BIP-Predictor

Marc Wildi
Zurich University of Applied Sciences (ZHAW)
8000 Zurich, Switzerland
marc.wildi@zhaw.ch

June 5, 2025

Abstract

1 Introduction

2 Data and Dependence

2.1 Data

- Choice of indicators: BIP, ip, ESI, ifo, spread
- Quarterly frequency (emphasize $h \geq 2$)
- Data transformations: Diff-log, standardization (merely for ease of visual inspection), trim Covid ourliers
- Revisions and publication lags: BIP lagged by quarter and ip lagged by two months in Jan-2025 data set
- Real-time data: all series are artificially aligned at end point. BIP and ip are advanced by their lags.

	GDP	ip	ifo_c	ESI	spr_10y_3m
2024-07-01	902.571	90.700	85.800	92.500	-1.300
2024-08-01		93.200	84.900	90.900	-1.400
2024-09-01		91.200	83.700	89.800	-1.400
2024-10-01	900.761	90.800	84.100	90.700	-1.200
2024-11-01		92.200	83.800	89.300	-0.900
2024-12-01			82.800	86.900	-0.800
2025-01-01			82.500	88.100	-0.200

Table 1: Ragged end.

The transformed quarterly series are displayed in Fig. 1, while Fig. 2 offers a close-up view of the financial crisis and the pandemic. The latter figure illustrates that the real-time GDP and industrial production series are synchronized at dips and peaks (coincident), whereas the

	GDP	ip	ifo_c	ESI	spr_10y_3m
2024-08-01		93.400	84.900	90.900	-1.400
2024-09-01		90.700	83.700	89.800	-1.400
2024-10-01	902.571	93.200	84.100	90.700	-1.200
2024-11-01		91.200	83.800	89.300	-0.900
2024-12-01		90.800	82.800	86.900	-0.800
2025-01-01	900.761	92.200	82.500	88.100	-0.200

Table 2: Data aligned at sample end.

other indicators tend to lead in relative terms. This characteristic of the additional explanatory variables can be leveraged in a multivariate approach to enhance forecasting performance compared to univariate benchmarks.

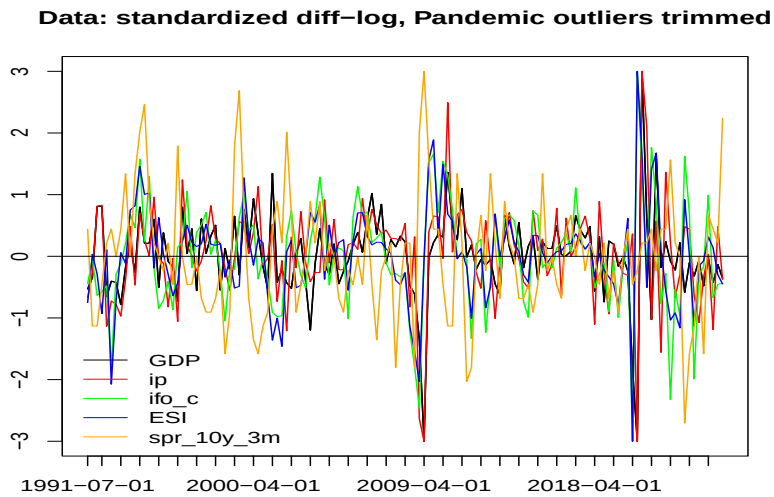

Figure 1: Data: standardized log-differences. Data trimmed to ±3 for ease of visual inspection.

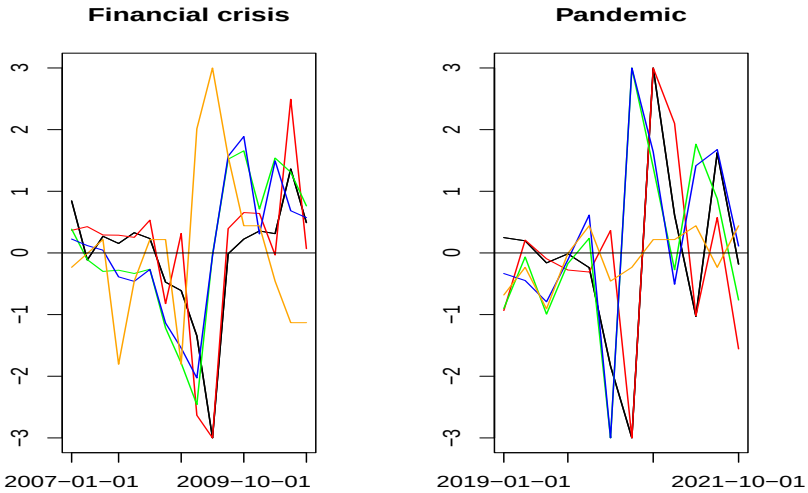

Figure 2: Data: leads and lags during financial crisis (left) and Pandemic (right). Data trimmed to ± 3 for ease of visual inspection.

2.2 Dependence

We analyze the dependence structure of the data and derive a simple VAR model to establish the multivariate filter. The sample cross-correlation function (CCF) is shown in Figs. 3 (entire dataset) and 4 (data prior to the pandemic).

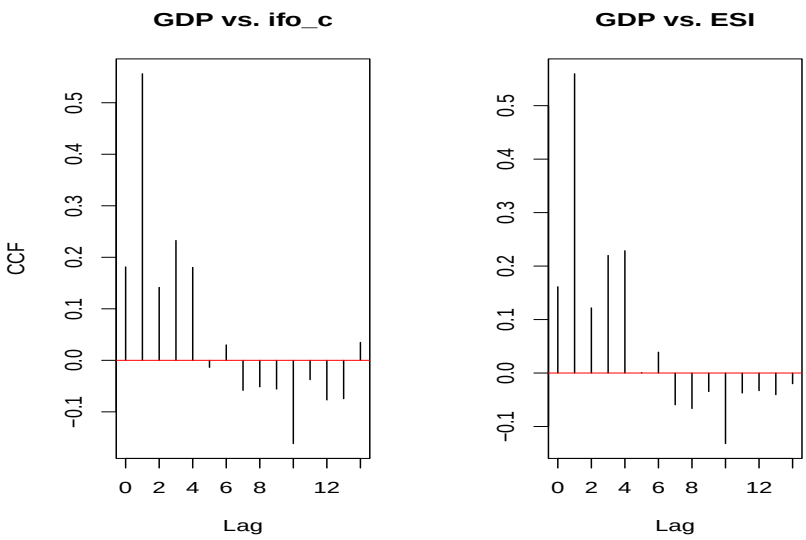

Figure 3: Sample cross-correlation function: GDP and ifo (left), GDP and ESI (right): full data set, including the pandemic. Correlations at positive lags signify that (real-time) GDP is lagging when referenced against the two (real-time) indicators.

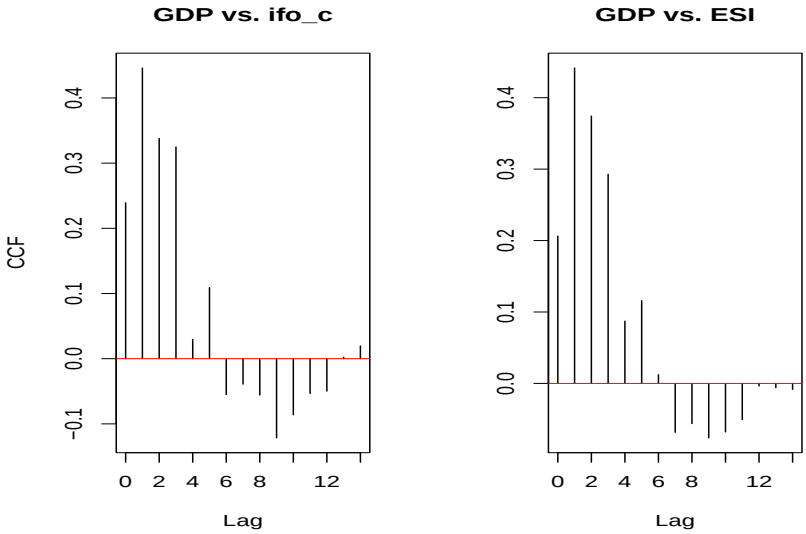

Figure 4: Sample cross-correlation function: GDP and ifo (left), GDP and ESI (right): without pandemic. Correlations at positive lags signify that (real-time) GDP is lagging when referenced against the two (real-time) indicators.

The CCF confirms the leads and lags illustrated in Fig. 2, although the strength of the dependence is influenced by the singular readings during the pandemic[1]. We can now fit a VARMA model, and according to standard diagnostic tests (not shown here), a simple VAR(1) specification aligns with the data. The resulting impulse responses of the model are presented in Figs. 5 (full data set) and (6) (prior the pandemic).

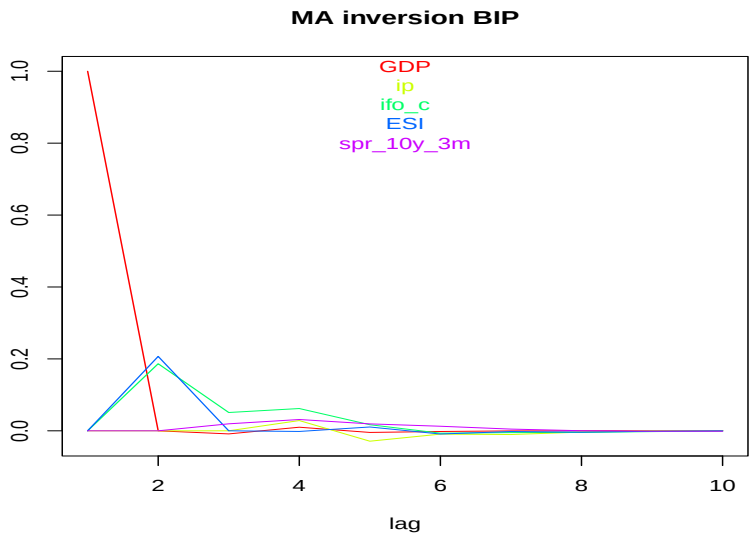

Figure 5: MA inversion of VAR(1): full data set.

[1]We discard the corresponding data when finalizing the predictor.

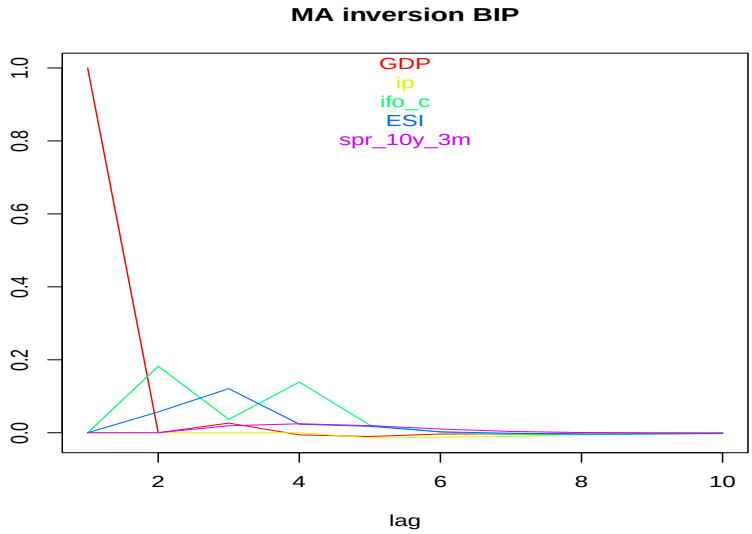

Figure 6: MA inversion of VAR(1): without Pandemic.

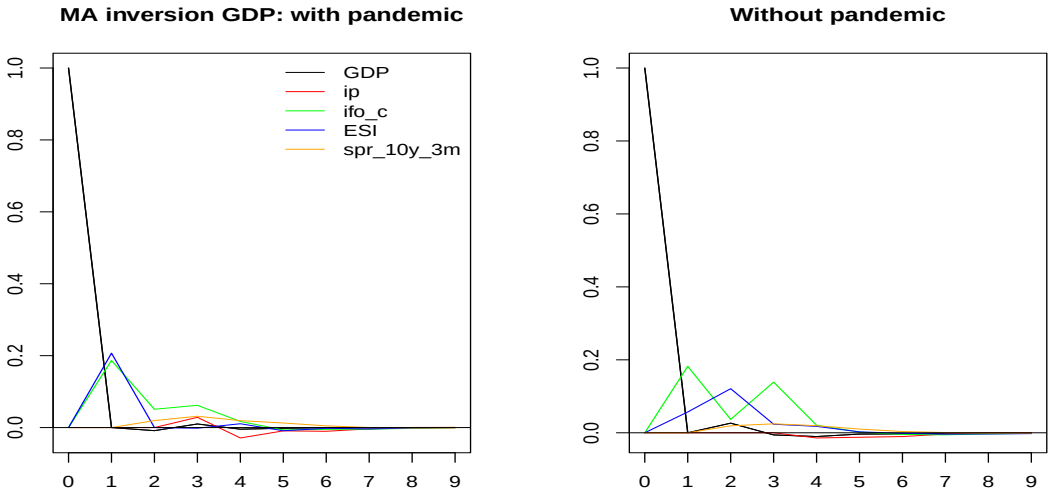

Figure 7: MA inversion of VAR(1) for GDP: with pandemic (left) and without pandemic (right).

The model summarizes key features of the data, where lagging series (GDP and industrial production, when accounting for publication lags) depend on indicators leading in relative terms. In particular, the impulse response of the spread (rightmost panel in the figure) suggests the relevance of a univariate model for this specific leading indicator.

3 Direct Forecasts

3.1 'Classic' Direct Forecast

So-called direct forecasts are obtained by regressing the relevant indicators on forward-shifted BIP, accounting for the additional publication lag. For illustration of the concept, we here rely on all indicators and full sample information, excluding the pandemic. In-sample forecasts and the shifted BIP, which serves as the target, are presented in Figs. 8 (full dataset without the pandemic) and 9 (financial crisis) for horizons up to three quarters ahead.

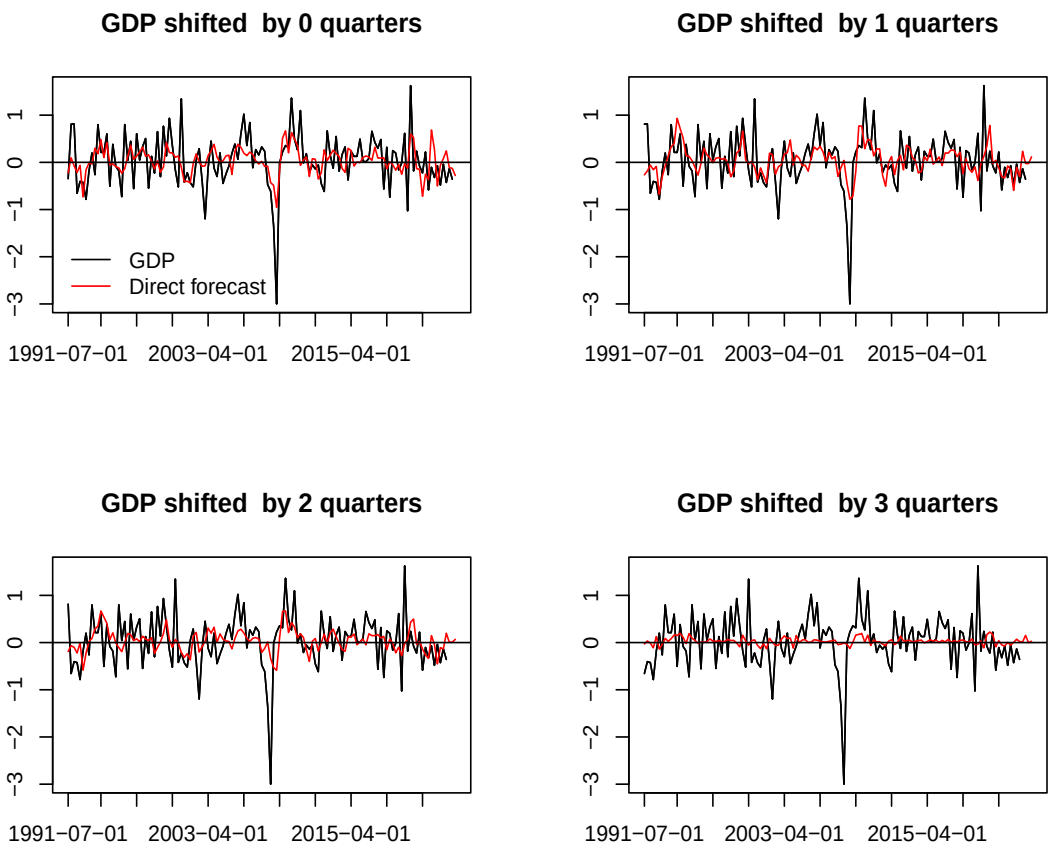

Figure 8: Direct forecasts (red lines) based on all indicators and full sample information, without the pandemic: GDP (black line) shifted by zero one, two and three quarters.

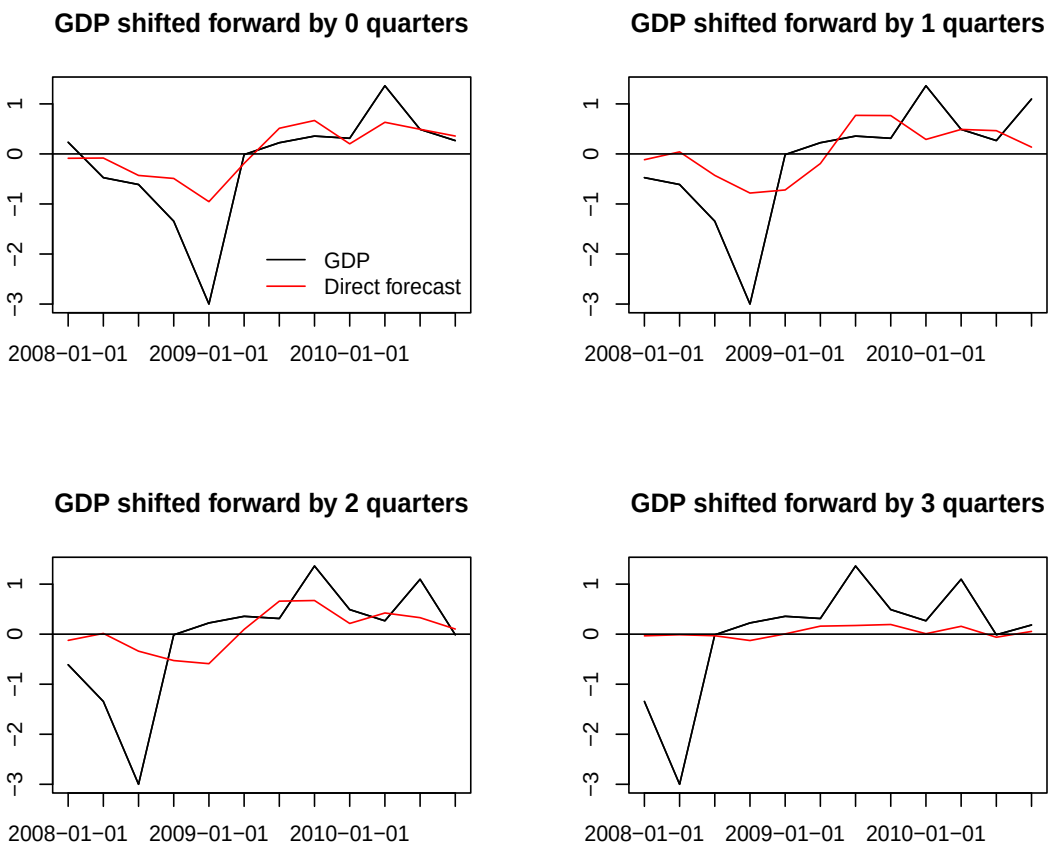

Figure 9: Direct forecast across the financial crisis: regression based on all indicators and full sample information, without the pandemic.

The results indicate that as h increases, the peaks and dips of the direct forecasts become increasingly right-shifted (delayed) relative to the target. Additionally, the F-statistics precipitously decline from $F = 3.21$, for $h = 2$, to $F = 0.28$, for $h = 3$, see Table (3), suggesting that the predictor becomes statistically insignificant for larger forecast horizons exceeding two quarters (link to Heinisch and Scheufele).

3.2 Filter: HP(160)

As shown, the performance of direct forecasts sharply declines for larger forecast horizons. We conjecture that erratic short-term fluctuations -—unpredictable high-frequency noise -— overlay the data, thereby obscuring the effective 'signal' and making a direct regression more susceptible to overfitting. We therefore envisage to highlight the signal by applying a filter to dampen high-frequency noise. The so-called Hodrick-Prescott filter is a classic tool used in business cycle analysis: the filter is specified by a single smoothing parameter λ and the value $\lambda = 1600$ is specifically recommended for quarterly data. However, Phillips and Jin (2021) suggest that the HP(1600) filter removes relevant information due to excessive smoothing. In our context, this oversmoothing issue would be further exacerbated when considering forecast horizons shorter than a year -—an interval inconsistent with the mean duration of up to several years of business cycles, as highlighted by the HP(1600). Consequently, we here select a more adaptive HP(160) 'target'

filter. While this choice may appear somewhat arbitrary, an analysis of the filter in the frequency domain confirms a more suitable profile for the corresponding amplitude function, assigning greater weight to yearly components -—relevant in our prediction framework-— than the classic quarterly HP filter. This approach also addresses and controls for undesirable high-frequency noise. See Fig. 10, which presents both two-sided and (classic) one-sided filters (top panels) along with the associated amplitude functions (bottom panels). We now posit the HP(160) filter as a means to achieve better forecasting performance at larger horizons, noting that a comprehensive technical analysis of the effect of λ on the resulting BIP predictor can be conducted using the M-SSA package.

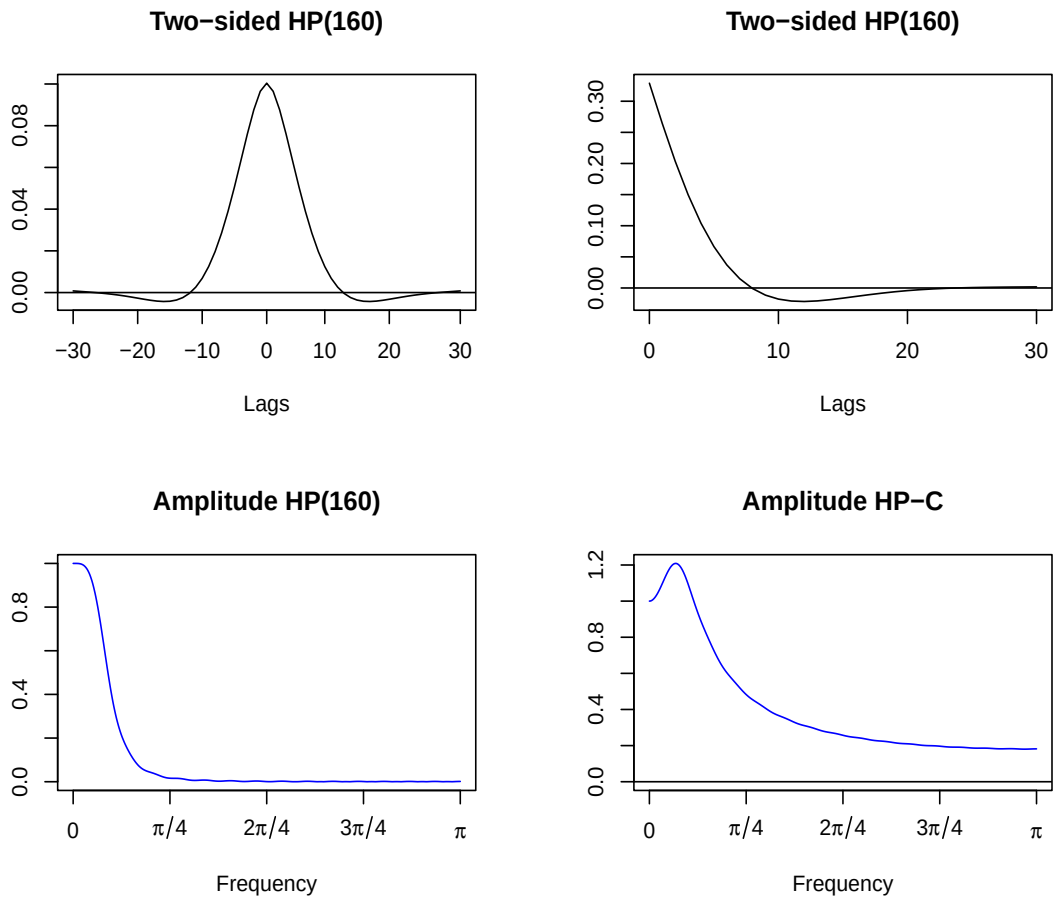

Figure 10: Two-sided HP-160) (top left) and one-sided (concurrent) HP-C filters (top right) with corresponding amplitude functions (bottom).

3.3 Direct HP Forecast

Extending the 'classic' direct forecast design, we here consider filtered indicators, based on the one-sided HP(160), denoted as HP-C: the filtered indicators are explanatory variables for the regressions on forward-shifted BIP[2]. As in the previous section, we rely on all indicators as well

[2]Katja and Simon: for each forecast horizon $h = 0, 1, 2, 3$ I'm using corresponding forecasts of the real-time filter output, assuming the data to be white noise (assuming an ARMA model does not lead to significantly different results). I think we might ignore these technical details in the paper.

as the full sample estimate for illustration, see Figs.(11) (all observations with exclusion of the pandemic) and (12) (financial crisis). A comparison with the classic direct forecasts in Fig.(9) indicates that the new predictors are slightly less retarded ('faster').

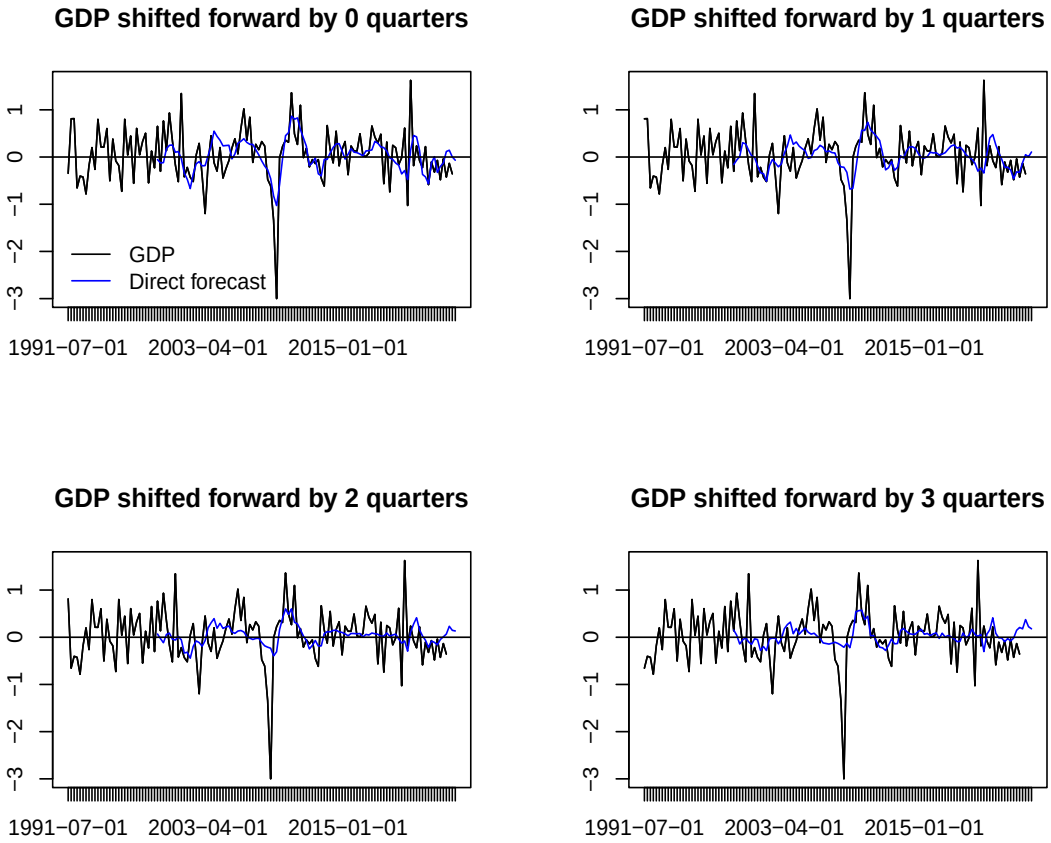

Figure 11: Direct HP forecasts: entire data set with exclusion of the pandemic.

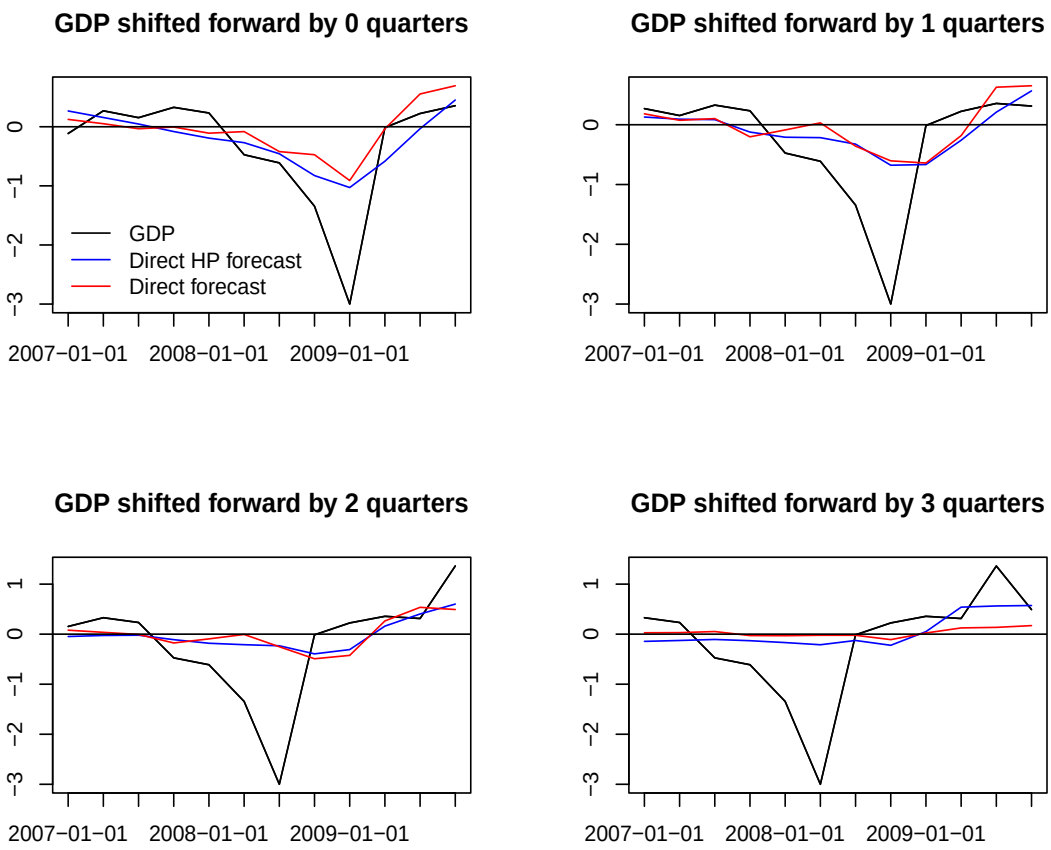

Figure 12: Shifted GDP (black), direct HP forecast (blue) and classic direct forecast (red) over the curse of the financial crisis: regression based on all indicators and full sample information (without pandemic).

Moreover, the F-statistics are larger than for the direct forecasts, see Table (3), but the predictor remains statistically insignificant at forecast horizons exceeding two quarters (even though the evaluation is conducted in-sample and based on all available indicators).

	h= 0	h= 1	h= 2	h= 3	h= 4	h= 5	h= 6
HP-C filtered	0.000	0.001	0.036	0.046	0.049	0.012	0.022
Unfiltered	0.028	0.000	0.001	0.910	0.509	0.091	0.178

Table 3: Statistical significance of unfiltered and HP-C filtered forecasts based on HAC-adjusted Wald test: all indicators and full sample (without pandemic).

We now propose a more refined causal filter design aiming for statistical significance of the resulting predictor at longer forecast horizons, out-of-sample.

4 Multivariate Causal Filter: M-SSA

The application of the one-sided HP-C filter to the indicators has demonstrated improvements in forecast performance. However, the univariate nature of this filter limits its ability to leverage information from the cross-section, as supported by leading indicators. Additionally, the noise suppression capability of the one-sided filter is compromised, as illustrated by high-frequency leakage of the amplitude function in Fig. (10) and further elaborated upon by Wildi (2025). To address these limitations, we propose an extension of the Smooth Sign Accuracy (SSA) framework introduced by Wildi (2025).

4.1 Optimization criterion

Let $\mathbf{X}_t$ (of dimension $t \times n$) denote a set of n explanatory series $\mathbf{x}_{1t}, ..., \mathbf{x}_{nt}$, with observations $\mathbf{x}_{it} = (x_{i1}, ..., x_{it})'$. Let $\mathbf{z}_t = (z_1, ..., z_t)$ denote a target series, which typically lies outside the linear space spanned by $\mathbf{X}_t$ (and is unknown at time t). For simplicity, we assume stationarity of all time series involved. In this framework, $\mathbf{X}_t$ corresponds to a matrix of selected economic indicators with $n = 5$, and z_{it} is the output of a two-sided HP(160) filter applied to the i-th indicator, $i = 1, ..., 5$. Since z_{it} depends on future observations $x_{it+1}, x_{it+2}, ...$, the prediction task involves 'tracking' z_{it} using the estimate $y_t^i = \sum_{j=1}^n y_{jt}^i$, where $y_{jt}^i = \mathbf{b}_j^i{}'\mathbf{x}_{jt} = \sum_{k=0}^L b_{jk}^i x_{jk}$ are the outputs of a multivariate (causal or one-sided) filter $\mathbf{B}^i$, with columns $\mathbf{b}_j^i$ of fixed length L, $j = 1, ..., 5$, assigning weight to the last L observations of the j-th indicator $\mathbf{x}_{jt}$ (the fixed-length assumption is merely used for ease of exposition). For the sake of clarity, we now omit the upper-case index i for all variables, assuming $z_t = z_t^{i_0}$ for some fixed i_0.The task of 'tracking' a target can be formalized in different ways; here, we focus on the target correlation $\rho(z, y, h)$ between y_t and z_{t+h}, where $h \geq 0$ denotes the forecast horizon (backcasting with $h < 0$ is ignored in this context). To streamline the notation and avoid unnecessary complexity, we assume a fixed horizon, say $h = h_0$, so that we may drop the reference to h in our notation. In the univariate case ($n = 1$), Wildi (2025) proposes the Simple Sign Accuracy (SSA) as an optimization criterion for this task:

$$\left.\begin{array}{c} \rho(z, y) \to \max \\ \rho(y, 1) = \rho_1 \end{array}\right\}. \tag{1}$$

Objective function and constraint are expressed in terms of correlations, namely the target correlation $\rho(z, y)$ between target and predictor and the lag-one autocorrelation $\rho(y, 1)$ of the predictor. The parameter ρ_1 governs the smoothness of the predictor: higher values of ρ_1 generally promote smoother trajectories of y_t. Formally, assuming that the process x_t is centered (zero-mean), Wildi (2024) establishes a connection between $\rho(y, 1)$ and the expected duration between consecutive zero-crossings (sign changes) of y_t, the so-called holding time, denoted by $ht(y)$:

$$ht(y) = \frac{\pi}{\arccos(\rho(y, 1))}.$$

Since the non-linear transformation involved is strictly monotonic, the SSA criterion can be reformulated as the optimization problem:

$$\left.\begin{array}{c} \rho(z, y) \to \max \\ \frac{\arccos(\rho(y,1))}{\pi} = 1/ht_1 \end{array}\right\}, \tag{2}$$

where the smoothness parameter $1/ht_1$ expresses the rate of zero-crossings of the predictor[3].

We now briefly discuss some implications of criterion (2). First, the objective function is indifferent to an affine transformation of the predictor. This ambiguity can be resolved by assuming an arbitrary scale and level for y_t (standardization). Alternatively, a mean-square error norm

[3]Formally, the exact relationship between the zero-crossing rate and the lag-one autocorrelation of the predictor assumes Gaussian time series, but Wildi (2024) demonstrates robustness to deviations from this assumption.

(MSE) may be substituted for the target correlation, as noted in Wildi (2025). In this case, the classic MSE predictor $y_{t,MSE}$ is obtained as a solution to the SSA criterion by insertion of $1/ht_1 := 1/ht_{MSE}$ in the constraint, where ht_{MSE} represents the holding time of $y_{t,MSE}$. Second, the concept can be extended to a multivariate framework, denoted as M-SSA, by generalizing the correlation functions, as detailed in Wildi (2025b). Lastly, sign changes in trend growth, represented by zero-crossings of the target z_t, are indicative of relevant changes in the trajectory of the economy, for example at transitions between expansions and recessions. Unfortunately, classic predictors often generate excessively many false 'noisy' crossings, due to high-frequency leakage as illustrated by the amplitude function in Fig. (10) (bottom right panel) and further elaborated upon by Wildi (2025). The SSA criterion helps control this phenomenon. More precisely, Wildi (2025) derives an equivalent dual formulation of criterion (2), which distinguishes the SSA solution as exhibiting the smallest rate of zero-crossings of any (linear) predictor with the same target correlation. We now rely on the M-SSA criterion to derive a 'smooth' predictor y_t of z_t.

4.2 M-MSE and M-SSA Nowcasts

For illustration, we consider a nowcast of the target z_t generated by the output of the two-sided HP(160) filter applied to BIP, which we denote by HP-BIP. We compare the classic multivariate mean squared error (MSE) predictor, denoted by M-MSE, with the M-SSA predictor, where we set $ht_1 = 1.5ht_{MSE}$ in the holding time constraint. This configuration ensures that the M-SSA predictor produces approximately 33% fewer zero-crossings than M-MSE over the long term. The purpose of this constraint is to compare the effect of a different level of smoothness on the prediction performance (target correlation). The estimation of the underlying VAR(1) model employed for the M-MSE and M-SSA is based on data observations up to January 2008. This approach ensures a lengthy out-of-sample period that encompasses significant events such as the financial crisis, the sovereign debt crisis, and the COVID-19 pandemic. Further details and background information are available in Wildi (2025b). All computations are performed utilizing the M-SSA package, as documented in Wildi (2025b).

Figure 13 displays the coefficients of the multivariate nowcasts, where increased smoothing achieved through the M-SSA method manifests as a distinct pattern in the filter weights, notably influencing the rate of decay across increasing lags. The resulting filtered series, presented in Figure 14, are standardized to enable direct comparisons. The M-SSA nowcast adheres to the specified constraint, exhibiting approximately one-third fewer zero-crossings than the M-MSE, confirming the effectiveness of the smoothness control imposed during the design. This is further corroborated by Table 4, which reports empirical holding times and target correlations. Increased smoothness achieved by the M-SSA comes at the cost of a slight reduction in tracking accuracy, as indicated by a marginally lower target correlation. This highlights the inherent trade-off—often referred to as the prediction dilemma—between smoothness (zero-crossing rate) and predictive accuracy (target correlation) within the M-SSA criterion.

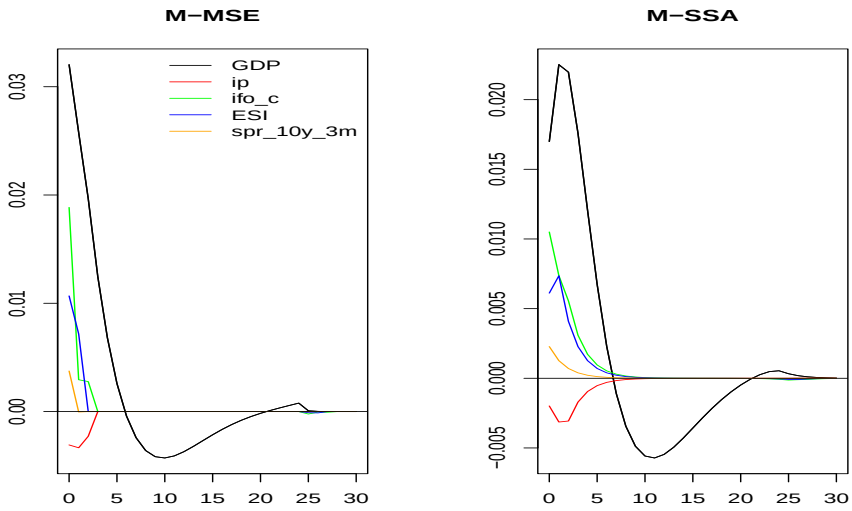

Figure 13: Multivariate M-MSE (left) and M-SSA (right) nowcast filters.

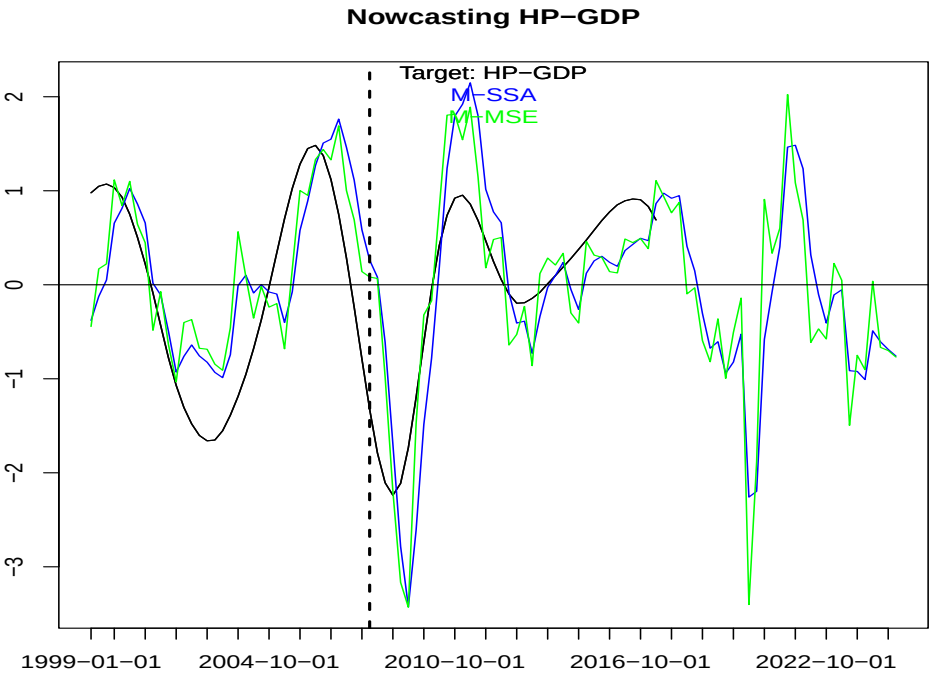

Figure 14: Two-sided HP(160) applied to GDP (target, black) and nowcasts: M-SSA (blue) and M-MSE (green). The dashed vertical line delimits in- and out-of-sample spans. The two-sided filter does not extend to the sample end.

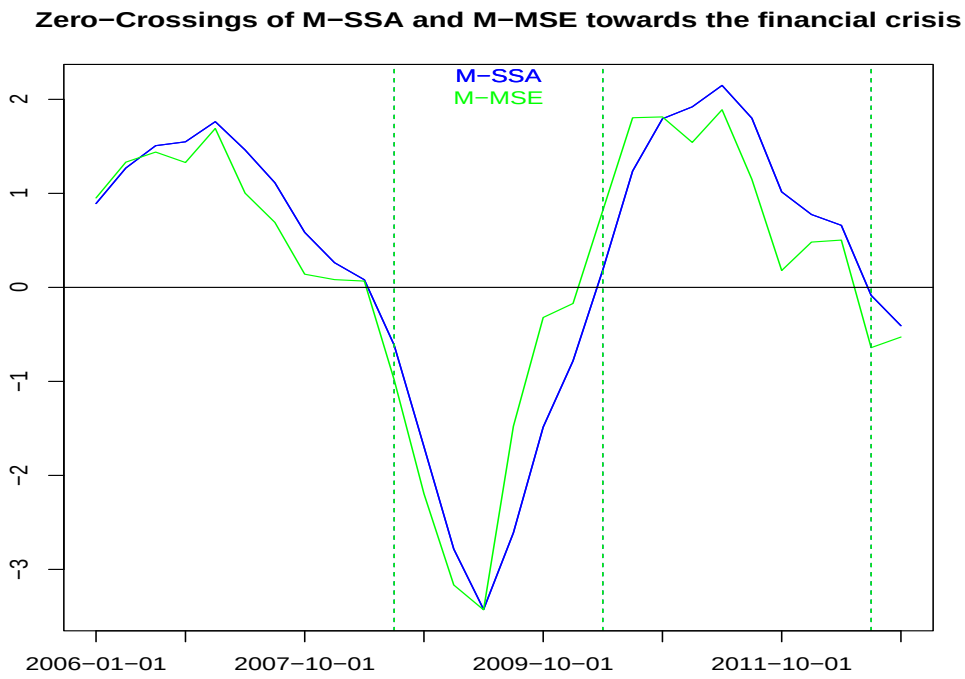

Figure 15: Sign changes of the M-MSE (green) and M-SSA (blue) nowcasts are examined in relation to the financial crisis, with sign changes of the M-MSE marked by green vertical lines. The sample holding time (HT) of each filter corresponds to the average duration between its consecutive zero-crossings. For each filter, its sample HT converges to the respective expected HT asymptotically, given sufficiently long samples. Here, we impose $HT_{MSSA} = 1.5 \times HT_{MSE}$: in the long run, we expect the M-MSE to exhibit approximately 50% more sign changes than the M-SSA.

	M-SSA	M-MSE
Target correlation	0.70	0.73
HT	7.43	5.78

Table 4: Sample target correlations and HTs of M-SSA and M-MSE nowcasts.

The results validate that the smoothing constraint incorporated into the multivariate extension of Criterion (1) effectively manages the zero-crossings of the predictor: sample estimates of the holding times (HTs) align with the expected values derived from Equation (2)[4]. An increase of 50% in the HT relative to the M-MSE predictor has only a modest effect on the target correlation, suggesting that this tradeoff could be beneficial in our application. Going forward, we now omit the M-MSE, as it is a special case of the more general M-SSA framework.

[4]Effective convergence of sample estimates to their theoretical expected values —assuming the true data-generating process is known— can be validated using the M-SSA package. This validation is typically performed through analyses of very long samples of simulated (multivariate) data, ensuring the consistency of the estimation procedures under idealized conditions.

4.3 M-SSA Forecasts

The previously described nowcast example can be systematically extended to generate forecasts of the target variable. This extension utilizes the same empirical framework, incorporating an assumed 50% increase in the HT over the M-MSE, and considers forward shifts of $h = 0, ..., 6$ quarters of HP-BIP (note that M-SSA does not target BIP explicitly). The influence of the forecast horizon on the M-SSA design is depicted in Fig. (16), which contrasts nowcast and one-year-ahead forecast: as h increases (right panel), the scaling factor diminishes (zero-shrinkage), attributable to increased forecast uncertainty, and there is a relative amplification of the influence of the relevant indicators, leading BIP in relative terms.Additionally, a phase shift associated with the filter applied to BIP is observed (red line, right panel), reflecting the cyclical characteristics inherent to the HP filter. While this phase effect might cause an effective sign change in the BIP filter output as h increases—something to be avoided in this context—the low-pass filters assigned to the additional 'leading' indicators (green and violet lines, right panel) continue to effectively track the low-frequency components of HP-BIP. The combination of the phase effect, applied to BIP, and level-tracking through these additional indicators enables more refined and effective tracking of the target compared to univariate forecasts, which rely solely on the phase-effect of the filter to look ahead of time.

M-SSA predictors are compared to the traditional univariate HP-C in Figs. (17) and (18). In the first figure, we compare both filters across the periods of the financial crisis (top panels) and the COVID-19 pandemic (bottom panels), for forecast horizons $h = 0, ..., 6$. The nowcast from the previous section is included for reference. In the second figure we compare both filters to the forward-shifted BIP for shifts larger than two quarters. All series are standardized to facilitate visual comparison. The primary distinction of the M-SSA approach is the progressive left-shift of the predictor as a function of h; this shift becomes more pronounced with increasing h and appears to be consistent across all levels, including the peaks and troughs of the series. Conversely, the positions of peaks and troughs in the HP-C series appear to remain largely unaffected by variations in h. The more pronounced and systematic left-shift observed in the M-SSA forecasts can be partly attributed to the increasing weight assigned to the additional indicators leading BIP, which dominate the BIP component at larger forecast horizons.

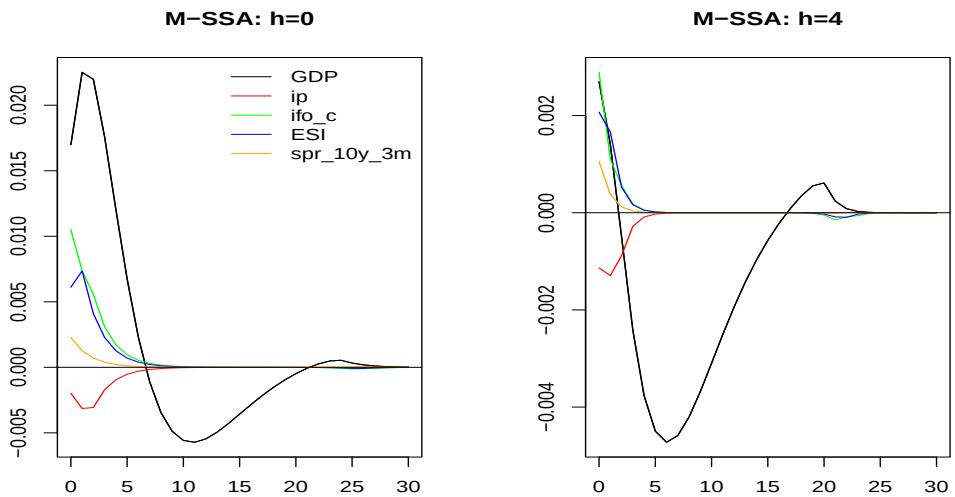

Figure 16: Multivariate M-SSA: $h = 0$ (nowcast, left) and $h = 4$ (one year forecast, right).

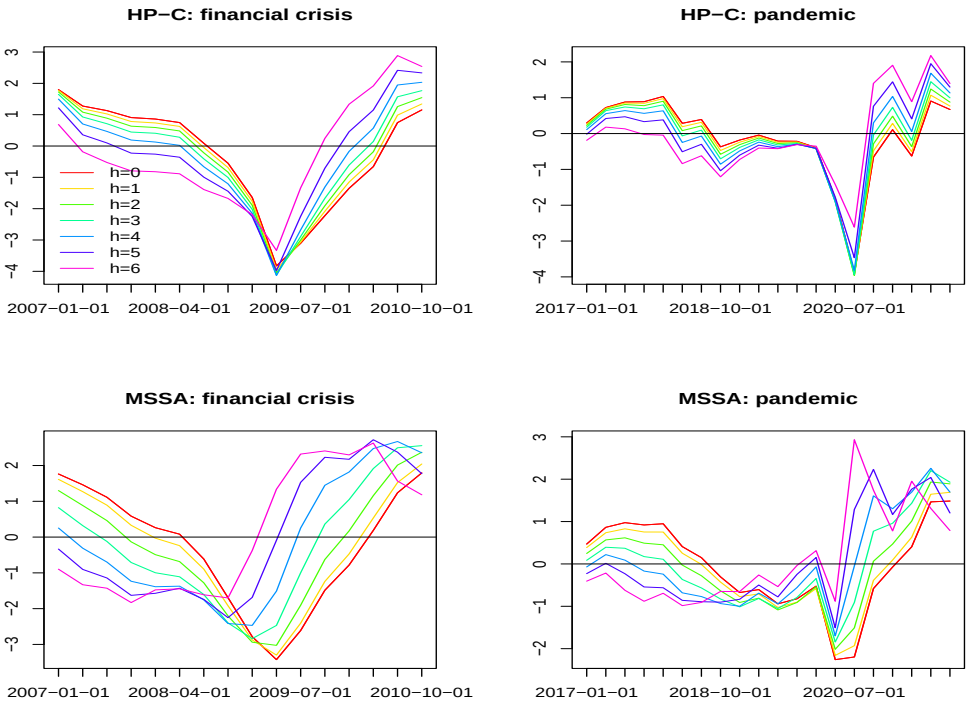

Figure 17: Now- and forecasts of HP-GDP for $h = 0, ..., 6$: HP-C (top) vs. M-SSA (bottom) across the financial crisis (left) and the pandemic (right).

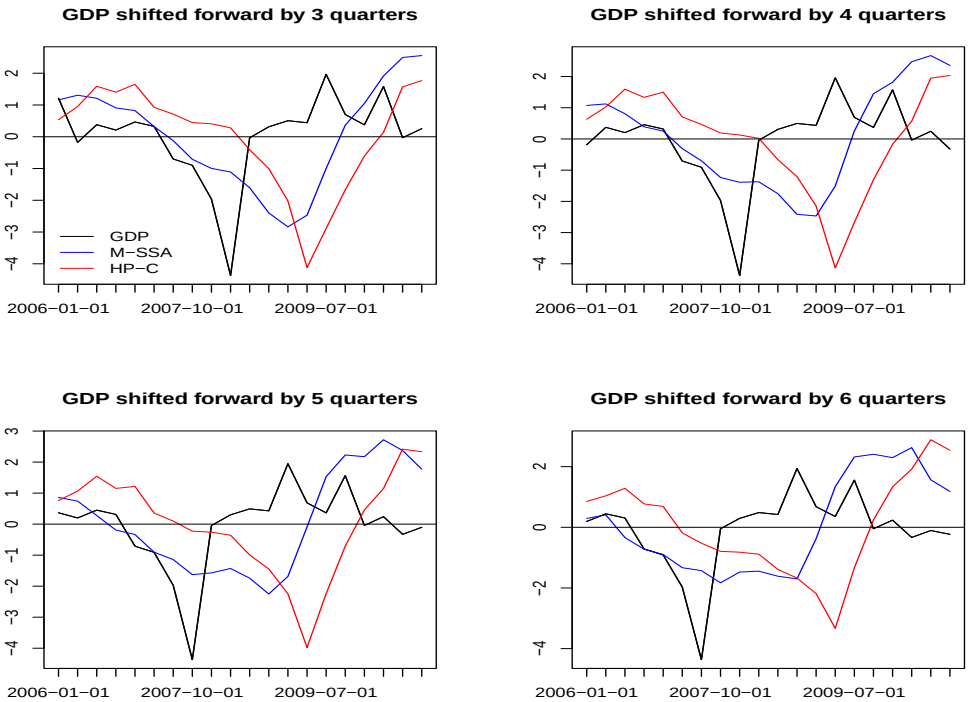

Figure 18: Forward-shifted GDP (black), M-SSA (blue) and HP-C (red) for $h = 3, 4, 5, 6$ across the financial crisis. All series are standardized for better visual inspection.

5 M-SSA Component Predictor

The targets of the multivariate filters are the smoothed outputs obtained from the acausal two-sided HP(160) filter applied to the five indicators. However, our objective is to develop predictors that emphasize the unfiltered BIP. To achieve this, we can utilize the M-SSA outputs, referred to as M-SSA components.

5.1 Standardized Predictor: Equal-Weighting

Assuming that all M-SSA components contribute equally to BIP prediction, an initial step involves averaging the standardized components to form an equally-weighted composite measure. Subsequently, the predictor is derived by regressing this standardized aggregate on the (forward-shifted) BIP series. Fig. (19) presents the BIP series alongside the equally-weighted M-SSA predictor. All series are standardized to facilitate visual comparison and to highlight the dynamic shifts, particularly the left-shift characteristic of the predictors.

To enhance interpretability and assessment of the M-SSA predictor, we propose analyzing its individual components. For illustration, Fig. (20) shows the nowcast at $h = 0$. The solid blue line (the nowcast) approaches the zero line near the end of the sample, with a recent upward movement mainly supported by the leading spread component. In contrast, the ifo- and ESI-components remain near the zero line, while the ip- and BIP-components appear to be awaiting additional evidence and confirmation before signaling a definitive movement. This type of assessment can be useful for evaluating the relevance of recent changes in the forecasted time series dynamics.

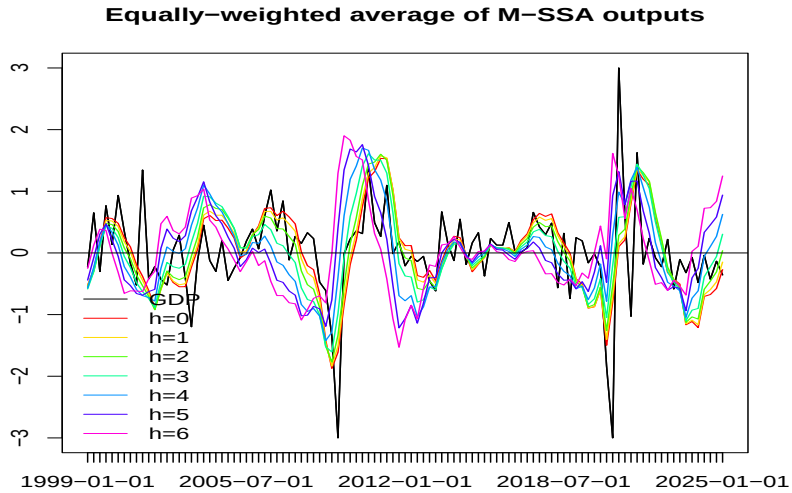

Figure 19: BIP and equally-weighted M-SSA predictor: all series standardized.

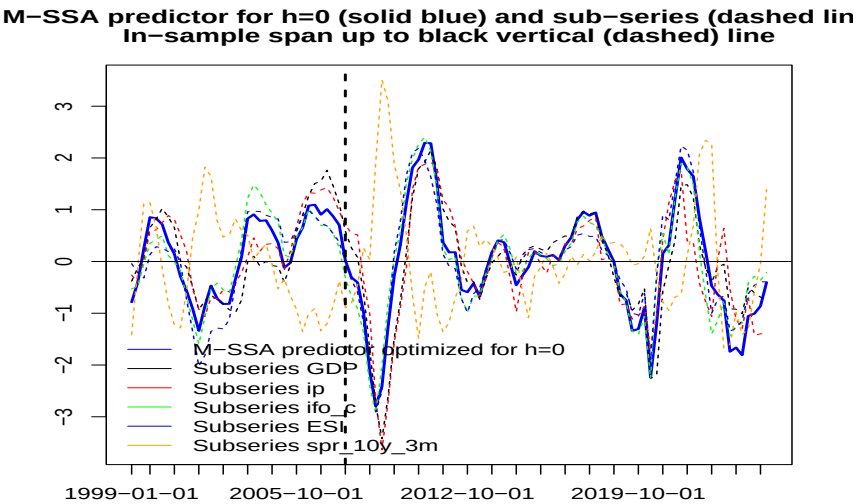

Figure 20: Equally-weighted M-SSA predictor and M-SSA components: all series standardized.

5.2 Tracking BIP: Optimal Weighting

We now utilize the M-SSA components as regressors on forward-shifted BIP, extending the previous (standardized) equally-weighted M-SSA predictor to incorporate optimal weighting of the components. While we keep the M-SSA fixed, as based on data up to January 2008, we implement an expanding window starting from January 2007 for deriving the component weights, which are up-dated on a quarterly basis (this contrasts with Section (3) which was based on the full sample information).

There are various combinations of M-SSA components for tracking BIP. For illustration, we here select the single BIP component resulting from the M-SSA filter tracking HP-BIP. Table (5) reports HAC-adjusted p-values from regressions of the out-of-sample predictor on HP-BIP, indicating a strong link between the predictor and the target, out-of-sample. Our choice—focusing

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.000	0.000	0.000	0.000	0.002	0.036	0.488
Shift=1	0.000	0.000	0.000	0.000	0.000	0.006	0.058
Shift=2	0.001	0.000	0.000	0.000	0.000	0.002	0.012
Shift=3	0.067	0.027	0.004	0.002	0.001	0.001	0.005
Shift=4	0.429	0.271	0.087	0.007	0.006	0.003	0.004
Shift=5	0.835	0.710	0.428	0.097	0.018	0.011	0.007

Table 5: HAC adjusted p-values of regressions of BIP (M-SSA) component optimized for forecast horizons $h = 0, ..., 6$ (the columns) on HP-BIP shifted forward by shift=0,...,5 (the rows). Out-of-sample span starting in Jan-2007, without pandemic.

on the single M-SSA BIP component—emphasizes simplicity and interpretability, as future BIP and HP-BIP are connected through their shared low-frequency component. Additionally, the M-SSA BIP component is the most important predictor in the regression of the M-SSA outputs on BIP. Finally, revisions due to quarterly updates of the regression equations are minimized by this predictor, and the estimates are more stable and robust over time[5] (see Section (6)). Note that information from the other indicators is incorporated into the M-SSA BIP component through the multivariate filter design.

This setup allows us to evaluate the out-of-sample performance of various predictor approaches, including: (i) the simple mean of BIP, (ii) the direct forecast method described in Section (3.1), (iii) the direct HP forecasts from Section (3.3), based on the univariate HP-C, and (iv) the new M-SSA BIP component predictor.

5.3 Out-of-Sample Forecast Performances

Tables 6 and 7 present HAC-adjusted p-values and rRMSEs for the GDP component predictor. They are obtained from the simple regressions

$$\text{GDP}_{t+shift} = \beta_0(shift, h) + \beta_1(shift, h) \cdot \text{M-SSA}_t^{\text{GDP}}(h) + \epsilon_t^{shift,h} \tag{3}$$

where $\text{M-SSA}_t^{\text{GDP}}(h)$ is the out-of-sample M-SSA predictor for GDP (not HP-GDP.) The p-values test $H_0 : 0 = \beta_1(shift, h)$ and do not account for potential bias or mis-scaling of the GDP predictor. In contrast, the rRMSEs, evaluated against the mean benchmark, incorporate eventual bias and mis-scaling effects. Additional results involving other predictor benchmarks are provided in the appendix, based on data sets with and without the pandemic.

The purposes of Equation 3 are twofold: first, the regression parameters enable affine transformations of the left-hand side variable, so that the standardized target $\text{GDP}_{t+shift}$ can be replaced by the original log-returns without affecting the p-values or rRMSEs reported in Tables 6 and 7 (noting that the benchmark predictors, i.e., the direct forecasts and the arithmetic mean, similarly adjust to affine transformations); second, the static scale adjustment by $\beta_1(shift, h)$ compensates for the increasing zero-shrinkage of $\text{M-SSA}_t^{\text{GDP}}(h)$, reflecting greater uncertainty as the forecast horizon h increases. This simple transformation allows $\text{M-SSA}_t^{\text{GDP}}(h)$ to be used as a predictor for $\text{GDP}_{t+shift}$ even when $h \neq shift$. Such flexibility enables assessment of the dynamic left-shift

[5] Other combinations can be experimented with using the M-SSA package. Regression weights of more complex designs, involving multiple components, tend to be more difficult to interpret due to the occurrence of negative weights and potential sign changes of the regressors over time, which may indicate overfitting or non-stationarity.

of M-SSA$_t^{\text{GDP}}(h)$, as illustrated in Figure **??**, at various forward-shifts of GDP$_{t+shift}$.

When excluding the pandemic, the direct forecasts outperform the mean forecast up to two quarters ahead, see Table 12.[6] However, the M-SSA component predictor, optimized for longer forecast horizons, outperforms the mean benchmark up to 6Q ahead. Additionally, the GDP component predictor optimized for large forecast horizons $h \geq 4$ surpass the direct forecasts for $shift > 2$, see Table 11. Statistical significance of the GDP component predictor, optimized for larger forecast horizons, can be established for at least up to one year ahead.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.077	0.049	0.017	0.004	0.002	0.001	0.079
Shift=1	0.249	0.186	0.108	0.039	0.004	0.001	0.008
Shift=2	0.805	0.694	0.436	0.181	0.029	0.003	0.005
Shift=3	0.905	0.949	0.878	0.585	0.187	0.034	0.023
Shift=4	0.722	0.944	0.984	0.900	0.471	0.075	0.015
Shift=5	0.591	0.892	0.995	0.958	0.588	0.232	0.031

Table 6: HAC adjusted p-values of regressions of M-SSA component predictors optimized for forecast horizons $h = 0, ..., 6$ (the columns) on BIP shifted forward by shift=0,...,5 (the rows). Out-of-sample span starting in Jan-2007, without pandemic.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.981	0.947	0.888	0.837	0.841	0.897	0.967
Shift=1	1.087	1.071	1.012	0.926	0.866	0.870	0.919
Shift=2	1.075	1.089	1.074	1.010	0.936	0.897	0.904
Shift=3	1.036	1.069	1.090	1.059	0.991	0.943	0.931
Shift=4	1.012	1.058	1.101	1.096	1.036	0.966	0.927
Shift=5	1.015	1.064	1.107	1.114	1.078	1.015	0.951

Table 7: rRMSEs of M-SSA components predictor benchmarked against the expanding mean of BIP. Out-of-sample span starting in Jan-2007, without pandemic.

6 Revisions

Revisions refer to changes in the historical values of a time series resulting from updates with new information. Revisions of the M-SSA component predictor can arise from data updates, which we do not analyze further here[7], or from updates to predictor parameters. The latter can be further separated into changes in VAR parameters and regression weights.

6.1 Revisions: Regression Equation Only

For illustration, we here focus on a forecast horizon $h = 4$ for the BIP component predictor and a forward-shift $shift = 4$ for BIP, noting that similar results would be obtained with other combinations of h and $shift$. Fig.(21) compares the final and real-time predictors: as the sample size increases, the real-time predictor converges steadily to the final predictor and the rate of convergence is partially determined by the simplicity of our design, which avoids incorporating additional M-SSA components. As a confirmation, Fig.(22) displays the evolution of the intercept and regression weights over time: with increasing sample size, these estimates seem to converge to fixed points, indicating both the stationarity of the process and the consistency of the estimates (note,

[6] However, the pandemic period generally obscures the relationships between predictors and GDP.

[7] Except for BIP, the other indicators are not or only slightly revised. Moreover, the smoothing effect of filters mitigates the effect of revisions when compared to direct forecasts.

however, the slightly elevated values of the regression weight assigned to the M-SSA component during the financial crisis and the pandemic, suggesting a stronger link between the target and the predictor). In contrast, more complex designs based on multiple M-SSA components can exhibit substantial fluctuations in the regression parameters, sometimes changing signs. This suggests instability and complicates interpretation.

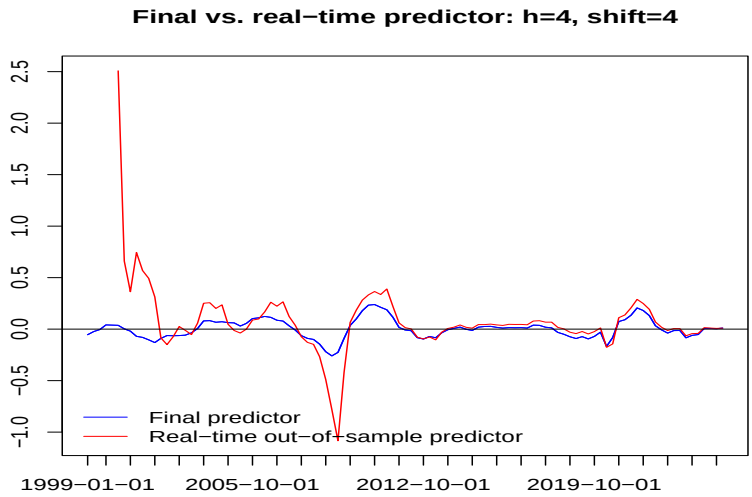

Figure 21: Revisions of the M-SSA GDP component predictor are shown, with the final predictor in blue and the real-time design in red. These revisions result from quarterly updates to the regression equations. The M-SSA filter itself remains fixed, based on data up to January 2008.

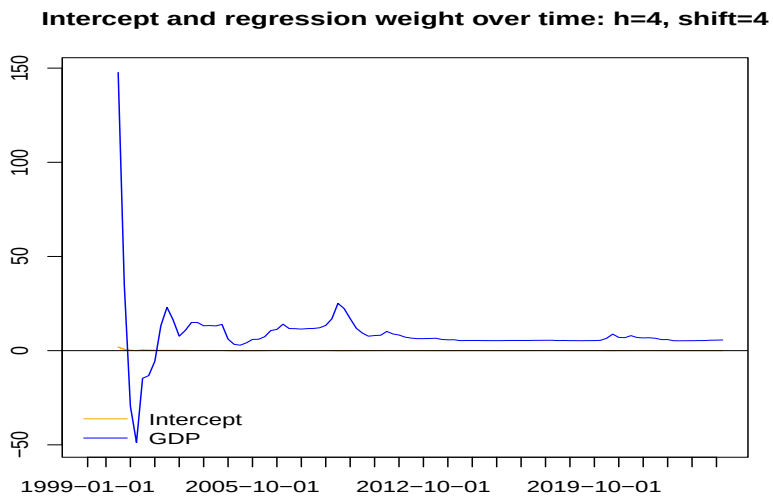

Figure 22: Revisions of intercept and regression weight.

6.2 Revisions: M-SSA Only

The second source of revisions in the M-SSA component predictor arises from updates to the VAR model within the M-SSA filter. We compare the original estimates, based on data up to January 2008, with the final estimates obtained using $h = 4$ and $shift = 4$ (similar findings apply to other values of forecast horizon and forward-shift). For consistency, we excluded the entire pandemic episode from the data. Fig.(23) compares the MA-inversions (impulse responses) of the BIP equation derived from the 'old' and the updated VAR models: unlike the former, the latter incorporates ESI as an additional explanatory variable for BIP. Fig.(24) illustrates the resulting effect on the M-SSA filter targeting HP-BIP at forecast horizon $h = 4$. The updated filter (right panel) assigns more weight to the additional 'leading' indicators relative to BIP. Finally, Fig.(25) compares the M-SSA BIP predictors. To isolate the effect of the regression revision, the figure displays standardized series. The update to the VAR model influences the left-shift of the predictor (blue line), which appears slightly 'faster' than the original predictor (red line). This suggests that the out-of-sample performance of the M-SSA component predictor may be better than reported in Section (5.3), which is based on the fixed and increasingly outdated version of M-SSA.

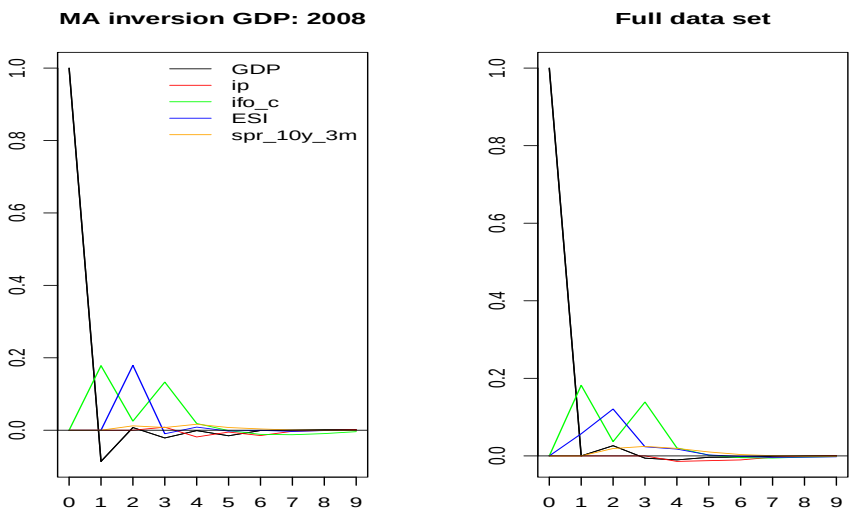

Figure 23: MA-inversion (impulse response) of VAR: data up to Jan-2008 (left) vs. entire data set without pandemic (right).

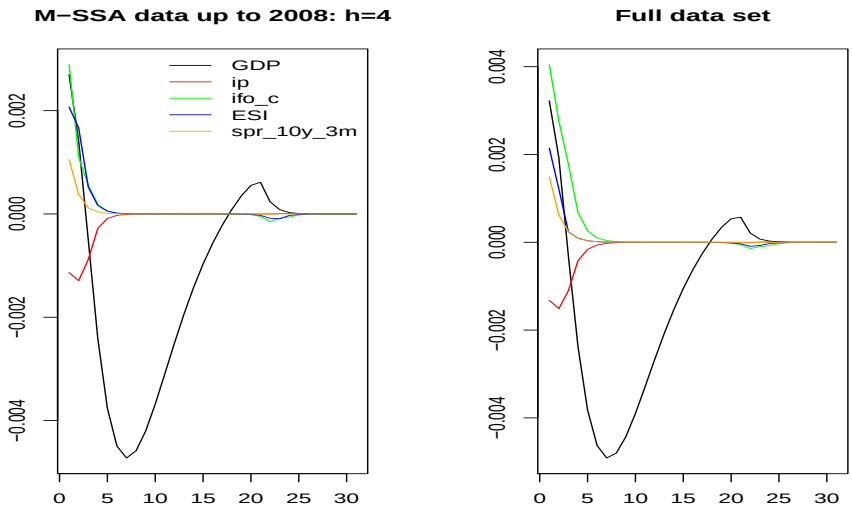

Figure 24: M-SSA filter tracking HP-GDP at forecast horizon $h = 4$: data up to Jan-2008 (left) vs. entire data set without pandemic (right).

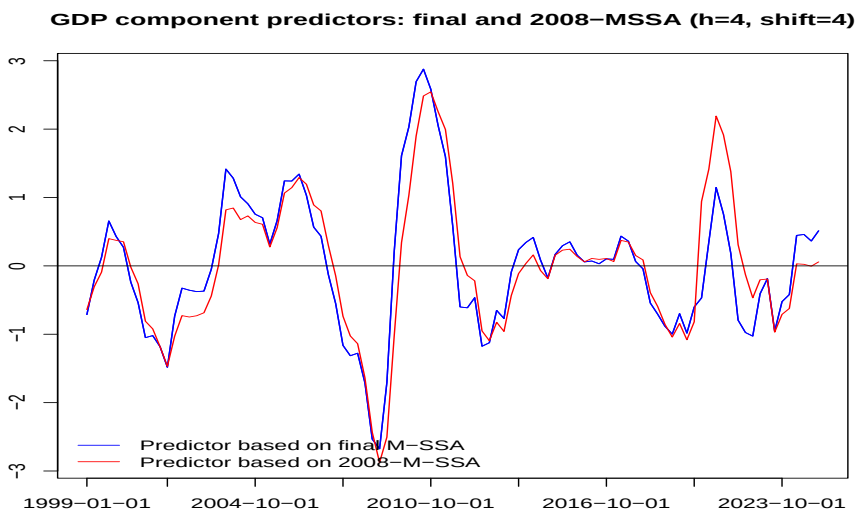

Figure 25: Revisions resulting from updates to the VAR model in M-SSA optimized for $h = shift = 4$. To isolate the effect of the regression revision, all series are standardized. The updated M-SSA (blue) is left-shifted at the zero-crossings and appears slightly faster. For consistency, the entire pandemic episode has been removed.

7 Hit Rate vs. False Alarm Rate

Remarks:

- Alternative to MSE, target correlation, holding-time: hit-rate, false alarm.

- We can link objective function (target correlation) with sign accuracy, see Wildi (2025). BUT: risk that reviewer asks for comparison with logit or whatever sign-forecast tool.
- Context ROC-curve

For the ROC curves, we consider predictors of GDP based on both original (unfiltered) and filtered indicators. For the direct forecast, we use the regression

$$GDP_{t+shift} = \beta_0^{shift} + \sum_{i=1}^{5} \beta_i^{shift} x_{ti} + \epsilon_t^{shift,direct} = G\hat{D}P_{t+shift}^{direct} + \epsilon_t^{shift,direct},$$

where x_{ti} designates the original indicators. For the filtered indicators, we consider similar regression equations:

$$GDP_{t+shift} = \beta_0^{shift,h,filt} + \sum_{i=1}^{5} \beta_i^{shift,h,filt} y_{ti}^{h,filt} + \epsilon_t^{h,shift,filt} = G\hat{D}P_{t+shift}^{h,filt} + \epsilon_t^{h,shift,filt},$$

where $y_{ti}^{h,filt}$ represents the i-th filtered indicator optimized at forecast horizon h, based on either HP-C, M-MSE, or M-SSA.
Remarks:

- I'm computing all combinations of shift and h.
- But we could emphasize the diagonal only (h=shift): simplifies notation/explanation.
- The table of AUCs (below) is based on diagonal only

For simplicity of exposition, we present results based on the full sample, excluding the pandemic period. Fig.26 compares ROCs of the predictors for small and mid-term forecast horizons and Table **??** presents AUCs (area under the curve) for forecast horizons 1-5 quarters ahead.

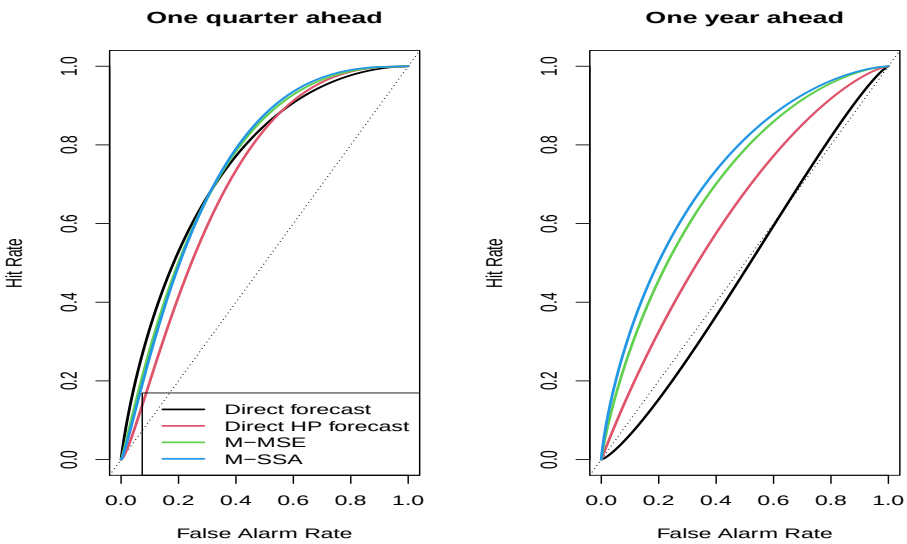

Figure 26: Hit rate vs. false alarm rate of the predictors at small (one quarter ahead:left panel) and yearly forecast horizons (right panel).

At forecast horizons less than or equal to two quarters, the direct forecast challenges the filters. However, for horizons $h > 2$, the filters outperform the direct approach: the univariate

	Direct forecast	Direct HP forecast	M-MSE	M-SSA
Forecast horizon 1	0.752	0.715	0.750	0.750
Forecast horizon 2	0.697	0.663	0.714	0.740
Forecast horizon 3	0.554	0.622	0.666	0.710
Forecast horizon 4	0.486	0.624	0.707	0.731
Forecast horizon 5	0.593	0.636	0.660	0.710

Table 8: AUCs of predictors against GDP at various forecast horizons.

HP-C design is surpassed by the classic multivariate MSE filter, which is in turn dominated by the M-SSA. These results confirm that controlling the rate of sign changes of the predictor, while maximizing its target correlation through M-SSA can enhance predictor performance in terms of sign accuracy and false alarm rate.

8 Summary and Conclusion

Classic direct forecasts of BIP can outperform simple benchmarks, such as the mean, up to two quarters ahead. However, this limited forecast horizon is often insufficient for certain applications. A key challenge with direct forecasts is that the relevant indicators tend to be noisy, which can mask the true signal and lead to overfitting. In particular, the predictor struggles to timely track dips (slowdowns, recessions) or peaks (recoveries, expansions). To address this issue, we apply a HP(160)-filter to smooth the data, with the smoothing parameter $\lambda = 160$ reflecting a more adaptive design than the classic quarterly HP, in line with an intended forecast horizon of up to one year ahead. The resulting predictor tends to outperform the classic direct forecasts in terms of statistical significance, mainly due to a slight left-shift noticeable at sharp recession dips. However, despite the filtering, the predictor remains somewhat noisy, potentially producing excessive noisy zero-crossings at the transitions between recessions and expansions. This issue is partly due to noise leakage inherent in the one-sided HP filter.

To address these problems, we propose a multivariate extension of SSA based on Wildi (2024). Unlike the univariate HP filter, the M-SSA can leverage cross-sectional information from indicators leading BIP in real time while controlling for the rate of zero-crossings of the predictor. Consequently, the multivariate filter outputs become increasingly left-shifted as the forecast horizon lengthens, allowing dips and peaks of the target series to be tracked more systematically, while reducing the occurrence of noisy zero-crossings. However, since M-SSA does not explicitly target BIP, an additional step is necessary to derive the final BIP predictor. We propose two approaches: a simple equally-weighted aggregate, assuming all M-SSA components are equally important for predicting BIP, and an optimally weighted predictor based on a regression of the M-SSA components on future BIP.

For illustration, we consider a simple approach based on regressing the single M-SSA BIP output on future BIP, thereby ignoring the other filter outputs. This predictor is intuitively appealing because the future HP-BIP—i.e., the target of M-SSA—is the low-frequency component of the future BIP target. Consequently, a strong link between M-SSA and future HP-BIP also indicates a connection with BIP, even if the statistical significance is obscured by the noise inherent in BIP. Out-of-sample performance suggests that this predictor is statistically significant at horizons up to one year. Moreover, designs optimized for larger forecast horizons outperform the mean benchmark at all considered forecast horizons, and the predictor exceeds the performance of direct forecasts—based on unfiltered or HP filtered indicators—at horizons longer than two quarters. An analysis of revision errors arising from quarterly updates of the proposed M-SSA BIP component predictor indicates that the regression parameters, the VAR model, and the resulting M-SSA filters stabilize after an initial burn-in period. This stability enhances the explainability and interpretability of the predictor.

In conclusion, the new predictor design integrates the traditional direct forecast approach with a novel multivariate filter to target BIP up to one year ahead. This multivariate approach leverages a set of economic indicators leading BIP in real time while controlling the smoothness of the predictor through the zero-crossing rate.

9 Appendix

9.1 Performances: Without Pandemic

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.077	0.049	0.017	0.004	0.002	0.001	0.079
Shift=1	0.249	0.186	0.108	0.039	0.004	0.001	0.008
Shift=2	0.805	0.694	0.436	0.181	0.029	0.003	0.005
Shift=3	0.905	0.949	0.878	0.585	0.187	0.034	0.023
Shift=4	0.722	0.944	0.984	0.900	0.471	0.075	0.015
Shift=5	0.591	0.892	0.995	0.958	0.588	0.232	0.031

Table 9: HAC adjusted p-values of regressions of M-SSA component predictors optimized for forecast horizons $h = 0, ..., 6$ (the columns) on BIP shifted forward by shift=0,...,5 (the rows). Out-of-sample span starting in Jan-2007, without pandemic.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.981	0.947	0.888	0.837	0.841	0.897	0.967
Shift=1	1.087	1.071	1.012	0.926	0.866	0.870	0.919
Shift=2	1.075	1.089	1.074	1.010	0.936	0.897	0.904
Shift=3	1.036	1.069	1.090	1.059	0.991	0.943	0.931
Shift=4	1.012	1.058	1.101	1.096	1.036	0.966	0.927
Shift=5	1.015	1.064	1.107	1.114	1.078	1.015	0.951

Table 10: rRMSEs of M-SSA components predictor benchmarked against the expanding mean of BIP. Out-of-sample span starting in Jan-2007, without pandemic.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.138	1.098	1.029	0.971	0.975	1.040	1.121
Shift=1	1.297	1.277	1.207	1.104	1.034	1.038	1.096
Shift=2	1.180	1.195	1.179	1.108	1.027	0.984	0.992
Shift=3	1.028	1.061	1.082	1.051	0.984	0.936	0.924
Shift=4	1.011	1.058	1.101	1.095	1.035	0.965	0.926
Shift=5	0.987	1.034	1.076	1.084	1.048	0.987	0.925

Table 11: rRMSEs of M-SSA components predictor benchmarked against the direct forecasts. Out-of-sample span starting in Jan-2007, without pandemic.

	Shift=0	Shift=1	Shift=2	Shift=3	Shift=4	Shift=5
RRMSE	0.862	0.838	0.912	1.008	1.001	1.028

Table 12: rRMSEs direct forecasts benchmarked against the expanding mean. Out-of-sample span starting in Jan-2007, without pandemic.

9.2 Performances: Including Singular Pandemic Data

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.087	0.046	0.011	0.001	0.000	0.002	0.011
Shift=1	0.414	0.303	0.188	0.098	0.055	0.063	0.125
Shift=2	0.736	0.614	0.364	0.111	0.017	0.008	0.028
Shift=3	0.947	0.949	0.872	0.568	0.154	0.026	0.009
Shift=4	0.572	0.874	0.957	0.894	0.555	0.210	0.108
Shift=5	0.506	0.869	0.971	0.927	0.597	0.232	0.037

Table 13: HAC adjusted p-values of regressions of M-SSA component predictors optimized for forecast horizons $h = 0, ..., 6$ (the columns) on BIP shifted forward by shift=0,...,5 (the rows). Out-of-sample span starting in Jan-2007, including the pandemic.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.001	0.982	0.952	0.924	0.912	0.918	0.948
Shift=1	1.057	1.050	1.028	0.997	0.980	0.984	0.989
Shift=2	1.031	1.036	1.027	0.995	0.962	0.951	0.968
Shift=3	1.022	1.035	1.043	1.025	0.991	0.962	0.948
Shift=4	1.002	1.023	1.045	1.046	1.024	1.001	0.989
Shift=5	1.009	1.032	1.054	1.060	1.043	1.013	0.981

Table 14: rRMSEs of M-SSA components predictor benchmarked against the expanding mean of BIP. Out-of-sample span starting in Jan-2007, including the pandemic.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.233	1.210	1.173	1.138	1.123	1.131	1.167
Shift=1	0.983	0.977	0.956	0.928	0.912	0.916	0.920
Shift=2	1.040	1.044	1.035	1.004	0.970	0.959	0.976
Shift=3	1.045	1.059	1.067	1.049	1.014	0.984	0.971
Shift=4	0.981	1.002	1.023	1.024	1.003	0.980	0.968
Shift=5	0.979	1.001	1.023	1.028	1.012	0.983	0.951

Table 15: rRMSEs of M-SSA components predictor benchmarked against the direct forecasts. Out-of-sample span starting in Jan-2007, including the pandemic.

	Shift=0	Shift=1	Shift=2	Shift=3	Shift=4	Shift=5
RRMSE	0.812	1.075	0.992	0.977	1.021	1.031

Table 16: rRMSEs direct forecasts benchmarked against the expanding mean. Out-of-sample span starting in Jan-2007, including the pandemic.