A Comparative study of Different Predictor Concepts on a HWI Data Set

Marc Wildi

Institute of data analysis and process design (ZHAW)

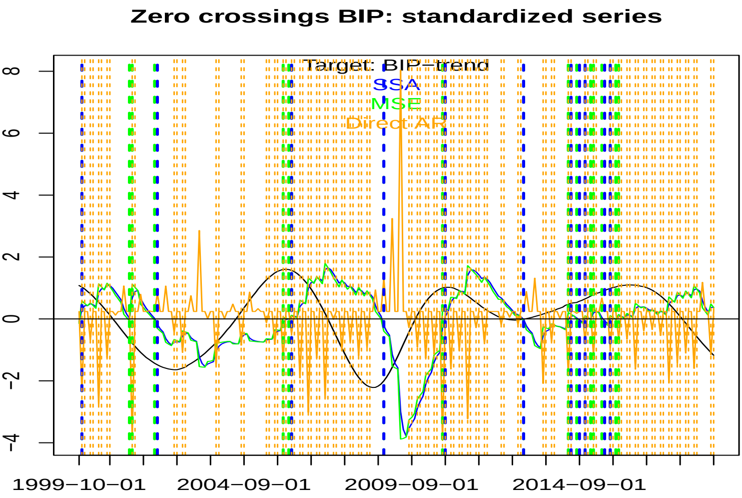
Table of Contents I

Table of Contents II

Introduction

- See slides on ip (ip-case) for a similar introduction: motivation of trends vs. direct forecasts and theses

Zero-crossings BIP: Standardized Series



Overview of (and Links to) Performance Metrics in this Study

- Relative **forecast mean-square** error (target: BIP-shifted): slide (25) (direct AR forecasts full sample), (27) (same but shorter sample), (38) (univariate filters, full sample), (42) (same but before Pandemic), (64) (multivariate filters)
- Relative **filter mean-square** error (target: BIP-trend): (66) (outperformance of multivariate over univariate)
- Target **correlations**: (70), (72), (89) and (100)
- **Sign accuracies**: (97) and (99)
- **t-statistics**: (40), (univariate filters and direct AR, full sample), (43) (same but prior pandemic), (64) (multivariate filters)
- **Holding times**: (46) (direct AR and univariate filters), (87) and (101) (M-SSA vs. M-MSE)

Section 2

Data

Data

- Selected (small) data set: BIP,ip,ifo-c,ifo-exp,ifo-l,ESI,spr-10y-3m (we did not consider pmi because the series is shorter: for consistency we thus would have to shorten the evaluation span: pmi is OK but not best overall)
- Transformations:
  - log returns (except spread which is differenced only)
  - scaling (we mostly work with standardized series)
  - trim scaled data to ± 5 (to regularize singular pandemic data)
- Target: BIP (shifted by forecast horizon and publication lag, see below)
- Explanatories: see the above selection
- Ragged end: see below

Ragged End

- Ragged end: prior to shifting:

	BIP	ip	ifo_c	ifo_exp	ifo_l	ESI	spr
2024-09-01		91.20	83.70	85.90	81.60	89.80	
2024-10-01	900.76	90.80	84.10	87.50	80.80	90.70	
2024-11-01		92.20	83.80	86.80	81.00	89.30	
2024-12-01			82.80	84.80	80.90	86.90	
2025-01-01			82.50	83.20	81.70	88.10	
2025-02-01							

Table: Ragged end

- Findings:
  - Publication lags: lag(BIP)=4, lag(ip) =2, lag(ESI)=1 (same as pmi); the other variables are contemporaneous
  - Real-time data: shift all series to be aligned at sample end
  - Introduce a new 'target' column based on BIP.
  - Target: real-time BIP shifted by sum of publication lag and forecast horizon $h = 3$, i.e., shift=$h + lag$ =7

Real-Time data: Forecast Horizon $h = 3$

- Real-time data matrix: shift of target depends on forecast horizon $h = 3$:

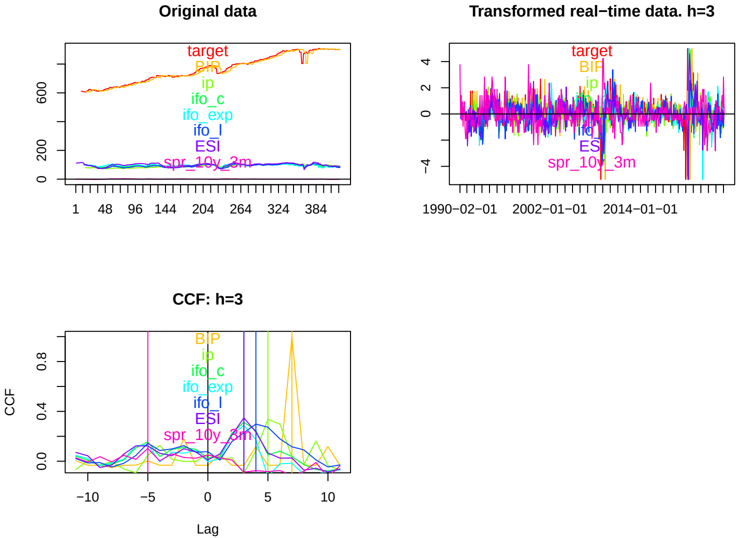
	target	BIP	ip	ifo_c	ifo_exp	ifo_l	
2024-05-01	902.57	904.30	95.10	88.30	89.80	86.90	9
2024-06-01	902.57	904.30	94.60	89.30	90.80	87.80	9
2024-07-01	900.76	904.30	94.70	87.60	88.20	86.90	9
2024-08-01	900.76	901.62	91.70	85.80	87.60	84.10	9
2024-09-01	900.76	901.62	93.40	84.90	86.80	83.00	9
2024-10-01	900.76	901.62	90.70	83.70	85.90	81.60	8
2024-11-01	900.76	902.57	93.20	84.10	87.50	80.80	9
2024-12-01	900.76	902.57	91.20	83.80	86.80	81.00	8
2025-01-01	900.76	902.57	90.80	82.80	84.80	80.90	8
2025-02-01	900.76	900.76	92.20	82.50	83.20	81.70	8

Table: Target column shifted upwards by h+publication lag=7. All other explanatories are shifted to be aligned at sample end.

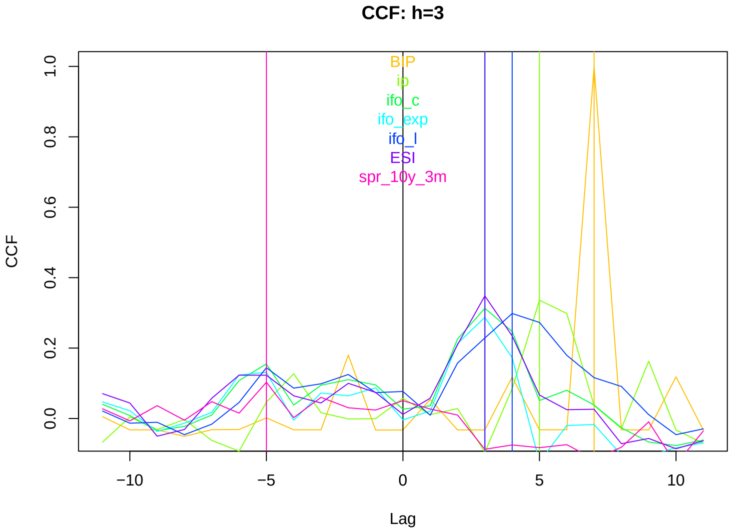
Leads/Lags

- Analysis of leads/lags in real-time data matrix
- Compute cross correlation function (CCF) between target and explanatory variables
  - Peak of CCF indicates lead (left-shift) or lag (right-shift)
  - Note: (shift of) target depends on h

Data and CCF for $h = 3$



CCF for $h = 3$


CCF: h=3
BIP
ip
ifo_c
ifo_exp
ifo_l
ESI
spr_10y_3m
CCF
1.0
0.8
0.6
0.4
0.2
0.0
−10
−5
0
5
10
Lag

Leads/Lags

	Peak correlation
BIP	7.00
ip	5.00
ifo_c	3.00
ifo_exp	3.00
ifo_l	4.00
ESI	3.00
spr_10y_3m	-5.00

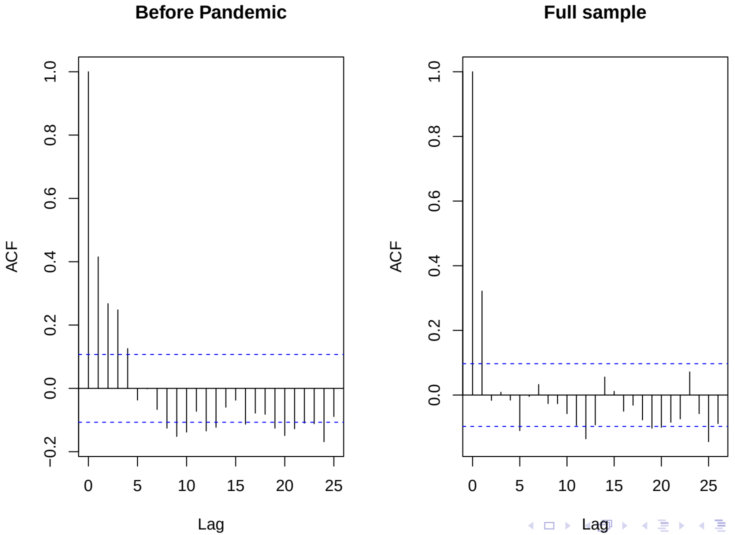
Table: Peak correlation: positive numbers signify a lag of the real-time explanatory (shifted backward by its publication lag) relative to the target (shifted forward by the forecast horizon).

- Stair-step interpolation of BIP makes an assessment of leads/lags less clear (than for ip)

Non-Stationarity (Structural Change?)

- The **dependence** structure of some of the series seems to **change** over time (in particular before/after Pandemic)
- This can have more or less severe impacts on data **modelling**, **filtering**, forecast **performances**, **holding times** (to be defined below),...
- Example: ACF of **ifo-exp** for data prior Pandemic (up to 2019) vs. full data set, see plot on next next

Non-Stationarity/Structural Change in Dependence: ifo-exp



Non-Stationarity/Structural Changes

- Before Pandemic the data has a stronger and longer dependence (ARMA(model)): the data is **smoother**
- On the **full** data set the previous dependence structure simplifies to a simple lag-one dependence (MA(1)-model): the data is **noisier** (higher vola and more frequent zero-crossings)

Section 3

Benchmarks and Direct AR Forecasts

Benchmarks and Direct AR Predictors

- Use data from 1990-02-01 to 2025-02-01
  - Singular readings during pandemic affect OLS estimation
  - Use trimmed scaled data (threshold ± 5)
- Denote $target_t(h)$: first column of real-time data matrix, i.e. real-time BIP shifted forward by lag+h=7
  - Stair-step extrapolation not ideal
- Benchmark 1: mean of BIP, i.e., $\frac{1}{t}\sum_{k=1}^{t} BIP_k$
- Benchmark 2: regression of target on BIP:

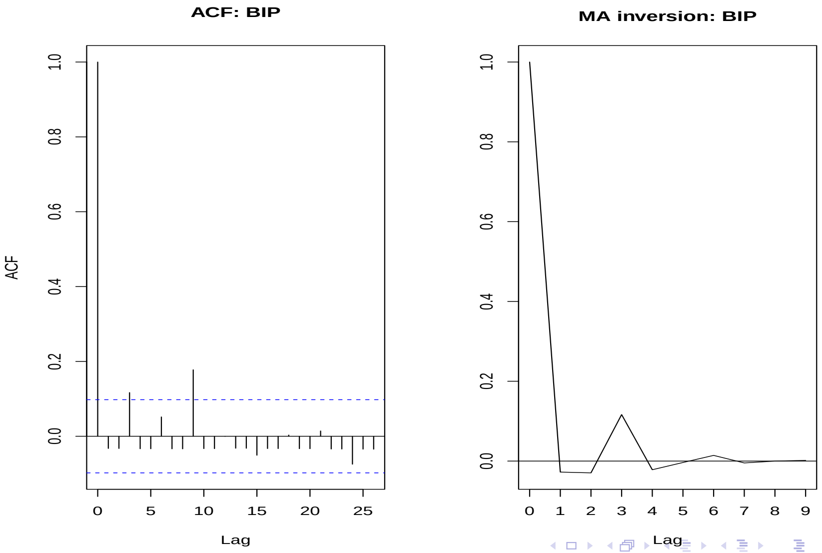
$$target_t(h) = c + a_1 BIP_t + ... + a_p BIP_{t-p} + \epsilon_t$$

- Additional benchmarks: regress any of the real-time indicators on target

ARMA-Models

- For each transformed series we can fit an ARMA-model: $p = 3$, $q = 1$
- The model is not used for direct forecasts (which are based on OLS regression, see previous slides)
- Instead, the ARMA-models are used for computing the real-time filters, see below
- For each model we can obtain its MA-inversion, see next plot in the case of BIP

Benchmark 2: ACF and MA-Inversion of BIP



Findings Benchmark 2

- The MA-inversion of BIP as well as the ACF are close to white noise, i.e., **Benchmark 1**$\approx$**Benchmark 2**.
  - Stair-step interpolation not ideal
  - The regressors in the direct AR predictors are insignificant, see t-values further down
  - This result is reliant on the span for estimation (with/without Pandemic), see below for further details

Forecast horizon

- We here emphasize a one quarter ahead forecast horizon $h = 3$
- Alternative horizons such as $h = 0$ or $h = 12$ lead to similar qualitative outcomes but significance levels (t-tests) are affected

Relative Forecast Performances

- Regress the (real-time) indicator y_{it} on the target to obtain the associated $MSE_i = E[\hat{\epsilon}_{it}^2]$, where $\hat{\epsilon}_{it}$ is the residual of the regression
- Relative forecast MSEs: $MSE_i/MSE(benchmark)$.
- We can consider Benchmark 1 or 2 (mean of BIP or AR forecast based on BIP: virtually the same)
- Numbers smaller one indicate outperformance of the corresponding indicator
- The following table summarizes performances for the **entire** sample (pmi would be in the middle)

Relative Forecast Performances: Forecast Benchmarks, $h = 3$, Full Sample

	Relative MSE h=3
BIP	1.00000
ip	0.99472
ifo_c	1.00007
ifo_exp	1.00054
ifo_l	0.99378
ESI	1.00078
spr_10y_3m	0.99615

Table: Relative mean-square performances against benchmark 2 (which is virtually identical with benchmark 1, the mean of BIP).

Effect of Sample Selection

- The above predictors and the evaluation span rely on the full sample, including Pandemic
- All results are statistically insignificant, see below for details
- What happens if we discard the singular (trimmed) Pandemic?
- The following table reports relative MSEs for data prior to 2019-01-01

Relative Forecast Performances: Evaluation on Data Before Pandemic

	Relative MSE h=3
BIP	1.00000
ip	0.99895
ifo_c	0.96738
ifo_exp	0.97755
ifo_l	0.97750
ESI	0.98354
spr_10y_3m	0.99520

Table: Relative mean-square performances against benchmark 2 (which is virtually identical with benchmark 1, the mean of BIP).

Findings

- Indicators perform less well **after** 2019-01-01 in relative terms
- Ifo series perform best (before Pandemic)
- Difference is statistically significant (see below for more detailed results)

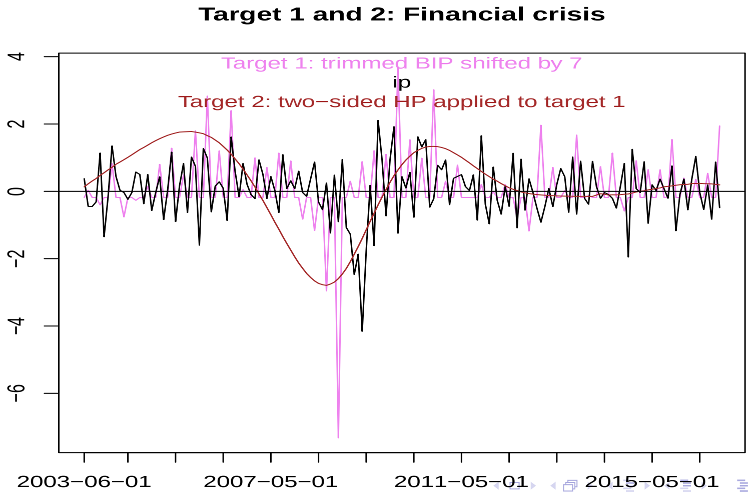
Section 4

Univariate Filters

Alternative Targets

- Instead of looking at BIP-returns shifted by $lag + h$ (7-steps ahead) we might be interested in looking at the **trend** of BIP-returns: *trend-growth*
- Which 'trend'?
  - Ideal trend, model-based trend (SEATS, state-space), Henderson, Hodrick Prescott (HP),...
- Due to its wide use in applications we look at **HP**(14400) (monthly design)
- On the next slide, we plot and compare the two targets: original shifted BIP (violet line) and HP applied to shifted BIP (brown)

Targets 1 and 2: BIP-Shifted (violet), BIP-Trend (brown), ip Real-Time (black): All Series Standardized



Findings

- The filter output (target 2: brown) tracks the (long-term/systematic) trend-growth of target 1 (violet)
  - Business-cycle or **growth-cycle**: depending on how the filter is implemented, a negative sign can be associated with negative economic growth (growth cycle), under-par growth (below potential output), recessions/crises (see SSA paper)
  - The sign of the filtered data, i.e., its *zero-crossings*, have a potentially meaningful *economic* interpretation
- **Problem**: two-sided (acausal) filter cannot be computed in real-time

Causal Filter Design

- We can apply the two-sided HP to each indicator (ip, ifo, ESI, spread,...)
- **Task**: for each indicator compute a *causal* filter which tracks the two-sided HP (applied to that indicator)
  - **MSE** and **SSA** designs: the latter encompasses the former (see SSA paper)
  - We account for h and the publication lag of each indicator (details left aside)
- New **predictors**: MSE and SSA (applied to indicators)
- Potentially interesting **targets**
  - We here continue to look at *BIP-shifted* (violet line in above plot) as our main target of interest
  - Later, we also consider *BIP-trend*, i.e. the two-sided HP applied to BIP-shifted (brown line), as an alternative target

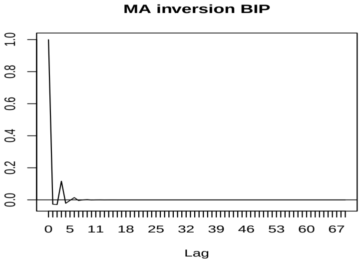
Forecast Performance

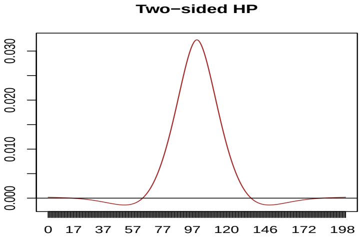
- For each indicator, we compute the mean-square forecast error $E[(BIP_{t+h} - \hat{T}_{ti})^2]$, where BIP_{t+h} is the target series (shifted BIP) and $\hat{T}_{ti}$ is either the MSE- or the SSA-output when applied to the $i-$th indicator
- Note that by design, $\hat{T}_{ti}$ does not target BIP_{t+h}. Instead it tracks the two-sided HP applied to the i-th indicator (accounting for h and publication lag of that indicator)
- **Idea**: if $\hat{T}_{ti}$ tracks the (future) trend-growth of the i-th indicator and if the indicator is informative about the effective target BIP_{t+h}, then $\hat{T}_{ti}$ should also track BIP_{t+h} 'somehow'
  - Validate MSE- and SSA-trends by computing performances against the 'indirect' target BIP_{t+h}

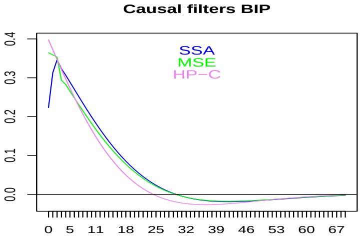
MSE and SSA for BIP

- The figure on the next slide displays
  - The **MA inversion** of the ARMA-model fitted to BIP (top left)
  - The **acausal** two-sided HP (top right)
  - The **causal** filters: MSE (green), SSA (blue) and classic concurrent HP-C (violet)

Univariate Filters for BIP







Relative Forecast Performances

- In the next table, we compare relative forecast performances of all predictors
  - direct **AR** forecasts based on indicators: first column
  - trend forecasts **SSA**: second column
  - trend forecasts **MSE**: third column
- Relative forecast error: all forecast mean-square errors are normalized by AR-forecast based on BIP (benchmark 2).
  - Numbers smaller one signify outperformance of the benchmark (significance is analyzed further down)
- We assume **calibrated** filters, i.e., we regress filters on target to account for (static) level and scale parameters (filters implicitly assume the data to be centered and the scale differs from BIP)

Relative Forecast Performances: $h = 3$ Full Sample

	Direct AR forecasts	SSA	MSE
BIP	1.0000	1.0048	1.0053
ip	0.9947	1.0066	1.0069
ifo_c	1.0001	0.9897	0.9908
ifo_exp	1.0005	0.9859	0.9888
ifo_l	0.9938	0.9947	0.9941
ESI	1.0008	0.9924	0.9945
spr_10y_3m	0.9961	0.9937	0.9931

Table: Relative mean-square performances against benchmark 2 (which is virtually identical with benchmark 1, the mean of BIP).

Discussion

- Calibrated trend forecasts (second and third columns) are generally not worse than 'direct' AR indicator forecasts (first column)
- MSE and SSA are similar (the latter is smoother: less zero-crossings, see below for details)
- **Significance**: t-values in regressions of predictors on target, see next slide

Significance ($h = 3$)

	t_stat direct AR	t-stat SSA	t-stat MSE
BIP	0.44	-0.79	-0.70
ip	1.40	-0.31	-0.08
ifo_c	0.40	2.30	2.22
ifo_exp	0.13	2.54	2.35
ifo_l	1.46	1.93	1.98
ESI	-0.18	2.11	1.95
spr_10y_3m	1.25	2.00	2.06
Consensus	2.11	2.51	2.48

Table: Significance of predictors: t-statistics from a regression of the predictors on the target. Consensus forecast is the equally weighted mean of all forecasts.

Findings

- Direct AR predictors are insignificant (on full sample, including Pandemic)
- SSA and MSE similar (with respect to t-statistics) and some of them are significant
- **Consensus** (equally-weighted) generally perform well

Relative Forecast Performances: Evaluation on Data Before Pandemic

	Direct AR forecasts	SSA	MSE
BIP	1.00	1.00	1.00
ip	1.00	1.01	1.01
ifo_c	0.97	0.98	0.97
ifo_exp	0.98	0.96	0.96
ifo_l	0.98	1.00	0.99
ESI	0.98	0.98	0.98
spr_10y_3m	1.00	0.99	0.99

Table: Relative mean-square performances against benchmark 2: data prior to Pandemic.

Significance: Data prior to Pandemic

	t_stat direct AR	t-stat SSA	t-stat MSE
BIP	0.81	-1.41	-1.37
ip	1.12	-1.21	-0.96
ifo_c	2.98	2.62	2.96
ifo_exp	2.49	3.35	3.67
ifo_l	2.52	1.72	1.98
ESI	2.13	2.68	2.86
spr_10y_3m	1.50	2.24	2.32
Consensus	3.25	2.82	3.14

Table: Significance of predictors: t-statistics from a regression of the predictors on the target. Consensus forecast is the equally weighted mean of all forecasts.

Findings

- **Main effect**: t-statistics **markedly larger** (despite shorter sub-span)
- Direct AR forecasts often strongly significant: performances similar to filters

Alternative Evaluation Criteria

- Until yet we emphasized the mean-square forecast error
- Question: what about **smoothness**, i.e., noise suppression and reliability?
  - Original justification for trend targets: for identical (relative) forecast performances we might prefer a **more reliable/regular** (smoother/less noisy) forecast rule
  - The above results suggest that MSE forecast performances are **not** negatively affected by filtering ('trends')
  - What about **zero-crossings** and shape/dynamics of filter outputs vs. direct AR forecasts
  - And what about **retardation** (lag/right-shift) of trend forecasts?

Zero-Crossings and Holding Times

- **Holding time** (HT): mean duration between consecutive zero-crossings
- **SSA**: impose 50% larger *expected* HT than MSE
  - Dual formulation of SSA **optimization** principle: *maximize HT* for given tracking accuracy (target correlation or MSE)
- Compute and compare *empirical* HTs (larger = smoother):

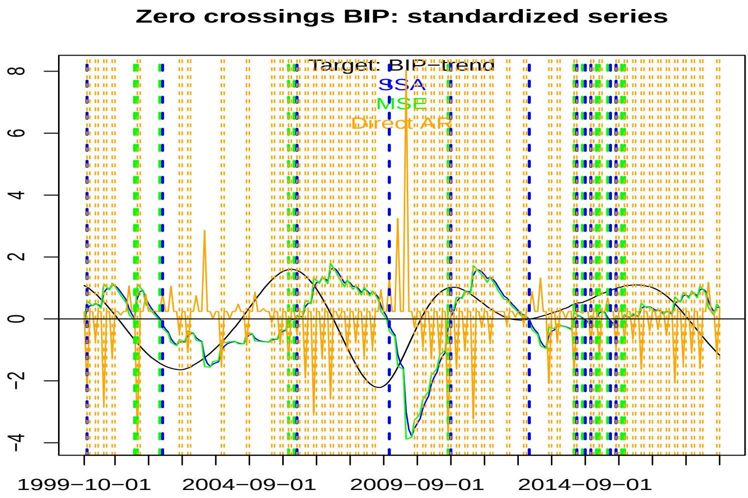
	HT AR	HT MSE	HT SSA
BIP	2.27	19.00	23.38
ip	1.68	7.07	11.26
ifo_c	3.01	11.26	14.48
ifo_exp	2.95	9.81	11.69
ifo_l	2.69	9.81	16.00
ESI	2.64	12.67	19.00
spr_10y_3m	2.81	9.50	12.67

Table: Empirical (sample) Holding times of predictors: direct AR, MSE and SSA.

Zero-Crossings and Holding Times

- Filters much smoother (less crossings) than direct AR forecasts
- Sample HTs of M-SSA approximately 50% larger than MSE (series are centered and therefore expected and sample HTs match better than in ip-case: see corresponding slides)
- On the folowing slide we display AR-forecasts (orange) and MSE- (green) and SSA-filters (blue)
  - Zero-crossings are marked by corresponding colored vertical lines
  - Series are standardized (scaled and centered)

Zero-crossings BIP: Standardized Series



Findings

- See similar comments in ip-case (slides ip)
- Sign of lagged BIP is negative (predictor is not significant): explains 'strange' readings of direct AR predictor (orange) during financial crisis

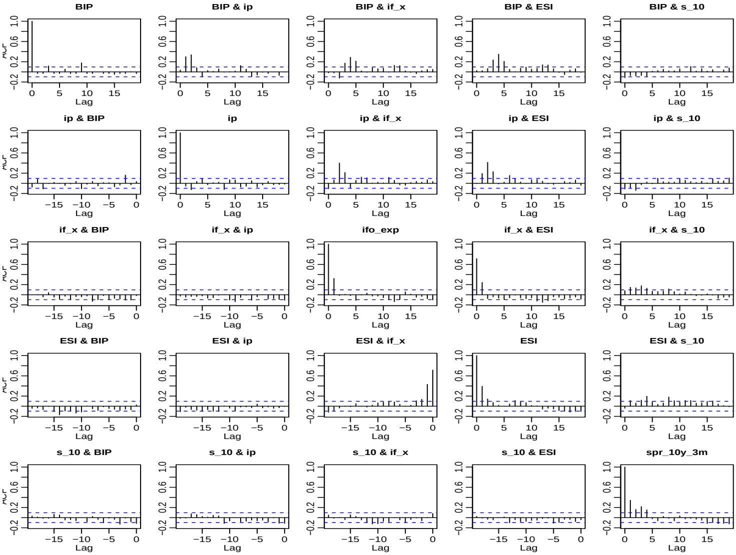
Section 5

Multivariate Filters

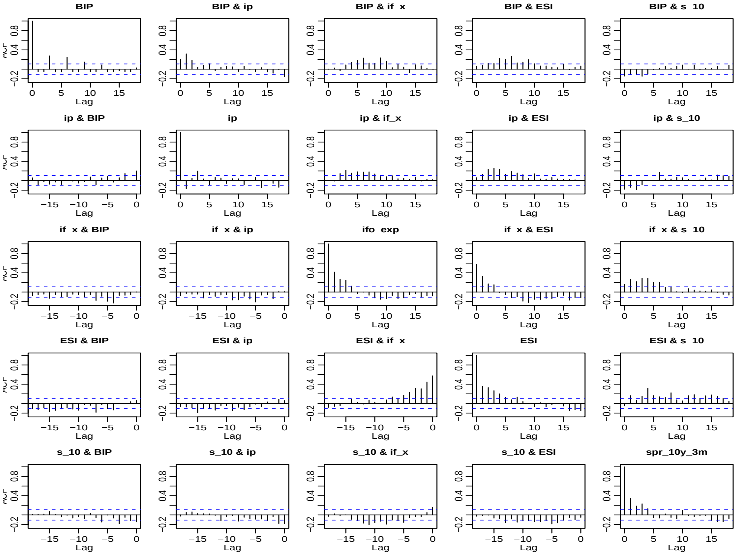
Dependence

- Multivariate M-MSE and M-SSA: rely on several time series to track the two-sided HP
- For the multivariate design we consider BIP, ip, ifo-exp, ESI, spr-10y-3m
  - Criterion for **selection**: the above series differ in terms of leads/lags (see CCF on slide (13))
  - Therefore a multivariate design can exploit the mutually differing (as well as the common) features of each series (see below)
- The next two slides display the **CCF**s corresponding to the full sample (first slide) and data up to the Pandemic (second slide)
  - The CCF refers to the *real-time* data: after accounting for publication lags

CCF: Full Data Set



CCF: Up to Pandemic



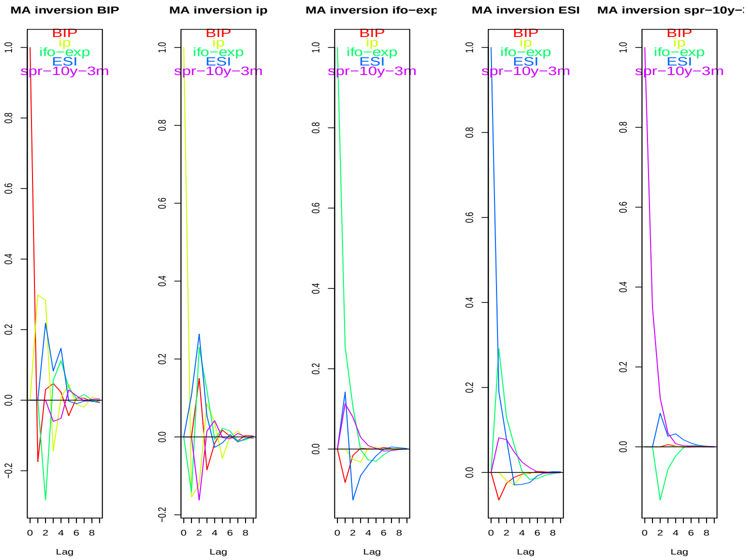
Findings

- Significant ACFs of BIP at multiples of three months: effect of **stair-step** interpolation
- **Non-stationarity**: dependence structure changes after pandemic (structural change?)
- Dependence: BIP cross-correlates with ip, ifo and ESI (first row in above plots).
- Leads/lags correspond to peak-CCF on slide (13) (cross-check)
- Additional findings:
  - **BIP** and **spread** are nearly **uncorrelated**! (at observed leads/lags)
  - **Spread** seems to correlate with (and to lead) **ifo**-series (but not ip). Maybe spread enters somehow into construction of ifo?

VARMA

- Fit a VARMA model to BIP, ip, ifo-exp, ESI, spr-10y-3m
  - Rely on R-package MTS by Ruey Tsay (we also adopt regularization to avoid overfitting)
- Purpose: VARMA model is used for M-SSA and M-MSE (need to account for leads/lags and CCF)
- Estimation span: from first observation to $2024 - 11 - 01$
- Model orders: VARMA(2,0), see below for details
- The (V)MA-inversion of the VARMA is displayed on the next slide

MA Inversion of VARMA



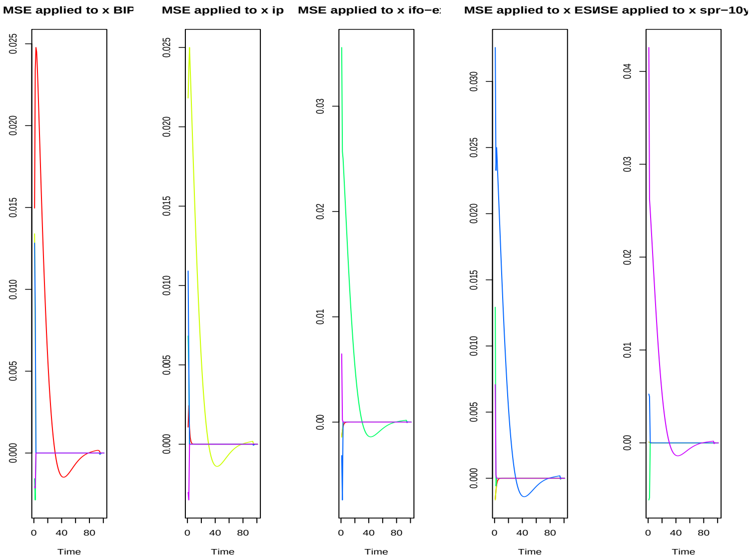
Findings

- Interpretation: see ip-case (slides ip)
- We can **interpret** VARMA model: simple and explainable

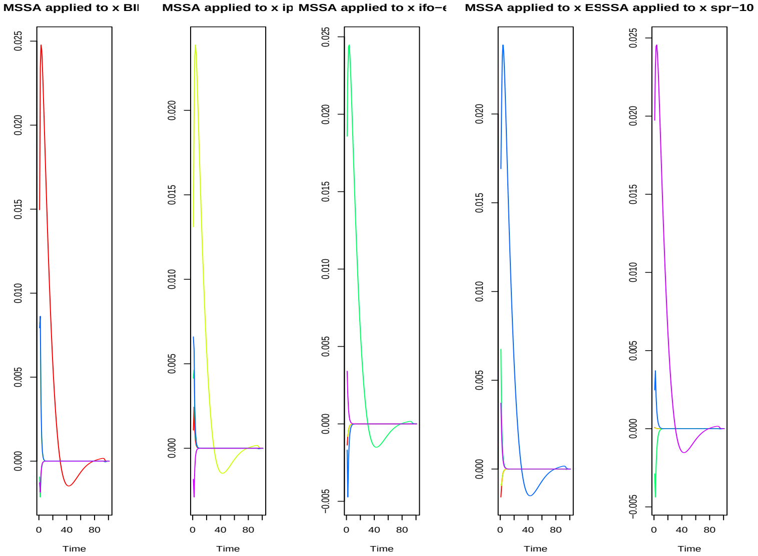
M-SSA and M-MSE Filters

- For each series y_{it}, $i = 1, ..., 5$ of the multivariate filter, the **target** T_{it} is specified by applying the acausal HP to y_{it}
  - **Same trend** targets as in univariate case
  - But each trend target can rely on *multiple* time series for deriving the *causal* filter
  - *VARMA*-model is used for computing the *causal* filters
- M-SSA: impose the **same HT** as in univariate case
- Overall design decisions: compare **apples with apples** (can compare uni- and multivariate designs in a consistent way)
- The following two slides display multivariate MSE (M-MSE) and SSA (M-SSA) filters

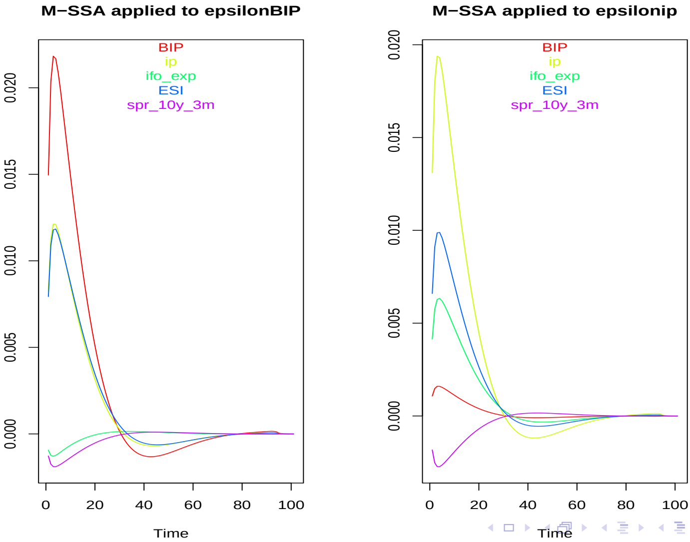
M-MSE


MSE applied to x BIP
MSE applied to x ip
MSE applied to x ifo−e:
MSE applied to x ES
MSE applied to x spr−10y
Time
0 40 80

M-SSA


MSSA applied to x BII
MSSA applied to x ip
MSSA applied to x ifo−e
MSSA applied to x ES
MSSA applied to x spr−10
0.000
0.005
0.010
0.015
0.020
0.025
−0.005
0
40
80
Time

MA-Inversion M-SSA: Filter as Applied to VARMA-Residuals



Findings

- Interpretation: see ip-case (slides ip)
- Intuitively **appealing** (interpretable) outcome

Relative (Mean Square) Forecast Error: Target is Shifted BIP

- The table on the next slide reports relative mean-square forecast errors
- Specifically:
    - Our **target** for forecast evaluation is the shifted BIP
    - The $i-$th output of the multivariate filter is designed to track the *acausal HP* applied to the *i−th indicator*, $i = 1, ..., 5$
        - By construction, M-MSE and M-SSA do not explicitly 'see' or target the shifted BIP.
        - Nevertheless, we expect that the filter outputs should track BIP 'somehow' (see discussion for univariate case above)
- Below, we extend the analysis to *acausal HP* targets.

Relative (Mean Square) Forecast Error (Full Sample)

	Relative MSE	t-statistic
M-SSA BIP	1.01	-0.64
M-SSA ip	1.00	0.07
M-SSA ifo_exp	0.98	2.75
M-SSA ESI	0.98	2.35
M-SSA spr_10y_3m	0.99	2.01
M-MSE BIP	1.01	-0.67
M-MSE ip	1.00	0.02
M-MSE ifo_exp	0.98	2.67
M-MSE ESI	0.99	2.20
M-MSE spr_10y_3m	0.99	2.03

Table: Relative mean-square forecast performances and t-statistics of multivariate designs (target is shifted BIP)

Relative (Mean Square) Filter Error: Target acausal HP

- Findings: the above performances are **comparable** to the previous univariate designs, see slide (40)
- Question: what is the **added-value** of multivariate designs?
- For this purpose we need to consider alternative **trend-targets**
- What is the mean-square **filter error** with respect to the **acausal HP** (applied to each indicator)?
- The table on the next slide computes the relative mean-square *filter errors* of *multivariate* against *univariate* filters
- **Added-value** of multivariate designs: numbers smaller 1 signify an outperformance by the multivariate design (with respect to the *trend* targets)

Relative (Mean Square) Filter Error

	M-MSE	M-SSA
BIP	0.91	0.91
ip	0.90	0.89
ifo_exp	0.93	0.93
ESI	1.04	1.05
spr_10y_3m	1.06	1.08

Table: Relative mean square filter errors: ratio of multivariate over univariate filters (number smaller one signify outperformance by multivariate designs). Targets of uni and multivariate filter are identical but multivariate designs can rely on additional explanatory series

Findings

- Findings similar to ip-case (slides ip forecasting)
- However, M-SSA performs worse than SSA for ESI and spread. Explanations:
  - For spread, M-SSA is worse because it is smoother: sample HT of M-SSA is substantially larger (not shown)
  - In the case of ESI, M-SSA is worse probably because of sample variations (and model misspecification: but the latter affects all (M-)SSA filters)
  - In fact M-SSA trend for ESI 'looks better' (not shown): it is slightly smoother and faster (left-shifted) than SSA because the multivariate filter relies mainly on ifo-exp (as additional explanatory variable), which is leading, due to its smaller publication lag

BIP-Shifted and BIP-Trend Targets

- Up to now, we considered different targets in this study: **shifted BIP** and two-sided **HP-trend** applied to each of the different indicators
- In the following evaluation we restrict this choice by emphasizing **BIP**. Moreover we consider an alternative performance measure.
- Evaluation framework:
  1. Two targets: BIP-shifted and BIP-trend, i.e., acausal HP applied to BIP-shifted. In particular, we discard HP applied to the other indicators.
  2. Performance measure: instead of mean-square forecast performances we compute the (sample) **target correlations**, i.e., the (sample) correlations between the outputs of M-MSE (or M-SSA) and each one of the two targets

BIP-Shifted and BIP-Trend Targets

- The table on the next slide reports corresponding *sample target correlations* for M-MSE and M-SSA against BIP-trend (first two columns) and shifted BIP (last two columns)
  - We consider the 5 filter outputs (the 5 rows in the table) corresponding to BIP, ip, ifo-exp, ESI, spr-10y-3m in M-MSE and M-SSA as predictors for each one of the two targets
  - Note that M-SSA (or M-MSE) outputs corresponding to ip, ESI, ifo-exp and spread (rows 2-5 in the table) do not target BIP-shift or BIP-trend explicitly (no overfitting)

Sample Target Correlations

	M-MSE HP	M-SSA HP	M-MSE BIP	M-SSA BIP
BIP	0.21	0.17	-0.04	-0.04
ip	0.30	0.28	0.00	0.00
ifo_exp	0.46	0.47	0.15	0.16
ESI	0.66	0.66	0.13	0.13
spread	-0.14	-0.09	0.12	0.11

Table: Target correlations: correlations of M-MSE and M-SSA filter outputs with BIP-trend (first two columns) and with BIP-shifted (last two columns).

Findings

- In general it is **'easier'** to track the (shifted) BIP-**trend** target than the (shifted) BIP: correlations in the first two columns (of the above table) tend to be larger
- Interestingly, the filter outputs corresponding to ip, ESI and ifo-exp correlate significantly with (shifted) BIP-trend (first two columns, rows 2-4), despite the singular Pandemic
- Our findings provide **additional justification** for considering **trend** targets (instead of shifted data)
- In particular the correlations in rows 2-4 (first two columns) suggest that the trend outputs corresponding to ip, ESI and ifo in the multivariate filter are *candidates* for predicting the future trend-growth of BIP
- On the next slide we compare target correlations of M-SSA and SSA with respect to BIP-trend

Sample Target Correlations Against BIP-trend : M-SSA vs. SSA

	SSA HP	M-SSA HP
BIP	0.08	0.17
ip	0.15	0.28
ifo_exp	0.50	0.47
ESI	0.66	0.66
spr_10y_3m	-0.11	-0.09

Table: Target correlations: SSA vs. M-SSA.

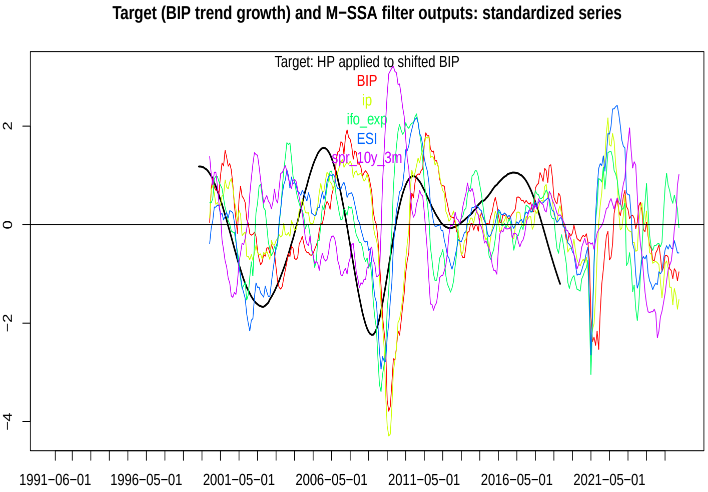
Findings

- Multivariate filters of lagged series (first two rows) improve over univariate designs: M-SSA can rely on 'faster' explanatory series (this information is missing in univariate designs)

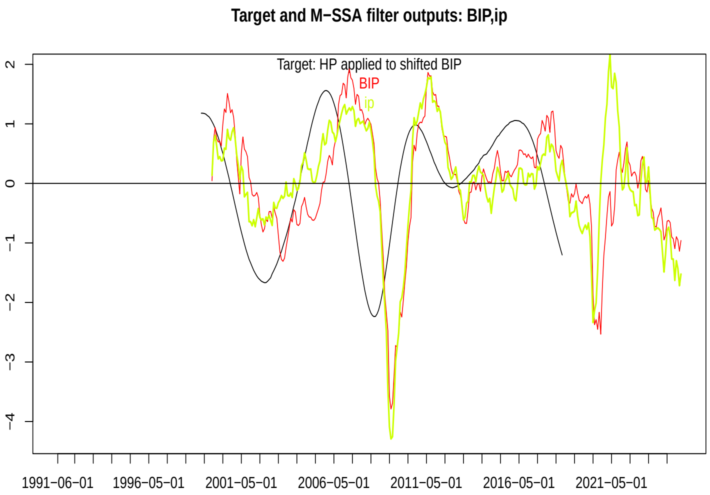
M-SSA Outputs as Predictors for BIP-Trend

- In the following five slides we display the M-SSA filter outputs together with BIP-trend (black line)
  - The first plot displays all M-SSA outputs together with the target (black line)
  - The second plot emphasizes M-SSA for BIP and **ip** (together with the target)
  - The third plot emphasizes M-SSA for BIP and **ESI** (together with the target)
  - The fourth plot emphasizes M-SSA for BIP and **ifo-exp** (together with the target)
  - The last plot emphasizes M-SSA for BIP and **spread** (together with the target)
- For ease of visual inspection, all series are **standardized** (standardization is equivalent to filter calibration)

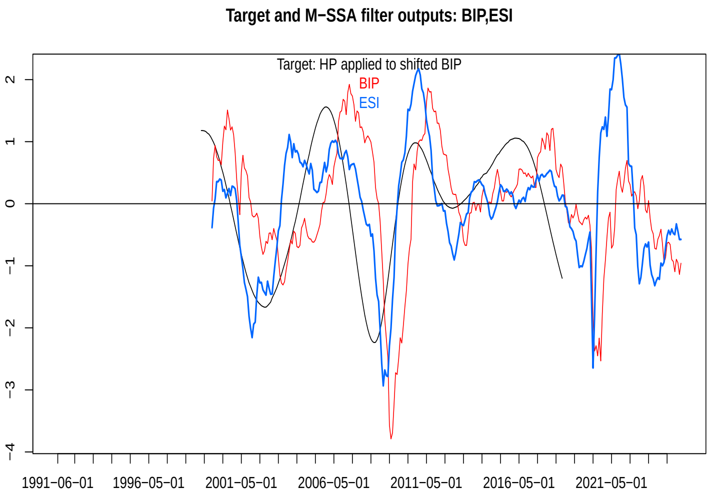
BIP-Trend (black) and All M-SSA Outputs



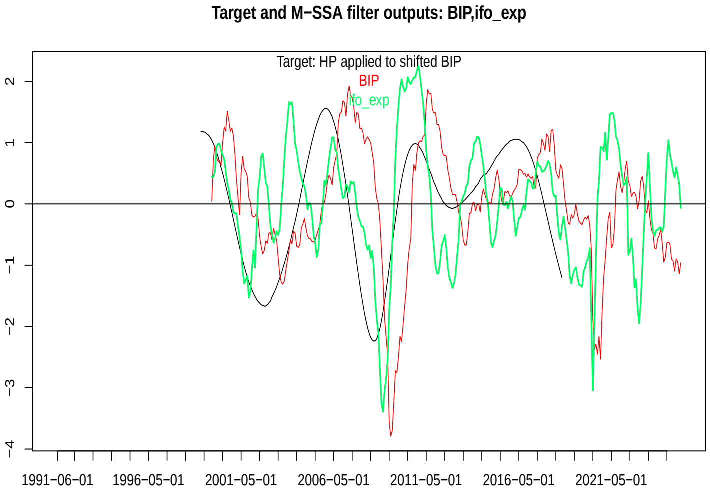
BIP-Trend (black) and M-SSA: BIP and ip Outputs



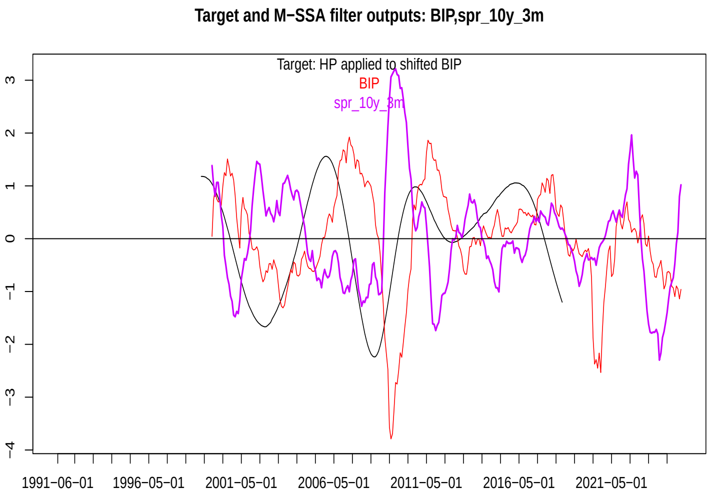
BIP-Trend (black) and M-SSA: BIP and ESI Outputs



BIP-Trend (black) and M-SSA: BIP and ifo-exp Outputs



BIP-Trend (black) and M-SSA: BIP and spread Outputs



Findings

- M-SSA BIP (red) is lagging target (black) due to its (largest) publication lag
- M-SSA ip (yellow-green) is also lagging but less than M-SSA BIP
- M-SSA ESI (blue) is both 'fast' (left-shifted) and smooth: this explains the large target correlation reported in the above table(s)
- M-SSA ifo-exp (green) is fastest but overlaid with 'noisy' cycles
- Spread (violet) does not seem to correlate with target (black)

Section 6

Addressing Retardation (and Smoothness)

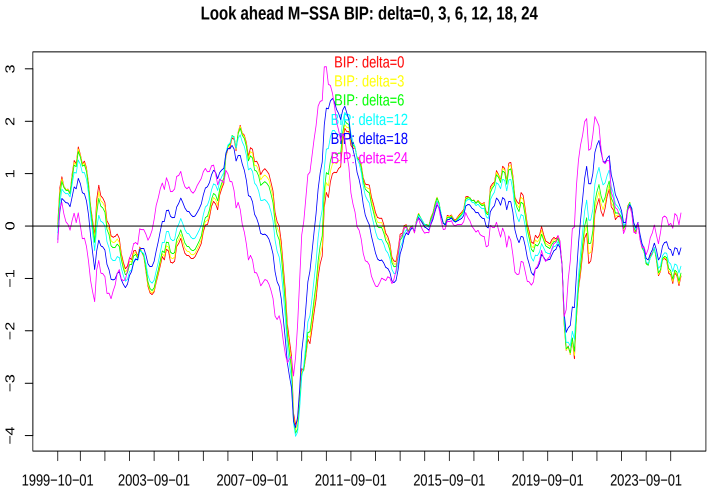
Addressing Retardation and Smoothness (HT)

- Stronger smoothing generally signifies a more pronounced retardation (right shift; lag): dilemma
- *Retardation* of our filter designs can be addressed by *increasing the forecast horizon* h to $h+\delta$, where $\delta \geq 0$
  - We here illustrate filters with *forecast excesses* $\delta = 0, 3, 6, 12, 18$ and 24 ($h = 3$ is maintained)
- In general, **increasing** δ leads ceteris paribus to a **left-shift** of the predictors (smaller retardation) and to **noisier** filters with **smaller HT** (more zero-crossings).
- However, in the **M-SSA** framework we can control and **fix the HT**: the M-SSA predictor will maintain a fixed HT irrespective of δ
- Therefore, we can **maintain smoothness** while **mitigating retardation** to some extent (by extending the classic forecast dilemma to a more general forecast trilemma, see SSA paper)

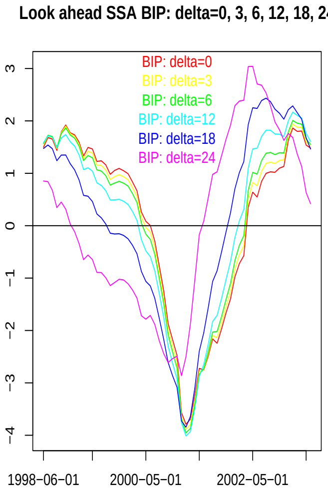
Addressing Retardation and Smoothness (HT)

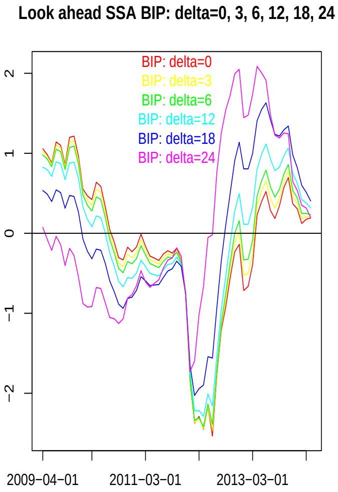
- In the previous slides we saw that M-SSA-BIP (the output of M-SSA corresponding to the shifted BIP-target: red lines) was lagging
- The next two plots illustrate the effect of $\delta = 0, 3, 6, 12, 18, 24$ on M-SSA-BIP for the whole sample (first plot) and for the financial crisis and the pandemic (second plot)
  - We **fix** the target, i.e. M-SSA-BIP, and we **vary** δ
  - Later, we will compare **all** M-SSA outputs (not only M-SSA-BIP) for $\delta = 12$ **fixed**
  - Finally, we will analyze **all combinations**: all selected M-SSA outputs for all selected δ
  - Note: we do not display the M-MSE outputs which are noisier (more 'false alarms'), see the ip-case for reference
- For ease of visual inspection we apply **standardization** (equivalent to filter calibration)

Look Ahead: Effect of δ on Advancement (Left Shift) of M-SSA-BIP



Financial Crisis (left) and Pandemic (right)





Findings

- The main effect of $\delta = 3, 6, 12, 18, 24$ on **M-SSA-BIP** is a **left-shift** of the filter output
  - A larger δ means that zero-crossings at onset and end of crises are detected earlier(advancement, left-shift)
  - Note that M-SSA keeps the (expected) HT fixed for increasing δ (in contrast to M-MSE which becomes noisier, see slides of ip-case)
  - The filters are also subject to zero-shrinkage for increasing δ (but this effect is masked by the standardization)
- We report **sample HTs** for M-MSE and M-SSA filters as a function of δ: see the table on the next slide

Holding-Times

	delta=0	delta=3	delta=6	delta=12	delta=18
M-MSE	9.20	9.80	8.70	7.60	6.90
M-SSA	16.00	13.20	13.80	16.90	19.00

Table: Holding times of M-MSE and M-SSA for various values of look ahead parameter δ

- The above HTs of M-SSA should be compared with the univariate SSA on slide (46), first row, last column: by construction expected numbers should match (sample HTs are subject to random deviations)

Findings

- M-SSA is substantially smoother (less crossings, i.e., less false alarms)
- The *expected* HT is independent of δ
  - The differences in the above *sample* HTs (of M-SSA) reveal mainly *finite sample* variations
  - The *sample* HTs converge to the *fixed expected* value in sufficiently long samples (see SSA tutorial)

Effect of δ on Target Correlation and Sign Accuracy

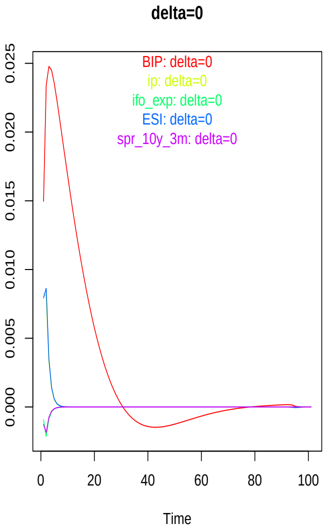
- We analyze the effect of δ on the target correlation and the sign accuracy, see the following table.

	Target correlation	Sign accuracy
delta=0	0.17	0.62
delta=3	0.22	0.63
delta=6	0.27	0.64
delta=12	0.38	0.68
delta=18	0.56	0.77
delta=24	0.69	0.77

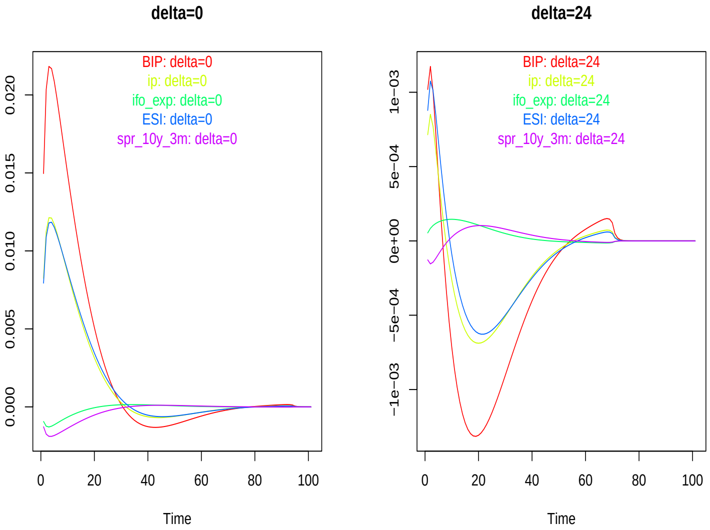
Table: Target correlations (first column) and sign accuracy (second column) of M-SSA-BIP for various $\delta \geq 0$ values

- Note: the target correlation for $\delta = 0$ (first row, first column) corresponds to slide (72) (first row, second column)

Effect of δ on Filter Weights



Same as Previous Slide but MA-Inversion of Filter



Findings

- The original M-SSA-BIP, based on $\delta = 0$, was lagging (due in part to VARMA model misspecification)
- Increasing δ leads to substantial improvements of *target correlation* and *sign accuracy*
- Increasing δ signifies a **left-shift** of the filter outputs: this left-shift *explains* mainly the improvement
- Increasing δ induces a zero-shrinkage of the filter weights and a modification of the general decay-pattern, see the plots just above
  - The shrinkage of the filter coefficients does not affect *standardized* filter outputs
  - The different decay pattern affects the left-shift of the (standardized) filter outputs; but it does not affect the (expected) HT

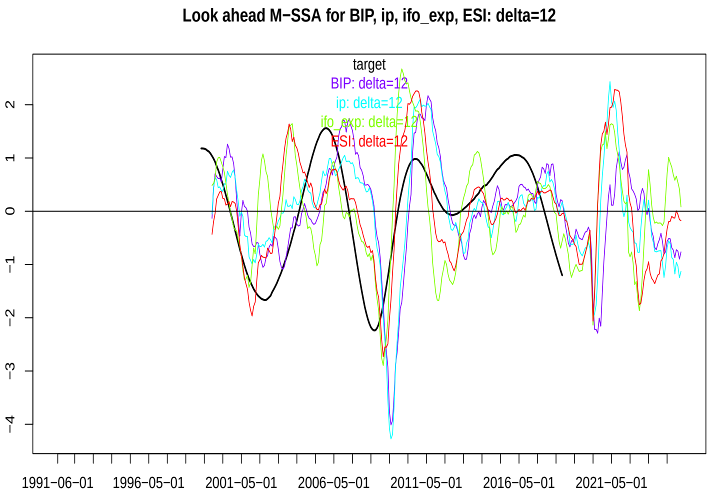
Variable (All) M-SSA Targets and Fixed Forecast Excess

- In the above slides we showed the effect of δ on the M-SSA-BIP output
- We now consider **all** M-SSA outputs for **fixed** $\delta = 12$ (all M-SSA outputs for $\delta = 0$ were analyzed previously)

All M-SSA Outputs for Fixed $\delta = 12$

- In contrast to the previous slides, we now **fix** $\delta := 12$ and we analyze **all** M-SSA outputs for that δ
- On the next slide we now display and compare the corresponding look-ahead M-SSA outputs for $\delta = 12$
- For reference, we also include BIP-trend, i.e., the acausal HP applied to BIP-shifted (black line)
- For easier visual inspection all series are standardized (equivalent to filter calibration)

All M-SSA Outputs for (Fixed) $\delta = 12$



Findings

- M-SSA outputs for ESI (red), ifo-exp (green), ip (cyan) and BIP (violet) differ mainly during the financial crisis, where especially the former two are left-shifted and faster
    - The financial crisis cannot be reconciled with the VARMA-model (misspecification) which explains the lesser performance of the BIP M-SSA output in all previous comparisons (details omitted).
- Towards the sample end, M-SSA ifo-exp (and to some extent M-SSA ESI) suggest evidence of a **recovery** whereas M-SSA ip and BIP still indicate negative growth.

Sign Accuracies

- In order to complete our evaluation metrics, the table below reports **sign accuracies** of the M-SSA outputs, i.e., the (sample) probability that a predictor matches the sign of the target (BIP-trend, black line in previous plot)

	BIP	ip	ifo_exp	ESI	spr_10y_3
sign acc.	0.68(0.03)	0.68(0.03)	0.65(0.03)	0.74(0.03)	0.39(0.03

Table: Sign accuracies of M-SSA outputs for fixed $\delta = 12$ with standard errors in parentheses

- For fixed $\delta = 12$, ESI (red line in previous plot) is best, followed closely by ip, BIP and ifo-exp. The sign accuracy of spread is below 50%

All Combinations of Forecast Excess and M-SSA Outputs

- Up yet we considered specific combinations of forecast excess and M-SSA outputs
- The tables in the following three slides report sample *sign-accuracy*, sample *target correlation* and sample *holding-times* for **all combinations** of M-SSA outputs and δ
  - Sign-accuracy and sample correlation are evaluated against BIP-trend, i.e., acausal HP applied to BIP-shift

Sample Sign Accuracy

	BIP	ip	ifo_exp	ESI	spr_10y_3m
delta=0	0.62	0.64	0.68	0.74	0.43
delta=3	0.63	0.65	0.67	0.73	0.42
delta=6	0.64	0.68	0.67	0.74	0.42
delta=12	0.68	0.68	0.65	0.74	0.39
delta=18	0.77	0.74	0.61	0.67	0.36
delta=24	0.77	0.71	0.50	0.52	0.42

Table: Sign accuracies of all combinations of M-SSA outputs (columns) and forecast excesses δ (rows): the sign accuracy is referenced against BIP-trend

Sample Target Correlation

	BIP	ip	ifo_exp	ESI	spr_10y_3m
delta=0	0.17	0.28	0.47	0.66	-0.09
delta=3	0.22	0.31	0.45	0.65	-0.10
delta=6	0.27	0.34	0.43	0.63	-0.12
delta=12	0.38	0.41	0.38	0.58	-0.18
delta=18	0.56	0.52	0.23	0.41	-0.31
delta=24	0.69	0.54	-0.10	-0.02	-0.40

Table: Target correlations of all combinations of M-SSA outputs (columns) and forecast excesses δ (rows): the target correlation is referenced against BIP-trend

Sample HT

	BIP	ip	ifo_exp	ESI	spr_10y_3m
delta=0	16.00	11.26	11.69	19.00	25.33
delta=3	13.22	11.26	12.67	21.71	25.33
delta=6	13.82	10.48	12.67	19.00	19.00
delta=12	16.89	12.67	11.69	19.00	21.71
delta=18	19.00	15.20	11.69	14.48	13.82

Table: Sample holding times of all combinations of M-SSA outputs (columns) and forecast excesses δ (rows)

- Remarks:
  - Expected HTs change along columns (from left to right) but are fixed along rows (top to bottom)
  - Sample HTs in the table follow this pattern (though subject to random variation)

Findings

- M-SSA outputs of lagging series (such as BIP and ip: first two columns in above tables) perform better for larger δ
- M-SSA outputs of coincident series (ESI, ifo-exp, columns 3 and 4) perform best for small values of δ
  - M-SSA ifo-exp is systematically outperformed by M-SSA ESI
- The above findings are intuitively appealing/interpretable
- M-SSA output of spread is not a good predictor of the target
- **Summary**: we can combine M-SSA **BIP** and **ip** with **large** δ and M-SSA **ESI** for **small** δ

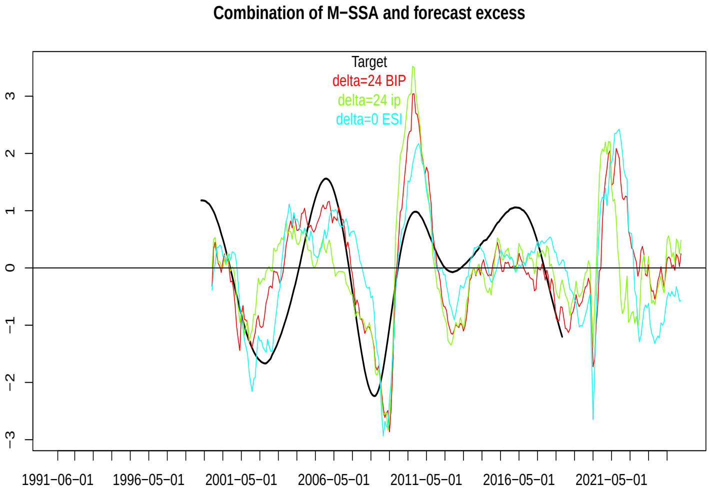
Section 7

(Construction of a) Forward-Looking Consensus BIP-Predictor

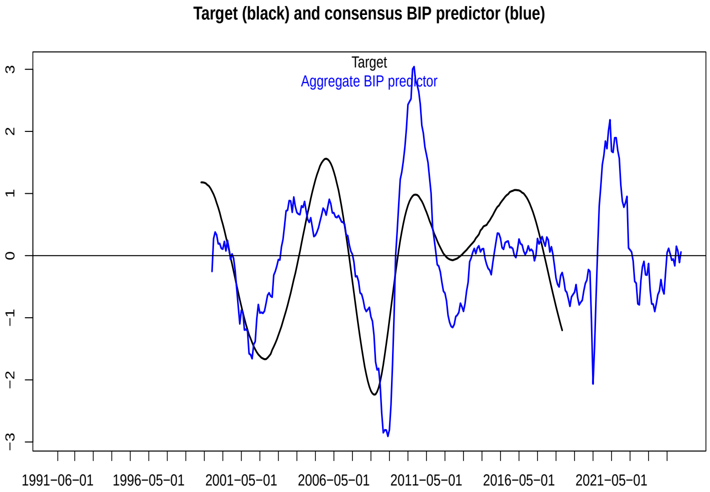
Findings

- As claimed, we can combine M-SSA **BIP** and **ip** with **large** δ and M-SSA **ESI** for **small** δ to construct a forward-looking ($h = 3$) BIP-predictor
  - The procedure cyn be extended to arbitrary h (hyperparameter)
- Given similar performances of the selected indicators we rely on an *equally-weighted* 'consensus' predictor
  - Equal-weighting of *standardized* predictors (standardization accounts for zero-shrinkage)
- On the following two slides we display the selected (standardized) indicators and their cross-sectional mean,i.e., the **forward-looking consensus BIP-predictor**

BIP-Trend (black) and Selected Predictors



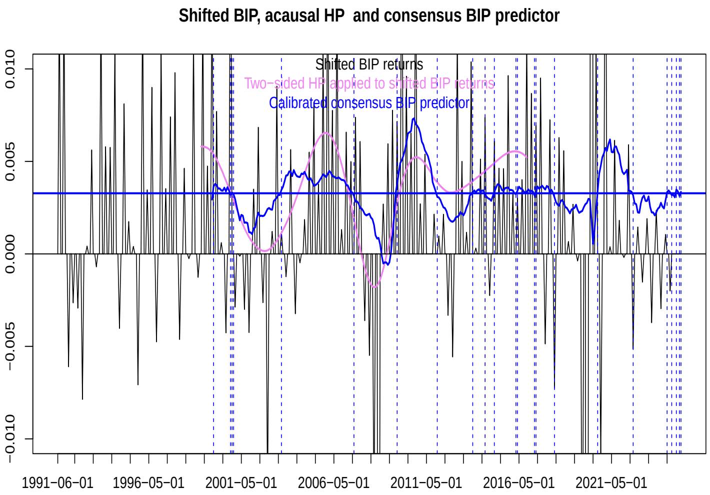
BIP-Trend (black) and Forward-Looking Consensus BIP-Predictor



Calibrated Forward-Looking Consensus BIP-Predictor

- We can also calibrate the predictor on level and scale of BIP-returns by simple linear regression, see next slide
- We display
  - **Quarterly BIP** returns shifted forward by $h = 3$ (black): interpolated with zero for months 2 and 3
  - **Two-sided HP** applied to shifted BIP (violet). We multiply HP by 3 to match the *quarterly growth* scale of BIP
  - **Calibrated consensus** indicator with long-term mean growth (horizontal blue line) and **mean-crossings** (vertical blue lines)
- Duration between mean-crossings is HT=13.82 (≈mean of HTs of selected indicators)

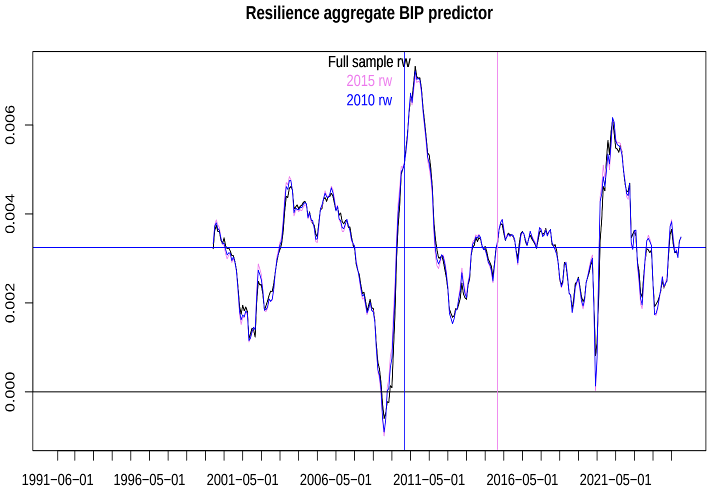
Calibrated Forward-Looking Consensus BIP-Predictor



Sensitivity to Estimation Sample

- Sources of revisions of the 'Calibrated Forward-Looking Consensus BIP-Predictor' are due to calibration (mean and level), to the VAR-model and to the data (BIP/ip)
- On the following slide we briefly analyze the effect of the first two source (calibration and estimation sample)
- Specifically, we recompute the VAR and recalibrate the resulting predictor based on full data (NOv-2024), data prior to 2015 and data prior to 2010

Sensitivity to Estimation Sample



Section 8

Conclusion

Conclusion

- Models and performances are generally reliant on the data sample (for estimation or evaluation): 'Pandemic' effect
- Trend targets can be interpreted meaningfully in terms of economic dynamics (recessions/expansions)
- Trend targets are easier to forecast (less noisy), see slide (71)
- Multivariate designs outperform mainly with respect to trend targets and lagged series (BIP and ip), see slide (66)
- Target correlations, sign accuracies and HT are potentially interesting alternative performance metrics
- Target correlations and sign accuracies referenced against BIP-filter are strongly significant (not random) even on full sample (including singular Pandemic), see slides (70), (72), (89) and (100) and (97) (exception: filtered spread)

Conclusion

- Filters (uni and multivariate) are similar to direct AR forecasts in terms of forecast mean-square error with respect to BIP-shifted, although they do not fit the target explicitly (no overfitting), see slides (40) and (64)
- Real-time indicators such as ESI and ifo (any of the selected ifo series) provide additional information due to *smaller publication lags* (left shifted)
- Spread is **ambiguous**:
  - It does not appear to correlate systematically with BIP, see CCFs on slide (52)
  - The target correlation of the corresponding M-SSA output with BIP-trend is slightly negative, see slide (72): this outcome is confirmed by the plot on slide (79)
  - The t-statistics on slide (40) suggest that the univariate filter based on spread tracks BIP-shifted (not BIP-trend) significantly
  - But the same univariate filter is insignificant when regressed on BIP-trend (value of the t-statistic is 1.3)

Conclusion

- (M-)SSA can control the terms of the forecast trilemma and address user/forecast priorities
  - The hyperparameters (δ, HT) (HT-constraint of M-SSA) allow to fine-tune priorities
  - M-SSA generalizes classic MSE-paradigm
- Possibility to combine various predictors:
  - Fast but noisier: early warning, look ahead indicator
  - A bit slower but more reliable: nowcast, flash indicator
  - Very smooth/reliable but retarded: confirmatory indicator
- Forecast combinations: direct AR with uni- and multivariate filters

Summary of (and Links to) Sample Performance Metrics

- Relative **forecast mean-square** error (target: BIP-shifted): slide (25) (direct AR forecasts full sample), (27) (same but shorter sample), (38) (univariate filters, full sample), (42) (same but before Pandemic), (64) (multivariate filters)
- Relative **filter mean-square** error (target: BIP-trend): (66) (outperformance of multivariate over univariate)
- Target **correlations**: (70), (72), (89) and (100)
- **Sign accuracies**: (97) and (99)
- **t-statistics**: (40), (univariate filters and direct AR, full sample), (43) (same but prior pandemic), (64) (multivariate filters)
- **Holding times**: (46) (direct AR and univariate filters), (87) and (101) (M-SSA vs. M-MSE)

Design Settings (Hyperparameters) in this Study

- Data: selected indicators
- Estimation sample (for models): full data vs. ante Pandemic
- Evaluation sample (forecast errors): full data vs. ante Pandemic
- (V)AR(MA) model orders: ARMA(3,1), VAR(2)
- Trend target: HP(14400)
- (M-)SSA settings: increase of HT by 50%, forecast excess $\delta = 0, 3, 6, 12, 18, 24$
- Forecast horizon: $h = 3$

Design Settings (Hyperparameters) in this Study

- Random-walk or linear interpolation of BIP: default is random walk (na.locf in read function)
- Use two-sided HP as filter target when evaluating sample mean-square filter errors (=TRUE)
  - One can rely on two-sided HP or on symmetrized HP-C: default is two-sided HP (T)
- Use two-sided HP (or one-sided HP-C) as (M-)SSA/(M-)MSE targets in (M-)SSA optimization (=TRUE)
  - One can rely on two-sided HP or on one-sided HP-C: default is two-sided HP (T)
  - One-sided HP-C is a potentially interesting target (HP-C has some interesting characteristics as a causal filter)

Sources of Revisions

- Data
- Models: (V)ARMA
- Calibration: ‘static’ level and scale parameters

To Dos

- Mixed-frequency approach for BIP
- Use all series (extension to dimension reduction and/or aggregation techniques)
- **Better VAR(MA)**-model
  - The weight assigned to 'faster' indicators (small or no publication lag) in multi-step ahead forecasts is too weak
- Additional forecast horizons: $h = 0$ (nowcast), $h = 12$ one-year forecast.
- Alternative holding-time constraints (smoothness)
- Alternative filter specification(s): stronger suppression of 'noisy cycles' of ifo-exp in previous plots?
- 'More' (forecast) combinations
- Portfolio (of predictors) construction: emphasize fast/smooth tradeoff (instead of classic return/risk approach)
- R-package, automatic up-dating with new data,...

Next Steps

- Worth pursuing?
  - Refresh (latest data)
- Conceptual: **priorities**
  - Target, forecast horizon, 'consensus' vs. separate predictors, ...
- **Method**: VAR, mixed-frequency, dimension reduction,...
- Work with **R-code** (cross-check, ROC,...): Github?
- **Paper**
- **Application** (e.g. shiny app, public domain,...)