Machine Learning for Economists

Federico Bertoia 1030753
Nicola Cassarino 1026567
Virginia Pagliero 1025340
Benedikt Scheid 1038296

December 2022

1 Research Question

Does accounting for non-pervasive shocks improve forecasting performance?
Factor models are often used by modern literature to account for the curse of dimensionality[1], since they allow writing each of the N variables as a sum of two orthogonal components $x_{it} = \chi_{it} + \xi_{it}$, where χ_{it} is the common component that is driven by the $q << N$ shocks common to all variables, i.e. *pervasive* shocks, and ξ_{it} is accounting for the variable-specific idiosyncrasies.
Traditionally, the idiosyncratic components ξ are treated as if $\Gamma^{\xi} = E[\xi_t \xi_t']$ was diagonal[2], even in the case of approximate factor models, i.e. when there is mild cross-sectional dependence. Hence, the linear projection on the whole information set, $\Omega_t = span\{x_{t-p}, p = 0, 1, ...\} = \Omega_t^u \otimes \Omega_t^{\xi} = span\{u_{t-p}, p = 0, 1, 2, ...\} \otimes span\{\xi_{t-p}, p = 0, 1, 2, ...\}$, is approximated by the sum of the linear projection on the space spanned by the common shocks and of the linear projection on the space spanned by the *variable-specific* idiosyncracies (see for example Schumacher (2007)[10], Stock & Watson (2002a)[11], and Stock & Watson (2002b)[12]).
In other words, the forecast $x_{i,t+h|t} = Proj\{x_{i,t+h}|\Omega_t\}$ is approximated by $Proj\{x_{i,t+h}|\Omega_t^u\} + Proj\{x_{i,t+h}|\Omega_{it}^{\xi}\}$, implying that factor models of this kind implicitly assume $\Omega_{it}^{\xi} \approx \Omega_t^{\xi}$ despite the presence of weak cross-correlation in the idiosyncracies (i.e. *non-pervasive* shocks).
Taking this into consideration, the central idea of the paper is that neglecting *non-pervasive* shocks also implies neglecting information that could be used for forecasting. To access this extra information the author develops a model that puts a sparse structure on the idiosyncratic component via shrinkage and the h-step ahead forecast for variable i becomes

$$x_{i,t+h|t} = Proj\{x_{i,t+h}|\Omega_t^u\} + Proj_s\{x_{i,t+h}|\Omega_t^{\xi}\}$$

where $Proj_s\{x_{i,t+h}|\Omega_t^{\xi}\}$ is the linear projection with shrinkage onto the space spanned by the idiosyncratic components *of all variables* (Luciani, 2014)[9].[3] Note that by using shrinkage to select the variables for predicting ξ_t we are also enabled to be agnostic about the structure within Γ^{ξ}, i.e. there is no need to have prior information on the non-pervasive shocks.
In order to assess the research question, the Mean Squared Errors (MSE) of forecasts obtained from this model are compared to MSE's of forecasts coming from a simple AR(1), from the traditional factor model (Forni et al., 2005)[7], a factor model that extends the traditional one by forecasting $\xi_{i,t+h|t}$ with a simple AR(1) (D'Agostino & Giannone, 2012)[4], and from models that don't impose any factor structure but forecast x_{it} by applying shrinkage on the x's directly (De Mol et al., 2008)[5].[4] Specifically, this exercise is done for the time

[1] When analysing data in high dimensions, problems of multicollinearity as well as overfitting become more pressing, which leads to issues regarding interpretation and estimation.
[2] Note that $E[\xi_t] = 0$ is assumed, such that $\Gamma^{\xi} = Var(\xi_t)$.
[3] The estimation of the shrinkage projection is discussed further below.
[4] For all models with factor structure the forecast for the common component is estimated with the method proposed by Forni et al. (2005)[6].

series of GDP, Inflation, and Residential Investment of which we expect the latter to have a particularly strong dependence on non-pervasive shocks.
Lastly, we tried to expand on the paper's methodologies by implementing a forecast based on Bayesian Additive Regression Trees (BART).

2 Data

To assess the research question we used quarterly US data on macroeconomic variables taken from FRED database (Federal Reserve Economic Data)[5] from 01.09.2000 (2000:Q3) until 01.12.2019 (2019:Q4) amounting to 78 observations: indeed, we focused our work on this time span omitting missing values and excluding COVID-19 period to allow a prediction over stable years. For all our methods, we used the first 52 observations as training-set (i.e. two thirds of the total sample) and the last 26 observations as test-set to forecast.
It is important to notice that Luciani's (2014)[9] work refers to a different period (up to 2010:3) with a higher number of observations to forecast (80 observations). The dataset includes 246 variables accounting for main features that allow to describe the economic condition of a country[6]. However, these variables show high correlation within their belonging group which allows to compute factors to summarize their underlying relationship.

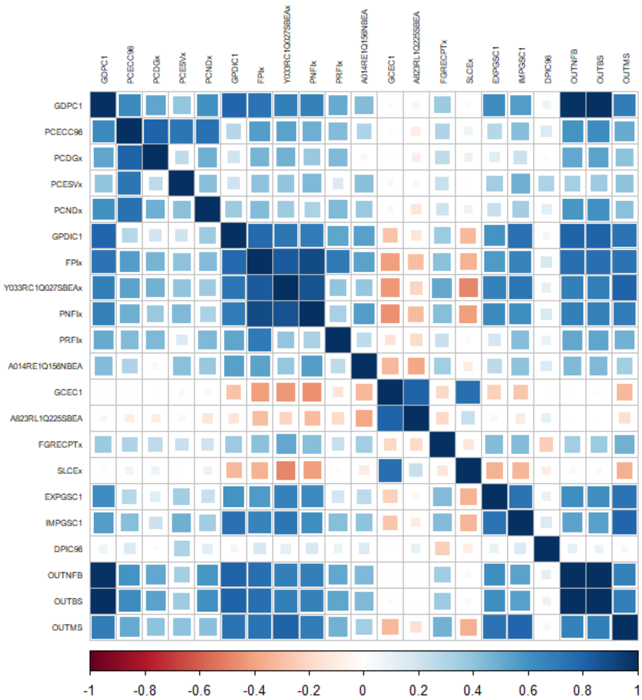


Figure 1: Correlogram depicting correlation relationships among the first group of variables (NIPA).

To perform our analysis we have decided, following Luciani (2014)[9], to first make the variables stationary and, secondly, standardize them (zero mean and unit variance) so that our factor model is not influenced by the different scales[7].

[5]We accessed the data at https://research.stlouisfed.org/econ/mccracken/fred-databases/, last accessed at 28.12.2022
[6]Some of the main groups are industrial production, employment, prices and productivity
[7]We used 7 different transformations for the variables to reach stationarity, following the example of McCracken.

3 Method & Models

We implemented several methods to forecast GDP, Inflation and Residential Investment in order to understand whether forecasts become more precise when non-pervasive shocks are taken into account. This is done by first evaluating the models' MSE's relative to the benchmark model's (simple AR(1)) MSE and then comparing the respective ratios. In order to evaluate the statistical significance of the ratios we perform a Diebold-Mariano test, that allows to check whether two loss functions (MSE and the tested model) differ from each other.

For most of the models we implemented a rolling scheme of forecasts, meaning that the training set gets shifted ahead by one observation (i.e. one quarter) after each forecast. The model is then refitted on the new training set which becomes the basis for the next forecast etc. Following the author we chose to perform 1- and 4-step ahead forecasts for each variable: accordingly, when forecasting on the test set with 26 observations (from 2013:3 to 2019:4) we obtain 26 forecasts for the 1-step ahead and 23 for the 4-step ahead forecast, where each forecast is based on the previous 52 quarters, i.e the first estimation is carried out on a sample from 2000:3 to 2013:2, while the last estimation uses a sample from 2006:4 to 2019:3 (resp. from 2006:1 to 2018:4 for the 4-step ahead).
As regards the forecasts based on factor models, we implemented the *ABC Criteria* to determine the number of common factors in Static Approximate Factor Model[8] that, following a less penalizing criterion, suggested to choose five static factors. In addition, we chose three dynamic factors according to the log criterion[9]: we get a total of eight factors that should account for the correlation among our regressors. This is consistent with the previous description of the dataset, that identified some groups of highly correlated variables: hence, we think that these factors manage to efficiently summarize the common (pervasive) component.
The models considered are the following:

- AR(1): Our benchmark model is a simple autoregression of lag one, where the variable in question is the sum of its weighted previous value and an error term, $y_{i,t} = \alpha \cdot y_{i,t-1} + \epsilon_{i,t}$ Hence, the h-step ahead forecast is: $\hat{y}_{i,t+h|t} = \hat{\alpha}^h \cdot y_{i,t}$

- FHLR1: this model considers only the common component χ_{it}, hence the forecasts for $\hat{x}_{i,t+h|t} = \hat{\chi}_{i,t+h|t}$. The forecasts of the common component are estimated following the two-step procedure of Forni et al. (2005)[7].[10] Therefore, this method allows catching the co-movement among all variables, i.e. common shocks.

- FHLR2: this method extends FHLR1 by forecasting the idiosyncratic component ξ_{it} through an AR(1) model as suggested by D'Agostino & Giannone (2012)[4] in addition to the forecast of χ_{it}. Thus, we obtain $\hat{x}_{i,t+h|t} = \hat{\chi}_{i,t+h|t} + \hat{\xi}_{i,t+h|t}$, where $\hat{\xi}_{i,t+h|t} = \hat{\beta}^h \cdot \xi_{i,t}$.

- ML1 & ML2: both of these models are based on Luciani's (2014)[9] idea of combining the factor model with shrinkage (see Section 1), but differ according to the way shrinkage is performed. Specifically, ML1 uses L1 penalized regressions (i.e. Bayesian shrinkage with double exponential prior, also called Lasso) while ML2 forecasts the idiosyncratic component via boosting.
  In greater detail, with Lasso the aim is to minimize jointly the RSS and the l_1 norm, which leads to shrinking unnecessary coefficients to zero. In order to do so, we use cross-validation and select the best tuning parameter $\lambda \geq 0$ that defines the penalty. Since we applied the rolling scheme, we use 26 and 23 Lasso regressions for the one and four-steps ahead forecasts. Therefore, each fitted training set will result in a unique and optimal λ, that is selected specifically for the set analysed, and different coefficients will be shrunk accordingly.
  On the other hand, Boosting allows for predicting by adding consequently new and updated trees to the

[8] as described in Alessi, Barigozzi and Capasso (2010)[1]
[9] as described in Hallin and Liska (2007) [8]
[10] The respective code was taken from http://www.barigozzi.eu/Codes.html, last accessed on 3.1.2023.

ensemble: following this idea, each new model is built on the original sample and trained according to the previous error. The construction of each new tree is filtered by a learning rate that allows getting small incremental improvements and therefore avoids overfitting due to the high number of trees[11]. The rolling scheme allows us to update the training set and by adjusting the tuning parameters, in particular the number of trees and the learning rate, we get the optimal values for our forecasting.
As these models do not treat the idiosyncracies as mutually orthogonal, we expect these approaches to perform better than the ones described before especially when forecasting a variable that is highly affected by non-pervasive shocks.

- DGR1 & DGR2: these models are based on De Mol et al. (2008)[5] and do the forecast with L1 penalized regressions and boosting respectively but without imposing any factor structure beforehand. The reason is that, as Luciani pointed out, in our case, it may happen that factor models and selected variables (through Lasso and Boosting) carry almost the same information.

- BCPS: finally, we decided to add our contribution by forecasting via BART (Bayesian Additive Regression Trees), called BCPS from now on. Our choice is motivated by the good performances shown for forecasting purposes, compared to other regression tree methods (i.e. Bagging, Random Forest, Boosting). This high relative prediction accuracy is due to the fact that BART sequentially applies Boosting and adds random perturbations at each tree, leading to considerable learning improvements. Another potential advantage is that we do not have to carefully choose specific penalty parameters but only the number of trees and iterations. In fact, we expect to find lower error rates than with boosting-based methods, especially DGR2: as the latter we do not assume any factor model for BCPS.

4 Results

GDP

Using the rolling scheme aforementioned we get positive and significant results for the autoregressive estimates[12] of GDP and the one and four-steps ahead MSE's are respectively 0.481 and 0.363.

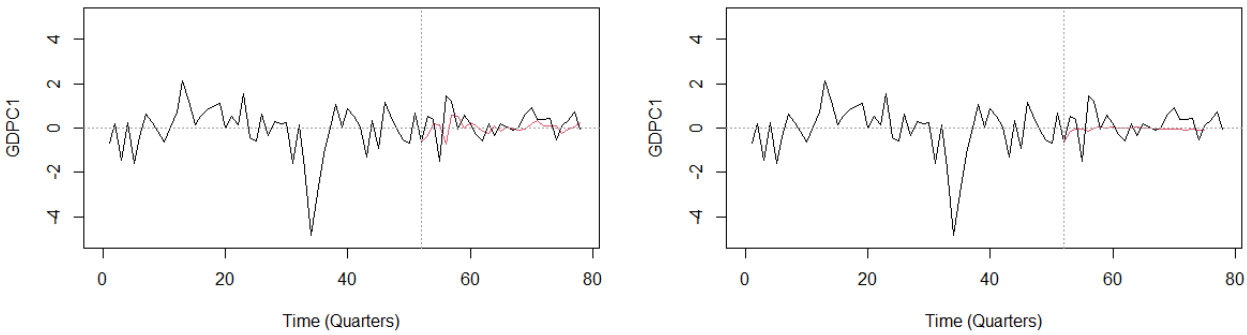


Figure 2: 1-step ahead AR forecast (left) and 4-steps ahead AR forecast (right) of GDP

However, this model isn't very good at predicting our test set and factor models may improve these results.
The rolling forecast of GDP through the FHLR1 outperforms the benchmark model and, if we implement FHLR2 by adding the AR(1) forecast of the idiosyncratic part ξ, we obtain similar MSE: for both horizons,

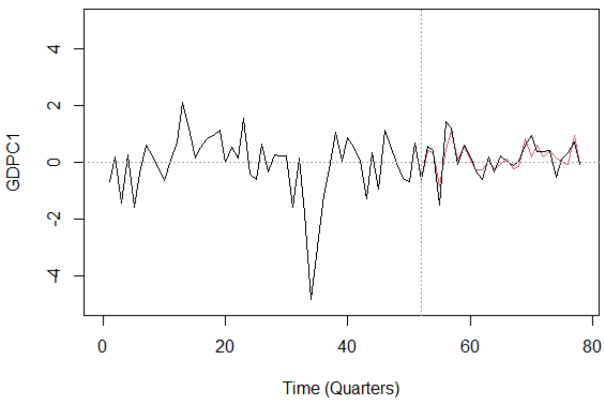
[11] The aim is also to improve with respect to Bagging and Random Forest methods that grow trees from a random sample and therefore tend to be highly correlated

[12] the autoregressive coefficients of the last fit for one and four-steps ahead are respectively $\alpha = 0.412^{**}$ and $\alpha = 0.389^{**}$

the two approaches manage to overcome the benchmark (see Table 1). This hints at the conclusion that using the co-movement within the data via factor models yields higher precision when forecasting GDP compared to estimates based on an AR(1).
However, when we test their relative predictive accuracy via the Diebold-Mariano test we get a p-value that does not allow us to reject the null that AR(1) is equally accurate as the tested model at any meaningful significance level, neither for one nor for four-steps ahead forecast MSE's[13].

Regarding differences in the predictive accuracy of the two FHLR's we draw on simulations of Luciani (2014)[9] showing that FHLR2 may yield a more precise prediction if the idiosyncratic component of the related variable is more persistent compared to the common part, i.e. the coefficient of the fitted AR(1) for ξ is high. In our case, there is no evidence for a strong difference between the two methods indicating that the idiosyncratic part doesn't have a persistent and significant autoregressive process[14] and that the common shocks account for the bulk of total variance.

When performing ML1 and ML2 we get very low MSE's, in particular regarding Lasso (see Table 1).
Since we select λ to get the smallest test MSE and there is high correlation between variables, we always get very low tuning parameters[15] which means that we are imposing an unrestrictive penalty on the parameters and we are close to OLS estimates. As we expected, the prediction is performed with the non-zero coefficients left, that are often the variables highly correlated with GDP, such as *OUTBS*, a measure for the output of the business sector or *GCEC1*, the real government consumption expenditures and gross investment.
On the other hand, Boosting is performed with a learning rate of 0.01 since it yields incremental improvements that minimize the CV error with 3405 trees: indeed, this reduces the risk of overfitting that may be caused by a large number of trees.



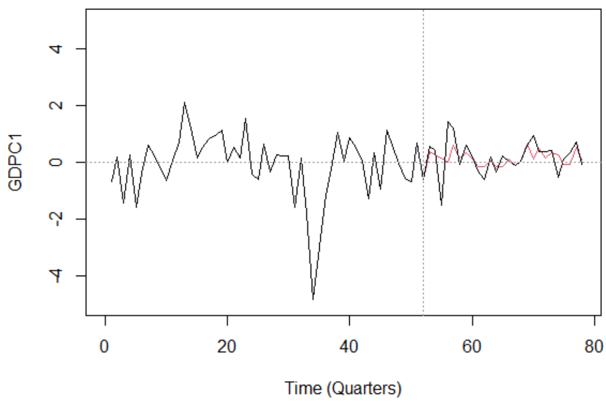


Figure 3: 1-step ahead ML1 forecast (left) and 1-steps ahead ML2 forecast (right) of GDP

Overall, the two models look very precise in their prediction (Figure 3): if we test their predictive accuracy we obtain that Lasso and Boosting forecasts are significantly more accurate for both horizons than the benchmark model (see Table 1). In fact, from Diebold and Mariano test we reject the null for the one and four-step ahead Lasso (given respectively p-value=0.01541 and p-value=0.0009513) as well as for the one and four-step ahead

[13] by doing the one-sided test for AR(1), i.e. check whether AR(1) is more accurate than the FHLR's, we got the same result, which indicates that there is no significant difference in predictive accuracy in neither direction

[14] indeed, by fitting the rolling AR(1) models to each training sample of ξ we always get very low coefficients that are not always strongly significant

[15] the minimum values for the one-step ahead is $\lambda = 0.00483$ whereas the maximum is $\lambda = 0.01438$

Boosting (p-value= 0.02006 and p-value= 0.01992).

Similarly, both DGR1 and DGR2 outperform the benchmark, as we obtain a lower MSE for every specification: in particular, Lasso yields peculiar results with MSE of 0.004582918 and 0.003613999 for one and four-steps ahead (we will further discuss this issue in the last section).

	FHLR1	FHLR2	ML1	ML2	DGR1	DGR2	BCPS
h = 1	0.80	0.78	0.24**	0.39**	0.0095	0.34*	0.26**
h = 4	0.72	0.72	0.30**	0.45**	0.0099	0.44	0.26*

(a) ** and * when the model is more accurate than benchmark at 5% and 10% significance level for Diebold-Mariano test.

Table 1: Relative mean squared errors of one and four steps ahead forecasts for GDP (AR(1) as benchmark).

They also do comparatively well compared to the other models: this may be due to the fact that, in our case, it is likely that factor models and selected variables (through Lasso and Boosting) carry almost the same information (Luciani, 2014)[9]. However, according to Diebold-Mariano test, DGR1 results are not significantly better than AR(1). Despite this, applying an alternative way to reduce the number of regressors without making use of factors or principal components may actually improve forecasting accuracy.

As MSE's ratios clearly show, our proposal, BCPS, not only does dramatically better than AR but also exceeds the other boosting-based models, ML2 and DGR2, even if the former accounts for the idiosyncratic shocks. Moreover, differently from these methods, it maintains a similar accuracy (compared to AR) also for the 4-step ahead prediction.

To conclude, we observe that every method outperforms the benchmark AR(1) and using variable selection methods provides considerably accurate estimates for the test sample of GDP, with or without a factor model. However, we need to consider that our forecasted observations are a very small number and therefore we may get biased results since the asymptotic properties are not always verified.

Inflation

For the analysis of the inflation we use the PCE Chain-type Price Index, which is the Federal Reserve's main measure for inflation across a large amount of expenses.

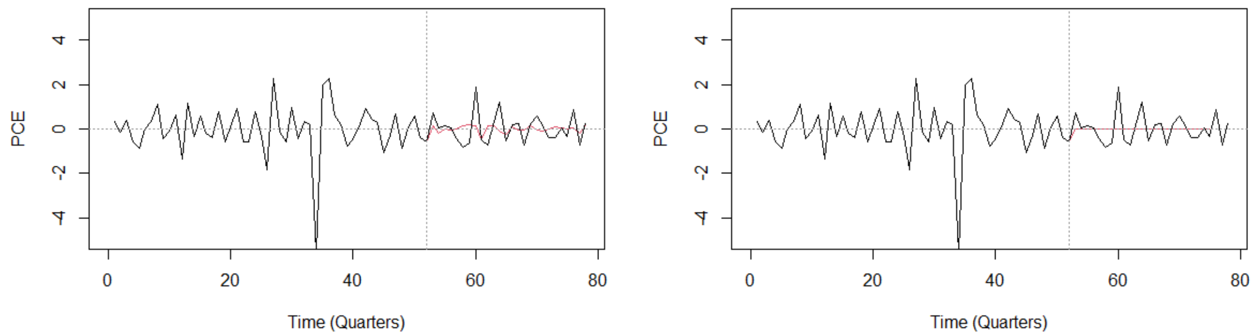


Figure 4: 1-step ahead AR forecast (left) and 4-steps ahead AR forecast (right) of Inflation

We implemented the same approaches with a rolling scheme, however, in this case, the AR(1) coefficients are

not significant: indeed, from the economic theory we may consider this variable as a random walk and we can't expect to get meaningful predictions.
While for the GDP we expected FHLR1 to be more accurate, the implemented model by D'Agostino & Giannone (2012), FHLR2, may perform better for Inflation if we think that it is characterized by more persistent idiosyncratic component. As a matter of fact, this is what we get by fitting the AR(1) of ξ[16], since it yields significant and more persistent coefficients than the ones obtained for GDP.
In addition, this is even more evident if we consider the four-steps ahead: we can imagine that the common part is less likely to provide a good forecast as time goes by, also given the random walks features of the variable analysed.
However, the accuracy test doesn't reject H_0 for any forecast horizon, hence FHLR models don't improve significantly the benchmark AR(1).

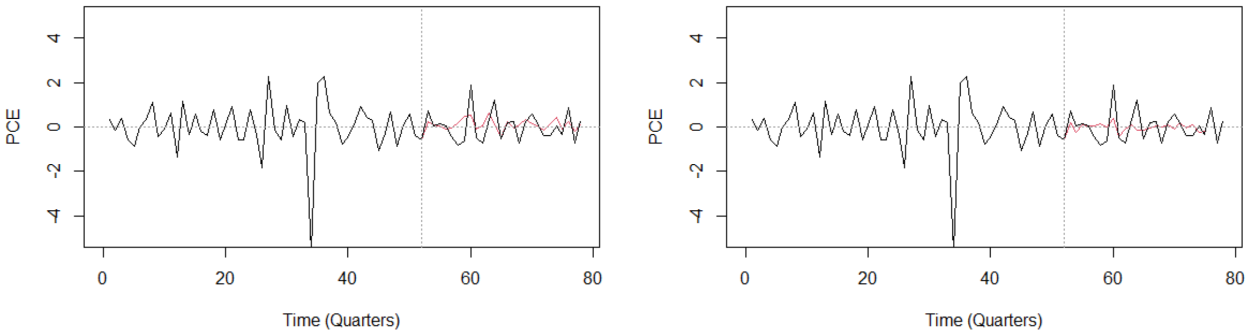


Figure 5: 1-step ahead ML1 forecast (left) and 4-step ahead ML1 forecast (right) of Inflation

If we try to forecast by accounting for non-pervasive shocks, we obtain that ML1 actually provides comparable results to GDP: not very restrictive λ's for every rolling set[17] and selection of the non zero predictors that are highly correlated to the Inflation such as the price index of goods or services.
Even by implementing ML2, it seems that taking into account local shocks does not importantly improve predictions results (see Table 2). Moreover, the Diebold-Mariano test for many of these models leads to not reject the null and therefore there is no statistical difference from the benchmark model.
Overall, factor-based models seem to behave similarly and accounting for non-pervasive shocks does not help forecasting.

Again, if we drop the factor model, similar conclusions can be drawn: in particular, ensamble methods on the variables such as Boosting and BART perform very good, whereas Lasso still gives unusual outcomes.

	FHLR1	FHLR2	ML1	ML2	DGR1	DGR2	BCPS
h = 1	0.84	0.89	0.85	0.76*	0.02	0.47	0.25**
h = 4	1.05	0.70	1.02	0.42	0.0019	0.68	0.21**

(a) ** and * when the model is more accurate than benchmark at 5% and 10% significance level for Diebold-Mariano test.

Table 2: Relative mean squared errors of one and four steps ahead forecasts for Inflation (AR(1) as benchmark)

[16] the autoregressive estimate is α= -0.693 for 1 step ahead and α= -0.476 for the four steps ahead and both are significant
[17] for the one-step ahead fit the minimum lambda is 0.001039874 whereas the maximum is 0.002846743

Overall, these results are very difficult to evaluate since implementing the accuracy test with respect to the AR(1) model has very low meaning. If we assume Inflation to be a random walk, hence affected by random shocks, we expect it to be unpredictable through ordinary methods such as autoregressive forecasts and therefore, the benchmark used is already not significant itself.

Residential Investment

We applied the same methods to Residential Investment and deflated it through PCE since it is one of the most volatile components of GDP [13] and should be highly affected by non-pervasive shocks: hence, we expect that the approaches suggested by Luciani [9] should improve forecast results.
The AR(1) estimates[18] are positive and statistically significant and predictions through the rolling scheme provide nice results for one-step ahead and slightly worse ones if the forecasting horizon is higher. Nevertheless, the AR model is clearly more adequate in explaining the behaviour of Residential Investment rather than inflation.

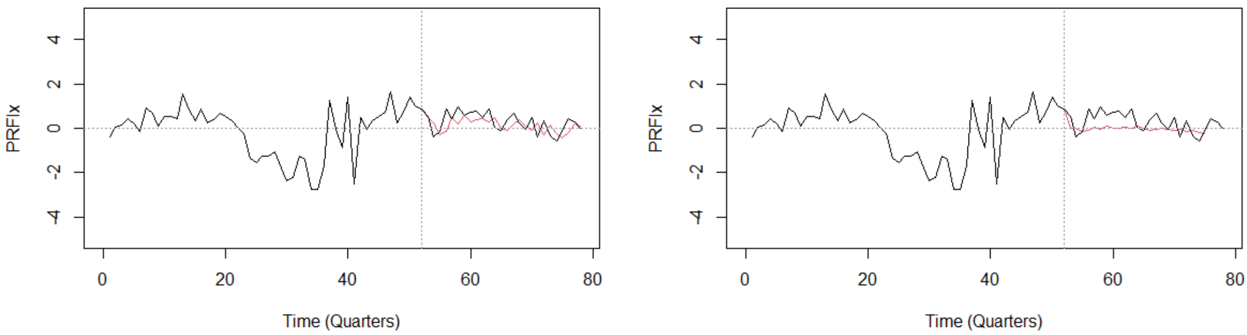


Figure 6: 1-step ahead AR forecast (left) and 4-steps ahead AR forecast (right) of Residential Investments

Next, the forecasts are computed by factors model and we notice that modeling the idiosyncratic component by an AR(1) slightly improves prediction accuracy for each forecast horizon, i.e there is not a significant difference between FHLR1 and FHLR2. This is due to the nature of the variable itself: in explaining residential investment, which is highly affected by regional shocks, the weakly cross-sectional correlation plays a fundamental role. A simple AR(1) is not able to capture the local correlation and that is why the predictions are close. Moreover, since the benchmark model had a poor long-horizon forecasting accuracy, for both models the one-step estimates are in line with the AR(1) while the four-step estimates are definitely better.

	FHLR1	FHLR2	ML1	ML2	DGR1	DGR2	BCPS
h = 1	1.09	1.07	0.79	0.61**	1.10	1.78	0.99
h = 4	0.61*	0.59*	0.54	0.51**	0.84	1.45	0.70

(a) ** and * when the model is more accurate than benchmark at 5% and 10% significance level for Diebold-Mariano test.

Table 3: Relative Mean Squared Errors of one and four steps ahead forecasts for residential investment (AR(1) as benchmark)

Through the models presented by Luciani, we notice that both Lasso and Boosting are able to account for non-pervasive shocks that affect Residential Investments: by modeling the idiosyncratic component via shrinkage we are able to capture the local correlation and improve the forecast.

[18] the autoregressive coefficients of the fit over the last training sample for one and four-steps ahead are repectively $\alpha = 0.552^{***}$ and $\alpha = 0.563^{***}$

For these reasons the MSE of ML1 and ML2 are relatively lower with respect to the other models.

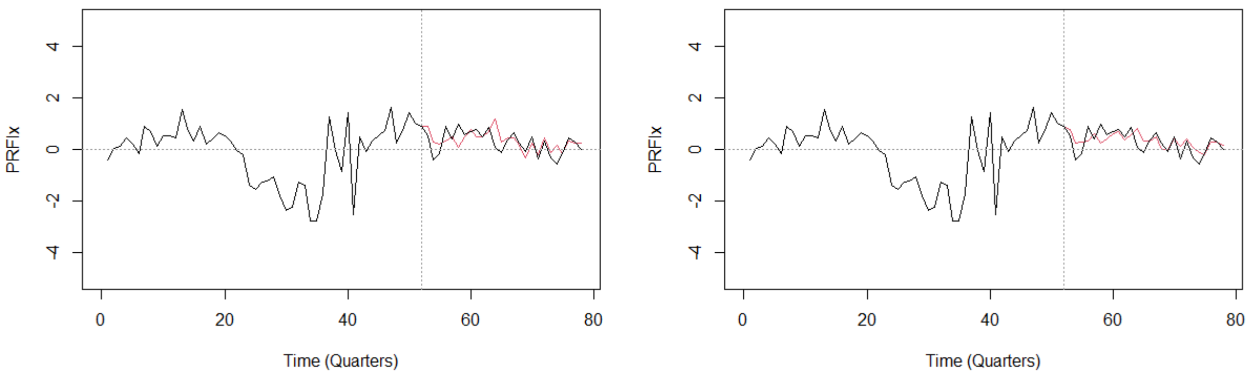


Figure 7: 1-step ahead ML1 forecast(left) and 1-step ahead ML2 forecast (right) of Residential Investment

Finally, we forecast Residential Investment via shrinkage but without assuming any factor model: if we perform lasso on the original set of variables then we get one-step forecasts which are as good as some of the factor models. When we predict four-step ahead we still get better predictions than the AR(1) model but worse than for the other models. However, if we consider boosting, without capturing the bulk of comovement before, we do not get accurate predictions.
A slightly better performance is given by BCPS which reaches lower error rates than its cousins, DGR1 and DGR2. Nevertheless, it is evident that for this variable, building a factor model and then using shrinkage leads to noticeable improvements.

In this context, the results of Luciani's approaches, ML1 and ML2, seem to confirm the local nature of residential investments probably influenced by non-pervasive shocks. Note that in this case we get to the same conclusion as the author.

5 Issues & Discussion

1. While implementing DGR1, we noticed that some groups of variable presented very low MSE whereas for some others they were comparable to the benchmark model. If we apply the same method to *ANDENOx*, i.e., Real value of manufacturers' new orders for capital goods or even for Residential Investments, predictions are worse than the same method applied for GDP or Inflation. This feature could be explained by the way in which lasso selects the variables to keep: it may be easier for some target variables to identify the most relevant regressors, such as GDP or inflation.

2. Another issue was the implementation of the rolling scheme, especially regarding the prediction of ξ. Our problem here was mainly due to the fact that the ξ's are, unlike variable values, not directly observable but model-based. For example, we decided to test our 1-step forecast against the estimate of the last ξ that results from a fit over the updated information-set i.e. the training set shifted one quarter ahead. However, we were insecure about whether this way or, for example, to use the ξ for the relevant quarter that was estimated with a fit spanning over the whole time frame, i.e. from 2000:3 to 2019:4, is more appropriate.

3. Following our results, in particular for GDP, we verified that predictions through Boosting without factor structure, i.e., DGR2, often performs much better than the benchmark model. Hence, we tried to further

improve these outcomes through Bayesian Additive Regression Trees (BART).

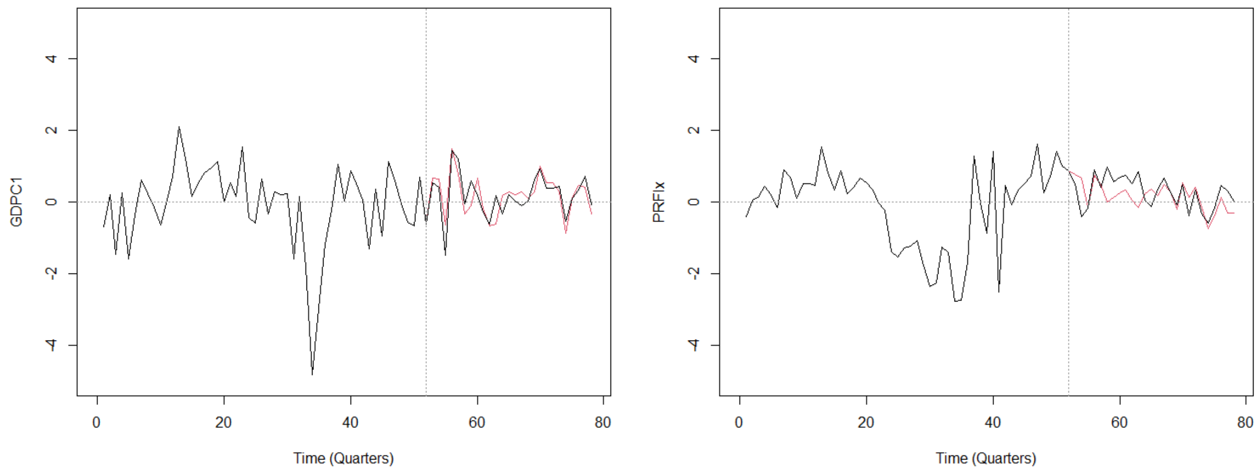


Figure 8: 1-step ahead BCPS forecast(left) for GDP and 1-step ahead BCPS forecast (right) of Residential Investment

As already reported in the results tables (Tables 2, 1 and 3), BART performs significantly better for the GDP, whereas it is comparable to benchmark for Residential Investment: indeed, since they are highly affected by non pervasive shocks, ML are better in capturing it.

4. To assess each methods proposed, computing the MSE is strictly necessary, since it provides an indicator of how close forecasts and true values are. However, even if we obtained very promising results, they sometimes couldn't beat the AR(1) model due to their non significance.
We can point out that it is because factor estimation may be weaken by high cross-correlation among idiosyncratic errors [3]: as a result of this, we could explain the low significance of factor-based models by recalling that our data shows a strong relationship within group variables.
In addition, since we included all the variables of the dataset to get consistent estimates (ensured if $N \longrightarrow \infty$), this may lead to unbalances, i.e. some groups result as more represented than others, and biased test criteria for the number of chosen factors.

5. In Luciani's paper the analysis is conducted both by simulating data (choosing different values of the parameter governing the persistence of the common shocks and for the non-pervasive ones) and by empirics. They showed that it is not always necessary to account for non-pervasive shocks when forecasting with factor models, but it depends whether the variable of interest responds primarily to macroeconomic (pervasive) shocks as GDP or there is the presence of regionals and sectoral shocks as in residential investments.

The main differences between his results and what we found can be summarized in:

- Models combining factors and shrinkage improves significantly GDP forecasts in our time span.
- Residential investments forecasts without assuming an underlying factor structure (DGR1/DGR2) give worse estimates than the ones obtained in Luciani's paper.
- The addition of BART estimates shows that there are ways in which forecasts can be improved. However, many different approaches could be used to explore the features of forecasting taking into account pervasive and non-pervasive shocks, as recurrent neural networks (RNN).
The presence of inflation as one of the main variables of interest shows that models prediction accuracy depends also on the nature of the variables itself.

- As far as we know part of the divergent results we got in our work can be explained by the different time span we chose to cover.

References

[1] Alessi, L., Barigozzi, M., & Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters, 80*, 1806-1813. https://doi.org/10.1016/j.spl.2010.08.005

[2] Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics, 146*, 304-317. https://doi.org/10.1016/j.jeconom.2008.08.010

[3] Bai, J., & Ng, S. (2008b), Large Dimensional Factor Analysis. *Foundations and Trends in Econometrics, 3*, 89-163. http://dx.doi.org/10.1561/0800000002

[4] D' Agostino, A. Giannone, D. (2012). Comparing Alternative Predictors Based on Large-Panel Factor Models. *Oxford Bulletin of Economics and Statistics, 74*(2), 306–326. https://doi.org/10.1111/j.1468-0084.2011.00642.x

[5] De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian regression a valid alternative to principal components? *Journal of Econometrics, 146*, 318-328. https://doi.org/10.1016/j.jeconom.2008.08.011

[6] Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The Generalized Dynamic-Factor Model: Identification and Estimation. *Review of Economics and Statistics, 82*(4), 540–554. https://doi.org/10.1162/003465300559037

[7] Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The generalized dynamic factor model: one sided estimation and forecasting. *Journal of the American Statistical Association, 100*, 830–840. https://doi.org/10.1198/016214504000002050

[8] Hallin, M., & Liska, R. (2007). Determining the Number of Factors in the General Dynamic Factor Model. *Journal of the American Statistical Association, 102*, 603-617. https://doi.org/10.1198/016214506000001275

[9] Luciani, M. (2014). Forecasting with approximate dynamic factor models: The role of non-pervasive shocks. *International Journal of Forecasting, 30*, 20-29. https://doi.org/10.1016/j.ijforecast.2013.05.001

[10] Schumacher, C. (2007). Forecasting German GDP Using Alternative Factor Models Based on Large Datasets. *Journal of Forecasting, 26*(4), 271–302. https://doi.org/10.1002/for.1026

[11] Stock, J. H. & Watson, M. W. (2002a). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association, 97*(460), 1167–1179. https://doi.org/10.1198/016214502388618960

[12] Stock, J. H. & Watson, M. W. (2002b). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business Economic Statistics, 20*(2), 147–162. https://doi.org/10.1198/073500102317351921

[13] Kohlscheen, E., Mehrotra, A., Mihaljek, D. (2018). Residential investment and economic activity: evidence from the past five decades. *BIS Working Papers No 726*. https://www.bis.org/publ/work726.pdf