Heinisch, Katja; van Norden, Simon; Wildi, Marc

Working Paper

Smooth and persistent forecasts of German GDP: Balancing accuracy and stability

IWH Discussion Papers, No. 1/2026

Provided in Cooperation with:
Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

Suggested Citation: Heinisch, Katja; van Norden, Simon; Wildi, Marc (2026) : Smooth and persistent forecasts of German GDP: Balancing accuracy and stability, IWH Discussion Papers, No. 1/2026, Halle Institute for Economic Research (IWH), Halle (Saale), https://doi.org/10.18717/dp99kr-7336

This Version is available at:
https://hdl.handle.net/10419/335681

IWH
Halle Institute for Economic Research
Member of the Leibniz Association

Discussion Papers

No. 1
January 2026

Smooth and Persistent Forecasts of German GDP: Balancing Accuracy and Stability

Katja Heinisch, Simon van Norden, Marc Wildi

https://doi.org/10.18717/dp99kr-7336

Authors

Katja Heinisch
Halle Institute for Economic Research (IWH) –
Member of the Leibniz Association,
Department of Macroeconomics
E-mail: katja.heinisch@iwh-halle.de
Tel +49 345 7753 836

Simon van Norden
HEC Montréal and CIREQ

Marc Wildi
Zurich University of Applied Sciences (ZHAW)

IWH Discussion Papers are indexed in RePEc-EconPapers and in ECONIS.

Editor

Halle Institute for Economic Research (IWH) –
Member of the Leibniz Association

Address: Kleine Maerkerstrasse 8
D-06108 Halle (Saale), Germany
Postal Address: P.O. Box 11 03 61
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60
Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188

Smooth and Persistent Forecasts of German GDP: Balancing Accuracy and Stability*

Abstract

Forecasts that minimize mean squared forecast error (MSE) often exhibit excessive volatility, limiting their practical applicability. We address this accuracy-smoothness trade-off by introducing a Multivariate Smooth Sign Accuracy (M-SSA) framework, which extracts smoothed components from leading indicators to enhance the signal-to-noise ratio and control the forecast volatility and timing. Applied to quarterly German GDP growth, our method yields smoothed forecasts that can improve forecasting accuracy, particularly over medium-term horizons. We find that while smoother forecasts tend to lag slightly around turning points, this can be offset by adjusting the forecast horizon. These findings highlight the practicality of the M-SSA framework for both forecasters and policymakers, especially in settings where forecast revisions or policy adjustments are costly.

Keywords: forecast smoothing, Smooth Sign Accuracy, time-series filtering

JEL classification: C53, E37, E66

* We would like to thank Barbara Rossi and the participants at the Vienna Forecasting Workshop 2025, at the Conference on Real-Time Data Analysis, Methods, and Applications in Macroeconomics 2025 as well as at the IWH-CIREQ-GW-BOKERI Macroeconometric Workshop 2025 for their comments and suggestions.

1 Motivation

> "As we parse the incoming information, we are focused on separating the signal from the noise as the outlook evolves."
> *Speech by Federal Reserve Board Chair Jerome H. Powell, March 07, 2025, University of Chicago Booth School of Business*

Economic forecasters face a dilemma: while informational efficiency requires that forecasts incorporate all available data, doing so often results in forecasts that are excessively volatile. This may reflect excessive "noise" due to overfitting (e.g. see Hastie et al., 2009). At the same time, forecast users (including public and private decision makers) may face constraints in their speed of adjustment to changing forecasts.[1]

Recent work by Wildi (2024, 2025) and McElroy and Wildi (2019, 2020) formalizes this problem and proposes a novel method that optimally smooths forecasts by controlling the expected frequency of sign changes.[2] Specifically, the Smooth Sign-Accuracy (SSA) method maximizes the correlation between the forecast target and the predictor subject to a constraint on the rate of sign changes of the predictor. In particular, SSA encompasses the classic mean-squared error (MSE) as a special case and yields increasingly smooth forecasts as sign changes are penalized more heavily. Applications of SSA to real-time business-cycle analysis for a wide range of countries with sufficiently long histories of officially dated recessions are provided in Wildi (2024).

This study extends the SSA-framework in two ways: First, we pre-smooth and filter leading indicators with the aim of improving their signal-to-noise ratio.[3] Second, we generalize SSA to a multivariate context (M-SSA), allowing us to exploit cross-sectional information among leading indicators to enhance forecast performance. We apply this framework to well-established German economic indicators to show its ability to improve forecasts of GDP growth (Lehmann and Reif, 2021; Heinisch and Scheufele, 2018) and GDP trends using a variety of sample periods and performance metrics. Previous research highlights the utility of direct forecasts, particularly for short-term predictions such as nowcasting and forecasts one quarter ahead. We find that filtered components obtained via M-SSA can provide better forecast performance. Further, we find that mitigating noise by smoothing indicators provides useful information for predicting GDP. Specifically, we show that increasing the forecast horizon while fixing the forecast's smoothness reduces the root mean squared error (RMSE) of quarterly GDP forecasts by roughly 5 to 15 percentage points at forecast horizons of two to five quarters. The inherent smoothness of M-SSA also mitigates the false-alarm rate when predicting the sign of future GDP growth, yielding consistent performance up to five quarters ahead and outperforming a range of benchmarks, including direct projections, univariate filters, and the classic multivariate Wiener-Kolmogorov (WK) method.

We show that pre-filtering indicators with univariate filters improves upon the classic (unfiltered) direct forecasts, particularly at horizons beyond two quarters. Conversely, classic direct forecasts can

[1] For example, fixed annual budgets may constrain spending, or transaction costs may limit portfolio adjustments. Others may attempt to smooth their reactions for reasons of signalling, such as dividend smoothing by corporate managers and interest rate smoothing by central banks.

[2] A stationary zero-mean (centered) predictor is assumed in Wildi (2024), while Wildi (2025) generalizes the concept to integrated processes.

[3] We conjecture that economic indicators are contaminated by unpredictable high-frequency noise, thereby obscuring the effective 'signal'. If so, it may be possible to dampen the noise and highlight the signal by filtering our indicators.

challenge the filtered designs at short and very short horizons, suggesting that the exclusion of high-frequency content is unnecessary or counterproductive for nowcasting. Our results also confirm the usefulness of some of the leading indicators for GDP forecasting.

We proceed as follows. Section 2 describes our data, the sample period used, and documents their dynamic behaviour. Section 3 provides benchmark forecasts using simple direct projections of GDP based on selected indicators. We demonstrate how in-sample forecast performance may improve when we "de-noise" the indicators prior to projection with simple smoothing filters. Section 4 defines the M-SSA method as an extension and generalization of the classic multivariate WK approach. Section 5 compares the forecast performance of M-SSA to that of direct projections for various forecast horizons, while Section 6 examines robustness and stability. Section 7 summarizes our results and concludes. Additional Figures and Tables are presented in the Appendix.

2 Data and Dependence

While our approach could be extended to a higher-dimensional framework – encompassing multiple indicators with potentially varying sampling frequencies – we propose here a simple, illustrative design based on a selection of well-established indicators to highlight key methodological aspects of the novel approach.

2.1 Data and Publication Lags

Our analysis employs a set of well-established leading indicators to forecast quarterly German GDP (Drechsel and Scheufele, 2012; Heinisch and Scheufele, 2018), including industrial production (IP), the ifo Business Climate Index (ifo_c), the Economic Sentiment Indicator (ESI), and the term spread between 10-year bond yields and the 3-month EURIBOR rate (spre_10y_3m) covering the period 1991 to 2024. We take into account the publication lag of the indicators and the GDP series (Table 1).[4] While GDP is available with a one-quarter delay, and IP is released with a two-month lag, business surveys and financial data are published contemporaneously.

Table 1: Ragged edge structure.

	GDP	IP	ifo_c	ESI	spr_10y_3m
2024-07-01		90.700	85.800	92.500	-1.300
2024-08-01		93.200	84.900	90.900	-1.400
2024-09-01	902.571	91.200	83.700	89.800	-1.400
2024-10-01		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

Note: Information set available to forecasters as of January 2025, before publication of the GDP flash estimate.

[4] Data revisions are not considered, given that previous studies (Heinisch and Scheufele, 2019) have shown that their overall effect on forecast performance is minor.

To allow for missing observations due to publication lags, we re-align series to reflect the information available to forecasters. Table 2 displays the artificial result as of the end of 2024, with GDP and IP columns shifted to reflect their release dates rather than the periods that they measure. This data configuration will be the used for nowcasting or forecasting GDP. For illustrative purposes, an additional target column is included on the right side of the table, representing the dependent variable for nowcasting applications.[5]

To ensure comparability and mitigate scale effects, all indicators (except the term spread, which is differenced only) are log-differenced and standardized. Additionally, extreme outliers associated with the COVID-19 pandemic are trimmed to prevent distortions in subsequent figures. Monthly series are aligned (Table 2) and averaged to quarterly series (Figure 1a), while Figure 1b offers a close-up view of the financial crisis and the pandemic. The latter figure shows that the real-time GDP and industrial production series are synchronized at turning points, whereas the other indicators tend to lead somewhat. The challenge for forecasters is to use the latter in a multivariate approach to improve forecasting performance over that of univariate benchmarks.

Table 2: Aligned dataset

	GDP	IP	ifo_c	ESI	spr_10y_3m		Target (nowcast)
2024-07-01		93.400	85.800	92.500	-1.300	—	
2024-08-01		90.700	84.900	90.900	-1.400	—	
2024-09-01	901.623	93.200	83.700	89.800	-1.400	—	902.571
2024-10-01		91.200	84.100	90.700	-1.200	—	
2024-11-01		90.800	83.800	89.300	-0.900	—	
2024-12-01	902.571	92.200	82.800	86.900	-0.800	—	?

Note: Indicators realigned to reflect real-time availability. The **GDP** column is shifted by its publication lag. The **Target (nowcast)** is the target variable for a GDP nowcast (i.e. the values of column 1 shifted upwards by the publication lag of GDP).

2.2 Dependence

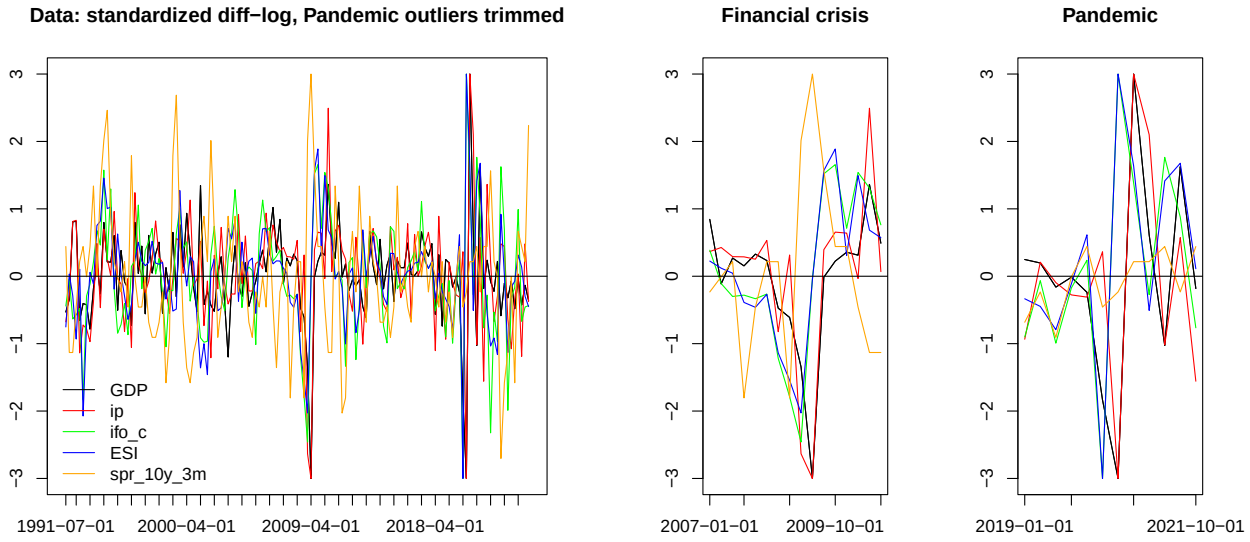
To examine the joint dynamics of our selected variables, we estimate their cross-correlations as well as their vector moving-average (VMA) representations. The latter is a key input to the multivariate filters we develop in Sections 4 and 5.3. The sample cross-correlation functions (CCF's) between current and lagged GDP and the two survey indicators show strong correlations that peak when GDP is slightly lagged.[6] Results for IP showed no such lag, while correlations for the term spread were weaker. However, the precise strengths of the correlations are influenced by the outliers associated with the COVID-19 pandemic. In the interest of model stability, we exclude data from Q4-2019 to Q4-2020 (pandemic) when modelling the dynamics for the multivariate filters developed below. However, some of our empirical results will emphasize crisis episodes, including the pandemic.

We then estimate a variety of parsimonious VARMA models to capture these dynamics, selecting parsimonious VAR(1) and VAR(3) specifications fitted to the transformed (diff-log) quarterly data.[7]

[5] In the case of an h-step ahead forecast, this column is shifted upwards by h quarters.

[6] See Figures A1a (full sample) and A1b (omitting the Pandemic period) in the Appendix.

[7] We used the `refVARMA()` procedure in the Tsay et al. (2022) `MTS` package, which sets VARMA coefficients with t-statistics below a critical threshold to zero (see Tsay, 2013). This led to a VAR(1) specification with a threshold



(a) Full Sample. (b) Leads & lags during financial crisis & Pandemic.

Figure 1: Quarterly German GDP and Indicators.

Note: standardized log-differences, trimmed to ±3 standard deviations to facilitate visual inspection.

The implied responses of GDP to shocks to each indicator variable are indicating that survey indicators have the strongest and most persistent impact on GDP, with pandemic data shortening the apparent lead time.[8]

GDP exhibits the strongest response to shocks in the two survey indicators—ifo and ESI.[9] The impact of the pandemic modifies the lag structure by emphasizing shorter dependencies triggered by the alternating outliers during that specific episode. Removing the pandemic from the analysis results in smoother, somewhat longer-tailed dependencies, consistent with the observed empirical autocorrelation functions. In particular, the VAR(3) model captures slightly longer lead times and more nuanced lag patterns.[10]

The delayed response of GDP to movements in the survey indicators suggests that multivariate models could yield better forecasts than those of univariate models, a hypothesis we explore in the next section.

of 1.5. The VAR(1) left some residual evidence of higher-order dynamics, but capturing these in our limited sample size was difficult without overfitting. We therefore explored VAR(p) models with coefficients estimated by Elastic Net using the Zou and Hastie (2020) `elasticnet` package. This led to a VAR(3) specification with 7 steps. For additional details, see Zou and Hastie (2020) and the references therein.

[8] See Figures A2a and A2b in the Appendix.

[9] There is also weaker evidence of a response to the term spread at longer lags.

[10] The VAR(3) model yields marginal improvements in multivariate filtering performance. However, due to its relative simplicity, this study primarily employs the VAR(1) model. The M-SSA package (accessible at `https://github.com/wiaidp/R-package-SSA-Predictor`) offers more advanced capabilities, including the implementation of the VAR(3) model and Bayesian VAR specifications.

3 Baseline Forecasting Approaches

In this section, we evaluate the performance of simple indicator-based forecasts of German GDP growth and examine how pre-filtering can improve their accuracy by removing noise. Our starting point is a direct in-sample forecasting framework in which GDP growth at horizon h is regressed on a set of leading indicators and lagged GDP, estimated over the pre-pandemic period.

$$u_{t+h} = \boldsymbol{\beta'} \cdot \mathbf{X}_t + \varepsilon_t \tag{1}$$

where

- u_{t+h} is the quarterly growth rate of GDP for forecast horizon $h = \{0, 1, 2, ...\}$ (labeled shifted GDP in the following)
- $\mathbf{X}_t \equiv [1, u_{t-1}, ip_t, ifo\_c_t, ESI_t, spr\_10y\_3m_t]'$ or their transformed counterparts, and u_{t-1} accounts for the publication lag of GDP.[11]
- $\hat{\boldsymbol{\beta}}$ is estimated by least-squares, omitting observations affected by the pandemic, i.e. Q4-2019 to Q4-2020.

We first estimate forecasts using unfiltered indicators (Figure 2), which reflect the raw information available to forecasters in real time. We then apply a one-sided concurrent smoothing filter to the same indicators and demonstrate how in-sample forecast performance improves when we "de-noise" the indicators before projection with simple smoothing filters.[12] Comparing these two designs allows us to isolate the effect of pre-filtering in a straightforward, transparent way.

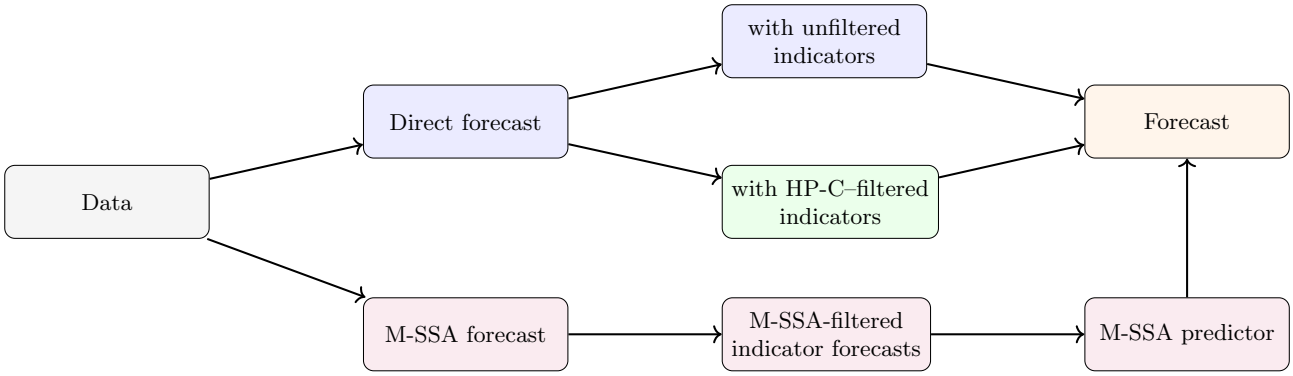


Figure 2: Forecasting strategies with unfiltered, HP-C–filtered, and M-SSA predictors

[11] Including lagged GDP enables the nesting of the AR forecast within the broader modeling framework.

[12] The smoothing filters are one-sided (concurrent) filters of length 31; therefore, estimation of both forecasting models omits the first 30 observations.

3.1 Direct Forecasts with Unfiltered Indicators

As a benchmark, we estimate direct projections of GDP growth at horizons $h = 0$ to $h = 6$ quarters ahead using all indicators without filtering. Figures 3a and 3b compare actual and fitted values for GDP at forecast horizons $h = 0, \ldots, 3$, using the four indicators plus lagged GDP. At short horizons, forecasts track turning points relatively well. As the forecast horizon increases, the forecasts become increasingly damped (zero-shrinkage) and delayed, with peaks and troughs progressively shifted to the right relative to the target series GDP.

Table 3: Statistical Significance of Direct Forecasts

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

Note: p-values for $H_0 : \boldsymbol{\beta} = 0$ in eq. (1) based on HAC-adjusted Wald test.
Estimation over full sample (without pandemic).

Table 3 reports p-values for tests of the joint null hypothesis that all coefficients on the indicators are zero (HAC-adjusted Wald test). Additionally, these results confirm that the predictive power of the unfiltered indicators declines sharply beyond $h = 2$, with significance levels falling to insignificance for horizons beyond two quarters.This finding motivates the use of pre-filtering to extract a clearer signal from the indicator set.

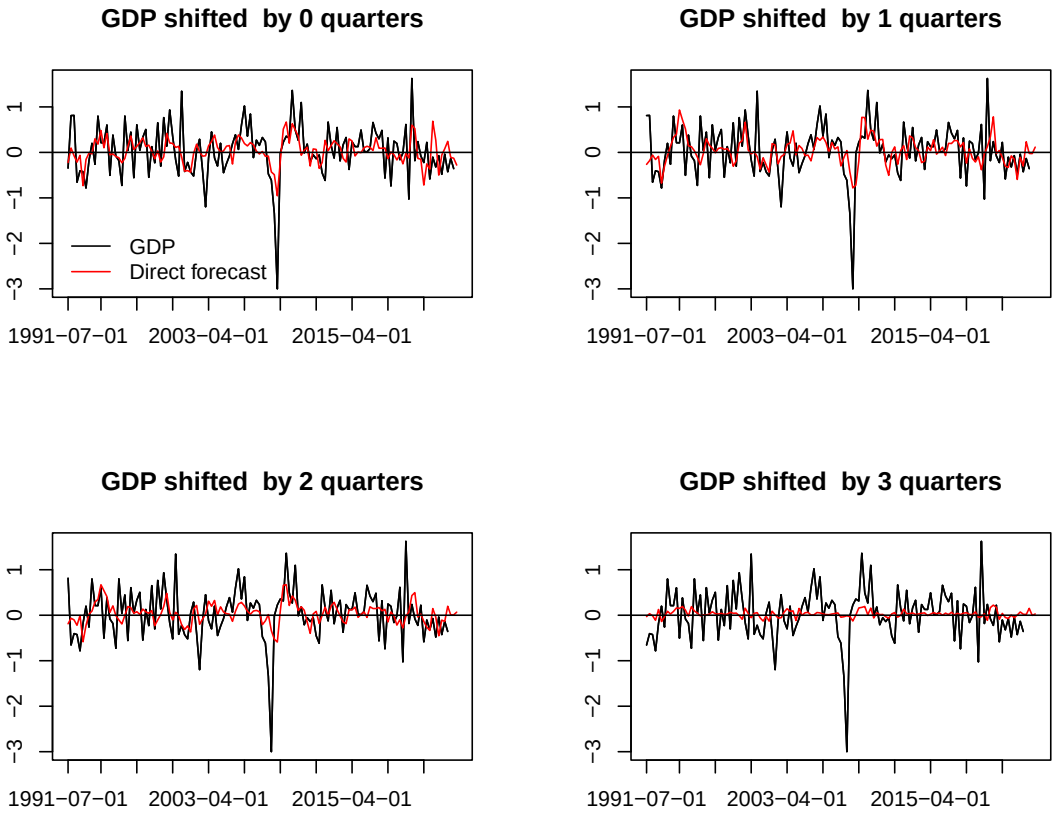
3.2 Direct Forecasts with HP-C-Filtered Indicators

We conjecture that economic indicators may be contaminated by unpredictable high-frequency noise, thereby obscuring the effective 'signal' and making a direct regression more susceptible to overfitting. If so, it may be possible to highlight the signal by filtering our indicators to dampen high-frequency noise.

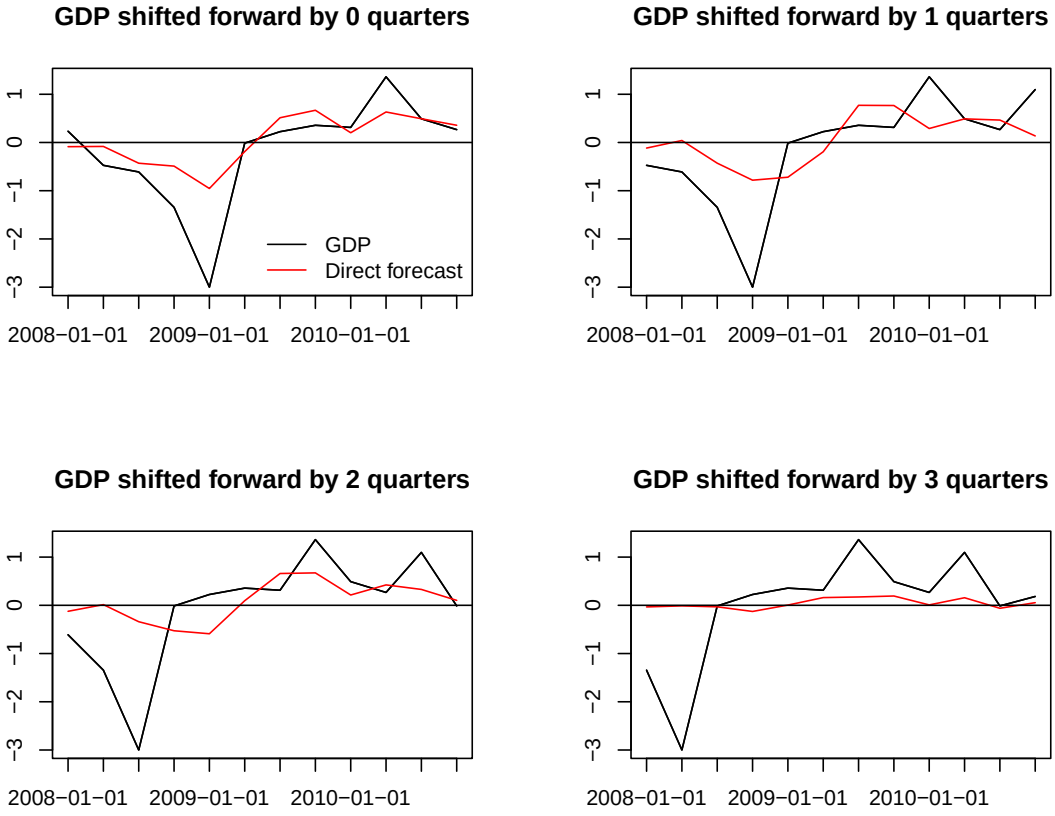
The Hodrick-Prescott (HP) filter is a classic tool used in business cycle analysis (Hodrick and Prescott, 1997), characterized by a single smoothing parameter, λ. While the conventional value $\lambda = 1600$ is widely applied to quarterly output levels, Phillips and Jin (2021) caution that this setting induces excessive smoothing and thereby removes economically relevant information. Given that our study focuses on quarterly growth rates, a lower smoothing parameter is more appropriate.

Accordingly, we instead use $\lambda = 160$ to allow for a more adaptive filter. This choice also addresses the critique by Ravn and Uhlig (2002), who emphasize that inappropriate calibrations of λ can distort cyclical dynamics. Hence, our parameter selection is sufficiently low to avoid eliminating relevant medium-term fluctuations, yet high enough to dampen volatility driven by measurement error and shocks.[13] Comparing classic two-sided and one-sided HP(160) (denoted HP-C for *concurrent*) filters, we find that the HP-C filter assigns greater weight to yearly components—relevant in a medium-term

[13] An analysis of forecast performance based on the M-SSA package (not shown) suggests that both the rRMSE as well as the hit rate progressively degrade for small ($\lambda < 16$) or large ($\lambda > 1600$) values of λ, and that this degradation occurs fairly evenly across forecast horizons. See (`https://github.com/wiaidp/R-package-SSA-Predictor`).



(a) Full sample information (without pandemic)



(b) Financial crisis

Figure 3: Direct forecasts of GDP (red lines) using unfiltered indicators and target (black line).

Note: GDP left-shifted by 0 - 3 quarters.

forecast context—than the classic quarterly HP filter, while also reducing undesirable high-frequency noise, see Figure 5 (bottom right panel).

Figure 4 provides results comparable to those of 3 but now using indicators smoothed with HP-C. The fitted values lag GDP slightly less and track turning points somewhat more quickly. Table 3,reporting the p-values for tests of the joint null hypothesis that all coefficients $\boldsymbol{\beta} = 0$, also shows that filtering improves the statistical significance of the indicators for horizons beyond two quarters, suggesting that much of the predictive content of the raw series is concentrated in their lower-frequency components.

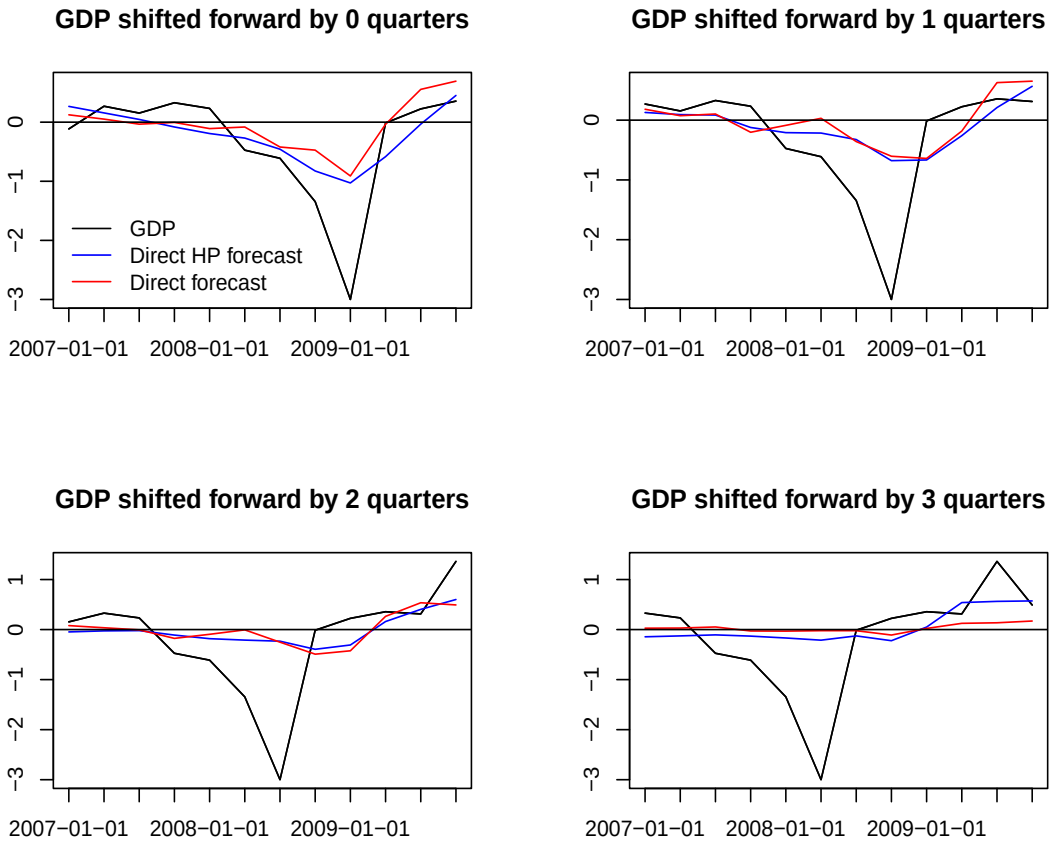


Figure 4: Forecasts During the Financial Crisis: HP-C vs. Unfiltered Indicators.
Note: Shifted GDP (black), forecasts based on HP-C filtered indicators (blue) and unfiltered indicators (red) during the financial crisis. Regression based on full sample information (without pandemic).

The application of the one-sided HP-C filter to the indicators yields some measurable improvements in forecast performance. However, the noise suppression capability of the one-sided filter is compromised, as explained by Wildi (2025) and illustrated by the high-frequency leakage of the amplitude function. This leakage is evident in Figure 5: the amplitude function of the one-sided filter (bottom-right panel) lies markedly above the zero line, towards higher frequencies, in contrast to the two-sided filter (bottom-left panel). To address these limitations, we turn to the Multivariate Smooth Sign Accuracy (M-SSA) framework proposed by Wildi (2025).

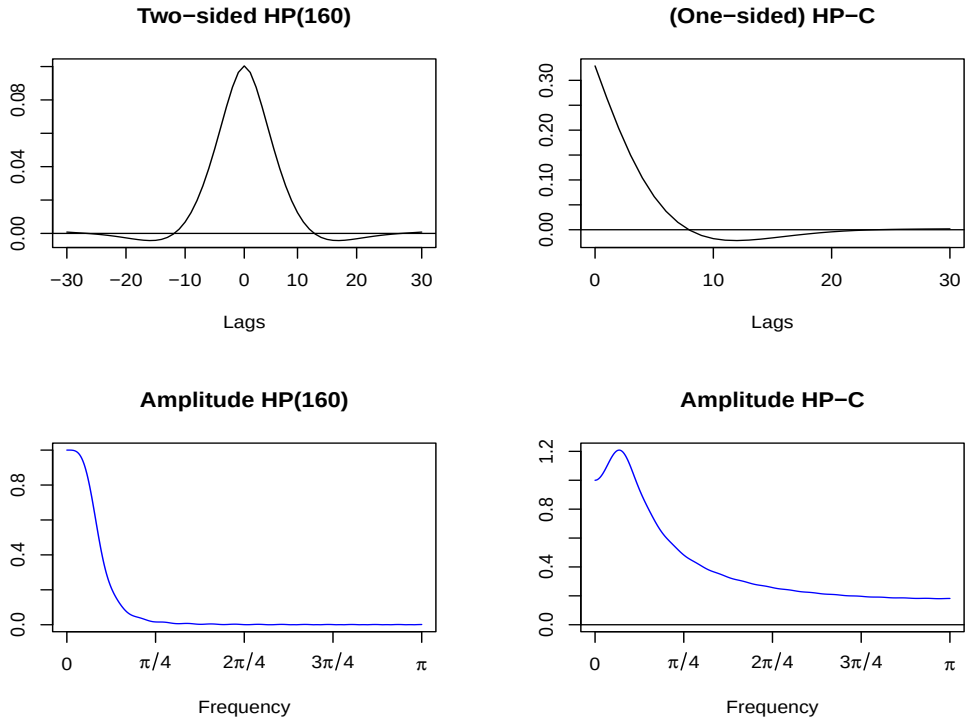


Figure 5: HP($\lambda = 160$) filters and Frequency Response.

Note: Left: HP (Two-sided, symmetric), Right: HP-C (One-sided, concurrent)
Top Row: Filter Weights, Bottom Row: Amplitude Functions.

4 Methodology: The M-SSA Framework

While pre-filtering improves the performance of direct forecasts by removing high-frequency noise, it does not directly control the smoothness of the resulting forecast series. Excessive sign changes in the forecast growth rate can occur near turning points, leading to spurious recession or recovery signals.

The Smooth Sign Accuracy (SSA) framework proposed by Wildi (2024, 2025) addresses this by explicitly controlling the forecast's holding time (HT), i.e. the expected duration between sign changes, while preserving a high correlation with the target series.[14] Hence, the SSA framework nests the classical mean squared error (MSE) predictor as a special case, corresponding to a specific choice of the smoothness (holding time) constraint. The Multivariate Smooth Sign Accuracy (M-SSA) framework extends SSA to multiple predictors.

4.1 Smooth Sign Accuracy (SSA)

Let $\mathbf{x}_t$ be a time series representing the data, with observations $(x_t, ..., x_1)'$. Let $\mathbf{z}_t = (z_t, ..., z_1)'$ denote a target series, which depends on $\mathbf{x}_t$ as well as on unknown future x_{t+k}, $k > 0$ (and possibly unknown past x_k, $k \leq 0$). For simplicity, we assume stationarity of all time series involved. In our current framework, z_t corresponds to the output of a two-sided Hodrick–Prescott filter with smoothing parameter $\lambda = 160$ applied to the doubly infinite sequence x_k, $k \in \mathbb{Z}$. The finite sample $\mathbf{x}_t$ is the portion effectively observed and represents the log-differenced GDP (the dependence on $x_0, x_{-1}, ...$

[14] See Wildi (2024) for further details on the SSA framework and its applications.

is neglected because the HP filter weights decay rapidly to zero; see Figure 5). The prediction task comprises tracking z_t using the (causal) predictor $y_t = \sum_{k=0}^{L-1} b_k x_{t-k} = \mathbf{b}'\mathbf{x}_t$, where the filter $\mathbf{b}$ with weights $b_0, ..., b'_{L-1}$ (actually of length $L = t$) is available for optimization purposes.

The task of tracking a target 'optimally' can be formalized in different ways; here, we focus on the correlation $\rho(z, y, h)$ between y_t and target z_{t+h}, where $h \geq 0$ denotes the forecast horizon (backcasting with $h < 0$ is ignored in this context). For simplicity, assume a fixed horizon, say $h = h_0$, so that we may drop the reference to h in our notation. In the univariate case, Wildi (2025) proposes the Smooth Sign Accuracy (SSA) as an optimization criterion for determining $\mathbf{b}$:

$$\left.\begin{array}{c} \max_{\mathbf{b}} \ \rho(z, y) \\ s.t. \quad \rho(y, 1) = \rho_1 \end{array}\right\} \tag{2}$$

where $\rho(y, 1)$ is the first-order autocorrelation of the predictor y_t. The parameter ρ_1 controls the smoothness of the predictor: higher values of ρ_1 favour smoother trajectories of y_t. Assuming that the process x_t has mean zero, Wildi (2024) establishes a connection between $\rho(y, 1)$ and the expected duration between consecutive sign changes (or zero-crossings) of y_t, the 'holding time', denoted by $HT(y)$[15]:

$$HT(y) = \frac{\pi}{\arccos(\rho(y, 1))}. \tag{3}$$

Since eq. (3) is strictly monotonic, the SSA criterion can be reformulated as the optimization problem:

$$\left.\begin{array}{c} \max_{\mathbf{b}} \ \rho(z, y) \\ \frac{\pi}{\arccos(\rho(y,1))} = HT_1 \end{array}\right\}, \tag{4}$$

where the smoothness parameter HT_1 expresses the desired duration between consecutive sign changes of the predictor.

One implication of criterion (4) is that the objective function as well as the constraint are 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 (MSE) can 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 $1/HT_1 \equiv 1/HT_{MSE}$ in the constraint, where HT_{MSE} represents the holding time of $y_{t,MSE}$.

Traditional predictors often yield an excessive number of false sign changes due to leakage of high-frequency components (high-frequency leakage) in the predicted series, as discussed above (see Figure 5). This noisiness is particularly problematic near the onset or termination of economic crises (recessions), where the target fluctuates around zero; reliable tracking of sign changes would facilitate

[15] Of course, when z_t is GDP growth, sign changes are indicative of expansions and recessions. Formally, the precise relationship between the expected duration between sign changes and the lag-one autocorrelation of the predictor is derived under the assumption of Gaussian time series. However, Wildi (2024) demonstrates that this relationship remains robust even when deviations from Gaussianity occur (central limit theorem applied to lowpass-filtered series).

real-time assessment of the occurrence and timing of such events. Wildi (2025) shows that the SSA criterion helps control this phenomenon (Figure 8).

Using the dual formulation of the criterion (4), the author demonstrates that the SSA solution yields the *lowest* sign change rate among all (linear) predictors with equivalent target correlation. The concept can be extended to a multivariate framework, denoted M-SSA, by generalizing the correlation functions as detailed in the next section. As we shall see, controlling the sign change rate within the M-SSA framework will have a quantifiable effect on the hit rate and false positive rate, thereby explicitly addressing the detection of present (nowcasting) and future (forecasting) economic contractions and expansions.

We now briefly elucidate the SSA optimization criterion outlined in Eq.(2). Consider a simplified univariate framework where $x_t = \epsilon_t$ represents white noise with variance σ^2 and assume, for notational simplicity, that $\sigma^2 = 1$. A generic target specification for z_t is then defined as a linear filter of x_t:

$$z_t = \sum_{k=-\infty}^{\infty} \gamma_k x_{t-k} = \sum_{k=-\infty}^{\infty} \gamma_k \epsilon_{t-k},$$

where it is assumed that $\gamma_k \neq 0$ for some $k < 0$ (implying an acausal filter), and $\sum_{|k|<\infty} \gamma_k^2 < \infty$ (squared summability), ensuring the stationarity of the target z_t. With these definitions, all processes are centered at zero, justifying the use of Equation (4) which links first-order ACF and HT. Extensions to non-stationary integrated processes are discussed in Wildi (2025). Suppose the goal is to predict z_{t+h}, $h \geq 0$. Let $y_{ht} = \sum_{k=0}^{L-1} b_{hk} x_{t-k}$ denote a causal predictor of length L. For notational convenience, we now omit the subscript h from the predictor $\mathbf{b} = \mathbf{b}_h$ and we denote $\boldsymbol{\gamma}_h \equiv (\gamma_h, \gamma_{h+1}, ..., \gamma_{h+L-1})'$. Then, under the postulated white noise assumption, we obtain:

$$\rho(z, y) = \frac{\mathbf{b}'\boldsymbol{\gamma}_h}{\sqrt{\mathbf{b}'\mathbf{b}}\sqrt{\boldsymbol{\gamma}_\infty'\boldsymbol{\gamma}_\infty}} \text{ and } \rho(y, 1) = \frac{\mathbf{b}'\mathbf{M}\mathbf{b}}{\mathbf{b}'\mathbf{b}},$$

where $\boldsymbol{\gamma}_\infty'\boldsymbol{\gamma}_\infty \equiv \sum_{|k|<\infty} \gamma_k^2 < \infty$ and where the so-called autocovariance generating matrix $\mathbf{M}$, of dimension $L \times L$, is defined as:

$$\mathbf{M} = \begin{pmatrix} 0 & 0.5 & 0 & 0 & 0 & ... & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & ... & 0 & 0 & 0 \\ ... & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & ... & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & ... & 0 & 0.5 & 0 \end{pmatrix}.$$

The matrix $\mathbf{M}$ is such that $\mathbf{b}'\mathbf{M}\mathbf{b} = \sum_{k=1}^{L-1} b_{k-1} b_k$, representing the first-order autocovariance of y_t. Using this, the optimization criterion (2) can be reformulated as:

$$\left.\begin{array}{cc} & \max_{\mathbf{b}} \mathbf{b}'\boldsymbol{\gamma}_h \\ \text{s.t.} & \mathbf{b}'\mathbf{M}\mathbf{b} = l\rho_1 \\ & \mathbf{b}'\mathbf{b} = l \end{array}\right\}, \tag{5}$$

where the additional length constraint $\mathbf{b}'\mathbf{b} = l$ serves two purposes: firstly, to ensure the uniqueness of the SSA solution; secondly, to normalize the objective function $\mathbf{b}'\boldsymbol{\gamma}_h$, making it proportional to the

target correlation $\rho(z, y)$, up to a fixed scaling factor $\sqrt{\mathbf{b}'\mathbf{b}}\sqrt{\boldsymbol{\gamma}_\infty'\boldsymbol{\gamma}_\infty} = \sqrt{l}\sqrt{\boldsymbol{\gamma}_\infty'\boldsymbol{\gamma}_\infty}$. For convenience, l is typically set to 1, yielding a standardized y_t. Consequently, the solution to the criterion is invariant to the choice between these two objective formulations. However, Criterion (5) is readily amenable to optimization, as it involves affine and quadratic functions of $\mathbf{b}$ (Wildi, 2024).[16]
Under the above assumptions, the classical MSE predictor of length L can be expressed via orthogonal projection as:

$$y_{MSE,t} = \sum_{k=0}^{L-1} \gamma_{k+h} x_{t-k}$$

which corresponds to setting $b_{MSE,k} = \gamma_{k+h}$, $k = 0, ..., L-1$. In general, to enhance the smoothness of the SSA predictor y_t, we select $\rho_1 > \rho_{MSE}$ in criterion (5), where ρ_{MSE} denotes the first-order ACF of the (benchmark) MSE predictor.[17] Finally, an extension to autocorrelated $x_t \neq \epsilon_t$ is discussed below.

4.2 Multivariate Extension of SSA (M-SSA)

The SSA criterion (5) is now extended to a multivariate setting. Let $\mathbf{X}_t$ (of dimension $t \times n$) denote a set of n explanatory series $\mathbf{x}_{1t}, ..., \mathbf{x}_{nt}$, with observations $\mathbf{x}_{it} = (x_{it}, ..., x_{i1})'$. Let $\mathbf{Z}_t$, of dimension $t \times m$, represent a set of m target series $\mathbf{z}_{it} = (z_{it}, ..., z_{i1})'$, $i = 1, ...m$, which typically lie outside the linear space spanned by $\mathbf{X}_t$ (and are unknown at time t but not independent of $\mathbf{X}_t$). In this framework, $\mathbf{X}_t$ represents a matrix of selected economic indicators for the German economy, with $n = 5$. The targets z_{it}, i=1,...,5, are obtained by applying a two-sided HP(160) filter to the doubly-infinite extensions x_{ik}, $k \in \mathbf{Z}$, of these indicators. In a subsequent step, as delineated in Section 5.3, the (filtered) predictors y_{hit} of z_{it+h}, with $h \geq 0$, will be utilized to construct predictors for the GDP. However, unlike the approach in Section 3, these predictors will be based on multivariate M-SSA filters rather than univariate HP-C designs.

For this purpose, to introduce the M-SSA framework, we begin by assuming that $\mathbf{X}_t = \boldsymbol{\epsilon}_t$ is an n-dimensional white noise sequence with a full-rank variance-covariance matrix $\boldsymbol{\Sigma}$. Consider the generalized target $\mathbf{z}_t$:

$$\mathbf{z}_t = \sum_{|k|<\infty} \boldsymbol{\Gamma}_k \mathbf{x}_{t-k},$$

where $\boldsymbol{\Gamma}_k$ is a sequence of filter matrices of dimension $n \times n$, satisfying the square-summability condition. Each $\boldsymbol{\Gamma}_k$ has entries (γ_{ijk}) for $i, j = 1, \ldots, n$. We focus on the estimation of $\mathbf{z}_{t+h}$, utilizing the $n-$dimensional predictor defined by

$$\mathbf{y}_t = \sum_{k=0}^{L-1} \mathbf{B}_k \mathbf{x}_{t-k},$$

[16] Detailed derivations are obtained in Wildi (2025).
[17] Selecting $\rho_1 < \rho_{MSE}$ would produce an SSA predictor with more frequent sign changes than the MSE benchmark, which is typically undesirable in practical applications.

with $n \times n$-dimensional filter matrices $\mathbf{B}_k$, $k = 0, ..., L-1$ (for notational simplicity we drop the subscript h referring to the forecast horizon in these expressions). Let b_{ijk} denote the entries of the matrix $\mathbf{B}_k$ for $k = 0, \ldots, L-1$. We define several vector notations as follows:

$$\begin{aligned}
\boldsymbol{\epsilon}_{it} = (\epsilon_{it}, \epsilon_{it-1}, ..., \epsilon_{it-(L-1)})' \quad &, \quad \boldsymbol{\epsilon}_{\cdot t} = (\boldsymbol{\epsilon}'_{1t}, ..., \boldsymbol{\epsilon}'_{nt})' \\
\boldsymbol{\gamma}_{ijh} = (\gamma_{ijh}, \gamma_{ijh+1}, ..., \gamma_{ijh+L-1})' \quad &, \quad \boldsymbol{\gamma}_{i\cdot h} = (\boldsymbol{\gamma}'_{i1h}, \boldsymbol{\gamma}'_{i2h}, ..., \boldsymbol{\gamma}'_{inh})' \\
\mathbf{b}_{ij} = (b_{ij0}, b_{ij1}, ..., b_{ijL-1})' \quad &, \quad \mathbf{b}_i = (\mathbf{b}'_{i1}, \mathbf{b}'_{i2}, ..., \mathbf{b}'_{in})' \\
\boldsymbol{\gamma}_h = (\boldsymbol{\gamma}_{1\cdot h}, \boldsymbol{\gamma}_{2\cdot h}, ..., \boldsymbol{\gamma}_{n\cdot h}) \quad &, \quad \mathbf{b} = (\mathbf{b}_1, ..., \mathbf{b}_n).
\end{aligned}$$

Here, $\boldsymbol{\gamma}_h$ (or $\mathbf{b}$) is a matrix of dimension $(L \cdot n) \times n$. The i-th column, $i = 1, ..., n$, contains the filter weights associated with the i-th target (or i-th predictor), and the filter weights corresponding to the $j = 1, ..., n$ series (indicators) are stacked into this column vector, of total length $n \cdot L$. The i-th column vector of $\boldsymbol{\gamma}_h$ (or $\mathbf{b}$) can be applied to the $L \cdot n$-dimensional $\boldsymbol{\epsilon}_{\cdot t}$, which stacks all n (artificial) 'indicators' into a single long data vector. Specifically, we define

$$y_{ijt} \equiv \mathbf{b}'_{ij} \boldsymbol{\epsilon}_{jt},$$

which allows us to express the i-th predictor as:

$$y_{it} = \mathbf{b}'_i \boldsymbol{\epsilon}_{\cdot t} = \sum_{j=1}^{n} y_{ijt}.$$

Similarly, the multivariate predictor vector can be written as:

$$\mathbf{y}_t = \mathbf{b}' \boldsymbol{\epsilon}_{\cdot t}.$$

Also, under the white noise assumption, the orthogonal projection leads to the MSE predictor expressed as:

$$\hat{\mathbf{y}}_{t,MSE} = \boldsymbol{\gamma}'_h \boldsymbol{\epsilon}_{\cdot t}.$$

Note that we here assume that the MSE predictor relies on sub-filters of length L for each subseries $\mathbf{x}_{it}$. [18] We define the Kronecker products:

$$\tilde{\mathbf{I}} \equiv \boldsymbol{\Sigma} \otimes \mathbf{I}_{L \times L}, \quad \text{and } \tilde{\mathbf{M}} \equiv \boldsymbol{\Sigma} \otimes \mathbf{M},$$

where $\boldsymbol{\Sigma}$ is the variance-covariance matrix of the (white noise) data, $\mathbf{I}_{L \times L}$ is the $L \times L$ identity matrix and $\mathbf{M}$ is the autocovariance generating matrix introduced earlier. Under the simplifying standardized white noise assumption, we can express the following expectations:

$$\begin{aligned}
E[z_{it+h} y_{it}] &= \boldsymbol{\gamma}'_{i\cdot h} \tilde{\mathbf{I}} \mathbf{b}_i \\
E[y_{it}^2] &= \mathbf{b}'_i \tilde{\mathbf{I}} \mathbf{b}_i \\
E[y_{it-1} y_{it}] &= \mathbf{b}'_i \tilde{\mathbf{M}} \mathbf{b}_i.
\end{aligned}$$

[18] Recall from the SSA criterion (5) that it is not necessary to rely on the doubly infinite representations of the target when formulating the optimization criterion, a point that is confirmed below.

Consequently, we propose the following multivariate M-SSA criterion for $i = 1, ..., n$:

$$\begin{aligned} \max_{\mathbf{b}_i} \boldsymbol{\gamma}'_{i\cdot h} \tilde{\mathbf{I}} \mathbf{b}_i \\ \text{s.t. } \mathbf{b}_i' \tilde{\mathbf{M}} \mathbf{b}_i = \rho_i \\ \mathbf{b}_i' \tilde{\mathbf{I}} \mathbf{b}_i = 1, \end{aligned} \tag{6}$$

for $i = 1, ..., n$, where we assume an arbitrary normalization, $l = 1$, in the length constraint $\mathbf{b}_i' \tilde{\mathbf{I}} \mathbf{b}_i = 1$. If explicitly necessary, the arbitrary scaling $l = 1$ can be revised and optimized separately at a later stage, after an optimal 'unity-length' M-SSA predictor has been determined. In analogy to the univariate case, the objective function is proportional to the target correlation $\rho(z_i, y_i)$ between the i-th predictor and the i-th target. Moreover, the criterion involves affine and quadratic forms of $\mathbf{b}_i$, which simplifies the process of numerical optimization. Finally, the HT constraint $\mathbf{b}_i' \tilde{\mathbf{M}} \mathbf{b}_i = \rho_i$, where we used $l = 1$, serves to control the first-order ACF or, equivalently, the HT of the predictor.

Finally, we can relax the white noise hypothesis and assume a stationary process

$$\mathbf{X}_t = \sum_{k=0}^{\infty} \boldsymbol{\Xi}_k \boldsymbol{\epsilon}_{t-k}, \tag{7}$$

where $\boldsymbol{\Xi}_0 = \mathbf{I}$. Then, the target and predictor can be expressed formally as:

$$\begin{aligned} \mathbf{z}_t &= \sum_{|k|<\infty} (\boldsymbol{\Gamma} \cdot \boldsymbol{\Xi})_k \boldsymbol{\epsilon}_{t-k} \\ \mathbf{y}_t &= \sum_{j \geq 0} (\mathbf{B} \cdot \boldsymbol{\Xi})_j \boldsymbol{\epsilon}_{t-j}. \end{aligned}$$

Here, $(\boldsymbol{\Gamma} \cdot \boldsymbol{\Xi})_k = \sum_{m \leq k} \boldsymbol{\Gamma}_m \boldsymbol{\Xi}_{k-m}$ and $(\mathbf{B} \cdot \boldsymbol{\Xi})_j = \sum_{m=0}^{\min(L-1,j)} \mathbf{B}_m \boldsymbol{\Xi}_{j-m}$ represent the convolutions of the sequences $\boldsymbol{\Gamma}_k$ and $\mathbf{B}_j$ with the Wold-decomposition $\boldsymbol{\Xi}_m$ of $\mathbf{x}_t$. For given $(\mathbf{B} \cdot \boldsymbol{\Xi})_j$, $j = 0, ..., L-1$, the original coefficients $\mathbf{B}_k$ can be derived through deconvolution:

$$\mathbf{B}_k = \begin{cases} (\mathbf{B} \cdot \boldsymbol{\Xi})_0 \boldsymbol{\Xi}_0^{-1} & , k = 0 \\ \left((\mathbf{B} \cdot \boldsymbol{\Xi})_k - \sum_{m=0}^{k-1} \mathbf{B}_m \boldsymbol{\Xi}_{j-m} \right) \boldsymbol{\Xi}_0^{-1} & , k = 1, ..., L-1 \end{cases} \tag{8}$$

In our parametrization, $\boldsymbol{\Xi}_0^{-1} = \mathbf{I}$ simplifies this expression. Denote the elements of the matrices as follows:

$$(\boldsymbol{\Gamma} \cdot \boldsymbol{\Xi})_{lmk} \quad \text{and} \quad (\mathbf{B} \cdot \boldsymbol{\Xi})_{lmj}.$$

We now assume $h \geq 0$ (note: analogous but slightly more complex expressions arise for $h < 0$, i.e., backcasting, which are omitted here). In addition, L should be chosen sufficiently large to ensure $(\boldsymbol{\Gamma} \cdot \boldsymbol{\Xi})_{h+k} \approx 0$ and $(\mathbf{B} \cdot \boldsymbol{\Xi})_j \approx 0$ for $k, j \geq L$.[19] This restriction is typically satisfied in practical applications;

[19] The top panels of Figure 5 indicate a relatively rapid decay of the filter coefficients, suggesting that a value of L greater than 30 may be appropriate in the present context.

otherwise, the criterion can be generalized to accommodate atypical long-memory prediction problems (see Wildi 2025). We define:

$$\begin{aligned}(\boldsymbol{\gamma}\cdot\boldsymbol{\xi})_{ijh} &= ((\boldsymbol{\Gamma}\cdot\boldsymbol{\Xi})_{ijh},(\boldsymbol{\Gamma}\cdot\boldsymbol{\Xi})_{ijh+1}...,(\boldsymbol{\Gamma}\cdot\boldsymbol{\Xi})_{ijh+L-1})'\\ (\boldsymbol{\gamma}\cdot\boldsymbol{\xi})_{ih} &= \left((\boldsymbol{\gamma}\cdot\boldsymbol{\xi})'_{i1h},(\boldsymbol{\gamma}\cdot\boldsymbol{\xi})'_{i2h},...,(\boldsymbol{\gamma}\cdot\boldsymbol{\xi})'_{inh}\right)'\\ (\mathbf{b}\cdot\boldsymbol{\xi})_{ij} &= ((\mathbf{B}\cdot\boldsymbol{\Xi})_{ij0},(\mathbf{B}\cdot\boldsymbol{\Xi})_{ij1},...,(\mathbf{B}\cdot\boldsymbol{\Xi})_{ijL-1})'\\ (\mathbf{b}\cdot\boldsymbol{\xi})_{i} &= \left((\mathbf{b}\cdot\boldsymbol{\xi})'_{i1},(\mathbf{b}\cdot\boldsymbol{\xi})'_{i2},...,(\mathbf{b}\cdot\boldsymbol{\xi})'_{in}\right)'\end{aligned}$$

It is important to note that if $\boldsymbol{\Xi}_k = \mathbf{0},\ k>0$, resulting in $\mathbf{X}_t = \boldsymbol{\epsilon}_t$, then we have $(\boldsymbol{\gamma}\cdot\boldsymbol{\xi})_{ih} = \boldsymbol{\gamma}_{i\cdot h}$ and $(\mathbf{b}\cdot\boldsymbol{\xi})_i = \mathbf{b}_i$, as defined in the preceding section. The generalized M-SSA criterion can then be expressed as:

$$\begin{aligned}&\max_{(\mathbf{b}\cdot\boldsymbol{\xi})_i}(\boldsymbol{\gamma}\cdot\boldsymbol{\xi})'_{ih}\tilde{\mathbf{I}}(\mathbf{b}\cdot\boldsymbol{\xi})_i\\ \text{s.t. }&(\mathbf{b}\cdot\boldsymbol{\xi})'_i\tilde{\mathbf{M}}(\mathbf{b}\cdot\boldsymbol{\xi})_i = \rho_i\\ &(\mathbf{b}\cdot\boldsymbol{\xi})'_i\tilde{\mathbf{I}}(\mathbf{b}\cdot\boldsymbol{\xi})_i = 1\end{aligned} \tag{9}$$

for $i = 1,...,n$. In analogy to the above white noise case, the criterion involves affine and quadratic functions of $(\mathbf{b}\cdot\boldsymbol{\xi})_i$ and thus can be solved for the convolution weights. The requested M-SSA coefficients $\mathbf{b}_i$ can subsequently be determined through deconvolution, as indicated in Eq. (8). The present discussion is confined to stationary processes, omitting a lengthier treatment of rank-defficient designs (cointegration) in the non-stationary case.

The subsequent analysis employs the M-SSA framework to construct 'smooth' predictors for forecasting z_{it+h} as well as for forecasting the original (unfiltered) GDP. In particular, it is shown that controlling the sign change rate within the M-SSA framework has a quantifiable effect on the hit rate and false positive rate.

5 Empirical Results: M-SSA Performance

5.1 M-MSE vs. M-SSA Nowcasts

For illustration, consider the problem of nowcasting HP-filtered GDP, denoted by HP-GDP, where the target z_t is generated using the two-sided HP(160) filter examined previously. We compare the classic multivariate mean squared error (MSE) predictor (denoted M-MSE) with the M-SSA predictor, where we set $HT_1 = 1.5 \times HT_{MSE}$. While this choice is somewhat arbitrary, it provides a smoothing that is both illustrative and sufficiently significant to be measurable while avoiding excessive distortions of the accuracy component in the accuracy–smoothness dilemma (our package can be used to experiment with alternative settings). This constraint ensures that the M-SSA predictor produces roughly 33% fewer sign changes than M-MSE in large samples. Comparisons with M-MSE allow us to see the cost of this constraint in terms of prediction performance (target correlation). Both predictors use the same underlying VAR(1) model introduced in Section 3 and estimated on data up to Q4-2007. This

ensures a lengthy out-of-sample period that includes significant events such as the financial crisis, the sovereign debt crisis, and the COVID-19 pandemic.[20]

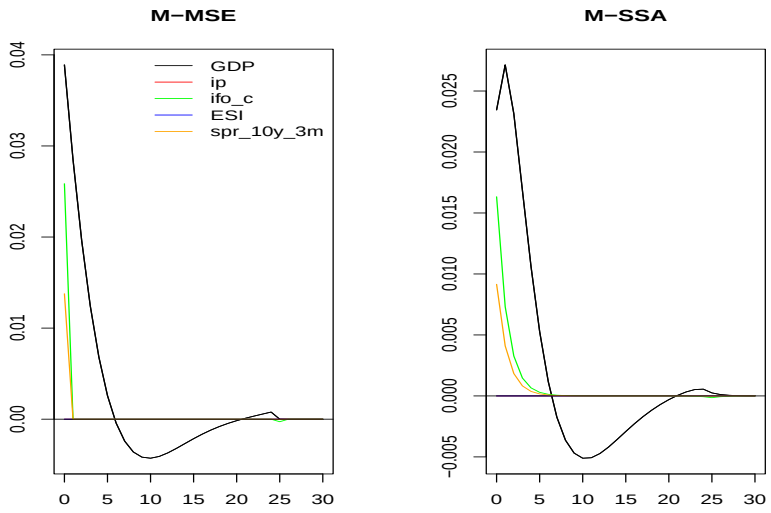


Figure 6: M-MSE (left) and M-SSA (right) nowcast filter weights.
Note: Note the difference in scales.

Figure 6 displays the coefficients of the multivariate nowcasts, where M-SSA's increased smoothing reduces the maximum amplitudes of the filter weights and slows their rate of decay at longer lags. The resulting filtered series, presented in Figure 7, are standardized to enable direct comparisons. The M-SSA nowcast is slightly smoother than that of the M-MSE, exhibiting approximately one-third fewer sign-changes. Figure 8 compares both nowcasts in relation to the financial crisis, highlighting that the classic (MSE-based) design tends to produce unwanted 'noisy' sign changes at (or in the vicinity of) the transitions between recessions and expansions.

[20] The economic sentiment indicator (ESI) does not receive any weight in the parsimonious VAR(1) model estimated over the shorter sub-sample ending in Q4-2007. Technical details and background information on the M-SSA are provided in Wildi (2025). All computations are conducted using the M-SSA package, as documented in Wildi (2025). The package can be accessed at `https://github.com/wiaidp/R-package-SSA-Predictor`.

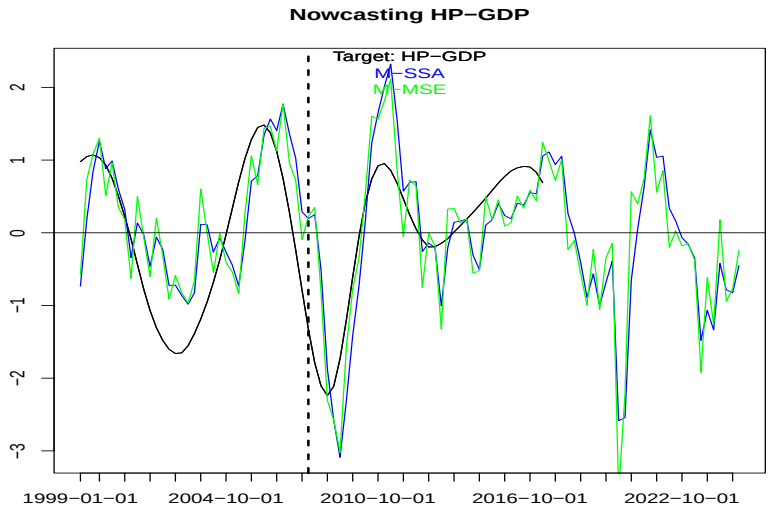


Figure 7: Nowcasts of Smoothed GDP.

Note: M-SSA (blue) and M-MSE (green) nowcasts are compared with the two-sided HP(160) GDP target (black). The dashed vertical line delimits in- and out-of sample periods. Note that the two-sided filter does not extend to the sample end.

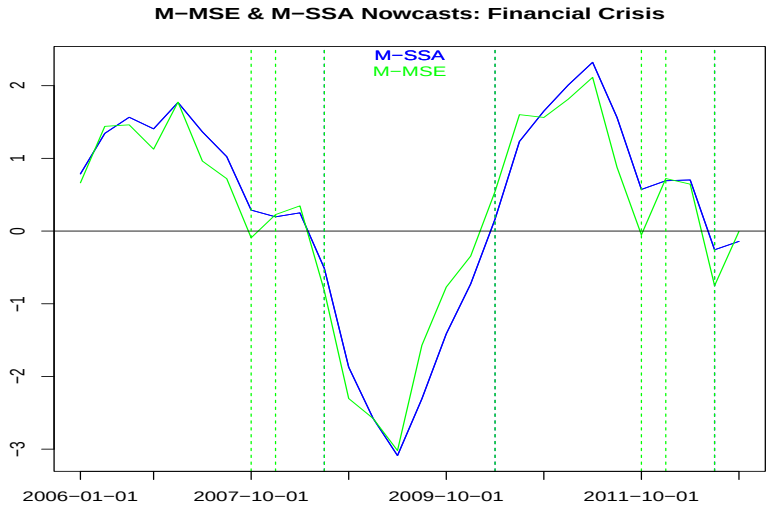


Figure 8: Nowcasts of Smoothed GDP during the financial crisis.

Note: M-SSA (blue) and M-MSE (green) nowcasts. Sign changes of the M-MSE and M-SSA are marked by green and blue vertical lines, respectively; overlap of these lines at start and end of the recession indicates no loss of timeliness despite a stronger smoothing. The holding time (HT) of each filter corresponds to the expected duration between its sign-changes. Here, we impose $HT_{MSSA} = 1.5 \times HT_{MSE}$. In large samples, we expect the M-SSA to exhibit approximately 33% fewer sign changes than the M-MSE.

We find that increased smoothness by the M-SSA, i.e., sample holding time increases from 3.7 to 6.5 quarters, comes at the cost of a slight reduction (0.68 vs 0.69) in target correlation. This highlights the inherent trade-off—often referred to as the prediction dilemma—between smoothness (sign change rate) and predictive accuracy (target correlation) within the M-SSA criterion. The results confirm that the smoothing constraint in M-SSA increases by 50% the HT relative to that of the M-MSE predictor[21] but has only a modest effect on the target correlation, suggesting that this trade-off could be beneficial in our application.

5.2 M-SSA Forecasts

Next, we consider forecasts of HP-GDP, changing the previous $h = 0$ (nowcast) to $h = 1, ..., 6$ quarters.

Figure 9 compares M-SSA filter weights for the nowcast and the one-year-ahead forecast($h = 4$). Noting the difference in vertical scales, we see that as h increases (right panel), the filter weights diminish, reflecting increased forecast uncertainty. We also see increased weight on the leading indicators relative to that on GDP. Additionally, a phase shift associated with the filter applied to GDP is observed (black line, right panel), reflecting the cyclical characteristics of the HP filter. While this phase effect might cause a sign change in the forecast as h increases — something to be avoided in this context — the low-pass filters assigned to the additional 'leading' indicators (green and yellow lines, right panel) enable 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.

While optimal in a specific mathematical sense, the above outcome may nevertheless appear to contradict the common notion that a predictor should assign the greatest weight to the most recent data points. However, the target HP-filter in our application can be forecast, to some extent, by incorporating its phase, as demonstrated in this example. This can lead to a counterintuitive situation in which more weight is assigned to past observations (those that are 'in phase' with the projected future). The ability to track the target during long monotonic (non-cyclical) expansion phases is in turn provided by the M-SSA sub-filters of the leading indicators, whose monotonically decaying weights align with our standard intuition. In this combined scenario, it is possible to look ahead and track the future target while maintaining controlled noise suppression. The effectiveness of this approach — the ability to anticipate the cyclical nature of the target by M-SSA — is illustrated in Figure 10, which contrasts M-SSA predictors with the traditional univariate HP-C for various forecast horizons.

[21] The convergence of the sample HT to its imposed expected value HT_1 can be verified utilizing the M-SSA package with very long time series samples.

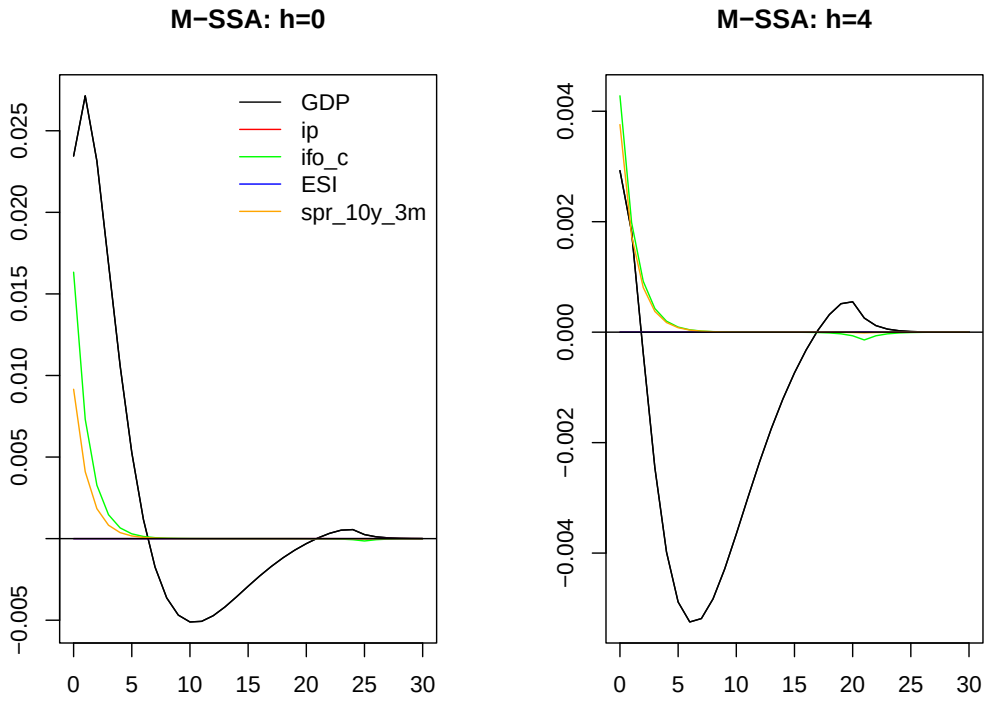


Figure 9: M-SSA filter weights.

Note: Note the difference in scales. Weights on *ip* and *ESI* are indistinguishable from zero. Left: $h = 0$ (nowcast), Right: $h = 4$ (one year ahead).

The figure compares the two filters across the periods of the financial crisis (left panels) and the COVID-19 pandemic (right panels) for forecast horizons h = 0, ..., 6. All series are standardized for ease of comparison.

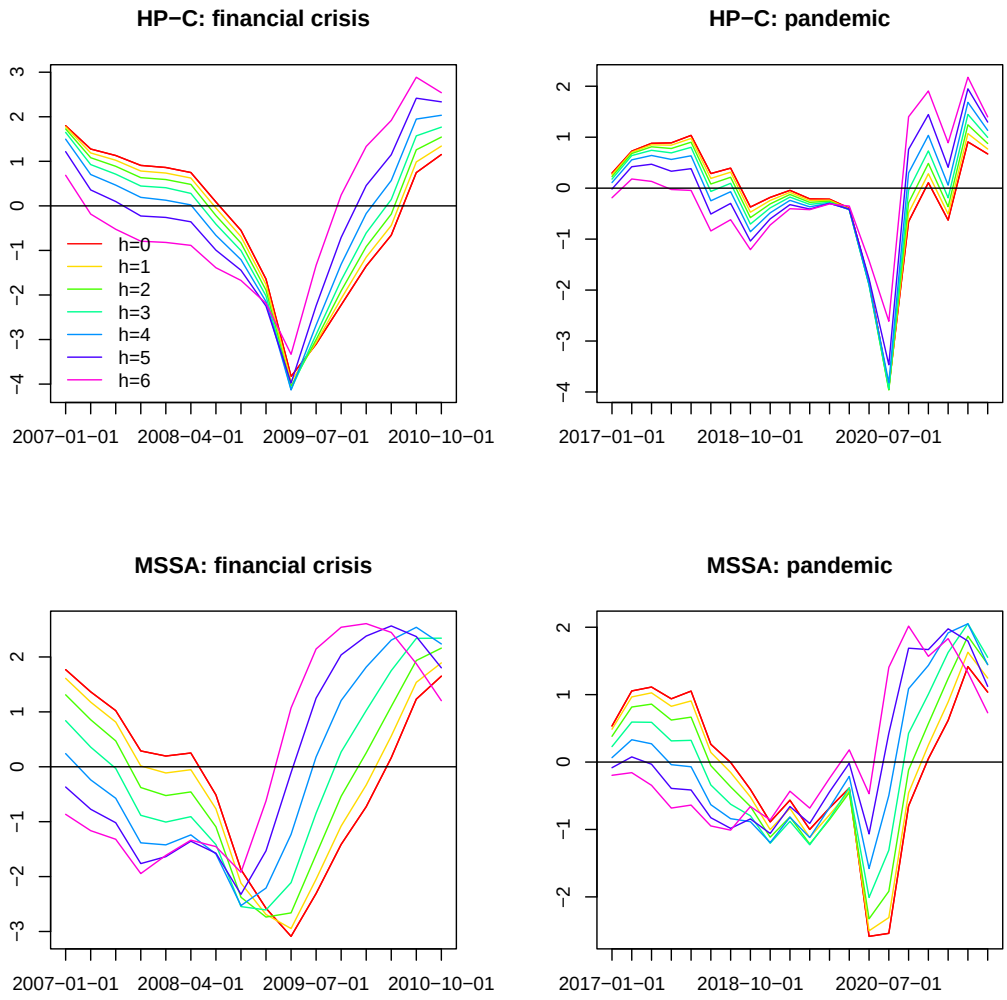


Figure 10: HP-C (top row) vs. M-SSA (bottom row) for horizons $h = 0, ..., 6$.

Note: Financial Crisis (left panels) and the Pandemic (right panels).
All series are standardized for ease of comparison.

Forecasts using the M-SSA approach consistently shift leftwards at all levels as h increases, including the peaks and troughs of the series. Conversely, the positions of peaks and troughs in the HP-C series are largely unaffected by increases in h. The leftward-shift of the M-SSA forecasts can be partly attributed to the increasing weight assigned to the additional leading indicators, which dominate the GDP component at longer forecast horizons. The systematic left-shift of the M-SSA is a key determinant of its out-of-sample forecast performances, as we see below.

It is worthwhile to remind that the M-SSA predictors maintain a constant level of smoothness, irrespective of the forecast horizon, due to the imposed fixed HT constraint —a feature that distinguishes them from the univariate forecasts.

5.3 M-SSA GDP Predictor

The final step in our framework uses the M-SSA filtered series as a predictor for actual GDP growth. This is achieved through a simple regression of forward-shifted GDP on the M-SSA predictor, where M-SSA is optimized for tracking (the acausal) HP-GDP at various forecast horizons. We distinguish between the forecast horizon h, for which the M-SSA filter is optimized, and the target shift $shift$, which determines how far ahead GDP is evaluated. This distinction allows predictors optimized for longer horizons to be evaluated at shorter horizons, exploiting the systematic left-shift induced by the M-SSA design. All the results that we present exclude the pandemic; further results including the pandemic are available in the Appendix.[22]

Tracking GDP: Optimal Weighting

Let $\text{M-SSA}_t^{\text{GDP}}(h)$ denote the M-SSA predictor of HP-GDP at time t for forecast horizon h. A GDP forecast can be obtained from the following regression specification[23]

$$\text{GDP}_{t+shift} = \beta_0(shift, h) + \beta_1(shift, h) \cdot \text{M-SSA}_t^{\text{GDP}}(h) + \epsilon_t^{shift,h}. \tag{10}$$

The corresponding out-of-sample predictor $\widehat{\text{GDP}}_{t+shift}(h)$ is given by:

$$\widehat{\text{GDP}}_{t+shift}(h) \equiv \hat{\beta}_0(shift, h, t) + \hat{\beta}_1(shift, h, t) \cdot \text{M-SSA}_t^{\text{GDP}}(h), \tag{11}$$

where the regression parameters $\hat{\beta}_0(shift, h, t), \hat{\beta}_1(shift, h, t)$ are estimated recursively (expanding window) using data up to time point $t-1$. The associated out-of-sample forecast error is defined as

$$\hat{\epsilon}_t^{shift,h} \equiv \text{GDP}_{t+shift} - \widehat{\text{GDP}}_{t+shift}(h). \tag{12}$$

In the empirical implementation, the M-SSA forecasts are held fixed, while the regression parameters $\hat{\beta}_0(\cdot)$ and $\hat{\beta}_1(\cdot)$ are updated each quarter using an expanding window beginning in 2008Q1.

Since $\text{M-SSA}_t^{\text{GDP}}(h)$ is based on data up to Q4-2007 and its weights are fixed thereafter, $\widehat{\text{GDP}}_{t+shift}(h)$ constitutes an out-of-sample predictor of $\text{GDP}_{t+shift}$ for all $shift \geq 0$ and all t larger than Q4-2007.[24] This procedure yields a sequence of optimally weighted M-SSA forecasts that adapt over time (via recursive estimation of $\hat{\beta}_0(shift, h, t), \hat{\beta}_1(shift, h, t)$) to evolving GDP dynamics.

[22] See Tables A2 through A5.

[23] By default, the M-SSA criterion (9) generates a standardized predictor. We calibrate the regression coefficients to match the level and scale of the forecast target.

[24] Equations (10) and (11) could be extended to incorporate additional M-SSA components, but for illustrative purposes we focus on the simplest (and most robust) design.

Table 4: Relative RMSE of M-SSA GDP Predictors vs. Mean Benchmark.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.973	0.937*	0.888*	0.850**	0.855**	0.900**	0.956*
Shift=1	1.073	1.051	0.997	0.918*	0.870**	0.873**	0.906**
Shift=2	1.069	1.077	1.063	1.006	0.940*	0.907**	0.907**
Shift=3	1.020	1.045	1.066	1.048	0.997	0.957*	0.941*
Shift=4	0.983	1.016	1.056	1.064	1.023	0.974	0.942*
Shift=5	0.988	1.026	1.071	1.088	1.057	1.007	0.963*

Note: Out-of-sample evaluation starting in Q1-2008, without pandemic.

M-SSA weights are fixed, based on data up to Q4-2007.

Values < 1 indicate superior predictions from Eq. (11).

Shift refers to the left-shift of the target $\text{GDP}_{t+shift}$

h refers to the forecast horizon for which $\text{M-SSA}_t^{\text{GDP}}(h)$ is optimized.

* and ** refer to the rejection of $H_0 : \alpha_1 = 0$ in Eq. (13) at the 5% and 1% levels respectively.

Relative RMSEs

Forecast accuracy is again evaluated in terms of relative root mean squared forecast errors (rRMSEs), where the performance of $\widehat{\text{GDP}}_{t+shift}(h)$ can be benchmarked against that of the mean predictor (Table 4) and that of the direct forecast (Table 5). For comparability, both benchmark predictors are out-of-sample designs that are updated quarterly. The statistical significance of the (out of sample) M-SSA GDP predictor in Eq. (11) can be evaluated by:

$$\text{GDP}_{t+shift} = \alpha_0 + \alpha_1 \widehat{\text{GDP}}_{t+shift}(h) + \nu_t^{shift,h}, \tag{13}$$

for observations after Q4-2007. Tests of the null hypothesis $H_0 : \alpha_1 = 0$ are based on heteroskedasticity- and autocorrelation-consistent (HAC) standard errors with results indicated in Table 4.[25]

Our results can be summarized as follows. When excluding the pandemic, the direct forecasts outperform the mean forecast up to two quarters ahead (Table A1 in the appendix). In contrast, when optimized for forecast horizons $h \geq 5$ the GDP predictors $\widehat{\text{GDP}}_{t+shift}(h)$ outperform the mean benchmark up to 5 quarters ahead (see Table 4.) Similarly, when optimized for $h \geq 4$, $\widehat{\text{GDP}}_{t+shift}(h)$ outperform the direct forecasts for $shift \geq 3$ (Table 5). In both cases the improvement in forecast performance is statistically significant at some of the longest forecast horizons. Finally, the strong performance of predictors optimized for larger forecast horizons ($h \geq 4$), at smaller shifts ($shifts < 4$) corroborates the importance of the left-shift of the corresponding predictors (see Figure 10.) Remarkably, these findings hold even though M-SSA was restricted to a relatively short in-sample window ending in Q4-2007, i.e., prior to the onset of the financial crisis.

5.4 Hit Rate vs. False Alarm Rate

To this point, indicators have been assessed by the extent to which they correlate with or lead movements in the target variable. An alternative approach would be to consider their ability to predict the

[25] * denotes significance at the 5% level; ** indicates significance at the 1% level. Note that statistical significance is influenced by the relatively short out-of-sample period of 17 years.

Table 5: Relative RMSEs of M-SSA GDP Predictors vs. Direct Forecasts.

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.128	1.086	1.030	0.985	0.992	1.044	1.109
Shift=1	1.280	1.254	1.189	1.096	1.037	1.041	1.081
Shift=2	1.173	1.181	1.166	1.104	1.031	0.995	0.995
Shift=3	1.012	1.037	1.058	1.040	0.989	0.949*	0.934*
Shift=4	0.982	1.016	1.056	1.063	1.022	0.974	0.941*
Shift=5	0.961	0.997	1.041	1.058	1.028	0.979	0.936*

Note: Out-of-sample evaluation starting in Q1-2008, without pandemic.

M-SSA weights are fixed, based on data up to Q4-2007.

Values < 1 indicate superior predictions from Eq. (11).

Shift refers to the left-shift of the target $\text{GDP}_{t+shift}$

h refers to the forecast horizon for which $\text{M-SSA}_t^{\text{GDP}}(h)$ is optimized.

* and ** refer to the rejection of $H_0 : \alpha_1 = 0$ in Eq. (13) at the 5% and 1% levels respectively.

correct "type" of outcome, e.g. whether growth will be stronger or weaker than average, or whether the economy will be in an expansion or a recession. This may be of particular interest to policymakers whose primary concern is to intervene "in the right direction" so as to stabilize the economy overall. Therefore we now analyse the M-SSA predictors as binary classifiers, focusing in particular on the ROC (Receiver Operating Characteristic Curve) curve and the AUC (Area Under the Curve).[26]

Consider a binary random variable $Z \in \{0, 1\}$ as our target, and a continuous variable Y. Our predictor for Z is given by the function

$$\mathbf{1}(Y \leq \tau) \equiv I(Y, \tau) \tag{14}$$

where $\mathbf{1}()$ is the indicator function.[27] The joint distribution of Z and Y allows us to define the Hit Rate $H(\tau)$ and the False Alarm Rate $F(\tau)$ as

$$H(\tau) \equiv Pr(Y \leq \tau | Z = 1) \tag{15}$$

$$F(\tau) \equiv Pr(Y \leq \tau | Z = 0) \tag{16}$$

The usefulness of a given variable Y as a predictor will depend on the value of τ selected, which in turn may be influenced by the relative importance attached to $H(\tau)$ and $F(\tau)$. To compare the range of possibilities associated with various choices for $Y\tau$, we plot $F(\tau)$ vs $H(\tau)$ on the unit square for all values of τ in the range of Y. This is called the ROC curve.

An ideal predictor would be one where $H(\tau) = 1$ and $F(\tau) = 0$ $\forall\tau$ (i.e. $Y > \tau$ whenever $Z = 0$ and $Y \leq \tau$ whenever $Z = 1$.) In this ideal case, the ROC curve would consist of the extreme left

[26] For example, see Yang et al. (2024) for a recent discussion of the ROC, the AUC and their relation to selected other binary classifiers. The discussion below follows their notation.

[27] The value of the indicator function is equal to 1 when the condition in its argument is satisfied; otherwise, the value of the function is equal to 0.

and top edges of the plot. Another extreme case would be one where $\{Y, Z\}$ are independent, so $H(\tau) = F(\tau)\ \forall \tau$. In this case, the ROC curve would simply be the 45-degree line through the origin. When comparing two indicators Y_a and Y_b, if the ROC curve of Y_a lies strictly above and to the left of that of Y_b, then we would always prefer to use Y_a as our predictor in place of Y_b since for any choice of τ it will always give us a superior trade-off between $H(\tau)$ and $F(\tau)$. When the ROC curves for two indicators cross, however, then our preferred indicator will depend on the relative importance that we attach to the hit and false alarm rates.[28]

A popular approximate measure of the overall performance of an indicator Y that abstracts from the choice of τ is the AUC, defined literally as the area under the ROC curve. Because the ROC is defined on the unit square, the maximum value of the AUC = 1. Above, we saw that when $\{Y, Z\}$ are independent, the ROC curve lies on the 45-degree line, so the area under the curve is then 0.5. The range of AUCs between 1 and 0.5 therefore provides some quantitative guidance as to which predictor variables Y have the potential to improve the prediction of Z.[29]

We construct ROC curves using the empirical frequency with which the sign of the target variable $GDP_{t+shift}$ corresponds to periods where the forecast $\widehat{GDP}_{t+shift}(h)$ exceeds a threshold τ. The full curve is constructed by varying τ across the range of the predictor space.[30] We examined all pairwise combinations of h and $shift$, but for simplicity of exposition the results below focus on the diagonal case where $h = shift$. Results are based on the entire sample, excluding the pandemic period. Figure 11 compares ROCs of the predictors for small and medium-term forecast horizons and Table 6 presents AUC (area under the curve) for forecast horizons 1-5 quarters ahead.

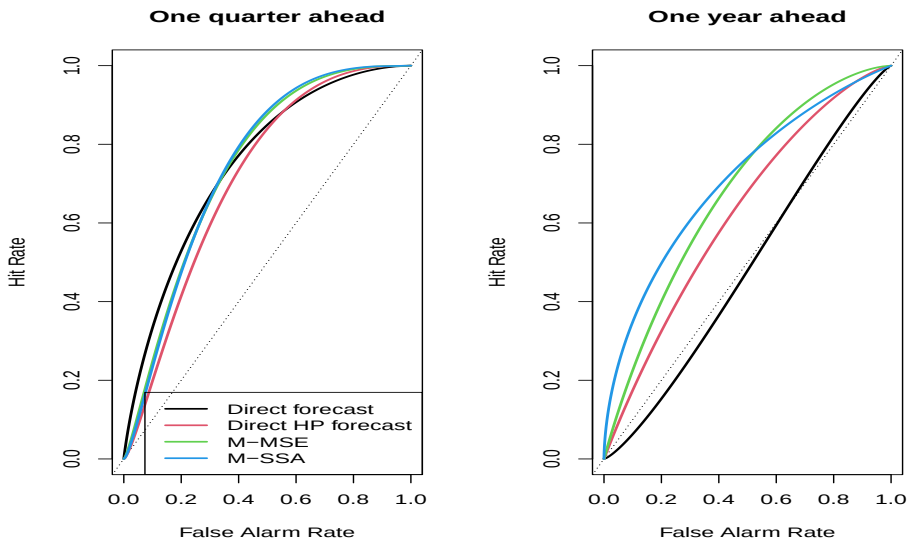


Figure 11: Hit rate vs. false alarm rate.

Note: ROC curves compare the classification performance of direct forecasts, HP-C forecasts, M-MSE, and M-SSA at one-quarter (left panel) and one-year (right panel) horizons.

[28] As shown in Wildi (2025), the sign accuracy can be related to the target correlation in the objective function of the (M-)SSA criterion; additionally, the rate of false alarms depends on the holding time (HT), so that the tradeoff underlying the ROC curve is to some extent captured by the M-SSA criterion.

[29] AUCs < 0.5 suggest that $-Y$ may give more useful predictions than Y.

[30] Although we are using standardized GDP, the regressions automatically account for scales and level shifts, so the results directly apply to the original GDP.

Table 6: AUCs of predictors against GDP at various forecast horizons.

	Direct forecast	Direct HP forecast	M-MSE	M-SSA
Forecast horizon 1	0.752	0.715	0.746	0.747
Forecast horizon 2	0.697	0.663	0.729	0.720
Forecast horizon 3	0.554	0.622	0.684	0.709
Forecast horizon 4	0.486	0.624	0.680	0.708
Forecast horizon 5	0.593	0.636	0.643	0.701

At forecast horizons less than or equal to two quarters, the direct forecasts perform almost as well or slight better than the M-SSA or M-MSE filters. However, for horizons $h > 2$, the filters outperform the direct approach: the univariate 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.

6 Robustness and Stability

Model parameters evolve as new observations become available, and forecasts are revised as incoming data update the information set. Similarly, revisions of the M-SSA GDP predictor can arise from data updates, which we do not analyze further here, or from updates to predictor parameters.[31] The latter can be further separated into changes in the regression weights and VAR parameters.

6.1 Parameter Stability and Revisions

A key concern for real-time applications is the stability of forecast parameters and the magnitude of forecast revisions as new data arrive. For illustration, we focus on a forecast horizon $h = 4$ for the M-SSA GDP predictor and a forward-shift $shift = 4$ for GDP, noting that similar results would be obtained with other combinations of h and $shift$.

Figure 12 compares the final and real-time predictors: as the sample size increases, the real-time predictor converges steadily to the final predictor.[32] In addition, Figure 13 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. 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.

[31] Except for GDP, 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.

[32] The rate of convergence is partially determined by the simplicity of our design, which avoids incorporating additional M-SSA components.

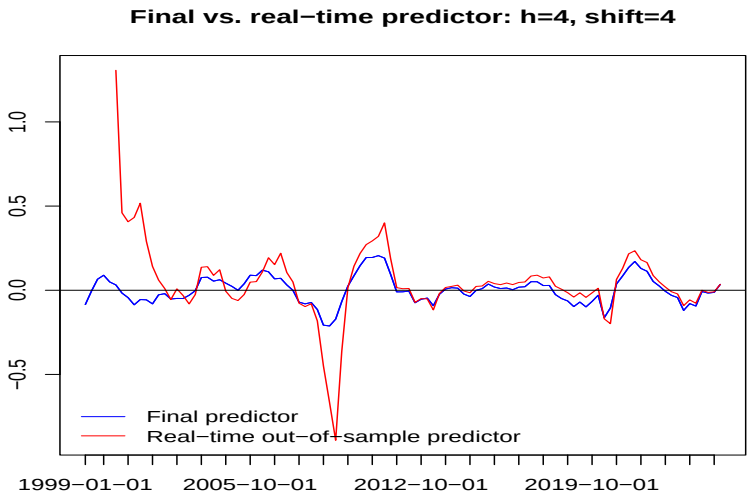


Figure 12: Revisions of the M-SSA GDP predictor.

Note: Comparison of final (blue) and real-time (red) M-SSA GDP predictors for h=4, shift=4. These revisions result from quarterly updates to the regression coefficients of equation (11). The M-SSA filter itself remains fixed, based on data up to Q4-2007.

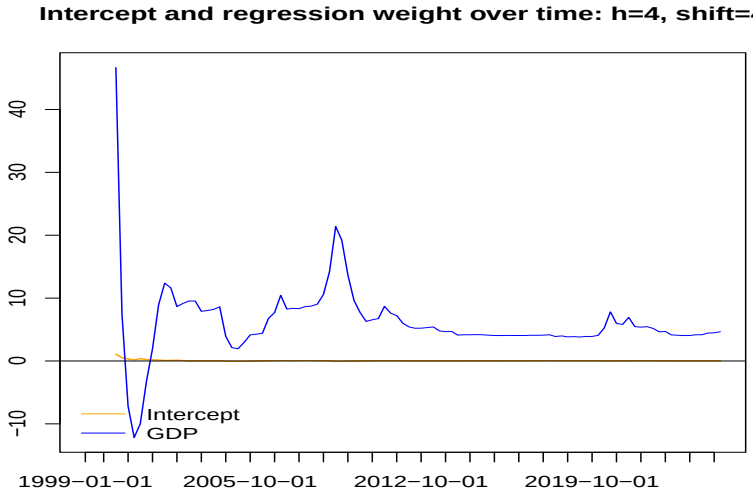


Figure 13: Revisions of intercept and regression weights.

Note: Quarterly estimates of the intercept and GDP coefficient in the M-SSA regression for h=4, shift=4 according to Equation (11).

6.2 VAR Model Updates and Filter Revisions

The second source of revisions in the M-SSA GDP predictor arises from updates to the VAR(1) model within the M-SSA filter.[33] We compare the original estimates, based on data up to Q4-2007, 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. Figure 14 compares the MA-inversions (impulse responses) of the GDP equation derived

[33] Similar findings apply to more sophisticated alternatives such as the VAR(3) discussed in Section 2.

from the 'old' and the updated VAR models. Distinctly, the updated model incorporates the ESI as an additional explanatory variable for GDP, reflecting evidence from the longer data sample. Figure 15 illustrates the resulting effect on the M-SSA filter targeting HP-GDP at forecast horizon $h = 4$. The updated filter (right panel) assigns more weight to the additional 'leading' indicators relative to GDP. Finally, Figure 16 compares the M-SSA GDP predictors from the initial and updated VAR models. To isolate the effect of the regression revision, the figure displays standardized series. The update to the VAR model results in a leftward shift of the predictor (represented by the blue line), which appears marginally more rapid compared to the original predictor (depicted by the red line). It is important to note that the smoothness (sign change rate) remains unchanged, though. This suggests that the out-of-sample performance of the M-SSA GDP predictor may be better than reported in Section 5.3, which is based on the fixed and increasingly outdated version of M-SSA.

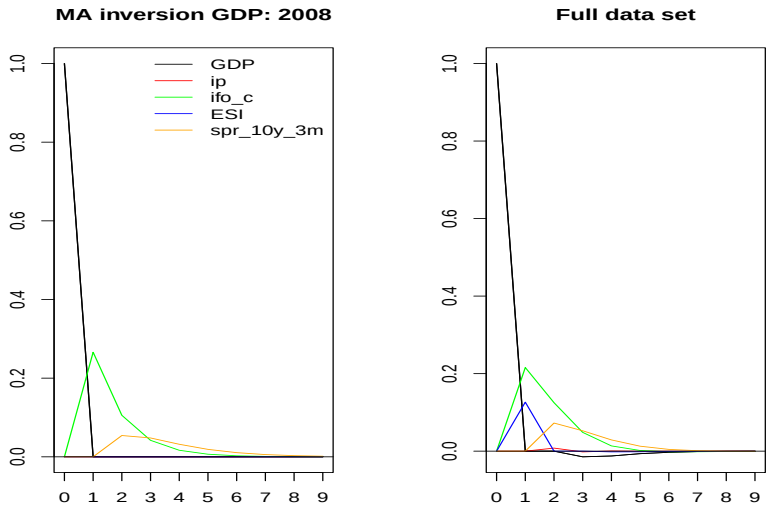


Figure 14: Impulse Responses from VAR Models: Initial vs. Full-Sample Estimates.

Note: MA-inverted impulse responses for GDP from VAR models estimated with data up to 2007Q4 (left) and the full sample excluding the pandemic (right).

7 Conclusion

Direct forecasts of GDP outperform simple benchmarks (such as the mean) up to two quarters ahead, largely due to noise in leading indicators. These forecasts struggle to track dips (slowdowns, recessions) or peaks (recoveries, expansions) in real time. To address this issue, we pre-filter the indicators to smooth-away fluctuations lasting much less than one year. The resulting predictor tends to forecast significantly better than the classic direct forecasts, due mainly to a slight left-shift most noticeable at sharp recessions such as the Great Financial Crisis or COVID-19-Pandemic. However, the predictor remains somewhat noisy, partly due to noise leakage in the simple one-sided filter we use.

To further improve performance, we propose a multivariate extension of the SSA framework of Wildi (2024). Unlike univariate filters, the M-SSA can extract information from leading indicators while controlling the smoothness of the predictor by targeting an expected time between sign changes (in

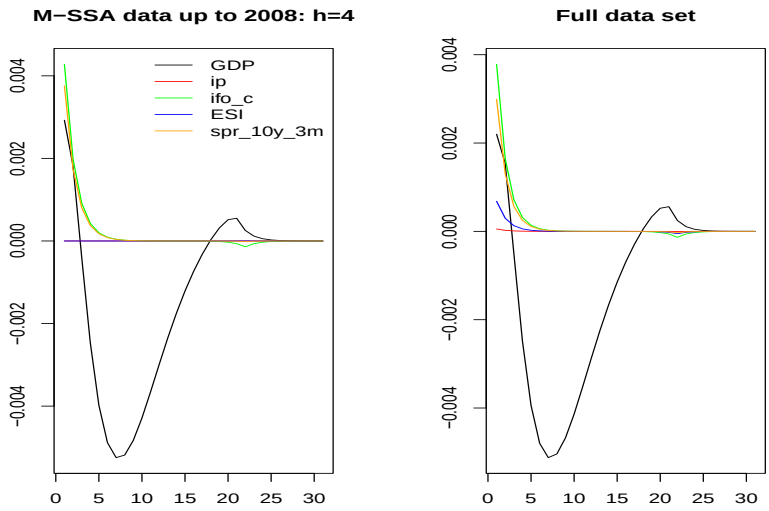


Figure 15: M-SSA Filter Weights: Initial vs. Full-Sample Estimates.

Note: M-SSA filters for HP-GDP at $h = 4$ using 2007Q4 (left panel) and full-sample (right panel) VAR models (without pandemic).

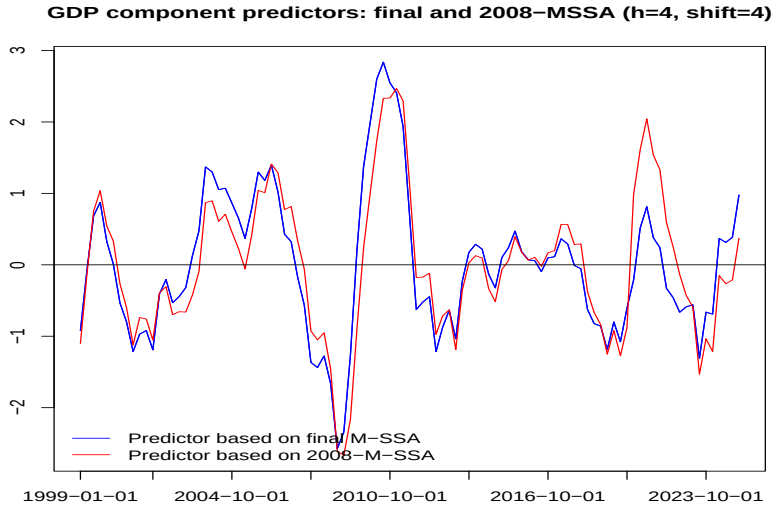


Figure 16: M-SSA GDP Predictors: Initial vs. Full-Sample Estimates.

Note: Standardized M-SSA GDP predictors from initial (red) and updated (blue) VAR models for $h = shift = 4$ (excluding pandemic).

the de-meaned series.) The resulting forecasts become increasingly left-shifted as the forecast horizon lengthens, allowing dips and peaks of the target series to be tracked more accurately, while limiting 'noisy' sign changes.

However, since M-SSA targets smoothed GDP, an additional step is necessary to map predictions for smoothed GDP into those for actual GDP growth. For simplicity, we just regress the M-SSA predictor on future GDP.[34]

Out-of-sample performance suggests that this predictor is statistically significant at horizons exceeding one year, and designs optimized for longer forecast horizons outperform the mean benchmark in terms of RMSE at all forecast horizons examined. The predictor also exceeds the performance of direct forecasts—based on unfiltered or HP filtered indicators—at horizons longer than two quarters. An ROC analysis confirms that filtering improves prediction of above- and below-average growth for forecast horizons exceeding two quarters, as the M-SSA framework allows more precise control of the rate of sign change, thereby improving the AUC. Remarkably, these results hold even though the M-SSA filter is optimized on a relatively short in-sample window ending prior to the onset of the financial crisis. Finally, an examination of revision errors associated with quarterly updates of the proposed M-SSA-based GDP predictor indicates that the regression parameters, VAR model, and resulting M-SSA filters stabilize after an initial burn-in period. This stabilization enhances the model's explainability and interpretability.

In conclusion, the new predictor design integrates the traditional direct forecast approach with a novel multivariate filter to target quarterly GDP growth up to one year ahead. This multivariate approach has significant forecasting power by leveraging a set of economic indicators leading GDP in real time while controlling the smoothness of the predictor through the sign change rate.

References

Drechsel, K. and Scheufele, R. (2012). The financial crisis from a forecaster's perspective. *Kredit und Kapital*, 45(1):1–26.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Heinisch, K. and Scheufele, R. (2018). Bottom-up or direct? forecasting German GDP in a data-rich environment. *Empirical Economics*, 54(2):705–745.

Heinisch, K. and Scheufele, R. (2019). Should forecasters use real-time data to evaluate leading indicator models for GDP prediction? German evidence. *German Economic Review*, 20(4):170–200.

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

[34] This predictor is intuitively appealing because the future HP-GDP, i.e., the target of M-SSA, is the low-frequency component of the future GDP target. Consequently, a strong link between M-SSA and future HP-GDP also indicates a connection with GDP, even if the statistical significance is obscured by the noise inherent in GDP.

Lehmann, R. and Reif, M. (2021). Predicting the German economy: Headline survey indices under test. *Journal of Business Cycle Research*, 17(2):215–232.

McElroy, T. and Wildi, M. (2019). The trilemma between accuracy, timeliness and smoothness in real-time signal extraction. *International Journal of Forecasting*, 35(3):1072–1084.

McElroy, T. and Wildi, M. (2020). The multivariate linear prediction problem: model-based and direct filtering solutions. *Econometrics and Statistics*, 14:112–130.

Phillips, P. C. B. and Jin, S. (2021). Business cycles, trend elimination, and the HP filter. *International Economic Review*, 62(2):469–520.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the hodrick-prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376.

Tsay, R. S. (2013). *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons.

Tsay, R. S., Wood, D., and Lachmann, J. (2022). *MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*. R package version 1.2.1.

Wildi, M. (2024). Business cycle analysis and zero-crossings of time series: a generalized forecast approach. *Journal of Business Cycle Research*.

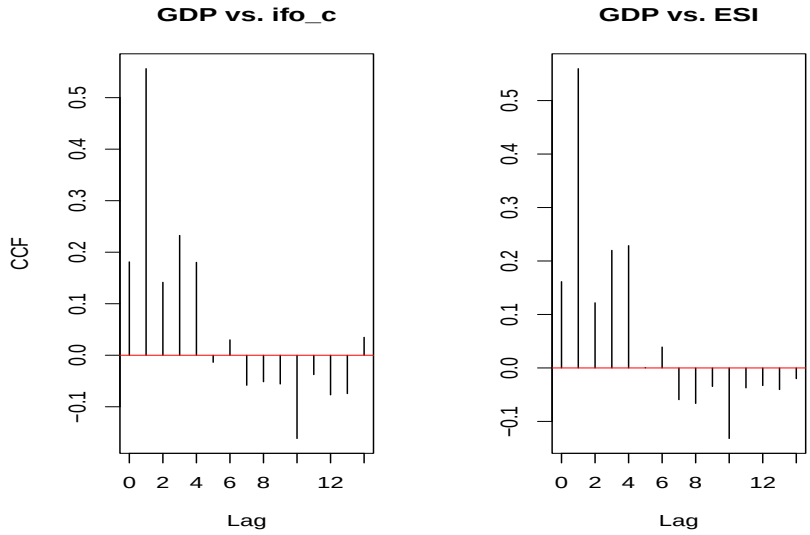
Wildi, M. (2025). Sign accuracy, mean-square error and the rate of zero crossings: a generalized forecast approach. mimeo.

Yang, L., Lahiri, K., and Pagan, A. (2024). Getting the ROC into Sync. *Journal of Business and Economic Statistics*, 42(1):109–121.

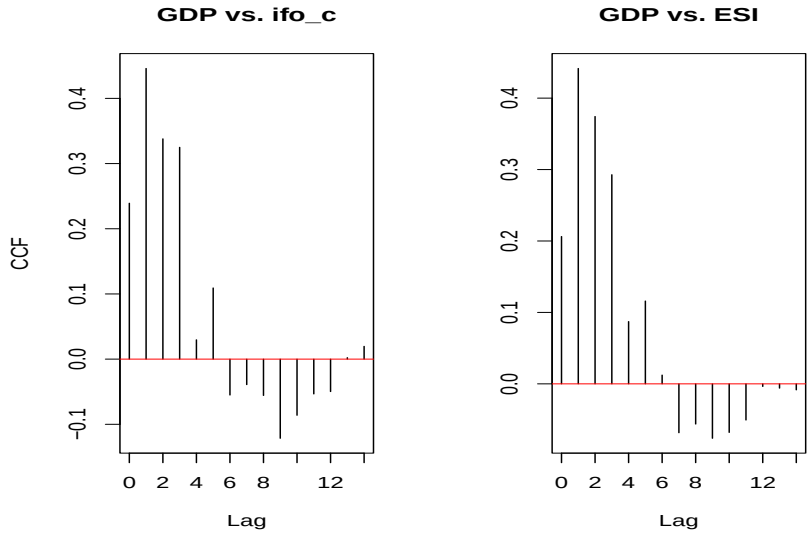
Zou, H. and Hastie, T. (2020). *elasticnet: Elastic-Net for Sparse Estimation and Sparse PCA*. R package version 1.3.

A Appendix

A Additional Figures and Tables

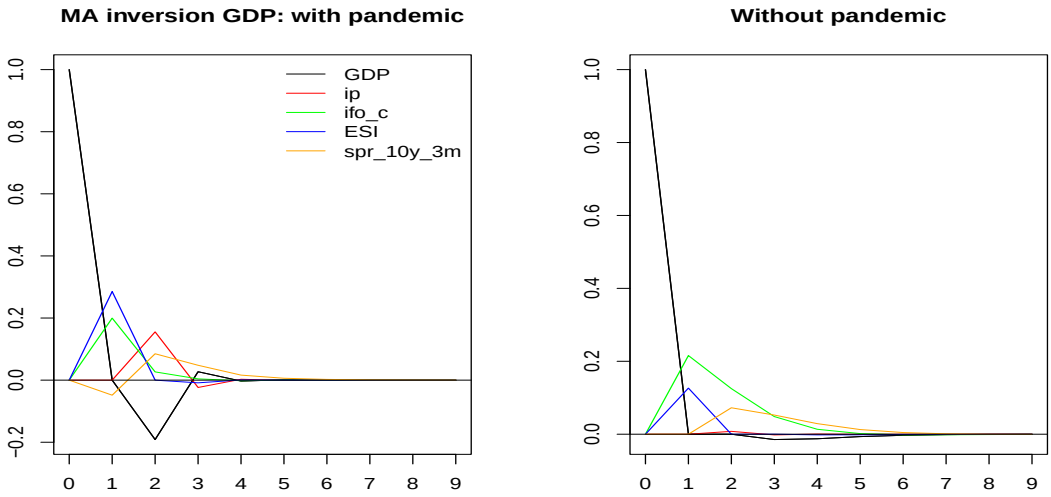


(a) Full data set from Q2-1991 to Q4-2024.

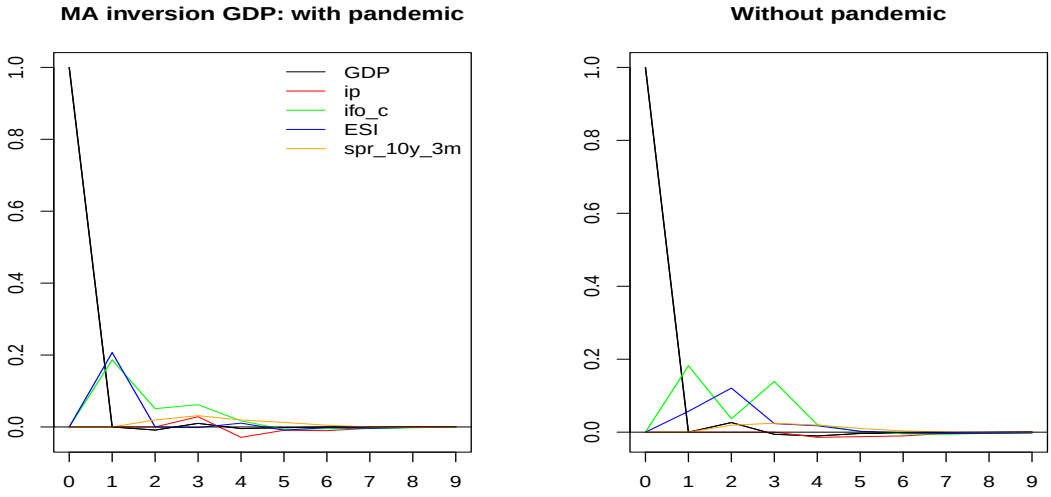


(b) Without Pandemic (omitting Q4-2019 to Q4-2020).

Figure A1: Sample cross-correlation functions (CCF) between GDP growth and Indicators.



(a) implied by VAR(1)



(b) implied by VAR(3)

Figure A2: Impulse Responses of GDP to Indicator Shocks from VAR Models.

Note: VMA coefficients for GDP implied by VAR models.
Estimation with pandemic (left) and without pandemic (right).

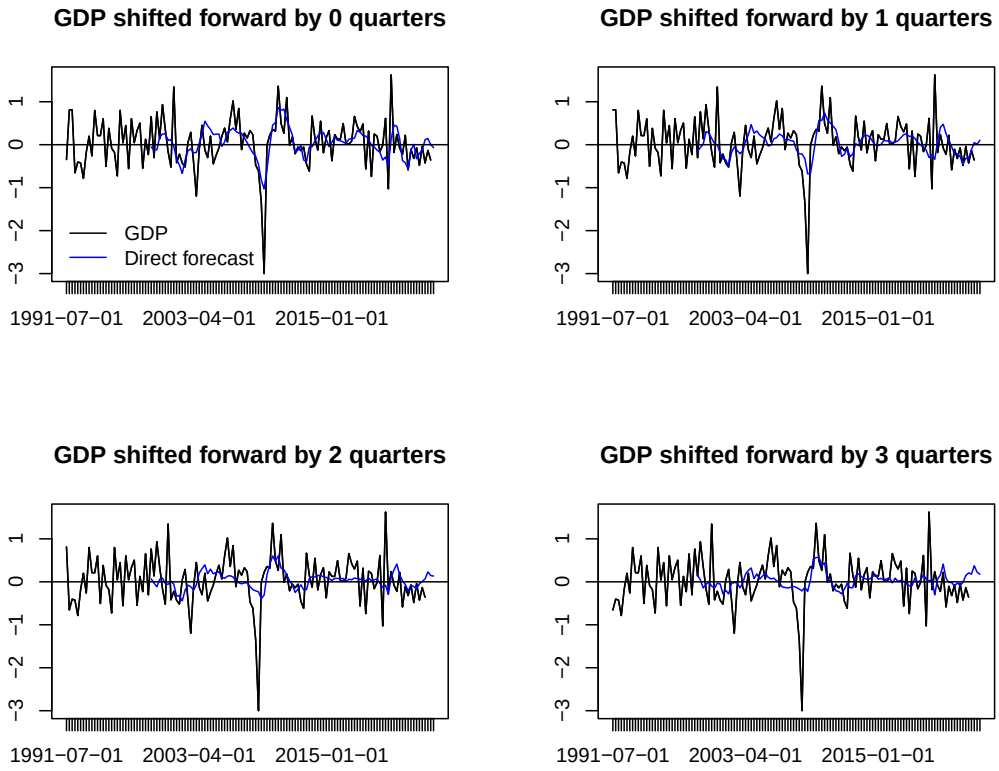


Figure A3: Direct GDP forecasts using HP-C filtered Indicators (blue) vs. GDP shifted forward (to the left) by 0-3 quarters (black line).

Note: Full data set (without pandemic).

Table A1: Relative RMSEs of Direct Forecasts vs. Mean Benchmark.

	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

Note: Out-of-sample rRMSEs for direct forecasts benchmarked against the expanding-mean.

Values below 1 indicate improvement over the mean.

Forecast sample excludes the pandemic.

B Additional Tables and Performance Metrics (including Pandemic).

Table A2: Significance of M-SSA GDP Component Predictor (Including Pandemic).

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	0.127	0.064	0.019	0.003	0.001	0.002	0.012
Shift=1	0.439	0.314	0.194	0.093	0.033	0.017	0.036
Shift=2	0.718	0.565	0.326	0.098	0.020	0.008	0.010
Shift=3	0.946	0.975	0.925	0.657	0.222	0.059	0.028
Shift=4	0.327	0.699	0.964	0.942	0.619	0.237	0.102
Shift=5	0.383	0.755	0.946	0.910	0.581	0.241	0.055

Note: p-values for $H_0 : \alpha_1 = 0$, HAC-adjusted Wald test, Equation (13).

Table A3: Relative RMSEs of M-SSA GDP Component Predictors vs. Mean Benchmark
(Including Pandemic).

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.005	0.986	0.960	0.935	0.927	0.939	0.964
Shift=1	1.060	1.054	1.032	0.998	0.973	0.965	0.970
Shift=2	1.028	1.029	1.020	0.993	0.963	0.950	0.952
Shift=3	1.018	1.027	1.034	1.022	0.994	0.972	0.961
Shift=4	0.990	1.006	1.025	1.032	1.017	0.998	0.985
Shift=5	1.000	1.016	1.037	1.046	1.032	1.008	0.985

Note: h denotes the optimization horizon of the M-SSA filter, while $shift$ denotes the evaluation horizon of GDP.

Table A4: Relative RMSEs of M-SSA GDP Component Predictors vs. Direct Forecasts (Including Pandemic).

	h=0	h=1	h=2	h=3	h=4	h=5	h=6
Shift=0	1.237	1.215	1.183	1.152	1.141	1.156	1.187
Shift=1	0.986	0.980	0.960	0.929	0.905	0.898	0.903
Shift=2	1.036	1.037	1.029	1.001	0.971	0.957	0.960
Shift=3	1.042	1.051	1.058	1.045	1.018	0.995	0.983
Shift=4	0.970	0.985	1.004	1.010	0.996	0.978	0.965
Shift=5	0.970	0.986	1.007	1.015	1.002	0.978	0.956

Note: h denotes the optimization horizon of the M-SSA filter, while $shift$ denotes the evaluation horizon of GDP.

Table A5: Relative RMSEs of Direct Forecasts vs Expanding Mean (Including 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

Halle Institute for Economic Research –
Member of the Leibniz Association

Kleine Maerkerstrasse 8

D-06108 Halle (Saale), Germany

Postal Adress: P.O. Box 11 03 61

D-06017 Halle (Saale), Germany

Tel +49 345 7753 60

Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188

The IWH is funded by the federal government and the German federal states.