of Composite Coincident and Leading Indexes for the UK, *International Journal of Forecasting*, 23, 219-236.

Clark, T., and K. West (2007), "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models", *Journal of Econometrics,* 138, 291-311.

D'Agostino A. and K. Whelan (2008). "Federal Reserve Information During the Great Moderation", *Journal of the European Economic Association*, 6, 2-3, 609-620.

D'Agostino, A. K. McQuinn and K. Whelan (2011). "Are some forecasters really better than others?", Forthcoming, *Journal Money Credit and Banking.*

D'Agostino, A., Giannone and Surico (2006), "(Un)Predictability and macroeconomic stability", ECB Working Paper Series 605, European Central Bank.

Elliott, G., I. Komunjer, and A. Timmermann (2005), "Estimation and Testing of Forecast Rationality under Flexible Loss," *Review of Economic Studies,* 72, 1107–25.

Genre, V., Kenny, G., Meyler, A., and A. Timmermann (2010), "Combining the forecasts in the ECB survey of professional forecasters: can anything beat the simple average?," ECB Working Paper Series 1277, European Central Bank.

Giannone, D., M. Modugno, L. Reichlin, and D. Small (2010), "Nowcasting in Real Time", mimeo.

Giannone, D., L. Reichlin, and D. Small (2008), "Nowcasting: The Real-time Informational Content of Macroeconomic Data", *Journal of Monetary Economics*, 55, 665-676.

Granger, C. W. J. and Ramanathan, R. (1984), "Improved Methods of Combining Forecasts", *Journal of Forecasting*, 3, 197-204.

Hahn E. and F. Skudelny (2008), "Early Estimates of Euro Area Real GDP Growth – A Bottom Up Approach From the Production Side", ECB Working Paper 975, Frankfurt am Main.

Harris E. (1991), "Tracking the Economy with the Purchasing Managers' Index", *Federal Reserve Bank of New York Quarterly Review*, Autumn.

Harris M., R. E. Owens, and P.-D. G. Sarte (2004), "Using Manufacturing Surveys to Assess Economic Conditions", *Federal Reserve Bank of Richmond Quarterly*, 90/4, Fall.

Jakaitiene A. and S. Dees (2009), "Forecasting the World Economy in the Short-Term" ECB Working Papers 1059, Frankfurt am Main.

Koenig E. F. (2002), "Using the Purchasing Managers' Index to Assess the Economy's Strength and the Likely Direction of Monetary Policy", *Federal Reserve Bank of Dallas Economic and Financial Policy Review,* 1, 6, 2-14.

Lahiri K. and G. Monokroussos (2011), "Nowcasting US GDP: The Role of ISM Business Sentiment", mimeo, University of Albany.

Liebermann, J. (2011), "Real-Time Nowcasting of GDP: Factor Model versus Professional Forecasters", Central Bank of Ireland Research Technical Paper 03/RT/11.

Parigi, G. and G. Schlitzer (1995), "Quarterly Forecasts of the Italian Business Cycle by Means of Monthly Economic Indicators", *Journal of Forecasting*, 14, 117 – 141.

Patton, A. J. and A. Timmermann (2010), "New Tests of Forecast Optimality Across Multiple Horizons", mimeo (http://research.stlouisfed.org/conferences/appliedeconometrics/2010/Patton_Timmermann_variance_tests_26mar10.pdf).

Romer, D, H. and C. Romer (2000), "Federal Reserve Information and the Behavior of Interest Rates", *American Economic Review,* 90, 429-457.

Stock, J. H. and M. W. Watson (2004), "Combination Forecasts of Output Growth in a Seven-Country Data Set", *Journal of Forecasting*, 23, 405-430.

Vermeulen, P. (2012), "Quantifying the Qualitative Responses of the Output Purchasing Managers Index in the US and the Euro Area", ECB Working Paper 1417, Frankfurt am Main.