The University of Manchester

MANCHESTER 1824

Centre for Growth & Business Cycle Research

CGBCR

Discussion Paper Series

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

By

Jaqueson K. Galimberti, Marcelo L. Moura

Centre for Growth and Business Cycle Research, Economic Studies, University of Manchester, Manchester, M13 9PL, UK

September 2011
Number 159

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

Jaqueson K. Galimberti

University of Manchester and The Capes Foundation

Marcelo L. Moura

Insper - Institute of Education and Research

Abstract

Incorporating survey forecasts to a forecast-augmented Hodrick-Prescott filter, we evidence a considerable improvement to the reliability of US output-gap estimation in real-time. Odds of extracting wrong signals of output-gap estimates are found to reduce by almost a half, and the magnitude of revisions to these estimates accounts to only three fifths of the output-gap average size, usually an one-by-one ratio. We further analyze how this end-of-sample uncertainty evolves as time goes on and observations accumulate, showing that a 90% rate of correct assessments of the output-gap sign can be attained with five quarters of delay using survey forecasts.

I. Introduction

The reliability of filtering methods to decompose a series into 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). Given the policy relevance of the accuracy of these assessments (see Croushore, 2011, pp. 95-6), efforts have been directed towards the search for ways to attenuate this uncertainty.

The main literature attempting to improve the real-time accuracy on filtering has been moving towards the incorporation of an increasing amount of information into the estimation procedure. Such attempts can be characterized by the employment of both statistically or economically-grounded priors in the search for signals correlated to the unobserved components being estimated. The most prominent attempts involve expanding the information set supplied to the filter in the temporal dimension with the use of forecasts (Kaiser and Maravall, 1999, 2001; Mise et al., 2005), in the covariates dimension with the employment of model-based multivariate representations (Planas and Rossi, 2004; Valle e Azevedo et al., 2006; Altissimo, 2010; Azevedo, 2011; Marcellino and Musso, 2011), or jointly in both of these dimensions (Garratt et al., 2008).

Departing from the previous literature, we propose and evaluate the use of forecasts from surveys, instead of those obtained from fitted models, to improve the reliability of real-time trend/gap decomposition. To keep the presentation succinct, we focus on the Hodrick-Prescott (HP) filter (Hodrick and Prescott, 1997), which despite of being subject to several criticims (Harvey and Jaeger, 1993; King and Rebelo, 1993; Blackburn et al., 1995; Cogley and Nason, 1995; Ehlgen, 1998, between others) it is one of the mostly used tools for business cycle empirical analysis. Besides, though the realm of the HP-filter has been threatened by the development of sophisticated band-pass filters (Christiano and Fitzgerald, 2003), these latter are also subject to the issue of the end-of-sample uncertainty as evidenced by the analysis of Watson (2007).

Overall, we find that the procedure of augmenting the HP-filter with survey forecasts is successful in attenuating end-of-sample uncertainty of output-gap estimates. We further enhance our understanding of how this uncertainty evolves as time goes on and new observations become available by proposing an evaluation of the lagged measures of output-gap obtained in real-time, which is an innovative focus to the analyses usually found in the related literature. In such an exercise we found, e.g., that it takes about eight periods of new observations, and re-estimations, for the end-of-sample uncertainty to be reduced to 25% of the level obtained in real-time.

Although the usefulness of survey forecasts can be interpreted into the statistical context of an improvement to the filter's information set, as in the previous literature, we can also support our results on the grounds of the role played by expectations in macroeconomic theory. Incorporating market participants judgements, data collected from professional forecasters into the form of surveys can be taken as proxies for those expectations that eminently introduce a self-referential[1] feature into the determination of an economy's dynamics. Drawing a parallel to the recent literature modelling expectations through adaptive learning (Evans and Honkapohja, 2001), our results can be seen as corroborating the recent findings that expectations shocks, as measured from survey data on forecasts, can help improving our understanding of business cycle fluctuations (see Eusepi and Preston, 2008, and Milani, 2011, for a recent account).

The remainder of this paper is organized as follows. Next section outlines the forecast-augmented HP-filter, which we use to obtain and evaluate the real-time estimates of the US output-gap in section III. We then conclude this paper with Section IV.

II. A forecast-augmented Hodrick-Prescott filter

Under the standard HP-filter 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 $\{g_t\}_{t=1}^{T}$ we solve the programming problem of choosing values for the first of these so as to minimize

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

[1]Self-referentiality in this context means that the actual path of the economy is determined in part by the path expected by its composing agents.

where Δ^2 stands for a twice-differenced lag operator[2], and λ is a parameter regulating the trade-off between the fit to the data, penalized in the first summation, and the smoothness of the trend, penalized in the second summation. There are other interpretations of λ depending on the rationalization of the HP-filter, the most famous perhaps being those of Harvey and Jaeger (1993) and King and Rebelo (1993), where it stands as the ratio between the variances of the gap and the change in the trend growth rate, but see also Kaiser and Maravall (1999, pp. 177-9).

For quarterly data it is customary to set λ equal to $1,600$, a value at which the HP-filter eliminates fluctuations at frequencies lower than about 8 years (Cooley and Prescott, 1995, pp. 27-8). For other frequencies some controversy exist on how much should this parameter be adjusted in order to keep time aggregation of the decomposed components consistent in view of the distortionary effects introduced by the filter (see Pedersen, 2001; Ravn and Uhlig, 2002; Maravall and del Río, 2007).

Our focus, however, is on the (in)accuracy of the HP-filter estimates of the trend/gap at the end-points[3] of the available sample of observations. By solving (1) to obtain an estimator of these components the HP-filter is a symmetric filter in the sense that the estimator of x_t uses both lags and leads of y_t. Besides, the second-differenced term in (1) entwines all the estimates of x_t with the observed values of y_t over the full sample. To see this, let ∇^2 denote a twice-differenced lead operator[4] and $x_{t|T}$ stand for the estimate of x_t conditional on information observable at T, and take the first order condition (FOC) to the minimization problem (1) at a mid-point t of the sample, which rearranging yields

$$y_t = x_{t|T} + \lambda \nabla^2 \Delta^2 x_{t|T}. \tag{2}$$

Clearly, (2) reveals how the HP procedure smooths the gross series by filtering its second-differences symmetrically and weighted by the λ parameter. We can compare this mid-point

[2]Specifically, $\Delta^2 g_t = (g_t - g_{t-1}) - (g_{t-1} - g_{t-2})$.

[3]The same issue and arguments presented here for the end-of-sample are applicable to the beginning-of-sample, but our focus goes solely to the former given our emphasis on real-time.

[4]Specifically, $\nabla^2 x_t = (x_{t+2} - x_{t+1}) - (x_{t+1} - x_t)$.

FOC to that at the end-point T of the sample,

$$y_T = x_{T|T} + \lambda \Delta^2 x_{T|T}, \tag{3}$$

and see that the higher uncertainty associated to the end-of-sample estimates of the trend/gap comes partly from the inability of the estimator to keep up with symmetry at these points due to the lack of forward information, i.e., an estimator-led uncertainty. Furthermore, once $T+2$ data becomes available the FOC for period T turns into the symmetric form as well,

$$y_T = x_{T|T+2} + \lambda \nabla^2 \Delta^2 x_{T|T+2}. \tag{4}$$

Comparing (3) to (4) we can see that the remaining part of the end-of-sample revision comes from the incorporation of new information from the newly available data (y_{T+1} and y_{T+2}), and we denote the associated uncertainty, due to the lack of this information, by data-led uncertainty. Notice, however, that this updating of the information set does not affect exclusively the sample's end-point estimates but also the mid-point estimates of trend/gap, though these latter have a lower sensitivity to this effect due to the exponential decay in the weights given to observations faraway of the estimator's current point. In spite of this (rough) theoretical distinction between the sources of the end-of-sample uncertainty of trend/gap estimates, at this stage we do not try to disentangle the sizes of each of these empirically and leave its precise identification for future research[5].

There are two natural ways to deal with this end-of-sample uncertainty on the trend/gap estimates obtained from the HP-filter. The first is to expand the information set supplied to the filter by the use of forecasts (Kaiser and Maravall, 1999, 2001; Mise et al., 2005; Garratt et al., 2008). The second involves modifying the formulation of the programming problem in (1) by re-scaling the parameter λ at the sample end-points so as to compensate the lack of symmetry with a higher penalty to deviations (Bruchez, 2003). This last approach has some similarities to the recent idea of augmenting the HP-filter with judgement (Jönsson, 2010), but these still lack a formal evaluation of optimality and how the incorporation of these priors may magnify the

[5]From a theoretical perspective, Monte Carlo exercises should provide a good tool for this task. Empirical identification, as usual, will be a more challenging issue.

inducement of spurious cycles in the filtered times series (Harvey and Jaeger, 1993; Blackburn et al., 1995; Cogley and Nason, 1995). For these reasons we adopt here the first approach.

Under the same circunstances as in the original formulation, suppose we have $k = 1,..,K$ forecasts of y_{T+k} at our disposal. Then a forecast-augmented HP-filter can be obtained as the solution to the problem of choosing $\left\{x_t^f\right\}_{t=1}^{T+K}$ so as to minimize

$$L\left(y_t, \lambda, x_t^f\right) + \sum_{t=T+1}^{T+K} \left(y_t^f - x_t^f\right)^2 + \lambda \sum_{t=T+1}^{T+K} \left(\Delta^2 x_t^f\right)^2, \tag{5}$$

where y_t^f stands for the forecasted value of y_t. Clearly, the use of forecasts in (5) will attenuate the end-of-sample bias introduced by the asymmetry of the filter in the estimate of x_T based solely on observable information. Notice, however, that by turning the filtering problem into a joint forecasting problem preservation of the optimality properties of the filter will depend on the optimality of the forecasts. From our previous discussion, it should now be clear that augmenting the HP-filter with inaccurate forecasts can have hazardous effects on the reliability of the trend/gap estimates if the negative effects of incorporating wrong signals is not offset by the gains from symmetry, specially at the end-of-sample but not exclusively. We now turn our focus to an empirical evaluation of this approach to deal with end-of-sample uncertainty of trend/gap estimates from the HP-filter.

III. Empirical application to US GDP

A. Data and definitions

In the spirit of Orphanides and van Norden (2002) we examine here the reliability of US GDP detrending in real-time by the standard HP-filter and compare it to the forecast-augmented version of this filter. Our definition of real-time data, however, does not correspond exactly to the usual meaning of the word as refering to the data/revision available at the time at which the estimation is carried for. Instead, our computations use an unique vintage of data from 2011:q2 such that our results are cleared from the effects of data revisions[6], which in most cases have

[6]In the notation of Orphanides and van Norden (2002, p. 571) our measure of real-time estimate refers to their quasi real estimate, i.e., “the rolling estimate based on the final data series”.

been found to play a minor role into real-time uncertainty (Orphanides and van Norden, 2002; Ince and Papell, 2010; Marcellino and Musso, 2011)[7]. Despite of that, we make the realistic assumption that data is observed with a 1-period delay[8], i.e., when estimating the output-gap at T the only historic values of the output series that are observable are those behind and up to $T-1$, such that the first forecast ($k=1$) to be used by the forecast-augmented HP-filter can actually be thought as a nowcast.

The output measure which we detrend consists of seasonally adjusted quarterly data on the US real GDP (in logs) from 1960:q1 to 2010:q4. Our data on this series come from the Philadelphia's Fed Real-Time Data Research Center. Other than being a thoroughly documented data set (Croushore and Stark, 2001), this source has also been showed fruitful into forecast evaluation contexts (Stark and Croushore, 2002). For the forecast-augmented HP-filter we extend the real-time horizon of our pre-filter series by adding one quarter of nowcast and four quarters of forecasts, where for the computation of these we adopt two approaches.

The first is to use three of the benchmark forecasting models as specified in Stark (2010), which consist of: (i) a no-change (NC), or random walk, model where the forecasts for real GDP growth are equal to this same measure from the previous quarter; (ii) an indirect autoregression (IAR) model where the forecasts for real GDP growth are obtained from the estimation of a one-period univariate autoregression and the forecasts are computed iteratively; (iii) a direct autoregression (DAR) model which modifies the previous model by estimating a different equation for each forecast horizon. Further descriptions about the IAR and DAR models can be found in Marcellino et al. (2006), and details on their estimation and adjustments to the gross data can be found in Stark (2010)[9].

The second approach is to use survey forecasts which may account for information not included in the historic values of the series used by the benchmark models. We use data from the Survey of Professional Forecasters (SPF), which are put available by the Philadelphia's Fed

[7]Cayen and van Norden (2005) analyze Canadian output-gap estimates and their findings suggest a slightly higher role for data revision.

[8]For the US quarterly GDP data, the Bureau of Economic Analysis (BEA) releases its first measure (Advance Report) by the end of the first month subsequent to the quarter, while its first revision is released by the end of the second month. See Figure 1 in Stark (2010, p. 15).

[9]All the data we used (including benchmark models forecasts) were constructed based on the data set of growth rates from Stark (2010), to whom we are grateful for putting them available through Philadelphia's Fed website. A worksheet with our own re-construction of level indices is readily available upon author's request.

as well and have forecast horizons compatible with the specifications of our benchmark models. Notice that, in contrast to our quasi real-time detrending procedure, the benchmark models are estimated using real-time data/revision in order to subject them to a information environment similar to that faced by a survey forecaster.

Given that SPF's series of forecasts for real GDP starts at 1968:q4, we start the recursive computations of the real-time HP-filter taking the realized values of output from 1960:q1 to 1968:q3 (35 obs.) as known. Then, computing the HP-filter as if we were in real-time from 1968:q4 and onwards we obtain a total of 169 quarters of real-time trend/gap estimates. Full-sample estimates are obtained by applying the filter to the whole series, and the revisions to the real-time estimates are computed as the difference between the full-sample estimates and those obtained in real-time. To facilitate the distinction between real-time and full-sample estimates, let $g_{t|T,K}$ stand for the output-gap estimate of g_t conditional on observations of $\{y_t\}_{t=1}^{T}$ and $\left\{y_{T+k}^{f}\right\}_{k=1}^{K}$. Then, if we have a total of T observations of output and K real-time forecasts, the full-sample, the standard real-time, and the forecast-augmented real-time estimates of the output-gap correspond to the sequences $\left\{g_{t|T,0}\right\}_{t=1}^{T}$, $\left\{g_{t|t,0}\right\}_{t=1}^{T}$, and $\left\{g_{t|t,K}\right\}_{t=1}^{T}$, respectively.

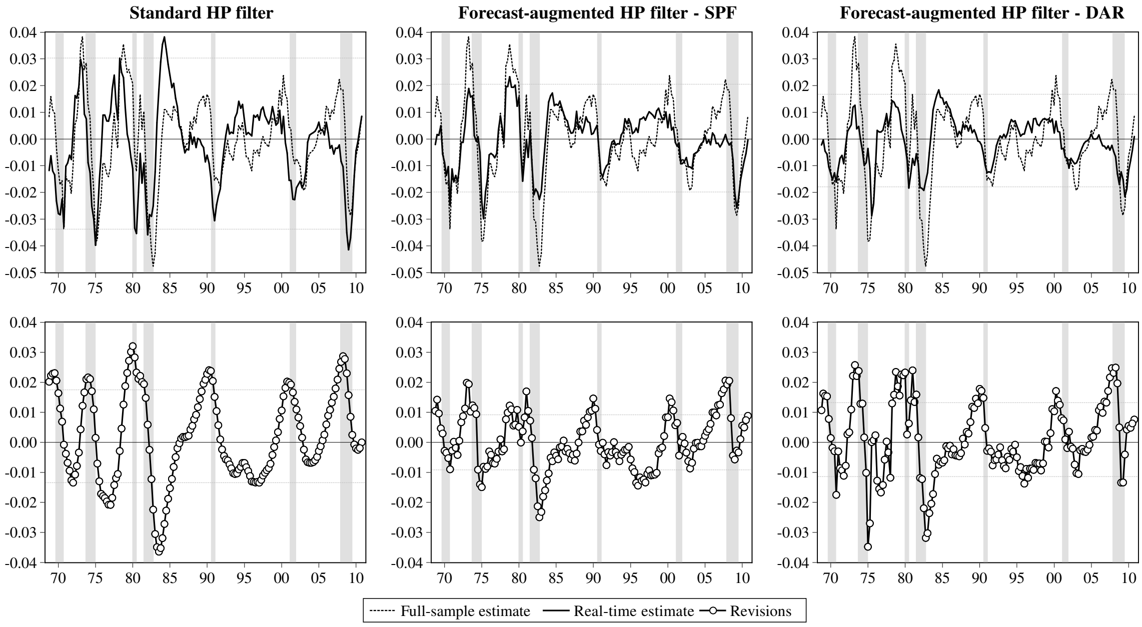
B. Results

We plot the estimates of the US output-gap in figure 1, with the first row of panels comparing the full-sample and the real-time estimates as obtained from the standard, the SPF's forecast-augmented, and the DAR's forecast-augmented HP-filters in columns, respectively[10].

The end-of-sample uncertainty into real-time estimates of output-gap can be observed into figure 1 by the detachment of the real-time estimates from those obtained with the full-sample. The difference between these series defines the concept of revisions to real-time estimates, which are plotted into the second row of panels in figure 1. Although the real-time estimates seem to move onto the same directions through time, notice how the volatility of the real-time output-gap estimates and, most importantly, of their revisions (two standard deviations bands are plotted as gray-colored horizontal lines) decrease with the use of the forecast-augmented

[10]Our graphical presentation abstracts from the plots for the NC and IAR forecast-augmented HP-filters, as the first has a poorer performance in general and the second behaves similarly to the DAR forecast-augmented filter. See also the statistics in tables 1 and 2.

Figure 1: Full-sample vs. real-time estimates of output-gap.



Panels in the first row plot the full-sample and the real-time estimates of output-gaps obtained from the HP-filter estimated in accordance to each specification (in columns). Panels in the second row plot the revisions in the real-time estimates with respect to the full-sample estimates. The gray horizontal lines refer to two standard deviations around the sample mean value of the (real-time) series, as computed in tables 1 and 2. The shaded areas refer to recessions as dated by the Business Cycle Dating Committee (BCDC) of the National Bureau of Economic Research (NBER).

specifications of the HP-filter.

Descriptive statistics and reliability indicators for the output-gaps estimated by each specification of the HP-filter and their revisions are summarized into tables 1 and 2, respectively. As our full-sample estimates may well incorporate end-of-sample biases into its end-points estimates, we exclude the last eight of these estimates in the computation of all our statistics for comparative purposes. We discuss this issue further later.

From the descriptive statistics, the standard deviations of the estimated output-gaps point out an interesting result from the usage of the forecast-augmented specifications. Specifically, notice that such a measure of the short-run volatility implied by the output-gap estimates is reduced substantially under the latter specifications in relation to that obtained under the standard HP-filter both in real-time and with the full-sample. Obviously, taking the full-sample estimates as the final measures of output-gap, this result is not favorable for the forecast-augmented specifications in terms of the magnitude of the gap measures that the filter provides in real-time. However, we can understand this as a scaling effect result of a reinforced smoothing, byproduct

Table 1: Summary statistics and reliability indicators for estimated output-gaps.

Method	Mean	Min.	Max.	St.Dv.	Corr	OSig
Full-sample	0.03	-4.76	3.83	1.57	1.00	0.00
Real-time						
Standard	-0.17	-3.98	3.82	1.60	0.53	0.44
SPF's forecasts	0.03	-2.98	2.33	1.01	0.83	0.25
NC's forecasts	0.12	-3.83	2.92	0.96	0.42	0.32
IAR's forecasts	-0.05	-2.67	1.80	0.85	0.62	0.37
DAR's forecasts	-0.06	-2.88	1.84	0.87	0.63	0.36

All statistics are computed over the estimates for the period of 1968:q4 - 2008:q4, and are presented in percentage points except for Corr, which denotes the correlation with the full-sample estimates of the output-gap, and OSig, which denotes the frequency of opposite signs of real-time and full-sample estimates of output-gap.

of the forecast-augmentation of the HP-filter from which the measures of output-gap are not taken from the (augmented) sample end-point anymore.

The reliability indicators presented into these tables are invariant to such scaling effects and can, therefore, be used for comparative purposes. Overall, these indicators corroborate the success of the procedure of augmenting the HP-filter with survey forecasts in attenuating end-of-sample uncertainty of output-gap estimates. The first of these indicators is the coefficient of correlation between the real-time and the full-sample estimates of the output-gap, Corr in table 1, which is increased by 30 percentage points (pp) with the use of SPF's forecasts while using model's forecasts this indicator is increased by not more than 10 pp. Also, notice that using the forecasts from the NC model this correlation coefficient even deteriorates in relation to the standard HP-filter. This last result, and similar inferences based on most of the other statistics, comes to support our previous statement that the improvement to the reliability of real-time estimates of output-gap from augmenting the HP-filter with forecasts depend on the quality of the forecasts. As showed by Stark (2010), the forecasts from the NC model are easily outperformed by SPF's forecasts.

A second key statistic is the frequency at which real-time and full-sample estimates of output-gap point to opposite directions in relation to its corresponding estimated trend, denoted by OSig in table 1. Again, the use of SPF's forecasts presented the bigger improvement to this statistic with a decrease of almost 20 pp in relation to the 44% frequency of mistakenly diagnosing the current side where the economy is situated in relation to potential output (if measured by the trend component) observed by the use of the standard HP-filter. Roughly speaking, our results indicate that in, say, four years we would have estimated a wrong output-gap sign in about seven quarters (on average) using the standard HP-filter, while we could have these measurement errors reduced to only four quarters if we used the SPF's forecast-augmented HP-filter in real-time.

Other key statistics are based on the series of revisions to real-time estimates, such as the ratio of its root mean squares to the standard deviation of the full-sample estimates of the output-gap, NSR in table 2, which captures the average size of the real-time measurement errors (or the revisions) relative to the average size of the full-sample output-gap estimates. As the results in table 2 reveal, while the revisions to real-time estimates obtained with the standard HP-filter are of the same order of magnitude as the output-gap full-sample estimates, with the SPF's forecast-augmented HP-filter the average relative size of these revisions is reduced to about three fifths (3/5) the average magnitude of the gap full-sample estimates. Furthermore, notice that the first-order serial correlation of these revisions, AR in table 2, which overall shows a high level of persistence into its dynamics, is reduced only slightly by the use of forecasts except for the NC forecasts. Given that this last had the higher size of revisions, we interpret the decreases to the degrees of persistence as a byproduct of the varying sizes of these revisions rather than a relevant change into its dynamical properties.

To further enhance our understanding on the dynamics of the revisions implied by each of these specifications we present in figure 2 a cross correlogram between the revisions taken by the standard HP-filter and those taken by the SPF and DAR's forecast-augmented HP-filters. A careful examination of these dynamics can shed some light into the question of which of the filtering specifications provides a more accurate leading indicator to the full-sample estimate of latest output-gaps. Clearly, the superior accuracy of the forecast-augmented specifications can be related to their ability to anticipate the revisions ultimately taken by the standard HP-filter. Also notice how the use of SPF's forecasts attains a higher degree of anticipation than using DAR's forecasts, which then carries these revisions further forward.

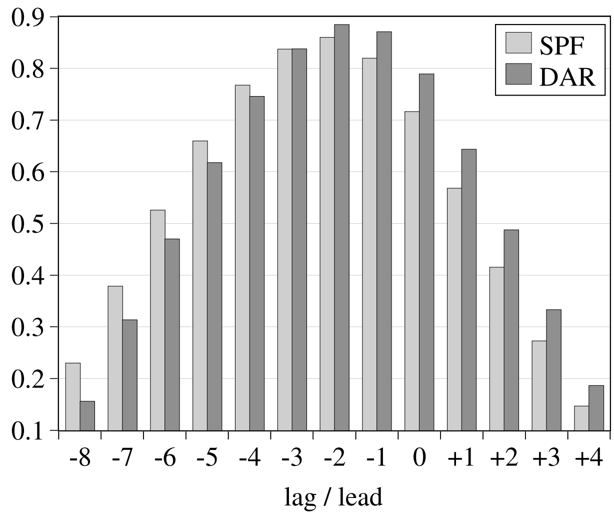
Another important question on the results presented so far is about the reliability of the real-

Table 2: Summary statistics and reliability indicators for revisions to real-time estimates.

Method	Mean	St.Dv.	RMS	AR	NSR
Standard	0.20	1.55	1.55	0.97	0.99
SPF's forecasts	-0.00	0.92	0.91	0.90	0.58
NC's forecasts	-0.09	1.46	1.46	0.58	0.93
IAR's forecasts	0.08	1.24	1.23	0.86	0.78
DAR's forecasts	0.09	1.23	1.23	0.86	0.78

All statistics are computed using the revisions to real-time estimates for the period of 1968:q4 - 2008:q4. RMS denotes the root mean square and AR denotes the first-order serial correlation of the corresponding revision series. NSR denotes the ratio of the RMS of the revisions to the standard deviation of the full-sample estimates of the output-gap. Means, standard deviations, and root mean squares are presented in percentage points.

Figure 2: Cross correlations between revisions by the standard and forecast-augmented HP-filters.



The plotted bars correspond to the coefficients of correlation between the revisions to real-time estimates incurred by the FHP-filter at t + lag/lead (indicated in the bottom axis) with those revisions incurred by the SHP-filter at t. Using previous section notation: $Correl\left(g_{t+i|T,0} - g^{f}_{t+i|t,K}, g_{t|T,0} - g_{t|t,0}\right)$, where i stands for the lag/lead.

time trend/gap estimates around business cycle turning points, as represented by the shaded areas in figure 1. It is well conceivable to presume an increase in the policy relevance of real-time estimates during these periods. Thus, in table 3 we computed some of the previous statistics focusing on the three quarters centered around each of the NBER registered peaks from 1969 to 2008. There, as expected, we can clearly see a deterioration on the reliability of real-time estimates of the output-gap for most of our specifications of the HP-filter. Still, the SPF's forecast-augmented HP-filter was again found to provide the most reliable real-time estimates and, perhaps surprisingly, presented revisions with relative magnitude (NSR=0.87) still lower than that obtained by the standard specification over the whole sample (NSR=0.99), and missed the sign of the full-sample estimate of the output-gap at almost the same frequency (OSig=0.24) that it did for the whole sample (OSig=0.25).

Table 3: Summary statistics and reliability indicators for revisions to real-time estimates during NBER peaks.

Method	Mean	St.Dv.	RMS	NSR	Corr.	OSig
Standard	2.21	0.44	2.25	1.93	0.93	0.67
SPF's forecasts	0.80	0.64	1.02	0.87	0.86	0.24
NC's forecasts	0.35	1.53	1.53	1.32	-0.13	0.38
IAR's forecasts	1.38	0.80	1.59	1.36	0.86	0.57
DAR's forecasts	1.41	0.79	1.61	1.38	0.87	0.57

The sample over which statistics are computed consists of 21 observations of the three quarters centered around each of the NBER peaks from 1968 to 2008, which are represented by the beginning of each shaded area in figure 1. Definitions of the statistics are explained in the notes to tables 1 and 2.

Lastly, it is important to mention that our statistics for the full-sample and the real-time estimates of output-gap using the standard HP-filter, and for their implied revisions, are overall consistent[11] with previous results in the literature (see tables 1-3, and 5 in Orphanides and van Norden, 2002, p. 574). Regarding the forecast-augmented real-time estimates the methodologically closest results previously reported to which we can make a comparison are those of Garratt et al. (2008, see table 1, p. 797) and Ince and Papell (2010, see tables 1 and 2, pp. 16-7), which used an eighth-order univariate autoregression to obtain up to two and twelve-periods ahead forecasts, respectively. Despite of discrepancies between some of the descriptive statistics of the output-gap estimates, which can be due to the use of different samples, we can see that our results on the use of survey expectations overcomes theirs with slightly higher Corr (0.83 against 0.78 and 0.76) and OSig (0.25 against 0.20)[12] statistics.

[11] All the comparisons we mention in this paragraph are in relation to the quasi real-time statistics in the cited papers.

[12] OSig was not presented in Ince and Papell (2010), and revision statistics as in table 2 were not presented in

C. Uncertainty decay time

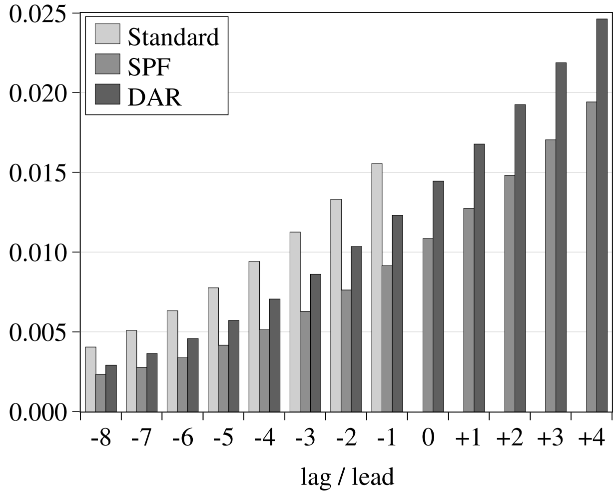
Our evaluation of the reliability of real-time estimates of output trend/gap has so far focused on the latest available of these estimates. In other terms, with the sequences $\left\{y_t, y_{t+1}^f, ..., y_{t+K}^f\right\}_{t=1}^N$ of observations and real-time forecasts, for $N \leq T$, we have recursively evaluated the reliability of the output-gap estimates at N in relation to full-sample estimates obtained with $\{y_t\}_{t=1}^T$. It has been showed that there is, indeed, a high degree of uncertainty associated with these estimates, though the use of survey forecasts is able to attenuate this problem considerably. A natural question ensuing from this finding is on how long it takes for this uncertainty to decay to reasonable levels. Besides, under the forecast-augmented approach we can also obtain forecasted measures of the output-gap directly from the filter, at least as far as goes the horizon of forecasts used into this specification (K). These measures are of special relevance for out-of-sample evaluation of forward-looking models, such as those for exchange rate determination (Engel et al., 2008; Galimberti and Moura, 2010). Thus, it is important to evaluate their reliability as well.

We answer to these questions by looking at lags/leads of the real-time estimated measures of the output-gap, ie., when estimating the real-time series of output-gaps at period t, $\left\{g_j|t,0\right\}_{j=1}^{j=t-1}$ for the standard HP-filter, and $\left\{g_j|t,K\right\}_{j=1}^{j=t+K-1}$ for the forecast-augmented HP-filter, we evaluate reliability indicators over the measures at $t+i$, where $i=-8,...,-1$ for the standard HP-filter, and $i=-8,...,+4$ for the forecast-augmented HP-filter. We focus on the RMS and OSig statistics, and represent them graphically in figures 3 and 4. Looking at the first of these we can see that, for both specifications of the filter, in the eighth quarter after the real-time estimate was firstly obtained the average magnitude of revisions has decreased to about 25% of that for the first estimates. Moreover, in despite of the improvements that the forecast-augmented specifications bring to the current real-time estimates of the output-gap, we can see that the forward measures that can be obtained from these specifications suffer with a (slightly exponential) deterioration of accuracy, approximately doubling their RMS at the fourth lead. Also notice that the magnitude of the revisions taken by the standard HP-filter achieves the latest level obtained in real-time by the SPF's forecast-augmented only with four quarters of

neither of these papers.

delay.

Figure 3: Root mean squares of revisions by lag/lead of the measured output-gap in relation to the real-time sample end-point.

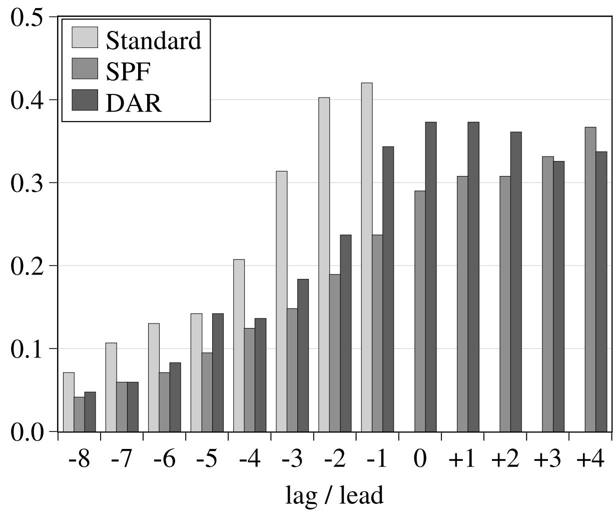


Root mean squares (RMS) are computed for different lags/leads of real-time measures of output-gap estimates. Using previous section notation: $RMS\left(g_{t+i|T,0} - g_{t+i|t,0}\right)$ for the SHP-filter and $RMS\left(g_{t+i|T,0} - g_{t+i|t,K}\right)$ for the FHP-filter, where i stands for the lag/lead indicated in the bottom axis. Notice that, from the assumption that data is observed with a 1-period delay, the latest available real-time estimate of output-gap obtained from the SHP-filter refers to that for $t-1$.

Similar conclusions can be drawn from the frequencies of opposite signs of real-time and full-sample estimates of the output-gap, plotted in figure 4, though the decays for this indicator are not as smooth as for the previous one. Despite of that, perhaps the most interesting aspect distinguishing these results refers to the forward measures obtained from the forecast-augmented specifications. Contrary to the previous results, the OSig statistic does not deteriorate so drastically, specially using DAR's forecasts which, at the fourth lead, ended up around the same level as of that attained for current real-time estimates. Also notice that, incidentally, the lag at which the SPF's forecast-augmented specification OSig crosses the 10% level, five quarters of delay, coincides with the average delay that the NBER's BCDC took to announce trough dates from 1980 to 2010, 15 months.

Other than providing us with enhanced insights about the timing of end-of-sample uncertainty in real-time estimates of the output-gap, the results presented in this section provide a

Figure 4: Frequencies of opposite signs of real-time and full-sample output-gap estimates by lag/lead of the measure.



The frequencies with which the signs of real-time and full-sample estimates of output-gap oppose (see the OSig definition in table 2) are computed for different lags/leads of the real-time estimated measures in a similar fashion to what we did in figure 3. See the notes to that figure for details.

justification for our concerns with the end-of-sample full-sample estimates used as benchmark for evaluation of the real-time estimates in the previous section. As Orphanides and van Norden (2002, p. 571) mentioned with respect to the measurement of output, we can say that the full-sample estimate "is very much an ephemeral concept in the measurement" of output-gap given that the same end-of-sample uncertainty affecting real-time estimates will be present at the end-points of the former. To attenuate possible biases from this source, we computed our evaluation statistics excluding the last two years of the estimates obtained with thefull-sample.

IV. Concluding remarks

In this paper we presented new results on the reliability of real-time trend-gap decomposition by the use of the famous Hodrick-Prescott filter. We showed, through an empirical application with US GDP data, that expanding the time horizon of the real-time pre-filtered series with survey forecasts considerably attenuates the end-of-sample uncertainty associated to these estimates. The improvement obtained with this approach was found to reduce the odds of real-time signal errors by almost a half, and the magnitude of revisions to these estimates accounts to only three fifths (3/5) of the output-gap average size, a figure usually amounting to a one-by-one ratio. Our analysis also took matters on the timing of this uncertainty, showing that a 90% rate of correct assessments of the output-gap sign can be achieved with five quarters of delay using the survey's forecast-augmented specification.

In what stands, these results provide favorable support for the forecast-augmentation of the HP-filter to attenuate the end-of-sample uncertainty inherent to the real-time estimates of trend-gap. In despite of turning the filtering procedure into a joint forecasting problem, the use of survey forecasts incorporating market participants judgements stands as a "cheaper" and at least as reliable as the rather harder, if not impossible, alternative of constructing optimal forecasts for empirical purposes. On the grounds of macroeconomic theory, our results also can be interpreted as providing favorable support to the view that expectations, here measured by survey data on forecasts, have an eminent role into the actual determination of an economy's dynamics.

References

Altissimo, F.a, C. R. F. M. L. M. V. G. (2010). New eurocoin: Tracking economic growth in real time. *Review of Economics and Statistics 92*(4), 1024–1034. cited By (since 1996) 2.

Azevedo, J. V. e. (2011). A multivariate band-pass filter for economic time series. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 60*(1), 1–30.

Blackburn, K., F. Orduna, and M. Sola (1995). Exponential smoothing and spurious correlation: a note. *Applied Economics Letters 2*(3), 76–79.

Bruchez, P.-A. (2003, August). A modification of the hp filter aiming at reducing the end-point bias. Working paper, Swiss Federal Finance Administration, Bundesgasse 3, CH-3003 Bern.

Cayen, J.-P. and S. van Norden (2005, December). The reliability of canadian output-gap estimates. *The North American Journal of Economics and Finance 16*(3), 373–393.

Christiano, L. J. and T. J. Fitzgerald (2003, 05). The band pass filter. *International Economic Review 44*(2), 435–465.

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.

Cooley, T. F. and E. C. Prescott (1995). Economic growth and business cycles. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, Chapter 1, pp. 1–38. Princeton University Press.

Croushore, D. (2011, March). 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.

Ehlgen, J. (1998). Distortionary effects of the optimal hodrick-prescott filter. *Economics Letters 61*(3), 345 – 349.

Engel, C., N. C. Mark, and K. D. West (2008, October). Exchange rate models are not as bad as you think. In *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pp. 381–441. National Bureau of Economic Research, Inc.

Eusepi, S. and B. Preston (2008, July). Expectations, learning and business cycle fluctuations. NBER Working Papers 14181, National Bureau of Economic Research, Inc.

Evans, G. W. and S. Honkapohja (2001). *Learning and expectations in macroeconomics*. Frontiers of Economic Research. Princeton, NJ: Princeton University Press.

Galimberti, J. K. and M. L. Moura (2010, October). Taylor rules and exchange rate predictability in emerging economies. Ibmec Working Papers 211, Ibmec Working Paper, Ibmec São Paulo.

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.

Harvey, A. C. and A. Jaeger (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics 8*(3), pp. 231–247.

Hodrick, R. J. and E. C. Prescott (1997, February). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking 29*(1), 1–16.

Ince, O. and D. H. Papell (2010, September). The (un)reliability of real-time output gap estimates with revised data. Mimeo.

Jönsson, K. (2010). Trend extraction with a judgement-augmented hodrick-prescott filter. *Empirical Economics 39*, 703–711. 10.1007/s00181-009-0325-3.

Kaiser, R. and A. Maravall (1999). Estimation of the business cycle: A modified hodrick-prescott filter. *Spanish Economic Review 1*, 175–206. 10.1007/s101080050008.

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.

Maravall, A. and A. del Río (2007). Temporal aggregation, systematic sampling, and the hodrick-prescott filter. *Computational Statistics & Data Analysis 52*(2), 975 – 998.

Marcellino, M. and A. Musso (2011, July). 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.

Milani, F. (2011). Expectation shocks and learning as drivers of the business cycle. *The Economic Journal 121*(552), 379–401.

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, 07). 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.

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, May). 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., S. J. Koopman, and A. Rua (2006, July). Tracking the business cycle of the euro area: A multivariate model-based bandpass filter. *Journal of Business & 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.