

Galimberti, Jaqueson K.; Moura, Marcelo L.

**Working Paper**

## Improving the reliability of real-time Hodrick-Prescott filtering using survey forecasts

KOF Working Papers, No. 360

**Provided in Cooperation with:**

KOF Swiss Economic Institute, ETH Zurich

*Suggested Citation:* Galimberti, Jaqueson K.; Moura, Marcelo L. (2014) : Improving the reliability of real-time Hodrick-Prescott filtering using survey forecasts, KOF Working Papers, No. 360, ETH Zurich, KOF Swiss Economic Institute, Zurich, <https://doi.org/10.3929/ethz-a-010185424>

This Version is available at:

<https://hdl.handle.net/10419/102973>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

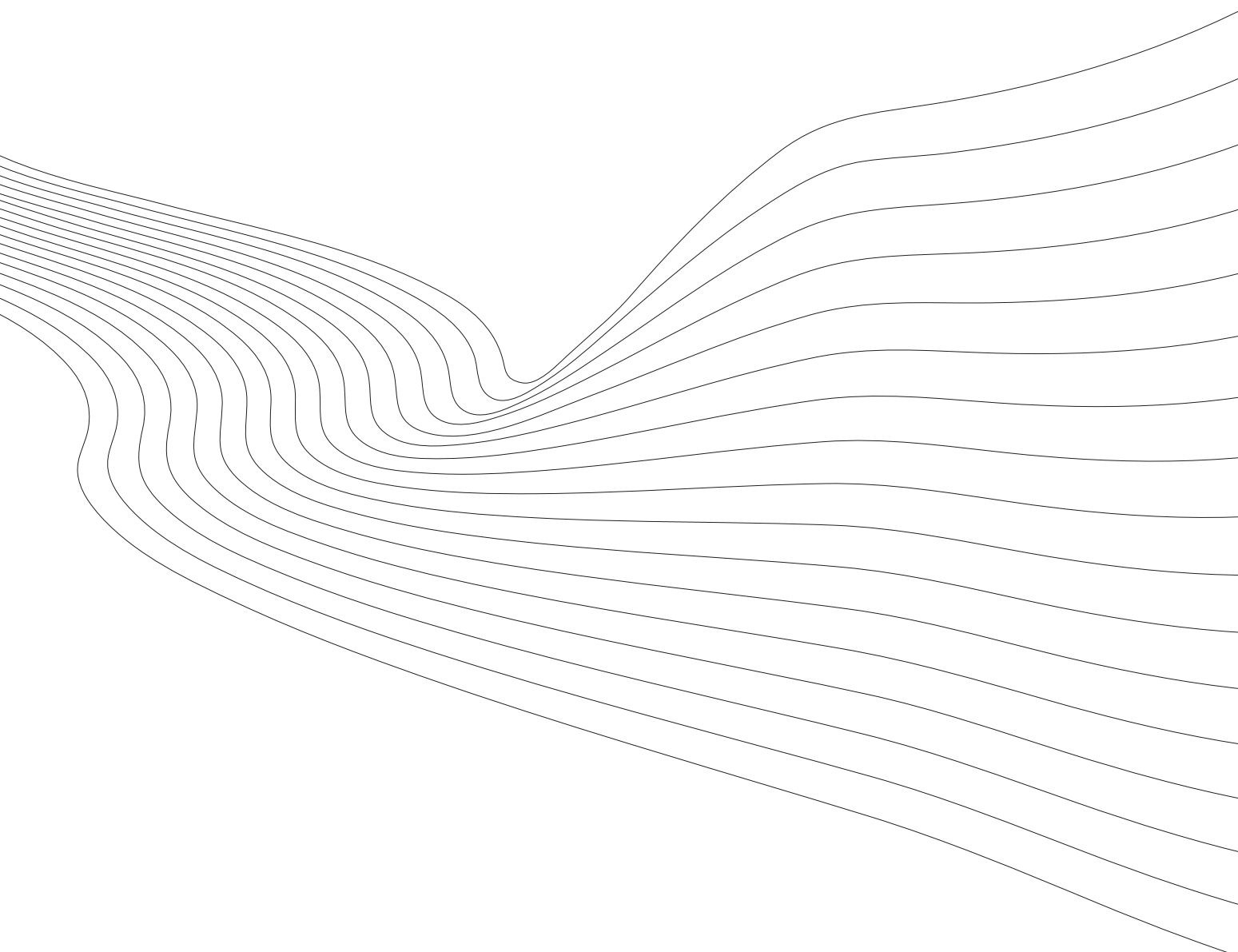
*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# KOF Working Papers

Improving the Reliability of Real-time  
Hodrick-Prescott Filtering Using Survey Forecasts

Jaqueson K. Galimberti and Marcelo L. Moura



# KOF

KOF Swiss Economic Institute  
ETH Zurich  
LEE G 116  
Leonhardstrasse 21  
8092 Zurich  
Switzerland

Phone +41 44 632 42 39  
Fax +41 44 632 12 18  
[www.kof.ethz.ch](http://www.kof.ethz.ch)  
[kof@kof.ethz.ch](mailto:kof@kof.ethz.ch)

# Improving the Reliability of Real-time Hodrick-Prescott Filtering Using Survey Forecasts\*

JAQUESON K. GALIMBERTI<sup>†</sup>

KOF Swiss Economic Institute – ETH Zurich

MARCELO L. Moura

Inspere – Institute of Education and Research

## Abstract

Measuring economic activity in real-time is a crucial issue in applied research and in the decision-making process of policy makers; however, it also poses intricate challenges to statistical filtering methods that are built to operate optimally under the auspices of an infinite number of observations. In this paper, we propose and evaluate the use of survey forecasts to augment one of those methods, namely the largely used Hodrick-Prescott filter so as to attenuate the end-of-sample uncertainty observed in the resulting gap estimates. We find that this approach achieves powerful improvements to the real-time reliability of these economic activity measures, and we argue that the use of surveys is preferable relative to model-based forecasts due to both an usually superior accuracy in predicting current and future states of the economy and its parsimony.

Keywords: business cycle measurement, end-of-sample uncertainty, gap and trend.

JEL Classifications: E32, E37.

---

\*An earlier version of this paper was presented at the 2012 EEA-ESEM joint meeting in Malaga. We thank for discussions with participants of that conference and also for the comments provided by two anonymous referees on that version of this paper.

<sup>†</sup>Corresponding author. E-mail: galimberti@kof.ethz.ch.

# 1 Introduction

Real-time estimates of the gap between output and its long-run trend level provide a lively measurement of the current state of economic activity within an economy. The information content in these estimates is of essential relevance for policymakers and market participants in general (Croushore, 2011, pp. 95-6, provides a review on this matter). Nevertheless, the reliability of filtering methods to decompose a series into its trend and gap components in real-time has been questioned and empirical results have pointed out the existence of a great uncertainty about these estimates at sample end-points (Orphanides and van Norden, 2002; Watson, 2007). The relevance of our study relies on the fact that improving the precision of those end-point estimates is imperative from a practical standpoint of applied researchers, policymakers and market practitioners regarding their need to infer the actual state of the economy for decision making. In this paper we provide an attempt to further understand and attenuate this uncertainty.

In general, previous attempts in the literature to improve real-time filtering accuracy have been characterised by the incorporation of an increasing amount of information into the estimation procedure. The most prominent attempts involve expanding the information set supplied to the filter in the temporal dimension, using either model-based forecasts (Kaiser and Maravall, 1999; Gomez, 2001; Kaiser and Maravall, 2001; Mise et al., 2005), or variables observed at a higher frequency (Aruoba et al., 2009); in the covariates dimension with the employment of model-based multivariate representations (Planas and Rossi, 2004; Valle e Azevedo et al., 2006; Altissimo et al., 2010; Valle e Azevedo, 2011; Marcellino and Musso, 2011), or jointly in both of these dimensions (Garratt et al., 2008; Clements and Galvão, 2012).

Complementing this literature, we propose and evaluate the use of forecasts from surveys to improve the reliability of real-time gap estimates. Our contribution is to show that these survey forecasts do a better job than model-based forecasts in attenuating the end-

of-sample uncertainty. We associate the success of surveys data to a better performance in providing signals of current conditions in the economy, a feature recently evidenced in other studies as well (Graff and Sturm, 2012; Leduc and Sill, 2013). We therefore recommend the use of survey forecasts as a more parsimonious alternative to model-based approaches.

Our focus goes to the famous Hodrick-Prescott (HP) filter (Hodrick and Prescott, 1997), which, despite being subject of considerable criticism (Harvey and Jaeger, 1993; King and Rebelo, 1993; Cogley and Nason, 1995, between others), is still one of the mostly used tools for business cycle analysis. Besides, the more recent and sophisticated band-pass filters (Christiano and Fitzgerald, 2003) are also subject to the issue of the end-of-sample uncertainty, as evidenced by the analysis of Watson (2007). Hence, our contribution may in principle hold true for more general model-based filter specifications that encompass both low-pass and band-pass filters (see Gomez, 2001; Harvey and Trimbur, 2003; Proietti, 2009).

We consider two different series of US output-related data as input to our filtering procedure: the quarterly real GDP and the monthly index of industrial production. By using these two different frequencies of observation we are also able to infer that filtering at a higher frequency (monthly) brings only slight improvements to the reliability of the real-time gap estimates usually obtained at the quarterly frequency. To the best of our knowledge, this is the first work that systematically tests the improvement of end-of-sample estimates of the output gap for US data on GDP and industrial production by employing market participant surveys, namely, the survey of professional forecasts (SPF) and the consensus economic forecasts (CEF). Other studies, as cited above, focused on expanding the economic series with forecasts based on time-series econometric models, a technique that is model specific and probably does not capture all the information contained in survey forecasts, which combine several market participants information sets.

Our empirical evidence indicate that not only are survey forecasts more parsimonious and easier to obtain than model based forecasts, but they are also providers of more pre-

cise gap estimates. Using the monthly series on industrial production, for example, the frequency within which the signals of the real-time gaps (i.e., whether economic activity is above or below the trend) are estimated incorrectly is reduced from 52%, without the aid of any forecast, to approximately 27% using the survey forecasts to augment the HP filter, whereas this statistic still amounts to 43% with model-based forecasts.

The remainder of this paper is organized as follows. Next section outlines the formulation of the forecast-augmented HP filter and discusses its implementation. In Section 3 we provide details about the data used in our empirical applications. The preparation of the forecasts for filtering is also described in this latter section and their relative performance is briefly evaluated. In Section 4 we present our main results on real-time filtering with the forecast-augmented HP filter. We then conclude this paper in Section 5.

## 2 Forecast-augmented Hodrick-Prescott filter

### 2.1 Formulation and end-of-sample inaccuracy

To decompose a series of observed values  $\{y_t\}_{t=1}^T$  into the sum of a trend component  $\{x_t\}_{t=1}^T$  and a gap component  $\{g_t\}_{t=1}^T$  the standard HP filter requires the choice of values for the first of these components so as to minimize

$$L(y_t, \lambda, x_t) \equiv \sum_{t=1}^T (y_t - x_t)^2 + \lambda \sum_{t=3}^T (\Delta^2 x_t)^2, \quad (1)$$

where  $\Delta^2$  stands for a twice-differenced lag operator (i.e.,  $\Delta^2 x_t = x_t - 2x_{t-1} + x_{t-2}$ ), and  $\lambda$  is a parameter regulating the trade-off between the fit of the trend to the data and its smoothness.

Under a time series model formulation of the HP filter (detailed in the implementation of the filter below),  $\lambda$  also stands as the ratio between the variances of the optimally filtered gap and the change in the trend growth. By exploring this model-based formulation, Harvey

and Jaeger (1993) and King and Rebelo (1993) established the asymptotic squared error optimality of the filter for series integrated up to the second order and with the same dynamical structure determining the changes in the trend growth and the innovations to the gap component. In spite of this optimality, critics of the use of the HP filter have pointed that it can induce spurious cyclical behavior into the filtered series (Blackburn et al., 1995; Cogley and Nason, 1995; Canova, 1998, between others). Nevertheless, later accounts have uncovered that such distortionary effects are mainly due to the unavailability of an infinite sample of observations (Ehlgren, 1998; Pedersen, 2001; Maravall and del R  o, 2007), an intrinsically applied restriction that affects any optimal signal extraction procedure. Hence, for applied purposes, the optimality of filtered gap estimates is valid only as an approximation at the midpoints of a sufficiently long series of data.

Our interest is in the inaccuracy of the HP filter estimates of the gap at the endpoints of a finite sample of observations. In contrast to the distortions affecting the optimality of filtering in finite samples, the so-called end-of-sample uncertainty in gap estimates is due to both the violation of the asymptotic requirement and the asymmetry of the filter's solution at the sample endpoints. To draw some intuition from these sources of uncertainty let  $x_{m|T}$  denote the estimate of the trend component for the sample midpoint period  $m < T$ , obtained from the solution of (1) using  $T$  observations of  $y_t$ . Also, for the moment, assume that period  $m$  is far enough from the sample endpoint  $T$  so as to validate the above mentioned approximation to optimality. Then, by taking the first order condition (FOC) to the minimization of (1) with respect to  $x_{m|T}$ , and rearranging terms, one finds that the value of the aggregate series at period  $m$  is decomposed as<sup>1</sup>

$$y_m = x_{m|T} + \lambda \nabla^2 \Delta^2 x_{m|T}, \quad (2)$$

where  $\nabla^2$  denotes a twice-differenced lead operator (i.e.,  $\nabla^2 x_t = x_{t+2} - 2x_{t+1} + x_t$ ). Now,

---

<sup>1</sup>Obviously, this does not fully determines the HP filter estimate of the components as we are focusing only on one of the FOCs of the problem in (1).



if we were restricted to use only observations up to period  $m$ , as it would be the case in real-time, this FOC would turn out to be

$$y_m = x_{m|m} + \lambda \Delta^2 x_{m|m}. \quad (3)$$

Comparison of the two decompositions for the same period  $m$  in (2) and (3) reveals one of the sources for the inaccuracy of gap estimates at sample endpoints, namely the inability of the filter to keep up with symmetry at these sample points.

This asymmetry is naturally solved as soon as data up to period  $m + 2$  becomes available, i.e., when (3) becomes

$$y_m = x_{m|m+2} + \lambda \nabla^2 \Delta^2 x_{m|m+2}. \quad (4)$$

Still, the decomposition in (4) may not yet be considered optimal since the finite sample approximation to optimality that we assumed from departure required  $m$  to be at a “sufficiently far” distance to the last observation being used for estimation. Therefore, the HP filter asymmetry at series endpoints does not account for the whole uncertainty associated with its real-time gap estimates.

Nevertheless, these derivations point to a natural way to deal with the problem of end-of-sample uncertainty in HP filtering. Namely, we can expand the information set supplied to the filter by use of forecasts, a solution that has been proved successful in previous literature (Kaiser and Maravall, 1999, 2001; Mise et al., 2005; Garratt et al., 2008). Under the same circumstances as in the original formulation, (1), suppose now we have  $h = 1, \dots, H$  forecasts of  $y_{T+h}$  at our disposal<sup>2</sup>. Then a forecast-augmented HP filter can be obtained as

---

<sup>2</sup>An alternative formulation can be developed with forecasts for the trend component, which may be particularly relevant for the case of survey forecasts where there is uncertainty about what respondents are actually targeting. Our experimentation with such formulation, however, did not yield good results. These are available upon request.

the solution to the problem of choosing  $\{\hat{x}_t\}_{t=1}^{T+H}$  so as to minimize

$$L(y_t, \lambda, \hat{x}_t) + \sum_{h=1}^H (\hat{y}_{T+h} - \hat{x}_{T+h})^2 + \lambda \sum_{h=1}^H (\Delta^2 \hat{x}_{T+h})^2, \quad (5)$$

where  $\hat{y}_t$  stands for the forecasted value of  $y_t$ ,  $\hat{x}_t$  for the trend estimate obtained from the forecast-augmented series, and  $L(y_t, \lambda, \hat{x}_t)$  comes from (1).

Clearly, our exercise in (2)-(4) shows that the use of forecasts in (5) can attenuate the end-of-sample inaccuracy introduced by the asymmetry of the filter in the estimate of  $x_t$  based solely on observable information. This solution take us back to the work of Burman (1980), which illustrated how optimal components estimates can be obtained from the augmentation of the original series with optimal forecasts. Still, notice that forecasts optimality can at best be defined within a given representation for the time series being forecast. Therefore, by turning the filtering problem into a joint forecasting problem, preservation of the filter optimality properties will also depend on the efficiency within which a time series representation of the aggregate series can be identified and estimated. In other words, augmenting the HP filter with inaccurate forecasts can have hazardous effects on the reliability of the cycle estimates if the negative effects of incorporating wrong signals are not offset by the gains from symmetry.

## 2.2 Implementation and calibration

Different approaches can be used for the computational implementation of the HP filter (see, e.g., Gomez, 1999). From the (quadratic) optimization problems in (1) and (5) one can derive the FOCs relative to  $\{x_t\}_{t=1}^T$  and  $\{\hat{x}_t\}_{t=1}^{T+H}$ , respectively, which would compose a system of linear equations in these unknown terms and could then be solved by standard matrix algebra for a given  $\lambda$  and a sample of observations (see Danthine and Girardin, 1989, pp. 49-50). One disadvantage of this approach is that solving this system of FOCs requires inversion of a  $T \times T$  matrix which, despite the computational power provided by

modern technology, may still constitute a costly procedure when dealing with the repeated computations required by a real-time framework.

A more practical alternative is provided by the formulation of the HP filter in a state-space estimation framework under which efficient Kalman filtering techniques can be employed for the estimation of the unobserved components. Following Harvey and Jaeger (1993, pp. 232-3), the state-space representation of the standard HP filter is given by

$$y_t = x_t + g_t, \quad (6)$$

$$x_t = x_{t-1} + z_{t-1}, \quad (7)$$

$$z_t = z_{t-1} + \varepsilon_t, \quad (8)$$

where  $g_t \sim NIID(0, \sigma^2)$ ,  $\varepsilon_t \sim NIID(0, \sigma^2/\lambda)$ , and, in the language of state-space models, (6) is the measurement equation, and (7)-(8) are state equations.

For the forecast-augmented version of the HP filter, the same formulation of (6)-(8) can be used by supplying the augmented series with forecasts instead of supplying them only with observables. Estimates for the trend component are obtained from the Kalmar smoother of state  $x_t$ , with  $\sigma^2$  estimated numerically from the data by maximum likelihood (ML)<sup>3</sup>, and the gap calculated as the residual from (6), i.e.,  $g_t = y_t - x_t$ .

Finally, we need to specify the calibration of the smoothing parameter. For quarterly data it is customary to set  $\lambda$  equal to 1,600, a value purposely chosen to filter fluctuations at frequencies lower than about 8 years (Prescott, 1986, p. 14). For monthly data we follow Ravn and Uhlig (2002) and set this parameter to 129,600 in order to keep the decomposed components consistent under time aggregation<sup>4</sup>.

<sup>3</sup>Although the inclusion of forecasts may cause distortions in the estimated  $\sigma^2$ , due to forecast errors, restricting the ML estimation to the sample of actual observations (yet smoothing over the entire augmented series) had virtually no numerical effects on our analysis.

<sup>4</sup>Our results were qualitatively robust in the variations of these parameters, such as under Hodrick and Prescott (1997) original recommendation of  $\lambda = 14,400$  for monthly data.

## 3 Data and forecasts

### 3.1 Sources and construction

We conduct two empirical exercises with real-time data on the US real GDP and the US index of total industrial production (IP), both sourced from the Philadelphia's Fed Real-Time Data Research Center. A real-time dataset provides snapshots of its variables for given dates in the past as it was available for an observer at that time, where each of these snapshots is usually denoted by a vintage date (Croushore and Stark, 2001). The data on real GDP is at the quarterly frequency covering the period from 1960q1 to 2013q4 with vintages from 1990q1 to 2014q1, whereas the data on IP is monthly collected and covers an equivalent period. The real-time structure of the dataset imposes a balanced restriction of one period lag in the observation of both variables<sup>5</sup>. To keep the analysis comparable, we close the gap of observations between these two series restricting our estimation with the monthly variable to consider only the vintages corresponding to the first month of each quarter, i.e., the real-time data we use for IP consists of vintages dated 1990m1, 1990m4, 1990m7, 1990m10, ..., until 2014m1<sup>6</sup>.

The initial vintages are selected to match with the first survey forecasts available in the Consensus Economics Forecasts (CEF), which is one of the sources we use for the forecast-augmented version of the HP filter. These forecasts are made by private sector economists and are collected and published monthly in the form of growth forecasts for the current and the following year. For the quarterly series on real GDP, we use the forecasts reported in the second month of each quarter. We take the mean over the individual forecasts and augment each series of observables interpolating the growth forecasts<sup>7</sup>.

---

<sup>5</sup>The only violation of this balance refers to the real GDP observation for 1995q4 which was missing in the 1996q1 vintage due to the US federal government shutdown in late 1995. We fulfill this gap by using the observation available in the March 1996 monthly vintage for the same series.

<sup>6</sup>Surely, a real-time practitioner would estimate the gaps as frequently as new observations become available. Here we prioritize the comparability of the estimates obtained from series observed at a different frequency by pinning down common estimation periods.

<sup>7</sup>Details about these interpolations are presented in the Appendix.

For simplicity we assume CEF forecasters have access to data on the variables being forecasted up to the same vintage for which the forecast was reported. The relative timing of information in this assumption is crucial for a realistic evaluation of the accuracy of these forecasts (see Stark and Croushore, 2002). The CEF reports are published in the second week of each month with survey forecasts collected in the previous two weeks. Our real-time vintages, in contrast, generally contain data released by government statistical agencies up to the middle of the vintage period. Hence there is some uncertainty about the information panelists already have available by the time they submit their forecasts to the CEF, particularly in the case of IP<sup>8</sup>.

One second source of forecasts we use to augment the HP filter comes from the Survey of Professional Forecasters (SPF), which is the oldest quarterly survey of macroeconomic forecasts in the US. Each quarter, this survey asks professional economists to give their forecasts for several macroeconomic variables, as well as over different forecasting horizons. Here we employ the median of the individual forecasts made for a total of five quarterly horizons, starting with a nowcast. In contrast to the CEF, the SPF forecasts refer to the level of the variables, which means they are ready to use for the augmentation of the quarterly series on real GDP in their original form. In the IP case, however, the forecasts refer to the average of the monthly levels over each quarter horizon, hence requiring again the use of interpolation to conciliate the frequency of the survey forecasts with that of the series actual observations.

Regarding the relative timing of information, we are fortunate to have an already constructed real-time dataset so as to be consistent with the SPF timing of information (see Stark, 2010, Figure 1, p. 15).

Following previous literature on forecast-augmentation of the HP filter (Kaiser and Maravall, 1999; Mise et al., 2005), we also benchmark the estimates obtained from these survey

---

<sup>8</sup>The Bureau of Economic Analysis releases the first measure (Advance Report) of the US quarterly GDP by the end of the first month subsequent to the quarter, while the US IP reports are released by the Federal Reserve Board around the middle of the month (varying between the 12th and the 18th day of each month).

forecasts by computing forecasts based on the  $ARIMA(p, 1, q)$  specifications. We select the lag orders,  $p$  and  $q$ , recursively by minimizing the Akaike Information Criterion associated with the estimates fitted under a rolling window covering 15 years of data. We then compute indirect forecasts iterating over the estimated specification for the same multiple horizons we have from the surveys data. Though theory indicates that direct forecasts, obtained from the estimation of horizon-specific models, are more robust to model misspecification, Marcellino et al. (2006) find that empirics tend to favor the use of iterated forecasts on US macroeconomic time series.

Finally, notice that all the data discussed above are already seasonally adjusted and transformed into logs before filter application.

### **3.2 Accuracy of forecasts**

Given the prominent role of the forecasts in the augmented version of the HP filter, it is useful to have an overview of their relative quality. For that purpose we compare the accuracy of the survey forecasts and those obtained from the benchmark model based on their associated series of forecast errors. It is important to note that our analysis is not directly aimed to assess the relative merits of each survey. One should recall that most of the survey forecasts we evaluate are reconstructed to match each series observation frequency, undeniably introducing some arbitrary distortions to the original forecasts. Therefore, our analysis is restricted to an assessment of the quality of our (mostly) interpolated forecasts to provide signals of upcoming filtered variable realizations.

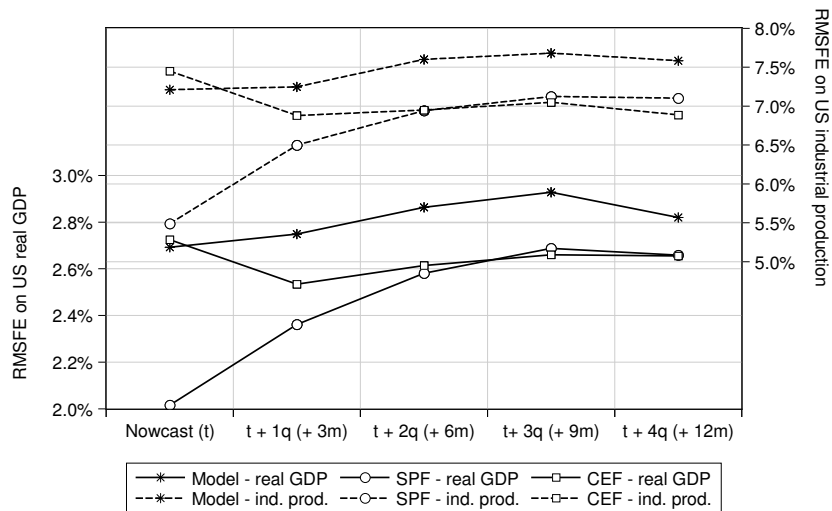
Since level values of the variables in the real-time dataset are not directly comparable, we assess the implied growth rates of the forecasts. As a measure of accuracy we calculate the root mean squared forecast errors (RMSFE), which reflects both the mean and the variance of the forecast errors. Regarding the vintage of data used as target to compute the forecast errors, we adopt the revised series, i.e., the series of growth rates obtained from the last vintage we have available (2014q1/m1). This contrasts to the literature on the

evaluation of survey forecasts, where the first-available measures are usually suggested as the appropriate target that survey respondents aim to forecast (see Stark and Croushore, 2002). Nevertheless, our evaluation of the HP filtered estimates (detailed in the next section) uses the revised series of data as a benchmark; hence, an evaluation of the forecasts in relation to the revised series of growth rates seems more consistent in our context.

The results of these forecast comparisons are presented in Figure 1, where the performance measures on forecasts for the US real GDP are scaled in the left axis while those for the US IP are scaled in the right axis. Clearly, the relative quality of the forecasts from the different sources, as well as their behavior through the forecasting horizons, is similar for the two forecasted variables. Overall, the evidence is favorable to the SPF forecasts, for which substantially smaller RMSFEs, relative to both the CEF and the model forecasts, are found over the first two forecasting horizons. The SPF forecasts also outperform the model forecasts over the remaining horizons, a result in conformity with the findings of Stark (2010).

The CEF forecasts, in contrast, perform slightly worse than the model forecasts for the first horizon, though they improve on the same levels as those of the SPF for the longer horizons. This result should provide an interesting case to check the relevance of forecast quality in shorter horizons for end-of-sample uncertainty attenuation in gap estimates; particularly for industrial production, where the model-based nowcast was able to outperform the CEF forecasts. Unfortunately, there is less agreement between the performance of our interpolation of the CEF forecasts and the previous studies in the literature evaluating the performance of this survey in its original format. Batchelor (2001), for example, finds that the CEF forecasts are more accurate than the forecasts performed by intergovernmental agencies, such as the IMF and the OECD. Also, the accuracy of these survey forecasts has been found to decrease when forecasting horizons are longer than one year (Isiklar and Lahiri, 2007; Ager et al., 2009), an observation not replicated in our measures.

Figure 1: Comparison of annualised growth forecasts accuracy.



Accuracy is measured by the root mean squared forecast error (RMSFE) of the annualised growth rates implicit in each source’s forecasts for the level of the variables. The target is given by the growth rates measured at the last vintage we have available for each variable. All forecasts presented refer to those computed in real-time. The forecast horizons in brackets refer to the monthly variable on industrial production.

## 4 HP filtering results

### 4.1 Measurement concepts

We follow Orphanides and van Norden (2002, p. 571) to distinguish between three concepts of gap estimates depending on the timing of the data supplied to the HP filter. The *real-time* estimates are those obtained by first applying the filter to each and every vintage of data and then recording the first estimate obtained for each period in the chronological order of the vintages. Thus, the real-time estimates represent the inferences one would have been able to compute using data available at the time for which the estimation is carried. The *quasi-real* estimates are those obtained by applying the filter to the revised series of data but still in a recursive procedure, i.e., using only the observations that would have been observed in real-time. The idea behind this quasi-real concept is to obtain real-time type of estimates, but without the effects of revisions on published data.



Forecast-augmentation of the HP filter for the quasi-real estimation requires some further adjustments for the case of survey forecasts. Since these forecasts are made in real-time they do not incorporate all the information available in the quasi-real estimation context. Most importantly, those forecasts made for the level of the variables, as in the SPF case, may be in a different base value. To circumvent this problem, we extrapolate the real-time forecasts to the quasi-real context by augmenting the series of observables with the forecasts (implied) growth rates. For the model-based forecasts, we simply use the revised series of data to estimate and select the ARIMA specifications in the same recursive fashion as we have done in real-time.

To benchmark the previous gap measures we define the *final* gap estimates, which are obtained by applying the filter directly to the revised series of data and recording this unique series of estimates. Clearly, the benchmark feature of these final estimates comes from the use of the latest information available within the real-time dataset. However, this measure may serve only as a provisional benchmark, since it will also be subject to end-of-sample uncertainty and data revisions as soon as new data becomes available. To avoid harmful effects of this provision we discard the last two years of estimates that we use to compute statistics evaluation. Hence, our analysis will focus on a total of 89 end-of-sample gap estimates covering the period from 1989q4 (1989m12) to 2011q4 (2011m12) for the US real GDP (for the US IP<sup>9</sup>).

The estimates we obtain under these different measurement concepts are visually presented in Figure 2, where for the the forecast-augmented HP filter case we focus on the estimates obtained using the SPF forecasts for clarity of presentation<sup>10</sup>. The gap estimates obtained from the series on real GDP are mostly consistent with those obtained from the IP series; the main discrepancy is observed during 1997 (year of the Asian financial crisis) in the final estimates, where IP anticipates the later observed boom in the real GDP gap;

---

<sup>9</sup>One should recall we use only one monthly vintage of IP for each quarter, so that we obtain the same number of gap estimates.

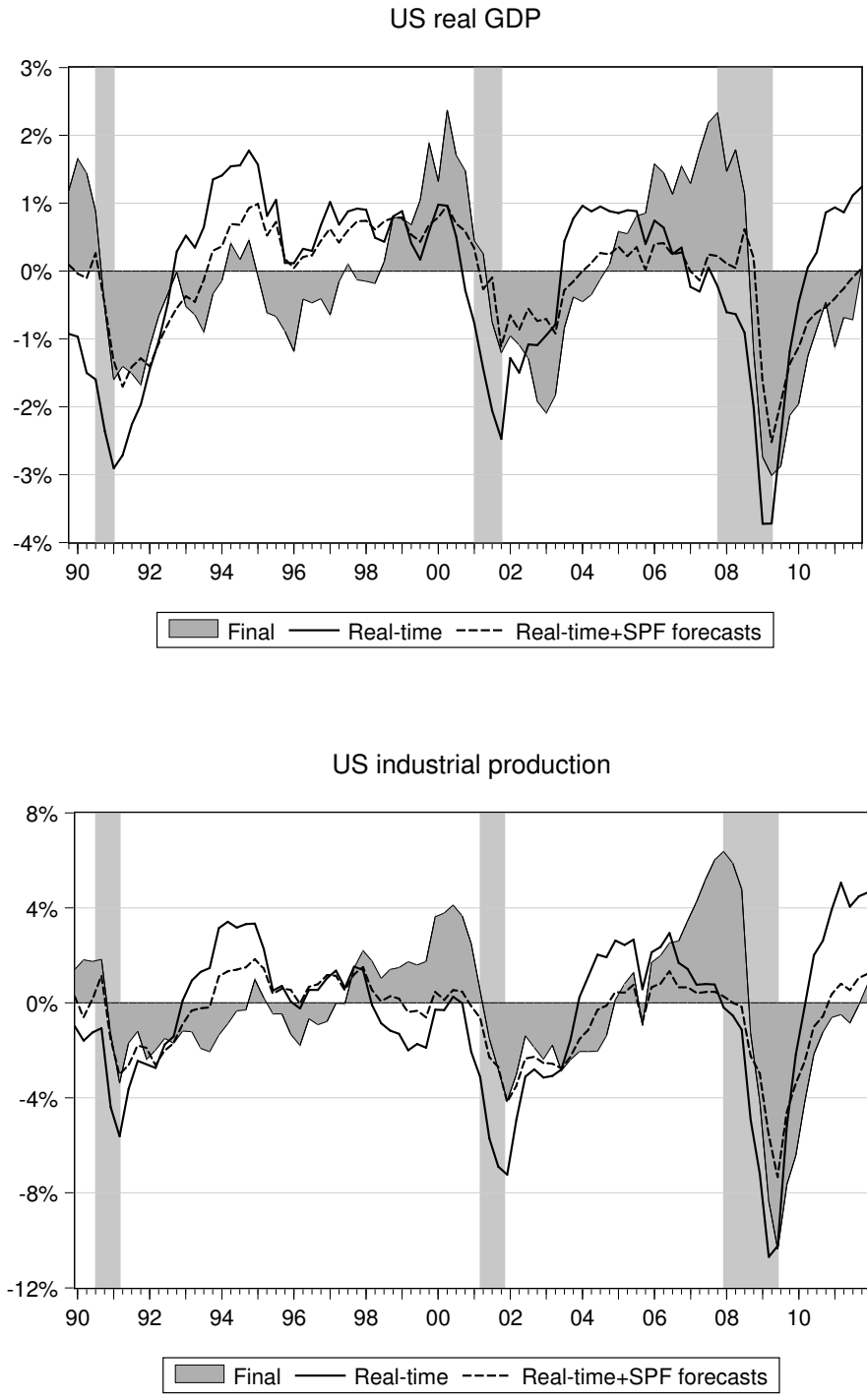
<sup>10</sup>Equivalent figures for the cases using the CEF forecasts and the model-based forecasts are available upon request.

however, it is also observed in the next two years for the real-time estimates, which provided virtually wrong signals based on the IP series. The unreliability of the real-time estimates of the gap is evident in Figure 2 by the detachment of these measures from the final ones.

Different concepts of measurement error are associated with these definitions of gap estimates. Namely, the *real-time measurement errors* denote the difference between the *real-time* and the *final* estimates and this occurs due to both revisions of published data and the filter inaccuracy at sample endpoints. To isolate these two sources we define the *quasi-real measurement errors* as the difference between the *quasi-real* and the *final* estimates to capture the effects of the filter, given that both these estimates were obtained from the same data measures. These different concepts of measurement error are illustrated in Figure 3, where the thickness of the area bands gives an idea of the amount of errors accrued by the filter due to the publication and revision process of data measures.

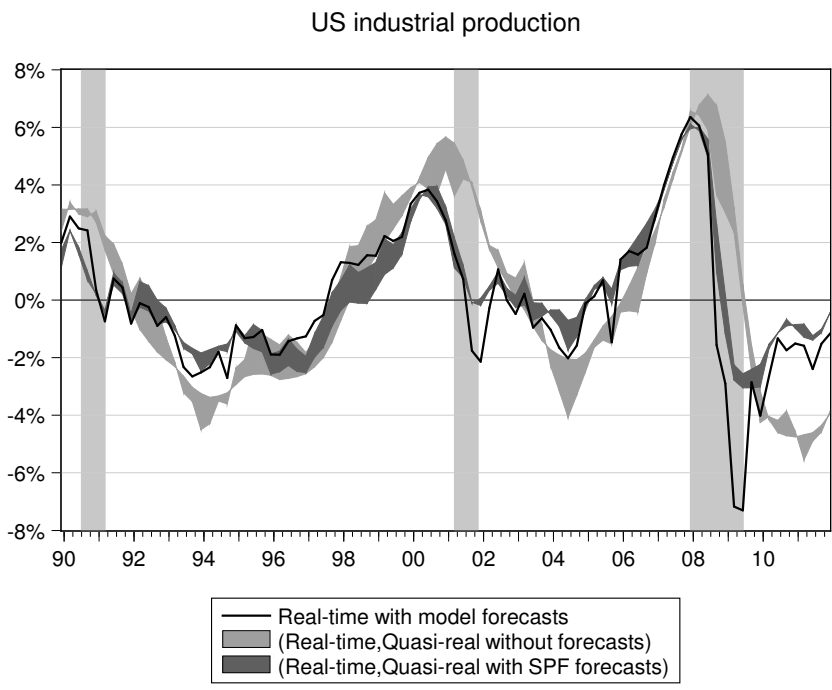
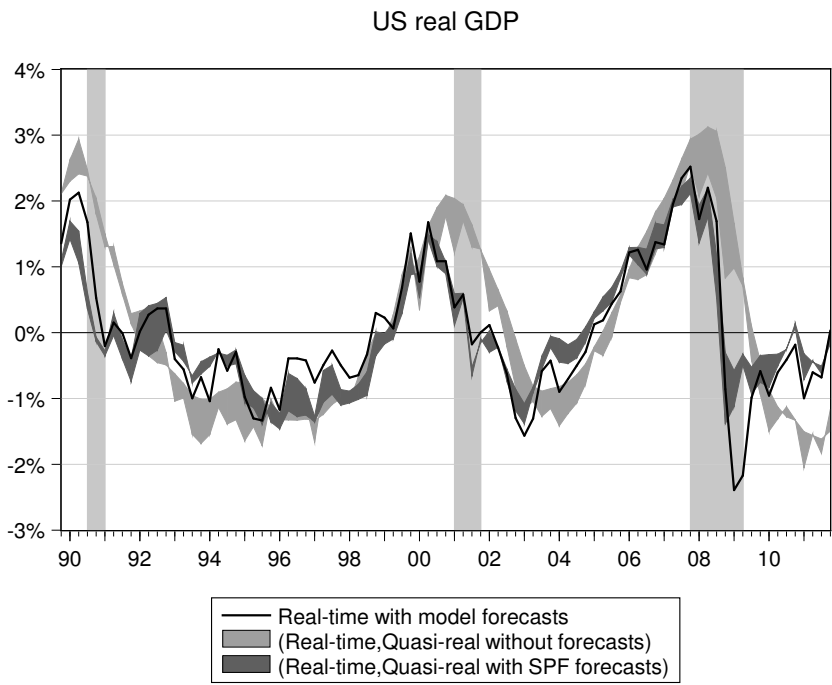
Previous results in the literature have often indicated that the effects of data revisions play a minor role in the determination of the uncertainty surrounding real-time estimates of trend and gap (Orphanides and van Norden, 2002; Marcellino and Musso, 2011; Ince and Papell, 2013). Our results corroborate this view as it can be seen from the generally thin bands between the real-time and the quasi-real measurement errors, presented in Figure 3. Most importantly, inspection of the dynamics of these series provides an appealing explanation for the improved performance we find in the reliability of real-time gap estimates with the forecast-augmented HP filter. Namely, we can see in Figure 3 that the errors incurred by the filtered estimates without the resource of the forecasts are anticipated by those obtained with the forecast-augmented HP filter. This effect seems particularly relevant for the periods preceding the start of recessions as dated by the Business Cycle Dating Committee (BCDC) of the National Bureau of Economic Research (NBER), highlighted in our figures with light gray shaded areas.

Figure 2: Final vs. real-time HP filter estimates of the US gap.



The light gray shaded areas refer to recessions (from peak to trough) as dated by the NBER's BCDC.

Figure 3: Measurement errors to real-time/quasi-real HP filter estimates of the US gap.



The area bands represent the differences between the measurement errors in the real-time and the quasi-real gap estimates. These can be drawn in the case of model forecasts but are not herein presented for reasons of clarity. The light gray shaded areas refer to recessions as dated by the NBER's BCDC.

## 4.2 Statistical results

We now turn our analysis to an assessment based on statistics calculated over the gap and measurement errors concepts defined above. Descriptive statistics for the gaps estimated by each specification of the HP filter and their measurement errors are summarized in Tables 2a and 2b, respectively. Firstly, we can confirm our observation from the visual analysis that there are only small differences between the properties the quasi-real and the real-time gap concepts and their corresponding measurement errors. Moreover, there is a clear increase in the magnitudes of the gap estimates obtained with the IP series compared with those obtained with the series of real GDP. This result, however, is not surprising: recall that from the model-based representation of the HP filter the variance of the gaps is directly related to the value of the smoothing parameter, which is higher for the series in the monthly frequency.

The standard deviations of the estimated gaps point out an interesting result in the usage of the forecast-augmented specifications. Particularly, notice that the short-run volatility implied by the gap estimates is substantially reduced under the latter specifications in relation to what is obtained under the standard HP filter. Though the less volatile gaps obtained with the forecasts for real GDP seem to get closer to the volatility obtained for the final estimates, it is unclear whether this additional smoothness is desirable. From a policymaker perspective, at the same time that a quickly responsive measure of economic activity may be useful to accelerate the diagnosis of macroeconomic malfunctioning, it also carries the perils of triggering unnecessary interventions too often.

To obtain a clearer assessment on the quality of the HP filtered gap estimates obtained in real-time we look at some reliability indicators<sup>11</sup>, presented in Table 2. The first of these indicators is a noise-to-signal ratio (NSR), calculated as the ratio between the measurement errors root mean squares (RMS) and the standard deviation of the gap final estimates. As such, this measure captures the average size of the real-time/quasi-real measurement

---

<sup>11</sup>Formulae used for the calculation of these indicators are presented in the Appendix.

Table 1: Descriptive statistics for estimated gaps and their measurement errors.

(a) Series of gaps.

Series	US real GDP				US industrial production			
	Mean	Min.	Max.	St.Dev.	Mean	Min.	Max.	St.Dev.
Final gaps	-0.11	-3.01	2.37	1.22	-0.26	-10.38	6.38	2.93
Quasi-real gaps								
No forecasts	-0.34	-4.41	1.57	1.32	-0.41	-11.56	4.75	3.31
Model forecasts	-0.22	-2.25	0.65	0.59	-0.27	-6.08	2.02	1.42
SPF forecasts	-0.12	-2.68	1.14	0.79	-0.51	-7.82	2.55	1.92
CEF forecasts	-0.12	-1.91	0.99	0.66	-0.40	-5.34	2.32	1.51
Real-time gaps								
No forecasts	-0.11	-3.73	1.78	1.26	-0.49	-10.70	5.06	3.13
Model forecasts	-0.13	-1.89	0.90	0.63	-0.28	-4.80	2.41	1.29
SPF forecasts	-0.08	-2.52	0.99	0.73	-0.57	-7.34	1.84	1.73
CEF forecasts	-0.08	-1.70	0.77	0.57	-0.46	-4.26	1.46	1.28

(b) Series of measurement errors.

Series	US real GDP				US industrial production			
	Mean	Min.	Max.	RMS	Mean	Min.	Max.	RMS
Quasi-real errors								
No forecasts	0.23	-1.62	3.12	1.39	0.15	-4.75	7.13	3.19
Model forecasts	0.11	-2.38	2.73	1.06	0.01	-6.49	6.40	2.44
SPF forecasts	0.01	-1.46	2.35	0.92	0.25	-2.57	6.01	2.10
CEF forecasts	0.01	-1.56	2.33	0.97	0.14	-5.98	6.10	2.33
Real-time errors								
No forecasts	0.00	-2.05	2.94	1.32	0.23	-5.55	6.58	3.06
Model forecasts	0.02	-2.39	2.52	1.03	0.03	-7.30	6.35	2.50
SPF forecasts	-0.03	-1.39	2.11	0.87	0.31	-3.05	6.11	2.04
CEF forecasts	-0.03	-1.85	2.10	0.92	0.20	-6.62	6.13	2.27

RMS denotes the root mean squares of the corresponding error series. All statistics are presented in percentage points and are computed covering the period from 1989q4 to 2011q4 for the quarterly series on the US real GDP, and every three months from 1989m12 to 2011m12, for the monthly series on the US index of industrial production.

Table 2: Reliability indicators of real-time/quasi-real estimates of the US gap.

Series	US real GDP			US industrial production		
	NSR	Corr.	Op.Sign	NSR	Corr.	Op.Sign
Quasi-real estimates						
No forecasts	1.14	42%	53%	1.09	48%	45%
Model forecasts	0.87	51%	36%	0.83	56%	43%
SPF forecasts	0.75	66%	30%	0.72	70%	29%
CEF forecasts	0.80	61%	29%	0.79	62%	31%
Real-time estimates						
No forecasts	1.08	44%	48%	1.05	49%	52%
Model forecasts	0.85	53%	31%	0.85	53%	43%
SPF forecasts	0.71	71%	26%	0.70	74%	27%
CEF forecasts	0.75	70%	25%	0.78	68%	29%

NSR (noise-to-signal ratio) is calculated as the ratio of the measurement errors RMS to the standard deviation of the final gap estimates. Corr. (correlation) denotes the correlation between the real-time/quasi-real gap estimates and the final gap estimates. Op. Sign (opposite sign) denotes the frequency that the real-time/quasi-real gap estimates are found with a sign different from that of the final gap estimate. The periods covered for each variable are the same as indicated in the notes of Table 1.

errors relative to the average size of the final gap estimates. As the results in Table 2 reveal, the size of the errors in the real-time and the quasi-real estimates obtained with the HP filter without forecasts are in general as big as the magnitude of the final gap estimates. The use of the monthly IP series tends to only slightly reduce the relative size of these measurement errors. The best performing specification is clearly that of using the SPF forecasts to augment the HP filter, which overall was able to reduce the NSRs to around 66% of what is obtained without forecasts.

The other two reliability indicators presented in Table 2 are based directly on the series of gap estimates: Corr. measures the correlation between the real-time/quasi-real gap estimates and the final estimates, whereas Op.Sign presents the frequency within which real-time/quasi-real gap estimates and the final gap estimates fall in different sides in relation to the long-run trend. Clearly, these two indicators confirm our findings favorable to the use of the SPF forecasts to augment the HP filter. For example, whereas using the standard HP filter one would have mistakenly inferred that the economy is currently operating

below/above its long-run trend about half the time this measurement is taken in real-time; with the assistance of the SPF forecasts these mistakes could have been reduced to only  $1/4$  of the times.

As we have previously alluded, the improvements to the reliability of real-time gap estimates brought by the forecast-augmentation of the HP filter are clearly related to the quality of the forecasts in use. Particularly, it seems compelling to argue that the better these forecasts are in signaling future changes in the final gap estimates, the better the improvements they should be able to bring to the real-time reliability of gap estimates. To complement our results regarding this point, we present in Figure 4 the correlations between the different forecasts we experimented with and the changes in the final gap measures. Consistent to our findings in this section, these correlations highlight the superiority of the survey forecasts in anticipating the movements in the final estimates of the gap: whereas the SPF forecasts provide the best signals in the shorter horizons, the model forecasts clearly degrade over the longer horizons. Thus, our results have been generally favorable to the more parsimonious approach of using survey forecasts to improve real-time measurements of economic activity with the HP filter.

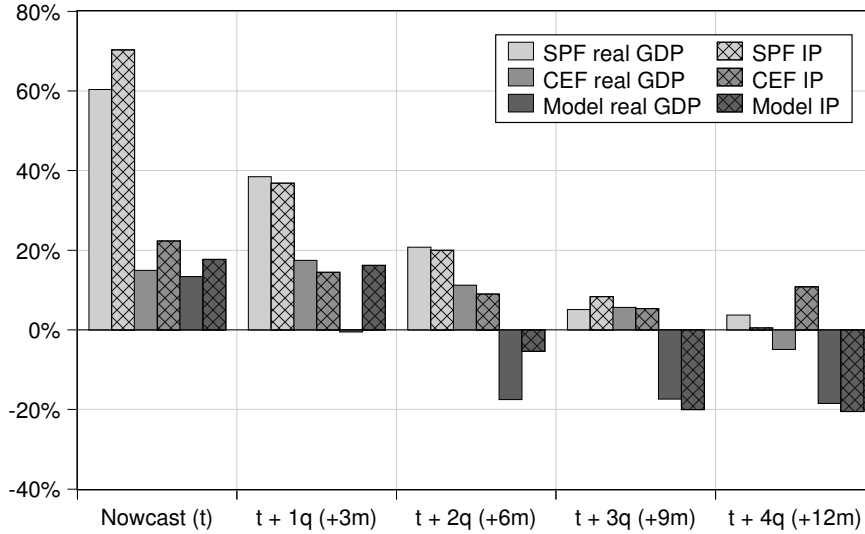
## **5 Concluding remarks**

In this paper we have presented renewed results in the reliability of real-time trend/gap decomposition by the use of the famous Hodrick-Prescott filter. We have shown, through two empirical applications with quarterly and monthly data for the US, that expanding the time horizon of the real-time pre-filtered series with survey forecasts considerably attenuates the end-of-sample uncertainty associated with these estimates.

As a benchmark for the use of survey forecasts we have also augmented the HP filter with model-based forecasts, as previously suggested in the literature. Despite turning the filtering procedure into a joint forecasting problem, our results provide favorable support to



Figure 4: Correlations between forecasts (growth) and final gap changes.



The correlations are calculated between the growth forecasts associated to each model/survey and the change in the final gap estimate (i.e.,  $\Delta g_{t|T}$ ) during the forecasted period.

the general approach of forecast-augmentation of the HP filter. A key element in this approach is the need of optimal forecasts. The use of survey forecasts, incorporating market participants judgements, turned out to be a “cheaper” and more parsimonious alternative to the hard task of approximating optimal forecasts for empirical purposes.

## A Appendix

### A.1 Survey forecasts interpolation formulae

**CEF:** Letting  $\hat{y}_{cy,t}$  and  $\hat{y}_{fy,t}$  denote the CEF growth forecasts for the current and the following years, respectively, we interpolate these forecasts in two steps:

1. We calculate, depending on the base period of the forecast, the constant growth rate ( $\hat{y}_{cy,t}$ ) which, when applied to the current year’s remaining months, results into an end-of-year growth rate equal to that predicted by the average CEF fore-

cast; i.e., formally we have

$$y_{t-1} (1 + \hat{y}_{cy,t})^{F-r} = y_{py,t} (1 + \hat{y}_{cy,t}) \iff \hat{y}_{cy,t} = \left[ \frac{y_{py,t} (1 + \hat{y}_{cy,t})}{y_{t-1}} \right]^{\frac{1}{F-r}} - 1, \quad (9)$$

where  $y_{t-1}$  is the last observation available in the vintage corresponding to the forecast,  $F$  is the frequency of observation (4 for real GDP, and 12 for IP), and  $r$  is the number of remaining periods of the year.

2. We compute the implied forecasts for the levels of each variable compounding  $\hat{y}_{cy,t}$  to the last observation available until the end of the current year is reached. From that point on we use the  $\hat{y}_{fy,t}$  growth forecasts instead, i.e.,

$$\begin{cases} \hat{y}_{t+h} = y_{t-1} (1 + \hat{y}_{cy,t})^{h+1}, & \text{if } r + h < F, \\ \hat{y}_{t+h} = \hat{y}_{t+h-1} (1 + \hat{y}_{fy,t}), & \text{otherwise,} \end{cases} \quad (10)$$

with  $h = 0, \dots, 4$  for the quarterly series, and  $h = 0, \dots, 14$  for the monthly series. Hence, though the number of forecasts varies with the frequency of observations, the time periods covered in the augmented series are the same for both variables.

**SPF:** We interpolate the SPF forecasts for IP arithmetically<sup>12</sup>, again in two steps:

1. We calculate the monthly in/decrement ( $\hat{y}_{q+h}$ ), in relation to the IP value observed (or forecasted for  $h > 0$ ) in the last month of the previous quarter ( $y_{q-1,m=3}$  for  $h = 0$ , or  $\hat{y}_{q+h-1,m=3}$  for  $h > 0$ ), necessary to make the average of the three months of forecasts equal to the average predicted by the SPF forecast for that quarter

---

<sup>12</sup>Alternatively we could have adopted a geometric interpolation so as to keep a constant rate of growth within each quarter, though that would require solving a cubic equation.

$(\hat{y}_{q+h})$ , i.e.,

$$\begin{cases} \hat{y}_{q+h,m=0} = y_{q-1,m=3}, & \text{if } h = 0, \\ \hat{y}_{q+h,m=0} = \hat{y}_{q+h-1,m=3}, & \text{if } h > 0, \end{cases} \quad (11)$$

$$\frac{\sum_{i=1}^3 \hat{y}_{q+h,m=0} + i\hat{y}_{q+h}}{3} = \hat{y}_{q+h} \iff \hat{y}_{q+h} = \frac{\hat{y}_{q+h} - \hat{y}_{q+h,m=0}}{2}.$$

2. We compute the IP monthly forecasts associated with each quarter SPF forecast accumulating  $\hat{y}_{q+h}$  recursively, taking as initial value the IP observation (forecast) in the last month of the quarter preceding that of forecast ( $\hat{y}_{q+h,m=0}$ ), so

$$\hat{y}_{q+h,m=i} = \hat{y}_{q+h,m=i-1} + \hat{y}_{q+h}, \quad (12)$$

for  $i = 1, 2, 3$ , and  $h = 0, \dots, 4$  reflecting the five quarterly horizons available in the SPF.

## A.2 Reliability indicators formulae

Let  $g_{t|t}$  and  $g_{t|T}$  denote the real-time (or quasi-real) and the final series of gap estimates, respectively. The real-time (quasi-real) measurement errors are given by  $g_{t|t} - g_{t|T}$ . Our reliability indicators are calculated according to the following expressions:

$$\text{Corr.} = \frac{\sum_{i=1}^T (g_{i|i} - \overline{g_{i|i}}) (g_{i|T} - \overline{g_{i|T}})}{\sqrt{\sum_{i=1}^T (g_{i|i} - \overline{g_{i|i}})^2 \sum_{i=1}^T (g_{i|T} - \overline{g_{i|T}})^2}}, \quad (13)$$

where  $\overline{g_{i|i}}$  and  $\overline{g_{i|T}}$  are the series averages,

$$\text{Op.Sign} = \frac{\sum_{i=1}^T |(\text{sign}(g_{i|i}) \neq \text{sign}(g_{i|T}))|}{T}, \quad (14)$$

where  $I(\bullet)$  is the indicator function assuming the value of 1 if the argument is true and 0 otherwise,

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^T (g_{i|i} - g_{i|T})^2}{T}}, \quad (15)$$

and

$$\text{NSR} = \frac{\text{RMS}}{\sqrt{\sum_{i=1}^T (g_{i|T} - \bar{g}_{i|T})^2 / T}}. \quad (16)$$

## References

- Ager, P., M. Kappler, and S. Osterloh (2009). The accuracy and efficiency of the consensus forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting* 25(1), 167–181.
- Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese (2010). New eurocoin: Tracking economic growth in real time. *The Review of Economics and Statistics* 92(4), 1024–1034.
- Aruoba, S. B., F. X. Diebold, and C. Scotti (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics* 27(4), 417–427.
- Batchelor, R. (2001). How useful are the forecasts of intergovernmental agencies? the imf and oecd versus the consensus. *Applied Economics* 33(2), 225–235.
- Blackburn, K., F. Orduna, and M. Sola (1995). Exponential smoothing and spurious correlation: a note. *Applied Economics Letters* 2(3), 76–79.
- Burman, J. P. (1980). Seasonal adjustment by signal extraction. *Journal of the Royal Statistical Society. Series A (General)* 143(3), pp. 321–337.
- Canova, F. (1998, May). Detrending and business cycle facts. *Journal of Monetary Economics* 41(3), 475–512.

- Christiano, L. J. and T. J. Fitzgerald (2003, 05). The band pass filter. *International Economic Review* 44(2), 435–465.
- Clements, M. P. and A. B. Galvão (2012). Improving real-time estimates of output and inflation gaps with multiple-vintage models. *Journal of Business & Economic Statistics* 30(4), 554–562.
- Cogley, T. and J. M. Nason (1995). Effects of the hodrick-prescott filter on trend and difference stationary time series implications for business cycle research. *Journal of Economic Dynamics and Control* 19(1-2), 253 – 278.
- Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of Economic Literature* 49(1), 72–100.
- Croushore, D. and T. Stark (2001). A real-time data set for macroeconomists. *Journal of Econometrics* 105(1), 111 – 130.
- Danthine, J.-P. and M. Girardin (1989, January). Business cycles in switzerland : A comparative study. *European Economic Review* 33(1), 31–50.
- Ehlgren, J. (1998). Distortionary effects of the optimal hodrick-prescott filter. *Economics Letters* 61(3), 345 – 349.
- Garratt, A., K. Lee, E. Mise, and K. Shields (2008). Real-time representations of the output gap. *The Review of Economics and Statistics* 90(4), 792 – 804.
- Gomez, V. (1999, January). Three equivalent methods for filtering finite nonstationary time series. *Journal of Business & Economic Statistics* 17(1), 109–16.
- Gomez, V. (2001, July). The use of butterworth filters for trend and cycle estimation in economic time series. *Journal of Business & Economic Statistics* 19(3), 365–73.
- Graff, M. and J.-E. Sturm (2012). The information content of capacity utilization rates for output gap estimates. *CESifo Economic Studies* 58(1), 220–251.

- Harvey, A. C. and A. Jaeger (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics* 8(3), pp. 231–247.
- Harvey, A. C. and T. M. Trimbur (2003, 03). General model-based filters for extracting cycles and trends in economic time series. *The Review of Economics and Statistics* 85(2), 244–255.
- Hodrick, R. J. and E. C. Prescott (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking* 29(1), 1–16.
- Ince, O. and D. H. Papell (2013). The (un)reliability of real-time output gap estimates with revised data. *Economic Modelling* 33(C), 713–721.
- Isiklar, G. and K. Lahiri (2007). How far ahead can we forecast? evidence from cross-country surveys. *International Journal of Forecasting* 23(2), 167–187.
- Kaiser, R. and A. Maravall (1999). Estimation of the business cycle: A modified hodrick-prescott filter. *Spanish Economic Review* 1, 175–206.
- Kaiser, R. and A. Maravall (2001). *Measuring Business Cycles in Economic Time Series*, Volume 154 of *Lecture Notes in Statistics*. Springer.
- King, R. G. and S. T. Rebelo (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control* 17(1-2), 207 – 231.
- Leduc, S. and K. Sill (2013). Expectations and economic fluctuations: an analysis using survey data. *The Review of Economics and Statistics* 95(4), 1352 – 1367.
- Maravall, A. and A. del Río (2007). Temporal aggregation, systematic sampling, and the hodrick-prescott filter. *Computational Statistics and Data Analysis* 52(2), 975 – 998.
- Marcellino, M. and A. Musso (2011). The reliability of real-time estimates of the euro area output gap. *Economic Modelling* 28(4), 1842–1856.

- Marcellino, M., J. H. Stock, and M. W. Watson (2006). A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics* 135(1-2), 499 – 526.
- Mise, E., T.-H. Kim, and P. Newbold (2005). On suboptimality of the hodrick-prescott filter at time series endpoints. *Journal of Macroeconomics* 27(1), 53 – 67.
- Orphanides, A. and S. van Norden (2002). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics* 84(4), 569–583.
- Pedersen, T. M. (2001, August). The hodrick-prescott filter, the slutzky effect, and the distortionary effect of filters. *Journal of Economic Dynamics and Control* 25(8), 1081–1101.
- Planas, C. and A. Rossi (2004). Can inflation data improve the real-time reliability of output gap estimates? *Journal of Applied Econometrics* 19(1), 121–133.
- Prescott, E. C. (1986, January). Theory ahead of business-cycle measurement. *Carnegie-Rochester Conference Series on Public Policy* 25(1), 11–44.
- Proietti, T. (2009). On the model-based interpretation of filters and the reliability of trend-cycle estimates. *Econometric Reviews* 28(1-3), 186–208.
- Ravn, M. O. and H. Uhlig (2002). On adjusting the hodrick-prescott filter for the frequency of observations. *The Review of Economics and Statistics* 84(2), 371–375.
- Stark, T. (2010). Realistic evaluation of real-time forecasts in the survey of professional forecasters. Research rap special report, Federal Reserve Bank of Philadelphia.
- Stark, T. and D. Croushore (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics* 24(4), 507 – 531.
- Valle e Azevedo, J. (2011). A multivariate band-pass filter for economic time series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 60(1), 1–30.

Valle e Azevedo, J., S. J. Koopman, and A. Rua (2006). Tracking the business cycle of the euro area: A multivariate model-based bandpass filter. *Journal of Business and Economic Statistics* 24, 278–290.

Watson, M. W. (2007). How accurate are real-time estimates of output trends and gaps? *Economic Quarterly* 93(3), 143–161.