The Role of Indicators' Selection in Nowcasting Euro Area GDP in Pseudo Real-Time

A. Girardi[1] R. Golinelli[2] C. Pappalardo[1]

[1]ISTAT
[2]University of Bologna

14th IWH$-$CIREQ Macroeconometric Workshop
Halle, December 2-3, 2013

Istat
Istituto Nazionale di Statistica

Forecasting with indicators

- **Focus**: to compare the performance of alternative approaches to nowcasting euro area GDP using several ways of indicators' selection.
- **Aim**: to extract the most valuable indicators from a flow of data issued every month.
- **Targeting predictors** (by pre-selecting information through pre-screening rules) is an effective way to improve forecast performance (Kim and Swanson, 2013; Bulligan et al., 2012).

Widely used methods

- Widely used methods for the extraction of reliable signals from noisy high frequency indicators:
- Bridge models (**BM**): link the forecast target to suitable indicators **selected a priori**. Baffigi et al. (2004), Diron (2008), Grassmann and Keereman (2001), and Rünstler and Sedillot (2003).
- The factor-based models (**FM**): **statistical summary** of **many** indicators through the extraction of some common factors. Angelini et al. (2011), Banbüra and Rünstler (2011), iannone et al. (2008, 2009), Marcellino et al. (2003), Rünstler et al. (2009).

Criticism to both approaches

- **BM**: may appear excessively *ad hoc* because of the strong exclusion restrictions;
- **FM**: Factors may be biased by unbalanced sources of information (Boivin and Ng, 2006). FM extracts factors blindly: as N increases, the average common component to explain the target could be smaller.
- **Remedy**, among the others, **to FM**: to use factors extracted from fewer but informative indicators (**targeted indicators**) can yield gains in terms of forecasting accuracy (compared to using large indicators' dataset).

Pre-screened Factor Models (PFM)

- Pre-screening before factor extraction leads to a mixed approach between BM and FM, which we will refer to as **pre-screened FM** (PFM).
- Compared to BM, PFM is less arbitrary (thresholding rules) and less extreme (k<n<N).
- Compared to FM, PFM extract factors but from a **panel of targeted indicators**, more likely to carry useful information than in FM. Recent literature (Caggiano et al., 2011; Bulligan et al., 2012) has documented the advantage of forecasting with PFM.
- Factors are estimated (in both the FM and PFM approaches) by following the SW method only.

Istat
Istituto Nazionale di Statistica

Empirical framework

$$\Delta y_t = \sum_{k=0}^{q_j} \Gamma_k P_{t-k} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \varepsilon_t$$

- After the selection/extraction of a vector of predictors, FM, PFM and BM are all **represented** by a dynamic model, ARDL (p, q$_j$) between the dependent variable y_t and the vector of predictors P_t.
- Specification is **adaptive** (Swanson and White, 1997).
- Usual reduction procedure applies: sequentially eliminating insignificant regressors and check that the model satisfies some a misspecification and parameter constancy tests (LSE general-to-specific modelling strategy, see Banerjee et al., 2005).
- But, how to select P_t?

Istat
Istituto Nazionale di Statistica

Indicators' selection by *rules*

- Aim: to proceed by ordering and selecting indicators to end up with a dataset of lower dimensions;
- 6 tools for data reduction methods: 1 *hard*- and 5 *soft*-thresholding rules.
- **Hard thresholding rules**: an indicator is selected according to the significance of its correlation coefficient with the target (below a given threshold).
- Shortcoming: it only takes into account the bivariate relationship between the target and each indicator, and disregards the information content of the other indicators. It selects highly collinear targeted predictors.

continue: soft rules

Soft thresholding rules are based on minimization of:

$$\underbrace{min}_{\beta}[\Phi(RSS) + \lambda\Psi(\beta_1, \ldots, \beta_j, \ldots, \beta_n)]$$

RSS of a regression of Y on the retained indicators, λ (the Lagrange multiplier) is the shrinkage parameter (the higher λ, the higher is the penalty for extra regressors), Φ and Ψ are functions of *RSS* and the regression coefficients and lead to:

1. Least angle regressions (LARS);
2. Ridge regression (RIDGE);
3. Least abs. shrinkage selection operator (LASSO);
4. Elastic net estimator (ENET);
5. Forward selection regressions (FWD).

Istat
Istituto Nazionale di Statistica

The procedure to PFM

Step 1 Dataset of N indicators over time span T. Partition T into a first portion, T_1, used for models' training, and (T-T_1) left to run the forecast exercise.

Step 2 Training: each rule runs over a rolling window (T_w<T_1) applied to the initial data (N$\times$(l+1)) to rank the targeted indicators.

Step 3 A binary variable is associated to each indicators and screening rule (=1 if indicator is ranked in the top 10 variables). 7^{th} screening rule: **union of all soft rules** (1= if is indicator selected at least by one soft rule).

- We get, for each indicator and rule, a collection of $(T_1 - T_w + 1)$ binary information.
- **Empirical probability distribution** for the indicators to be selected, conditional on a given rule (seven distributions).

Indicators' selection by *thresholds*

- The quantiles of each distribution are used as an **additional criteria** to pre-screen the indicators.
- For **each pair** (rule, quantile), we have a dataset of targeted indicators: the number of selected indicators decreases as the reference quantile increases.
- Q = [min, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}], where *min* means indicator picked at least once.

Data: Composition

- Forecasting performed using a dataset of short-term indicators for euro area and US economy. It consists of 259 time series ranging from Jan.1990 to Dec.2012.
- Four typologies of indicators:
  - soft indicators (60.2% of total information)
  - hard indicators (11.2% of total information)
  - financial variables (interest rates, stock market indices; nominal and effective exchange rates; public debt; 23.6% of total information)
  - price variables (inflation; energy, commodities; 5% of total information)
- Furthermore, 69.9 per cent of the whole data consists of euro area indicators, 44.4 per cent are at monthly frequency, (32.0 per cent at a quarterly frequency), 72.2 per cent of series are transformed to get stationarity.

Istat Istituto Nazionale di Statistica

Data: Treatment

- Ragged-edge issue: each forecast round is carried out near the half of each month (IP figures for euro area are released).
- Predictions of some missing indicators using auxiliary AR models. We use a four-step procedure:
  - transformation (logs and/or first-differences);
  - univariate (AR) modelling using rolling windows of 132 months (11 years);
  - monthly extrapolation depending on both publication lags and the specific forecast round;
  - quarterly averages and seasonal adjustment.

Data: Vintages and pseudo real-time

- Data are organized to exactly mimic the time schedule of monthly/quarterly indicators along each quarter. Each monthly forecast is classified into 3 cases according to **indicators' timeliness.**
- *First vintage*: (March, June, September, December). Only one month of hard indicators and 2-3 months of soft indicators are known; 51.7% of indicators are not available;
- *Second vintage*: (January, April, July, October). Two months for hard indicators and three months for soft ones are known; 12.7% of indicators are not available;
- *Third vintage*: (February, May, August, November). All the indicators are available for all the months of the GDP quarter to be forecast (*nowcast*).

Training with rolling estimates

- The quarterly data span over 1990q1-2013q1 and consists of 148 monthly vintages. Each vintage spans 44 quarters: the first sample from 1990q1 to 2000q4, the last one from 2002q2 to 2013q1.
- We split the 148 vintages into two parts:
  - first 85 vintages devoted to models' training;
  - the remaining 63 vintages devoted to the assessment of the pseudo real-time forecasting ability;
- The forecasting sample is 2008q1-2013q1.

Summary of models

- For each quantile, we can identify seven different PFM.
- For example, if we set Q = 90^{th} percentile, we select the lowest set of indicators in N. The models that exploit with factors these *n* targeted indicators are labeled PFMhard*90*, PFMlars*90*, ...
- Given that 6 quantiles times 7 pre-screening rules leads to 42 different PFM, we summarize the outcomes by reporting **PFM**hard*xx* and **PFM**soft*xx* forecasting performances.
- The predictors for both FM and PFM are the first three factors extracted by applying the SW principal components method and $q_j = 1$.
- BM specification is that used in Rünstler and Sedillot (2003).

Istat
Istituto Nazionale di Statistica

The role of shrinkage

- Soft rule tends to select contemporaneous variables, the hard includes lagged variables up to the fourth order;
- When selected according to the soft method, variables are included in the majority of the vintages over the training period.
- The selection related to the soft approach tends to include a very small number of regressors when compared to the hard one. This negatively affect the forecast performance of PFMhard specifications.

Selected indicators by rule

Model	*Threshold*	Qualitative	Quantitative	Financial	Price	Total
		Indicator type				
PFM*hard*	90	16	9	7	1	33
	75	58	22	27	2	109
	50	92	32	35	4	163
	25	126	42	58	11	237
	10	126	42	58	11	237
	min	126	42	58	11	237
PFM*soft*	90	4	4	0	0	8
	75	9	6	4	0	19
	50	21	9	7	0	37
	25	38	13	13	0	64
	10	38	13	13	0	64
	min	41	13	17	0	71

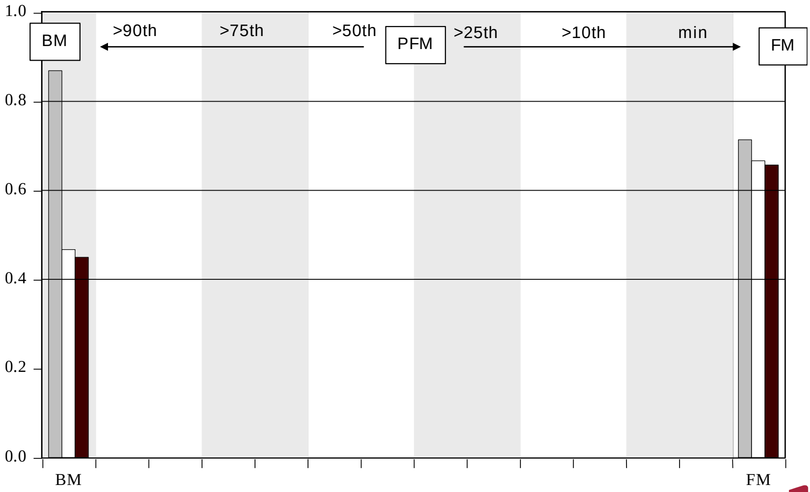
Istat
Istituto Nazionale di Statistica

Models' forecasting performance (RMSE over AR)

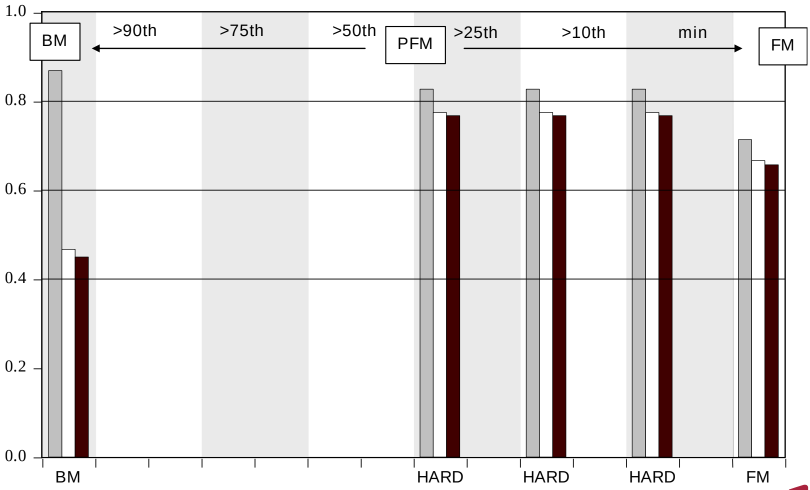
Code [a]	Model [b, c]	1st vintage	sig [d]	2nd vintage	sig [d]	3rd vintage (nowcast)	sig [d]	# of indicators [e]
	AR [f]	*0.0075*		*0.0073*		*0.0073*		*0*
1	BM	0.8689	*	0.4668	**	0.4499	**	4
2	PFM*hard*90	0.8407		0.6692	*	0.6755	*	33
3	PFM*soft*90	0.6932	**	0.4253	**	0.3161	**	8
6	PFM*hard*75	0.8179		0.7040		0.6819		109
7	PFM*soft*75	0.7729	*	0.5595	*	0.4978	*	19
10	PFM*hard*50	0.7573		0.6516		0.6324		163
11	PFM*soft*50	0.7489	*	0.5281	*	0.4862	*	37
14	PFM*hard*25	0.8269		0.7748		0.7676		237
15	PFM*soft*25	0.7005	*	0.6094	*	0.5736	*	64
18	PFM*hard*10	0.8269		0.7748		0.7676		237
19	PFM*soft*10	0.7005	*	0.6094	*	0.5736	*	64
22	PFM*hard*min	0.8269		0.7748		0.7676		237
23	PFM*soft*min	0.6689	*	0.5169	*	0.4914	*	71
26	FM	0.7134	*	0.6666	*	0.6571	*	1,295

Istat
Istituto Nazionale di Statistica

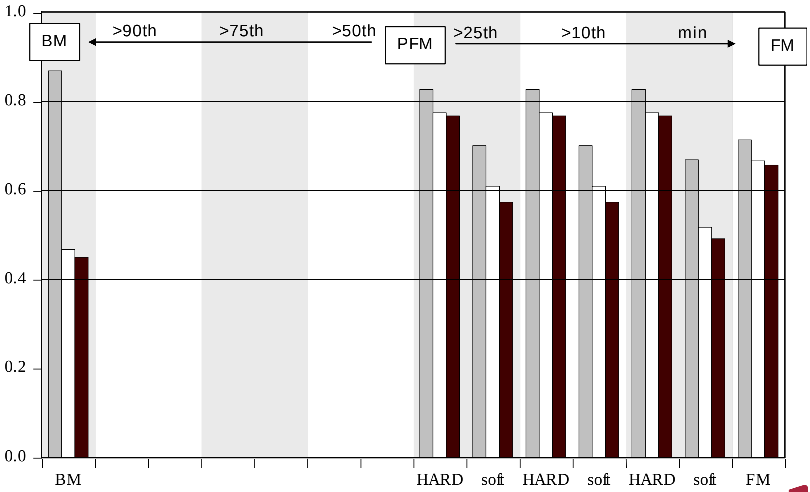
Models' forecasting performance (RMSE over AR)



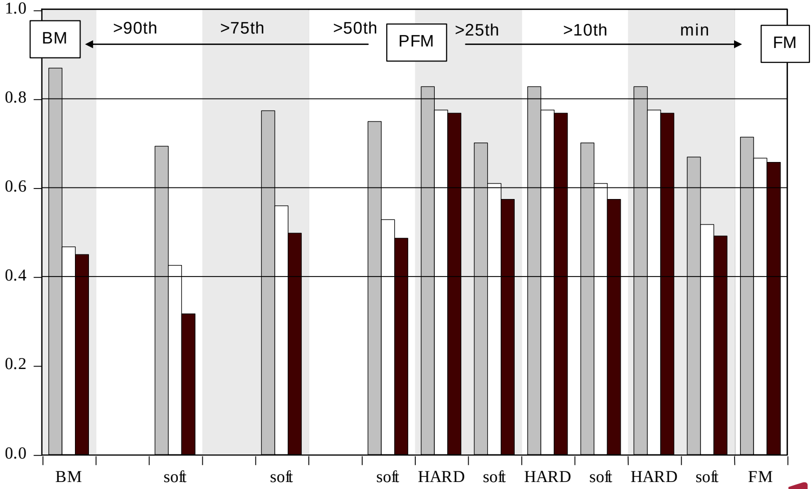
Models' forecasting performance (RMSE over AR)



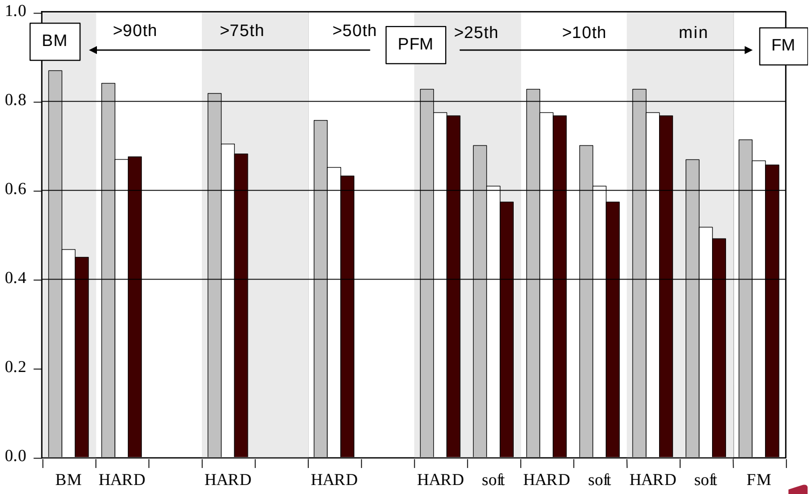
Models' forecasting performance (RMSE over AR)



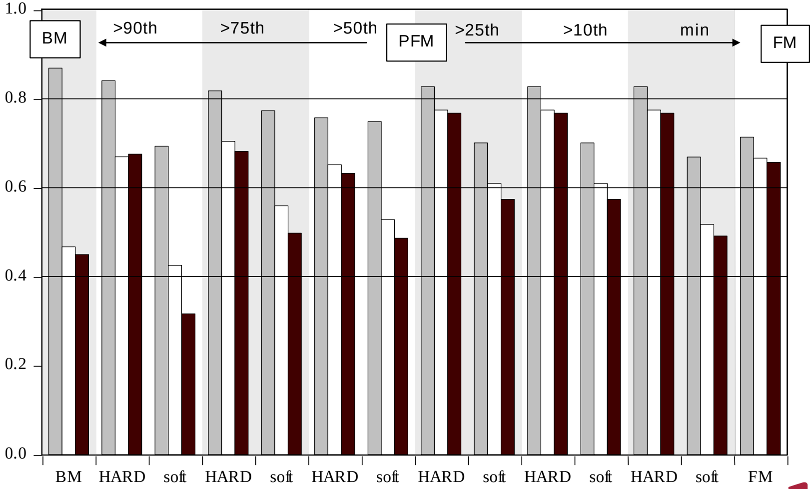
Models' forecasting performance (RMSE over AR)



Models' forecasting performance (RMSE over AR)



Models' forecasting performance (RMSE over AR)



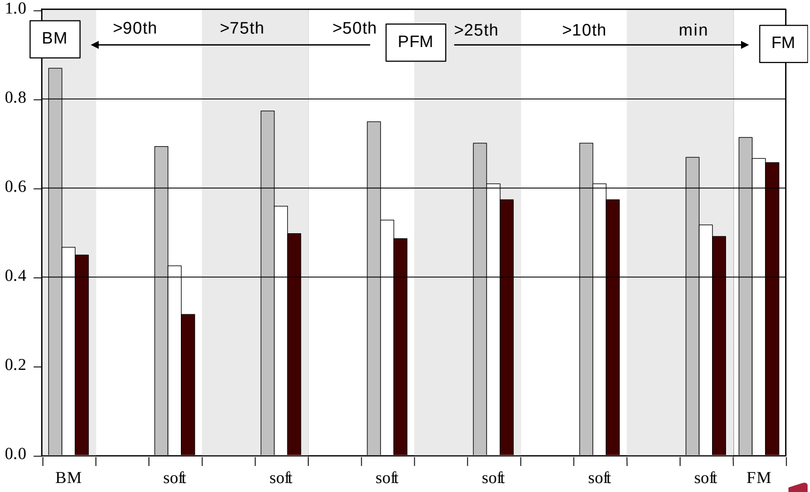
Forecast averaging and combination

- Given the large set of nowcasts obtained PFM with 5 soft-threshold rules, averages and combinations of single forecasts are computed. The aim is to further reduce the forecast error of soft-rules (Hendry and Clements, 2002; Stock and Watson, 2004).
- Challenging the average-view applying a procedure which aims at selecting a subset of available forecasts for combination by exploiting the complementarity between RMSE and the encompassing test (Costantini and Pappalardo, 2010; Kisinbay, 2010).

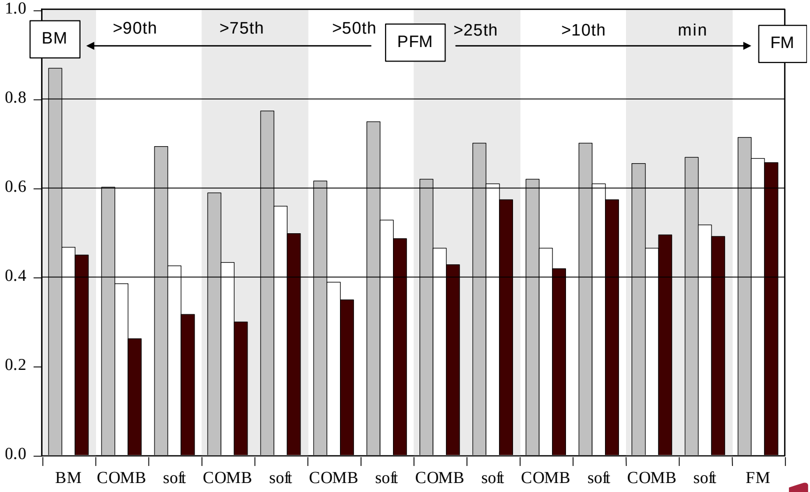
Forecast combination (RMSE over PFM*soft*)

Code [a]	Model [b, c]	1st vintage	sig[d]	2nd vintage	sig[d]	3rd vintage (nowcast)	sig[d]	# of indicators [e]
3[f]	*PFMsoft90*	*0.0052*		*0.0031*		*0.0023*		*8*
4	PFM*avg*90	0.8719		0.9944		0.9611		5
5	PFM*comb*90	0.8686		0.9052		0.8285		5
7[f]	*PFMsoft75*	*0.0058*		*0.0041*		*0.0036*		*19*
8	PFM*avg*75	0.7683		0.7498		0.6828		13
9	PFM*comb*75	0.7619		0.7734		0.6016		13
11[f]	*PFMsoft50*	*0.0056*		*0.0039*		*0.0035*		*37*
12	PFM*avg*50	0.8737		0.7830		0.7585		27
13	PFM*comb*50	0.8224		0.7355		0.7182		27
15[f]	*PFMsoft25*	*0.0053*		*0.0044*		*0.0042*		*64*
16	PFM*avg*25	0.9352		0.8175		0.7962		42
17	PFM*comb*25	0.8848		0.7629		0.7456		42
19[f]	*PFMsoft10*	*0.0053*		*0.0044*		*0.0042*		*64*
20	PFM*avg*10	0.9657		0.8740		0.8532		45
21	PFM*comb*10	0.8848		0.7629		0.7300		45
23[f]	*PFMsoftmin*	*0.0050*		*0.0038*		*0.0036*		*71*
24	PFM*avg*min	1.0341		1.0443		1.0077		49
25	PFM*comb*min	0.9792		0.8994		1.0065		49

Models' forecasting performance (RMSE over AR)



Models' forecasting performance (RMSE over AR)



Sensitivity analysis

Robustness is assessed to alterations coming from the following sources:

- disaggregate modelling of GDP demand components (i.e. demand-side GDP forecasts);
- changes in some settings in implementing the baseline FM approach: number of lags in expanding the panel of indicators (l = 4 in the baseline); partitioning in blocks b the input panel of indicators (b=1); an alternative number k of common (k = 3).

GDP forecasting ability: supply *vs.* demand-side models

Code [a]	z Models [a]	1st vintage S [b]	 D [b]	 **S/D [c]**	2nd vintage S [b]	 D [b]	 S/D [c]	3rd vintage S [b]	 D [b]	 S/D [c]
	Benchmark [d]	*0.0075*	*0.009*	***0.8333***	*0.0073*	*0.0088*	***0.8295***	*0.0073*	*0.0088*	***0.8295***
1	BM	0.8689	0.6667	**1.3033**	0.4668	0.6023	**0.7750**	0.4499	0.5682	**0.7918**
2	PFM*hard*90	0.8407	0.9667	**0.8697**	0.6692	0.8977	**0.7455**	0.6755	0.8864	**0.7621**
3	PFM*soft*90	0.6932	0.6444	**1.0757**	0.4253	0.4886	**0.8704**	0.3161	0.4091	**0.7727**
6	PFM*hard*75	0.8179	0.9000	**0.9088**	0.704	0.9318	**0.7555**	0.6819	0.9432	**0.7230**
7	PFM*soft*75	0.7729	0.8556	**0.9033**	0.5595	0.6023	**0.9289**	0.4978	0.625	**0.7965**
10	PFM*hard*50	0.7573	0.8111	**0.9337**	0.6516	0.7727	**0.8433**	0.6324	0.75	**0.8432**
11	PFM*soft*50	0.7489	0.7889	**0.9493**	0.5281	0.8523	**0.6196**	0.4862	0.8636	**0.5630**
14	PFM*hard*25	0.8269	0.900	**0.9188**	0.7748	0.8636	**0.8972**	0.7676	0.8295	**0.9254**
15	PFM*soft*25	0.7005	0.7333	**0.9553**	0.6094	0.8182	**0.7448**	0.5736	0.8409	**0.6821**
18	PFM*hard*10	0.8269	0.9111	**0.9076**	0.7748	0.875	**0.8855**	0.7676	0.8409	**0.9128**
19	PFM*soft*10	0.7005	1.0556	**0.6636**	0.6094	1.0000	**0.6094**	0.5736	0.9205	**0.6231**
22	PFM*hard*min	0.8269	0.9111	**0.9076**	0.7748	0.875	**0.8855**	0.7676	0.8409	**0.9128**
23	PFM*soft*min	0.6689	1.1111	**0.6020**	0.5169	1.0568	**0.4891**	0.4914	1.0568	**0.4650**
26	FM	0.7134	0.7000	**1.0191**	0.6666	0.6705	**0.9942**	0.6571	0.6364	**1.0325**

Istat
Istituto Nazionale di Statistica

GDP forecasting ability: RMSE ratios of alternative FM

FM(*l, b, k*) models	1st vintage	2nd vintage	3rd vintage
FM(0, 1, 1)	1.0185	1.0205	1.0209
FM(0, 1, 2)	1.0556	0.9796	0.9791
FM(0, 1, 3)	1.0926	1.0000	1.0000
FM(0, 1, 5)	1.1667	0.9184	0.8125
FM(1, 1, 1)	1.0185	1.0409	1.0416
FM(1, 1, 2)	0.9260	0.9388	0.9374
FM(1, 1, 3)	1.0185	1.0817	1.0833
FM(1, 1, 5)	0.9630	1.0000	1.0000
FM(2, 1, 1)	1.7408	1.8572	1.8958
FM(2, 1, 2)	1.0740	1.0817	1.0625
FM(2, 1, 3)	1.0371	1.0409	1.0416
FM(2, 1, 5)	1.0740	1.0817	1.0833
FM(4, 1, 1)	1.2778	1.3470	1.3542
FM(4, 1, 2)	1.0185	0.9796	0.9791
FM(4, 1, 3) [b]	*0.0054*	*0.0049*	*0.0048*
FM(4, 1, 5)	0.9815	1.0205	1.0416
FM(4, 4, 1)	1.2593	1.2245	1.2291
FM(4, 4, 2)	1.2963	1.3470	1.1667

Istat
Istituto Nazionale di Statistica

Usual results

- The forecast accuracy significantly improves over the AR benchmark as indicators' information increases, i.e. from the first to the third vintage.
- When v=1, i.e. in the less favourable situation, FM outperforms BM. This because FM includes indicators promptly available in v=1. By contrast, BM include **pre-selected hard** indicators, not yet available for a largest portion of the quarter to be forecast.
- As soon as monthly hard information is released (i.e. when v= 2, 3), BM performance improves considerably and its RMSFE ratio over the AR benchmark drops from 0.87 (for v=1) to 0.45 (for the nowcast, v=3).

New results

- RMSE ratios of the PFM for quantiles smaller than the 25^{th} percentile are systematically above 0.4.
- As we move towards higher quantiles, the performance of PFM *soft* rules show a **monotonic improvement**.
- PFM based on *soft rules* systematically **outperform** the PFM build on hard rule.
- It is **more effective** using **soft-thresholding** rules as they combine both variables selection and parameters estimation. The hard thresholding rule retains much more and less useful information.
- Best model (PFM*soft* at 90^{th} quantile) is close to BM in purpose, but shrinkage techniques help to reduce over-fitting problems in model specification.

Istat
Istituto Nazionale di Statistica

continue: New results

- There is some room for combination of forecast: the simple average of the five PFM based on the *soft* rules; the hierarchical procedure, which performs a selection of models based on both RMSE and encompassing tests (Costantini and Pappalardo; 2010 and Kisinbay, 2010).
- Focusing on forecast combination approaches, the hierarchical procedure (PFM*comb*) shows lower RMSFE ratios than the unweighted average of individual forecasts (PFM*avg*).
- Moving from Q = min to Q = 90th leads to monotonic improvements of the hierarchical forecast. This tends to vanish for the less stringent thresholding rule (Q = min). The larger the bulk of information, the less effective the combination approach is in lowering RMSE ratios compared to the unweighted average approach.

Istat
Istituto Nazionale di Statistica

Conclusions

- Overall, the first message coming from our results can be summarised by ***less modelling and more shrinking***: the exclusion of the indicators up to 90th percentile reduces at the same time the degree of freedom of the researcher and her unavoidable *non optimal* choices.
- The second message is that alternative very stringent soft rules lead to simple models that, however, are prone to be misspecified. In this, hierarchical combination of forecast and/or their average can help in preventing biases of the simplest models.