Monitoring the direction of the short-term trend of economic indicators

Monitoring the direction of the short-term trend of economic indicators

Estela Bee Dagum
Department of Statistical Sciences, University of Bologna
and
Silvia Bianconcini
Department of Statistical Sciences, University of Bologna

Abstract

Socioeconomic indicators have long been used by official statistical agencies to analyse and assess the current stage at which the economy stands via the application of linear filters used in conjunction with seasonal adjustment procedures. In this study, we propose a new set of symmetric and asymmetric weights that offer substantial gains in real-time by providing timely and more accurate information for detecting short-term trends with respect to filters commonly applied by statistical agencies. We compare the new filters to the classical ones through application to indicators of the US economy, which remains the linchpin of the global economic system. To assess the superiority of the proposed filters we develop and evaluate explicit tests of the null hypothesis of no difference in revision accuracy of two competing filters. Furthermore, asymptotic and exact finite-sample tests are proposed and illustrated to assess if two compared filters have equal probabilities of failing to detect turning points at different time horizons after their occurrence.

Keywords: data revisions, turning points, Diebold-Mariano test, McNemar test.

1 Introduction

The purpose of short-term economic statistics is to provide a comprehensive and timely picture of economic processes, such as production, income distribution, financing and expenditure, to evaluate the stage of the cycle at which the economy stands. Economic analysts and policymakers not only require accurate and timely information on the direction and magnitude of the trend of main economic variables, but they must also be confident that these estimates are unlikely to change significantly as more complete data becomes available. This is particularly important nowadays when the fallout from the COVID-19 pandemic and the ongoing war in Ukraine has led to a significant economic slump in 2022. Prices for some common commodities are reaching record levels, economic growth is slowing and inflation is rising. In July 2022, the Gross Domestic Product (GDP) growth in the United States fell for a second quarter, which has stirred up debate over whether the country is or will soon be in a recession. Hence, the problem of identifying the direction of the short-term trend of major socioeconomic indicators has become of great interest. Reliable trend-cycle estimates in real-time are of key importance to aid and assist governments and Central Banks in informed decision making.

Research efforts by official statisticians have recently been devoted to improving existing procedures to help reveal better the movements in the short-term trend and the occurrence of turning points. This differs from business cycle studies where cyclical fluctuations are measured around a long-term trend to estimate complete cycles (see among others Azevedo (2011); Azevedo et al. (2006); de Carvalho et al. (2012); de Carvalho and Rua (2017)).

National bureaus generally rely on the symmetric Henderson (1916) filter used together with asymmetric filters developed by Musgrave in 1964. However, the use of the latter introduces large revisions as new observations are added to the series. From a policy-making viewpoint, they are too slow in detecting true turning points (see Dagum and Bianconcini (2015, 2016), and reference therein). To overcome these main limitations, Dagum and Luati (2009) developed a linear filter based on the convolution of several noise suppression, trend estimation, and extrapolation filters. Main statistical agencies around the world nowadays use this linear cascade filter and the Henderson/Musgrave weights to provide information on trend-cycle movements for several economic indicators

(see https://www.statcan.gc.ca/eng/dai/btd/trend-cycle, and US Census Bureau (2017)).

The moving average used to estimate the trend-cycle component is selected based on the amplitude of irregular variations in the data relative to the amplitude of long-term systematic variations, known as noise-to-signal ratio (Shiskin et al., 1967). Following the so-called X11 variable trend-cycle routine (Findley et al., 1998), threshold values of this ratio are used to guide the selection of the most appropriate trend-cycle filter for monthly and quarterly series. For monthly data, 9-, 13- or 23-term filters are selected, whereas for quarterly data, 5- and 7-term filters are used. Threshold values for high-frequency data, such as weekly or daily observations, are not available but are the subject of current research by official statistical agencies (Ladiray et al., 2018).

Differently from the Henderson and Musgrave filters, only the 13-term cascade symmetric filter and corresponding asymmetric weights have been developed. Due to the application of a specific *ad hoc* mixed normalisation performed to ensure that the thirteen weights added up to unity, cascade filters of shorter length cannot be derived. The corresponding asymmetric weights, fundamental to obtaining real-time trend-cycle estimates, are not available for longer cascade filters.

This paper aims to provide a set of symmetric and asymmetric weights to be used in conjunction with seasonal adjustment procedures based on a kernel representation of the cascade filter that is twofold. First, we derive kernel weights that closely reproduce the 13-term symmetric cascade ones but, differently from the latter, symmetric kernel filters of any length can be determined. Secondly, asymmetric weights can be obtained by selecting time-varying bandwidths, specific for each asymmetric filter (Dagum and Bianconcini, 2015). We show that, for the 13-term symmetric filter, these asymmetric weights have similar properties to the cascade ones in terms of gain and phaseshift functions. However, differently from the latter, asymmetric filters of any length can be derived.

We illustrate the proposed asymmetric and symmetric weights using socioeconomic indicators of the US economy, which remains the linchpin of the global economic system. Specifically, we show that when the noise-to-signal ratio selects the 13-term filter, the proposed kernels and the cascade filter provide similar results in revisions and turning point detection. Furthermore, we illustrate that these real-time trend-cycle kernels always have

to be preferred to the Musgrave filters. We propose two widely applicable tests to assess the significance of our key empirical findings. One for the null hypothesis of no difference in revision accuracy of two competing filters, and the second to evaluate differences in the timely detection of turning points. For the former, our approach is in the spirit of the Diebold and Mariano (1995) test developed to compare predictive accuracy. For the latter, we consider different time horizons after the occurrence of the turning point and develop a test for the null hypothesis of equal probabilities of failing its detection for two compared filters. We show that it is similar to the McNemar (1947) test used for testing marginal homogeneity of paired binomial proportions.

2 Official statistical methods for short-term trend estimation

Linear filters developed by Henderson (1916) are the classical method to estimate the trend-cycle component of seasonally adjusted economic indicators used together with non-parametric seasonal adjustment estimates from the US Bureau of the Census X11 method (Shiskin et al., 1967) and its variants, X11/X12ARIMA (Findley et al., 1998) and X13ARIMA (US Census Bureau, 2017).

Assuming that the input series $\{y_t, t = 1, \ldots, N\}$ is seasonally adjusted, it can be decomposed into the sum of a systematic component g_t, usually referred to as the trend-cycle for they are estimated jointly, plus an erratic component u_t, called the noise, such that

$$y_t = g_t + u_t, \quad t = 1, \ldots, N.$$

The noise u_t is assumed to be either a white noise, $WN(0, \sigma_u^2)$, or, more generally, to follow a stationary and invertible autoregressive moving average (ARMA) process.
The Henderson trend-cycle estimates $\hat{g}_t$ for the central observations, $t = m+1, \ldots, N-m$, are obtained through a weighted moving average as follows

$$\hat{g}_t = \sum_{j=-m}^{m} w_j y_{t+j}, \tag{1}$$

where the weights $w_j, j = -m, \ldots, m$, are derived by the Henderson ideal formula:

$$w_j = \frac{315[(m+1)^2 - j^2][(m+2)^2 - j^2][(m+3)^2 - j^2][3(m+2)^2 - 16 - 11j^2]}{8(m+2)[(m+2)^2 - 1][4(m+2)^2 - 1][4(m+2)^2 - 9][4(m+2)^2 - 25]}. \quad (2)$$

Using the reproducing kernel Hilbert space methodology, Dagum and Bianconcini (2008) have shown that these weights can be equivalently derived by

$$w_j = \frac{K_3^B(j/(m+1))}{\sum_{j=-m}^{m} K_3^B(j/(m+1))}, \qquad j = -m, \ldots, m, \quad (3)$$

where K_3^B is a third order kernel derived from the biweight density function $f_{0B}(t) = (15/16)(1-t^2)^2, t \in [-1, 1]$, and corresponding Jacobi orthonormal polynomials $P_i, i = 0, \ldots, 3$, that is $K_3^B(t) = \sum_{i=0}^{3} P_i(t)P_i(0)f_{0B}(t), t \in [-1, 1]$.

At the end (beginning) of the series, asymmetric weights need to be applied. The asymmetric Henderson smoothers currently in use were developed by Musgrave (1964). They are based on the minimisation of the mean squared revision between final and preliminary estimates subject to the constraint that the sum of the weights is equal to one (Doherty, 2001). The assumption made is that at the end (beginning) of the series, the seasonally adjusted values follow a linear trend-cycle plus a purely random irregular component.

Dagum and Bianconcini (2008, 2013) introduced a kernel representation of the Musgrave filters. In this framework, given the biweight density function, once the length of the symmetric filter is chosen, the statistical properties of the asymmetric filters are strongly affected by the bandwidth parameter of the kernel function from which the weights are derived. The authors made the bandwidth parameters equal for all the asymmetric filters (global time-invariant bandwidth) to closely approximate the Musgrave filters, but in another paper (Dagum and Bianconcini, 2015) they have proposed time-varying bandwidth parameters since the asymmetric filters are time-varying.

Under common economic conditions, the 13-term filter and corresponding asymmetric weights are often applied for monthly data. These filters have the excellent property of fast detection of true turning points but the limitation of producing large revisions to the most recent estimates when new observations are added to the series. To overcome these main limitations, Dagum (1996) proposed a nonlinear semiparametric predictor that consisted of

(1) extending the seasonally adjusted series, modified by extreme values with zero weights, with ARIMA extrapolations, and (2) applying the 13-term Henderson filter to the extended series using stricter sigma limits for the identification and replacement of extreme values ($\pm 0.7\sigma$ and $\pm 1\sigma$ were recommended). Given its excellent properties, Dagum and Luati (2009) provided a linear approximation for both symmetric and asymmetric components using the convolution of several noise suppression, trend estimation, and extrapolation linear filters. The general matrix representation of the symmetric cascade linear filter is given by

$$\mathbf{H}[\mathbf{H} + \mathbf{M}_{7,(0.143)}(\mathbf{I}_N - \mathbf{H})][\mathbf{H} + \mathbf{M}_{5,(0.25)}(\mathbf{I}_N - \mathbf{H})], \tag{4}$$

where $\mathbf{H}$ refers to the Henderson filter, $\mathbf{M}_{5,(0.25)}$ is the matrix representation of a 5-term moving average with weights (0.250,0.250,0.000,0.250,0.250), and $\mathbf{M}_{7,(0.143)}$ is the matrix representation of a 7-term filter with all weights equal to 0.143. Each element of the convolution matrix (4) depends on single, pairs and triplets of Henderson weights, and, based on the Henderson ideal formula (2), the central weight w_0^{CLF} can be expressed as a direct function of the filter length, as detailed in the Supplementary Material.

When the 13-term Henderson filter is considered, the convolution (4) produces a symmetric filter of thirty one terms with very small weights at both ends. Dagum and Luati (2009) truncated this filter to thirteen terms and performed an *ad hoc* mixed normalisation to ensure that the weights added up to unity. The total weight discrepancy (equal to -0.065) was distributed over the thirteen weights, $w_j^{CLF}, j = -6, \ldots, 6$, by mostly allocating it to the central value w_0^{CLF} (+36.4%). Conversely, the values of w_3^{CLF} and w_{-3}^{CLF} were reduced (-34.1%) to preserve as much as possible the same area under the positive weights as the Henderson filter, without modifying the negative ones, for a proper estimation of points of maxima and minima. Asymmetric filters were applied to the last six data points, which are crucial for current analysis. They were obtained by the convolution of the symmetric filter with linear extrapolation filters for the last six data points from an ARIMA(0, 1, 1) model with $\theta = 0.40$.

2.1 Relationship between the cascade and Henderson filter

The main limitation of the cascade filter is that only the 13-term symmetric weights are available since the performed *ad hoc* normalisation, applied to ensure that the weights added to the unity, cannot be extended to different filter lengths. To provide a general formulation of the filter, the behaviour of the convolution (4) should be analysed when the normalisation is not needed, that is, for a filter of length greater than or equal to 21-term, and in comparison with the Henderson filter of the same length. Particularly for the latter, the following results hold.

Proposition 2.1. *For a Henderson filter of length* $2m+1$*, the weights* w_j *are negative (or null) if* $|j| \geq \lceil \sqrt{\frac{3(m+2)^2-16}{11}} \rceil$*, with*

$$\lceil \frac{m}{2} \rceil + 1 \leq \lceil \sqrt{\frac{3(m+2)^2-16}{11}} \rceil \leq \lceil \frac{m}{2} \rceil + 2,$$

where $\lceil x \rceil$ *denotes the function mapping* x *to the least integer greater than or equal to* x*.*

Corollary 2.1. *Given the Henderson filter of length* $2m+1$*, if* $2 \leq m \leq 7$*, the weights are negative when* $|j| \geq \lceil \frac{m}{2} \rceil + 1$*, if* $7 \leq m \leq 24$*,* w_j *is negative for* $|j| \geq \lfloor \frac{m}{2} \rfloor + 2$*, whereas, for* $m \geq 24$*,* w_j *is negative when* $|j| \geq \lceil \frac{m}{2} \rceil + 2$*, being* $\lfloor x \rfloor$ *the floor function mapping* x *to the greatest integer less than or equal to* x*.*

As widely discussed by Dagum and Luati (2009), the *ad hoc* normalisation applied to derive the 13-term symmetric weights were developed to preserve the same area under positive and negative weights as the Henderson filter for a proper estimation of points of maxima and minima.

Proposition 2.2. *The area under the positive Henderson weights is approximately equal to* $1.1 + \mathcal{O}\left(\frac{1}{m}\right)$*, and, consequently, the one under the negative weights is almost* $-0.1 + \mathcal{O}\left(\frac{1}{m}\right)$*.*

Differently from the Henderson filter, the unnormalised cascade weights have the following property.

Proposition 2.3. *For the cascade linear filter of length* $2m+1$*, the weights* w_j^{CLF} *are negative (or null) if* $j = \pm\left(\lceil \frac{m}{2} \rceil + 2\right), \ldots, \pm m$*, and positive if* $j = -\left(\lceil \frac{m}{2} \rceil + 1\right), \ldots, \left(\lceil \frac{m}{2} \rceil + 1\right)$*.*

Formal proofs of Proposition 2.1, 2.2, and 2.3 are provided in the Supplementary Material. It is evident that the Henderson weights get negative before, that is, for smaller (absolute) values of j, respect to the cascade filter (*Regularity 1.*). The discrepancies between the two filters reduce as the filter length increases as derived by Corollary 2.1, according to which both filters have negative weights when $|j| \geq \lceil \frac{m}{2} \rceil + 2$ as $m \geq 24$ (*Regularity 2.*). Finally, independently on the filter length (see Proposition 2.2), both the cascade and Henderson filters cover the same area under positive and negative weights, approximately equal to 1.1 and -0.1, respectively (*Regularity 3.*).

3 A reproducing kernel filter for real-time analysis

Based on the proven regularities that relate the convolution (4) to the Henderson weights, a generalisation of the 13-term cascade filter can be obtained by looking for a kernel representation within the same family from which the Henderson filter is derived (Dagum and Bianconcini, 2008). The latter belongs to the Beta family based on the density

$$f_{0Beta}(t) = \frac{r}{2B(s+1,1/r)}(1- \mid t \mid^r)^s, \quad t \in [-1,1] \tag{5}$$

where $B(s+1,1/r)$ is the Beta function. The corresponding third-order kernel is given by

$$K_3(t) = R_3(t) f_{0Beta}(t) = \left(\frac{\mu_4 - \mu_2 t^2}{\mu_4 - \mu_2^2} \right) f_{0Beta}(t), \quad t \in [-1,1] \tag{6}$$

where μ_2 and μ_4 are the second and fourth order moments of the Beta density, and $R_3(t)$ is the reproducing kernel based on orthonormal Jacobi polynomials. It defines the sign of the kernel function that results negative for any $\mid t \mid \geq \sqrt{\frac{\mu_4}{\mu_2}}$.

Different values of r and s in eq. (6) define different kernels. The Henderson kernel is obtained when $r = s = 2$, for which the ratio between the fourth and second moment is equal to $\frac{\mu_4}{\mu_2} = \frac{1}{3}$. The uniform ($r = 1, s = 0$), triangle ($r = 1, s = 1$), and Epanechinov ($r = 2, s = 1$) kernels are the other ones for which this ratio is greater than $\frac{1}{3}$. However, among them, the triangle kernel is the only one covering the same area, as the biweight

kernel, under negative and positive values (see Proposition 2.2). It is given by

$$K_3^T(t) = \left(\frac{12}{7} - \frac{30}{7}t^2\right)(1- \mid t \mid), \quad t \in [-1, 1],$$

and it can be considered as the kernel representation of the linear cascade filter.

3.1 Bandwidth selection

When applied to real data, the symmetric filter weights are derived from the triangle kernel K_3^T as follows:

$$w_j^T = \frac{K_3^T(j/b)}{\sum_{j=-m}^{m} K_3^T(j/b)}, \quad j = -m, \ldots, m, \tag{7}$$

where b is a time-invariant global bandwidth parameter (same for all $t = m+1, \ldots, N-m$) selected to ensure a symmetric filter of length $2m+1$. Its selection has to be done to reproduce the behaviour of the cascade filter correctly. For this purpose, the bandwidth b is chosen such that the central kernel weight w_0^T is equal to the central cascade weight w_0^{CLF}. The latter is the greatest weight of the symmetric filter, such that if we guarantee to reproduce it exactly due to the shape of the triangle kernel, all the remaining cascade weights are closely represented. This is true when the cascade weights add to one, that is, for filters of length greater than or equal to 21-term, since the kernel weights in eq. (7) sum up to one. Based on eq. (7),

$$w_0^T = \frac{12b^3}{12b^3(2m+1) - 12b^2m(m+1) - 10bm(2m^2+3m+1) + 15m^2(m+1)^2},$$

whereas w_0^{CLF} is a direct function of the filter length. This implies that an analytical relationship between the bandwidth parameter and the filter length can be derived. By imposing w_0^T to be equal to w_0^{CLT}, the bandwidth parameter is obtained by solving the following equation

$$12[1-w_0^{CLT}(2m+1)]b^3+12w_0^{CLT}m(m+1)b^2+10w_0^{CLT}m(m+1)(2m+1)b-15w_0^{CLT}m^2(m+1)^2 = 0. \tag{8}$$

The three roots of eq. (8) can be real or complex. Among them, the greatest real one is selected since, given the properties of the bandwidth parameter, it is the only one that is (a) real, (b) positive, and (c) greater than m.

Table 1 shows the bandwidths selected for cascade filters of length greater than or equal to 21-term. The application of these bandwidth parameters allows us to derive triangle kernel weights $w_j^T, j = -m, \ldots, m$, that closely reproduce the corresponding cascade ones. This is illustrated in Table 2 for the 23-term filters, which are almost identical to the third digit. Similar results for the other filter lengths are obtained, not reported for space reasons.

- TABLE 1 NEAR HERE -

Looking at the parameters shown in Table 1, it is evident that the relationship between the filter length, that is, m, and the corresponding bandwidth is strictly linear. Specifically, for all m, the bandwidth parameter is equal to

$$b = 2.327 + 0.913m. \tag{9}$$

Based on this result, kernel filters of any length can be derived. For example, the 13-term symmetric kernel weights are obtained by fixing $b = 7.808$. They are reported in Table 2 together with the classical cascade filter. It can be noted that the two weight systems are similar, with the main discrepancy being in the central weights. These differences are because the bandwidth is selected to reproduce w_0^{CLF}, but for the 13-term, and in general, for shorter (than 21-term) filters, it refers to not normalised symmetric weights.

- TABLE 2 NEAR HERE -

On the other hand, if we select the bandwidth b to exactly reproduce the normalised central weight obtained by Dagum and Luati (2009) - equal to 7.409 - we obtain kernel weights almost identical to the classical cascade filter weights. These weights are reported in Table 2. This implies that the triangle kernel K_3^T is the kernel representation of the cascade filter. In the following, all the symmetric kernel weights, including the 13-term filter, are derived by selecting the bandwidth (9).

3.2 Asymmetric weights

The derivation of the symmetric triangle filter has assumed the availability of $2m+1$ input values centred at t. However, at the end of the sample period, that is $t = N-(m+1), \ldots, N$, only $2m, \ldots, m+1$ observations are available, and asymmetric filters of the same length have to be considered. At the boundaries, Dagum and Bianconcini (2015) have suggested following the so-called "cut and normalise" method, according to which the kernels $K_3^{Tq^*}$ are obtained as follows

$$K_3^{Tq^*}(t) = \frac{K_3^T(t)}{\int_{-1}^{q^*} K_3^T(t)dt}, \qquad t \in [-1, q^*], q^* << 1. \tag{10}$$

Applied to real data, the asymmetric weights result

$$w_{q,j} = \frac{K_3^{Tq^*}(j/b_q)}{\sum_{j=-m}^{q} K_3^{Tq^*}(j/b_q)}, \tag{11}$$

for $j = -m, \ldots, q$, and $q = 0, \ldots, m-1$, where b_q is the local bandwidth, specific for each asymmetric filter. For each q, it relates the discrete domain of the filter, that is $\{-m, \ldots, q\}$, to the continuous domain of the kernel function, that is $[-1, q^*]$. Dagum and Bianconcini (2015) derive a class of optimal asymmetric filters based on bandwidth parameters selected as follows

$$b_{q,G} = \min_{b_q} \sqrt{2 \int_0^{1/2} |G_q(\omega) - G(\omega)|^2 d\omega}, \tag{12}$$

where $G(\omega)$ is the gain function of the symmetric filter, whereas $G_q(\omega)$ is the one corresponding to the asymmetric weights $w_{q,j}, j = -m, \ldots, q$. This study considers these time-varying bandwidth parameters since they determine optimal filters that minimise revisions and time lag to detect a true turning point (Dagum and Bianconcini, 2015).

- FIGURE 1 NEAR HERE -

Figure 1 illustrates the gain functions of the asymmetric cascade filters developed by Dagum and Luati (2009) (*left*) with the triangle kernels based on $b_{q,G}$ and relative to the 13-term symmetric filter obtained using the bandwidth (9) (*right*). The filters behave similarly,

with a fast convergence to the corresponding symmetric filter. The similar behaviour of these filters is true not only in revisions but also in terms of time lag in detecting a true turning point, as highlighted by Figure 2. The latter shows the phaseshifts (in months) of the last point cascade and triangle filters, and it is clear that both take, on average, two months to signal the upcoming peak or a trough.

- FIGURE 2 NEAR HERE -

4 Statistical tests for short-term trend filtering comparison

To evaluate the goodness of the proposed short-term trend filters, we first address the problem of the accuracy of their estimates in real-time, that is, in every period t. Real-time accuracy has been a great concern of socioeconomic time series analysts (see Orphanides and Norden (2002), and references therein). In a real-time application, the value of the short-term trend at time t relies on the information available up to that time. This value is, however, updated over time, and revisions occur for different reasons. The data itself undergoes revisions or/and new data becomes available as time passes, thus leading to a larger information set which improves the accuracy of the series. We neglect the former source of revisions since the dataset does not involve different data vintages corresponding to each time point in the sample and concentrate on the revisions due to filter changes.

4.1 Accuracy in real time

We compare the real-time performance of the triangle filters (TK) with the 13-term symmetric cascade filter and corresponding asymmetric weights (C), when applicable, and with the Henderson and Musgrave filters (H). For each method, we compute the real-time trend-cycle estimates $A^0_{it}, t = m+1, \ldots, N; i = TK, H, C$, that are conditioned on the sample up to period t, obtained by applying the last point filter, and the final estimates $S_{it}, t = m+1, \ldots, N-m$, based on the full sample. The comparisons are based on the relative filter revisions between the final and the last point estimates at each point in time,

that is,

$$e_{it} = \frac{S_{it} - A^0_{it}}{S_{it}}, \quad t = m+1, \ldots, N-m; \quad i = TK, H, C. \tag{13}$$

For each series, we assess the loss associated with each estimator by considering the squared revision errors e^2_{it} since we are interested in the revisions' size and not their sign. Their mean $\bar{e^2_i}$ provides a point estimate of the revision accuracy, $E(e^2_{it})$, for the i-th estimator. Following the idea behind the Diebold and Mariano (1995) test developed to evaluate differences in the accuracy of two competing forecasts, we account for the sampling uncertainty of $\bar{e^2_i}$ and $\bar{e^2_j}$ to test the null hypothesis of equal revision accuracy of filters i and l, that is

$$H_0 : E(e^2_{it}) = E(e^2_{lt}) \quad \text{or} \quad H_0 : E(d^{il}_t) = 0,$$

with $i \neq l, i, l = TK, H, C$, and where $d^{il}_t = e^2_{it} - e^2_{lt}$ is defined as the loss differential.

Under the assumption that the process $\{d^{il}_t\}_{t\in T}$ is weakly stationary, the large sample $N(0,1)$ statistic for testing the null hypothesis $H_0 : E(d^{il}_t) = 0$ versus the alternative $H_1 : E(d^{il}_t) \neq 0$ is given by

$$DM^{il} = \frac{\bar{d}^{il}}{\sqrt{\frac{2\pi \hat{f}_{d^{il}}(0)}{N-2m}}}$$

where $\bar{d}^{il} = \frac{1}{N-2m}\sum_{t=m+1}^{N-m} d^{il}_t$ is the sample mean loss differential, being $d^{il}_t = e^2_{it} - e^2_{lt}$. $\hat{f}_{d^{il}}(0)$ is a consistent estimate of the spectral density of $\{d^{il}_t\}_{t\in T}$ at frequency zero

$$f_{d^{il}}(0) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_{d^{il}}(k),$$

where $\gamma_{d^{il}}(k)$ is the autocovariance function at lag k, defined as $\gamma_{d^{il}}(k) = E[(d^{il}_t - \mu)(d^{il}_{t-k} - \mu)]$, being μ the expected value of the process $\{d^{il}_t\}_{t\in T}$. Following standard practice, a consistent estimate of $f_{d^{ij}}(0)$ is obtained by taking a weighted sum of the available sample autocovariances as

$$\hat{f}_{d^{il}}(0) = \frac{1}{2\pi}\left(\hat{\gamma}_{d^{il}}(0) + 2\sum_{k=1}^{S(N)}\left(1 - \frac{k}{S(N)}\right)\hat{\gamma}_{d^{il}}(k)\right),$$

where $S(N)$ is the truncation lag. Diebold and Mariano (1995) fixed it to $h-1$, under

the assumption that optimal h-step-ahead forecast errors are at most $(h-1)$-dependent. Analogously, we choose $S(N)$ to include only those estimated autocovariances that are significantly different from zero.

4.2 Timely detection of turning points

Most literature has recently discussed nowcasting based on real-time data (Giannone et al., 2008; Banbura et al., 2013). Indeed, timeliness in identifying true turning points plays an essential role in the decision-making process, as the information on the current state of the economy gives a more detailed picture of the future economic situation. The faster the upcoming turning point is detected, the faster new policies can be applied to counteract the impact of the business cycle stage.

In analysing the short-term trend of socioeconomic indicators, given the smoothness of the trend-cycle data, Zellner et al. (1991) have defined that a turning point occurs at time t if (*downturn*):

$$y_{t-3} \leq y_{t-2} \leq y_{t-1} > y_t \geq y_{t+1}$$

or (*upturn*)

$$y_{t-3} \geq y_{t-2} \geq y_{t-1} < y_t \leq y_{t+1}.$$

For a short-term trend estimator, the reduction of revisions in real-time trend-cycle estimates has not to be achieved at the expense of increasing the time lag to detect the upcoming true turning point, being the latter affected by the convergence path of its asymmetric filters to the symmetric one. To determine the time lag needed by a filter to detect a true turning point, we calculate the number of months it takes for the real-time trend-cycle estimate to signal a turning point in the same position as in the final trend-cycle series. From this latter, the sequence of true turning points $\{R_{it}, t = m+1, \ldots, N-m\}$ is identified by applying the Zellner et al. (1991) definition given above, such that $R_{it} = 1$ if $S_{i(t-3)} \leq S_{i(t-2)} \leq S_{i(t-1)} > S_{it} \geq S_{i(t+1)}$ *(downturn)* or $S_{i(t-3)} \geq S_{i(t-2)} \geq S_{i(t-1)} < S_{it} \leq S_{i(t+1)}$ *(upturn)*, whereas $R_{it} = 0$ otherwise. Since symmetric filters do not introduce phase shift effects in the estimates, the same set of true turning points is generally identified by considering different filters, that is $R_{it} = R_t$ for all $t = m+1, \ldots, N-m$, and $i = TK, H, C$.

The comparison between different estimators is performed in terms of average delay (in months) for each method in identifying all the peaks and troughs in the series. A formal statistical test for comparing the time lag in turning point detection is needed to assess the different behaviour of the competing filters. At this regard, preliminary trend-cycle estimates $\{A_{it}^q, t = m+1, \ldots, T-m\}, i = TK, H, C$, are obtained by applying the q-th asymmetric filters to the sample conditioned up to period $(t+q)$, for $q = 1, \ldots, m-1$. Binary sequences $\{R_{it}^q, t = m+1, \ldots, T-m\}$ are then derived by each preliminary series to describe at which point in time the corresponding asymmetric filter detects turning points, that is $R_{it}^q = 1$ if $A_{i(t-3)}^q \leq A_{i(t-2)}^q \leq A_{i(t-1)}^q > A_{it}^q \geq A_{i(t+1)}^q$ or $A_{i(t-3)}^q \geq A_{i(t-2)}^q \geq A_{i(t-1)}^q < A_{it}^q \leq A_{i(t+1)}^q$, and $R_{it}^q = 0$ otherwise.

To evaluate the error committed by each asymmetric filter in the identification of a turning point, we focus only on those occasions t in which a peak or trough occurred, that is $T^* = \{t_1, t_2, ..., t_{T_p}\}$. In this context, q can be interpreted as the time horizon (in months or quarters) after the occurrence of the turning point. Hence, we define the turning point detection error as

$$e_{it_j}^q = R_{it_j}^q - R_{t_j}, \qquad j = 1, \ldots, T_p.$$

It will be equal to -1 if the turning point is not signalled by the q-th asymmetric filter, and zero otherwise. By assessing the loss in accuracy in terms of squared detection errors $e_{it_j}^{q^2}$, the expected value $E\left(e_{it_j}^{q^2}\right)$ is the probability p_i^q that, q months after the turning point occurred, the filter i still does not identify it.

Since the time points $t_j, j = 1, \ldots, T_p$, are distant over time (more than 11 months or three quarters) and not necessarily equally spaced, the squared errors $\{e_{it_j}^{q^2}, j = 1, \ldots, T_p\}$ are assumed to be independent. In comparing two different filters, i and l, we want to test if q observations after a turning point occurred, the two methods have the same probability to fail in detecting it. That is,

$$H_0 : p_i^q = p_l^q \quad \text{or} \quad H_0 : E\left(d_t^{il^q}\right) = 0. \tag{14}$$

$\{d_t^{il^q}\}_{t \in T^*}$ is the loss differential process, whose realisations are $d_{t_j}^{il^q} = e_{it_j}^{q^2} - e_{lt_j}^{q^2}$, $j = 1, \ldots . T_p$. Given the independence of the error series, $\{d_t^{il^q}\}_{t \in T^*}$ is an independent

and identically distributed process, where each variable $d_t^{il^q}$ assumes value -1 if the filter l fails in the identification of a turning point but filter i does not, value 0 if both the filters behave similarly, or value 1 if the filter i does not identify the turning point but filter l does. Given the joint distribution of the bivariate error process $\{(e_{it}^{q^2}, e_{lt}^{q^2})\}_{t\in T^*}$, the expected value and the variance of the loss differential process are equal to

$$E(d_t^{il^q}) = p_{10}^q - p_{01}^q$$

and

$$V(d_t^{il^q}) = E(d_t^{il^{q^2}}) - [E(d_t^{il^q})]^2 = (p_{10} + p_{01}) - (p_{10} - p_{01})^2,$$

where $p_{kk'}^q$ corresponds to the joint probability that $e_{it_j}^{q^2} = k$ and $e_{lt_j}^{q^2} = k'$, being $k, k' = 0, 1$. The null hypothesis (14) of equal probability in the identification of a true turning point at the time horizon q can be reformulated as $H_0 : E(d_t^{il^q}) = p_{10}^q - p_{01}^q = 0$ or $H_0 : p_{10}^q = p_{01}^q$, and tested against the alternative hypothesis $H_1 : E(d_t^{il^q}) \neq 0$ or, equivalently, $H_1 : p_{10}^q \neq p_{01}^q$. An unbiased estimator of the expected value of the loss differential process $\{d_t^{il^q}\}_{t\in T^*}$ is given by

$$\bar{d}^{il^q} = \frac{1}{T_p}\sum_{j=1}^{T_p} d_{t_j}^{il^q} = \frac{\sum_{j=1}^{T_p}(d_{t_j}^{il^q} = 1) - \sum_{j=1}^{T_p}(d_{t_j}^{il^q} = -1)}{T_p} = \frac{T_{10}^q - T_{01}^q}{T_p},$$

where $T_{kk'}^q$ corresponds to the number of observed pairs $\{(e_{it_j}^{q^2}, e_{lt_j}^{q^2}), j = 1, \ldots, T_p\}$, with outcome $e_{it_j}^{q^2} = k$ and $e_{lt_j}^{q^2} = k'$, being $k, k' = 0, 1$. It follows that

$$E(\bar{d}^{il^q}) = \frac{1}{T_p}\sum_{j=1}^{T_p} E(d_{t_j}^{il^q}) = p_{10} - p_{01}$$

and

$$V(\bar{d}^{il^q}) = \frac{1}{T_p^2}\sum_{j=1}^{T_p} V(d_{t_j}^{il^q}) = \frac{1}{T_p}[(p_{10} + p_{01}) - (p_{10} - p_{01})^2].$$

Proposition 4.1. *To test the null hypothesis $H_0 : E(d_t^{il^q}) = 0$ versus the alternative $H_1 : E(d_t^{il^q}) \neq 0$, the test statistic, under H_0, is given by*

$$Z_d^{il^q} = \frac{\bar{d}^{il^q}}{\sqrt{V(\bar{d}^{il^q})}} = \frac{T_{10}^q - T_{01}^q}{\sqrt{T_{10}^q + T_{01}^q}}. \tag{15}$$

It is equivalent to the McNemar (1947) test statistic, which is the statistic used for testing marginal homogeneity of paired binomial proportions (Fagerland et al., 2014).

The formal proof of Proposition 4.1 is provided in the Supplementary Material.
Under the null hypothesis, the McNemar test statistic, and also $Z_D^{il^q}$, is approximated by the standard normal distribution. The asymptotic McNemar test is often presented as the equivalent statistic

$$\chi^2 = Z_d^{il^{q^2}} = \frac{(T_{01}^q - T_{10}^q)^2}{T_{01}^q + T_{10}^q}$$

that, under the null hypothesis, is approximately chi-squared distributed with one degree of freedom. The common advice (Agresti, 1990) has been to apply the McNemar test in its classical (asymptotic) form unless the values in the two discordant cells were small, that is $T_{01}^q + T_{10}^q < 25$. In this latter situation, which generally occurs even over 60 years, the classical version of the test cannot be guaranteed to preserve the type I error rate, and the exact binomial test or the continuity correction to the classical test (Edwards, 1948) should be applied. However, the use of these latter has recently been challenged on the evidence of extensive simulations by Fagerland et al. (2013, 2014). The authors have pointed out that these two tests do control the type I error rate below the nominal value but tend to be overly conservative. They suggest the use of a mid-p (or quasi-exact) version of the binomial test (Hirji, 2005) since they have found that it is almost as powerful as the asymptotic McNemar test but without exceeding the nominal significance level. While it has not been demonstrated analytically in all circumstances, it never exceeded the nominal level in their extensive simulations.

5 Illustrations

In this section, the proposed methodology is used to estimate the short-term trend of two main indicators commonly used to study the current socioeconomic conditions of the US economy. The evaluation of groups of cyclical indicators to assess the stage at which the

economy stands has been performed by the National Bureau of Economic Research (NBER) program since the work of Burns and Mitchell (1946), and also used by the Bureau of Economic Analysis (BEA) of the US Department of Commerce, and the Conference Board. The aim is to monitor and analyse the position of the economy and the current phase of the business cycle. Some indicators, called leading indicators, can predict the economy's development and promptly identify upcoming turning points. Policymakers, the press, and the public analyse these series to gauge whether a recession is forthcoming. On the other hand, coincident indicators are those whose time course refers to the development of the gross domestic product. They confirm or refute the economy's current stage and thereby contribute to characterising its stability and sustainability. The importance of using these indicators to assess the current stage at which the economy stands is evident when we consider, as an example, the series of the average weekly hours in manufacturing in the United States from January 1960 to December 2020 (Figure 3).

- FIGURE 3 NEAR HERE -

There is no clear long-term trend in this series over the last half-century, but there is quite a dramatic tendency for the average weekly hours to decrease sharply during periods characterised as economic recessions, which are represented as shaded regions in Figure 3. These are those determined by the Business Cycle Dating Committee of the National Bureau of Economic Research. One of the first things we want to know about these series is how their fluctuations are related to movements in and out of economic recessions.

In addition to this series, we also consider as illustrative example the index of industrial production. The span of both series extends from January 1960 to December 2020 to cover several periods of recession and expansion. Following the X11 variable trend-cycle routine (Findley et al., 1998), the variability of these series is measured by the noise-to-signal ratio. For monthly data, the ratio is computed by obtaining a preliminary estimate of the trend-cycle $\hat{g}_t$ via the application of the 13-term Henderson filter to the seasonally adjusted data. The irregular component $\hat{u}_t$ results from removing this trend estimate from the seasonally adjusted series. With $\bar{C}$ denoting the sample mean of the available values of the absolute trend changes $\mid \hat{g}_t - \hat{g}_{t-1} \mid$ and $\bar{I}$ the sample mean of $\mid \hat{u}_t - \hat{u}_{t-1} \mid$, the value of the noise to signal ratio, $\bar{I}/\bar{C}$, is used to determine the length of the filters based on specific threshold

values (Shiskin et al., 1967). For monthly data, if the noise-to-signal ratio is smaller than 1, the 9-term filter is selected; if $1 \leq \bar{I}/\bar{C} < 3.50$, the 13-term filter is chosen, whereas when $\bar{I}/\bar{C} \geq 3.5$, the 23-term is used.

5.1 Average weekly hours in manufacturing

The average weekly hours in manufacturing series comes from the Current Employment Statistics program, also known as the payroll or establishment survey, published by the US Bureau of Labor Statistics. It is a monthly survey of approximately 140000 businesses and government agencies representing about 440000 worksites throughout the United States. From the sample, the survey produces and publishes employment, hours, and earnings estimates for the nation, states, and metropolitan areas at detailed industry levels. Average weekly hours relate to the average hours per worker for which pay was received. As a sensitive measure of labour demand, this series serves as an indicator in The Conference Board's Leading Economic Index for the United States. For this monthly series, observed over the span January 1960 - December 2020, the noise-to-signal ratio is equal to 2.456, such that 13-term symmetric triangle, Henderson and cascade filters have been selected.
First of all, we assess the revision accuracy of the three considered estimators under square error loss. In terms of point estimates, the last point triangle filter provides the most accurate estimates. Indeed, its mean square revision error $\bar{e}^2_{TK}$ is equal to 0.0000358 as opposed to the last point Musgrave filter whose mean square revision error $\bar{e}^2_H$ is 0.0001. The ratio between these two quantities is around 0.370, implying that, applied to these real data, the triangle filter produces, on average, a reduction of almost sixty percent of the revisions introduced in the real-time trend-cycle estimates concerning the Musgrave filter. Figure 4 (*left*) shows the corresponding loss differential series, in which no obvious nonstationarity is visually apparent. Approximate stationarity is also supported by its sample autocorrelation function shown in Figure 4 (*right*), which decays quickly, being only the first three autocorrelation coefficients significantly different from zero.

- FIGURE 4 NEAR HERE -

We test the null hypothesis of equal expected revision accuracy between the two filters, that is $H_0 : E(d_t^{TK,H}) = 0$ versus the alternative assumption $H_1 : E(d_t^{TK,H}) \neq 0$. We focus

on the test statistic $DM^{TK,H}$ setting the truncation lag at three in light of the preceding discussion. We obtain $DM^{TK,H}$= -5.302, implying a p-value of 0.000. Thus, we reject the hypothesis of equal expected square revision errors; that is, the last point triangle filter is significantly better than the Musgrave one in terms of revisions.

Similarly, in comparison with the cascade filter, the mean square revision error for the latter results in 0.000043, which implies that, for this specific series, the last point triangle kernel provides similar revisions. However, it still shows an average reduction of around ten percent. The loss differential series associated with the triangle and cascade filter is illustrated in Figure 5 (*left*) together with the corresponding autocorrelation function (*right*), both supporting the stationarity of the series. The test statistic $DM^{TK,C}$ results equal to -4.989, with a p-value of 0.000, such that the last point triangle filter results also significantly better in terms of revisions than the last point cascade one.

- FIGURE 5 NEAR HERE -

The comparison between the three filters is then performed in terms of the average time taken to detect the turning points of the average weekly hours in manufacturing series. For the latter, from all three final estimates, the same set of sixteen turning points has been identified. Over the entire set of peaks and troughs, the triangle filter takes about two months (2.3125 on average) to detect a turning point. In contrast, the cascade and Henderson filters require around three months, with an average delay of 2.9375 and 3.3125, respectively. To assess if the triangle kernel significantly outperforms respect to the Henderson and cascade filters, the squared turning point detection errors $e_{it_j}^{q^2}, j = 1, \ldots, 16$, are computed for each asymmetric filter and estimator. To assess the independence over time of each of these error sequences, we apply the exact test developed by Hirji (2005) (pp.438-439) used to assess the null hypothesis of independence of a binary series against the alternative hypothesis of a first-order Markovian dependence structure. Table 3 (a) reports the associated p-values. As expected, the null hypothesis is accepted for all the error sequences $\{e_{it_j}^{q^2}, j = 1, \ldots, 16\}, q = 1, \ldots, 5, i = TK, H, C$.

- TABLE 3 NEAR HERE -

At each time horizon q, the probability of failing in detecting a turning point is compared for the triangle filter and the Henderson one. Table 3 (b) shows that at each horizon q,

the number of discordant cells, that is $T_{01}^q + T_{10}^q$, is always smaller than 25. Hence, the one-sided exact binomial test is applied to assess the null hypothesis of equal detection accuracy, $H_0 : p_{TK}^q = p_H^q$, against the alternative hypothesis of the worst performance for the Henderson filter, that is $H_1 : p_H^q > p_{TK}^q$. Under the null hypothesis, the one-sided exact p-value is then given by

$$\sum_{k=T_{01}^q}^{T_{01}^q+T_{10}^q} \left(\begin{array}{c} T_{01}^q + T_{10}^q \\ k \end{array} \right) 0.5^{T_{01}^q+T_{10}^q}$$

Following the suggestions by Fagerland et al. (2013, 2014), we also computed the (quasi-exact) mid-p-value equal to

$$\sum_{k=T_{01}^q}^{T_{01}^q+T_{10}^q} \left(\begin{array}{c} T_{01}^q + T_{10}^q \\ k \end{array} \right) 0.5^{T_{01}^q+T_{10}^q} - 0.5 \left(\begin{array}{c} T_{01}^q + T_{10}^q \\ T_{01}^q \end{array} \right) 0.5^{T_{01}^q+T_{10}^q}.$$

It is evident that one and two months after the turning point occurred, the two filters have the same probability to fail in identifying it correctly. In contrast, after three (or more) months of the occurrence of the turning point, at 5% level, the Henderson filter has a significantly higher probability of failing to correctly signal the turning point with respect to the triangle kernel, which results to be superior in terms of turning point detection.

On the other hand, comparing the triangle and cascade filter, there are no significant differences in turning point detection independent of the time horizon considered. This is illustrated in Table 3 (c), which reports the one-sided exact binomial and mid- p-values to test the null hypothesis $H_0 : p_{TK}^q = p_C^q$ against the alternative one $H_1 : p_C^q > p_{TK}^q$, for $q = 1, \ldots, 5$.

5.2 Index of industrial production

The Index of Industrial Production measures the real output of all stages of production in the manufacturing, mining, gas and electric utility industries for all facilities in the United States. This coincident indicator is published by the US Federal Reserve and compiled every month to bring attention to short-term changes in industrial production. From January 1960 to December 2020, the seasonally adjusted series is characterised by a smoother pattern than the average weekly hours series, as highlighted by its noise-to-

signal ratio equal to 0.913. Hence, the comparison in terms of both revisions and turning point detection is performed between the 9-term triangle and Henderson filters, whereas the cascade filter is not applied, being available only with 13 terms.

The last point triangle filter provides a point estimate of the revision accuracy, as expressed by the mean square error $\bar{e}^2_{TK}$, equal to 0.0002, whereas, for the last point Musgrave filter, it is equal to 0.0006. Also, in this case, the triangle filter produces, on average, a reduction of more than fifty percent of the revisions introduced in the real-time trend-cycle estimates. To assess the significance of this empirical finding, we consider the corresponding loss differential series illustrated in Figure 6 (*left*), together with its sample autocorrelation function (*right*). In both these figures, stationarity is evident.

- FIGURE 6 NEAR HERE -

We then test the null hypothesis of equal expected revision accuracy between the two filters and focus on the test statistic $DM^{TK,H}$ setting the truncation lag at four since only four autocorrelation coefficients are significantly different from zero. We obtain $DM^{TK,H}$= -3.379, with p-value of 0.001. Thus, we can conclude that the last point triangle filter is significantly more accurate than the Musgrave one.

To compare the performance of the two filters in terms of detection of turning points, 19 peaks and troughs have been identified from both the final estimates. To detect all these points, the triangle filter takes, on average, 1.68 months, whereas the Musgrave filters more than two months (2.47 on average). To evaluate if the triangle filter is a significantly faster detector of turning points with respect to the Henderson one, the squared error series $e^{q^2}_{it_j}, j = 1, \ldots, 19$, are derived for each asymmetric filter, $q = 1, 2, 3$, and estimator $i = TK, H$. As for the previous series, we assess the independence over time of each of these sequences through the exact test by Hirji (2005), for which Table 4 (a) reports the associated p-values. As expected, the null hypothesis is accepted for all the error sequences $\{e^{q^2}_{it_j}, j = 1, \ldots, 19\}, q = 1, 2, 3, i = TK, H$.

- TABLE 4 NEAR HERE -

We then compare the probability of failing in detecting a turning point one, two or three months after its occurrence for the triangle and the Henderson filter. Based on the one-sided exact binomial and the corresponding quasi-exact p-values, it is evident that one and

two months after the turning point occurred, the two filters have the same probability of failing to identify it correctly. On the other hand, the Henderson filter has a significantly higher probability of failing its identification three months after the turning point occurred. Hence, the triangle filter is significantly more timely in detecting true turning points.

6 Discussion

This study deals with the problem of assessing, in real-time, the direction of the short-term trend with an application to some key indicators of the US economy due to its important role from an international macroeconomic perspective. Official statistical bureaus, central banks and organisations that produce real-time short-term trends apply the Musgrave or cascade linear filters. We have proposed a reproducing kernel representation of the cascade filter that allows deriving filters of any length and we have shown that they significantly perform better than the cascade and Musgrave filters in terms of revisions and time lag in detecting true turning points. Specific asymptotic and exact tests have been developed and illustrated to assess the significance of these main empirical findings.

In the paper, the properties of the proposed filters have been discussed theoretically and empirically, focusing on the revisions due to filter changes when new observations are added to the series. This has allowed us to directly compare our proposed filters with those available in the literature and account just for the filter effects. However, it would be important to evaluate if and how using different data vintages affects our main findings.

Following Croushore and Stark (2003) and Arouba (2008), different data vintages of a set of fifteen leading, coincident and lagging indicators of the US economy have been analysed. For each series, we have considered their first releases - as available in the ALFRED database (https://alfred.stlouisfed.org) - and their last available vintage at the time of the analysis - October 2022 - as final observation. For each series, final relative data revisions have been computed by accounting for the difference between the latest available observation and its first available release.

- TABLE 5 NEAR HERE -

The series considered are all seasonally adjusted. We have about fifty years of obser-

vations for the quarterly series, while we have between 20 and 60 years of data for the monthly indicators. The first column of Table 5 reports the mean of the final relative data revisions for each variable. All of them are different from zero, indicating that the first available releases are biased estimates of the final values. Greater values are observed for the index of industrial production, manufacturing and trade sales, and labour cost per unit of output in manufacturing, which have been subject to benchmark revisions due to changes in the base year in the two considered releases. For the remaining series, the means range from -0.366% to 6.249%. We also computed each variable's minimum and maximum final relative revision, reported in the second and third columns. The range of the data revisions is quite large for all variables, as also shown by their standard deviations - in the fourth column.

For each series at each vintage release, the performance of the proposed triangle kernel filters is compared with the Musgrave and - when applicable - cascade filters in terms of both accuracy in real-time and delay in detecting true turning points. In the sample, for the monthly series, the noise-to-signal ratio ranges from 0.572 to 3.307. Hence, filters of lengths 9 and 13 terms are applied. On the other hand, it is always smaller than 3.50 for the quarterly series, such that 5-term filters are selected.

To evaluate the accuracy in real-time, the comparisons are based on the relative filter revisions $e_{it}, i = TK, H, C; t = 1, \ldots, T$, as defined in eq. (13). In particular, we calculate the ratio between the mean square revision error corresponding to our last point filters, $\bar{e}^2_{TK}$, and those corresponding to the last point Musgrave, $\bar{e}^2_H$, and - when applicable - cascade filters, $\bar{e}^2_C$.

Independently on the data vintages, for all series, the results - illustrated in the sixth and seventh column in Table 5 - show that the ratio is always smaller than one, indicating that the kernel last point predictors introduce smaller revisions than the Musgrave and - when available - cascade filters. In particular, the ratio with the last point Musgrave filter is always less than one-half, and when the 13-term filter is selected, the ratio with the cascade filter is, on average, around 0.90. All these differences are statistically significant based on the test discussed in Section 4.1.

The last three columns of Table 5 show that, independently of the data release, the

newly proposed filters can capture turning points faster than or with the same delay as their competitors. For all the series and data releases considered, the proposed filter has a better or not statistically different performance than the others, as evaluated using the exact tests discussed in Section 4.2.

In summary, data vintages do not affect the better statistical properties of the proposed method both in terms of accuracy in revisions and timely detection of turning points.

SUPPLEMENTARY MATERIAL

Elementwise formula of the cascade filter: we show how each element of the convolution matrix (4) depends on single, pairs and triplets of Henderson weights.

Proofs: detailed proofs of Propositions 2.1, 2.2, 2.3, and 4.1.

References

Agresti, A. (1990). *Categorical Data Analysis*. John Wiley and Sons, New York.

Arouba, B. S. (2008). Data revisions are not well-behaved. *Journal of Money, Credit and Banking. 40*, 319–340.

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

Azevedo, J., S. 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. 81*, 575–593.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data flow. In *Working Paper Series*. European Central Bank.

Burns, A. and W. Mitchell (1946). Measuring business cycles. In *Technical report*. NBER. New York.

Croushore, D. and T. Stark (2003). A real-time data set for macroeconomists: does the data vintage matter? *Review of Economics and Statistics. 85*(3), 605–617.

Dagum, E. B. (1996). A new method to reduce unwanted ripples and revisions in trend-cycle estimates from x11arima. *Survey Methodology. 22*, 77–83.

Dagum, E. B. and S. Bianconcini (2008). The henderson smoother in reproducing kernel hilbert space. *Journal of Business and Economic Statistics. 26*(4), 536–545.

Dagum, E. B. and S. Bianconcini (2013). A unified probabilistic view of nonparametric predictors via reproducing kernel hilbert spaces. *Econometric Reviews. 32*(7), 848–867.

Dagum, E. B. and S. Bianconcini (2015). A new set of asymmetric filters for tracking the short-term trend in real time. *The Annals of Applied Statistics. 9*(3), 1433–1458.

Dagum, E. B. and S. Bianconcini (2016). *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation*. Springer International Publishing.

Dagum, E. B. and A. Luati (2009). A cascade linear filter to reduce revisions and turning points for real time trend-cycle estimation. *Econometric Reviews. 28*(1-3), 40–59.

de Carvalho, M., P. Rodrigues, and A. Rua (2012). Tracking the us business cycle with a singular spectrum analysis. *Economic Letters. 114*, 32–35.

de Carvalho, M. and A. Rua (2017). Real-time nowcasting the us output gap: singular spectrum analysis at work. *International Journal of Forecasting* (33), 185–198.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics. 13*(3), 253–263.

Doherty, M. (2001). The surrogate henderson filters in x-11. *Australian and New Zealand Journal of Statistics. 43*(4), 385–392.

Edwards, A. L. (1948). Note on the correction for continuity in testing the significance of the difference between correlated proportions. *Psychometrika. 13*(3), 185–187.

Fagerland, M. W., S. Lydersen, and P. Laake (2013). The mcnemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology. 13*(91).

Fagerland, M. W., S. Lydersen, and P. Laake (2014). Recommended tests and confidence intervals for paired binomial proportions. *Statistics in Medicine. 33*, 2850–2875.

Findley, D. F., B. C. Monsell, W. R. Bell, M. C. Otto, and B.-C. Chen (1998). New capabilities and methods of the x12arima seasonal adjustment program. *Journal of Business and Economic Statistics. 16*, 127–152.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: the real-time informational content of macroeconomic data. *Journal of Monetary Economics. 55*, 665–676.

Henderson, R. (1916). Note on graduation by adjusted average. *Transaction of Actuarial Society of America. 17*, 43–48.

Hirji, K. F. (2005). *Introduction to the exact analysis of discrete data*. Chapman and Hall.

Ladiray, D., J. Palate, G. Mazzi, and T. Proietti (2018). Seasonal adjustment of daily and weekly data. In *Handbook of seasonal adjustment*, pp. 759–781. European Union.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika. 12*(2), 153–157.

Musgrave, J. C. (1964). A set of end weights to end all end weights. In *Working paper*. U.S. Bureau of Census.

Orphanides, A. and S. V. Norden (2002). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics. 84*(4), 569–583.

Shiskin, J., A. Young, and J. Musgrave (1967). The x-11 variant of the census method ii seasonal adjustment program. In *Technical paper*. US Department of Commerce, Bureau of the Census.

US Census Bureau (2017). X-13arima-seats reference manual. Technical report, US Census Bureau.

Zellner, A., C. Hong, and C. Min (1991). Forecasting turning points in international output growth rates using bayesian exponentially weighted autoregression, time-varying parameter and pooling techniques. *Journal of Econometrics. 48*, 275–304.

Table 1: Bandwidth parameters.

Filter length	21	23	25	27	29	31	33	35	37	39
m	10	11	12	13	14	15	16	17	18	19
b	11.46	12.37	13.29	14.20	15.11	16.03	16.94	17.85	18.77	19.68

Table 2: Weight Systems for 13- and 23-term cascade filters and kernels

Length	Filter												
13	Cascade	-0.027	-0.007	0.031	0.067	0.136	0.188	**0.224**					
	K_3^T, $b = 7.409$	-0.027	-0.010	0.028	0.079	0.134	0.185	**0.224**					
	K_3^T, $b = 7.808$	-0.023	-0.002	0.035	0.081	0.130	0.175	**0.209**					
23	Cascade	-0.015	-0.016	-0.012	-0.005	0.007	0.028	0.046	0.071	0.095	0.110	0.125	**0.133**
	Triangle Kernel	-0.015	-0.016	-0.010	-0.000	0.014	0.031	0.049	0.068	0.087	0.105	0.121	**0.133**

Table 3: Average weekly hours in manufacturing: (a) p-values associated with the exact test for assessing the null hypothesis of independence of the sequence of the errors $\{e_{it_j}^{q^2}, j = 1, \ldots, 16\}, q = 1, \ldots, 5, i = TK, H, C$, versus the alternative of first order Markov dependence (Hijri, 2005); (b) one-sided exact binomial and (quasi-exact) mid- p-values to test $H_0 : p_{TK}^q = p_H^q$ versus $H_1 : p_H^q > p_{TK}^q$; and (c) to test $H_0 : p_{TK}^q = p_C^q$ versus $H_1 : p_C^q > p_{TK}^q$.

(a) independence test for $\{e_{it_j}^{q^2}, j = 1, \ldots, 16\}$

$i \quad \backslash \quad q$	1	2	3	4	5
TK	0.516	0.545	1.000	1.000	1.000
H	0.267	0.294	0.725	0.725	0.593
C	0.516	0.595	0.363	0.629	0.629

(b) $H_0 : p_{TK}^q = p_H^q$ versus $H_1 : p_H^q > p_{TK}^q$

q	1	2	3	4	5
$T_{01}^q + T_{10}^q$	6	5	5	5	5
Exact binomial p-value	0.8906	0.8125	0.0312	0.0312	0.0312
(Quasi-exact) mid-p-value	0.7734	0.656	0.0156	0.0156	0.0156

(c) $H_0 : p_{TK}^q = p_C^q$ versus $H_1 : p_C^q > p_{TK}^q$

q	1	2	3	4	5
$T_{01}^q + T_{10}^q$	0	1	3	3	3
Exact binomial p-value	0	0.50	0.1250	0.1250	0.1250
(Quasi-exact) mid-p-value	0	0.25	0.0625	0.0625	0.0625

Table 4: Index of industrial production: (a) p-values associated with the exact test for assessing the null hypothesis of independence of the sequence of the errors $\{e_{it_j}^{q^2}, j = 1, \ldots, 19\}, q = 1, 2, 3, i = TK, H$, versus the alternative of first-order Markov dependence (Hijri, 2005); (b) one-sided exact binomial and (quasi-exact) mid- p-values to test $H_0 : p_{TK}^q = p_H^q$ versus $H_1 : p_H^q > p_{TK}^q$.

(a) independence test for $\{e_{it_j}^{q^2}, j = 1, \ldots, 19\}$

$i \quad \backslash \quad q$	1	2	3
TK	0.499	1.000	1.000
H	0.501	0.439	0.583

(b) $H_0 : p_{TK}^q = p_H^q$ versus $H_1 : p_H^q > p_{TK}^q$

q	1	2	3
$T_{01}^q + T_{10}^q$	4	5	8
Exact binomial p-value	0.6875	0.1875	0.004
(Quasi-exact) mid-p-value	0.5000	0.109	0.002

Table 5: Summary statistics of the final data revisions and comparison between short-term filters in terms of accuracy and turning point detection.

	Final relative data revisions (%)				I/C	Filter revisions		Turning points			
	mean	min	max	std.dev.	(filter length)	$\frac{\bar{e}^2_{TK}}{\bar{e}^2_H}$	$\frac{\bar{e}^2_{TK}}{\bar{e}^2_C}$	number	Avg TK	Avg H	Avg C
Average weekly hours in manufacturing											
Period: 1939/01-1961/10											
First release: 1961/11					1.743 (13)	0.4121*	0.9066*	8	2.375	2.375	2.125
Vintage: 2022/10	0.217	-1.028	2.089	0.361	1.696 (13)	0.4103*	0.9054*	8	3.000	2.625	3.250
New orders for consumer goods											
Period: 1992/02-2011/04											
First release: 2011/06					1.629 (13)	0.3999*	0.8930*	5	2.800	4.200**	3.000
Vintage: 2022/10	0.091	-0.832	1.997	0.402	1.582 (13)	0.4056*	0.8950*	6	3.167	3.667	3.500
New orders for nondefense capital goods											
Period: 1992/02-2010/01											
First release: 2010/03					3.307 (13)	0.4110*	0.9045*	3	3.333	4.333	4.333
Vintage: 2022/10	0.476	-3.206	5.996	1.608	3.176 (13)	0.4106*	0.9037*	2	2.500	3.000	3.000
New private housing permits											
Period: 1960/01-1999/07											
First release: 1999/08					2.000 (13)	0.3937*	0.8935*	15	2.933	2.733	3.067
Vintage: 2022/10	0.018	-2.587	3.717	0.347	1.989 (13)	0.4106*	0.9037*	15	2.933	2.733	3.067
Capacity utilisation											
Period: 1967/01-1996/10											
First release: 1996/11					1.050 (13)	0.4022*	0.9015*	11	2.909	5.091**	3.636
Vintage: 2022/10	-0.084	-2.736	2.537	0.964	1.025 (13)	0.4053*	0.9043*	8	2.750	3.750	2.750
Employees on non-agricultural payroll											
Period: 1939/01-1990/04											
First release: 1990/05					0.572 (9)	0.4200*	NA	11	1.730	2.270	NA
Vintage: 2022/10	0.1753	-0.788	0.844	0.194	0.578 (9)	0.4216*	NA	12	1.920	2.000	NA
Industrial production index											
Period: 1921/01-1989/12											
First release: 1990/01					0.833 (9)	0.3560*	NA	26	2.077	2.500**	NA
Vintage: 2022/10	-52.550	-56.850	-51.080	1.045	0.842 (9)	0.3559*	NA	25	2.000	2.360**	NA
Manufacturing and trade sales											
Period: 1967/01-2013/05											
First release: 2013/06					1.713 (13)	0.4511*	0.9156*	9	2.111	3.000	2.566
Vintage: 2022/10	22.25	21.53	25.73	1.004	1.728 (13)	0.4510*	0.9155*	9	2.222	3.333	2.660
Retail sales											
Period: 1992/01-2010/05											
First release: 2010/06					1.734 (13)	0.4632*	0.9195*	1	4.000	6.000	6.000
Vintage: 2022/10	-0.366	-2.982	0.391	0.533	1.673 (13)	0.4536*	0.9156*	1	4.000	6.000	6.000
Gross National Product											
Period: 1947/Q1-1992/Q2											
First release: 1990/Q2					0.355 (5)	0.4173*	NA	2	1.000	1.000	NA
Vintage: 2022/Q2	6.249	4.367	9.906	1.110	0.349 (5)	0.4191*	NA	2	1.000	1.000	NA
Average duration of unemployment											
Period: 1948/01-1971/12											
First release: 1972/02					1.719 (13)	0.4122*	0.9057*	11	2.364	3.000	2.730
Vintage: 2022/10	-0.189	-8.51	2.542	1.074	1.769 (13)	0.4124*	0.9060*	11	2.727	3.909**	2.909
Inventory to sales ratio											
Period: 1992/01-2010/01											
First release: 2010/02					1.865 (13)	0.4038*	0.9017*	3	3.333	5.667	4.000
Vintage: 2022/10	-0.034	-1.504	2.362	0.510	2.077 (13)	0.4052*	0.9000*	7	2.715	4.143	3.000
Labour cost per unit of output - manufacturing											
Period: 1960/Q1-2013/Q1											
First release: 2013/Q2					0.655 (5)	0.3758*	NA	3	1.000	1.000	NA
Vintage: 2022/Q2	-10.092	-14.953	-7.049	1.538	0.803 (5)	0.3738*	NA	3	1.000	1.000	NA
Commercial and industrial loans											
Period: 1947/01-2010/11											
First release: 2010/12					0.490 (9)	0.421*	NA	14	1.929	2.143	NA
Vintage: 2022/10	-0.159	-2.971	0.390	0.489	0.494 (9)	0.421*	NA	14	1.786	2.071	NA
Unemployment rate											
Period: 1948/01-1990/02											
First release: 1990/03					1.364 (13)	0.433*	0.9192*	19	2.895	3.211**	2.474
Vintage: 2022/10	-0.004	-1.887	1.887	0.267	1.367 (13)	0.433*	0.9192*	19	2.895	3.211**	2.684

*: significant at 5% level based on the statistic DM^{il} (see Section 4.1).
**: significant at 5% level based on the quasi-exact binomial test (see Section 4.2).

Figure 1:

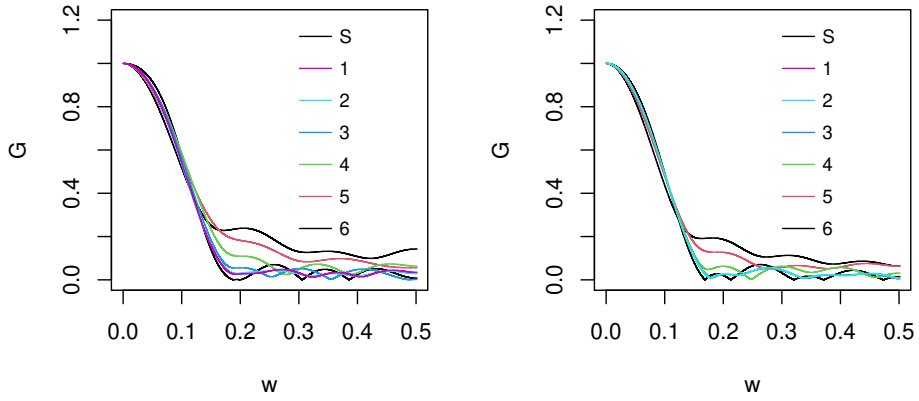

Figure 2:

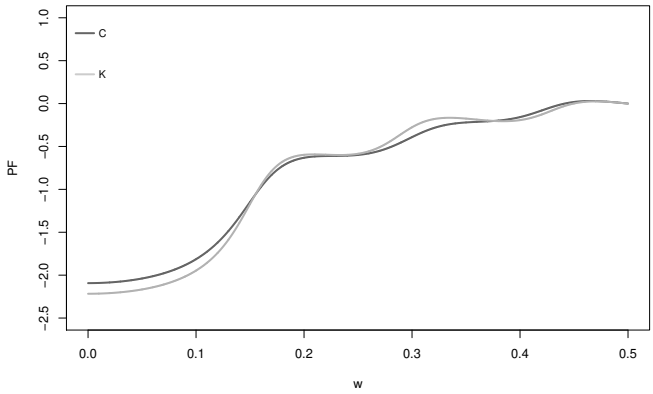

Figure 3:

Figure 4:

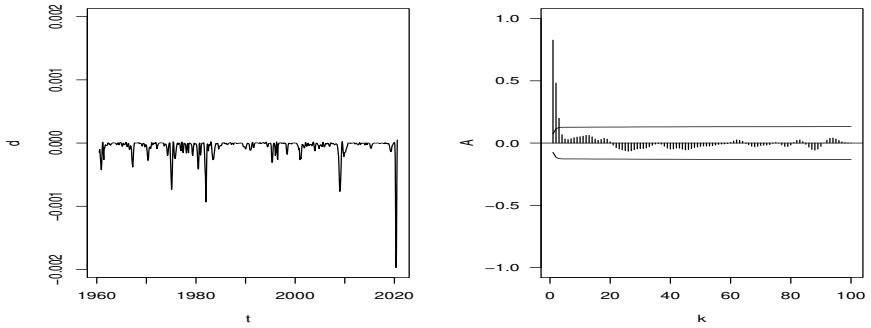

Figure 5:

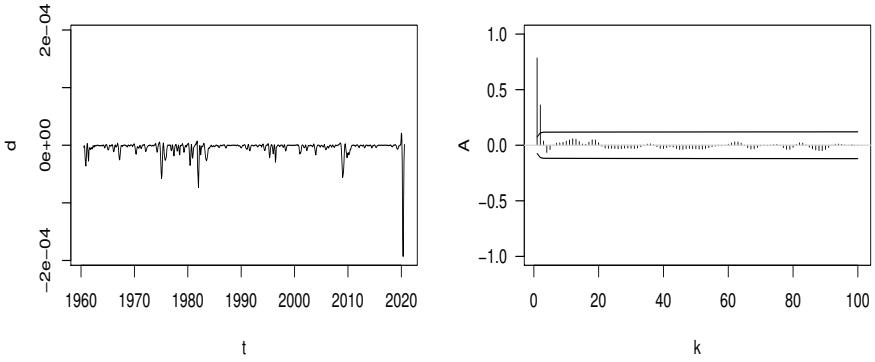

Figure 6:

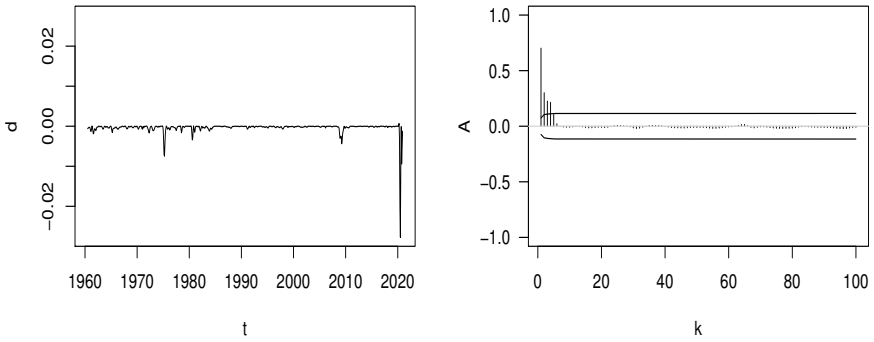

Figure 1: Gain functions of the 13-term cascade symmetric and asymmetric filters (*left*) and 13-term symmetric and asymmetric triangle kernels (*right*).

Figure 2: Phaseshift functions of the last point cascade and kernel filters corresponding to 13-term symmetric weights (*right*).

Figure 3: Average weekly hours in manufacturing in the United States. Shaded areas represent contractions of business cycle according to NBER chronology.

Figure 4:*Average weekly hours in manufacturing*: loss differential (triangle vs Henderson filter) series (*left*) and autocorrelations (*right*). The estimated autocorrelations are graphed, together with Bartlett's approximate 95% confidence interval (series: average week.

Figure 5: *Average weekly hours in manufacturing*: loss differential (triangle vs cascade filter) series (*left*) and autocorrelations (*right*). The estimated autocorrelations are graphed, together with Bartlett's approximate 95% confidence interval.

Figure 6: *Index of industrial production*: loss differential (triangle vs Henderson filter) series (*left*) and autocorrelations (*right*). The estimated autocorrelations are graphed, together with Bartlett's approximate 95% confidence interval.