



Gabriel Filipe Rodrigues Vasconcelos

**Forecasting in high-dimension: Inflation and
other economic variables**

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Engenharia Elétrica.

Advisor : Prof. Álvaro de Lima Veiga Filho
Co-Advisor: Prof. Marcelo Cunha Medeiros

Rio de Janeiro
July 2018



Gabriel Filipe Rodrigues Vasconcelos

**Forecasting in high-dimension: Inflation and
other economic variables**

Thesis presented to the Programa de Pós-graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Engenharia Elétrica. Approved by the undersigned Examination Committee.

Prof. Álvaro de Lima Veiga Filho

Advisor

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Marcelo Cunha Medeiros

Co-Advisor

Departamento de Economia – PUC-Rio

Prof. Cristiano Augusto Coelho Fernandes

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Eduardo Fonseca Mendes

Escola de Matemática Aplicada – FGV-Rio

Prof. Márcio Gomes Pinto Garcia

Departamento de Economia – PUC-Rio

Prof. Rodrigo dos Santos Targino

Escola de Matemática Aplicada – FGV-Rio

Prof. Márcio da Silveira Carvalho

Vice Dean of Graduate Studies

Centro Técnico Científico – PUC-Rio

Rio de Janeiro, July 6th, 2018

All rights reserved.

Gabriel Filipe Rodrigues Vasconcelos

Graduated in Economics at the Federal University of Juiz de Fora in 2011 and obtained his M.Sc. Degree in Economics from the Federal University of Juiz de Fora in 2013.

Ficha Catalográfica

Filipe Rodrigues Vasconcelos, Gabriel

Forecasting in high-dimension: Inflation and other economic variables / Gabriel Filipe Rodrigues Vasconcelos; orientador: Álvaro de Lima Veiga Filho; co-orientador: Marcelo Cunha Medeiros. – 2018.

v., 169 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Big data;. 3. Econometria em alta dimensão;. 4. LASSO;. 5. Modelos de Fatores;. 6. Previsão;. I. de Lima Veiga Filho, Álvaro. II. Cunha Medeiros, Marcelo. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Acknowledgments

After an intensive period, writing this note of thanks is the finishing touch of my PhD Thesis. I would like to reflect on the people who have supported and helped me so much throughout this period.

I would like first of thank my advisor, Álvaro Veiga, for accepting me as his student and for guiding my through this process. He is one of the most insightful people I have ever met. I wish to give a very special thanks to Marcelo Medeiros, who gave me a chance almost five years ago when I had almost nothing to show. If I have something to show now it is thanks to his guidance and he will be a friend for life.

I would also like to thank my girlfriend Gláucia for staying with me through the entire process and for being patient and caring with me through every obstacle. My friends Henrique Helfer, Yuri Fonseca, Marcelo Ruas, Luiz Claudio and may others also deserve a special thanks. I also would like to thank the professors Luiz Brandão, Leonardo Lima and Fernanda Perobelli.

I would like to thank my mother and my family for supporting and caring for me through my PhD and my entire life.

Last but not least, I would like to thank PUC-Rio, CAPES, CNPq, Lojas Americanas and NUPEI for the support.

Abstract

Filipe Rodrigues Vasconcelos, Gabriel; de Lima Veiga Filho, Álvaro (Advisor); Cunha Medeiros, Marcelo (Co-Advisor). **Forecasting in high-dimension: Inflation and other economic variables**. Rio de Janeiro, 2018. 169p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis is made of four articles and an R package. The articles are all focused on forecasting economic variables on high-dimension. The first article shows that LASSO models are very accurate to forecast the Brazilian inflation in small horizons. The second article uses several Machine Learning models to forecast a set of US macroeconomic variables. The results show that a small adaptation in the LASSO improves the forecasts but with high computational costs. The third article is also on forecasting the Brazilian inflation, but in real-time. The main results show that a combination of Machine Learning models is more accurate than the FOCUS specialist forecasts. Finally, the last article is about forecasting the US inflation using a very large set of models. The winning model is the Random Forest, which opens the discussion of nonlinearity in the US inflation. The results show that both nonlinearity and variable selection are important features for the Random Forest performance.

Keywords

Big data; High-dimension econometrics; LASSO; Factor models; Forecasting;

Resumo

Filipe Rodrigues Vasconcelos, Gabriel; de Lima Veiga Filho, Álvaro; Cunha Medeiros, Marcelo. **Previsão em alta dimensão: Inflação e outras variáveis econômicas**. Rio de Janeiro, 2018. 169p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese é composta de quatro artigos e um pacote de R. Todos os artigos têm como foco previsão de variáveis econômicas em alta dimensão. O primeiro artigo mostra que modelos LASSO são muito precisos para prever a inflação brasileira em horizontes curtos de previsão. O segundo artigo utiliza vários métodos de Machine Learning para prever um grupo de variáveis macroeconômicas americanas. Os resultados mostram que uma adaptação no LASSO melhora as previsões com um alto custo computacional. O terceiro artigo também trata da previsão da inflação brasileira, mas em tempo real. Os principais resultados mostram que uma combinação de modelos de Machine Learning é mais precisa do que a previsão do especialista (FOCUS). Finalmente, o último artigo trata da previsão da inflação americana utilizando um grande conjunto de modelos. O modelo vencedor é o Random Forest, que levanta a questão da não-linearidade na inflação americana. Os resultados mostram que tanto a não-linearidade quanto a seleção de variáveis são importantes para os bons resultados do Random Forest.

Palavras-chave

Big data; Econometria em alta dimensão; LASSO; Modelos de Fatores; Previsão;

Table of contents

1	Introduction	12
2	Survey of Methods and Framework	14
2.1	General Framework	14
2.2	Models	15
2.2.1	Benchmark Models	16
2.2.2	Regularization and Shrinkage	17
2.2.2.1	Ridge Regression (RR):	17
2.2.2.2	Least Absolute Shrinkage and Selection Operator (LASSO):	18
2.2.2.3	Adaptive LASSO (adaLASSO):	18
2.2.2.4	Elastic Net (EINet)	18
2.2.3	Factor Models	19
2.2.3.1	Target Factors	19
2.2.3.2	Factor Boosting	20
2.2.4	Ensemble Methods	21
2.2.4.1	Bagging	21
2.2.4.2	Complete Subset Regressions	21
2.2.4.3	Jackknife Model Averaging	22
2.2.5	Regression Trees and Random Forests	23
2.2.6	Hybrid Linear-Random Forests Models	25
2.2.7	Bayesian Vector Autoregressive Model	26
3	Forecasting Brazilian inflation with high-dimensional models	27
3.1	Introduction	28
3.2	LASSO models	31
3.3	Main Results	33
3.3.1	IPCA Inflation Index	34
3.3.2	IGP-M Inflation Index	41
3.4	Final Remarks	47
3.5	Chapter Appendix	50
3.5.1	Estimation Procedures	50
3.5.1.1	Differentiation	50
3.5.1.2	LASSO and adaLASSO estimation	50
3.5.2	List of Variables	53
4	Forecasting Macroeconomic Variables in Data-Rich Environments	57
4.1	Introduction	58
4.2	Econometric Models for Data-Rich Environments	58
4.2.1	LASSO	58
4.2.2	Bagging	59
4.2.3	Target Factors	59
4.2.4	Complete Subset Regression	60
4.3	Estimation Procedures and Data	61
4.3.1	Estimation Procedures	61

4.3.2	Data	62
4.4	Results	62
4.5	Conclusion	64
5	Real-Time Inflation Forecasting with High-Dimensional Models: The Case of Brazil	65
5.1	Introduction	65
5.2	Empirical Methods	69
5.2.1	Factor Models with Targeted Predictors	70
5.2.2	LASSO and adaptive-LASSO	72
5.2.3	Random Forest	73
5.2.4	Complete Subset Regression with Targeted Predictors	75
5.3	The Data	76
5.4	Main Results	77
5.4.1	Forecasting Errors	77
5.4.2	Model Confidence Sets and Model Combination	81
5.4.3	Look Ahead Bias on the MCS Combined Forecasts	83
5.4.4	Different Window Size	84
5.4.5	Bayesian Alternatives	85
5.4.5.1	Unobserved Component Stochastic Volatility Model	86
5.4.5.2	Bayesian Vector Autoregressive Model	87
5.4.5.3	Results	87
5.4.6	Density Forecasts	89
5.5	Conclusion	93
5.6	Data Appendix	93
6	Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods	97
6.1	Introduction	97
6.1.1	Main takeaways	100
6.1.2	A brief comparison of the recent literature	102
6.2	Data	104
6.3	Methodology	105
6.4	Results	107
6.4.1	Overview	108
6.4.2	Results: Random Forests versus Benchmarks	109
6.4.3	Results: The Full Picture	110
6.4.4	Opening the Black Box: Variable Selection	113
6.5	Conclusions	115
7	References	161

List of figures

2.1	Example of a regression tree. Reproduction of part of Figure 9.2 in (1).	23
3.1	Shrinkage coefficients - LASSO	32
3.2	IPCA and IGM-M time series	34
3.3	Autocorrelation function: IPCA	35
3.4	RMSE and MAE of all models: IPCA	36
3.5	Number of selected variables: IPCA	37
3.6	Cumulative squared errors: IPCA	39
3.7	Variables selected by the LASSO: IPCA	41
3.8	Variables selected by the adaLASSO: IPCA	42
3.9	Autocorrelation function: IGP-M	43
3.10	RMSE and MAE of all models: IGP-M	44
3.11	Number of selected variables: IGP-M	45
3.12	Cumulative squared errors: IGP-M	47
3.13	Variables selected by the LASSO: IGP-M	48
3.14	Variables selected by the adaLASSO: IGP-M	49
3.15	Model selection using the BIC	52
5.1	Brazilian Consumer Prices Index and Focus Top5 Forecasts	77
5.2	Average Number of Selected Variables by the Shrinkage Methods	80
5.3	Forecasting Errors Correlation	81
6.1	Inflation rate (CPI, PCE and CPI core) from 1960 to 2015.	118
6.2	Rolling RMSE.	125
6.3	Rolling MAE.	126
6.4	Rolling MAD.	127
6.5	Eigenvalues of the matrix of contemporaneous regressor.	129
6.6	Correlation of the Forecasts for the CPI from 1990 to 2015	130
6.7	Variable importance	135
6.8	Word clouds for the adaLASSO method (1990–2000).	143
6.9	Word clouds for the adaLASSO method (2001–2015).	144
6.10	Word clouds for the Random Forest model (1990-2000).	145
6.11	Word clouds for the Random Forest model (2001—2015).	146

List of tables

3.1	Forecast correlations and GW test p-values: IPCA	38
3.2	Most selected variables: IPCA	43
3.3	Forecast correlations and GW test p-values: IGP-M	46
3.4	Most selected variables: IGP-M	46
4.1	Descriptive Statistics	62
4.2	RMSE across the rolling windows and Giacomini-White test for $h = 1$	63
4.3	RMSE across the rolling windows and Giacomini-White test for $h = 4$	64
5.1	Forecasts Mean Absolute Errors and Root Mean Squared Errors	79
5.2	Proportion each Model was in each Position of the Error Ranking	80
5.3	Model Confidence Set	83
5.4	Combined Forecasts Mean Absolute Errors and Mean Squared Errors	83
5.5	Combined Forecasts Mean Absolute Errors and Mean Squared Errors without Look Ahead Bias	84
5.6	Forecasts Mean Absolute Errors and Root Mean Squared Errors for 24 Rolling Windows	85
5.7	Model Confidence Set - 24 Rolling Windows	86
5.8	Forecasts Mean Absolute Errors and Root Mean Squared Errors for BVAR and UCSV	88
5.9	Giacomini and White test p -values comparing the UCSV and the Bayesian VAR to all other models	89
5.10	Average log-scores for all models and all forecasting horizons	91
5.11	Amisano and Giacomini test p -values comparing the forecasting densities of all models for all 12 horizons of forecast	92
5.12	Macroeconomic Variables	95
5.13	Focus Expectation Variables	96
6.1	Descriptive Statistics	119
6.2	Forecasting Results: Summary statistics for the out-of-sample period from 1990–2015	120
6.3	Forecasting Results: Summary statistics for the out-of-sample period from 1990–2000	121
6.4	Forecasting Results: Summary statistics for the out-of-sample period from 2001–2015	122
6.5	Forecasting Results: RMSE, MAE and MAD Ratios (1990–2015)	123
6.6	Forecasting Results: Ranking of Models (1990–2015)	124
6.7	Forecasting Results: Superior Predictive Ability Test (1990–2015)	125
6.8	Forecasting Errors for the CPI from 1990 to 2015	128
6.9	Forecasting Errors for the CPI from 1990 to 2000	131
6.10	Forecasting Errors for the CPI from 2001 to 2015	132
6.11	Conditional Giacomini-White Test for Superior Predictive Ability: Recession and Expansions	133
6.12	Conditional Giacomini-White Test for Superior Predictive Ability: Macroeconomic Uncertainty	134

6.13	Data Description: Output and Income	136
6.14	Data Description: Labor Market	137
6.15	Data Description: Housing	138
6.16	Data Description: Consumption, Orders and Inventories	138
6.17	Data Description: Money and Credit	139
6.18	Data Description: Interest and Exchange Rates	140
6.19	Data Description: Prices	141
6.20	Data Description: Stock Market	141
6.21	Forecasting Results PCE: Summary statistics for the out-of-sample period from 1990–2015	148
6.22	Forecasting Results PCE: Summary statistics for the out-of-sample period from 1990–2000	149
6.23	Forecasting Results PCE: Summary statistics for the out-of-sample period from 2001–2015	150
6.24	Forecasting Errors for the PCE from 1990 to 2015	152
6.25	Forecasting Errors for the PCE from 1990 to 2000	153
6.26	Forecasting Errors for the PCE from 2001 to 2015	154
6.27	Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 1990–2015	155
6.28	Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 1990–2000	156
6.29	Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 2001–2015	157
6.30	Forecasting Errors for the CPI-Core from 1990 to 2015	158
6.31	Forecasting Errors for the CPI-Core from 1990 to 2000	159
6.32	Forecasting Errors for the CPI-Core from 2001 to 2015	160

1 Introduction

In the dynamic world of today, forecasting important variables has become a difficult challenge because of the infinite amount of data being generated every day. The first problem is computational, it is not trivial to work and visualize big datasets. In many cases, simple personal computers might not be enough to deal with this data. The second problem concerns the selection of the most suited method to construct the analysis and the forecasts. New methods are published in a very high frequency and it is unclear for the researcher if there is an overall superior method or which method is more adequate to each problem. To name some methods we have the LASSO family (2, 3), factor models (4, 5, 6), boosting (7), bagging (8), regression trees (9) and many more.

Sometimes the theory on these models helps the researcher to select the best suited model. For example, LASSO and factor models are supposed to be used on opposite situations. The first requires sparsity, i.e, only a small subset of the potential variables are in fact useful. The second needs all variables to have some common behavior that can be extracted, therefore, most variables are important but the common behavior may be enough to describe them. However, the situation that best fits the data is not always clear. For example, if we are dealing with a large set of economic variables, it is reasonable to assume that these variables are all linked through a common trend. But we could also have only a few really important variables and all the rest with coefficients very close to zero that could simply be ignored.

Another important issue to consider is whether to iterate forecasts or to estimate them directly to the forecasting horizon of interest. For example, if we want to forecast the inflation five months ahead we could regress the inflation in $t + 5$ on potential variables in t . However, we could also use the $t + 1$ inflation and iterate the forecast until $t + 5$. The second case creates another complication on the covariates, because they would also need to be predicted. Fortunately VAR models can deal with this situation, even on high-dimension, if we estimate the equations by LASSO or using Bayesian VARs such as (10).

This thesis is composed of four articles, and one R package with

implementations. Three of these articles are already published and the other two are working papers. The R package is available on GitHub¹ and is yet to be submitted to CRAN.

The first article is on the forecasting of the Brazilian inflation using high-dimensional data (11). In this case we do not use real-time² data. Instead, we use variables in their corresponding period even when they are published with delay. Additionally, we do not consider expectation variables. The results show that the LASSO are more accurate for small forecasting horizons (until four months ahead), but on long horizons neither the LASSO nor factor models are able to beat simple autorregressives.

The second article aims to forecast the Brazilian inflation using real-time data in a much more sophisticated environment (12). We included several models such as Bayesian VARs, expectation variables, density forecasting, model confidence set and forecasting combinations. The expectation variables improve the forecasting accuracy in a way that for all forecasting horizons (five days to eleven months plus five days) there is at least one high-dimensional model that beats the naive models such as ARs and Random Walks.

The third article uses a very large set of high-dimensional models to forecast the US inflation (13). We also try to give an interpretation to the selected variables and rank them using importance measures. Moreover, we found evidence that nonlinear tree models such as Random Forests are very accurate to forecast US inflation and we explored this feature proposing two mixed linear models to understand if the Random Forest performance is due to variable selection or nonlinearity.

Finally, the R package HDeconometrics has implementations of many models used across the articles. Some models such as LASSO or Random Forests already have very good implementations on R, therefore, we only made small adjustments when required. However, to the best of our knowledge models such as Bagging, Complete Subset Regressions, some types of Bayesian VARs and others do not have implementations on CRAN or the implementations do not meet our needs. Additionally, our package has the advantage to combine many models in a single platform with a unifying use and notation.

The remainder of this thesis is divided in five chapters besides this introduction. Chapter two is a methodological survey and the remaining chapters are one for each article.

¹[gabrielrvsc/HDeconometrics](https://github.com/gabrielrvsc/HDeconometrics)

²By real-time we mean only the data available to the econometrician in the day the forecast is computed.

2 Survey of Methods and Framework

Although all articles in this thesis are about forecasting economic variables, the models and the notation are not exactly the same across all articles. This goal of this chapter is to unify the notation, framework and models used in all papers.

2.1 General Framework

Consider the following model:

$$y_{t+h} = T_h(\mathbf{x}_t) + u_{t+h}, \quad h = 1, \dots, H, \quad t = 1, \dots, T, \quad (2-1)$$

where y_{t+h} is the response variable at time $t+h$; $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})'$ is a n -vector of covariates, possibly containing lags of y_t and/or common factors as well as a large set of potential predictors; $T_h(\cdot)$ is the mapping between covariates and future values of the response variable; and u_t is a zero-mean random error. The target function $T_h(\mathbf{x}_t)$ can be a single model or an ensemble of different specifications. There is a different mapping for each forecasting horizon, h .

The direct forecasting equation is given by

$$\hat{y}_{t+h|t} = \hat{T}_{h,t-R_h+1:t}(\mathbf{x}_t), \quad (2-2)$$

where $\hat{T}_{h,t-R_h+1:t}$ is the estimated target function based on data from time $t - R_h + 1$ up to t and R_h is the window size, which varies according to the forecasting horizon and the number of lagged variables in the model.

The framework above is called **direct forecast**. This approach estimates one different model for each forecasting horizon and it has the advantage that we do not need to forecast the covariates for forecasting horizons larger than 1. The alternative way is called **recursive forecast**. In this case we only need to estimate the model one time but the forecasts must be iterated until the desired horizon, which requires forecasts of the covariates. I adopted direct forecasts because most of machine learning models are built to estimate single equations and forecasting the covariates would be a complicated problem. The only exception is the Bayesian VAR, which estimates a system of equations for all covariates at the same time.

The forecasts are based on a **rolling window** framework of fixed length. Suppose that we have T observations, define $t_0 \ll T$, as the number of observations in each window, K as the number of windows, the rolling window procedure follows the steps below:

- for $k = 1, \dots, K$
 1. estimate the model in (2-1) using the observations $k, \dots, t_0 + k - 1$,
 2. compute the forecast in (2-2)

The alternative to the rolling window approach is the **expanding window**, which follows the steps below:

- for $k = 1, \dots, K$
 1. estimate the model in (2-1) using the observations $1, \dots, t_0 + k - 1$,
 2. compute the forecast in (2-2)

The difference between the two methods is that the rolling window keeps the window size fixed and the expanding window increases it as we move on k . I adopted rolling windows because it is a requirement for some statistical tests to compare the forecasts and also because economic variables may have significant changes on their dynamics over long periods of time. Therefore, if we grow the estimation sample in a expanding window framework we will probably have multiple regimes in the sample and the forecasts would be less accurate.

2.2 Models

I used the following classification for models:

- **Benchmarks** : Autorregressive model (AR), Random Walk (RW), Unobserved Component Stochastic Volatility model (UCSV) (for inflation only).
- **Shrinkage**: Ridge Regression (RR), LASSO, adaLASSO, Elastic-Net, adaptive Elastic-Net.
- **Ensemble**: Bagging, Complete Subset Regression (CSR), Jackknife model averaging (JMA).
- **Factor based**: Factor model, target factor model, boosting of factors.
- **Nonlinear**: Random Forest (RF)

– **Bayesian:** Large Bayesian VAR (LBVAR)

The objective in this section is to do a brief presentation of all the models above. All the information contained here is also in the articles in the next chapters. The objective here is to put the models and the framework together. If the reader chooses to read this chapter he/she may skip all methodology sections in the articles.

2.2.1

Benchmark Models

The first benchmark is the RW model, where for $h = 1, \dots, 12$, the forecasts are computed as follows:

$$\hat{y}_{t+h|t} = y_t. \quad (2-3)$$

For the accumulated twelve-month forecast, we consider the following equation:

$$\hat{y}_{t+1:t+12|t} = y_{t-11:t}, \quad (2-4)$$

where $y_{t-11:t}$ is the accumulated inflation over the previous twelve months.

The second benchmark is the autoregressive (AR) model of order p , where p is determined by the Bayesian information criterion (BIC) and the parameters are estimated by Ordinary Least Squares (OLS). The forecast equation is

$$\hat{y}_{t+h|t} = \hat{\phi}_{0,h} + \hat{\phi}_{1,h}y_t + \dots + \hat{\phi}_{p,h}y_{t-p+1}. \quad (2-5)$$

There is a different model for each horizon. The accumulated forecasts are computed by aggregating the individual forecasts.

Finally, the third benchmark is the UCSV model, which is described as follows:

$$\begin{aligned} y_t &= \tau_t + e^{h_t/2}\varepsilon_t, \\ \tau_t &= \tau_{t-1} + u_t, \\ h_t &= h_{t-1} + v_t, \end{aligned} \quad (2-6)$$

where $\{\varepsilon_t\}$ is a sequence of independent and normally distributed random variables with zero mean and unit variance and $\varepsilon_t \sim \mathbf{N}(0, 1)$, u_t and v_t are also normal with zero mean and variance given by inverse-gamma priors. $\tau_1 \sim \mathbf{N}(0, V_\tau)$ and $h_1 \sim \mathbf{N}(0, V_h)$, where $V_\tau = V_h = 0.12$. The model is estimated by Markov chain Monte Carlo (MCMC) methods. The h -steps-ahead forecast is computed as $\hat{y}_{t+h} = \hat{\tau}_{t|t}$.

For accumulated forecasts, the UCSV is estimated with the twelve-month inflation as the dependent variable.

2.2.2 Regularization and Shrinkage

This subsection shows a discussion of all models in the Shrinkage Sparse group plus the Ridge from the Shrinkage Non-Sparse group. These are linear models where $T_h(\mathbf{x}_t) = \boldsymbol{\beta}'_h \mathbf{x}_t$ and

$$\hat{\boldsymbol{\beta}}_h = \arg \min_{\boldsymbol{\beta}} \left[\sum_{t=1}^{T-h} (y_{t+h} - \boldsymbol{\beta}' \mathbf{x}_t)^2 + \lambda \sum_{i=1}^n p(\beta_i; \omega_i, \alpha) \right], \quad (2-7)$$

where $p(\beta_i; \omega_i, \alpha)$ is a penalty function that depends on the penalty parameter λ and on a weight $\omega_i > 0$. We consider different choices for the penalty functions as described below.

2.2.2.1 Ridge Regression (RR):

RR shrinkage was proposed by (14, 15) and consists of the following penalty function:

$$\lambda \sum_{i=1}^n p(\beta_i; \omega_i, \alpha) := \lambda \sum_{i=1}^n \beta_i^2. \quad (2-8)$$

RR has the advantage of having an analytical solution that is easy to compute and shrinks the irrelevant variables to zero. However, given the geometry of the penalty, the coefficients rarely reach exactly zero for any size of λ . Therefore, RR is not an sparsity-inducing method.

One interesting fact about RR is its relation to principal component (factor) models. Let \mathbf{X} be the centered $T \times n$ predictor matrix and consider its singular value decomposition $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}'$ with \mathbf{S} being a diagonal matrix with diagonal elements s_i , $i = 1, \dots, n$.

The forecasts of the dependent variable are given by

$$\mathbf{y}_{\text{ridge}} = \mathbf{X} \hat{\boldsymbol{\beta}}_{\text{ridge}} = \mathbf{X} (\mathbf{X}' \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}' \mathbf{y} = \mathbf{U} \text{diag} \left(\frac{s_i^2}{s_i^2 + \lambda} \right) \mathbf{U}' \mathbf{y},$$

whereas for the factor model with k factors are given by

$$\mathbf{y}_{\text{PC}} = \mathbf{X}_{\text{PC}} \hat{\boldsymbol{\beta}}_{\text{PC}} = \mathbf{U} \text{diag} \left(\underbrace{1, \dots, 1}_{k \text{ ones}}, \underbrace{0, \dots, 0}_{n-k \text{ zeroes}} \right) \mathbf{U}' \mathbf{y}.$$

However, this parallel to factor models does not hold exactly in our implementation as the variable set for the RR is larger than the one for the principal component factor construction as it includes four lags of each variable, autoregressive terms and the factors as well. Nevertheless, the comparison is useful to understand the potential differences in performance between RR and factor alternatives.

2.2.2.2

Least Absolute Shrinkage and Selection Operator (LASSO):

LASSO was originally proposed by (2). LASSO is similar to RR but penalizes the ℓ_1 norm of the coefficients as follows:

$$\lambda \sum_{i=1}^n p(\beta_i; \omega_i, \alpha) := \lambda \sum_{i=1}^n |\beta_i|. \quad (2-9)$$

LASSO shrinks the irrelevant variables to zero and has some good properties both in variable selection and goodness of fit. In order to achieve consistent variable selection, LASSO requires the irrepresentable condition¹ (IRC) to be satisfied (3). However, even if the IRC is not satisfied, LASSO still has the variable screening property, i.e., LASSO selects the relevant variables with high probability, but it may also select some extra variables.

2.2.2.3

Adaptive LASSO (adaLASSO):

adaLASSO was proposed by (16), who showed that the inclusion of some additional information regarding the importance of each variable could considerably improve the results. The adaLASSO does not need the IRC to have variable selection consistency and also has oracle properties, i.e., it not only selects the correct set of variables with high probability, but the coefficient distribution of these variables is also the same as the OLS estimation using only the correct set of variables. adaLASSO uses the same penalty as LASSO with the inclusion of a weighting parameter that comes from a first-step model that can be LASSO or even OLS:

$$\lambda \sum_{i=1}^n p(\beta_i; \omega_i, \alpha) := \lambda \sum_{i=1}^n \omega_i |\beta_i|, \quad (2-10)$$

where $\omega_i = |\beta_i^*|^{-1}$ and β_i^* are the coefficients from the first-step model. Finally, LASSO has some good properties for high-dimensional data. LASSO can handle many more variables than observations and works well in nonGaussian environments and under heteroskedasticity (17).

2.2.2.4

Elastic Net (EINet)

Elastic net (EINet) is a generalization that includes LASSO and RR as special cases. EINet is a convex combination of the ℓ_1 and the ℓ_2 norms (18). EINet also does regularization and selects the most relevant variables. Since

¹The irrepresentable condition imposes some restrictions on the correlation structure between the relevant and the irrelevant variables. In other words, the correlation between the two groups is bounded and must be small.

its penalty is between that of LASSO and RR, ElNet normally selects more variables than LASSO, at least for the same value of λ . The ElNet penalty is defined as follows:

$$\lambda \sum_{i=1}^n p(\beta_i; \omega_i, \alpha) := \alpha \lambda \sum_{i=1}^n \beta_i^2 + (1 - \alpha) \lambda \sum_{i=1}^n |\beta_i|; \quad (2-11)$$

where $\alpha \in [0, 1]$. We also consider an adaptive version of ElNet (adaElNet). This version works in the same way as the adaptive LASSO, i.e., we estimate a first-step model and use it to calculate the weights ω_i .

2.2.3

Factor Models

Factor models using principal components are very popular approaches to avoid the curse of dimensionality when the number of predictors is potentially large. The idea is to extract common components from all variables, thus reducing the model dimension.

Factors are computed as principal components of a large set of variables \mathbf{z}_t such that $\mathbf{F}_t = \mathbf{A}\mathbf{z}_t$, where \mathbf{A} is a rotation matrix and \mathbf{F}_t is the vector of the principal components. Consider equation (2-1). In this case, \mathbf{x}_t is given by y_{t-j} , $j = 0, 1, 2, 3$ plus \mathbf{f}_{t-j} , $j = 0, 1, 2, 3$, where \mathbf{f}_t is the vector with the first four principal components of \mathbf{z}_t . The assumptions and the theory behind factor models and when can we treat factors as observed variables can be found in (4, 5, 6).

2.2.3.1

Target Factors

To improve the forecasting performance of factor models, (6) proposed targeting the predictors. The idea is that if many variables in \mathbf{z}_t are irrelevant predictors of y_{t+h} , factor analysis using all variables may result in noisy factors with poor forecasting ability. The target factors are regular factor models with a pretesting procedure to select only relevant variables to be included in the factor analysis. Let $z_{i,t}$, $i = 1, \dots, q$ be the candidate variables and \mathbf{w}_t a set of fixed regressors that will be used as controls in the pretesting step. We follow (6) and use \mathbf{w}_t as AR terms of y_t . The procedure is described as follows.

1. For $i = 1, \dots, q$, regress y_{t+h} on \mathbf{w}_t and $z_{i,t}$ and compute the t statistics for the coefficient corresponding to $z_{i,t}$.
2. Sort all t statistics calculated in step 1 in descending order.
3. Choose a significance level α and select all variables that are significant using the computed t statistics.

4. Let $\mathbf{z}_t(\alpha)$ be the selected variables from steps 1–3. Estimate the factors \mathbf{F}_t from $\mathbf{z}_t(\alpha)$ by principal components.
5. Regress y_{t+h} on \mathbf{w}_t and \mathbf{f}_{t-j} , $j = 0, 1, 2, 3$, where $\mathbf{f}_t \subset \mathbf{F}_t$. The number of factors in \mathbf{f}_t is selected using the BIC. (6) also selected the number of lagged factors using the BIC.

The same procedure was used by (13). The authors showed that in most cases, target factors slightly reduce the forecasting errors compared to factor models without targeting.

2.2.3.2 Factor Boosting

The optimal selection of factors for predictive regressions is an open problem in the literature. Even if the factor structure is clear in the data, it is not obvious that only the most relevant factors should be included in the predictive regression. I adopt the boosting algorithm as proposed by (7) to select the factors and the number of lags that must be considered in the predictive regression. Define $\mathbf{z}_t \in \mathbb{R}^q$, the set of all n factors computed from the original n variables plus four lags of each factor. Therefore, $q = 5n$. Define also SSR as the sum of the squared residuals.

The algorithm is defined as follows:

1. Let $\Phi_{t,0} = \bar{y}$ for each t , where $\bar{y} = \frac{1}{t} \sum_{i=1}^t y_i$.
2. For $m = 1, \dots, M$:
 - (a) Compute $\hat{u}_t = y_t - \Phi_{t-h,m-1}$.
 - (b) For each candidate variable $i = 1, \dots, q$, regress the current residual on $z_{i,t}$ to obtain \hat{b}_i and compute $\hat{e}_{t,i} = \hat{u}_t - z_{i,t}\hat{b}_i$. Calculate $SSR_i = \hat{\mathbf{e}}_i' \hat{\mathbf{e}}_i$.
 - (c) Select i_m^* as the index of the variable which delivers the smallest SSR and define $\hat{\phi}_{t,m} = z_{i_m^*,t} \hat{b}_{i_m^*}$,
 - (d) Update $\hat{\Phi}_{t,m} = \hat{\Phi}_{t,m-1} + v\hat{\phi}_{t,m}$, where v is the step length. We set $v = 0.2$.
3. Stop the algorithm after the M th iteration or when the BIC of the last residuals starts to increase.

2.2.4

Ensemble Methods

Ensemble forecasts are constructed from a (weighted) average of the predictions of an ensemble of methods. In this section, I describe the techniques considered in this paper.

2.2.4.1

Bagging

The term “bagging” comes from bootstrap aggregation, which was proposed by (8). The idea is to combine forecasts from several unstable models. Normally, there is much more to gain from combinations of models if they are very different. The first source of instability is generated by re-estimating the model using bootstrap samples, and the second source comes from a pretesting step prior to the estimation, which for each bootstrap sample selects a subset of variables based on their statistical significance. The bagging steps are as follows:

1. For each bootstrap sample b , run a regression with all candidate variables and select those with $|t| \geq c$, where c is a pre-defined critical value.
2. Estimate a new regression only with the variables selected in the previous step.
3. The coefficients from the second regression are finally used to compute the forecasts **on the actual sample**.
4. Repeat the first three steps for B bootstrap samples and compute the final forecast as the average of the B forecasts.

We used $B = 100$. Note that in our case, the number of observations may be smaller than the number of variables, which makes the regression in the first step unfeasible. We solve this issue by introducing a new source of instability in the pretesting step. For each bootstrap sample we randomly divide all variables in groups and run the pretesting step for each one of the groups.

2.2.4.2

Complete Subset Regressions

CSR was developed by (19, 20). The motivation for developing CSR was that selecting the optimal subset of \mathbf{x}_t to predict y_{t+h} by testing all possible combinations of regressors is computationally very demanding, and in most cases, even unfeasible. Supposing that we have n candidate variables, the CSR

selects a number $q \leq n$ and computes all combinations of regressions using only q variables. The forecast of the model will be the average forecast of all regressions in the subset.

CSR deals well with a small number of candidate variables. However, for large sets, the number of regressions to be estimated increases very fast. For example, with $n = 25$ and $q = 4$, we need to estimate 12,650 regressions. As the number of candidate variables is much larger, we adopt a pretesting procedure similar to that used with the target factors. We start fitting a linear regression of y_{t+h} on each of the candidate variables (including lags) and save the t-statistics of each variable². The t-statistics are ranked by absolute value, and we select the \tilde{n} variables that are more relevant in the ranking. The CSR forecast is calculated on these variables. We used $\tilde{n} = 25$ and $q = 4$.

2.2.4.3 Jackknife Model Averaging

JMA is a different way to combine forecasts from several small models. Instead of using the naive average of the forecasts, JMA uses leave-one-out cross-validation to estimate optimal weights. The procedure I followed is that of (21) with some adjustments for time series as discussed in (22).

Suppose we have M candidate models that we want to average from and write the forecast of each model as $\hat{y}_{t+h}^{(m)}$, $m = 1, \dots, M$. Set the final forecast as

$$\hat{y}_{t+h} = \sum_{m=1}^M \omega_m \hat{y}_{t+h}^{(m)},$$

where $0 \leq \omega_m \leq 1$ for all $m \in \{1, \dots, M\}$ and $\sum_{m=1}^M \omega_m = 1$.

The JMA procedure is as follows:

1. For each observation of (\mathbf{x}_t, y_{t+h}) :
 - (a) Estimate all the candidate models leaving the selected observation out of the estimation. Since we are in a time series framework with k lags in the model, we also removed k observations before and k observations after (\mathbf{x}_t, y_{t+h}) .
 - (b) Compute the forecasts from each model for the observations that were removed in the previous step.
2. Choose the weights that minimize the cross-validation errors subject to the constraints previously described.

²We did not use a fixed set of controls, \mathbf{w}_t , in the pretesting procedure like we did for the target factors.

The minimization problem above is quadratic and has the restriction that w must be positive and sum to 1. The problem does not have a closed solution but can be easily solved using the quadprog package in R. Given our set of candidate variables, each candidate model in the JMA has four autoregressive lags of the dependent variable and four lags of one candidate variable.

2.2.5

Regression Trees and Random Forests

The RF methodology was initially proposed by (9) as a solution to reducing the variance of regression trees and is based on bootstrap aggregation (bagging) of randomly constructed regression trees. In turn, regression trees are flexible nonparametric predictive models that recursively partition the set of explanatory variables, \mathbb{X} , into subsets, each modeled using regression methods; see (8).

To understand how a regression tree works, an example from (1) is useful. Consider a regression problem in which X_1 and X_2 are explanatory variables, each taking values in some given interval, and Y is the dependent variable. We first split the space into two regions, at $X_1 = s_1$, and then, the region to the left (right) of $X_1 = s_1$ is split at $X_2 = s_2$ ($X_1 = s_3$). Finally, the region to the right of $X_1 = s_3$ is split at $X_2 = s_4$. As illustrated in the right plot of Figure 2.1, the end result is a partitioning of \mathbb{X} into five regions: R_m , $m = 1, \dots, 5$. In each region R_m , we assume that the model predicts Y with a constant c_m , which could be estimated, for example, as the sample average of realizations of Y that “fall” within region R_m . A key advantage of this recursive binary partition is that it can be represented as a single tree, as illustrated in the left plot of Figure 2.1. Each region corresponds to a terminal node of the tree.

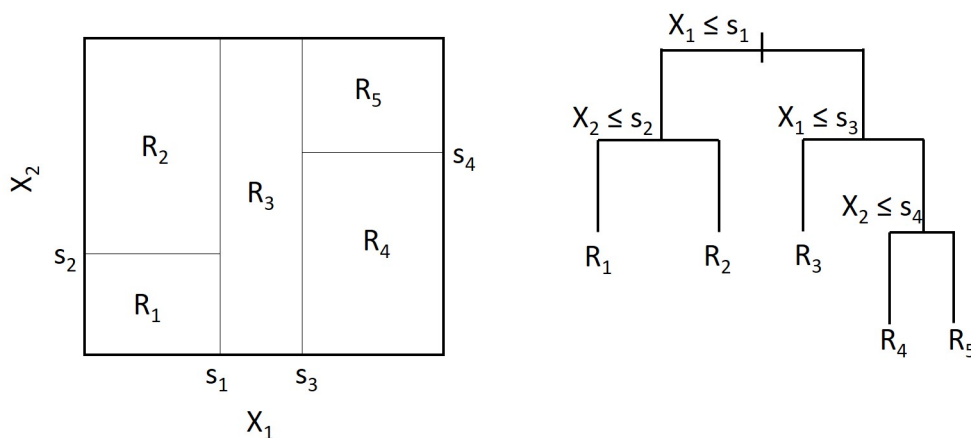


Figure 2.1: Example of a regression tree. Reproduction of part of Figure 9.2 in (1).

Now we turn to the question as to how to choose splitting variables and split points, i.e., how to grow a tree, when there are p explanatory variables. Let $x_t = (x_{1,t}, x_{2,t}, \dots, x_{p,t})$, for $t = 1, \dots, T$, where $x_{i,t}$ is the realization of variable X_i in period t .

We proceed backwards. Suppose that after choosing the splitting variables and split points, we reach M regions. If we adopt the sum of squared errors as our minimization criterion, the prediction of Y at T , \hat{c}_m , is simply the average of previous realizations y_t such that x_t belongs to R_m . Algebraically, for $m = 1, \dots, M$,

$$\hat{c}_m = \arg \min \sum_{t=1}^T \mathbf{I}(x_t \in R_m)(y_t - c_m)^2 = \frac{\sum_{t=1}^T \mathbf{I}(x_t \in R_m)y_t}{\sum_{t=1}^T \mathbf{I}(x_t \in R_m)}, \quad (2-12)$$

where \mathbf{I} is the indicator function.

The idea is to use the sum of squared errors to inform how to grow the regression tree. To begin, consider a splitting variable j and a split point s to partition \mathbb{X} into two regions, namely, $R_1(j, s) = \{X|X_j \leq s\}$ and $R_2(j, s) = \{X|X_j > s\}$. Then, seek the pair (j, s) that solves

$$\min_{j,s} \left[\min_{c_1} \sum_{t=1}^T \mathbf{I}(x_t \in R_1(j, s))(y_t - c_1)^2 + \min_{c_2} \sum_{t=1}^T \mathbf{I}(x_t \in R_2(j, s))(y_t - c_2)^2 \right].$$

Once the best split is found, we proceed iteratively, repeating this process on each of the resulting regions.

A natural question arises: when should we stop this process? A very large tree might overfit the data, which would be highly unstable. However, a tree that is too small might not capture a complex nonlinear relation between variables in the data. One possibility to address this trade-off is the *cost-complexity pruning* method described in Hastie et al (2009). Instead, we follow the RF method, which applies the essential idea of bagging, i.e., RF reduces the variance by averaging many noisy and unbiased models. The drawback is the loss of interpretability.

An RF is a collection of regression trees, each specified in a bootstrapped subsample of the original data. Suppose there are B bootstrapped subsamples. For each subsample, obtain a prediction for Y by applying a modified version of the aforementioned splitting iterative process until a prespecified minimum number of observations, say five, is reached in any resulting region. In particular, the modification is to select q variables at random from the p explanatory variables at each step of the process. Finally, simply average the predictions of Y across the B bootstrapped subsamples. Since we are dealing with time series, bootstrapped samples are calculated

using block bootstrapping.

The main advantages of the RF method are twofold: RF can handle both a very large number of explanatory variables and complex nonlinear relationships between variables.

2.2.6 Hybrid Linear-Random Forests Models

RF/OLS and adaLASSO/RF deserve some special attention because these are adaptations made specifically to answer how important the variable selection is and the nonlinearity in forecasting the US inflation. RF/OLS is estimated using the following steps:

1. For each bootstrap sample b :
 - (a) Grow a single tree with k nodes (we used $k = 20$) and save the $N \leq k$ split variables,
 - (b) Run an OLS on the selected splitting variables,
 - (c) Compute the forecast \hat{y}_{t+h}^b .
2. The final forecast will be $\hat{y}_{t+h} = B^{-1} \sum_{b=1}^B \hat{y}_{t+h}^b$ where B is the number of bootstrap samples.

The main objective of the RF/OLS is to check the performance of a linear model using variables selected from the RF. If the results are very close to the full RF, we understand that nonlinearity is not an issue, and the RF is superior solely because of variable selection. However, if we see some improvement on accuracy compared to other linear models, especially bagging³, but if RF/OLS is still less accurate than RF, we have evidence that both nonlinearity and variables selection play an important role.

The second adapted model is LASSO/RF, where we use the adaptive LASSO for variable selection and then estimate a fully grown RF with the variables selected by adaptive LASSO. If LASSO/RF performs similarly to RF, we understand that the variable selection in RF is irrelevant, and the only thing that matters is the nonlinearity. LASSO/RF and RF/OLS together create an "if and only if" situation where we test the importance of variable selection and nonlinearity from both sides. Our results point to the middle case where nonlinearity and variable selection are both important. The two adapted models perform very well compared to other linear specifications, but RF is more accurate than both. In other words, the good performance of RF is driven by both variable selection and nonlinearity.

³Bagging and RF are bootstrap-based models, the first of which is linear and the second is nonlinear.

2.2.7

Bayesian Vector Autoregressive Model

The Bayesian VAR that we used is capable of estimating large models with many variables and lags with a small computational cost. The idea is to set the priors of most coefficients in the VAR to zero ⁴. The trick is to inflate the model with dummy observations as in (10) in a way that the coefficients of the inflated model are estimated by OLS and the model replicates a VAR estimated with the desired prior.

Let $\mathbf{Y}_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})'$ be described as the following VAR model:

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \mathbf{u}_t, \quad (2-13)$$

where \mathbf{c} is a n -dimensional vector of constants, \mathbf{A}_i , $i = 1 \dots p$ are $(n \times n)$ matrices of coefficients and \mathbf{u}_t is the n -dimensional error vector. The same model may be written as a system of equations:

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{U} \quad (2-14)$$

where $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_T)'$ is a $(T \times n)$ matrix, $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_T)'$ is a $(T \times k)$ matrix with $k = np + 1$ and $\mathbf{X}_t = (1, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_{t-p})'$, $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_T)'$, and $\mathbf{B} = (\mathbf{c}, \mathbf{A}_1, \dots, \mathbf{A}_p)'$.

The model is estimated using dummy observations \mathbf{Y}_d and \mathbf{X}_d of dimensions $T_d \times n$ and $T_d \times k$ respectively (for details on creating the dummy observations see (10)). Using these dummies is equivalent to imposing the normal inverted Wishart prior on the covariance matrix of \mathbf{B} . The dummy observations are used to create $\mathbf{Y}^* = (\mathbf{Y}', \mathbf{Y}'_d)'$ and $\mathbf{X}^* = (\mathbf{X}', \mathbf{X}'_d)'$. The posterior mean of \mathbf{B} is the same as the ordinary least-squares (OLS) estimates of the regression of \mathbf{Y}^* on \mathbf{X}^* and also the same as the Minnesota prior. Additionally, by using the dummies we ensure that for each regression of the VAR the number of observations is larger than the number of variables, which makes the OLS estimation of \mathbf{B} feasible. Another important issue is the choice of the expected value of the priors for the diagonal of the \mathbf{A}_1 matrix. We choose the value of 0.5 for all elements.

⁴These priors have a shrinkage interpretation in the model. The coefficients only deviate from the prior if the data is informative.

3

Forecasting Brazilian inflation with high-dimensional models

JEL: C22.

Keywords: Emerging economies, Monetary policy, Brazilian inflation, Forecasting, LASSO, Shrinkage, Model selection.

Abstract: In this paper, we use high dimensional models, estimated by the least absolute shrinkage and selection operator (LASSO) to forecast Brazilian inflation. The models are compared to benchmark specifications such as linear autoregressive (AR) and the factor models based on principal components. Our results show that the LASSO-based specifications have the smallest errors for short-horizon forecasts. However, for long horizons, the AR benchmark is the best model with respect to point forecasts, even though there is no significant difference between them. The factor model also produces some good long-horizon forecasts in a few cases. We estimated all the models for the two most important Brazilian inflation measures, the IPCA and the IGP-M indexes. The results also show that there are differences in the selected variables for both measures. Finally, the most important variables selected by the LASSO-based models are, in general, related to government debt and money. On the other hand, variables such as unemployment and production are rarely selected by the LASSO. Therefore, our evidence is against the Phillips curve as the driving mechanism of Brazilian inflation.

3.1 Introduction

Inflation is a hard-to-forecast economic variable. However, reliable forecasts provide us with a more transparent economic environment. For example, companies need to know the future inflation to make their investment decisions; investors use inflation forecasts to know their real gains; wages are adjusted based on future inflation; and contracts such as rent and mortgages use it as an index to correct future prices. Furthermore, Central Banks usually need reliable inflation forecasts in order to conduct monetary policy (inflation targeting). For countries like Brazil, which had a long and recent hyperinflation experience, it is even more important to understand the components and future behavior of inflation.

In 1999, the Brazilian government adopted inflation-targeting policies to make the economy more transparent and stable. This changed the dynamics of inflation and created new challenges to keep it on the target. For example, volatility in the exchange rate had to be controlled since it is directly connected to inflation (23). (24) argues that inflation targets may be a good alternative for monetary policies in emerging countries; however, they have to be handled carefully. In the Brazilian case, they have brought lower inflation rates and more stable prices (25, 26).

There is an extensive recent literature on the determinants of Brazilian inflation, especially on the Phillips curve and its relation to unemployment (27, 28, 29). But some of these results are controversial, especially when different proxies and different periods of analysis are considered (29). The literature on inflation forecast is smaller, but also rich. More recently, (30) used several linear and nonlinear models and the Phillips curve to forecast inflation. The authors showed that some nonlinear models and the simple autoregressive (AR) model produce smaller forecast errors than the Phillips curve. (31) used long-memory heteroskedastic models to show that Brazilian inflation has long-range dependence both on the mean and on the variance. However, they do

not exclude the importance of the short-term AR component. The relevance of past inflation is also pointed out by (32).

Although the Phillips curve is the main theoretical framework to model inflation, there is an increasing interest in finding other variables that may describe its dynamics. (33) did an extensive work with a large list of macroeconomic variables and showed that the best variables to forecast inflation are, in general, linked to production. However, these variables can be very different from what is expected in the traditional Phillips curve unemployment framework. Additionally, (34) showed that the Phillips curve cannot produce better forecasts when compared with simpler naive models. These results led to the search for models and variables to forecast and explain inflation, such as commodity prices (35), financial variables (36), and several economic activity and expectation variables (37). In many cases, these variables were relevant to inflation forecast.

The main goal of this paper is to forecast inflation considering a large number of potential variables. We also analyze the selected variables and show their possible interpretation in the model. We use three types of models. First, the simple AR model, which uses only past inflation as a predictor. Second, since our dataset has more than 100 variables, we also used factor models. These models use principal components as explanatory variables. We selected the number of components using (4) criterion¹. The last class of models is based on shrinkage, more precisely on the LASSO (least absolute shrinkage and selection operator, (2)) and on the adaLASSO (adaptive least absolute shrinkage and selection operator, (3)). These models try to select the relevant variables using a penalty on the loss function. Although they may produce estimates with in-sample bias, they significantly reduce out-of-sample forecast errors and variance. The adaLASSO is an improved version that corrects some of the problems with the LASSO.

¹We adopted the BIC_3 criterion, which selected five factors. The PC_i and IC_i selected around eight factors. We selected the BIC_3 for parsimony, since our data have a small T and we are also using lagged factors and autoregressive components.

Our research contributes to the literature in several ways. First, we use high dimensional models to forecast Brazilian inflation in the same spirit as in (13). Most of the previous works use only Phillips curve related variables or simple univariate time-series models. Our results show that LASSO and adaLASSO models produce reliable inflation forecasts for horizons up to four months ahead. However, even with the inclusion of many candidate variables, the best model for longer horizons is the simple AR model. Second, we analyze and forecast the two most widely used Brazilian inflation measures and we are able to point several differences between them, especially for short-horizon models, where the selected candidate variables are different for both measures. Third, although our main goal is to forecast inflation, we have some new evidence on which variables are relevant to explain the Brazilian inflation dynamics. We show that for the official inflation (IPCA), nearly all the selected variables are related to government debt. Although we do not establish any causality relationship, it is important to point out that these variables were selected among more than 100 candidates. We also show that for the alternative inflation measure (IGP-M), variables related to the amount of money and credit in the economy are relevant predictors as well. These findings are evidence against the Phillips curve as the driving mechanism of Brazilian inflation. The IPCA is driven basically by an inertial component and variations in government debt. The IGP-M has other driving factors; however, they are mostly linked to debt through the issue of government bonds.

This chapter is organized as follows. Section 3.2 presents the LASSO and the adaLASSO models and how they select the relevant variables amongst all the candidates; in Section 3.3, we discuss the characteristics of the inflation measures used and show the main results. Section 3.4 concludes.

3.2

LASSO models

The least absolute shrinkage and selection operator (LASSO) was initially proposed by (2). The idea behind the method is to shrink irrelevant coefficients in a regression to zero. This is done by adding a penalty to the loss function that penalizes directly the estimated parameters, excluding those which are irrelevant. The LASSO estimator is defined as:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\hat{\boldsymbol{\beta}}} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \sum_{j=1}^p |\beta_j|, \quad (3-1)$$

where $\boldsymbol{\beta}$ is the $n \times 1$ vector of parameters, $\mathbf{Y} = (y_1, \dots, y_n)'$, \mathbf{X} is the $P \times n$ data matrix and λ is the shrinkage parameter. Figure 3.1 shows how the LASSO works: on the x axis, we have the λ and on the y axis, the coefficients; each line represents a different variable. The figure shows that the number of selected variables and the size of the coefficients decrease as we increase the shrinkage parameter. Moreover, the LASSO can be used when one is dealing with more variables than observations (16, 38).

If we define $\hat{\boldsymbol{\beta}}_0$ as the ordinary least squares (OLS) estimator without any shrinkage and λ_{ols} as the largest penalty parameter that delivers the same result as the OLS (no shrinkage), there will be shrinkage for any $\lambda > \lambda_{ols}$. Additionally, let λ_c be the smallest penalty parameter that results in a model only with a constant term. Then, we must choose λ from the interval $\lambda_{ols} \leq \lambda \leq \lambda_c$. The best λ is chosen using cross-validation or some information criterion such as the Bayesian Information Criterion (BIC).

(16) and (3) showed that the LASSO does not have the oracle property as defined by (39), i.e., the LASSO may fail to select the correct subset of relevant variables and its parameters may not have the same asymptotic distribution as the OLS estimator with only the relevant variables. To solve these issues, (3) proposed the adaptive LASSO (adaLASSO). It consists of a two-step estimation that uses a first model to generate different weights for each candidate variable. These weights are then used in the LASSO as

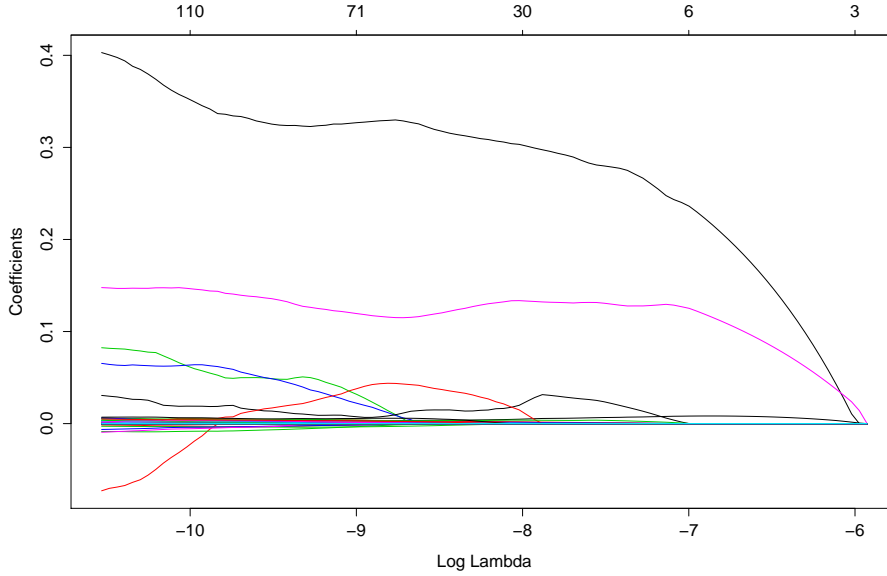


Figure 3.1: Shrinkage coefficients - LASSO

additional information. The adaLASSO estimator is defined as:

$$\hat{\beta} = \arg \min_{\hat{\beta}} \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|_2^2 + \lambda \sum_{j=1}^p w_j |\beta_j|, \quad (3-2)$$

where $w_j = |\hat{\beta}_j^*|^{-\tau}$, $\hat{\beta}_j^*$ is a first-step estimate and τ is a parameter chosen by using the same criterion as λ .

In order to maximize the model’s predictive power, it is usual to select the parameters λ and τ by using cross-validation. However, in a time-series case when the data are not independent and identically distributed (i.i.d.), (40) and (41) showed that the BIC is a reliable alternative since cross-validation in a non-i.i.d. framework may be complicated. Moreover, (17) showed that selecting the parameters λ and τ in a time-series environment by using the BIC and the LASSO as the first step for the adaLASSO yields estimates that have the oracle property even in adverse situations with heteroskedasticity and t distributed errors. Moreover, the authors allow the number of candidate variables to increase with the number of observations and show that, based on these conditions, the adaLASSO has model selection consistency, i.e., it chooses the most parsimonious model asymptotically. These are very strong results since, by assuming that we have the oracle in our set of variables and

asymptotic conditions are met, the adaLASSO will select the right variables and their distribution will be the same as the OLS estimator with the correct variables. Moreover, it will not select useless variables.

3.3

Main Results

In this section, we show the results for the two most important Brazilian price indexes: the IPCA, calculated by the Brazilian Institute of Geography and Statistics (IBGE); and the IGP-M, calculated by the Getúlio Vargas Foundation (FGV). Figure 3.2 shows the inflation calculated by both indexes. Our dataset consists of 102 monthly variables that cover production, government debt, price indexes, financial markets, taxes, import and export of goods and services, government accounts, savings, investment, wages, international variables that may be related to the Brazilian economy, etc. These variables came from the Brazilian Central Bank, the FGV, the IBGE, the IPEADATA, and the Bloomberg database. The period of analysis spans from January 2000 to December 2013. We selected these data because the inflation-targeting policy in Brazil started only in July 1999, and before that, the inflation-generating process was probably different. Additionally, the Brazilian economy was still adapting to the new currency, implemented in 1994, and to the economic policy during the late 1990s.

We estimated simple AR models, factor models using (4) BIC_3 criterion to select the number of principal components, and the LASSO and adaLASSO models using the BIC to choose the pair² (λ, τ) and the LASSO as the first step for the adaLASSO. All variables were tested for unit-root and first-differentiated when necessary. Moreover, we used four lags of the candidate variables and monthly dummies as possible predictors of inflation. The models were estimated using a rolling window of 132 observations and the forecasts were compared using the (42) test. The out-of-sample period goes from January

²The pair (λ, τ) is determined in-sample (See the Appendix for more information).

2011 to December 2013. Finally, we estimated models and forecasts for 1 to 12 steps ahead. The estimated equation is defined in (3-3):

$$\pi_{t+h} = \alpha_0 + \sum_{i=0}^3 \gamma_i \pi_{t-i} + \sum_{i=0}^3 \beta' \mathbf{x}_{t-i} + u_{t+h}, \quad (3-3)$$

where, π_{t+h} is the inflation calculated as $\frac{(P_{t+h} - P_{t+h-1})}{P_{t+h-1}}$, α_0 is a constant term, \mathbf{x}_t is the vector containing all candidate variables and u_t is an error term. Equation (3-3) states that the monthly inflation h periods ahead is a function of today's and of the previous 3 months' inflation and of other control variables.

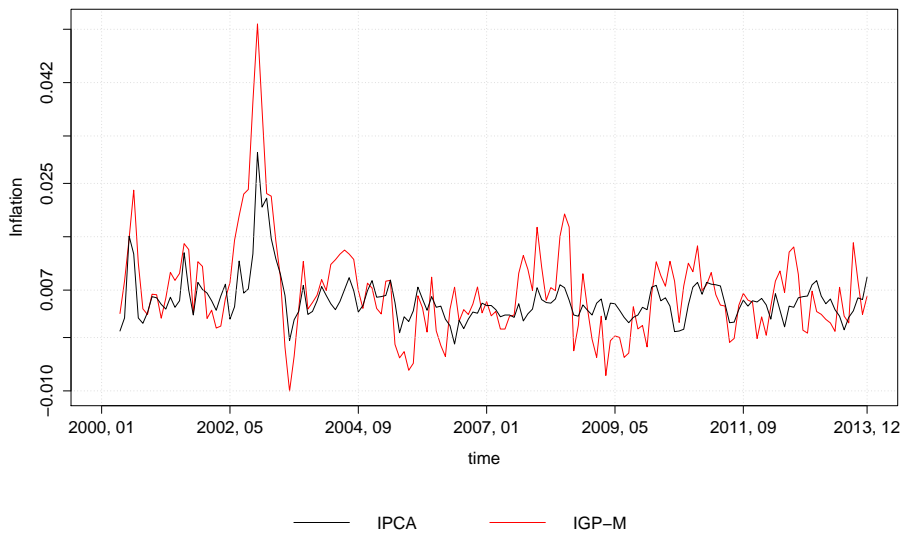


Figure 3.2: IPCA and IGM-M time series

3.3.1 IPCA Inflation Index

The IPCA is the Brazilian official price index. It is calculated monthly by IBGE using data from families that earn between 1 and 40 monthly minimum wages and live in urban areas.

The first noteworthy result is that Brazilian inflation is highly autocorrelated, which means that simple AR models tend to provide good forecasts (30, 31). Figure 3.3 shows that the first-order autocorrelation of the IPCA inflation exceeds 0.6 and remains high up to the fourth lag. Moreover,

there is also a small seasonal behavior. We included the monthly dummies as candidate variables to control for this seasonality. However, it is more likely that the dummies will be excluded by the LASSO and the seasonality will be captured by other candidate variables with a similar behavior. To contextualize, the autocorrelation of US inflation, if we consider the CPI, is also around 0.6 in the first lag for monthly data. However, it decreases significantly slower than the Brazilian autocorrelation, which is less than 0.3 in the third lag and statistically zero in the fifth one. In the US case, autocorrelation remains close to 0.6 up to the 10th lag and, in the 25th lag, it gets close to 0.4.

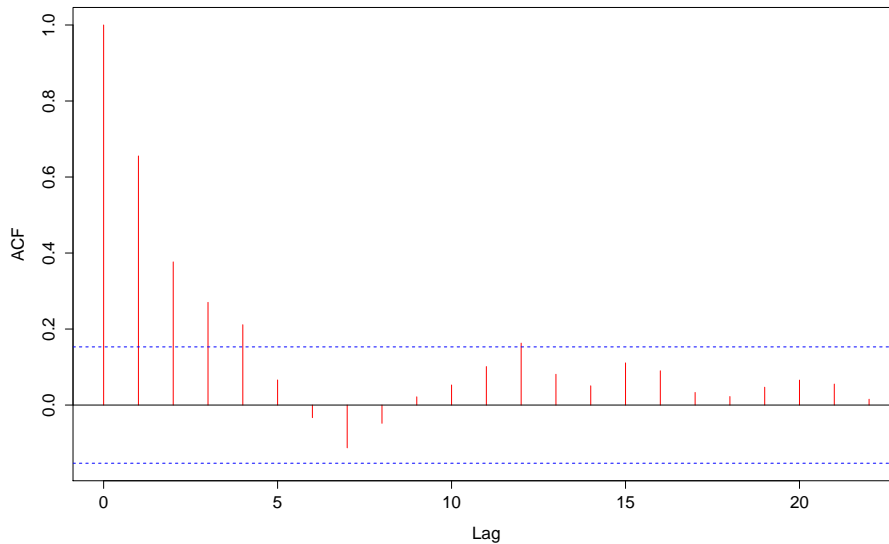


Figure 3.3: Autocorrelation function: IPCA

Figure 3.4 shows the root mean squared error (RMSE) and the mean absolute error (MAE) for the AR, the factor model, the LASSO, and the adaLASSO specifications for all the 12 forecast horizons. Regarding the RMSE in (a), the adaLASSO has smaller error for short horizons (up to 4 months) and the AR is the best model on most of the remaining horizons. The LASSO has errors slightly bigger than the adaLASSO and the factor model performed badly up to period 8, but it presented the smallest errors for 9 and 10 months ahead. The results for the MAE (panel (b)) are similar to those of the

RMSE; the most important differences are that the AR has a slightly better performance in the $t + 2$ forecasts and the LASSO and the adaLASSO have basically the same error in $t + 6$. The adaLASSO had a few large errors, which are more penalized by the RMSE. The mean absolute error for the adaLASSO in $t + 1$ is approximately 0.13%. If we consider all the forecast horizons, the average MAE is approximately 0.166% and 0.167% for the adaLASSO and the AR, respectively. If we consider the first four horizons, the errors are 0.151% and 0.167%; on the last 8 horizons, they are 0.171% and 0.164%.

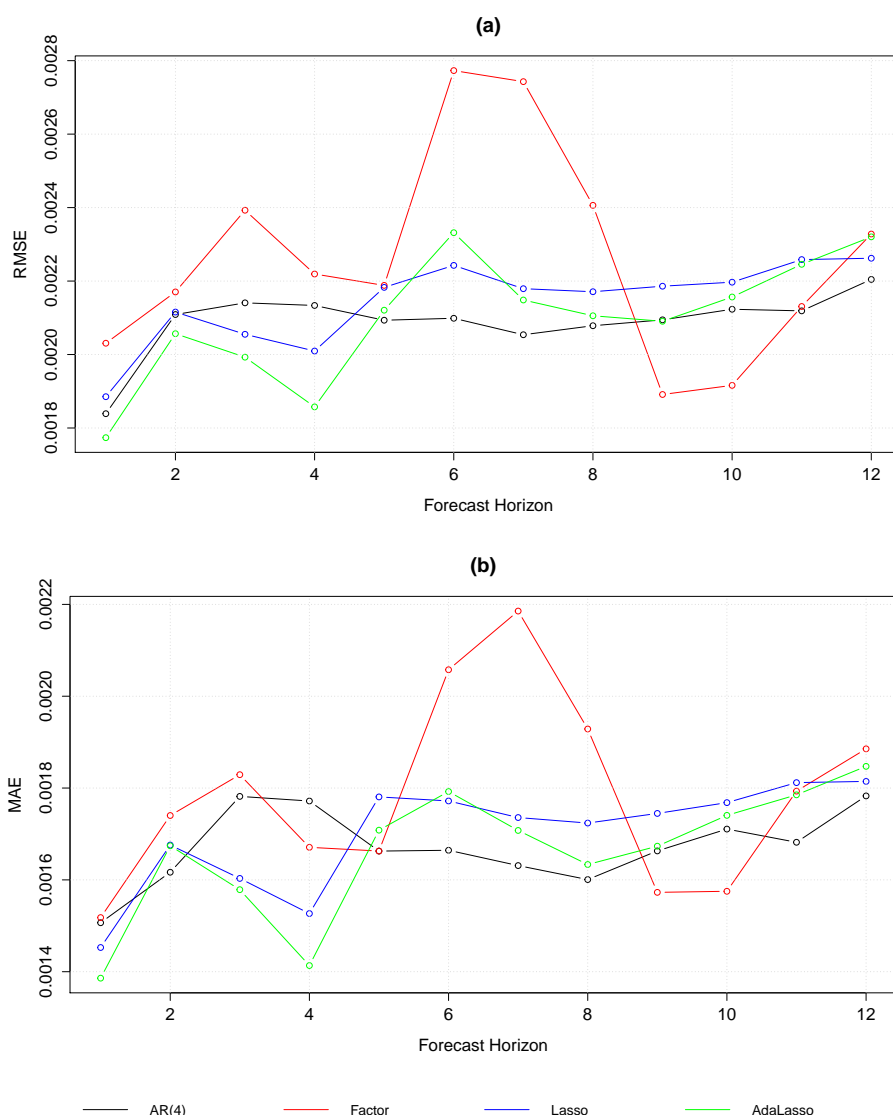


Figure 3.4: RMSE and MAE of all models: IPCA

Figure 3.5 shows the average number of variables selected by the LASSO

and the adaLASSO on all forecast horizons. The adaLASSO is, in general, more parsimonious, which is natural, since its first step was the LASSO. However, starting in $t + 6$, both models select basically the same variables; and from $t + 8$ to $t + 12$, most models select zero variables, becoming constant-only models. Considering that the adaLASSO has the oracle property, there are three possible conclusions: 1) the Brazilian monthly inflation cannot be predicted on long horizons using macroeconomic variables available today, i.e., no variable is relevant; 2) the relevant variables are not in our database, which is very unlikely since we covered all types of variables; 3) many variables are relevant for long horizons, but their coefficients are very small and the LASSO forces them to be zero. The most plausible hypotheses are 1 and 3. To test the third hypothesis, we re-estimated the adaLASSO using the ridge regression as the first step, but the results were the same. The ridge regression is similar to the LASSO; however, it penalizes the squared parameters instead of their absolute value. The squared penalization never excludes variables, i.e., their coefficients can be very small, but they do not reach zero. Using the ridge as the first step gives more liberty for the adaLASSO to choose the variables, since none of them were previously excluded.

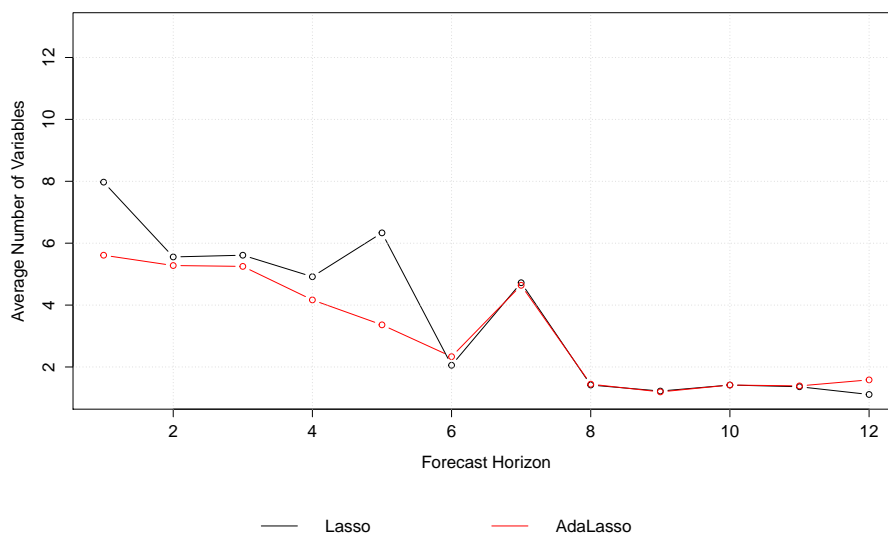


Figure 3.5: Number of selected variables: IPCA

Table 3.1 shows the correlation between the forecasts and the Giacomini and White (GW) test (42) p -values to check if the forecasts are statistically equal. We show the results for $t+1$, $t+3$, $t+6$, and $t+12$. The results show that, although some models perform better than others on different horizons, their forecasts are indeed statistically equal even when their correlation is small. Unfortunately, our dataset does not allow us to have more than 36 out-of-sample periods to check the robustness of these results. However, since the forecasts are statistically very similar, the adaLASSO forecasts are at least as good as the best forecast. This evidence supports that the variables selected by the model are indeed relevant to generate the dynamics of the IPCA.

Figure 3.6 shows the cumulative squared errors for the same horizons as in Table 3.1. It shows that, even though some models have smaller errors, they have a very similar behavior (excluding the factor model) especially on the horizons $t+6$ and $t+12$. While the factor model forecasts are apparently very different from the others, especially on $t+6$, the GW test failed to capture this difference in most cases.

$t+1$	AR(4)	Factors	LASSO	adaLASSO	$t+3$	AR(4)	Factors	LASSO	adaLASSO
AR(4)	1	0.48	0.75	0.74		1	0.22	0.46	0.33
	-	(0.037)	(0.268)	(0.405)		-	(0.168)	(0.161)	(0.155)
Factors	0.48	1	0.36	0.41		0.22	1	0.35	0.41
	(0.037)	-	(0.198)	(0.390)		(0.168)	-	(0.343)	(0.286)
LASSO	0.75	0.36	1	0.96		0.46	0.35	1	0.89
	(0.268)	(0.198)	-	(0.380)		(0.161)	(0.343)	-	(0.245)
adaLASSO	0.74	0.41	0.96	1		0.33	0.41	0.89	1
	(0.405)	(0.390)	(0.380)	-		(0.155)	(0.286)	(0.245)	-
$t+6$					$t+12$				
AR(4)	1	0.23	0.49	0.13		1	0.50	0.77	0.27
	-	(0.297)	(0.172)	(0.349)		-	(0.329)	(0.458)	(0.496)
Factor	0.23	1	0.14	0.30		0.50	1	0.20	0.12
	(0.279)	-	(0.426)	(0.459)		(0.329)	-	(0.212)	(0.151)
LASSO	0.49	0.14	1	0.68		0.77	0.20	1	0.37
	(0.172)	(0.426)	-	(0.185)		(0.458)	(0.212)	-	(0.340)
adaLASSO	0.13	0.30	0.68	1		0.27	0.12	0.37	1
	(0.349)	(0.459)	(0.185)	-		(0.496)	(0.151)	(0.340)	-

This table shows the correlation and the p -value of the Giacomini and White test for the forecasts in Figure 3.4.

The p -values are shown in parentheses. The null hypothesis of the GW test is that the forecasts are statistically equal.

Table 3.1: Forecast correlations and GW test p -values: IPCA

Figures 3.7 and 3.8 present the selected variables in the LASSO and

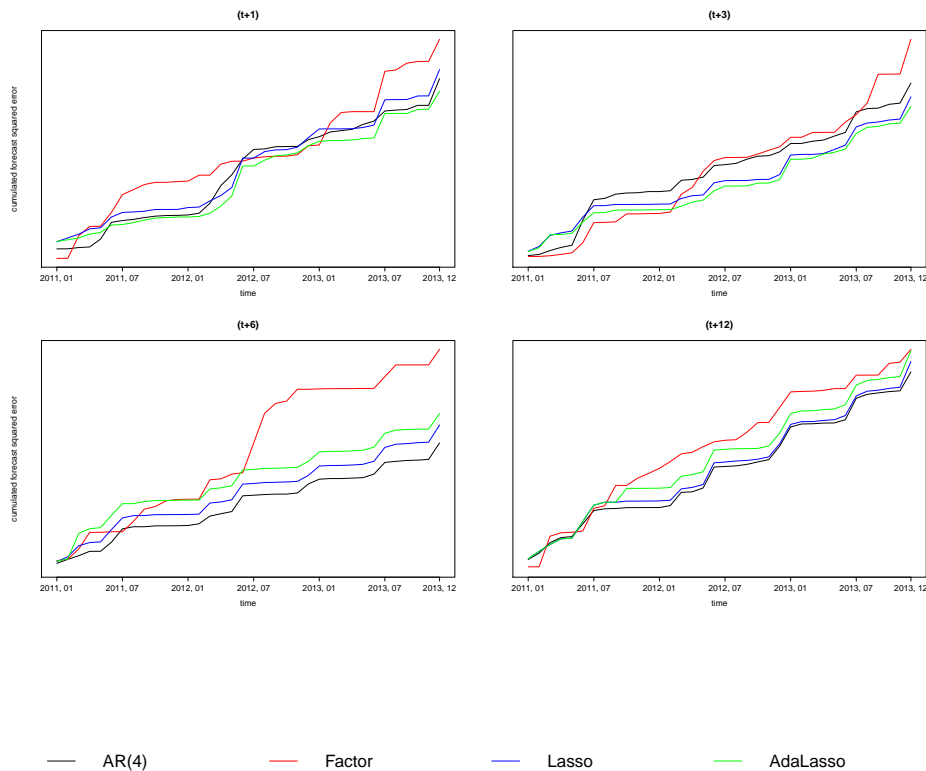


Figure 3.6: Cumulative squared errors: IPCA

adaLASSO for several forecast horizons³. The $t + 1$ forecast models have more variables than the others and the only horizon with AR terms. Factors were not selected, except for one period in the LASSO. Amongst the selected variables in the adaLASSO case are other inflation indexes, such as the IGP-DI, the federal government debt, the dollar exchange rate, state and municipal debt, among others. Government debt variables are selected very often for the $t + 1$ and occasionally for $t + 3$ models, especially state and municipal debt. Since this type of variable is first differentiated, when the variation of the debt is very big and positive, it probably means that the government is spending more money and putting upward pressure on inflation. The results for $t + 6$ and $t + 12$ show the same information as Figure 3.5, i.e., most of the models are constant-only, with government debt variables and unemployment appearing a

³We do not have space to show exactly which variables were selected in each model in the paper. The exact variables are available from the authors upon request.

few times. Finally, there was an unusual result in the $t + 12$ adaLASSO, which selected 17 variables. This model in particular had a forecast error of nearly 40%.

Table 3.2 shows how the selected variables are distributed. The first row displays the proportion of other inflation indexes, AR components, and dummy variables. In many cases, the AR terms were replaced with other indexes, especially with the IGP-DI. The second row shows the proportion of the variables which are not included in the first row, i.e., variables that are mostly unrelated to prices. We define this group as leading economic variables. They are selected more often than those in the first row, meaning that there is relevant information which is not the simple AR component of the inflation. The second group of rows (3-7) shows how the economic variables are distributed. In the $t + 1$ case, government debt (especially municipal and state debt) accounts for nearly 70% of all the economic variables selected by the model. Therefore, an increase in government debt is incorporated very fast into inflation. If we consider all the horizons, money variables, exchange rates and unemployment become more representative. However, these three categories together are selected approximately in the same proportion as the government debt variables alone. Money variables are mostly M3 and M4; in a very general way, M3 considers all the paper money held by the population, demand and time deposits, applications in investment funds and savings whereas M4 is M3 plus government bonds.

Brazil historically has a very interventionist government and our results are evidence that government debt is the most important variable to forecast inflation, excluding its AR component. This result may be related to the poor performance of the Phillips curve in forecasting inflation in (30) and to the ambiguous results in different studies of Brazilian inflation pointed out by (29). Finally, our results suggest that government debt is incorporated very fast into inflation. It is more expressive on the first forecast horizon than if we

look at all the horizons together.

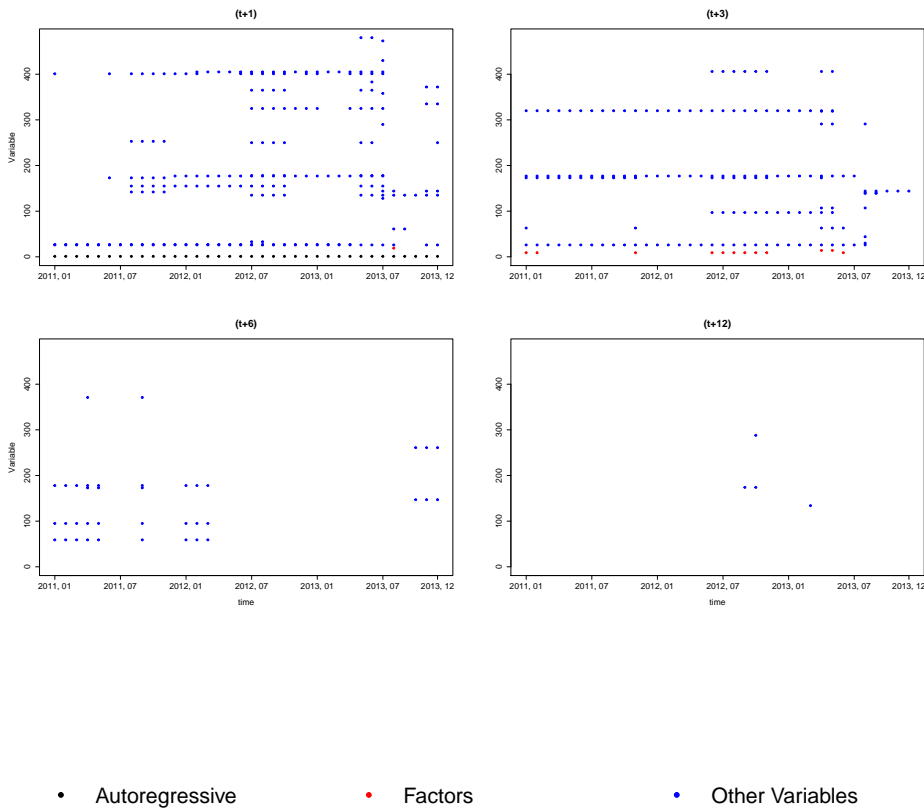


Figure 3.7: Variables selected by the LASSO: IPCA

3.3.2 IGP-M Inflation Index

The IGP-M is also a very important index for measuring inflation in Brazil. It is used in several contracts, such as energy prices and rent. Its construction is basically the same as that of the IGP-DI; the only difference is that the IGP-M refers to the 21st day of the previous month until the 20th day of the current month, while the IGP-DI covers all days of the current month.

The IGP-M is more volatile than the IPCA. Its coefficient of variation is 1.24, much larger than that of the IPCA, which is 0.73. Moreover, the IGP-M is also larger. In our sample, it has an average of 0.65% per month against

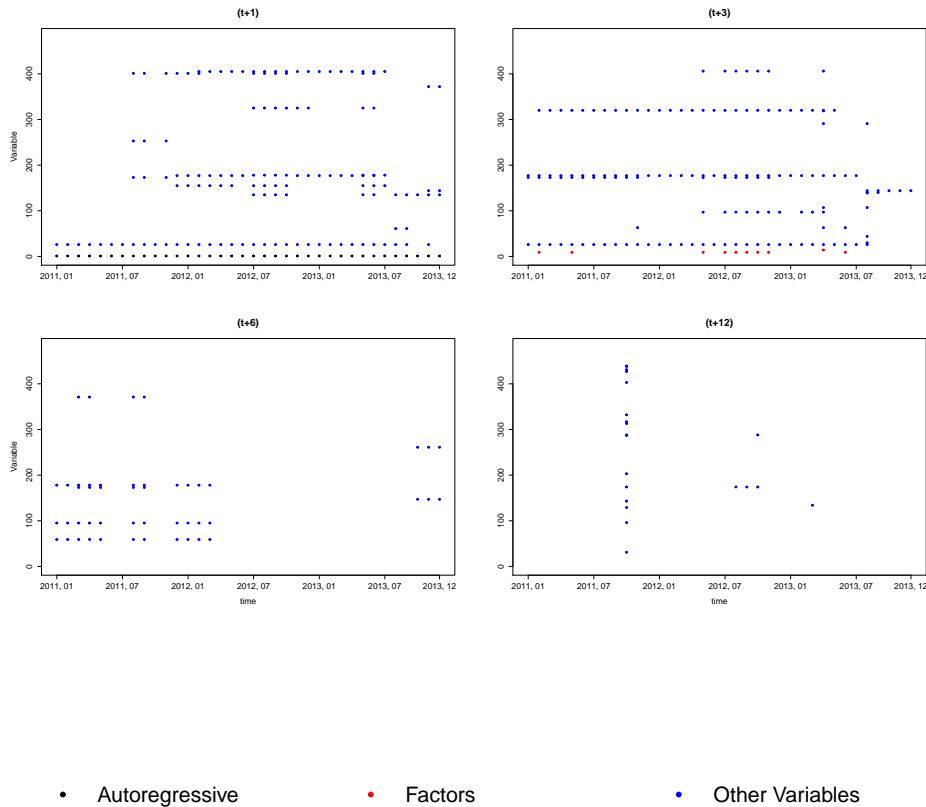


Figure 3.8: Variables selected by the adaLASSO: IPCA

0.52% of the IPCA. Finally, the correlation between the two indexes is 0.73. This is a low correlation considering that the two most important American inflation measures, the CPI and the PCE, have a correlation of approximately 0.90. The IGP-M has an autocorrelation even bigger than that of the IPCA. Figure 3.9 shows that the first-order autocorrelation is approximately 0.75.

Figure 3.10 shows the RMSE and the MAE for all IGP-M models on all the 12 forecast horizons. The first important thing to point out is that the factor model is more stable in the IGP-M than in the IPCA, even though it still produces larger errors. The LASSO had a very bad performance on the horizons $t + 6$ and $t + 11$. However, the adaLASSO successfully corrected the big errors of the LASSO. The adaLASSO has smaller errors than the AR only on $t + 1$, $t + 3$ and $t + 7$, and its performance starts to deteriorate on

	LASSO $t + 1$	adaLASSO $t + 1$	LASSO-all	adaLASSO-all
P., AR, Dummy	0.46	0.35	0.29	0.31
Other Variables	0.54	0.65	0.71	0.69
G. debt	0.64	0.69	0.45	0.42
Unemployment	0.05	0.05	0.02	0.19
Exchange	0.17	0.16	0.16	0.02
Money	0.10	0.10	0.21	0.23
Others	0.04	0	0.16	0.14

The table shows how often a variable from each of the groups was selected. The first row shows how often autorregressive terms, other inflation measures, and dummy variables were selected. The second row shows the proportion of the variables which are not considered in the first row.

Rows 3-7 show how the variables from row 2 are distributed.

Table 3.2: Most selected variables: IPCA

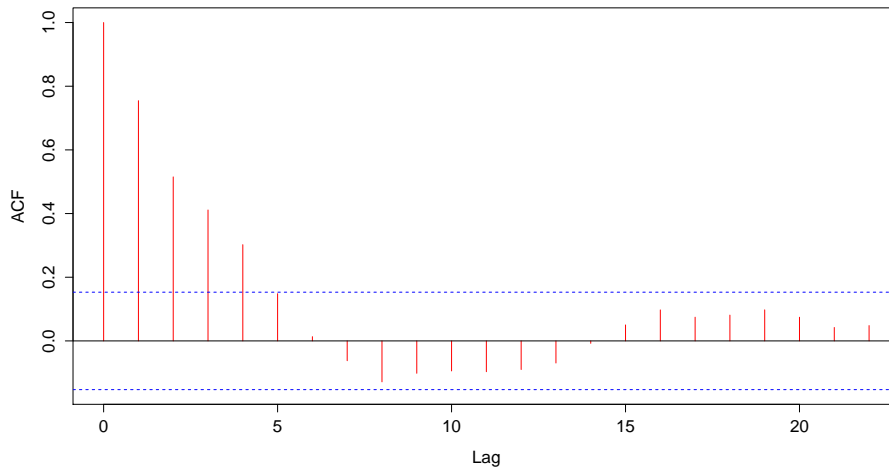


Figure 3.9: Autocorrelation function: IGP-M

$t + 8$. However, this time, the adaLASSO $t + 1$ forecast is statistically different from the AR and factor model forecasts (see Table 3.3). The average IGP-M inflation is 0.50% in the out-of-sample periods and the average MAE for all horizons is 0.37% for the AR and 0.40% for the adaLASSO. These are very large errors compared to the IPCA and to the average out-of-sample inflation.

The average number of selected variables by the LASSO and the adaLASSO on all forecast horizons is shown in Figure 3.11. The LASSO has around five selected variables up to the $t + 7$ horizon, when it increases to 20 variables. Note that the poor performance of the LASSO was also on the horizons $t + 6$ and $t + 11, t + 12$. These are exactly the same forecast horizons

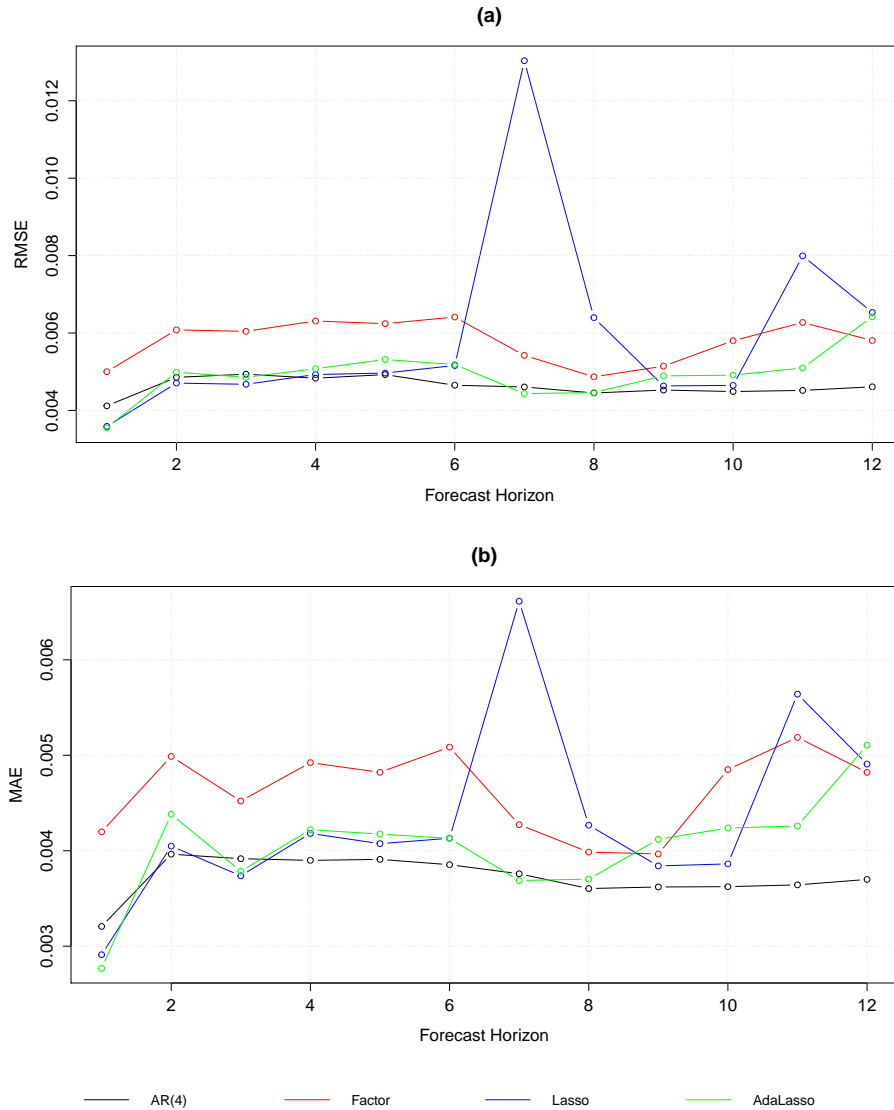


Figure 3.10: RMSE and MAE of all models: IGP-M

whose number of variables is high.

Table 3.3 and Figure 3.12 show the correlations between the forecasts, the GW p -values and the cumulative squared errors of all the models on the horizon $t + 1$, $t + 3$, $t + 6$ and $t + 12$. The results here are very different from the IPCA case. First, there are many statistically different forecasts, according to the GW test, on $t + 1$, $t + 6$ and $t + 12$. Second, there are some negative correlations between the forecasts. Third, the LASSO and adaLASSO large errors are caused by one single out-of-sample period on $t + 12$. Finally, the LASSO and the adaLASSO are statistically the best models to forecast the

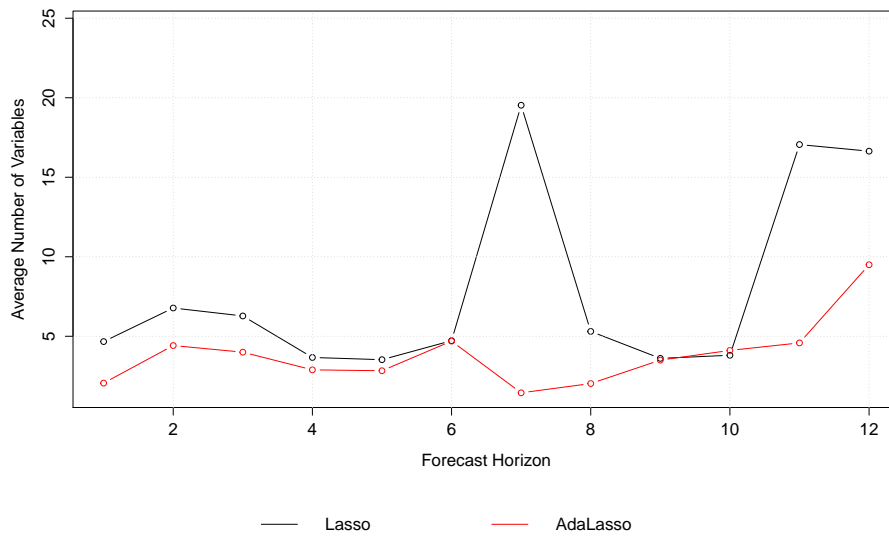


Figure 3.11: Number of selected variables: IGP-M

$t + 1$ inflation.

Figures 3.13 and 3.14 show the LASSO and adaLASSO selected variables. There were more unusual results with many selected variables for the IGP-M than for the IPCA. The results were unfavorable in all cases except for the $t + 6$ adaLASSO forecast, which had an RMSE slightly smaller than the AR. Both the LASSO and the adaLASSO selected more variables on the $t + 12$ horizon and their errors for this horizon were very large. On $t + 1$ and $t + 3$, the LASSO selected the IGP-DI instead of the AR component. Government debt and exchange rates were also selected, but mostly in a smaller proportion than in the IPCA case. Additionally, the models consistently selected M3 and M4 variables. The adaLASSO on $t + 1$ uses basically only the IGP-DI as a predictor for the inflation; the other two dots refer to government debt. On $t + 6$ and on some of the omitted horizons, we obtained the constant-only model. Finally, in the IGP-M case, the government also plays an important role in inflation, especially when M4 is included. However, its importance is smaller than it was in the IPCA case.

Table 3.4 shows the proportion of the selected variables. As we mentioned before, the most frequently selected variables are related to money. Government

$t + 1$	AR(4)	Factors	LASSO	adaLASSO	$t + 3$	AR(4)	Factors	LASSO	adaLASSO
AR(4)	1	0.58	0.92	0.93	1	0.28	0.40	0.69	
	-	(0.044)	(0.182)	(0.023)	-	(0.320)	(0.434)	(0.453)	
Factors	0.58	1	0.62	0.55	0.28	1	0.33	0.36	
	(0.044)	-	(0.001)	(0.003)	(0.320)	-	(0.156)	(0.180)	
LASSO	0.92	0.62	1	0.98	0.40	0.33	1	0.81	
	(0.182)	(0.001)	-	(0.329)	(0.434)	(0.156)	-	(0.176)	
adaLASSO	0.93	0.55	0.98	1	0.69	0.36	0.81	1	
	(0.023)	(0.003)	(0.329)	-	(0.453)	(0.180)	(0.176)	-	
$t + 6$	$t + 12$								
AR(4)	1	0.11	0.12	0.10	1	0.39	-0.04	-0.15	
	-	(0.019)	(0.498)	(0.489)	-	(0.12)	(0.022)	(0.006)	
Factor	0.11	1	-0.12	-0.12	0.39	1	-0.05	-0.09	
	(0.019)	-	(0.128)	(0.127)	(0.016)	-	(0.116)	(0.384)	
LASSO	0.12	-0.12	1	0.99	-0.04	-0.05	1	0.88	
	(0.498)	(0.128)	-	(0.030)	(0.022)	(0.116)	-	(0.465)	
adaLASSO	0.10	-0.12	0.99	1	-0.15	-0.09	0.88	1	
	(0.489)	(0.127)	(0.030)	-	(0.006)	(0.384)	(0.465)	-	

This table shows the correlation and the p-value of the Giacomini and White test for the forecasts in Figure 3.10.

The p-values are shown in parentheses. The null hypothesis of the GW test is that the forecasts are statistically equal.

Table 3.3: Forecast correlations and GW test p-values: IGP-M

debt came second and unemployment and exchange rates were rarely selected.

The $t + 1$ adaLASSO selected only government debt variables. However, Figure 3.14 and the first row of the table show that this is not very important. The continuous segment of dots on the $t + 1$ adaLASSO is the IGP-DI and debt variables were selected only twice (the two dots in the middle of the $t + 1$ plot), accounting for 6% of the selected variables.

	LASSO $t + 1$	adaLASSO $t + 1$	LASSO-all	adalasso total
P., AR, Dummy	0.27	0.94	0.07	0.14
Other Variables	0.73	0.06	0.93	0.86
G. debt	0.24	1	0.25	0.41
Unemployment	0	0	0.02	0.09
Exchange	0.15	0	0.02	0
Money	0.56	0	0.43	0.46
Others	0.05	0	0.28	0.04

The table shows how often a variable from each of the groups was selected. The first row shows how often autorregressive terms, other inflation measures, and dummy variables were selected. The second row shows the proportion of the variables which are not considered in the first row. Rows 3-7 show how the variables from row 2 are distributed.

Table 3.4: Most selected variables: IGP-M

3.4

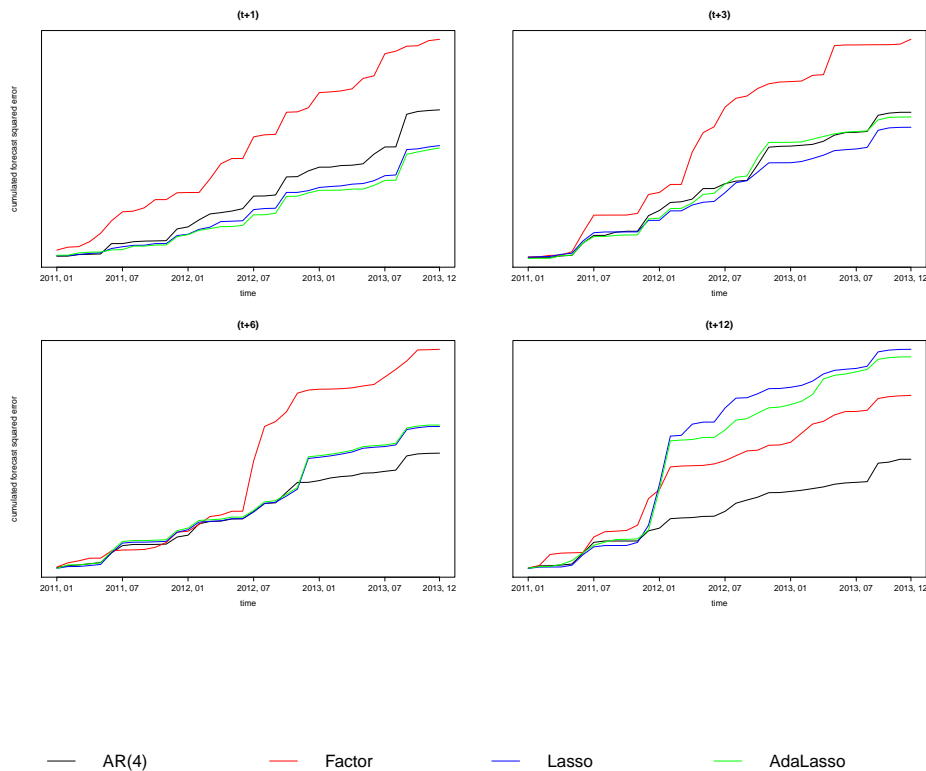


Figure 3.12: Cumulative squared errors: IGP-M

Final Remarks

We analyzed the two most important Brazilian inflation measures, the IPCA, which is the official measure, and the IGP-M, which is used in several contracts.

Our main objective was to forecast inflation. We used several econometric models such as AR models, factor models, LASSO, and adaLASSO. The latter two are models based on shrinkage estimations. This sort of estimation procedure selects only a few regressors in a high dimensional framework with hundreds of candidate variables.

We used all the models to produce forecasts for several horizons. The adaLASSO was the best model to forecast the IPCA inflation from one up to four months ahead. On the other hand, for longer horizons, the AR and the factor models have shown smaller errors. The Giacomini and White test for

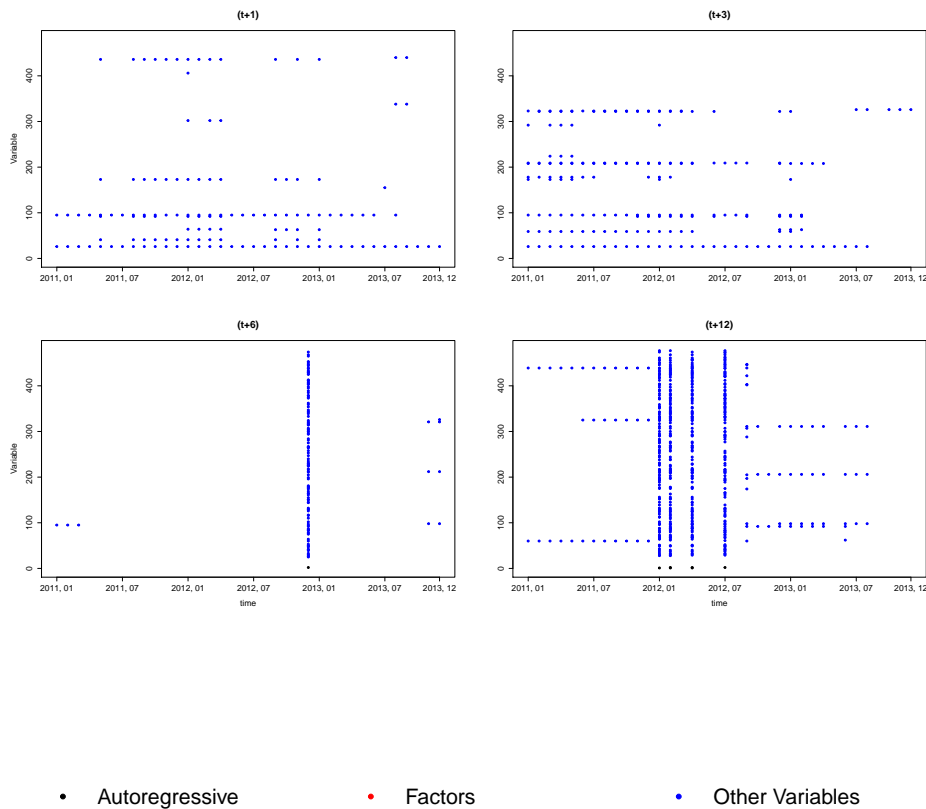


Figure 3.13: Variables selected by the LASSO: IGP-M

equal predictive accuracy showed that, in the IPCA case, the forecasts of all the models were not statistically different. However, this finding may be due to the restricted sample size (only 36 observations). Finally, most of the variables selected by the LASSO and adaLASSO were related to government debt and to the AR component of inflation.

Regarding the IGP-M inflation, the adaLASSO was statistically the best model for one-month-ahead forecasts. However, the AR dominated all the other models for longer horizons. Government debt variables were also selected, but in a smaller proportion than in the IPCA case. Additionally, the M3 and M4 variables were consistently selected.

We have three main results. First, economic variables other than price indexes are not good at predicting inflation on longer horizons. Second, apart

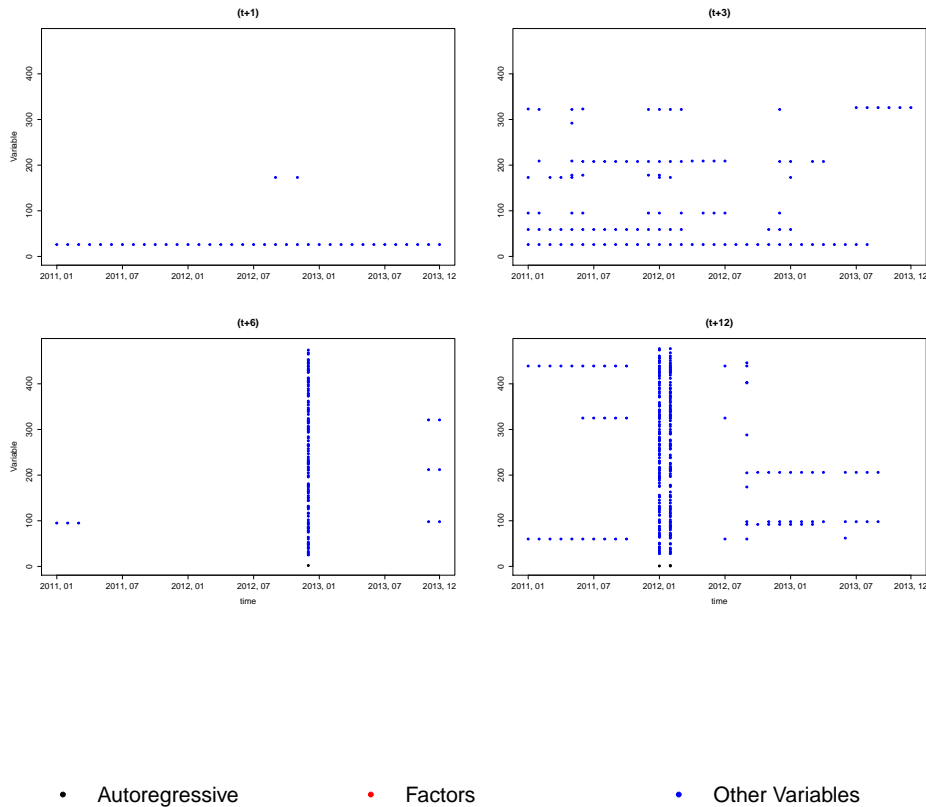


Figure 3.14: Variables selected by the adaLASSO: IGP-M

from variables related to inflation persistence, government debt variables represent the most important predictors to forecast short-term inflation. Third, our evidence is that the inflation mechanisms in Brazil are not those stated by the Phillips curve, especially when it comes to the unemployment-inflation relationship. Although Brazilian unemployment indexes are very low, the proportion of the population that does not work is big due to some public policies that target only those who are not currently working. These people are not considered unemployed because they are not looking for jobs. This type of policy may change the relationship that variables related to production have with inflation.

Finally, studies such as this one are especially important for emerging countries like Brazil. It helps academics and practitioners to understand the

inflation determinants and it also produces reliable forecasts. Many of these countries, including Brazil, have experienced periods of hyperinflation, and even though Brazilian inflation is much more under control now than in the late 1980s and early 1990s, it is still bigger than the inflation in developed countries and it is much easier for emerging countries to lose control over it.

3.5 Chapter Appendix

3.5.1 Estimation Procedures

In this Appendix, we show some technical details on the estimation procedure.

3.5.1.1 Differentiation

We used the Augmented Dickey-Fuller test (ADF) and the Kwiatkowski–Phillips–Schmidt–Shin test (KPSS) to decide whether to differentiate the variables. An important thing to mention is that we assumed that the IPCA and the IGP-M indexes are stationary. The ADF strongly rejected the null, with p-values smaller than 0.01. The KPSS p-value was greater than 0.1 for both indexes (recall that the null hypothesis for the KPSS is stationary). Several results for both tests in our dataset indicated the opposite or that the p-values were very close to the frontier of rejection. We adopted the criterion of differentiating the time-series when the test results were not reliable. All variables were differentiated, except the inflation indexes.

3.5.1.2 LASSO and adaLASSO estimation

The LASSO was estimated using the R package GLMNET package (43). Let λ_{ols} be the penalty parameter that implies the model with the biggest possible number of variables, and λ_c the penalty parameter that results in a

constant-only model such that $\lambda_{ols} < \lambda_c$. We estimate the LASSO for 100 different values of λ such that the first model uses λ_{ols} and the last model has only a constant. All the other 98 models have at least one selected variable. This is the standard procedure in the GLMNET. We select the best 100 models using the BIC. Figure 3.15 shows the BIC as a function of the penalty parameter. The dashed line is the best model and the upper axis shows the number of selected variables in each model (including the constant). Although the function is not convex, it clearly has a global minimum.

Recall that the adaLASSO estimator is defined as:

$$\hat{\beta} = \arg \min_{\hat{\beta}} \|Y - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^p (\hat{\beta}_j^*)^{-\tau} |\beta_j|, \quad (3-4)$$

(13) choose parameter τ computationally. After estimating the LASSO's first step, they estimate the adaLASSO for different τ 's and select the best model using the BIC in a very similar way to the one implemented to select the best λ . This increases considerably the computational cost since for each τ we have to estimate 100 models to select the λ . We used the same procedure to estimate the λ , as this is the standard in the R package GLMNET. As for the τ , we selected one fixed value for each index (0.3 for the IPCA and 0.5 for the IGP-M). The values were selected using the BIC in the first subsample of the rolling window implementation. This procedure saves time as one τ is selected for each window with a negligible variation in the results, given that the best τ does not change much across the rolling window samples.

PUC-Rio - Certificação Digital N° 1421640/CA

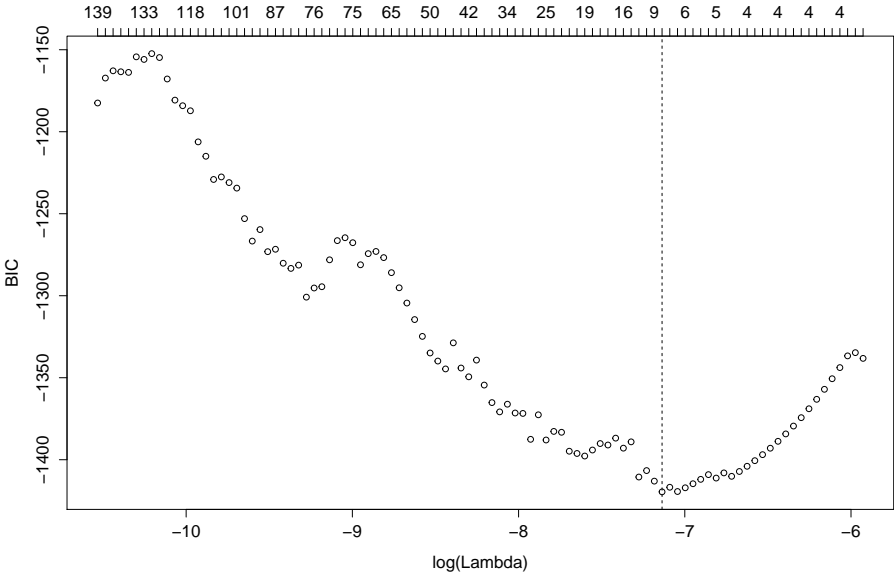


Figure 3.15: Model selection using the BIC

3.5.2

List of Variables

Prices

IPCA - general -index (dec. 1993 = 100)

IGP-M - general - index (aug. 1994 = 100)

IGP-DI - general - index (aug. 1994 = 100)

IGP-OG - general - index (aug. 1994 = 100)

IGP-10 - index (ago. 1994 = 100)

Employment and Wages

Unemployment Rate - RMSP

Unemployment Rate - Open - RMSP

Unemployment Rate - Hidden - RMSP

Unemployment Rate - Hidden - Precarious - RMSP

Employment Personnel - Industry

Working Hours (Hours Worked???) - Production - Industry

Number of Employees

Employment level - Industry

Average Income

Minimum Wage - Industry

Real Minimum Wage

Payroll - Industry

Minimum Wage - Purchasing Power Parities

International Transactions and Government Debt
Balance of Payments - Total Result - US\$
Financial and Capital Accounts - US\$
Current Transactions - US\$
Unrequited Transfers - US\$
External Debt - States and Cities
Fiscal Debt - Public Sector
Internal Debt - States and Cities
Internal Debt - Federal Government
Total Net Debt - States and Cities
Total Net Debt - Federal Government and Central Bank

Economic Activity and Production
Apparent Consumption - Alcohol
Apparent Consumption - Oil and Derivatives
Apparent Consumption - Capital Goods
Apparent Consumption - Intermediate Goods
Apparent Consumption - Consumer Goods - All
Apparent Consumption -Durable Consumer Goods
Apparent Consumption -Semi-durable and Non-durable Goods
Apparent Consumption -Industry - General
Apparent Consumption - Manufacturing Industry
Real Income - Industry
Real Sales - Industry
Electricity - Consumption
Default - Number of Queries
Default - New Registers - Net
Economic Condition Index
Bad Checks
Capacity Utilization
Industrial Production - Intermediate Goods
Industrial Production - Plastic and Rubber
Industrial Production - General
Industrial Production - Capital Goods
Industrial Production - Metallurgy
Industrial Production - Cellulose and Paper
Industrial Production - Oil
Industrial Production - Steel
Industrial Production - Motor Vehicles
Slaughter Cattle and Poultry
Slaughter Pigs
Construction Index

Taxes and Government Income

Cofins - total - Gross Income
Social Contribution over Net Profits
PIS / Pasep - total - Gross Income
Gross Collection of Federal Revenues
Taxes on Goods Circulation
Financial Execution
Import Taxes
Income Taxes - Legal Entity
Income Taxes - Individual
Total Gross Income Taxes
Taxes on Rural Properties
Social Security Cash Flow
Tax on Motor Vehicles
Other Taxes
Payment Rates
Financial Taxes
Industry Taxes

Exchange Rates and Finance

Exchange Rate - Dollar - Commercial
Exchange Rate - Dollar - Tourism
BOVESPA Stock Index
Dow Jones Stock Index
Average Return on Investment Funds
Return on Savings
Returns on Gold
Interest Rate - CDI / Over
Interest Rate - TJLP
Interest Rate - Over / Selic

Money

M0 - Monetary Base
M1
M2
M2 - Savings Deposits
M2 - Corporate Securities
M3 - Fixed Income Funds
M3 - New Concept
M4 - New Concept
M4 - Federal Bonds

Economic Confidence
Consumer Confidence Index
Index of Economic Expectations
Government Evaluation - Excellent and Good
Government Evaluation - Regular
Government Evaluation - Poor
Way of Governing - Approves
Way of Governing - Disapproves
Confidence in the President - Good
Confidence in the President - Poor

4 Forecasting Macroeconomic Variables in Data-Rich Environments

JEL: C22.

Keywords: Big Data, forecasting, LASSO, Shrinkage, Model selection.

Abstract: We show that high-dimensional models produce, on average, smaller forecasting errors for macroeconomic variables when we consider a large set of predictors. Our results showed that, empirically, a good selection of the adaptive LASSO hyperparameters also reduces forecast errors.

4.1

Introduction

Recent advances in computer science, statistics and econometrics allow us to work with large and complex datasets. In this paper we consider high-dimensional econometric models to forecast macroeconomic variables in situations when is hard to select predictors. Our results show that in most cases, more complex models provide smaller forecast errors than simple autoregressive and factor models.

4.2

Econometric Models for Data-Rich Environments

4.2.1

LASSO

The Least Absolute Shrinkage and Selection Operator (LASSO) was proposed by (2) as an alternative to ridge regression. Instead of imposing a quadratic penalty on the parameters, the LASSO penalizes their absolute value. This type of penalization allows irrelevant variables to be shrunked exactly to zero. The LASSO estimator is defined as

$$\hat{\beta} = \arg \min_{\beta} \|\mathbf{Y} - \mathbf{X}\beta\|_2^2 + \lambda \sum_{j=1}^p |\beta_j|, \quad (4-1)$$

where β is a N -dimensional vector of parameters, \mathbf{X} is a $T \times N$ matrix of candidate variables,¹ \mathbf{Y} is the dependent variable and λ is the shrinkage parameter. The second part of (5-4) controls the shrinkage and depends directly on the size of λ , which is selected using the BIC.

However, (3) showed that the LASSO does not have the oracle property and its performance tends to deteriorate as we increase the number of candidate variables. (3) showed that the oracle could be achieved by using a weighted penalization with weights estimated from a first-step model. This new estimator is the adaptive LASSO (adaLASSO), defined as:

¹The LASSO can be used even when $N > T$ (16).

$$\hat{\beta} = \arg \min_{\hat{\beta}} \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|_2^2 + \lambda \sum_{j=1}^p w_j |\beta_j|, \quad (4-2)$$

where $w_j = |\hat{\beta}_j^*|^{-\tau}$, $\hat{\beta}_j^*$ is a parameter estimated in a first-step, and $\tau > 0$ determines how much we want to emphasize the weights.

We estimated the adaLASSO using the LASSO as the first-step and $\tau = 1$, or using the elastic-net² as first step, selecting τ using the BIC. We call this last model flex-adaLASSO.

4.2.2

Bagging

The use of bagging to forecast time-series was proposed by (44). It consists on using several bootstrap samples to estimate the parameters, selecting the relevant variables on each sample using t-statistics. After we computed the forecasts using the estimated parameters from each bootstrap sample, we use the average forecast as the final object.

The bootstrap samples were drawn using block-bootstrap and the number of samples was 100. Since we are dealing with 130 variables, and we also use lags and factors as candidate variables, we did a pre-testing to select the relevant variables in two steps. First, we divided the sample into five arbitrary groups of variables and selected those with $|t| > 1.96$. Then we did another pre-testing with the selected variables from the five groups. The ones that remained with $|t| > 1.96$ were selected.

4.2.3

Target Factors

(6) showed that one could achieve better forecasting results using factor models considering which variable is going to be predicted. In other words, if we have a very large set of potentially relevant variables, we could improve our

²The elastic-net is a convex combination of the LASSO penalization and the ridge penalization. We determined how much weight to put on each penalization also using the BIC. This procedure has a very high computational cost.

forecasts simply by doing some type of pre-test to select which variables might be more relevant to forecast our targeted variable.

Let y_t be the variable we want to forecast, h is the forecasting horizon and \mathbf{W}_t contains a set of controls in the pre-test. In our case, \mathbf{W}_t contains only lags of y_t . Furthermore, X_{it} , $i = 1, \dots, N$ represents the candidate variables. The details of the pre-test are as follows:

1. For each $i = 1, \dots, N$, fit a regression³ of y_{t+h} on \mathbf{W}_t and X_{it} and save the t-statistics for all X_{it} .
2. Sort the absolute t-statistics of each X_{it} in descending order.
3. Select a significance level α . The variables considered as relevant will be those with $|t_i|$ bigger than the critical value.
4. Let $\mathbf{x}_t(\alpha)$ be the selected relevant variables. Compute factors \mathbf{F}_t from $\mathbf{x}_t(\alpha)$ using principal components.
5. Fit a regression of y_{t+h} on \mathbf{W}_t and $\hat{\mathbf{f}}_t \subset \hat{\mathbf{F}}_t$.

In the last step, we selected \mathbf{f}_t using the criteria in (4). Additionally, instead of selecting lags for \mathbf{f}_t using the BIC like (6), we allowed four lags of each candidate variable in the pre-testing procedure.

4.2.4 Complete Subset Regression

The Complete Subset Regression (CRS) was proposed by (45). It consists of selecting a number $k < N$, where N is the number of variables in the database, and fit regressions for all possible combinations of k variables. In our case we have 25 variables and $k = 4$, therefore we had to estimate 12,650 regressions. The final forecasts is the average forecast computed from all regressions.

³We followed (6) and used \mathbf{W}_t as four lags of y_t .

Note that, initially, we have 131 potential variables, and as in the target factors case, we allowed four lags of each variable. This number of variables makes the CSR computationally infeasible, therefore, we had to use a pre-testing procedure to select only a few variables (25). Following the same notation we used for the target factors, the pre-testing is described below:

1. For each $i = 1, \dots, N$, fit a regression of y_{t+h} on $X_{i,t}$. This time we did not use fixed controls \mathbf{W}_t , the autoregressive variables were treated as variables in \mathbf{X}_t .
2. Select the K variables with biggest $|t|$. We used $K = 25$.
3. Fit regressions for y_{t+h} using all possible subsets of k variables from the K pre-selected variables. In our case, $k = 4$.
4. Compute the final forecast as the average forecast from all the fitted regressions.

4.3 Estimation Procedures and Data

4.3.1 Estimation Procedures

We estimated models directly for the horizons we wanted to forecast, in this case, one and four steps ahead. Furthermore, we used a rolling window scheme of 144 windows to compare the forecasts. Considering the particularities of each model, the estimated equation was:

$$y_{t+h} = \alpha_0 + \sum_{i=1}^N \beta_i x_{i,t} + \varepsilon_{t+h} \quad (4-3)$$

where, y_t is the variable we want to forecast, h is the forecast horizon, α_0 is a constant, β_i are the regression parameters, x_{it} represents all variables, or in some cases, candidate variables and ε_t is an error term.

We used autoregressive terms in all models. They are candidate variables on the LASSO models, CSR and Bagging. In the target factors,

	Mean	SD	Min	Max
CPI	0.0033	0.0032	-0.0179	0.0181
emptot	0.0014	0.0026	-0.0099	0.0147
iptot	0.0022	0.0077	-0.0413	0.0309
ncom	63.0539	14.8183	18.0000	97.6000
nempl	49.1782	6.9385	25.3000	67.8000
nprod	55.4676	7.4651	26.3000	80.6000

Table 4.1: Descriptive Statistics

four autoregressive lags were included in the regression, they are represented by \mathbf{W}_t , which is used as fixed control on the pre-testing and on the factor regression.

4.3.2 Data

We used 131 macroeconomic variables related to the US economy. The period of analysis is from January 1960 to December 2011. This data-set was also used by (46). We selected six variables from the data to be forecasted using all the remaining 130 variables. The selected variables were the CPI inflation, variations on the total total number of employees on the private sector except by the farm (emptot), variations on the aggregated industrial production index (iptot) and NAPM indexes for commodities (ncom), employment (nempl) and production (nprod).

4.4 Results

In this section we show the results for $h = 1$ and $h = 4$. First, Table 4.1 shows descriptive statistics for the forecasted variables. The forecast results for $h = 1$ are displayed in Table 4.2 where the values in parenthesis are the p-value of the test of (42) using the AR(4) as benchmark. The results show that the flex-adaLASSO produces smaller forecast errors for all variables. The high-dimensional models provided better forecasts for all variables, except for the emptot, where these models had forecast errors equal to the AR(4) on average. The CSR was the high-dimensional model with the worst performance.

	CPI	EMPTOT	IPTOT	NCOM	NEMPL	NPROD
AR(4)	0.0033	0.0011	0.0068	6.13	2.44	3.21
	-	-	-	-	-	-
AR(4) + Factors	0.0033 (0.760)	0.0012 (0.271)	0.0061 (0.068)	6.04 (0.844)	2.29 (0.288)	3.06 (0.414)
T. Factors	0.0032 (0.914)	0.0012 (0.192)	0.0060 (0.049)	6.37 (0.099)	2.34 (0.509)	3.04 (0.736)
Lasso	0.0029 (0.021)	0.0012 (0.081)	0.0065 (0.036)	5.90 (0.370)	2.40 (0.992)	3.02 (0.336)
adaLasso	0.0031 (0.008)	0.0011 (0.663)	0.0065 (0.457)	11.49 (0.000)	3.86 (0.000)	3.90 (0.002)
PO adaLasso	0.0031 (0.223)	0.0011 (0.849)	0.0065 (0.288)	10.94 (0.000)	2.33 (0.750)	3.52 (0.151)
Flex-adaLasso	0.0028 (0.079)	0.0011 (0.562)	0.0059 (0.001)	5.72 (0.078)	2.26 (0.315)	2.75 (0.013)
PO flex-adaLasso	0.0029 (0.126)	0.0011 (0.929)	0.0059 (0.002)	5.78 (0.126)	2.27 (0.314)	2.73 (0.009)
Bagging	0.0030 (0.114)	0.0011 (0.414)	0.0064 (0.329)	6.16 (0.822)	2.29 (0.999)	2.90 (0.182)
CSR	0.0032 (0.061)	0.0013 (0.002)	0.0065 (0.019)	8.26 (0.000)	3.16 (0.000)	3.90 (0.001)

Values in parenthesis are the GW p-values using the AR(4) as benchmark and rows with PO are post-ols estimation.

Table 4.2: RMSE across the rolling windows and Giacomini-White test for $h = 1$

Table 4.3 shows the results for four steps ahead forecasts. The Flex-adaLASSO had the smallest forecast errors for four of the six variables. The target factor had the smallest errors for the variable *emptot*, however it was not statistically different from that of the AR(4). The AR(4) generated the smaller errors for the *iptot* variable, but it was also not statistically different from the forecasts of many other models.

Additionally, there was no significant gain in estimating the post-ols for the adaLASSO and the flex-adaLASSO. Finally, the parameter τ from the adaLASSO seems to make a big difference on the results. For example, on the *ncom* variable for $h = 1$, the adaLASSO had very big errors, however, when we selected the τ using the BIC as in the flex-adaLASSO, the adaLASSO went from the worst model to the best model.

	CPI	EMPTOT	IPTOT	NCOM	NEMPL	NPROD
AR(4)	0.0038	0.0017	0.0072	14.59	4.78	6.37
	-	-	-	-	-	-
AR(4) + Factors	0.0037 (0.764)	0.0016 (0.501)	0.0073 (0.617)	14.18 (0.898)	4.69 (0.369)	6.07 (0.594)
T. Factors	0.0036 (0.513)	0.0015 (0.949)	0.0074 (0.865)	14.17 (0.661)	4.58 (0.371)	6.05 (0.298)
Lasso	0.0036 (0.099)	0.0019 (0.000)	0.0078 (0.075)	13.18 (0.004)	4.38 (0.127)	6.28 (0.668)
adaLasso	0.0035 (0.020)	0.0016 (0.201)	0.0073 (0.866)	14.12 (0.830)	6.44 (0.000)	7.39 (0.004)
PO adaLasso	0.0035 (0.071)	0.0017 (0.169)	0.0077 (0.349)	14.21 (0.809)	6.38 (0.000)	7.28 (0.007)
Flex-adaLasso	0.0035 (0.051)	0.0017 (0.056)	0.0075 (0.047)	12.69 (0.015)	4.08 (0.031)	5.67 (0.023)
Po flex-adaLasso	0.0035 (0.256)	0.0017 (0.070)	0.0079 (0.005)	12.87 (0.041)	3.99 (0.015)	5.69 (0.055)
Bagging	0.0037 (0.917)	0.0017 (0.440)	0.0076 (0.107)	14.19 (0.571)	4.50 (0.312)	5.84 (0.156)
CSR	0.0036 (0.011)	0.0020 (0.000)	0.0079 (0.126)	13.15 (0.016)	5.08 (0.005)	6.80 (0.243)

Values in parenthesis are the GW p-values using the AR(4) as benchmark and rows with PO are post-ols estimation.

Table 4.3: RMSE across the rolling windows and Giacomini-White test for $h = 4$

4.5 Conclusion

We showed that high-dimensional models such as LASSO and Bagging provide smaller forecast errors for some macroeconomic variables. The results for target factors and complete subset regression were less conclusive. The adaLASSO had the smaller forecast errors of all models when we allowed the hyper

5

Real-Time Inflation Forecasting with High-Dimensional Models: The Case of Brazil

JEL: C22.

Keywords: real-time inflation forecasting, emerging markets, shrinkage, factor models, LASSO, regression trees, random forests, complete subset regression, machine learning, model confidence set, forecast combination, expert forecasts.

Abstract: We show that high-dimensional econometric models, such as shrinkage and complete subset regression, perform very well in real time forecasting of inflation in data-rich environments. We use Brazilian inflation as an application. It is an ideal example because it exhibits high short-term volatility and several agents devote extensive resources to forecast its short-term behavior. Therefore, precise specialist's forecasts are available both as a benchmark and as an important candidate regressor for the forecasting models. Furthermore, we combine forecasts based on model confidence sets and we show that model combination can achieve superior predictive performance.

5.1

Introduction

Forecasting inflation in real-time is a difficult task and it has been extensively studied in the literature. At least since (47) introduced the concept of real interest rates, forecasting inflation has been a crucial issue for both academics and practitioners. We estimate models to forecast inflation in real-time and in data-rich environments. By real-time we mean that forecasts are

computed by using solely the information available to the econometrician at the time the forecasts are made. A data-rich environment is the one where the number of potential predictors is large, possibly larger than the sample size. We consider the case of an emerging economy with inflation targeting, where precise inflation forecasts are of utmost importance for monetary policy and investment strategies (48).

Emerging markets usually exhibit higher and more volatile inflation, which tend to shorten the investment horizon. In Brazil, a country that conquered hyperinflation only in 1994, most fixed income assets are still very short. Therefore, the importance of forecasting short-term inflation is higher than in advanced economies, with more resources being devoted by financial institutions to such endeavor. Forecasting short-term inflation in Brazil is not only hard enough an exercise, with lots of data, but also one where extremely good expert forecasts exist with which different econometric techniques may be compared.

The literature on inflation forecasting is vast and there is substantial evidence that models based on the Philips curve do not provide good inflation forecasts. Although (33) showed that many production-related variables are useful predictors of the US inflation, (34) showed that in many cases the Philips curve fails to beat even simple naive models. These results inspired researchers to look for different models and variables in order to improve inflation forecasts. Among the variables used are financial variables (36), commodity prices (35) and expectation variables (37).

Real-time inflation forecasting has been recently considered by several authors. (48) evaluate forecasts made in real time to support monetary policy decisions at the Swedish Central Bank from 2007 to 2013. The authors compare Dynamic Stochastic General Equilibrium (DSGE) models with Bayesian Vector Autoregressive (BVAR) models. (49) propose a mixed-frequency model for daily forecasts of euro area inflation in real-time.

The authors showed that the mixed-frequency model has superior predictive performance with respect to forecasts based only on economic derivatives. (50) consider real-time inflation forecasts by AR models and with revised data. Finally, (37) evaluate the use of Bayesian Model Averaging (BMA) to forecast inflation in real-time. However, none of these authors have considered the use of large-dimensional machine learning models.

There is also a growing literature on inflation forecasting in Brazil. (30) used several linear and nonlinear models and the Phillips curve to forecast inflation. The authors showed that some nonlinear models and the simple autoregressive (AR) model produce smaller forecast errors than the Phillips curve. (31) used long-memory heteroskedastic models to show that Brazilian inflation has long-range dependence both on the mean and on the variance. However, they do not exclude the importance of the short-term AR component. The relevance of past inflation is also pointed out by (32). More recently, (11) considered different high-dimensional models to forecast Brazilian inflation. The authors showed that techniques based on the *Least Absolute Shrinkage and Selection Operator* (LASSO) have the smallest forecasting errors for short horizon forecasts. For longer horizons, the AR benchmark is the best model with respect to point forecasts, even though there is no significant differences between them. Factor models also produces some good long-horizon forecasts in a few cases. However, none of the above papers consider real-time forecasts.

In this paper, we make use of the most important advances in econometric modeling to estimate real-time forecasts of the Brazilian CPI inflation (IPCA). This is not only the most widely used inflation measure in Brazil, but is also the index used to set the inflation target for central bank policy.

As far as we know, this is the first paper to use high dimensional and machine learning models to forecast inflation in real-time for an emerging economy, using expert survey forecasts as potential candidate predictors. The models used here may be classified into shrinkage models, such as the LASSO

(2), the adaptive LASSO (3), or the Post-Ordinary Least Squares (51), and models that combine information, such as target factors (6) and Complete Subset Regression (45, 20). We also included AR models and Random Walk forecasts as benchmarks and the Random Forest model (9) as a nonlinear alternative. As a robustness check, we also compare the high-dimensional models with the Unobserved Component Stochastic Volatility (UC-SV) model advocated by (52) and a Bayesian vector autoregression with priors from (10). Furthermore, we use specialist's forecasts compiled by the Brazilian Central Bank (BCB) both to gauge the quality of our forecasts and to include them as potential variables in our models. The specialist's forecasts are obtained in the FOCUS report, which contains expectations for several variables regarding the Brazilian economy (53). The FOCUS is an online environment that collects projections from more than a hundred professional forecasters about key Brazilian macroeconomic variables. The report was created to support the inflation target regime and it is published by the Brazilian Central Bank weekly, every Monday. The information is collected from several agents in the market such as banks, fund managers, and consulting companies. We use the median, the mean and the standard deviation of these market expectations in our models. Additionally, the FOCUS report also publishes the Top5 expectations, which includes only the five agents which were more accurate on previous periods. The expectations are collected daily, but many forecasters update their forecasts only on Friday, since the survey is published every Monday. Besides inflation, the report also publishes expectations on GDP, industrial production, exchange rates and other variables. All this information is used by the Brazilian Central Bank to gauge its monetary policy. Finally, following (54), we use a forecast combination strategy based on the model confidence sets proposed by (55). The idea is to compute the average of the forecasts from the models included in a given confidence set. We show that this delivers superior forecasts than all individual models as well as the simple

average of all models.

We estimated forecasts for the following forecast-horizons: five days before the CPI index is published to 11 months plus five days (12 forecasts on total). For five-day-ahead, the LASSO and the FOCUS (expert forecast) are virtually the same. For the second horizon, the adaptive LASSO is superior than any other model. For the remaining horizons, the Complete Subset Regression dominates all other alternatives. The results are the same if we either use the root mean squared error or the mean absolute error. In terms of accumulated inflation, the Complete Subset Regression is the model which delivers the most precise forecasts. However, most of the forecasts from different models are not statistically different according the model confidence set. In light of this finding, we construct the final forecast as the average of the models included in the confidence set. This approach delivers the best forecasts among all the competing alternatives. Finally, we also compute density forecasts for each model based on bootstrap re-sampling. According to the log-score statistic, the CSR has superior performance for most of the forecasting horizons except the first two where LASSO based methods are ranked as the best models.

Following this Introduction, this paper has four sections. In Section 2, we describe the models that were used, as well as the empirical procedures. In Section 3, we explain the dataset. The main results are presented and discussed in Section 4. Finally, the main conclusions are summarized in Section 5. A description of the dataset is included in the appendix.

5.2 Empirical Methods

In this section we describe the methods used in this paper to forecast future inflation. We consider a direct forecast approach where the inflation h -periods-ahead, π_{t+h} , is modeled as a function of a set of predictors measured at time t such as:

$$\pi_{t+h} = T(\mathbf{x}_t) + u_{t+h}, \quad (5-1)$$

where $T(\mathbf{x}_t)$ is a possibly nonlinear mapping of a set of q predictors and u_{t+h} is the forecasting error, $\mathbf{x}_t = (x_{1t}, \dots, x_{qt})' \in \mathbb{X} \subseteq \mathbb{R}^q$ may include weakly exogenous predictors, lagged values of inflation and a number of factors computed from a large number of potential covariates. Importantly, as we focus on real-time forecasts, \mathbf{x}_t contains only variables that are observed and available to the econometrician at time t . Many variables are published months after their period of reference. These variables are not included in the dataset at time t . Note further that by considering direct forecasts models for each horizon, this avoids the necessity to estimate a model for the evolution of \mathbf{x}_t .

For most of the methods considered in this paper, the mapping $T(\cdot)$ is linear, such that:

$$\pi_{t+h} = \boldsymbol{\beta}' \mathbf{x}_t + u_{t+h}, \quad (5-2)$$

where $\boldsymbol{\beta} \in \mathbb{R}^q$ is a vector of unknown parameters.

5.2.1 Factor Models with Targeted Predictors

Factor models using principal components are a very popular approach to avoid the curse of dimensionality when the number of predictors is potentially large. The idea is to extract common components from all variables, thus reducing the model dimension.

Consider equation 5-2. When the number of candidate predictors q is large, potentially larger than the sample size T , ordinary least squares (OLS) is either infeasible or have a very large variance. One solution to circumvent this drawback is to use factors as predictors instead of \mathbf{x}_t . The factors can be observed as in Fama and French (1993,1996) or unobserved as in (56) and (57). Our focus are on unobserved factors. Consider the following forecasting model:

$$\pi_{t+h} = \sum_{i=1}^p \gamma_i' \mathbf{f}_{t-i} + u_{t+h}, \quad (5-3)$$

where, \mathbf{f}_t is a vector of k of common factors extracted from \mathbf{x}_t and k is

much smaller than q . Note that \mathbf{f}_t is not observed and must be estimated by principal components. The assumptions and the theory behind factor models and when can we treat factors as observed variables can be found in Bai and Ng (2002,2006,2008).

In order to improve the forecasting performance of factor models, (6) proposed targeting the predictors. The idea is that if many variables in \mathbf{x}_t are irrelevant predictors of π_{t+h} , factor analysis using all variables may result in noisy factors with poor forecasting ability. The target factors are regular factor models with a pre-testing procedure to select only relevant variables to be included in the factor analysis. We show the steps of this procedure pointing out where our methodology differs from that proposed by (6). Let $x_{i,t}$, $i = 1, \dots, q$, be the candidate variables and \mathbf{w}_t a set of fixed regressors that will be used as controls in the pre-testing. We follow (6) and use \mathbf{w}_t as AR terms of π_t . The procedure is described as follows.

1. For $i = 1, \dots, q$, regress π_{t+h} on \mathbf{w}_t and $x_{i,t}$ and compute the t -statistics for the coefficient corresponding to $x_{i,t}$. We include four lags of each candidate variable in the pre-testing. (6) uses only the variables in t and select the lags latter.
2. Sort all t -statistics calculated in Step 1 in descending order.
3. Choose a significance level α , and select all variables which are significant using the computed t -statistics.
4. Let $\mathbf{x}_t(\alpha)$ be the selected variables from Steps 1–3. Estimate the factors \mathbf{F}_t from $\mathbf{x}_t(\alpha)$ by principal components.
5. Regress π_{t+h} on \mathbf{w}_t and $\mathbf{f}_t \subset \mathbf{F}_t$. The number of factor in \mathbf{f}_t is selected using the BIC. (6) selected also the number of lagged factors using the BIC. However, since we use lagged variables as regressors in the pre-testing, we did not use lagged factors.

The same procedure was used by (58). The authors showed that, in most cases, target factors slightly reduce the forecasting errors compared to factor models without targeting.

5.2.2 LASSO and adaptive-LASSO

A successful alternative to factor models to estimate parameters in large dimensions is to use *shrinkage* methods. The idea is to shrink to zero the parameters corresponding to irrelevant variables. Under some conditions, it is possible to handle more variables than observations. Among shrinkage methods, the Least Absolute Shrinkage and Selection Operator (LASSO), introduced by (2), and the adaptive LASSO (adaLASSO) of (3), have received particular attention. It has been shown that the LASSO can handle more variables than observations and the correct subset of relevant variables can be selected (59, 16, 38). As noted in (16) and (3), for attaining model selection consistency, the LASSO requires a rather strong condition denoted “Irrepresentable Condition” and does not have the oracle property. (3) proposes the adaLASSO to amend these deficiencies. The adaLASSO is a two-step methodology which uses a first-step estimator, usually the LASSO, to weight the relative importance of the regressors.

The LASSO estimator is defined as

$$\hat{\beta} = \arg \min_{\hat{\beta}} \left[\sum_{t=1}^T (\pi_{t+h} - \beta' \mathbf{x}_t)^2 + \lambda \sum_{j=1}^q |\beta_j| \right], \quad (5-4)$$

where λ controls the amount of shrinkage and is determined by data-driven techniques such as cross-validation or the use of information criteria.

The adaLASSO is defined as:

$$\hat{\beta} = \arg \min_{\hat{\beta}} \left[\sum_{t=1}^T (\pi_{t+h} - \beta' \mathbf{x}_t)^2 + \lambda \sum_{j=1}^q w_j |\beta_j| \right], \quad (5-5)$$

where $w_j = |\hat{\beta}_j^*|^{-\tau}$ represents different weights on the penalization of each variable, $\hat{\beta}_j^*$ is the parameter estimated on a first step, and $\tau > 0$ determines

how much we want to emphasize the difference of the weights. (17) showed that the conditions that must be satisfied on the adaLASSO are very general. The model works even when the number of variables increases faster than the number of observations and when errors are non-Gaussian and heteroskedastic.

The most usual is to make $\tau = 1$. However, (58) showed that selecting τ using the BIC reduces the forecasting errors. They refer to this model as Flex-adaLASSO. The τ is not bounded on both sides as the λ . If $\tau \rightarrow 0$ then we have the traditional LASSO without weights, however, we do not have an upper bound. Note that if $\tau \rightarrow \infty$, then the $w_i \rightarrow 0$ and we have no penalty. Thus, in order to select the τ using an information criterion, one must establish an upper bound or the problem becomes computationally infeasible. If we use the LASSO as the first model, than some weights will be infinite. To deal with this issue computationally, we sum $T^{-\frac{1}{2}}$ to all coefficients from the first model.

(51) showed that the estimating a linear regression with the variables selected by the LASSO (post-OLS) is at least as good as the LASSO itself in terms of rate of convergence to the oracle and it also has a smaller bias. We estimated the post-OLS regression for the Flex-adaLASSO in order to check if it reduces forecasting errors.

5.2.3 Random Forest

The Random Forest (RF) methodology was initially proposed by (9) as a solution to reduce the variance of regression trees and is based on bootstrap aggregation (Bagging) of randomly constructed regression trees.

A regression tree is a nonparametric model based on the recursive binary partitioning of the covariate space \mathbb{X} where the function $T(\cdot)$ is a sum of local models (usually just a constant), each of which is determined in $K \in \mathbb{N}$ different regions (partitions) of \mathbb{X} . The model is usually displayed in a graph which has the format of a binary decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (also called leaves), and which grows from

the root node to the terminal nodes. Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the *split variable*. Hence, conditionally to the knowledge of the subregions, the relationship between π_{t+h} and \mathbf{x}_t in (5-1) is approximated by a piecewise constant model, where each leaf (or terminal node) represents a distinct regime.

To mathematically represent a complex regression-tree model, we introduce the following notation. The root node is at position 0 and a parent node at position j generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively. Every parent node has an associated split variable $x_{s_j t} \in \mathbf{x}_t$, where $s_j \in \mathbb{S} = \{1, 2, \dots, q\}$. Furthermore, let \mathbb{J} and \mathbb{T} be the sets of indexes of the parent and terminal nodes, respectively. Then, a tree architecture can be fully determined by \mathbb{J} and \mathbb{T} .

The forecasting model based on regression tree can be mathematically represented as

$$\pi_{t+h} = H_{\mathbb{J}\mathbb{T}}(\mathbf{x}_t; \boldsymbol{\psi}) + u_{t+h} = \sum_{i \in \mathbb{T}} \beta_i B_{\mathbb{J}i}(\mathbf{x}_t; \boldsymbol{\theta}_i) + u_{t+h} \quad (5-6)$$

where

$$B_{\mathbb{J}i}(\mathbf{x}_t; \boldsymbol{\theta}_i) = \prod_{j \in \mathbb{J}} I(x_{s_j, t}; c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} [1 - I(x_{s_j, t}; c_j)]^{(1-n_{i,j})(1+n_{i,j})}, \quad (5-7)$$

$$I(x_{s_j, t}; c_j) = \begin{cases} 1 & \text{if } x_{s_j, t} \leq c_j \\ 0 & \text{otherwise,} \end{cases} \quad (5-8)$$

$$n_{i,j} = \begin{cases} -1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\ 0 & \text{if the path to leaf } i \text{ includes the right-child node of the parent node } j; \\ 1 & \text{if the path to leaf } i \text{ includes the left-child node of the parent node } j. \end{cases} \quad (5-9)$$

Let \mathbb{J}_i be the subset of \mathbb{J} containing the indexes of the parent nodes that form

the path to leaf i . Then, θ_i is the vector containing all the parameters c_k such that $k \in \mathbb{J}_i$, $i \in \mathbb{T}$. Note that $\sum_{j \in \mathbb{J}} B_{\mathbb{J}_j}(\mathbf{x}_t; \theta_j) = 1$, $\forall \mathbf{x}_t \in \mathbb{R}^{q+1}$.

A Random Forest is a collection of regression-trees each of which is specified in a bootstrapped sub-sample of the original data. Suppose there are B bootstrapped sub-samples and denote $H_{\mathbb{J}_b \mathbb{T}_b}(\cdot; \psi_b)$ as the estimated regression-tree for each one of the sub-samples. The final prediction is defined as:

$$\hat{\pi}_{t+h} = \frac{1}{B} \sum_{b=1}^B H_{\mathbb{J}_b \mathbb{T}_b}(\mathbf{x}_t; \psi_b). \quad (5-10)$$

For each of the bootstrapped sub-samples a regression-tree is estimated by recursively repeating the following steps for each terminal node of the tree until the minimum number of observations at each node is achieved.

1. Randomly select m out of q covariates as possible split variables.
2. Pick the best variable/split point among the m candidates.
3. Split the node into two children nodes.

Random Forests can deal with a very big number of explanatory variables and the predicted model is highly nonlinear. It is important to notice that since we are dealing with time-series, bootstrap samples are calculated using block bootstrap.

5.2.4 Complete Subset Regression with Targeted Predictors

The Complete Subset Regression (CSR) was developed by (45, 20). The motivation is that selecting the optimal subset of \mathbf{x}_t to predict π_{t+h} by testing all possible combinations of regressors is computationally very demanding and in most of the times even unfeasible. Suppose that we have q candidate variables, the CSR selects a number $n \leq q$ and computes all combinations of regressions using only n variables. The forecast of the model will be the average forecast of all regressions in the subset.

The CSR deals well with a small number of candidate variables. However, for large sets the number of regressions to be estimated increases very fast. For example, with $q = 25$ and $n = 4$ we need to estimate 12,650 regressions. As the number of candidate variables is much larger, we adopt a pre-testing procedure similar to the one used with the target factors. We start fitting a linear regression of π_{t+h} on each of the candidate variables (including lags) and saving the t-statistics of each variable¹. The t-statistics are ranked in absolute value and we select the \tilde{q} variables which are more relevant on the ranking. The CSR forecast is calculated on these variables. We used $\tilde{q} = 25$ and $n = 4$.

5.3 The Data

Inflation is measured by the Brazilian Consumer Price Index (IPCA), which is the official inflation index in Brazil. Furthermore, a sizeable amount of inflation-linked bonds use the IPCA as reference. The dataset is obtained from Bloomberg and from the Central Bank of Brazil, covering the period from January 2003 to December 2015, a total of 156 observations. We have 59 macroeconomic variables and 34 variables linked to specialist forecasts. The number of macroeconomic variables is smaller than that of (11) because we are using only variables available on the period the forecast is computed. The dataset also includes expert forecasts from the FOCUS survey produced by the Central Bank of Brazil. Among expectation variables we consider the median of the h -period-ahead specialist forecasts, the median of the top five (Top5) experts, i.e., the five experts who produced the best forecasts in the previous period, and, finally, the mean and the standard deviation of the Top5. The macroeconomic variables cover several inflation and industry indexes, unemployment and other variables related to labour, energy consumption, exchange rates, stock markets, government accounts, expenditure and debt, taxes, monetary variables and exchange of goods and services. The inflation

¹We did not use a fixed set of controls, w_t , in the pre-testing like we did on the target factors.

series as well as the Top5 median are presented in Figure 5.1. As can be seen from Figure 5.1, the Top5 delivers the smallest RMSE for $h = 1$ (five-day-ahead) but rapidly loses performance as h grows.

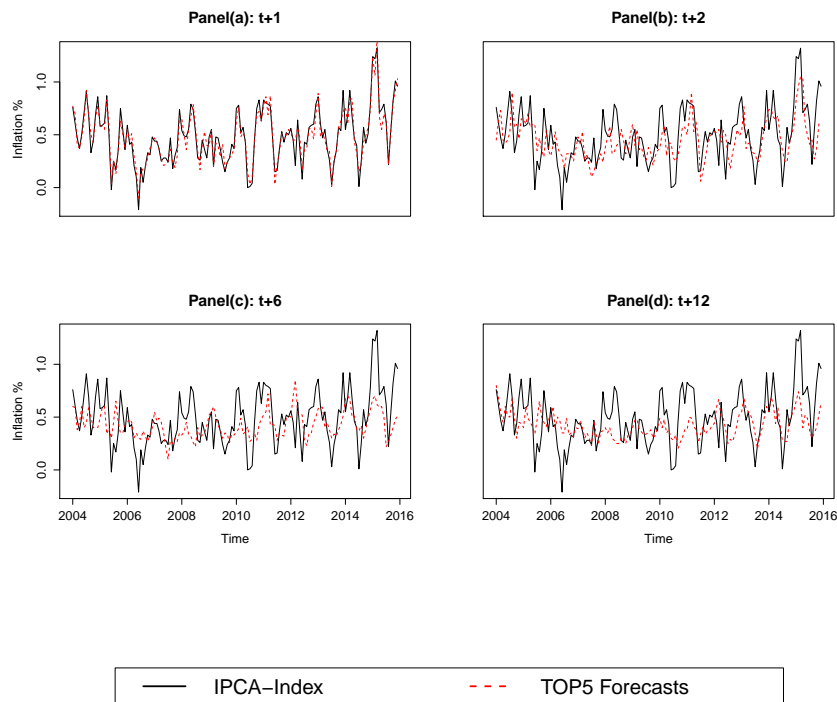


Figure 5.1: Brazilian Consumer Prices Index and Focus Top5 Forecasts

5.4 Main Results

5.4.1 Forecasting Errors

We estimate all models described in Section 5.2 for $h = 1, \dots, 12$. Recall that $h = 1$ is five days before the IPCA inflation is published, $h = 2$ is one month plus five days and $h = 12$ is 11 months plus five days. In this section we show the results and compare the forecasting errors of all models. We also estimate autoregressive models with lags selected by the BIC and include as well Random Walk forecasts in the comparison. All models are estimated in a nine-year rolling-window scheme and the first forecast is for January 1, 2012.

Therefore, the models are evaluated based on 48 point forecasts with the last forecast being for December 2015². This period covers different situations of the Brazilian economy. In 2012 and 2013 the Brazilian GDP increased 1.9% and 3% respectively, 2014 had an increase of 0.1% and 2015 had a decrease of 3.7%. Figure 5.1 shows that the state of the economy does not affect the precision of the short term forecasts. However, for longer forecasting horizons the errors are bigger in 2015, which was an year of 10.67% inflation. Note that the inflation target is 4.5% and its ceiling is 6.5%.

Table 5.1 shows the root mean squared error (RMSE) and the Mean Absolute Error (MAE) for all forecasting models. The model with the smallest forecasting error in each horizon is displayed in bold font. The last column of Table 5.1 shows the cumulative error for the 11 months inflation. The LASSO and the Flex-adaLASSO have the smallest errors for $h = 1$ and $h = 2$ and the CSR has the smallest errors for all other horizons. However, for $h = 1$, the forecasts from the LASSO are not statistically different from the expert forecasts. On the other hand, for larger horizons there is a substantial gain from using the CSR models. The target factor models become more competitive as h increases. The Random Walk and the autoregressive models have both a poor performance. The Random Forest was not the best model in any horizon, however, its performance was not bad overall. Its cumulative forecasting error was smaller than that of the FOCUS, the Top5 and the LASSO. Additionally, the Post-OLS estimation with the variables selected by the Flex-adaLASSO delivers larger errors than the Flex-adaLASSO itself.

The reason the LASSO and the Flexible adaLASSO are the best models for small horizons is due to the fact that expert forecasts are very precise for $h = 1$ and $h = 2$. As previously mentioned, market players devote considerable amount of resources to inflation forecasting. Therefore, variable

²We start producing forecasts in 2012 in order to have a reasonably number of point forecasts for each forecasting horizon and still have enough observations for the in-sample estimation of the models. As the models are for direct forecasts, we have 108 observations to estimate models for $h = 1$, 107 for $h = 2$, and so on. We also show results for 24 rolling windows.

The table shows the root mean squared error and the mean absolute deviation, in parenthesis, of the forecasts. The values in bold represent the best model in each measure of error and each forecasting horizon. All values are multiplied by 1000. The column Acc. shows the errors of the forecast accumulated over 11 months.

Brazilian Consumer Price Index													
RMSE × 1000 (MAE × 1000)	Forecast Horizon											Acc.	
	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10	t+11		t+12
RW	2.41 (1.99)	3.23 (2.63)	3.68 (3.01)	4.10 (3.38)	4.40 (3.44)	4.62 (3.64)	4.76 (3.71)	4.33 (3.41)	3.76 (3.04)	3.40 (2.73)	3.03 (2.58)	2.75 (2.13)	33.94 (26.11)
AR	2.30 (1.93)	2.89 (2.21)	3.26 (2.47)	3.31 (2.60)	3.23 (2.54)	3.18 (2.49)	3.04 (2.37)	2.82 (2.16)	2.72 (2.13)	2.70 (2.07)	2.67 (2.06)	2.64 (2.01)	20.75 (16.14)
Factors	1.33 (0.98)	2.19 (1.75)	2.42 (1.88)	2.48 (1.93)	2.44 (1.83)	2.49 (1.91)	2.48 (1.96)	2.37 (1.89)	2.29 (1.79)	2.50 (2.01)	2.38 (1.86)	2.44 (1.93)	14.31 (9.63)
LASSO	0.95 (0.74)	1.85 (1.46)	2.85 (2.28)	3.21 (2.44)	2.75 (2.11)	2.83 (2.24)	2.79 (2.12)	3.33 (2.65)	2.80 (2.15)	3.33 (2.69)	3.51 (2.89)	3.33 (2.71)	17.09 (12.42)
F. aL	0.98 (0.75)	1.58 (1.30)	2.20 (1.75)	2.43 (1.94)	2.39 (1.82)	2.42 (1.89)	2.53 (2.04)	2.86 (2.33)	2.48 (1.94)	2.56 (2.06)	2.54 (2.01)	2.46 (1.88)	13.50 (9.39)
P. OLS	0.98 (0.75)	1.62 (1.34)	2.23 (1.80)	2.23 (1.80)	2.49 (1.89)	2.52 (1.97)	2.53 (2.02)	3.08 (2.48)	2.52 (1.94)	2.66 (2.11)	2.61 (2.06)	2.46 (1.89)	14.02 (9.58)
RF	1.43 (0.97)	1.95 (1.45)	2.56 (1.93)	2.54 (1.93)	2.66 (2.06)	2.88 (2.30)	2.82 (2.21)	2.85 (2.25)	2.71 (2.09)	2.65 (1.96)	2.64 (1.99)	2.46 (1.82)	15.67 (12.36)
CSR	1.05 (0.88)	1.64 (1.33)	2.04 (1.69)	2.23 (1.75)	2.25 (1.79)	2.29 (1.80)	2.29 (1.80)	2.26 (1.80)	2.26 (1.81)	2.26 (1.79)	2.27 (1.77)	2.25 (1.78)	11.93 (8.41)
FOCUS	0.95 (0.76)	1.83 (1.50)	2.39 (1.87)	2.48 (1.91)	2.53 (1.93)	2.57 (1.97)	2.56 (1.94)	2.53 (1.91)	2.55 (1.93)	2.57 (1.93)	2.58 (1.94)	2.60 (1.96)	16.82 (12.51)
Top 5	0.96 (0.74)	1.69 (1.39)	2.32 (1.83)	2.48 (1.90)	2.62 (1.99)	2.70 (2.07)	2.77 (2.06)	2.67 (2.03)	2.51 (1.99)	2.65 (1.97)	2.56 (1.91)	2.55 (1.89)	16.69 (12.12)

Table 5.1: Forecasts Mean Absolute Errors and Root Mean Squared Errors

selection models such as the LASSO perform better than those methods that combine information from many variables such as target factors and CSR. However, as we increase the forecast horizons the expert forecasts lose their prediction power of prediction and many variables become more relevant. Models that combine information can successfully extract common information on all variables that are useful to forecast inflation. Figure 5.2 shows the average number of variables selected by the LASSO and the Flexible adaLASSO in all horizons. For $h = 1$ and 2, the number of selected variables is very small for both models, and it gets bigger for other horizons, especially in the case of the LASSO. For shorter horizons, the Flexible adaLASSO is mostly a combination of specialist forecasts.

Frequently, the model with the smallest average squared error may not be the model with smallest errors in most of the 48 rolling windows. Table 5.2 shows the ranking of models for each forecasting window. The table reports the proportion of cases where each model is in each position of the ranking. The results are aggregated for all horizons. Surprisingly, the Random Walk, which performed badly in terms of average errors, is the best model in 24% of the cases, the same proportion as the Top5. However, the same two models deliver the worst forecasts in 19% and 17% of the cases, respectively. The CSR model,

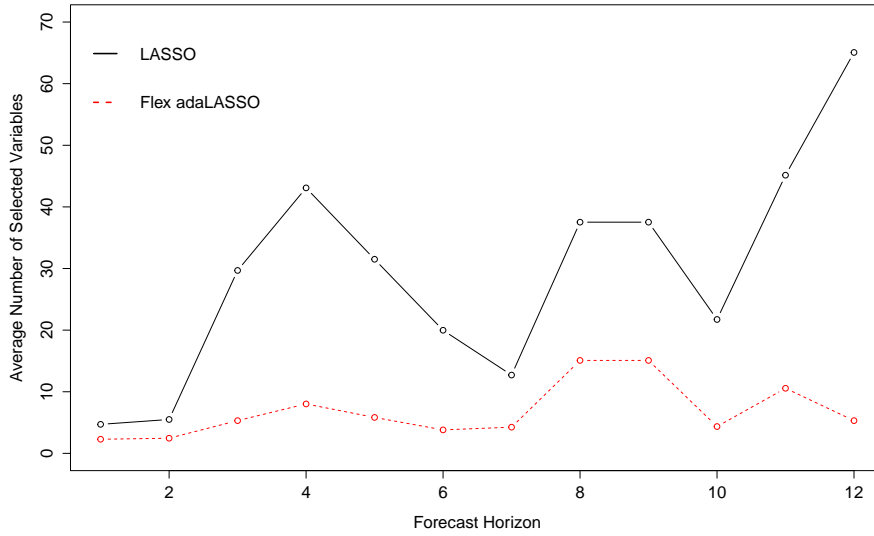


Figure 5.2: Average Number of Selected Variables by the Shrinkage Methods

The table shows the proportion which each model is in each ranking position. The results are aggregated for all forecasting horizons.

Brazilian Consumer Price Index										
	<u>Model Position</u>									
	1	2	3	4	5	6	7	8	9	10
RW	0.24	0.01	0.08	0.07	0.02	0.05	0.03	0.02	0.28	0.19
AR	0.04	0.06	0.06	0.06	0.03	0.07	0.06	0.06	0.24	0.31
Factors	0.14	0.05	0.18	0.07	0.09	0.15	0.09	0.10	0.06	0.08
LASSO	0.06	0.08	0.12	0.08	0.15	0.15	0.14	0.18	0.02	0.02
F. aL	0.03	0.07	0.10	0.14	0.18	0.15	0.09	0.20	0.02	0.02
P. OLS	0.05	0.10	0.11	0.12	0.18	0.12	0.11	0.17	0.03	0.02
RF	0.05	0.14	0.11	0.16	0.17	0.09	0.11	0.13	0.03	0.03
CSR	0.07	0.14	0.10	0.15	0.10	0.10	0.15	0.09	0.06	0.04
FOCUS	0.08	0.16	0.08	0.07	0.05	0.06	0.16	0.04	0.19	0.12
Top5	0.24	0.20	0.06	0.08	0.02	0.05	0.09	0.02	0.07	0.17

Table 5.2: Proportion each Model was in each Position of the Error Ranking

which is the best model on average in most horizons has the smallest errors only on 7% of the forecasts, and the Flexible adaLASSO model, which is the second best model considering the cumulative inflation, is the best model only in 3% of the cases. The models with smallest errors on average are the ones that perform well when most models perform bad. However, when all models are doing well they are not the best models anymore.

We show the correlation of the forecasting errors in Figure 5.3. The figure displays the heat-maps for horizons 1, 2, 6, and 12. The pattern is very similar for all horizons. The FOCUS and the Top5 are positively correlated. However,

their correlation with all other models is negative. The remaining forecasts are all positively correlated. The two best models, which are the Flexible adaLASSO and the CSR models, have a strong negative correlation with both expert forecasts considered in this paper. This result shows that even though some models and the expert forecasts have small forecasting errors, these forecasts are considerably different, and that opens the possibility of improving the results using combinations of these forecasts, which will be discussed on the next section.

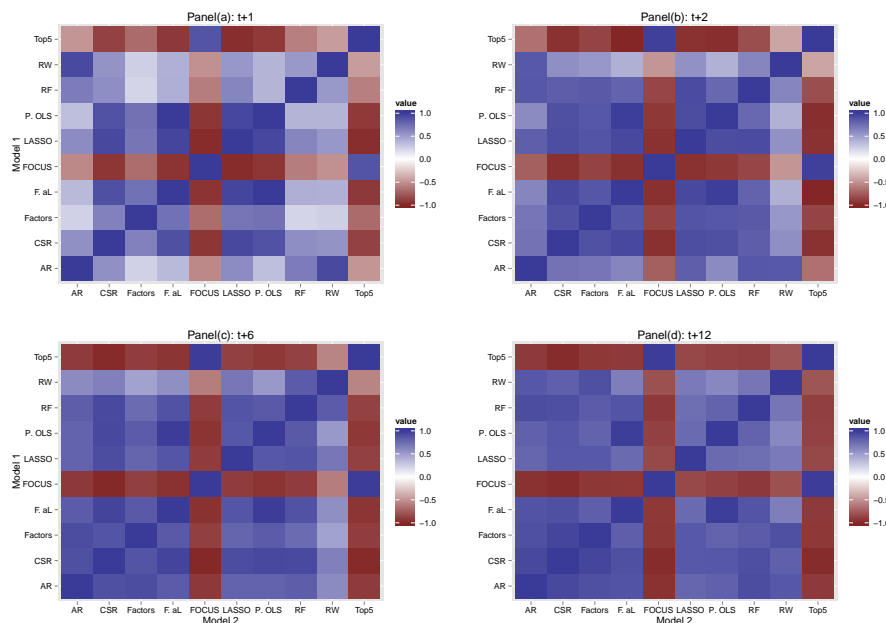


Figure 5.3: Forecasting Errors Correlation

5.4.2 Model Confidence Sets and Model Combination

In this section we report the Model Confidence Set (MCS) developed by (55). The MCS allows us to compare a large number of models at the same time. The test returns a confidence set that includes the best model with probability $(1 - \alpha)$. As we decrease α the set becomes wider (with more models) and for large values of α we may even have a set with only one single model.

The MCS uses bootstrapped samples of a given loss function, in our case squared error, to create the test statistics. The confidence set estimates p -

values for all models using the bootstrapped samples and uses α to select which models are inside the set. Since models are removed from the set interactively, the MCS also generates a ranking. The best model has a p -value equals 1 by definition, since it can only be as good as itself and there is no other model to compare. If model 1 is removed from the set with p -value equals k_1 and model 2 is removed afterwards with p -value equals k_2 , the test p -value if only model 1 and model 2 are excluded will be $\max\{k_1, k_2\}$. Therefore, the p -value may not decrease when a new model is excluded from the confidence set. We exclude models until the null hypotheses is not rejected. There are two statistics proposed by (55) to be used as a decision rule, the $T_{max,M}$ and the $T_{R,M}$. We adopted the first, since it is simple and easy to compute. The second statistic compares all the models two by two to create the set, making the procedure more intensive.

The MCS p -values are presented in Table 5.3. Values in bold represent models that remained in the confidence set with $\alpha = 20\%$. Autoregressive models and the Random Walk were removed from the set on most forecasting horizons. The only models which are in the confidence set for all horizons are the Flex-adaLASSO, the Random Forest, the Complete Subset Regression and the FOCUS forecast. If we include the cumulative forecasts than we remain only with the Flex-adaLASSO and the CSR as the models which are always in the set. If we look at the ranking, the CSR is the model with more p -values equal to 1.

We use the results from Table 5.3 to generate combined forecasts from the models in the confidence set. These results are displayed in Table 5.4. The first row of the table shows the forecasting errors from averaging the forecasts from all models. The second row shows the forecasting errors from averaging the forecasts of the models in the confidence set and the last row shows the forecasting error of the best model from Table 5.1 on each forecasting horizon³.

³The cumulative errors are calculated considering the 95% confidence set in order to include the specialist forecasts. This was done because of the results in Figure 5.3, which show that the specialist forecasts are negatively correlated with the other forecasts.

The table shows the Model Confidence Set p -values for all forecasting horizons and the 12 month accumulated inflation. Values in bold are included in the $\alpha = 20\%$ or 80% confidence set. The p -values can be used to rank the models. The model with p -value equals 1 is the best model, or the model that remains in all confidence sets.

Brazilian Consumer Price Index													
	Forecast Horizon												Acc
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
RW	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.03	0.01	0.05	0.15	0.68	0.01
AR	0.00	0.03	0.03	0.03	0.02	0.11	0.42	0.56	0.43	0.75	0.74	0.70	0.03
Factors	0.16	0.03	0.48	0.80	0.68	0.71	0.62	0.52	0.79	0.54	0.35	0.75	0.35
LASSO	0.94	0.48	0.07	0.31	0.28	0.22	0.42	0.13	0.43	0.05	0.02	0.01	0.06
F. aL.	0.79	1.00	0.24	0.65	0.68	0.71	0.90	0.46	0.66	0.67	0.90	0.95	0.34
P. OLS	0.76	0.74	0.19	0.91	0.64	0.59	0.76	0.42	0.41	0.85	0.79	0.93	0.35
RF	0.27	0.48	0.28	0.91	0.61	0.22	0.27	0.56	0.39	0.76	0.81	0.93	0.06
CSR	0.31	0.74	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
FOCUS	1.00	0.48	0.48	0.80	0.64	0.59	0.90	0.34	0.72	0.58	0.90	0.70	0.05
Top5	0.94	0.49	0.29	0.88	0.42	0.20	0.45	0.19	0.72	0.85	0.79	0.95	0.05

Table 5.3: Model Confidence Set

The table shows the forecasting errors of the average forecast of all models and of the models in the confidence set. The last row shows the best individual model to compare with the combined forecasts. All values are multiplied by 1000.

Brazilian Consumer Price Index													
RMSE \times 1000	Forecast Horizon												Acc.
	MAE \times 1000	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	
All Models	0.74	1.18	1.48	1.53	1.52	1.58	1.58	1.63	1.49	1.55	1.52	1.42	9.72
	(0.62)	(0.89)	(1.15)	(1.17)	(1.15)	(1.18)	(1.18)	(1.32)	(1.15)	(1.24)	(1.25)	(1.12)	(7.13)
MCS Models	0.42	0.71	0.80	1.22	1.15	1.73	1.38	1.81	1.33	1.22	1.19	1.27	9.69
	(0.33)	(0.58)	(0.63)	(0.97)	(0.85)	(1.35)	(1.08)	(1.47)	(1.05)	(0.97)	(0.95)	(0.99)	(6.72)
Best ind. Model	0.96	1.58	2.04	2.23	2.25	2.29	2.29	2.26	2.26	2.27	2.25	2.26	11.93
	(0.74)	(1.30)	(1.69)	(1.75)	(1.79)	(1.80)	(1.80)	(1.80)	(1.79)	(1.79)	(1.77)	(1.78)	(8.41)

Table 5.4: Combined Forecasts Mean Absolute Errors and Mean Squared Errors

The results in Table 5.4 show that the simple average of all models improves the results from the best individual model. The combined forecast from the MCS improves the results even more, especially considering the first forecasting horizons. Even on the horizon $h = 1$, which is only five days before the IPCA is published, the forecasting errors are considerably smaller when we combine forecasts. In many cases the forecasting errors are less than half the errors from the best individual model.

5.4.3 Look Ahead Bias on the MCS Combined Forecasts

Our combined forecasts based on the MCS is contaminated with look ahead bias as we need the forecasting errors in order to estimate the confidence set. However, the selected models in the confidence set tend to be stable over the time. To test how stable and to provide results free of look ahead bias we split the sample of 48 observations into two sub-samples: one with 36 observations to estimate the confidence set and another with 12 observations to

This table shows the forecasting errors of the average forecast of all models and of the models in the confidence set. The last row shows the best individual model to compare with the combined forecasts. All values are multiplied by 1000.

Brazilian Consumer Price Index													
RMSE \times 1000	Forecast Horizon												
(MAE \times 1000)	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	Acc.
All Models	0.75 (0.67)	1.62 (1.33)	2.23 (1.93)	2.24 (1.84)	2.32 (1.95)	2.39 (2.01)	2.29 (1.89)	2.27 (1.89)	2.20 (1.78)	2.10 (1.73)	2.03 (1.73)	2.09 (1.81)	15.67 (13.68)
MCS Models	0.43 (0.35)	0.93 (0.83)	1.02 (0.86)	1.71 (1.43)	1.71 (1.43)	1.35 (1.10)	1.80 (1.51)	1.61 (1.39)	1.85 (1.48)	1.64 (1.32)	1.71 (1.40)	1.86 (1.61)	12.08 (9.97)

Table 5.5: Combined Forecasts Mean Absolute Errors and Mean Squared Errors without Look Ahead Bias

estimate the combined forecasts. We also estimate the simple average forecast for these 12 months.

The results are displayed in Table 5.5. They show that the combined MCS forecasting errors calculated without look ahead bias are still smaller than those calculated with a simple average across all models. Note that these results are only for the last 12 months in the sample (January to December, 2015), which was the worst year in terms of GDP growth for the Brazilian economy in our sample.

5.4.4 Different Window Size

Given the length of the dataset, it is not viable to test the models on a completely different sample. However, we can check if changing the size of the rolling window, and consequently, the number of forecasts, has a significant impact on our results.

Increasing the window size from 9 to 10 years, reduces the number of forecasts (windows) from 48 to 24. Table 5.6 shows the forecasting RMSEs and MAEs when the models are estimated with a larger window of observations. The results are similar to the case of 48 windows. However, the errors in Table 5.6 are in general larger because the forecasts are just for 2014 and 2015, two years of more instability in the Brazilian economy (specially 2015). As we mentioned before, the forecasting errors are larger in 2015 for longer horizons, and that is what shifted the errors up. The target factor are the model with the smallest errors in several forecasting horizons. The other models worth mentioning are the LASSO and Flexible adaLASSO, which performed well on

The table shows the root mean squared error and the mean absolute deviation, in parenthesis, of the forecasts based on 24 rolling windows. The values in bold represent the best model in each measure of error and each forecasting horizon. All values are multiplied by 1000. The column Acc. shows the errors of the 12 month cumulative forecast build using the monthly forecasts.

Brazilian Consumer Price Index - 24 Rolling Windows													
RMSE * 1000	Forecast Horizon												Acc.
MAE * 1000	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10	t+11	t+12	
RW	2.69 (2.25)	3.80 (3.13)	4.37 (3.68)	4.81 (4.02)	5.14 (4.10)	5.46 (4.45)	5.72 (4.69)	5.29 (4.36)	4.40 (3.72)	3.77 (3.00)	3.45 (2.97)	3.17 (2.44)	43.16 (34.70)
AR	2.61 (2.24)	3.41 (2.66)	3.29 (2.68)	3.55 (2.81)	3.82 (3.12)	3.77 (3.06)	3.60 (2.90)	3.38 (2.68)	3.33 (2.71)	3.26 (2.62)	3.22 (2.61)	3.26 (2.54)	32.59 (29.57)
Factors	1.50 (1.15)	2.45 (1.98)	2.72 (2.32)	2.96 (2.40)	3.15 (2.49)	3.11 (2.41)	2.95 (2.20)	2.85 (2.33)	2.56 (2.09)	2.45 (2.01)	2.45 (1.97)	2.55 (2.07)	21.83 (18.04)
LASSO	0.95 (0.76)	2.15 (1.75)	2.87 (2.35)	3.17 (2.48)	3.21 (2.45)	3.24 (2.56)	3.23 (2.54)	3.21 (2.51)	3.20 (2.48)	3.32 (2.64)	3.79 (2.91)	3.55 (2.97)	25.27 (23.04)
F. aL	1.03 (0.83)	1.76 (1.46)	2.50 (2.08)	2.84 (2.32)	2.84 (2.24)	2.97 (2.32)	3.32 (2.64)	3.30 (2.60)	3.32 (2.59)	3.28 (2.56)	3.31 (2.60)	3.35 (2.70)	24.84 (22.83)
P. OLS	1.04 (0.83)	1.77 (1.48)	2.58 (2.08)	2.58 (2.08)	2.95 (2.28)	3.06 (2.38)	3.36 (2.72)	3.27 (2.58)	2.96 (2.33)	2.93 (2.31)	2.87 (2.17)	2.81 (2.15)	22.74 (20.02)
RF	1.65 (1.08)	2.30 (1.74)	3.03 (2.36)	3.11 (2.46)	3.36 (2.60)	3.69 (2.94)	3.57 (2.80)	3.64 (2.79)	3.39 (2.52)	3.10 (2.27)	3.08 (2.31)	2.99 (2.22)	25.17 (23.00)
CSR	1.05 (0.91)	1.87 (1.55)	2.44 (2.09)	2.71 (2.18)	2.77 (2.23)	2.75 (2.20)	2.68 (2.11)	2.71 (2.11)	2.69 (2.05)	2.59 (1.93)	2.62 (1.99)	2.63 (2.02)	18.04 (16.40)
FOCUS	0.97 (0.83)	2.14 (1.74)	2.99 (2.38)	3.13 (2.48)	3.20 (2.51)	3.25 (2.60)	3.22 (2.56)	3.20 (2.52)	3.22 (2.55)	3.25 (2.58)	3.27 (2.58)	3.28 (2.60)	26.95 (24.86)
Top 5	0.99 (0.78)	1.92 (1.55)	2.80 (2.21)	3.09 (2.42)	3.29 (2.57)	3.34 (2.64)	3.46 (2.64)	3.35 (2.63)	3.05 (2.48)	3.35 (2.58)	3.27 (2.54)	3.24 (2.48)	26.94 (24.72)

Table 5.6: Forecasts Mean Absolute Errors and Root Mean Squared Errors for 24 Rolling Windows

shorter horizons, and the Complete Subset Regression, which has good results for longer horizons. We already detected an improvement on target factor models on longer horizons in the results for 48 rolling windows. The difference it that for 24 rolling windows, factor models are able to beat the CSR in some cases.

We show the Model Confidence Set results for the 24 rolling window analysis in Table 5.7. The results are similar to those of the 48 windows. However, the only model in the confidence set on the accumulated inflation is the CSR. If we look at the monthly horizons individually, the models that are included in the 80% confidence set on all horizons are the Flexible adaLASSO, the Post-OLS estimated with the variables selected by the Flexible adaLASSO, the Random Forest and the CSR. The CSR was the last remaining model in six cases, against four of the target factors. The LASSO and the Flexible adaLASSO are the last remaining models in one case each.

5.4.5 Bayesian Alternatives

In this section we show the results using two alternative Bayesian models. First, the Unobserved Component Stochastic Volatility (UC-SV) model proposed by (52), which is a very popular model for the U.S. inflation;

The table shows the Model Confidence Set p-values for all forecasting horizons and the 12 month cumulative inflation using 24 rolling windows. Values in bold are included in the $\alpha = 20\%$ or 80% confidence set. The size of the p-values can be used to rank the models. The model with p-value equals 1 is the best model, or the model that remains in all confidence sets.

Brazilian Consumer Price Index - 24 Rolling Windows													
	Forecast Horizons												
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	Acc
RW	0.00	0.00	0.00	0.00	0.02	0.03	0.01	0.01	0.06	0.44	0.50	0.64	0.12
AR	0.00	0.08	0.23	0.39	0.04	0.03	0.14	0.72	0.26	0.27	0.43	0.43	0.03
Factors	0.09	0.10	0.56	0.75	0.35	0.67	0.33	0.60	1.00	1.00	1.00	1.00	0.06
LASSO	1.00	0.40	0.43	0.63	0.60	0.38	0.44	0.59	0.54	0.27	0.63	0.17	0.03
F. aL	0.86	1.00	0.50	0.34	0.55	0.57	0.58	0.55	0.26	0.44	0.63	0.20	0.05
P. OLS	0.91	0.60	0.50	0.75	0.45	0.67	0.44	0.59	0.28	0.35	0.33	0.45	0.06
RF	0.34	0.53	0.48	0.69	0.67	0.38	0.58	0.72	0.54	0.35	0.33	0.45	0.05
CSR	0.91	0.29	1.00	1.00	1.00	1.00	1.00	1.00	0.71	0.66	0.59	0.78	1.00
FOCUS	0.79	0.40	0.41	0.63	0.67	0.23	0.19	0.23	0.49	0.19	0.50	0.29	0.02
Top 5	0.79	0.53	0.56	0.59	0.39	0.38	0.37	0.55	0.51	0.19	0.19	0.64	0.12

Table 5.7: Model Confidence Set - 24 Rolling Windows

and second, a large Bayesian Vector Autorregressive (BVAR) using all variables in the dataset ad with priors defined as in (10).

5.4.5.1

Unobserved Component Stochastic Volatility Model

The UC-SV model is described by the following equations:

$$\begin{aligned}\pi_t &= \tau_t + e^{h_t/2} \varepsilon_t, \\ \tau_t &= \tau_{t-1} + u_t, \\ h_t &= h_{t-1} + v_t,\end{aligned}\tag{5-11}$$

where $\{\varepsilon_t\}$ is a sequence of independent and normally distributed random variables with zero mean and unit variance, $\varepsilon_t \sim N(0, 1)$, u_t and v_t are also normal with zero mean and variance given by inverse-gamma priors. $\tau_1 \sim N(0, V_\tau)$ and $h_1 \sim N(0, V_h)$, where $V_\tau = V_h = 0.12$. The model is estimated by Markov Chain Monte Carlo (MCMC) methods. The h -steps-ahead forecast is computed as $\hat{\pi}_{t+h} = \hat{\tau}_{t|t}$. We computed forecasts for the same forecasting horizons as the models from the previous sections and the forecasting errors are calculated for 48 months out-of-sample as before.

5.4.5.2

Bayesian Vector Autoregressive Model

Let $\mathbf{Y}_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})'$ be described as the following VAR model:

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \mathbf{u}_t, \quad (5-12)$$

where \mathbf{c} is a n -dimensional vector of constants, \mathbf{A}_i , $i = 1 \dots p$ are $(n \times n)$ matrices of coefficients and \mathbf{u}_t is the n -dimensional error vector. The same model may be written as a system of equations:

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{U} \quad (5-13)$$

where $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_T)'$ is a $(T \times n)$ matrix, $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_T)'$ is a $(T \times k)$ matrix with $k = np + 1$ and $\mathbf{X}_t = (1, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_{t-p})'$, $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_T)'$, and $\mathbf{B} = (\mathbf{c}, \mathbf{A}_1, \dots, \mathbf{A}_p)'$.

The model is estimated using dummy observations \mathbf{Y}_d and \mathbf{X}_d of dimensions $T_d \times n$ and $T_d \times k$ respectively (for details on creating the dummy observations see (10)). Using these dummies is equivalent to imposing the normal inverted Wishart prior on the covariance matrix of \mathbf{B} . The dummy observations are used to create $\mathbf{Y}^* = (\mathbf{Y}', \mathbf{Y}'_d)'$ and $\mathbf{X}^* = (\mathbf{X}', \mathbf{X}'_d)'$. The posterior mean of \mathbf{B} is the same as the ordinary least-squares (OLS) estimates of the regression of \mathbf{Y}^* on \mathbf{X}^* and also the same as the Minnesota prior. Additionally, by using the dummies we ensure that for each regression of the VAR the number of observations is larger than the number of variables, which makes the OLS estimation of \mathbf{B} feasible. Another important issue is the choice of the expected value of the priors for the diagonal of the \mathbf{A}_1 matrix. We choose the value of 0.5 for all elements.

5.4.5.3 Results

The results for the UC-SV and the large BVAR are in Table 5.8. The forecasting errors are calculated for the last 48 observations of the dataset using a rolling window scheme. The UC-SV does not use any information other

than the past inflation. Therefore, its forecasting errors are larger than the multivariate models that use the FOCUS and other macroeconomic variables as regressors. Nevertheless, the UC-SV is comparable to the other univariate models even though it cannot beat the AR model. The BVAR is much more accurate than the UC-SV and it competes directly with the high-dimensional models, especially for small forecasting horizons. However, it is not the best model in any horizon. Note that a h -steps-ahead for the BVAR is iterated, while the models in Table 5.1 are estimated directly for the horizon of interest. Table 5.9 shows the (42) (GW) test p -values, the null hypothesis is that both models have the same forecasting accuracy. The null is rejected in most cases for the UC-SV except when it is compared to the random walk and the AR models. It has the same forecasting ability as the random walk for small horizons and it is similar to the AR model for long horizons. Note that these models are all univariate and very simple. Therefore, it is natural that with only 48 rolling widows the test fails to detect significant differences between them. The Bayesian VAR (second panel of Table 5.9) is just as accurate as any other multivariate model five days before the CPI is published ($h = 1$). The reason is that all models have specialist's forecasts, which are very accurate right before the CPI is published, therefore, the forecasting errors are small for all multivariate models and any difference between them is not detected by the GW test. For larger forecasting horizons, the performance of the BVAR is clearly inferior. The CSR, which is the best models in most cases, is statistically different from the BVAR for all horizons larger than one.

The table shows the root mean squared error and the mean absolute deviation, in parenthesis, of the forecasts. The values in bold represent the best model in each measure of error and each forecasting horizon. All values are multiplied by 1000. The column Acc. shows the errors of the 12 month accumulated forecasts. The order of the BVAR is 4 and the UCSV is estimated by MCMC.

Brazilian Consumer Price Index													
RMSE \times 1000 (MAE \times 1000)	Forecast Horizon												Acc.
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	
UCSV	2.56 (2.14)	3.17 (2.66)	3.54 (2.91)	3.86 (3.11)	4.06 (3.19)	4.09 (3.24)	3.95 (3.06)	3.51 (2.75)	3.08 (2.47)	2.79 (2.28)	2.59 (2.10)	2.61 (2.00)	27.74 (21.79)
BVAR	1.19 (0.92)	2.14 (1.66)	2.53 (1.91)	2.73 (2.04)	2.88 (2.20)	2.98 (2.28)	3.05 (2.33)	3.09 (2.37)	3.11 (2.39)	3.12 (2.39)	3.13 (2.41)	3.14 (2.41)	19.27 (14.6)

Table 5.8: Forecasts Mean Absolute Errors and Root Mean Squared Errors for BVAR and UCSV

The table shows the Giacomini and White test p -values for all models compared to the UCSV (first block) and the Bayesian VAR (second block).

Brazilian Consumer Price Index - GW p -values													
Forecast Horizon - All models against UCSV													
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	Acc.
RW	0.423	0.763	0.515	0.255	0.075	0.020	0.001	0.001	0.000	0.001	0.010	0.358	0.183
AR	0.020	0.107	0.275	0.079	0.037	0.017	0.024	0.053	0.124	0.683	0.566	0.854	0.178
Factors	0.000	0.003	0.000	0.001	0.001	0.000	0.001	0.005	0.004	0.183	0.082	0.282	0.073
LASSO	0.000	0.000	0.029	0.203	0.001	0.002	0.001	0.668	0.250	0.061	0.004	0.004	0.043
F. al.	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.106	0.006	0.185	0.841	0.532	0.079
P. OLS	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.342	0.043	0.553	0.940	0.577	0.083
RF	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.004	0.023	0.447	0.794	0.503	0.059
CSR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.042	0.064	0.025
BVAR	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.085	0.912	0.223	0.059	0.042	0.002

Forecast Horizon - All models against BVAR													
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$	Acc.
RF	0.000	0.001	0.000	0.000	0.001	0.001	0.000	0.003	0.013	0.378	0.770	0.231	0.001
AR	0.000	0.001	0.001	0.001	0.045	0.232	0.968	0.188	0.051	0.035	0.023	0.008	0.003
Factors	0.518	0.810	0.453	0.311	0.061	0.043	0.027	0.014	0.009	0.035	0.010	0.025	0.092
LASSO	0.165	0.027	0.118	0.317	0.462	0.378	0.045	0.491	0.184	0.474	0.291	0.505	0.021
F. al.	0.200	0.019	0.070	0.123	0.024	0.012	0.032	0.516	0.011	0.020	0.058	0.011	0.041
P. OLS	0.210	0.031	0.127	0.330	0.128	0.057	0.049	0.987	0.051	0.100	0.205	0.021	0.042
RF	0.202	0.065	0.737	0.116	0.095	0.432	0.089	0.068	0.016	0.018	0.012	0.004	0.002
CSR	0.399	0.027	0.018	0.011	0.005	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.145
UCSV	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.085	0.912	0.223	0.059	0.042	0.002

Table 5.9: Giacomini and White test p -values comparing the UCSV and the Bayesian VAR to all other models

The main conclusion from the above results is that the two Bayesian alternatives considered in the paper have a performance inferior than the one from the machine learning methods.

5.4.6 Density Forecasts

So far we analyzed only point forecasts for several models. In this section we turn our attention to density forecasts. The point forecasts in a rolling window scheme provide good information on which model is more accurate on average. However, it does not tell us anything about the forecasting uncertainty. We obtain the predictive densities by bootstrapping the in-sample residuals. For each model, in each rolling window, we randomly selected 100,000 observations of the in-sample residuals and added it to the point forecast, resulting on an empirical predictive density⁴.

The predictive densities are used to estimate average log-scores following

⁴The procedure for the BVAR and the UCSV is slightly different because in these models we did not estimate direct forecasts. The $t+1$ densities were obtained in the same way as all other models. We used these densities to estimate bootstrap point forecasts, which were used to obtain the $t+2$ densities. We kept iterating the bootstrap point forecasts until $t+12$.

(60). Suppose we estimate predictive densities $\hat{f}(\cdot)$ for a given model. Let Y be the observed value of the variable on the period we aim to forecast. The log-scores are calculated as $S(\hat{f}, Y) = \log \hat{f}(Y)$. The log-scores will be larger when the probability of the observed Y is high. For each model, in each forecasting horizon, we compute average log-scores across all the rolling windows. The best model in each case is the one with the highest average log-score. (60) also propose a test to check whether the predictive densities are statistically equal for two models, which will be referred as AG test from here on. If we reject the null, then the two models have different densities.

The average log-scores and the AW test p -values are presented in Tables (5.10) and (5.11). The first interesting result in Table (5.10) is that although the LASSO and the Flex-adaLASSO provide similar point forecasts, the Flex-adaLASSO densities have larger log-scores, especially for long forecasting horizons. The CSR, which is the best model on the point forecast for most forecasting horizons, is also the one with the largest log-scores. However, for $t + 1$ and $t + 2$ the POLS and the Flex-adaLASSO are the best models even though the CSR has a good performance on these horizons too. The BVAR worked well for $t + 1$ but its performance deteriorates very fast for longer horizons. Note that for the BVAR, forecasts for horizons larger than one are calculated by iterating previous forecasts. As a result, point forecasts and variances converge fast to their respective unconditional values. Table (5.11) shows that when there is a large difference between the average log-scores of two models we, in general, reject the null and obtain statistically different models.

The table shows the average log-scores as described by (60). The log-scores were calculated on the empirical densities generated using bootstrap. Each bootstrap forecast is the sum of the point forecast and a random observation of the in-sample residuals. Each empirical density was constructed using 100000 bootstrap forecasts.

Brazilian Consumer Price Index												
Forecast Horizon - Average log-scores												
	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$	$t+9$	$t+10$	$t+11$	$t+12$
RW	-0.01	-0.99	-1.27	-1.79	-1.79	-1.74	-1.90	-1.78	-1.51	-1.17	-0.62	-0.61
AR	-0.05	-0.34	-0.50	-0.57	-0.48	-0.50	-0.49	-0.33	-0.28	-0.23	-0.18	-0.20
Factor	0.33	-0.01	-0.17	-0.29	-0.23	-0.21	-0.28	-0.18	-0.05	-0.18	-0.14	-0.16
LASSO	0.87	0.20	-0.79	-1.00	-0.44	-0.48	-0.33	-0.85	-0.47	-1.14	-1.62	-1.87
F. aL.	0.90	0.42	0.03	-0.02	-0.08	-0.06	-0.15	-0.46	-0.10	-0.30	-0.23	-0.04
P. OLS	0.93	0.39	0.02	-0.11	-0.15	-0.20	-0.30	-0.62	-0.06	-0.43	-0.35	-0.08
RF	0.14	-0.07	-0.24	-0.22	-0.20	-0.22	-0.18	-0.26	-0.28	-0.16	-0.20	-0.29
CSR	0.85	0.38	0.16	-0.04	-0.09	-0.10	-0.07	0.00	-0.02	0.00	-0.01	-0.02
UCSV	-0.85	-0.83	-0.83	-0.82	-0.80	-0.80	-0.80	-0.79	-0.78	-0.76	-0.76	-0.75
BVAR	0.62	-0.20	-0.40	-0.68	-0.69	-0.71	-0.81	-0.79	-0.82	-0.85	-0.99	-0.94

Table 5.10: Average log-scores for all models and all forecasting horizons

The table shows the Amisano and Giacomini test p -values. Each block shows the p -values for two different horizons of forecast (gray and white cells). The test statistics were calculated on the log-scores estimated from bootstrap densities. Each bootstrap forecast is the sum of the point forecast and a random observation of the in-sample residuals. Each empirical density was constructed using 100000 bootstrap forecasts.

Brazilian Consumer Price Index																						
AG test p -values																						
$t+2 \setminus t+1$	RW	AR	Factor	LASSO	F. a.L.	P. OLS	RF	CSR	UCSV	BVAR	$t+4 \setminus t+3$	RW	AR	Factor	LASSO	F. a.L.	P. OLS	RF	CSR	UCSV	BVAR	
RW	0.064	0.798	0.215	0.002	0.001	0.001	0.699	0.000	0.009	0.001	RW	0.031	0.029	0.007	0.244	0.004	0.005	0.012	0.007	0.163	0.031	
AR	0.013	0.201	0.160	0.001	0.001	0.001	0.603	0.001	0.003	0.003	AR	0.012	0.139	0.036	0.323	0.005	0.008	0.125	0.005	0.149	0.609	
Factor	0.004	0.003	0.210	0.020	0.033	0.027	0.660	0.043	0.011	0.293	Factor	0.194	0.196	0.049	0.019	0.181	0.225	0.477	0.122	0.011	0.174	
LASSO	0.002	0.011	0.032	0.230	0.768	0.513	0.066	0.829	0.001	0.141	LASSO	0.005	0.024	0.265	0.013	0.016	0.018	0.037	0.016	0.874	0.142	
F. a.L.	0.003	0.017	0.050	0.335	0.277	0.212	0.053	0.608	0.001	0.046	F. a.L.	0.007	0.051	0.449	0.020	0.045	0.863	0.090	0.315	0.005	0.081	
P. OLS	0.056	0.186	0.861	0.223	0.203	0.248	0.049	0.414	0.001	0.034	P. OLS	0.005	0.185	0.744	0.025	0.338	0.600	0.131	0.325	0.007	0.091	
RF	0.002	0.014	0.035	0.331	0.404	0.886	0.242	0.113	0.026	0.217	RF	0.006	0.031	0.153	0.007	0.916	0.642	0.294	0.088	0.020	0.313	
CSR	0.002	0.051	0.013	0.002	0.004	0.005	0.021	0.003	0.001	0.001	CSR	0.039	0.389	0.071	0.624	0.014	0.027	0.024	0.010	0.005	0.078	
UCSV	0.612	0.319	0.527	0.055	0.082	0.109	0.346	0.098	0.030	0.001	UCSV	0.025	0.652	0.172	0.317	0.036	0.073	0.045	0.039	0.022	0.152	
BVAR	0.026	0.057	0.108	0.017	0.004	0.007	0.004	0.006	0.022	0.054	BVAR	0.007	0.015	0.005	0.005	0.003	0.006	0.003	0.003	0.022	0.034	
$t+6 \setminus t+5$	RW	0.022	0.009	0.017	0.004	0.007	0.004	0.006	0.022	0.054	$t+8 \setminus t+7$	RW	0.015	0.005	0.005	0.005	0.003	0.006	0.003	0.003	0.022	0.034
AR	0.008	0.057	0.108	0.814	0.040	0.185	0.196	0.080	0.227	0.379	AR	0.007	0.286	0.286	0.550	0.277	0.556	0.300	0.177	0.389	0.200	
Factor	0.023	0.914	0.169	0.188	0.418	0.770	0.908	0.460	0.047	0.072	Factor	0.008	0.239	0.009	0.819	0.611	0.946	0.636	0.348	0.075	0.095	
LASSO	0.005	0.062	0.415	0.031	0.051	0.222	0.245	0.046	0.158	0.331	LASSO	0.100	0.063	0.009	0.240	0.240	0.816	0.293	0.097	0.056	0.122	
F. a.L.	0.005	0.249	0.975	0.183	0.127	0.610	0.401	0.929	0.011	0.067	F. a.L.	0.030	0.659	0.181	0.120	0.053	0.309	0.883	0.410	0.018	0.073	
P. OLS	0.005	0.246	0.956	0.153	0.390	0.936	0.814	0.753	0.039	0.190	P. OLS	0.061	0.383	0.075	0.318	0.445	0.250	0.413	0.121	0.060	0.150	
RF	0.007	0.115	0.547	0.045	0.702	0.524	0.461	0.413	0.024	0.070	RF	0.005	0.600	0.624	0.047	0.445	0.022	0.041	0.450	0.022	0.045	
CSR	0.046	0.293	0.024	0.282	0.013	0.042	0.024	0.014	0.014	0.767	CSR	0.004	0.065	0.143	0.007	0.029	0.022	0.041	0.008	0.010	0.047	
UCSV	0.045	0.354	0.082	0.374	0.054	0.147	0.058	0.066	0.769	0.001	UCSV	0.045	0.092	0.037	0.861	0.297	0.626	0.040	0.008	0.040	0.950	
BVAR	0.009	0.708	0.169	0.021	0.005	0.007	0.005	0.004	0.090	0.104	BVAR	0.055	0.046	0.047	0.890	0.426	0.696	0.059	0.022	0.979		
$t+10 \setminus t+9$	RW	0.009	0.008	0.021	0.005	0.007	0.005	0.004	0.090	0.104	$t+12 \setminus t+11$	RW	0.068	0.029	0.021	0.094	0.211	0.087	0.026	0.617	0.437	
AR	0.021	0.708	0.169	0.246	0.144	0.141	0.981	0.069	0.071	0.063	AR	0.079	0.631	0.007	0.798	0.444	0.874	0.095	0.037	0.033		
Factor	0.936	0.009	0.010	0.033	0.656	0.898	0.236	0.794	0.019	0.044	Factor	0.047	0.774	0.004	0.553	0.238	0.644	0.260	0.026	0.046		
LASSO	0.013	0.549	0.382	0.017	0.062	0.061	0.343	0.014	0.264	0.222	LASSO	0.005	0.002	0.001	0.006	0.006	0.006	0.007	0.004	0.089		
F. a.L.	0.030	0.259	0.179	0.045	0.117	0.443	0.241	0.423	0.014	0.038	F. a.L.	0.069	0.317	0.484	0.001	0.539	0.085	0.850	0.110	0.075		
P. OLS	0.009	0.471	0.879	0.009	0.237	0.143	0.252	0.714	0.013	0.040	P. OLS	0.088	0.482	0.668	0.001	0.240	0.339	0.469	0.063	0.174		
RF	0.007	0.090	0.137	0.004	0.024	0.030	0.159	0.115	0.056	0.075	RF	0.270	0.626	0.423	0.003	0.866	0.642	0.124	0.044			
CSR	0.319	0.042	0.041	0.368	0.074	0.233	0.015	0.010	0.011	0.023	CSR	0.045	0.180	0.227	0.001	0.020	0.028	0.088	0.015	0.012		
UCSV	0.511	0.040	0.069	0.530	0.088	0.205	0.048	0.022	0.776	0.882	UCSV	0.670	0.062	0.034	0.016	0.020	0.028	0.069	0.020	0.618		
BVAR											BVAR	0.349	0.016	0.050	0.086	0.026	0.030	0.069	0.020	0.618		

Table 5.11: Amisano and Giacomini test p -values comparing the forecasting densities of all models for all 12 horizons of forecast

5.5 Conclusion

We have tested several high-dimensional econometric models to forecast inflation in real-time and with a large number of predictors. We have also considered a forecasting combination mechanism based on the Model Confidence Sets. We have evaluated the methods discussed here with Brazilian inflation data (IPCA). The results can be summarized as follows.

For five-day-ahead, the LASSO and the FOCUS (expert forecast) are virtually the same and deliver the best forecasts. For the second horizon, the adaptive LASSO is superior than any other model considered. For the remaining horizons, the Complete Subset Regression dominates all other alternatives. The results are the same if we either use the root mean squared error or the mean absolute error. In terms of accumulated inflation, the Complete Subset Regression is the best model. However, most of the forecasts from different models are not statistically different according to the Model Confidence Set. We construct the final forecast as the average of the models included in the confidence set. This approach delivers the best forecasts among all the competing alternatives. The Bayesian VAR also produced accurate forecasts for shorter horizons but not as good as some of the high-dimensional models.

Finally, we computed predictive densities for all individual models using bootstrap and estimated log-scores to compare the models. The results are coherent with the point forecasts. Models from the LASSO family are better for $t + 1$ and $t + 2$ and the *CSR* is the best model for longer horizons.

5.6 Data Appendix

The dataset and the computer codes are available from <https://github.com/gabrielrvsc/hdeconometrics>. The “HD econometrics” repository is an R package with implementations of the models used in this

paper and the data in a .rda file. The package contains a number of functions used in the paper such as a function that selects the best LASSO model using the BIC, a function for the complete subset regression with several arguments to control for pre-testing and a function for the Bayesian VAR model. Documentation is available in markdown and follows the same format as traditional R documentation. It can also be viewed in R if the package is installed. Every function comes with an example in the documentation.

All the variables included in the models are listed in Tables 5.12 and 5.13. The first table shows the macroeconomic variables. They are all obtained on Bloomberg⁵. The second table shows the variables from the expectations database of the Brazilian Central Bank.

Most variables in our data are published for month t before the Brazilian CPI, which is made public around the 10th day of month $t + 1$. Some variables have some delay or may be available only after the CPI is published. In that case, we use the last available observation of such variables.

The first group of Table 5.12 covers prices and money. The CPI IPCA is the variable of interest and the CPI IPCA-15 is another index, which is released earlier and used as an indicator to the final CPI IPCA. These two indexes are officially adopted by the government. The FGV indexes are calculated by the Getúlio Vargas Foundation (FGV). They are also important measures of inflation. The second group of the table covers employment variables, the third group is for exchange rates, financial variables, savings and interest rates. IBOVESPA is the Brazilian most important stock index, BNDES is the national bank of investment, which lends money at lower rates and have significant impact on the national investment. The Selic is the target interest rate published by the Central Bank. The last group of variables in Table 5.12 covers government and international transactions.

All variables in Table 5.13 are obtained in the Brazilian Central Bank

⁵The names of the variables in Table 5.12 are the same names as in the Bloomberg database.

60	$t+1$ median	77	Top 5 $t+5$ median
61	$t+2$ median	78	Top 5 $t+6$ median
62	$t+3$ median	79	Top 5 $t+7$ median
63	$t+4$ median	80	Top 5 $t+8$ median
64	$t+5$ median	81	Top 5 $t+9$ median
65	$t+6$ median	82	Top 5 $t+10$ median
66	$t+7$ median	83	Top 5 $t+11$ median
67	$t+8$ median	84	Top 5 $t+12$ median
68	$t+9$ median	85	Top 5 $t+13$ median
69	$t+10$ median	86	$t+1$ median ²
70	$t+11$ median	87	$t+1$ mean
71	$t+12$ median	88	$t+1$ mean ²
72	$t+13$ median	89	$t+1$ Std
73	Top 5 $t+1$ median	90	$t+12$ median ²
74	Top 5 $t+2$ median	91	$t+2$ mean
75	Top 5 $t+3$ median	92	$t+2$ mean ²
76	Top 5 $t+4$ median	93	$t+2$ Std

Table 5.13: Focus Expectation Variables

6

Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods

JEL: C22, E31.

Keywords: Big data, inflation forecasting, LASSO, random forests, machine learning.

Abstract: Inflation forecasting is an important but difficult task. We explore the advances in machine learning (ML) methods and the availability of new datasets to forecast US inflation. Despite the skepticism in the previous literature, we show that ML models with a large number of covariates are systematically more accurate than the benchmarks. The ML method that deserves more attention is the random forest, which dominated all other models. Its good performance is due not only to its specific method of variable selection but also the potential nonlinearities between past key macroeconomic variables and inflation.

6.1

Introduction

It is difficult to overemphasize the importance of forecasting inflation in rational economic decision-making. Many contracts concerning employment, sales, tenancy, and debt are set in nominal terms. Therefore, inflation forecasting is of great value to households, businesses and policymakers. In addition, central banks rely on inflation forecasts not only to inform monetary policy but also to anchor inflation expectations and thus enhance policy efficacy. Indeed, as part of an effort to improve economic

decision-making, many central banks release their inflation forecasts on a regular basis.

Despite the great benefits of forecasting inflation accurately, improving simple benchmark models has been proven to be a major challenge for both academics and practitioners. As (61) emphasize, “it is exceedingly difficult to improve systematically upon simple univariate forecasting models, such as the (34) random walk model [...] or the time-varying unobserved components model in (52).” This conclusion is supported by a large literature; see (62) for a recent survey. Nonetheless, this literature has so far largely ignored the recent machine learning (ML) and “big data” boom in economics.¹ Moreover, previous works either focused on a restrictive set of variables or were based on a small set of factors computed from a larger number of potential predictors known as “diffusion indexes”; see, for example, (65).

“Big data” and ML methods are not passing fads, and investigating whether the combination of these two methods is able to provide more accurate forecasts is of paramount importance. (66), for example, show recently that machine learning methods coupled with hundreds of potential predictors improve substantially out-of-sample stock return predictions. In a similar spirit, and despite the previous skepticism, we argue that these methods lead to more accurate inflation forecasts. Moreover, this new set of models can also help to uncover the main predictors for future inflation, possibly shedding light on the drivers of price dynamics.

In this paper, we contribute to the literature in a number of ways. First, we robustly show that it is possible to beat the usual univariate benchmarks for inflation forecasting, namely, random walk (RW), autoregressive (AR) and unobserved components stochastic volatility (UCSV) models. We consider several ML models in a data-rich environment with hundreds of variables from

¹See (63) and (64) for discussions of ML methods and big data in economics. In this paper, we call ML models any statistical model that is able to either handle a large set of covariates and/or describe nonlinear mappings nonparametrically. Some of the methods have been around even before the “machines”.

the FRED-MD, a monthly database put together by (67), to forecast the US consumer price index (CPI) inflation during more than twenty years of out-of-sample observations and we show that the gains can be as large as 30% in terms of mean squared errors.²

Second, we highlight the main set of variables responsible for these forecast improvements. Our results indicate that such set of variables is not sparse, which corroborates the findings of (68) warning against the use of sparse predictive models. Indeed, we find that ML models that do not impose sparsity are the best performing ones. In contrast, the high level of aggregation of factor models, which has been one of the most popular models for macroeconomic forecasting, is not adequate.

Finally, we aim to give a guidance for the choice of which class of ML methods should be used for inflation forecasting. Throughout the paper, we pay special attention to a particular ML model, the random forest (RF) of (9), which systematically outperforms the benchmarks, factor models and ten additional ML methods covering a wide class of specifications: the least absolute shrinkage and selection operator (LASSO) family, which includes LASSO, adaptive LASSO, elastic net and the adaptive elastic net; ridge regression (RR); Bayesian vector autoregressions (BVAR); and linear ensemble methods such as bagging, boosting, complete subset regressions (CSR) and jackknife model averaging (JMA). RF models are highly nonlinear nonparametric models that have a tradition in statistics but have only recently attracted attention in economics. This late success is partly due to the new theoretical results developed by (69) and (70). Notably, (66) also find that RF, by allowing for nonlinearities, substantially improves stock return predictions.

6.1.1

²In the appendix, we show similar results for both the personal consumption expenditures (PCE) inflation and the core CPI inflation.

Main takeaways

First, as mentioned before, and contrary to the previous evidence in (33, 52), (34), and many others, our results show that consistently beating the benchmark specifications is possible. The ML models outperform the univariate alternatives, especially if we consider the 2001–2015 period, when the US inflation was more volatile compared to the 1990s. This is a robust finding for both individual horizons and the accumulated twelve-month forecasts. Second, although there is strong evidence of the existence of a small number of factors that drive the joint dynamics of the potential predictors, factor models deliver inferior forecasts compared to ML alternatives and are inferior to the RW method for the accumulated twelve-month horizon. Furthermore, either replacing standard principal component factors with target factors, as advocated by (6), or using boosting to select factors as discussed in (6), improves the results only marginally. Third, RR has a superior performance once compared to the other linear ML methods, especially for short horizons. However, the RF model delivers the smallest errors for most of the forecasting horizons for the CPI inflation. The gains, in terms of mean squared error reduction, can be, on average, of the order of 30%. This is a robust finding that is independent of the sample considered, the state of the economy or the level of either macroeconomic, financial uncertainty or real uncertainty. The RF model is an ensemble of fully grown regression trees estimated on different bootstrap subsamples of the original data. Therefore, the RF model is a nonsparse, highly nonlinear specification that aims to reduce the high variance of a single regression tree.

To open the black box of ML methods, we compare the variables selected by the adaptive LASSO method, RR, and the RF alternative. Following (67), we classify variables into eight different groups: (i) output and income; (ii) labor market; (iii) housing; (iv) consumption, orders and inventories; (v) money and credit; (vi) interest and exchange rates; (vii) prices; and (viii) stock market. In

addition, we also consider autoregressive terms and the principal component factors computed from the full set of potential predictors. The most important variables for RR and RF models are stable across forecasting horizons but are quite different between the two specifications. While for RR, autoregressive terms, prices and employment are the most important predictors, resembling a sort of backward-looking Phillips curve, RF models give more importance to prices, interest-exchange rates, employment and housing. LASSO selection is quite different across forecasting horizons and is, by construction and in opposition to RF and RR models, sparse. Only AR terms retain their relative importance independent of the horizon and prices gradually lose their relevance until up to six months ahead but partially recover for longer horizons. Output-income variables are more important for medium-term forecasts. Finally, none of the three classes of models selects either factors or stocks. Not even RR or RF which produce nonsparse variable selection. This result may indicate that the high level of cross-section aggregation of the factors is one possible cause for the poor performance of the factor models.

To disentangle the effects of variable selection from nonlinearity, we also consider alternative models. In the first method, we use the variables selected by the RF model and estimate a linear specification by OLS. In the second method, we estimate an RF specification with only the regressors selected by the adaptive LASSO method. Both models outperform the RF only for one-month-ahead forecasting. For longer horizons, the RF model is still the winner, which provides evidence that both nonlinearity and variable selection play a key role in the superiority of the RF model.

There are many sources of nonlinearities relating the variables selected and inflation that could justify the superiority of the RF model. For instance, the relationship between inflation and employment is nonlinear to the extent that it depends on the degree of slackness in the economy. Another source of nonlinearities is economic uncertainty as this uncertainty increases the

option value of economic decision delays if they have an irreversible component (71). For example, if it is expensive to dismiss workers, hiring should be nonlinear on uncertainty. In addition, this real option argument also makes households and businesses less sensitive to changes in economic conditions when uncertainty is high. Hence, the responses of employment and inflation to interest rate decisions are arguably nonlinear on uncertainty. The presence of a zero lower bound on nominal interest rates and the implications of this bound for unconventional monetary policy is another source of nonlinearity among inflation, employment and interest rate variables (72, 73). Finally, to the extent that houses serve as collateral for loans, it interacts with monetary policy (74) and financial intermediation (75, 76). As in the Great Recession, a housing bubble can form, resulting in a deep credit crash (77, 78). Needless to say, these interactions are highly nonlinear and arguably have nonlinear effects on inflation, employment and interest rates.

6.1.2

A brief comparison of the recent literature

The literature on inflation forecasting is vast, and there is substantial evidence that models based on the Phillips curve do not provide good inflation forecasts. Although (33) showed that many production-related variables are potential predictors of US inflation, (34) showed that in many cases, the Phillips curve fails to beat even simple naive models. These results inspired researchers to look for different models and variables to improve inflation forecasts. Among the variables used are financial variables (36), commodity prices (35) and expectation variables (37). However, there is no systematic evidence that these models outperform the benchmarks.

More recently, due to the advancements in computational power, theoretical developments in ML, and availability of large datasets, researchers have started to consider the usage of high-dimensional models on top of the well-established (dynamic) factor models of (65), (4, 5), and (79, 80).

However, most of these studies have either focused only on a very small subset of ML models or presented a restrictive analysis. For example, (44) considered bagging, factor models and other linear shrinkage estimators in an exercise to forecast US inflation with a small set of real economic activity indicators. Their study is more limited than ours both in terms of the pool of models considered and richness of the set of predictors. Nevertheless, the authors were among the few voices that suggested that ML techniques can deliver nontrivial gains over univariate benchmarks. (17) provided evidence that LASSO-based models outperform both factor and AR benchmarks in forecasting US CPI. However, the analysis in the paper is restricted to a single ML method for just one-month-ahead inflation forecasting.

Most of the previous papers in the inflation forecasting literature have explored only linear ML models but ignored nonlinear alternatives. One possible explanation for this limitation is that most of the papers in the early days considered only univariate nonlinear models that were, in most cases, outperformed by simple benchmarks; see (81) for an example.³ The message of our paper is that the combination of a rich dataset with modern ML tools is responsible for the nontrivial forecasting gains over traditional univariate benchmarks.⁴

Finally, we do not apply ML methods as pure black-boxes specifications. In fact, this paper is different from many “horse-races” in the literature, as we not only compare a large number of different models, but we also try to clarify the mechanisms why a given class of models is superior than others.

6.2

³An exception is (82), who show that neural networks outperform univariate autoregressive models for short horizons.

⁴More recently, (12) applied a large number of ML methods, including RFs, to real-time inflation forecasting in Brazil. The results were very promising and indicated a clear superiority of the CSR method of (45, 20). However, an important question is whether this is a particular result for Brazil or if similar findings can be replicated for the US economy.

Data

Our data consist of variables from the FRED-MD database, which is a large monthly macroeconomic dataset designed for empirical analysis in data-rich environments. The dataset is updated in real-time through the FRED database and is available from Michael McCracken's webpage.⁵ For further details, we refer to (67).

In this paper, we use the vintage as of January 2016. Our sample goes from January 1960 to December 2015 (672 observations), and only variables with all observations in the defined sample period are used (122 variables). In addition, we include as potential predictors the four principal component factors computed from this set of variables. We consider four lags of all variables, as well as four autoregressive terms for inflation, so the analysis contemplates 508 potential predictors. The out-of-sample window is from January 1990 to December 2015. All variables are transformed as described in Appendix 6.5. The price indexes are log-differenced only one time. Therefore, π_t is the inflation in month t computed as $\pi_t = \log(P_t) - \log(P_{t-1})$, and P_t is a given price index in period t . The baseline price index is the CPI, but in Appendix 6.5, we also report results considering both the PCE and the core CPI. Figure 6.1 displays the evolution of these inflation measures during the full sample period.

We compare performances not only across models in the out-of-sample window but also in two subsample periods, namely, 1990 to 2000 (132 out-of-sample observations) and 2001 to 2015 (180 out-of-sample observations). In Table 6.1, we report the mean, standard deviation (Sd), median, maximum, minimum, first-order autocorrelation (AC1), and sum of the first 36 autocorrelations (AC36) for several macroeconomics variables. These variables include CPI monthly inflation (π_t), CPI twelve-month inflation ($\pi_{12,t}$), monthly growth of the industrial production (ΔIP_t) and

⁵<https://research.stlouisfed.org/econ/mccracken/fred-databases/>.

twelve-month growth of industrial production ($\Delta_{12}IP_t$). We also report these statistics for measures of macroeconomic, financial and real uncertainty computed as in (46), which are the conditional volatility of the unforecastable part of macroeconomic, financial and firm-level variables, respectively. In particular, the authors consider forecasting horizons of one, three and twelve months ahead.⁶

The statistics in Table 6.1 give an overview of the economic scenario in each subsample. The first sample corresponds to a period of low inflation volatility ($\sigma = 0.17\%$), while in the second sample, inflation is much more volatile ($\sigma = 0.32\%$). However, on average, inflation is higher during 1990-2000 than 2001-2015 and much more persistent as well. Relative to the 1990-2000 period, inflation was more volatile near the recession in the early 1990s. The monthly growth in industrial production is on average higher and less volatile during the first subsample. Finally, uncertainty measures are uniformly higher during 2001-2015, mainly due to the Great Recession.

6.3 Methodology

Consider the following model:

$$\pi_{t+h} = T_h(\mathbf{x}_t) + u_{t+h}, \quad h = 1, \dots, H, \quad t = 1, \dots, T, \quad (6-1)$$

where π_{t+h} is the inflation in month $t + h$; $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})'$ is a n -vector of covariates, possibly containing lags of π_t and/or common factors as well as a large set of potential predictors; $T_h(\cdot)$ is the mapping between covariates and future inflation; and u_t is a zero-mean random error. The target function $T_h(\mathbf{x}_t)$ can be a single model or an ensemble of different specifications. There is a different mapping for each forecasting horizon.

The direct forecasting equation is given by

⁶These uncertainty measures are available at Sydney C. Ludvigson's webpage (<https://www.sydneyludvigson.com/>).

$$\hat{\pi}_{t+h|t} = \hat{T}_{h,t-R_h+1:t}(\mathbf{x}_t), \quad (6-2)$$

where $\hat{T}_{h,t-R_h+1:t}$ is the estimated target function based on data from time $t - R_h + 1$ up to t and R_h is the window size, which varies according to the forecasting horizon and the number of lagged variables in the model. We consider direct forecasts as we do not make any attempt to predict the covariates. The only exception is the case of the BVAR model, where joint forecasts for all predictors are computed in a straightforward manner following the procedure described in (10).

The forecasts are based on a rolling window framework of fixed length. However, as mentioned before, the actual in-sample number of observations depends on the forecasting horizon. For example, for the 1990–2000 period, the number of observations is $R_h = 360 - h - p - 1$, where p is the number of lags in the model. For 2001–2015, $R_h = 492 - h - p - 1$.

In addition to three benchmark specifications (RW, AR and UCSV models), we consider factor-augmented AR models, sparsity-inducing shrinkage estimators (LASSO, adaptive LASSO, elastic net and adaptive elastic net), other shrinkage methods that do not induce sparsity (RR and BVAR with Minnesota priors), averaging (ensemble) methods (bagging, CSR and JMA)⁷ and RF. With respect to the factor-augmented AR models, we consider in addition to the standard factors computed with principal component analysis a set of target factors as advocated by Bai and Ng (6) and boosted factors as in Bai and Ng (7). A detailed discussion of the models implemented in this paper can be found in Appendix 6.5. Finally, we also include in the comparison three different model combination schemes, namely, simple average, trimmed average and the median of the forecasts.

We find that the RF, a highly nonlinear method, robustly outperforms other methods. To disentangle the role of variable selection from nonlinearity, we also consider a linear model where the regressors are selected by the RFs

⁷Bagging and CSR can also be viewed as nonsparsity-inducing shrinkage estimators.

(RF/ordinary least squares, OLS) and an RF model with regressors preselected by adaptive LASSO (adaLASSO/RF).

Forecasts for the accumulated inflation over the following twelve months is computed, with the exception of the RW and UCSV models, by aggregating the individual forecasts for each horizon. In the case of the RW and UCSV models, a different specification described in Appendix 6.5 is used to construct the forecast of the 12-month inflation.

6.4 Results

In this section, we describe our main results for the CPI. More detailed results, robustness checks and a similar set of results for the PCE and the CPI-core can be all found in Appendix 6.5.

The models are compared according to three different statistics, namely, the root mean squared error (RMSE), the mean absolute error (MAE) and the median absolute deviation from the median (MAD), which are defined for each model and forecasting horizon as follows:

$$\begin{aligned} \text{RMSE}_{m,h} &= \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T \hat{e}_{t,m,h}^2, \\ \text{MAE}_{m,h} &= \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T |\hat{e}_{t,m,h}|, \\ \text{MAD}_{m,h} &= \text{median} [|\hat{e}_{t,m,h} - \text{median}(\hat{e}_{t,m,h})|], \end{aligned}$$

where $\hat{e}_{t,m,h} = \pi_t - \hat{\pi}_{t,m,h}$ and $\hat{\pi}_{t,m,h}$ is the inflation forecast for month t made by model m with information up to $t - h$. The first two performance measures above are the usual ones in the forecasting literature. MAD, which is less commonly used in empirical papers, is robust to outliers.

To test whether the forecasts from different models are different, we consider a number of tests, namely, the model confidence sets (MCS) as proposed in (55), the superior predictive ability (SPA) tests of (83), and the multi-horizon uniform SPA test of (84).

6.4.1 Overview

In this section, we report an overview of the main findings of the paper.

Tables 6.2–6.4 report a number of statistics for each model across all the forecasting horizons, including the accumulated twelve-month horizon. The first three columns report the average RMSE, the average MAE and the average MAD. Columns (4), (5) and (6) report the number of times (out of thirteen possible horizons)⁸ each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present, for square and absolute losses, the average p -values of the MCS based either on the **range** or the t_{\max} statistics. Columns (11) and (12) show the average p -values of the SPA test for the squared and absolute errors, respectively. Finally, columns (13) and (14) display the p -value of the uniform multi-horizon test for superior predictive ability and the p -value of the MCS based on the multi-horizon comparison of the models, respectively. The uniform SPA test is designed to check for superior performance at each individual horizon. Table 6.2 shows the results for the full out-of-sample period (1990–2015), whereas Tables 6.3 and 6.4 present the results for the subsample periods 1990–2000 and 2001–2015, respectively. The bold figures highlight the best-performing model. The following facts emerge from the tables:

1. Machine learning models and the use of a large set of predictors are able to systematically improve the quality of inflation forecasts over traditional benchmarks in the literature. This is a robust and statistically significant result.
2. The RF model outperforms all the other alternatives in terms of point statistics. The superiority of RF is due both to the variable selection mechanism induced by the method as well as the presence of nonlinearities in the relation between inflation and its predictors. RF

⁸To be precise, monthly inflation from one month up to twelve months ahead and yearly inflation accumulated over the following twelve months.

has the lowest RMSEs, MAEs, and MADs across the horizons and the highest MCS p -values. The RF model also has the highest p -values in the SPA test, multi-horizon SPA test and multi-horizon MCS. The improvements over the RW in terms of RMSE, MAE and MAD are almost 30% and are more pronounced during the second subsample, where inflation volatility is much higher.

3. Shrinkage methods also produce more precise forecasts than the benchmarks. Sparsity-inducing methods are slightly worse than nonsparseness-inducing shrinkage methods. Overall, the forecasting performance among shrinkage methods is very similar, and ranking them is difficult.
4. Factor models are strongly outperformed by other methods. The adoption of boosting and target factors improves the quality of the forecasts produced by factor models only marginally. The poor performance of factor models is more pronounced during the first subsample (low volatility period).
5. CSR and JMA do not perform well either and are comparable to the factor models.
6. Forecast ensembles do not bring any significant improvements in any of the performance criteria considered.
7. Among the benchmark models, both AR and UCSV outperform the RW alternative. Furthermore, the UCSV model is slightly superior to the AR specification.

6.4.2

Results: Random Forests versus Benchmarks

Tables 6.5–6.7 show the results of the comparison between the RF and the benchmark models. Table 6.5 presents the RMSE, MAE and MAD ratios

of the AR, UCSV and RF models with respect to the RW alternative for all the forecasting horizons as well as for the accumulated forecasts over twelve months. The models with the smallest ratios are highlighted in bold. It is clear from the table that the RF model has the smallest ratios for all forecasting horizons.

To check whether this is a robust finding across the full out-of-sample period, we also compute rolling RMSEs, MAEs, and MADs over windows of 48 observations. Table 6.6 reports the results. The table shows the frequency with which each model achieved the lowest RMSEs, MAEs and MADs as well as the frequency with which each model was the worst-performing alternative among the four competitors. The RF model is the winning specification and is superior to the competitors for the majority of time periods, including the Great Recession. In contrast, the RW model delivers the worst forecasts most of the time. Figures 6.2, 6.3, and 6.4 show the rolling RMSEs, MAEs, and MADs, respectively, over the out-of-sample period. As expected, the performance of the RW deteriorates as the forecasting horizon increases. However, the accomplishments of the RFs seem rather robust.

Finally, Table 6.7 reports the p -values of the unconditional Giacomini and White (2000) test for superior predictive ability for squared (panel (a)) and absolute errors (panel (b)). Rejections of the null mean that the forecasts are significantly different. It is evident from the table that the RF has forecasts that are significantly different from and superior to the three benchmark models.

6.4.3

Results: The Full Picture

In this section, we compare all models. The main results are shown in Tables 6.8–6.10. Table 6.8 presents the results for the full out-of-sample period, whereas Tables 6.9 and 6.10 present the results for the 1990–2000 and 2000–2015 periods, respectively. The tables report the RMSEs and, between parenthesis, the MAEs for all models relative to the RW specification. The

error measures were calculated from 132 rolling windows covering the 1990–2000 period and 180 rolling windows covering the 2001–2015 period. Values in bold denote the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% MCS using the squared error (absolute error) as the loss function. The MCSs were constructed based on the maximum t statistic. The last column in the table reports the number of forecast horizons in which the model was included in the MCS for the square (absolute) loss. The last two rows in the table report the number of models included in the MCS for the square and absolute losses.

Several conclusions come out from the tables and we start by analyzing the full out-of-sample period. Apart from a few short horizons, where either the RF/OLS or the adaLASSO/RF alternatives are the winning models, the RF alternative delivers the smallest ratios in most of the cases. The RF is followed closely by shrinkage models, where RR seems be superior to the other alternatives. RR, RF and the hybrid linear-RF models are the only specifications included in the MCS for all forecasting horizons. Neither RF nor RR impose sparsity, which may corroborate the conclusions of (68), who provide evidence against sparsity in several applications. Factor models have very poor results and are almost never included in the MCS. When factors are combined with boosting, there is a small gain, but the results are still greatly inferior to those from RF and shrinkage models. This is particularly curious as there is a correspondence between factor models and RR: RR predictions are weighed combinations of all principal component factors of the set of predictors. Several reasons might explain the difference. Firstly, lack of clear factor structure in the regressors. This is not the case as shown in Figure 6.5, where we display the eigenvalues of the covariance matrix of regressors over the forecasting period. As can be seen, there is a small number of dominating factors. Secondly, there might be factors which explain only a small portion of the total variance of the regressors but have a high predictive power on

inflation. Again, we do not think this is the case as target factors as well as boosting are specifically designed to enhance the quality of the predictions but, in this case, do not bring any visible improvement. Furthermore, we allow the ML methods to select factors as well and, as we are going to show later, they are never selected. Lastly, we believe the most probable explanation is that although sparsity can be questioned, factor models are a too aggregated representation of the potential predictors. The results of JMA are not encouraging either. Nevertheless, all the competing models outperform the RW for almost all horizons. Finally, forecast combination does not provide any significant gain, which can be explained by the empirical fact that most of the forecasts are positively correlated, as depicted in Figure 6.6.

Focusing now on the two subsamples, the following conclusions stand out from the tables. The superiority of RF is more pronounced during the 2000–2015 period, when inflation is much more volatile. During this period, RF achieves the smallest RMSE and MAE ratios for almost all horizons. From 1990–2000, the linear shrinkage methods slightly outperform the RF for short horizons. However, RF dominates for long horizons and for the twelve-month forecasts. Among the shrinkage models and during the first period, there is no clear evidence of a single winner. Depending on the horizon, different models perform the best. Another salient fact is that there are fewer models included in the MSC during the first subsample.

Finally, we test whether the superiority of the RF model with respect to alternative models depends on the state of the economy. We consider two cases, namely, recessions versus expansions and high versus low macroeconomic uncertainty.⁹ The results of the test proposed by (42) for squared loss functions are presented in Tables 6.11 and 6.12. The tables report the value of the test statistic as well as the respective p -values. As usual, one, two and three asterisks represent rejection of the null hypothesis at 10%, 5%, and 1% significance

⁹Since results barely change if we consider either financial or real, rather than macroeconomic, uncertainty, we do not report them for brevity. They are available upon request.

levels, respectively. In Table 6.11, the results for expansion periods versus recessions are presented, whereas in Table 6.12, we consider periods of high macroeconomic uncertainty versus periods of low macroeconomic uncertainty. Periods of high (low) macroeconomic uncertainty are those where uncertainty is higher (lower) than the historical average. For conciseness, we display only the results for the most relevant models.

Inspecting the tables, it is clear that the majority of the statistics are negative, meaning that the RF model is superior than its competitors. For instance, out of 72 entries in each table, the values of the statistics are positive only in four and seven cases in Tables 6.11 and 6.12, respectively. However, the differences are not statistically significant during recessions. This result is not surprising as only 34 of the 312 out-of-sample observations are labeled as recessions. Nevertheless, the magnitudes of the differences are much higher during recessions. Turning attention now to periods of low and high macroeconomic uncertainty, it is evident from Table 6.12 that the RF model is statistically superior to the benchmark models for both periods, and as in the previous case, the differences are higher in periods of greater uncertainty. As argued above, both the degrees of slackness and uncertainty might be sources of nonlinearities in the economy. The fact that the RF model outperforms competitors in these states of the economy suggests that allowing for nonlinearities is key to improving macroeconomic forecasts.

6.4.4 Opening the Black Box: Variable Selection

In this section, we compare the predictors selected by some of the ML methods, namely, adaptive LASSO (adaLASSO), ridge regression (RR) and random forest (RF). We select these three models for two reasons. First, they are generally the three best-performing models. Second, they have quite different characteristics. While adaLASSO is a true sparsity-inducing method, RR and RF models are only approximately sparse. In addition, RR is a linear

model, whereas RF is a highly nonlinear specification.

In principle, this analysis is straightforward with sparsity-inducing shrinkage methods such as the adaLASSO, as the coefficients of potentially irrelevant variables are automatically set to zero.¹⁰ For the other ML methods, the analysis is more complex. To keep the results among models comparable, we adopt the following strategy. For RR and adaLASSO, the relative importance measure is computed as the average coefficient size (divided by the respective standard deviations of the regressors). To measure the importance of each variable for the RF models, we use out-of-bag (OOB) samples.¹¹ When the b th tree is grown, the OOB samples are passed down the tree and the prediction accuracy is recorded. Then, the values of the j th variable are randomly permuted in the OOB sample, and the accuracy is again computed. The decrease in accuracy due to the permutation is averaged over all trees and is the measure of the importance of the j th variable in the RF.

Due to space constraints, we cannot show the relative importance for each variable, each lag, each horizon or each estimation window. Therefore, as described in Appendix 6.5, and following (67), we categorize the variables, including lags, into the following nine groups: (i) output and income; (ii) labor market; (iii) housing; (iv) consumption, orders and inventories; (v) money and credit; (vi) interest and exchange rates; (vii) prices; and (viii) stock market. We also consider two additional groups, namely, the principal component factors computed from the full set of potential predictors and autoregressive terms. Furthermore, the results are averaged across all estimation windows.

Figure 6.7 shows the importance of each variable group for the adaLASSO, RR and RF methods for all the twelve forecasting horizons. For all different methods, the values in the plots are re-scaled to sum one.

¹⁰(17) showed, for example, that under sparsity conditions, the adaLASSO model selection is consistent for high-dimensional time series models in very general settings, i.e., the method correctly selects the “true” set of regressors.

¹¹For a given data point (y_t, \mathbf{x}'_t) , the OOB sample is the collection of all bootstrap samples that do not include (y_t, \mathbf{x}'_t) .

The set of the most relevant variables for RR and RF models is quite stable across forecasting horizons but is remarkably different between the two specifications. While for RR, autoregressive terms, prices and employment are the most important predictors, RF models give more importance to prices, interest-exchange rates, employment and housing. For adaLASSO, selection is quite different across forecasting horizons, and only autoregressive terms retain their relative importance independent of the horizon. Prices gradually lose their relevance until up to six-months-ahead and partially recover relevance when longer horizons are considered. Output-income variables are more important for medium-term forecasts. Finally, none of the three classes of models selects either factors or stocks. This result may indicate that the high level of cross-section aggregation of the factors is causing the poor performance of factor models.

To compare the degree of sparsity of each model, we report word clouds of the selected variables in Appendix 6.5.

6.5 Conclusions

We show that with the recent advances in ML methods and the availability of new and rich datasets, it is possible to improve inflation forecasts. Models such as LASSO, bagging, RF and others are able to produce more accurate forecasts than the standard benchmarks. These results highlight the benefits of ML methods and rich datasets for macroeconomic forecasting. Although our paper focuses on inflation forecasting in the US, one can easily apply ML methods to forecast other macroeconomic series in a variety of countries. We leave for further research the question as to whether ML methods can systematically outperform standard methods when other macroeconomic series, such as industrial production, and other countries are considered.

The RF method deserves special attention as it delivers the smallest

errors for most forecasting horizons in the two out-of-sample periods (1990–1999 and 2001–2015). The good performance of the RF is due both to potential nonlinearities in the relationship between inflation and its predictors and the variable selection mechanism of such a model.

The selection of variables for RF models is quite stable across forecasting horizons. These variables are mostly selected from the following four groups of variables: prices, exchange and interest rates, housing and labor market. Although it is difficult to disentangle the precise sources of nonlinearities that the RF method uncovers, this variable selection may shed light on them. In fact, there are many theoretical reasons that nonlinearities may be induced among inflation, interest rate, labor market outcomes and housing. For example, the relationship between inflation and employment depends on the degree of slackness in the economy. In addition, as we argued above, uncertainty might induce nonlinearities among these variables. Finally, part of the out-of-sample window encompasses quarters when the zero lower bound on nominal interest rates is binding, which is another source of nonlinearity. This out-of-sample window also encompasses a period in which a housing bubble led to a credit crunch, which are events with highly nonlinear consequences.

The RF is the winning method not only in the full sample but also in the periods of expansion and recession as well as low uncertainty and high uncertainty. Relative to other methods, the RF performs particularly well in periods of high uncertainty. In addition, the RF also outperforms other methods during and after the Great Recession, when uncertainty skyrocketed and when the zero lower bound was binding. Altogether, these results suggest that the relationships among key macroeconomic variables might be highly nonlinear. If this is the case, the various linear methods widely applied in the profession not only to forecast variables but also to achieve other objectives such as approximate DSGE models might lead to inaccurate results.

Finally, in this paper, we focus on the RF model due to its flexibility

and scalability for very large datasets. Nevertheless, alternative nonlinear methods such as deep learning and other semiparametric models should also be considered in future work.

TABLES AND FIGURES

The figure shows the time evolution of the consumer price index (CPI), the personal consumption expenditures (PCE) and the core CPI inflation measures from January 1960 to December 2015 (672 observations). Inflation is computed as $\pi_t = \log(p_t) - \log(p_{t-1})$, where p_t represents each one of the price measures considered in this paper. Shaded areas represent recession periods.

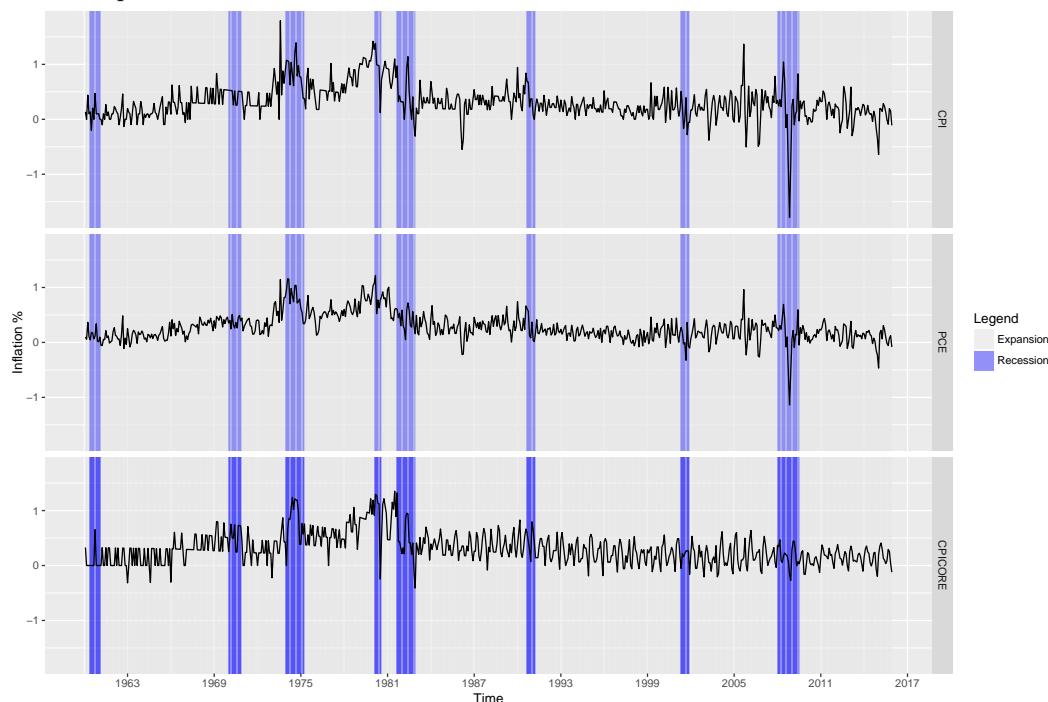


Figure 6.1: Inflation rate (CPI, PCE and CPI core) from 1960 to 2015.

The table reports the mean, standard deviation (Sd), median, maximum, minimum, first-order autocorrelation (AC1), and sum of the first 36 autocorrelations (AC36) for several macroeconomics variables, namely: CPI monthly inflation (π_t), CPI twelve-month inflation ($\pi_{12,t}$), monthly growth of industrial production (ΔIP_t), twelve-month growth of industrial production ($\Delta_{12}IP_t$), and measures of macroeconomic, financial and real uncertainty for one, three and twelve months.

Statistic	π_t	$\pi_{12,t}$	ΔIP_t	$\Delta_{12}IP_t$	$\mathcal{U}_{fn}(1)$	$\mathcal{U}_{fn}(3)$	$\mathcal{U}_{fn}(12)$	$\mathcal{U}_{macro}(1)$	$\mathcal{U}_{macro}(3)$	$\mathcal{U}_{macro}(12)$	$\mathcal{U}_{real}(1)$	$\mathcal{U}_{real}(3)$	$\mathcal{U}_{real}(12)$
Full Sample: 1990–2015													
Mean	0.203	2.395	0.154	1.929	0.891	0.934	0.979	0.640	0.779	0.905	0.613	0.724	0.862
Sd	0.268	1.199	0.642	4.181	0.186	0.149	0.055	0.081	0.083	0.047	0.046	0.041	0.022
Median	0.201	2.569	0.187	2.661	0.827	0.882	0.964	0.623	0.763	0.897	0.601	0.715	0.858
Max	1.367	6.067	2.026	8.298	1.549	1.432	1.137	1.022	1.174	1.121	0.826	0.935	0.971
Min	-1.786	-1.978	-4.398	-16.721	0.641	0.729	0.903	0.551	0.686	0.847	0.539	0.661	0.826
AC1	0.415	0.926	0.245	0.973	0.987	0.988	0.992	0.987	0.989	0.992	0.969	0.979	0.985
AC36	0.812	5.079	1.727	6.980	14.358	14.743	15.879	15.460	15.535	14.906	11.495	11.640	11.960
First Sample: 1990–2000													
Mean	0.245	2.789	0.302	3.898	0.897	0.941	0.985	0.609	0.748	0.891	0.597	0.711	0.857
Sd	0.168	0.815	0.515	2.432	0.184	0.150	0.058	0.052	0.055	0.031	0.029	0.026	0.016
Median	0.214	2.782	0.335	4.163	0.869	0.922	0.979	0.587	0.723	0.879	0.594	0.707	0.859
Max	0.946	6.067	2.026	8.298	1.277	1.232	1.081	0.777	0.902	0.955	0.673	0.771	0.889
Min	-0.058	1.367	-1.173	-3.552	0.659	0.745	0.909	0.551	0.686	0.853	0.539	0.661	0.826
AC1	0.229	0.893	0.075	0.939	0.983	0.985	0.990	0.981	0.987	0.991	0.926	0.956	0.976
AC36	1.678	5.482	1.311	9.489	21.454	21.680	22.257	15.273	16.836	16.476	6.560	8.632	10.321
Second Sample: 2001–2015													
Mean	0.172	2.130	0.045	0.605	0.886	0.928	0.975	0.663	0.802	0.915	0.624	0.734	0.866
Sd	0.320	1.337	0.703	4.577	0.188	0.149	0.053	0.090	0.092	0.053	0.053	0.047	0.024
Median	0.181	2.074	0.120	2.066	0.812	0.869	0.952	0.643	0.779	0.900	0.608	0.720	0.858
Max	1.367	5.352	1.545	8.148	1.549	1.432	1.137	1.022	1.174	1.121	0.826	0.935	0.971
Min	-1.786	-1.978	-4.398	-16.721	0.641	0.729	0.903	0.554	0.691	0.847	0.555	0.681	0.840
AC1	0.441	0.928	0.264	0.972	0.977	0.978	0.981	0.986	0.988	0.991	0.974	0.981	0.985
AC36	0.339	4.195	0.237	1.933	8.074	8.195	8.573	11.493	11.287	11.995	9.321	9.460	10.816

Table 6.1: Descriptive Statistics

PUC-Rio - Certificação Digital N° 1421640/CA

The table reports for each model a number of different summary statistics across all the forecasting horizons, including the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present for square and absolute losses, the average p values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p values of the superior predictive ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.111	0.017	0.116	0.053	0.010	0.000	0.000	0.000
AR	0.834	0.838	0.749	0	0	0	0.276	0.047	0.541	0.235	0.056	0.008	0.000	0.000
UCSV	0.815	0.815	0.836	0	0	0	0.202	0.143	0.532	0.346	0.047	0.029	0.000	0.000
LASSO	0.780	0.786	0.724	0	0	0	0.310	0.349	0.741	0.540	0.160	0.192	0.729	0.729
adaLASSO	0.783	0.775	0.746	0	0	0	0.332	0.449	0.701	0.637	0.171	0.355	0.184	0.184
ElNet	0.780	0.791	0.720	0	0	2	0.320	0.320	0.734	0.530	0.178	0.174	0.889	0.889
adaElnet	0.781	0.775	0.738	0	0	2	0.367	0.477	0.722	0.650	0.204	0.386	0.048	0.048
Ridge	0.762	0.765	0.774	0	0	0	0.589	0.572	0.781	0.676	0.496	0.505	0.398	0.398
BVAR	0.799	0.812	0.780	0	0	0	0.278	0.355	0.713	0.502	0.135	0.145	0.036	0.036
Bagging	0.795	0.830	0.875	0	0	0	0.322	0.153	0.621	0.313	0.111	0.019	0.037	0.037
CSR	0.818	0.813	0.779	0	0	0	0.353	0.187	0.629	0.392	0.140	0.063	0.000	0.000
JMA	0.870	0.915	0.969	0	0	0	0.100	0.018	0.219	0.065	0.018	0.000	0.000	0.000
Factor	0.835	0.857	0.833	0	0	0	0.179	0.054	0.382	0.137	0.049	0.011	0.000	0.000
T. Factor	0.829	0.862	0.862	0	0	0	0.082	0.022	0.313	0.100	0.011	0.001	0.000	0.000
Boosting	0.828	0.878	0.940	0	0	0	0.234	0.102	0.635	0.278	0.027	0.004	0.000	0.000
RF	0.737	0.731	0.692	9	11	6	0.712	0.749	0.943	0.963	0.933	0.942	1	1
Mean	0.769	0.765	0.752	0	0	0	0.526	0.532	0.766	0.607	0.451	0.461	0.830	0.889
T. Mean	0.768	0.763	0.730	0	0	0	0.431	0.517	0.758	0.685	0.414	0.551	0.826	0.826
Median	0.769	0.764	0.731	0	0	0	0.415	0.502	0.751	0.674	0.367	0.509	0.804	0.889
RF/OLS	0.767	0.770	0.786	1	1	0	0.529	0.542	0.782	0.684	0.504	0.516	0.963	0.963
adaLASSO/RF	0.751	0.746	0.713	3	1	3	0.621	0.714	0.837	0.776	0.589	0.651	0.807	0.889

Table 6.2: Forecasting Results: Summary statistics for the out-of-sample period from 1990–2015

PUC-Rio - Certificação Digital N° 1421640/CA

The table reports for each model a number of different summary statistics across all the forecasting horizons, including the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present for square and absolute losses, the average p values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p values of the superior predictive ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.176	0.137	0.373	0.419	0.014	0.019	0.000	0.000
AR	0.871	0.924	0.709	0	0	2	0.436	0.211	0.701	0.519	0.236	0.097	0.011	0.011
UCSV	0.879	0.889	0.844	0	0	0	0.408	0.486	0.667	0.731	0.184	0.278	0.036	0.036
LASSO	0.879	0.947	0.740	0	0	0	0.187	0.122	0.617	0.489	0.041	0.021	0.000	0.000
adaLASSO	0.838	0.864	0.772	1	1	1	0.615	0.625	0.850	0.896	0.486	0.534	0.773	0.775
ElNet	0.890	0.971	0.725	0	0	1	0.162	0.107	0.538	0.355	0.073	0.028	0.007	0.007
adaElnet	0.843	0.874	0.753	0	2	0	0.605	0.624	0.833	0.883	0.479	0.532	0.732	0.794
Ridge	0.826	0.872	0.805	2	0	0	0.539	0.576	0.809	0.770	0.482	0.463	0.594	0.794
BVAR	0.939	1.017	0.795	0	1	1	0.397	0.311	0.562	0.368	0.130	0.105	0.183	0.183
Bagging	0.907	0.943	0.949	0	0	0	0.287	0.415	0.501	0.614	0.077	0.162	0.005	0.011
CSR	0.849	0.900	0.808	2	0	0	0.396	0.333	0.716	0.605	0.306	0.218	0.794	0.794
JMA	1.055	1.092	1.105	0	0	0	0.093	0.077	0.202	0.184	0.003	0.002	0.000	0.000
Factor	0.970	1.057	0.902	0	0	0	0.075	0.087	0.284	0.185	0.009	0.016	0.000	0.000
T. Factor	1.007	1.121	0.873	0	0	0	0.024	0.009	0.163	0.077	0.001	0.000	0.000	0.000
Boosting	1.004	1.111	1.164	0	0	0	0.328	0.230	0.469	0.363	0.035	0.020	0.000	0.000
RF	0.793	0.821	0.655	5	6	5	0.915	0.857	0.976	0.927	0.981	0.927	1	1
Mean	0.820	0.862	0.771	0	0	0	0.699	0.625	0.873	0.844	0.694	0.579	0.433	0.479
T. Mean	0.825	0.870	0.745	1	0	0	0.613	0.445	0.868	0.846	0.579	0.401	0.775	0.775
Median	0.827	0.873	0.743	0	0	0	0.566	0.447	0.855	0.828	0.479	0.401	0.741	0.794
RF/OLS	0.822	0.867	0.835	0	1	0	0.691	0.623	0.880	0.857	0.653	0.510	0.479	0.479
adaLASSO/RF	0.812	0.846	0.703	2	2	3	0.780	0.726	0.878	0.876	0.725	0.646	0.127	0.794

Table 6.3: Forecasting Results: Summary statistics for the out-of-sample period from 1990–2000

PUC-Rio - Certificação Digital N° 1421640/CA

The table reports for each model a number of different summary statistics across all the forecasting horizons, including the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present for square and absolute losses, the average p values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p values of the superior predictive ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.137	0.007	0.137	0.037	0.014	0.000	0.000	0.000
AR	0.828	0.807	0.794	0	0	0	0.308	0.107	0.514	0.228	0.067	0.022	0.000	0.000
UCSV	0.805	0.788	0.799	0	0	0	0.272	0.226	0.596	0.395	0.087	0.063	0.000	0.000
LASSO	0.762	0.728	0.670	0	0	2	0.483	0.702	0.876	0.798	0.347	0.630	0.972	0.992
adaLASSO	0.773	0.743	0.704	0	0	0	0.424	0.474	0.784	0.669	0.205	0.300	0.181	0.186
ElNet	0.760	0.726	0.663	0	0	3	0.513	0.726	0.875	0.813	0.423	0.711	0.974	0.992
adaElnet	0.771	0.739	0.699	0	0	0	0.460	0.534	0.838	0.716	0.264	0.355	0.132	0.132
Ridge	0.751	0.727	0.682	0	0	1	0.672	0.755	0.844	0.803	0.626	0.694	0.206	0.206
BVAR	0.773	0.737	0.688	0	0	1	0.475	0.690	0.861	0.822	0.290	0.488	0.766	0.766
Bagging	0.775	0.789	0.785	0	0	0	0.560	0.247	0.760	0.315	0.234	0.051	0.186	0.186
CSR	0.813	0.781	0.714	0	0	0	0.418	0.306	0.624	0.468	0.164	0.127	0.000	0.000
JMA	0.837	0.852	0.861	0	0	0	0.291	0.077	0.446	0.130	0.077	0.011	0.002	0.002
Factor	0.810	0.783	0.742	0	0	0	0.377	0.268	0.597	0.440	0.131	0.087	0.001	0.001
T. Factor	0.795	0.767	0.734	0	0	2	0.381	0.315	0.630	0.489	0.175	0.152	0.012	0.012
Boosting	0.795	0.792	0.763	0	0	1	0.467	0.399	0.821	0.582	0.147	0.087	0.002	0.002
RF	0.728	0.698	0.696	9	8	3	0.837	0.862	0.952	0.962	0.922	0.909	1	1
Mean	0.761	0.729	0.694	0	1	0	0.560	0.569	0.860	0.795	0.506	0.578	0.976	0.986
T. Mean	0.759	0.724	0.680	0	0	0	0.548	0.724	0.874	0.833	0.515	0.810	0.983	0.983
Median	0.759	0.725	0.683	0	0	0	0.539	0.650	0.870	0.824	0.474	0.719	0.986	0.986
RF/OLS	0.758	0.735	0.684	1	1	0	0.640	0.745	0.894	0.789	0.535	0.623	0.971	0.992
adaLASSO/RF	0.742	0.711	0.688	3	3	0	0.684	0.728	0.913	0.888	0.614	0.655	0.992	0.992

Table 6.4: Forecasting Results: Summary statistics for the out-of-sample period from 2001–2015

The table reports the root mean squared error (RMSE), mean absolute error (MAE) and median absolute deviation from the median (MAD) ratios with respect to the random walk model for the full out-of-sample period (1990–2015). The statistics for the best-performing model are highlighted in bold.

Panel (a): RMSE Ratio													
Forecasting Horizon													
Model	1	2	3	4	5	6	7	8	9	10	11	12	Acc.
AR	0.902	0.809	0.790	0.805	0.786	0.791	0.783	0.764	0.779	0.824	0.837	0.753	1.218
UCSV	0.954	0.816	0.797	0.813	0.783	0.777	0.784	0.776	0.770	0.804	0.832	0.781	0.908
RF	0.844	0.731	0.706	0.738	0.711	0.715	0.718	0.712	0.722	0.763	0.773	0.685	0.766

Panel (b): MAE Ratio													
Forecasting Horizon													
Model	1	2	3	4	5	6	7	8	9	10	11	12	Acc.
AR	0.874	0.791	0.782	0.805	0.802	0.806	0.777	0.760	0.807	0.847	0.861	0.764	1.220
UCSV	0.911	0.817	0.786	0.803	0.801	0.795	0.796	0.787	0.784	0.799	0.851	0.777	0.894
RF	0.811	0.721	0.711	0.749	0.727	0.728	0.699	0.681	0.717	0.747	0.767	0.668	0.774

Panel (c): MAD Ratio													
Forecasting Horizon													
Model	1	2	3	4	5	6	7	8	9	10	11	12	Acc.
AR	0.738	0.703	0.815	0.822	0.828	0.755	0.664	0.685	0.767	0.697	0.769	0.600	0.889
UCSV	0.876	0.770	0.832	0.906	0.878	0.790	0.761	0.835	0.857	0.829	0.884	0.777	0.876
RF	0.698	0.633	0.772	0.841	0.750	0.728	0.653	0.639	0.728	0.685	0.706	0.575	0.587

Table 6.5: Forecasting Results: RMSE, MAE and MAD Ratios (1990–2015)

The table reports the frequency with which each model achieved the best (worst) performance statistics over a rolling window period of four years (48 observations).

Panel (a): Lowest Rolling RMSE													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.038	0.000	0.083
AR	0.083	0.049	0.000	0.158	0.011	0.011	0.098	0.177	0.117	0.128	0.113	0.000	0.000
UCSV	0.023	0.049	0.211	0.098	0.181	0.109	0.004	0.030	0.192	0.109	0.094	0.000	0.236
RF	0.894	0.902	0.789	0.743	0.808	0.879	0.898	0.792	0.691	0.762	0.755	1.000	0.681

Panel (b): Lowest Rolling MAE													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.094	0.000	0.039
AR	0.166	0.034	0.000	0.049	0.000	0.023	0.098	0.136	0.132	0.064	0.049	0.023	0.000
UCSV	0.151	0.177	0.257	0.226	0.151	0.155	0.000	0.072	0.242	0.226	0.094	0.023	0.201
RF	0.683	0.789	0.743	0.725	0.849	0.823	0.902	0.792	0.626	0.709	0.762	0.955	0.760

Panel (c): Lowest Rolling MAD													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.136	0.000	0.034	0.000	0.091	0.045	0.004	0.000	0.034	0.053	0.026	0.045	0.020
AR	0.234	0.162	0.234	0.230	0.121	0.234	0.264	0.321	0.109	0.147	0.423	0.234	0.039
UCSV	0.038	0.192	0.268	0.328	0.094	0.117	0.034	0.008	0.219	0.087	0.034	0.019	0.051
RF	0.592	0.645	0.464	0.442	0.694	0.604	0.698	0.672	0.638	0.713	0.517	0.702	0.890

Panel (d): Highest Rolling RMSE													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.823	0.985	0.936	0.992	0.977	0.996	1.000	1.000	0.864	0.804	0.713	0.849	0.000
AR	0.000	0.004	0.057	0.008	0.023	0.004	0.000	0.000	0.136	0.189	0.287	0.151	0.969
UCSV	0.177	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.031
RF	0.000	0.011	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000

Panel (e): Highest Rolling MAE													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.898	0.943	0.864	1.000	0.992	0.966	1.000	0.943	0.819	0.728	0.687	0.770	0.031
AR	0.083	0.034	0.128	0.000	0.008	0.034	0.000	0.057	0.181	0.272	0.283	0.230	0.862
UCSV	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.106
RF	0.000	0.023	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.030	0.000	0.000

Panel (f): Highest Rolling MAD													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.683	0.940	0.853	0.804	0.796	0.657	0.811	0.943	0.774	0.804	0.789	0.921	0.315
AR	0.053	0.026	0.117	0.098	0.102	0.068	0.042	0.019	0.098	0.034	0.019	0.008	0.512
UCSV	0.215	0.034	0.026	0.098	0.042	0.260	0.140	0.038	0.106	0.143	0.192	0.072	0.173
RF	0.049	0.000	0.004	0.000	0.060	0.015	0.008	0.000	0.023	0.019	0.000	0.000	0.000

Table 6.6: Forecasting Results: Ranking of Models (1990–2015)

The table reports the p -values of the unconditional Giacomini-White test for superior predictive ability between the random forest models and each of the benchmark models. The test is based on the full out-of-sample period. Panel (a) presents the results for squared errors, while panel (b) shows the results for absolute errors.

Panel (a): Giacomini-White Test (Sq. Errors)													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.003	0.000	0.000	0.001	0.006	0.012	0.010	0.003	0.003	0.027	0.024	0.001	0.049
AR	0.002	0.010	0.023	0.045	0.024	0.024	0.056	0.075	0.047	0.062	0.008	0.000	0.021
UCSV	0.003	0.003	0.013	0.055	0.055	0.024	0.001	0.000	0.003	0.038	0.002	0.000	0.072

Panel (b): Giacomini-White Test (Abs. Errors)													
Model	Forecasting Horizon												Acc.
	1	2	3	4	5	6	7	8	9	10	11	12	
RW	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023
AR	0.000	0.000	0.001	0.012	0.002	0.005	0.009	0.017	0.007	0.004	0.000	0.000	0.000
UCSV	0.000	0.000	0.010	0.029	0.003	0.007	0.000	0.000	0.007	0.054	0.008	0.000	0.078

Table 6.7: Forecasting Results: Superior Predictive Ability Test (1990–2015)

The figure displays the root mean squared errors (RMSE) computed over rolling windows of 48 observations. Panel (a) displays the results for one-month-ahead forecasts ($h = 1$), panel (b) displays the results for six-months-ahead forecasts ($h = 6$), panel (c) displays the results for twelve-months-ahead forecasts ($h = 12$), and finally, Panel (d) displays the results for the accumulated twelve month forecasts.

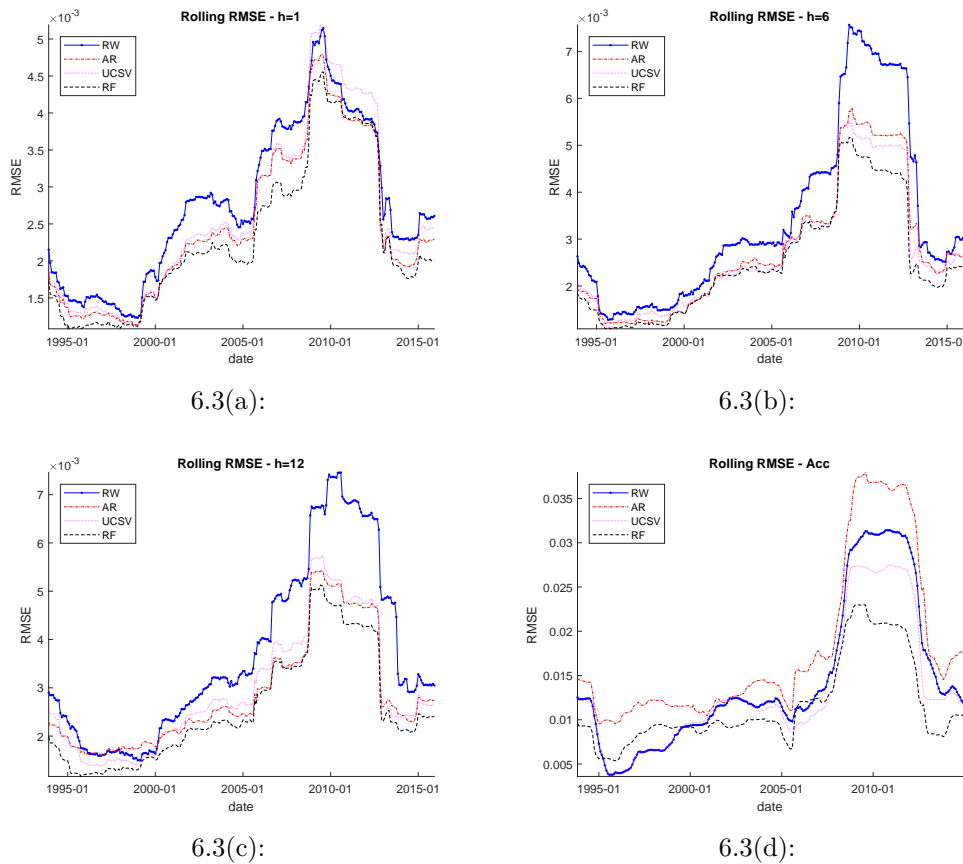
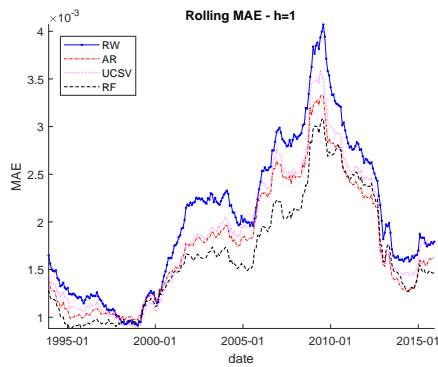
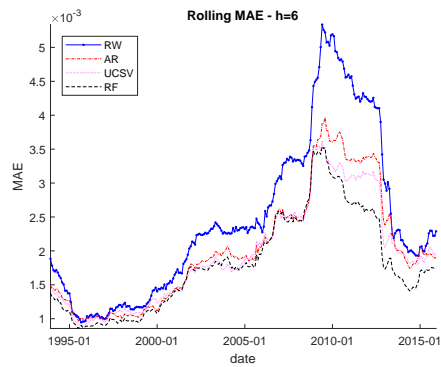


Figure 6.2: Rolling RMSE.

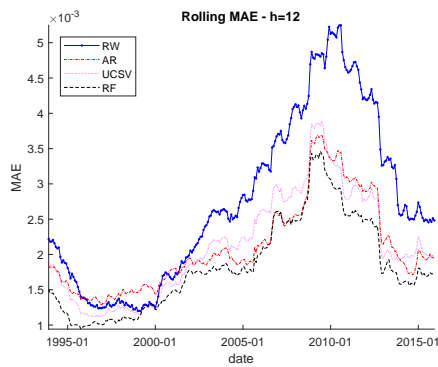
The figure displays the mean absolute errors (MAE) computed over rolling windows of 48 observations. Panel (a) displays the results for one-month-ahead forecasts ($h = 1$), panel (b) displays the results for six-months-ahead forecasts ($h = 6$), panel (c) displays the results for twelve-months-ahead forecasts ($h = 12$), and finally, panel (d) displays the results for the accumulated twelve month forecasts.



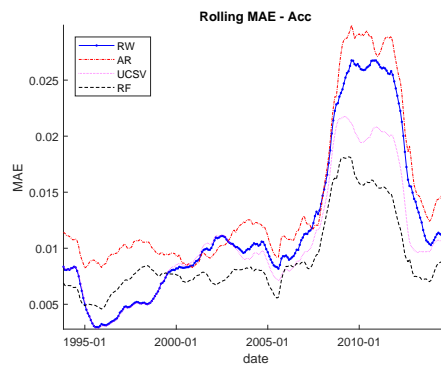
6.4(a):



6.4(b):



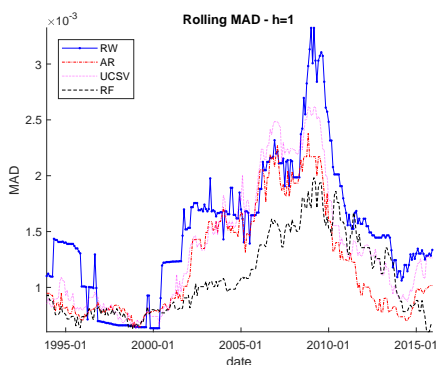
6.4(c):



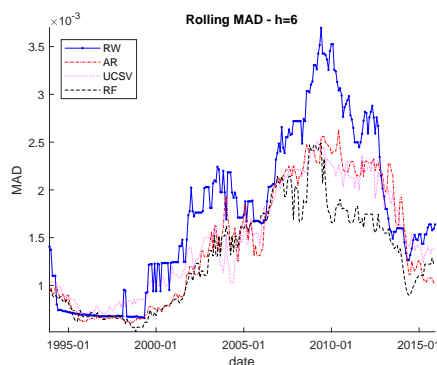
6.4(d):

Figure 6.3: Rolling MAE.

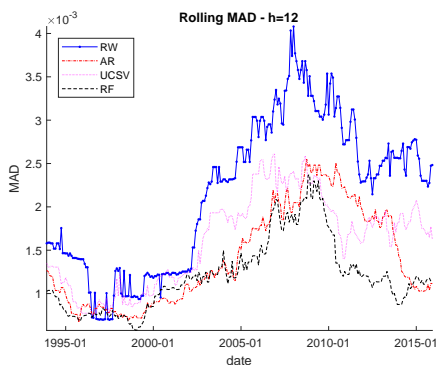
The figure displays the mean absolute deviation from the median (MAD) computed over rolling windows of 48 observations. Panel (a) displays the results for one-month-ahead forecasts ($h = 1$), panel (b) displays the results for six-months-ahead forecasts ($h = 6$), Panel (c) displays the results for twelve-months-ahead forecasts ($h = 12$), and finally, panel (d) displays the results for the accumulated twelve month forecasts.



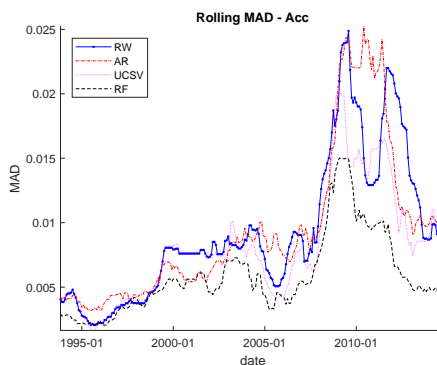
6.5(a):



6.5(b):



6.5(c):



6.5(d):

Figure 6.4: Rolling MAD.

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the random walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2000 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum t -statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table report how many models were included in the MCS for square and absolute losses.

Consumer Price Index 1990-2015														
RMSE/(MAE)	Forecasting Horizon												Acc.	RMSE count (MAE count)
	1	2	3	4	5	6	7	8	9	10	11	12		
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1
AR	0.90 (0.87)	0.81 (0.79)	0.79 (0.78)	0.81 (0.81)	0.79 (0.80)	0.79 (0.81)	0.78 (0.78)	0.76 (0.76)	0.78 (0.81)	0.82 (0.85)	0.84 (0.86)	0.75 (0.76)	1.22 (1.22)	9 (0)
UCSV	0.95 (0.91)	0.82 (0.82)	0.80 (0.79)	0.81 (0.80)	0.78 (0.80)	0.78 (0.79)	0.78 (0.80)	0.78 (0.79)	0.77 (0.78)	0.80 (0.80)	0.83 (0.85)	0.78 (0.78)	0.91 (0.89)	8 (3)
LASSO	0.83 (0.82)	0.75 (0.74)	0.73 (0.73)	0.76 (0.77)	0.74 (0.77)	0.75 (0.75)	0.75 (0.74)	0.73 (0.71)	0.75 (0.76)	0.80 (0.81)	0.82 (0.84)	0.73 (0.74)	0.98 (1.04)	11 (9)
adaLASSO	0.84 (0.81)	0.76 (0.75)	0.74 (0.72)	0.77 (0.77)	0.75 (0.75)	0.75 (0.74)	0.76 (0.73)	0.75 (0.71)	0.76 (0.75)	0.80 (0.79)	0.83 (0.84)	0.72 (0.73)	0.96 (0.96)	11 (11)
ElNet	0.83 (0.82)	0.75 (0.74)	0.73 (0.73)	0.76 (0.78)	0.75 (0.78)	0.74 (0.76)	0.75 (0.75)	0.74 (0.71)	0.76 (0.77)	0.81 (0.81)	0.82 (0.85)	0.73 (0.75)	0.98 (1.05)	11 (9)
adaElnet	0.84 (0.82)	0.75 (0.74)	0.73 (0.72)	0.77 (0.76)	0.75 (0.75)	0.75 (0.74)	0.75 (0.73)	0.74 (0.71)	0.76 (0.75)	0.80 (0.79)	0.81 (0.83)	0.73 (0.75)	0.96 (0.97)	11 (11)
Ridge	0.85 (0.83)	0.73 (0.72)	0.72 (0.72)	0.75 (0.77)	0.74 (0.76)	0.75 (0.76)	0.75 (0.73)	0.73 (0.71)	0.74 (0.74)	0.77 (0.77)	0.78 (0.79)	0.70 (0.71)	0.89 (0.93)	13 (13)
BVAR	0.86 (0.87)	0.76 (0.73)	0.75 (0.75)	0.77 (0.79)	0.74 (0.78)	0.76 (0.78)	0.77 (0.76)	0.76 (0.76)	0.77 (0.76)	0.82 (0.81)	0.83 (0.85)	0.74 (0.76)	1.07 (1.09)	11 (8)
Bagging	0.83 (0.84)	0.76 (0.78)	0.76 (0.79)	0.80 (0.87)	0.78 (0.86)	0.79 (0.85)	0.83 (0.83)	0.81 (0.80)	0.78 (0.80)	0.82 (0.84)	0.83 (0.86)	0.74 (0.78)	0.82 (0.88)	11 (4)
CSR	0.85 (0.84)	0.77 (0.76)	0.76 (0.75)	0.79 (0.79)	0.77 (0.79)	0.79 (0.79)	0.79 (0.76)	0.79 (0.74)	0.79 (0.79)	0.83 (0.83)	0.84 (0.84)	0.76 (0.77)	1.13 (1.11)	11 (3)
JMA	0.99 (0.99)	0.82 (0.85)	0.84 (0.89)	0.85 (0.94)	0.84 (0.96)	0.81 (0.90)	0.91 (0.91)	0.86 (0.87)	0.84 (0.93)	0.95 (0.96)	0.92 (0.96)	0.80 (0.83)	0.88 (0.91)	2 (1)
Factor	0.87 (0.88)	0.78 (0.80)	0.78 (0.80)	0.79 (0.82)	0.79 (0.82)	0.78 (0.80)	0.78 (0.78)	0.80 (0.80)	0.81 (0.87)	0.82 (0.87)	0.84 (0.87)	0.78 (0.82)	1.17 (1.21)	4 (0)
T. Factor	0.88 (0.87)	0.79 (0.82)	0.78 (0.81)	0.80 (0.84)	0.77 (0.83)	0.79 (0.84)	0.79 (0.80)	0.80 (0.80)	0.80 (0.84)	0.80 (0.87)	0.82 (0.86)	0.83 (0.80)	1.17 (1.23)	3 (0)
Boosting	0.95 (0.96)	0.77 (0.80)	0.76 (0.81)	0.78 (0.85)	0.77 (0.84)	0.79 (0.86)	0.79 (0.84)	0.78 (0.82)	0.79 (0.82)	0.83 (0.85)	0.84 (0.86)	0.74 (0.75)	1.17 (1.32)	10 (1)
RF	0.84 (0.81)	0.73 (0.72)	0.71 (0.71)	0.74 (0.75)	0.71 (0.73)	0.72 (0.73)	0.72 (0.70)	0.71 (0.68)	0.72 (0.72)	0.76 (0.75)	0.77 (0.77)	0.68 (0.67)	0.77 (0.77)	13 (13)
Mean	0.83 (0.81)	0.75 (0.74)	0.73 (0.73)	0.76 (0.76)	0.74 (0.76)	0.74 (0.75)	0.75 (0.73)	0.74 (0.71)	0.75 (0.75)	0.77 (0.76)	0.78 (0.78)	0.71 (0.70)	0.95 (0.97)	12 (12)
T.Mean	0.84 (0.81)	0.74 (0.74)	0.73 (0.72)	0.75 (0.76)	0.74 (0.75)	0.74 (0.74)	0.75 (0.72)	0.73 (0.70)	0.74 (0.74)	0.78 (0.77)	0.79 (0.79)	0.71 (0.70)	0.95 (0.96)	12 (12)
Median	0.84 (0.81)	0.75 (0.74)	0.72 (0.72)	0.76 (0.76)	0.74 (0.76)	0.74 (0.74)	0.75 (0.73)	0.73 (0.70)	0.74 (0.74)	0.78 (0.77)	0.79 (0.79)	0.71 (0.71)	0.94 (0.97)	12 (12)
RF/OLS	0.81 (0.79)	0.73 (0.73)	0.72 (0.72)	0.75 (0.76)	0.74 (0.76)	0.75 (0.76)	0.75 (0.73)	0.74 (0.72)	0.74 (0.75)	0.78 (0.78)	0.79 (0.81)	0.71 (0.72)	0.94 (0.97)	13 (13)
adaLASSO/RF	0.85 (0.82)	0.76 (0.73)	0.72 (0.72)	0.73 (0.74)	0.73 (0.74)	0.72 (0.73)	0.72 (0.71)	0.71 (0.68)	0.72 (0.72)	0.79 (0.79)	0.82 (0.82)	0.70 (0.68)	0.80 (0.82)	13 (13)
RMSE count	14	15	16	17	19	19	17	15	17	18	19	8	8	
MAE count	(12)	(11)	(12)	(11)	(10)	(12)	(13)	(11)	(10)	(15)	(16)	(7)	(9)	

PUC-Rio - Certificação Digital Nº 1421640/CA

Table 6.8: Forecasting Errors for the CPI from 1990 to 2015

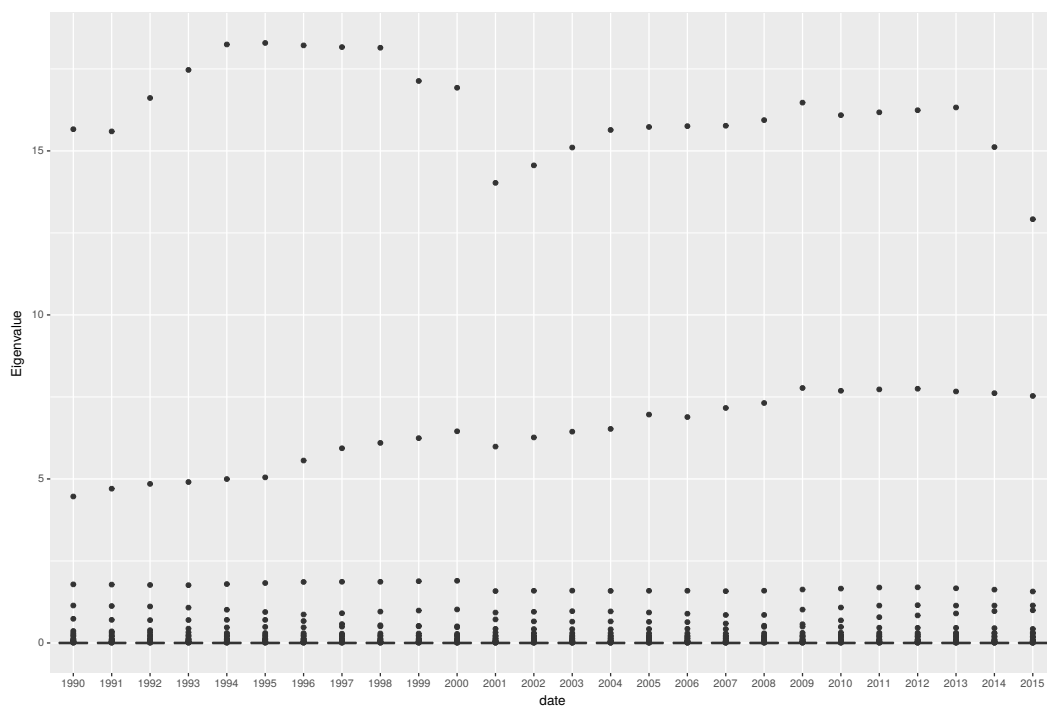


Figure 6.5: Eigenvalues of the matrix of contemporaneous regressor.

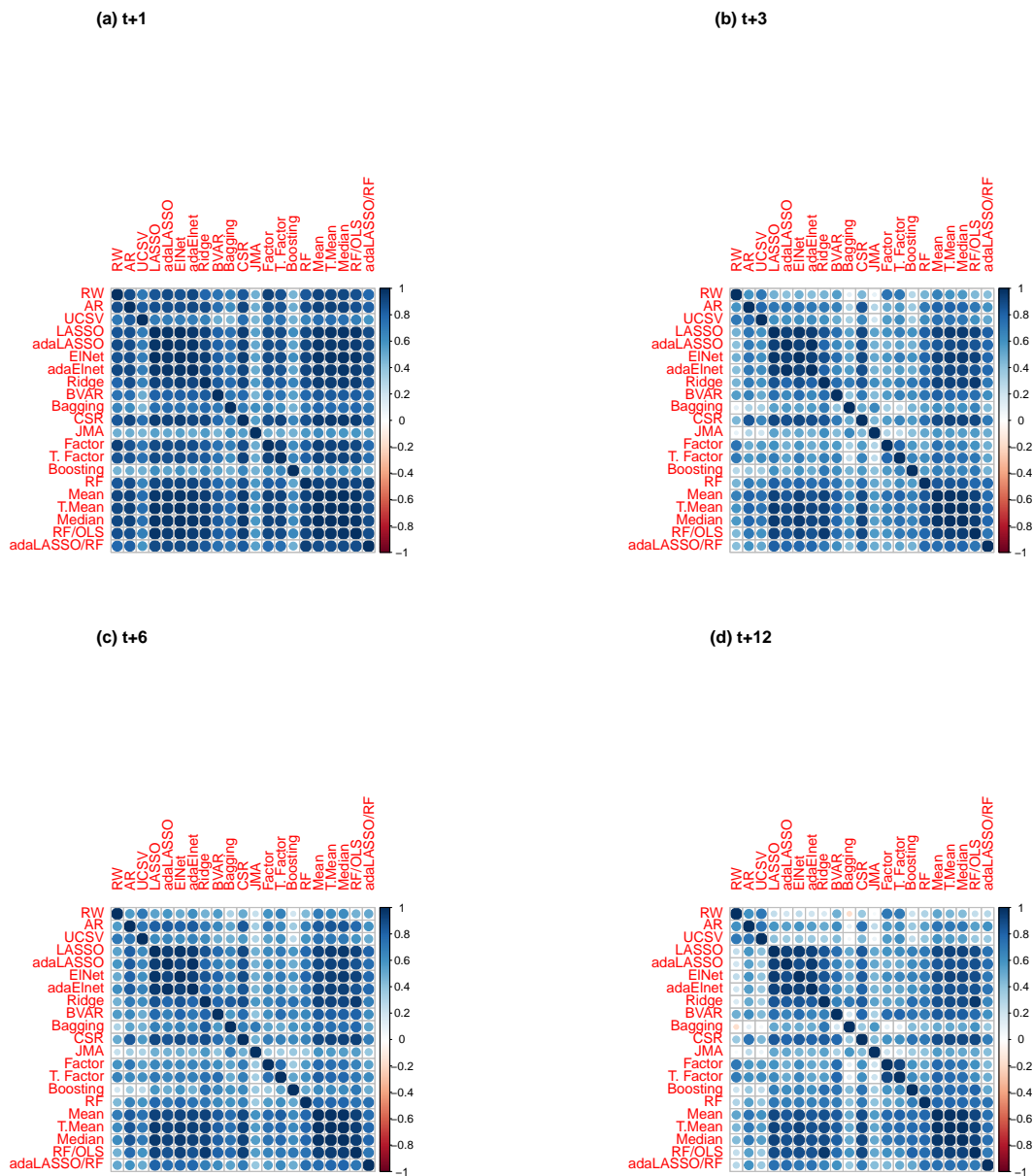


Figure 6.6: Correlation of the Forecasts for the CPI from 1990 to 2015

The table shows the root mean squared error (RMSE), and between parenthesis, the mean absolute errors (MAE) for all models relative to the random walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2000 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss functions. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table report how many models were included in the MCS for square and absolute losses.

Consumer Price Index 1990-2000														
RMSE/(MAE)	Forecasting Horizon												Acc.	RMSE count (MAE count)
	1	2	3	4	5	6	7	8	9	10	11	12		
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	3 (4)
AR	0.84 (0.88)	0.82 (0.83)	0.88 (0.92)	0.82 (0.83)	0.78 (0.81)	0.79 (0.84)	0.79 (0.84)	0.80 (0.80)	0.87 (0.94)	0.89 (0.98)	0.95 (1.04)	0.85 (0.94)	1.24 (1.38)	10 (6)
UCSV	0.86 (0.88)	0.84 (0.85)	0.87 (0.88)	0.87 (0.87)	0.85 (0.86)	0.85 (0.86)	0.86 (0.87)	0.85 (0.84)	0.86 (0.88)	0.89 (0.91)	0.94 (0.96)	0.88 (0.89)	1.00 (1.02)	8 (11)
LASSO	0.83 (0.88)	0.82 (0.84)	0.88 (0.92)	0.83 (0.84)	0.79 (0.83)	0.78 (0.84)	0.80 (0.88)	0.81 (0.83)	0.88 (0.96)	0.92 (1.02)	0.97 (1.08)	0.85 (0.96)	1.24 (1.41)	9 (5)
adaLASSO	0.81 (0.84)	0.82 (0.82)	0.87 (0.86)	0.83 (0.80)	0.75 (0.73)	0.75 (0.77)	0.77 (0.81)	0.77 (0.77)	0.85 (0.90)	0.87 (0.92)	0.92 (1.00)	0.82 (0.89)	1.03 (1.08)	13 (13)
ElNet	0.81 (0.86)	0.81 (0.84)	0.88 (0.92)	0.83 (0.86)	0.80 (0.86)	0.80 (0.85)	0.79 (0.92)	0.82 (0.83)	0.81 (1.02)	0.92 (1.02)	0.92 (1.14)	1.00 (1.02)	0.89 (1.47)	7 (4)
adaElnet	0.81 (0.85)	0.82 (0.83)	0.86 (0.86)	0.80 (0.77)	0.74 (0.73)	0.75 (0.78)	0.77 (0.81)	0.78 (0.78)	0.87 (0.92)	0.87 (0.93)	0.92 (1.00)	0.87 (0.95)	1.06 (1.13)	12 (12)
Ridge	0.79 (0.83)	0.77 (0.78)	0.86 (0.90)	0.80 (0.81)	0.76 (0.78)	0.80 (0.84)	0.80 (0.85)	0.80 (0.79)	0.86 (0.90)	0.85 (0.92)	0.88 (0.96)	0.76 (0.82)	0.99 (1.15)	12 (12)
BVAR	0.97 (1.00)	0.80 (0.77)	0.92 (0.96)	0.83 (0.88)	0.77 (0.84)	0.84 (0.93)	0.87 (0.98)	0.90 (0.95)	1.00 (1.12)	0.98 (1.10)	1.02 (1.16)	0.88 (1.01)	1.43 (1.56)	6 (1)
Bagging	0.85 (0.86)	0.86 (0.87)	1.02 (1.04)	0.92 (0.95)	0.92 (0.93)	0.90 (0.95)	0.91 (0.92)	0.90 (0.82)	0.86 (0.94)	0.91 (0.95)	0.91 (0.99)	0.93 (0.87)	1.02 (1.15)	8 (8)
CSR	0.83 (0.89)	0.85 (0.89)	0.89 (0.92)	0.81 (0.82)	0.77 (0.79)	0.76 (0.81)	0.76 (0.82)	0.76 (0.76)	0.85 (0.91)	0.88 (0.95)	0.91 (0.97)	0.81 (0.89)	1.11 (1.25)	10 (8)
JMA	0.94 (1.00)	1.01 (1.02)	1.17 (1.19)	0.99 (1.01)	1.03 (1.07)	1.01 (1.05)	1.06 (1.06)	1.03 (1.01)	1.21 (1.29)	1.13 (1.19)	1.13 (1.20)	0.93 (0.98)	1.00 (1.08)	1 (2)
Factor	0.87 (0.96)	0.85 (0.92)	0.98 (1.05)	0.90 (0.97)	0.89 (0.92)	0.86 (0.90)	0.84 (0.88)	0.90 (0.91)	1.02 (1.14)	0.97 (1.09)	1.04 (1.15)	0.98 (1.14)	1.51 (1.72)	1 (1)
T. Factor	0.87 (0.93)	0.91 (0.98)	1.01 (1.13)	0.98 (1.07)	0.92 (1.02)	0.94 (1.05)	0.91 (0.94)	0.91 (0.93)	1.04 (1.16)	1.02 (1.18)	1.02 (1.15)	0.95 (1.10)	1.62 (1.91)	0 (0)
Boosting	0.96 (1.09)	0.90 (0.98)	1.05 (1.16)	0.91 (0.98)	0.88 (0.97)	0.95 (1.06)	0.95 (1.06)	0.97 (1.03)	1.02 (1.12)	0.96 (1.06)	0.97 (1.07)	0.81 (0.89)	1.66 (1.92)	5 (3)
RF	0.79 (0.82)	0.78 (0.78)	0.85 (0.88)	0.77 (0.77)	0.73 (0.76)	0.76 (0.79)	0.76 (0.78)	0.77 (0.75)	0.82 (0.86)	0.82 (0.86)	0.85 (0.86)	0.72 (0.76)	0.87 (0.94)	13 (12)
Mean	0.80 (0.83)	0.79 (0.81)	0.85 (0.87)	0.79 (0.80)	0.76 (0.79)	0.77 (0.81)	0.77 (0.81)	0.77 (0.76)	0.84 (0.90)	0.84 (0.91)	0.87 (0.94)	0.78 (0.85)	1.02 (1.11)	13 (12)
T.Mean	0.80 (0.84)	0.80 (0.82)	0.85 (0.87)	0.79 (0.79)	0.75 (0.77)	0.76 (0.80)	0.77 (0.81)	0.77 (0.78)	0.85 (0.91)	0.84 (0.91)	0.89 (0.97)	0.79 (0.87)	1.04 (1.15)	13 (12)
Median	0.80 (0.84)	0.80 (0.83)	0.85 (0.88)	0.79 (0.79)	0.75 (0.78)	0.76 (0.80)	0.77 (0.82)	0.77 (0.77)	0.85 (0.91)	0.85 (0.91)	0.89 (0.97)	0.79 (0.87)	1.05 (1.16)	13 (12)
RF/OLS	0.80 (0.82)	0.80 (0.82)	0.86 (0.89)	0.78 (0.79)	0.74 (0.76)	0.77 (0.81)	0.77 (0.82)	0.78 (0.78)	0.85 (0.90)	0.85 (0.92)	0.88 (0.96)	0.76 (0.83)	1.01 (1.14)	13 (12)
adaLASSO/RF	0.79 (0.84)	0.81 (0.81)	0.91 (0.94)	0.77 (0.77)	0.72 (0.73)	0.77 (0.81)	0.77 (0.81)	0.77 (0.77)	0.82 (0.86)	0.89 (0.94)	0.90 (0.99)	0.72 (0.76)	0.89 (0.95)	12 (12)
RMSE count	12	18	14	16	11	14	10	15	14	18	16	11	13	
MAE count	(12)	(15)	(9)	(10)	(3)	(14)	(13)	(15)	(15)	(16)	(14)	(13)	(13)	

Table 6.9: Forecasting Errors for the CPI from 1990 to 2000

The table shows the root mean squared error (RMSE), and between parenthesis, the mean absolute errors (MAE) for all models relative to the random walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2000 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss functions. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table report how many models were included in the MCS for square and absolute losses.

Consumer Price Index 2001–2015														
RMSE/(MAE)	Forecasting Horizon												Acc.	RMSE count (MAE count)
	1	2	3	4	5	6	7	8	9	10	11	12		
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1
AR	0.92 (0.87)	0.81 (0.78)	0.78 (0.74)	0.80 (0.79)	0.79 (0.80)	0.79 (0.80)	0.78 (0.75)	0.76 (0.75)	0.77 (0.76)	0.81 (0.80)	0.82 (0.79)	0.73 (0.70)	1.21 (1.17)	7 (0)
UCSV	0.98 (0.93)	0.81 (0.81)	0.79 (0.76)	0.80 (0.77)	0.79 (0.78)	0.77 (0.77)	0.77 (0.77)	0.76 (0.77)	0.76 (0.75)	0.79 (0.76)	0.81 (0.81)	0.76 (0.73)	0.89 (0.85)	9 (5)
LASSO	0.84 (0.79)	0.74 (0.71)	0.71 (0.67)	0.75 (0.75)	0.74 (0.74)	0.74 (0.72)	0.75 (0.69)	0.72 (0.67)	0.74 (0.69)	0.78 (0.73)	0.79 (0.75)	0.70 (0.65)	0.91 (0.91)	13 (12)
adaLASSO	0.84 (0.80)	0.75 (0.72)	0.72 (0.68)	0.76 (0.76)	0.75 (0.76)	0.75 (0.73)	0.76 (0.71)	0.74 (0.69)	0.75 (0.71)	0.79 (0.75)	0.81 (0.78)	0.70 (0.67)	0.93 (0.92)	13 (11)
ElNet	0.84 (0.80)	0.74 (0.70)	0.71 (0.67)	0.74 (0.74)	0.73 (0.74)	0.74 (0.72)	0.74 (0.69)	0.73 (0.67)	0.73 (0.69)	0.79 (0.73)	0.79 (0.74)	0.70 (0.64)	0.91 (0.92)	13 (12)
adaElNet	0.85 (0.81)	0.74 (0.71)	0.72 (0.68)	0.76 (0.76)	0.75 (0.76)	0.75 (0.73)	0.75 (0.70)	0.74 (0.68)	0.74 (0.70)	0.79 (0.74)	0.80 (0.76)	0.70 (0.67)	0.93 (0.92)	12 (11)
Ridge	0.86 (0.83)	0.72 (0.70)	0.70 (0.67)	0.75 (0.75)	0.73 (0.74)	0.74 (0.74)	0.74 (0.69)	0.72 (0.68)	0.72 (0.69)	0.76 (0.71)	0.77 (0.73)	0.69 (0.67)	0.86 (0.86)	12 (12)
BVAR	0.83 (0.81)	0.75 (0.72)	0.72 (0.68)	0.75 (0.75)	0.74 (0.75)	0.74 (0.73)	0.75 (0.69)	0.74 (0.69)	0.74 (0.70)	0.79 (0.73)	0.79 (0.74)	0.72 (0.67)	0.99 (0.93)	13 (12)
Bagging	0.82 (0.84)	0.74 (0.74)	0.72 (0.71)	0.78 (0.83)	0.76 (0.84)	0.77 (0.82)	0.81 (0.80)	0.80 (0.79)	0.76 (0.76)	0.80 (0.80)	0.81 (0.82)	0.73 (0.74)	0.77 (0.80)	11 (4)
CSR	0.86 (0.82)	0.75 (0.71)	0.74 (0.69)	0.78 (0.78)	0.78 (0.79)	0.79 (0.78)	0.80 (0.74)	0.77 (0.73)	0.78 (0.75)	0.82 (0.78)	0.83 (0.80)	0.75 (0.73)	1.12 (1.07)	10 (5)
JMA	1.00 (0.99)	0.78 (0.78)	0.79 (0.79)	0.83 (0.92)	0.80 (0.91)	0.77 (0.84)	0.89 (0.85)	0.83 (0.82)	0.79 (0.81)	0.91 (0.88)	0.88 (0.87)	0.77 (0.78)	0.84 (0.85)	5 (1)
Factor	0.87 (0.84)	0.77 (0.76)	0.75 (0.72)	0.77 (0.77)	0.76 (0.78)	0.77 (0.76)	0.79 (0.74)	0.80 (0.76)	0.79 (0.78)	0.81 (0.79)	0.81 (0.77)	0.74 (0.69)	1.10 (1.04)	9 (5)
T. Factor	0.88 (0.85)	0.76 (0.75)	0.74 (0.71)	0.76 (0.74)	0.74 (0.75)	0.76 (0.76)	0.78 (0.75)	0.78 (0.75)	0.78 (0.74)	0.78 (0.76)	0.80 (0.76)	0.74 (0.69)	1.05 (1.00)	9 (6)
Boosting	0.95 (0.91)	0.75 (0.72)	0.72 (0.70)	0.76 (0.79)	0.74 (0.78)	0.76 (0.79)	0.77 (0.76)	0.75 (0.75)	0.76 (0.76)	0.81 (0.79)	0.81 (0.79)	0.73 (0.69)	1.03 (1.13)	12 (8)
RF	0.86 (0.81)	0.72 (0.70)	0.69 (0.66)	0.73 (0.74)	0.71 (0.71)	0.71 (0.70)	0.71 (0.67)	0.70 (0.66)	0.71 (0.67)	0.75 (0.70)	0.76 (0.72)	0.68 (0.63)	0.74 (0.72)	13 (13)
Mean	0.84 (0.80)	0.74 (0.71)	0.72 (0.69)	0.75 (0.74)	0.74 (0.75)	0.74 (0.73)	0.75 (0.70)	0.74 (0.70)	0.73 (0.70)	0.76 (0.71)	0.77 (0.72)	0.69 (0.65)	0.93 (0.92)	13 (11)
T.Mean	0.85 (0.80)	0.73 (0.71)	0.71 (0.67)	0.75 (0.74)	0.73 (0.74)	0.74 (0.72)	0.74 (0.69)	0.73 (0.68)	0.73 (0.69)	0.77 (0.72)	0.78 (0.72)	0.70 (0.64)	0.92 (0.90)	13 (12)
Median	0.85 (0.80)	0.73 (0.70)	0.71 (0.67)	0.75 (0.74)	0.73 (0.75)	0.74 (0.72)	0.74 (0.69)	0.73 (0.68)	0.73 (0.69)	0.77 (0.72)	0.78 (0.73)	0.70 (0.65)	0.92 (0.90)	13 (12)
RF/OLS	0.81 (0.78)	0.72 (0.70)	0.71 (0.67)	0.75 (0.75)	0.74 (0.76)	0.75 (0.74)	0.75 (0.70)	0.73 (0.69)	0.73 (0.70)	0.77 (0.73)	0.78 (0.76)	0.70 (0.68)	0.92 (0.91)	13 (12)
adaLASSO/RF	0.87 (0.81)	0.75 (0.70)	0.69 (0.66)	0.72 (0.73)	0.74 (0.74)	0.71 (0.71)	0.72 (0.68)	0.70 (0.65)	0.71 (0.67)	0.77 (0.73)	0.80 (0.75)	0.70 (0.66)	0.77 (0.77)	13 (13)
RMSE count	11	17	15	19	18	18	18	17	19	19	20	19	17	
MAE count	(13)	(15)	(16)	(16)	(16)	(16)	(15)	(11)	(12)	(16)	(13)	(12)	(6)	

Table 6.10: Forecasting Errors for the CPI from 2001 to 2015

The table reports the p -value of the Giacomini & White's (2005) test of superior predictive ability of the Random Forest model with respect to each one of the alternative models considered in the paper. The test is for squared loss functions and is conducted for expansion versus recessions periods as labeled by the NBER.

	RW	AR	UCSV	adaLASSO	adaElnet	Ridge	Boosting	RF/OLS	adaLASSO/RF
Expansion	$-2.56 \times 10^{-6***}$ (5.32×10 ⁻⁷)	$-8.48 \times 10^{-7***}$ (2.51×10 ⁻⁷)	$-1.04 \times 10^{-6***}$ (3.24×10 ⁻⁷)	-1.45×10^{-7} (1.09×10 ⁻⁷)	1-month-ahead -1.61×10^{-7} (1.08×10 ⁻⁷)	-1.72×10^{-7} (1.53×10 ⁻⁷)	$-1.16 \times 10^{-6***}$ (3.65×10 ⁻⁷)	-9.68×10^{-9} (1.33×10 ⁻⁷)	-6.07×10^{-8} (1.58×10 ⁻⁷)
	Recession	-1.06×10^{-6} (4.02×10 ⁻⁶)	-7.66×10^{-7} (1.58×10 ⁻⁶)	-6.63×10^{-6} (5.11×10 ⁻⁶)	2.21×10^{-6} (1.86×10 ⁻⁶)	1.25×10^{-6} (1.28×10 ⁻⁶)	9.32×10^{-7} (1.40×10 ⁻⁶)	-5.71×10^{-6} (3.56×10 ⁻⁶)	4.94×10^{-6} (3.50×10 ⁻⁶)
Expansion		$-5.25 \times 10^{-6***}$ (1.36×10 ⁻⁶)	$-7.33 \times 10^{-7***}$ (2.49×10 ⁻⁷)	$-1.30 \times 10^{-6***}$ (4.44×10 ⁻⁷)	-3.01×10^{-7} (1.75×10 ⁻⁷)	3-month-ahead -1.72×10^{-7} (1.71×10 ⁻⁷)	-1.69×10^{-7} (1.66×10 ⁻⁷)	$-1.24 \times 10^{-6***}$ (3.78×10 ⁻⁷)	-1.54×10^{-7} (1.58×10 ⁻⁷)
	Recession	$-2.39 \times 10^{-5*}$ (1.41×10 ⁻⁵)	-1.07×10^{-5} (6.80×10 ⁻⁶)	-7.48×10^{-6} (5.00×10 ⁻⁶)	-3.56×10^{-6} (2.23×10 ⁻⁶)	-3.76×10^{-6} (2.38×10 ⁻⁶)	-1.34×10^{-6} (1.90×10 ⁻⁶)	-1.35×10^{-6} (1.03×10 ⁻⁶)	-2.16×10^{-6} (1.77×10 ⁻⁶)
Expansion		$-8.01 \times 10^{-6***}$ (2.45×10 ⁻⁶)	$-1.32 \times 10^{-6***}$ (2.60×10 ⁻⁷)	$-1.78 \times 10^{-6***}$ (4.38×10 ⁻⁷)	$-7.36 \times 10^{-7**}$ (3.15×10 ⁻⁷)	12-month-ahead $-8.92 \times 10^{-7***}$ (2.90×10 ⁻⁷)	-2.00×10^{-7} (2.68×10 ⁻⁷)	$-1.03 \times 10^{-6***}$ (3.36×10 ⁻⁷)	-4.90×10^{-7} (3.14×10 ⁻⁷)
	Recession	-1.17×10^{-5} (8.29×10 ⁻⁶)	-3.36×10^{-6} (3.79×10 ⁻⁶)	-5.98×10^{-6} (4.88×10 ⁻⁶)	$-2.01 \times 10^{-6**}$ (1.01×10 ⁻⁶)	$-2.07 \times 10^{-6*}$ (1.22×10 ⁻⁶)	$-1.49 \times 10^{-6*}$ (8.28×10 ⁻⁷)	$-3.22 \times 10^{-6*}$ (1.97×10 ⁻⁶)	$-1.39 \times 10^{-6***}$ (4.93×10 ⁻⁷)
Expansion		$-4.34 \times 10^{-5*}$ (2.57×10 ⁻⁵)	$-1.18 \times 10^{-4} **$ (1.90×10 ⁻⁵)	-1.50×10^{-5} (2.11×10 ⁻⁵)	$-4.90 \times 10^{-5***}$ (1.10×10 ⁻⁵)	accumulated $-5.05 \times 10^{-5***}$ (1.12×10 ⁻⁵)	$-4.68 \times 10^{-5***}$ (1.80×10 ⁻⁵)	$-2.27 \times 10^{-4***}$ (7.11×10 ⁻⁵)	$-4.92 \times 10^{-5***}$ (1.79×10 ⁻⁵)
	Recession	$-6.30 \times 10^{-4***}$ (2.52×10 ⁻⁴)	$-11.73 \times 10^{-4**}$ (5.25×10 ⁻⁴)	$-4.43 \times 10^{-4**}$ (2.00×10 ⁻⁴)	$-3.91 \times 10^{-4**}$ (1.96×10 ⁻⁴)	$-3.85 \times 10^{-4**}$ (1.90×10 ⁻⁴)	$-9.59 \times 10^{-5} **$ (1.45×10 ⁻⁴)	$-4.98 \times 10^{-5} **$ (2.08×10 ⁻⁴)	-3.06×10^{-4} (2.30×10 ⁻⁴)

Table 6.11: Conditional Giacomini-White Test for Superior Predictive Ability: Recession and Expansions

The table reports the p -value of the Giacomini & White's (2005) test of superior predictive ability of the Random Forest model with respect to each one of the alternative models considered in the paper. The test is for squared loss functions and is conducted for low versus high macroeconomic uncertainty periods.

	RW	AR	UCSV	adaLASSO	adaElnet	Ridge	Boosting	RF/OLS	adaLASSO/RF
	<u>1-month-ahead</u>								
Low	$-1.28 \times 10^{-6***}$ (3.75×10^{-7})	$-4.89 \times 10^{-7***}$ (1.36×10^{-7})	$-6.79 \times 10^{-7***}$ (2.50×10^{-7})	-6.82×10^{-8} (9.38×10^{-8})	-1.06×10^{-7} (8.98×10^{-8})	-2.87×10^{-8} (9.91×10^{-8})	$-1.13 \times 10^{-6***}$ (3.29×10^{-7})	9.40×10^{-8} (1.01×10^{-7})	-2.14×10^{-8} (9.46×10^{-8})
High	$-2.59 \times 10^{-6***}$ (1.45×10^{-6})	$-1.30 \times 10^{-6**}$ (6.33×10^{-7})	$-2.93 \times 10^{-6**}$ (1.46×10^{-6})	3.50×10^{-7} (5.30×10^{-7})	1.22×10^{-7} (3.92×10^{-7})	-8.12×10^{-8} (4.63×10^{-7})	$-2.36 \times 10^{-6**}$ (1.16×10^{-6})	1.11×10^{-6} (9.50×10^{-7})	-2.44×10^{-7} (3.71×10^{-7})
	<u>3-month-ahead</u>								
Low	$-2.52 \times 10^{-6***}$ (8.53×10^{-7})	$-2.86 \times 10^{-7**}$ (1.70×10^{-7})	-5.54×10^{-7} (4.05×10^{-7})	-1.58×10^{-7} (1.40×10^{-7})	-1.47×10^{-7} (1.33×10^{-7})	-1.72×10^{-7} (1.27×10^{-7})	$-1.39 \times 10^{-6***}$ (3.44×10^{-7})	-1.27×10^{-7} (1.06×10^{-7})	$-2.61 \times 10^{-7**}$ (1.36×10^{-7})
High	$-1.41 \times 10^{-5***}$ (4.64×10^{-6})	$-4.02 \times 10^{-6**}$ (2.25×10^{-6})	$-4.02 \times 10^{-6**}$ (1.78×10^{-6})	-1.37×10^{-6} (7.31×10^{-7})	-1.16×10^{-6} (7.84×10^{-7})	-4.75×10^{-7} (5.51×10^{-7})	-1.04×10^{-6} (7.19×10^{-7})	-7.26×10^{-7} (5.35×10^{-7})	-1.23×10^{-7} (5.76×10^{-7})
	<u>12-month-ahead</u>								
Low	$-3.37 \times 10^{-6***}$ (7.17×10^{-7})	$-1.09 \times 10^{-6***}$ (2.74×10^{-7})	$-1.14 \times 10^{-6***}$ (3.91×10^{-7})	$-1.09 \times 10^{-6***}$ 3.45 $\times 10^{-7}$	$-1.32 \times 10^{-6***}$ (3.28×10^{-7})	-2.91×10^{-7} (2.75×10^{-7})	$-6.05 \times 10^{-7} **$ (3.05×10^{-7})	-4.37×10^{-7} (3.58×10^{-7})	-4.31×10^{-7} (2.46×10^{-7})
High	$-1.55 \times 10^{-5***}$ (5.23×10^{-6})	$-2.18 \times 10^{-6**}$ (1.07×10^{-6})	$-3.78 \times 10^{-6***}$ (1.38×10^{-6})	-5.67×10^{-7} (5.27×10^{-7})	-5.96×10^{-7} (5.13×10^{-7})	-4.10×10^{-7} (5.24×10^{-7})	$-2.19 \times 10^{-6***}$ (8.03×10^{-7})	-8.01×10^{-7} (4.67×10^{-7})	-4.18×10^{-7} (5.38×10^{-7})
	<u>accumulated</u>								
Low	-2.32×10^{-5} (2.46×10^{-5})	$-9.19 \times 10^{-5***}$ (2.11×10^{-5})	-1.81×10^{-5} (2.25×10^{-5})	$-3.91 \times 10^{-5***}$ (1.39×10^{-5})	$-4.34 \times 10^{-5***}$ (1.49×10^{-5})	-2.09×10^{-5} (1.43×10^{-5})	$-1.65 \times 10^{-4*}$ (8.84×10^{-5})	$-2.24 \times 10^{-5*}$ (1.32×10^{-5})	$4.13 \times 10^{-6*}$ (-1.32×10^{-5})
High	-2.30×10^{-4} 1.24 $\times 10^{-4*}$	$-4.39 \times 10^{-4**}$ 2.21 $\times 10^{-4}$	-1.26×10^{-4} 8.67 $\times 10^{-5}$	$7.35 \times 10^{-5**}$ -1.55 $\times 10^{-4}$	$7.35 \times 10^{-5**}$ -1.51 $\times 10^{-4}$	$-9.59 \times 10^{-5**}$ 5.06 $\times 10^{-5}$	$1.21 \times 10^{-4**}$ -2.64 $\times 10^{-4}$	$-1.56 \times 10^{-4**}$ 8.15 $\times 10^{-5}$	-1.28×10^{-5} 1.66 $\times 10^{-5}$

Table 6.12: Conditional Giacomini-White Test for Superior Predictive Ability: Macroeconomic Uncertainty

The picture shows.

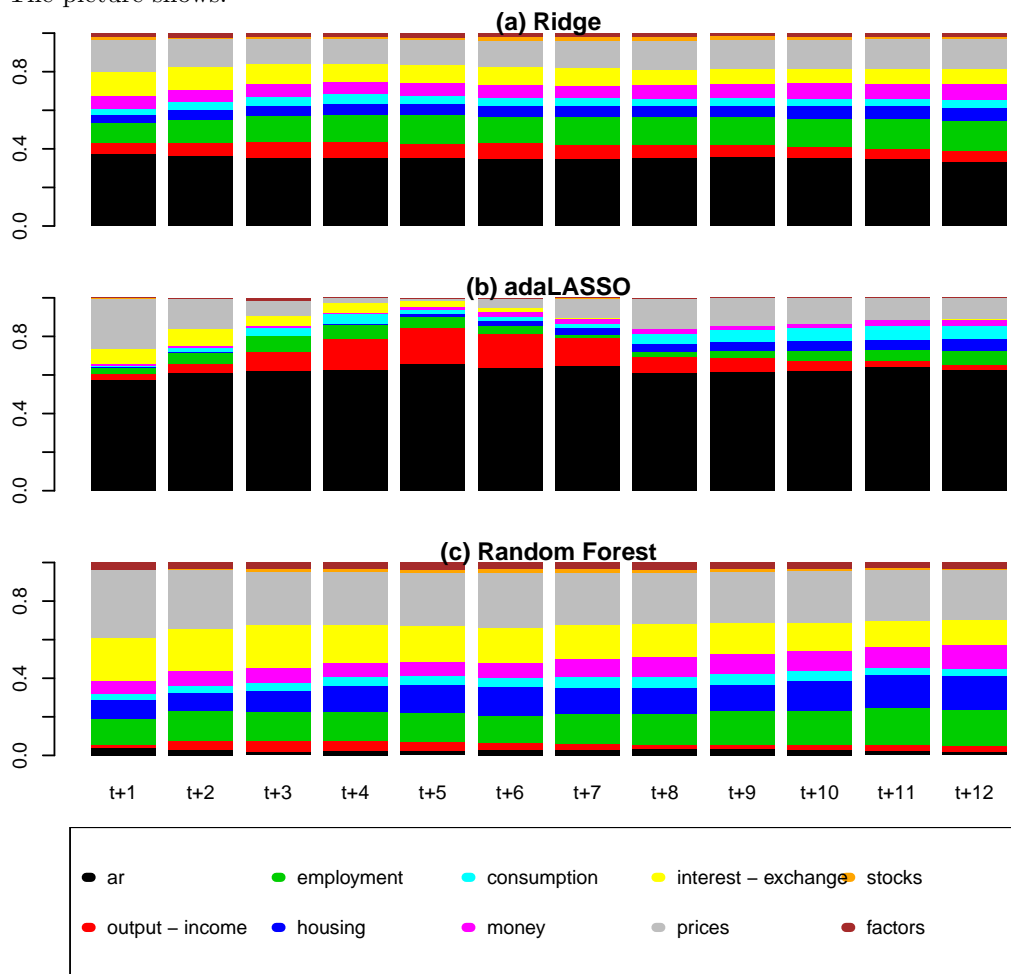


Figure 6.7: Variable importance

Data Appendix

In this section, we present a description of the dataset used in this paper. Tables 6.13–6.20 describe the data and the transformations that were applied to each variable. Each table considers one of the eight different sectors in which the variables are grouped. The column *tcode* denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; and (7) $\Delta(x_t/x_{t-1}-1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in global insight is given in the column GS.

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 1: Output and income						
	id	tcode	fred	description	gsi	gsi:description
1	1	5	RPI	Real Personal Income	M_14386177	PI
2	2	5	W875RX1	Real personal income ex transfer receipts	M_145256755	PI less transfers
3	6	5	INDPRO	IP Index	M_116460980	IP: total
4	7	5	IPFPNSS	IP: Final Products and Nonindustrial Supplies	M_116460981	IP: products
5	8	5	IPFINAL	IP: Final Products (Market Group)	M_116461268	IP: final prod
6	9	5	IPCONGD	IP: Consumer Goods	M_116460982	IP: cons gds
7	10	5	IPDCONGD	IP: Durable Consumer Goods	M_116460983	IP: cons dble
8	11	5	IPNCONGD	IP: Nondurable Consumer Goods	M_116460988	IP: cons nondble
9	12	5	IPBUSEQ	IP: Business Equipment	M_116460995	IP: bus eqpt
10	13	5	IPMAT	IP: Materials	M_116461002	IP: matls
11	14	5	IPDMAT	IP: Durable Materials	M_116461004	IP: dble matls
12	15	5	IPNMAT	IP: Nondurable Materials	M_116461008	IP: nondble matls
13	16	5	IPMANSICS	IP: Manufacturing (SIC)	M_116461013	IP: mfg
14	17	5	IPB51222s	IP: Residential Utilities	M_116461276	IP: res util
15	18	5	IPFUELS	IP: Fuels	M_116461275	IP: fuels
16	19	1	NAPMPI	ISM Manufacturing: Production Index	M_110157212	NAPM prodn
17	20	2	CUMFNS	Capacity Utilization: Manufacturing	M_116461602	Cap uti

Table 6.13: Data Description: Output and Income

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 2: Labor market						
id	tcode	fred	description	gsi	gsi:description	
1	21*	2	HWI	Help-Wanted Index for United States		Help wanted indx
2	22*	2	HWIURATIO	Ratio of Help Wanted/No. Unemployed	M_110156531	Help wanted/unemp
3	23	5	CLF16OV	Civilian Labor Force	M_110156467	Emp CPS total
4	24	5	CE16OV	Civilian Employment	M_110156498	Emp CPS nonag
5	25	2	UNRATE	Civilian Unemployment Rate	M_110156541	U: all
6	26	2	UEMPMEAN	Average Duration of Unemployment (Weeks)	M_110156528	U: mean duration
7	27	5	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	M_110156527	U < 5 wks
8	28	5	UEMP5TO14	Civilians Unemployed for 5-14 Weeks	M_110156523	U 5-14 wks
9	29	5	UEMP15OV	Civilians Unemployed - 15 Weeks & Over	M_110156524	U 15+ wks
10	30	5	UEMP15T26	Civilians Unemployed for 15-26 Weeks	M_110156525	U 15-26 wks
11	31	5	UEMP27OV	Civilians Unemployed for 27 Weeks and Over	M_110156526	U 27+ wks
12	32*	5	CLAIMSx	Initial Claims	M_15186204	UI claims
13	33	5	PAYEMS	All Employees: Total nonfarm	M_123109146	Emp: total
14	34	5	USGOOD	All Employees: Goods-Producing Industries	M_123109172	Emp: gds prod
15	35	5	CES1021000001	All Employees: Mining and Logging: Mining	M_123109244	Emp: mining
16	36	5	USCONS	All Employees: Construction	M_123109331	Emp: const
17	37	5	MANEMP	All Employees: Manufacturing	M_123109542	Emp: mfg
18	38	5	DMANEMP	All Employees: Durable goods	M_123109573	Emp: dble gds
19	39	5	NDMANEMP	All Employees: Nondurable goods	M_123110741	Emp: nondbles
20	40	5	SRVPRD	All Employees: Service-Providing Industries	M_123109193	Emp: services
21	41	5	USTPU	All Employees: Trade, Transportation & Utilities	M_123111543	Emp: TTU
22	42	5	USWTRADE	All Employees: Wholesale Trade	M_123111563	Emp: wholesale
23	43	5	USTRADE	All Employees: Retail Trade	M_123111867	Emp: retail
24	44	5	USFIRE	All Employees: Financial Activities	M_123112777	Emp: FIRE
25	45	5	USGOVT	All Employees: Government	M_123114411	Emp: Govt
26	46	1	CES0600000007	Avg Weekly Hours : Goods-Producing	M_140687274	Avg hrs
27	47	2	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	M_123109554	Overtime: mfg
28	48	1	AWHMAN	Avg Weekly Hours : Manufacturing	M_14386098	Avg hrs: mfg
29	49	1	NAPMEI	ISM Manufacturing: Employment Index	M_110157206	NAPM empl
30	127	6	CES0600000008	Avg Hourly Earnings : Goods-Producing	M_123109182	AHE: goods
31	128	6	CES2000000008	Avg Hourly Earnings : Construction	M_123109341	AHE: const
32	129	6	CES3000000008	Avg Hourly Earnings : Manufacturing	M_123109552	AHE: mfg

Table 6.14: Data Description: Labor Market

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 3: Housing						
id	tcode	fred	description	gsi	gsi:description	
1	50	4	HOUST	Housing Starts: Total New Privately Owned	M_110155536	Starts: nonfarm
2	51	4	HOUSTNE	Housing Starts, Northeast	M_110155538	Starts: NE
3	52	4	HOUSTMW	Housing Starts, Midwest	M_110155537	Starts: MW
4	53	4	HOUSTS	Housing Starts, South	M_110155543	Starts: South
5	54	4	HOUSTW	Housing Starts, West	M_110155544	Starts: West
6	55	4	PERMIT	New Private Housing Permits (SAAR)	M_110155532	BP: total
7	56	4	PERMITNE	New Private Housing Permits, Northeast (SAAR)	M_110155531	BP: NE
8	57	4	PERMITMW	New Private Housing Permits, Midwest (SAAR)	M_110155530	BP: MW
9	58	4	PERMITS	New Private Housing Permits, South (SAAR)	M_110155533	BP: South
10	59	4	PERMITW	New Private Housing Permits, West (SAAR)	M_110155534	BP: West

Table 6.15: Data Description: Housing

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 4: Consumption, orders, and inventories						
id	tcode	fred	description	gsi	gsi:description	
1	3	5	DPCERA3M086SBEA	Real personal consumption expenditures	M_123008274	Real Consumption
2	4*	5	CMRMTSPLx	Real Manu. and Trade Industries Sales	M_110156998	M&T sales
3	5*	5	RETAILx	Retail and Food Services Sales	M_130439509	Retail sales
4	60	1	NAPM	ISM : PMI Composite Index	M_110157208	PMI
5	61	1	NAPMNOI	ISM : New Orders Index	M_110157210	NAPM new ordrs
6	62	1	NAPMSDI	ISM : Supplier Deliveries Index	M_110157205	NAPM vendor del
7	63	1	NAPMII	ISM : Inventories Index	M_110157211	NAPM Invent
8	64	5	ACOGNO	New Orders for Consumer Goods	M_14385863	Orders: cons gds
9	65*	5	AMDMMNOx	New Orders for Durable Goods	M_14386110	Orders: dble gds
10	66*	5	ANDENOX	New Orders for Nondefense Capital Goods	M_178554409	Orders: cap gds
11	67*	5	AMDMMUOX	Unfilled Orders for Durable Goods	M_14385946	Unf orders: dble
12	68*	5	BUSINVx	Total Business Inventories	M_15192014	M&T invent
13	69*	2	ISRATIOx	Total Business: Inventories to Sales Ratio	M_15191529	M&T invent/sales
14	130*	2	UMCSENTx	Consumer Sentiment Index	hhsntn	Consumer expect

Table 6.16: Data Description: Consumption, Orders and Inventories

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 5: Money and credit						
id	tcode	fred	description	gsi	gsi:description	
1	70	6	M1SL	M1 Money Stock	M_110154984	M1
2	71	6	M2SL	M2 Money Stock	M_110154985	M2
3	72	5	M2REAL	Real M2 Money Stock	M_110154985	M2 (real)
4	73	6	AMBSL	St. Louis Adjusted Monetary Base	M_110154995	MB
5	74	6	TOTRESNS	Total Reserves of Depository Institutions	M_110155011	Reserves tot
6	75	7	NONBORRES	Reserves Of Depository Institutions	M_110155009	Reserves nonbor
7	76	6	BUSLOANS	Commercial and Industrial Loans	BUSLOANS	C&I loan plus
8	77	6	REALLN	Real Estate Loans at All Commercial Banks	BUSLOANS	DC&I loans
9	78	6	NONREVSL	Total Nonrevolving Credit	M_110154564	Cons credit
10	79*	2	CONSPI	Nonrevolving consumer credit to Personal Income	M_110154569	Inst cred/PI
11	131	6	MZMSL	MZM Money Stock	N.A.	N.A.
12	132	6	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding	N.A.	N.A.
13	133	6	DTCTHFNM	Total Consumer Loans and Leases Outstanding	N.A.	N.A.
14	134	6	INVEST	Securities in Bank Credit at All Commercial Banks	N.A.	N.A.

Table 6.17: Data Description: Money and Credit

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 6: Interest and exchange rates						
id	tcode	fred	description	gsi	gsi:description	
1	84	2	FEDFUNDS	Effective Federal Funds Rate	M_110155157	Fed Funds
2	85*	2	CP3Mx	3-Month AA Financial Commercial Paper Rate	CPF3M	Comm paper
3	86	2	TB3MS	3-Month Treasury Bill:	M_110155165	3 mo T-bill
4	87	2	TB6MS	6-Month Treasury Bill:	M_110155166	6 mo T-bill
5	88	2	GS1	1-Year Treasury Rate	M_110155168	1 yr T-bond
6	89	2	GS5	5-Year Treasury Rate	M_110155174	5 yr T-bond
7	90	2	GS10	10-Year Treasury Rate	M_110155169	10 yr T-bond
8	91	2	AAA	Moody's Seasoned Aaa Corporate Bond Yield		Aaa bond
9	92	2	BAA	Moody's Seasoned Baa Corporate Bond Yield		Baa bond
10	93*	1	COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS		CP-FF spread
11	94	1	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS		3 mo-FF spread
12	95	1	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS		6 mo-FF spread
13	96	1	T1YFFM	1-Year Treasury C Minus FEDFUNDS		1 yr-FF spread
14	97	1	T5YFFM	5-Year Treasury C Minus FEDFUNDS		5 yr-FF spread
15	98	1	T10YFFM	10-Year Treasury C Minus FEDFUNDS		10 yr-FF spread
16	99	1	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS		Aaa-FF spread
17	100	1	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS		Baa-FF spread
18	101	5	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies		Ex rate: avg
19	102	* 5	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	M_110154768	Ex rate: Switz
20	103	* 5	EXJPUSx	Japan / U.S. Foreign Exchange Rate	M_110154755	Ex rate: Japan
21	104	* 5	EXUSUKx	U.S. / U.K. Foreign Exchange Rate	M_110154772	Ex rate: UK
22	105	* 5	EXCAUSx	Canada / U.S. Foreign Exchange Rate	M_110154744	Ex rate: Canada

Table 6.18: Data Description: Interest and Exchange Rates

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 7: Prices						
id	tcode	fred	description	gsi	gsi:description	
1	106	6	WPSFD49207	PPI: Finished Goods	M110157517	PPI: fin gds
2	107	6	WPSFD49502	PPI: Finished Consumer Goods	M110157508	PPI: cons gds
3	108	6	WPSID61	PPI: Intermediate Materials	M_110157527	PPI: int matls
4	109	6	WPSID62	PPI: Crude Materials	M_110157500	PPI: crude matls
5	110*	6	OILPRICEx	Crude Oil, spliced WTI and Cushing	M_110157273	Spot market price
6	111	6	PPICMM	PPI: Metals and metal products	M_110157335	PPI: nonferrous
7	112	1	NAPMPRI	ISM Manufacturing: Prices Index	M_110157204	NAPM com price
8	113	6	CPIAUCSL	CPI : All Items	M_110157323	CPI-U: all
9	114	6	CPIAPPSL	CPI : Apparel	M_110157299	CPI-U: apparel
10	115	6	CPIITRNSL	CPI : Transportation	M_110157302	CPI-U: transp
11	116	6	CPIMEDSL	CPI : Medical Care	M_110157304	CPI-U: medical
12	117	6	CUSR0000SAC	CPI : Commodities	M_110157314	CPI-U: comm.
13	118	6	CUUR0000SAD	CPI : Durables	M_110157315	CPI-U: dbles
14	119	6	CUSR0000SAS	CPI : Services	M_110157325	CPI-U: services
15	120	6	CPIULFSL	CPI : All Items Less Food	M_110157328	CPI-U: ex food
16	121	6	CUUR0000SA0L2	CPI : All items less shelter	M_110157329	CPI-U: ex shelter
17	122	6	CUSR0000SA0L5	CPI : All items less medical care	M_110157330	CPI-U: ex med
18	123	6	PCEPI	Personal Cons. Expend.: Chain Index	gmde	PCE defl
19	124	6	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	gmded	PCE defl: dlbes
20	125	6	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	gmden	PCE defl: nondble
21	126	6	DSERRG3M086SBEA	Personal Cons. Exp: Services	gmdcs	PCE defl: service

Table 6.19: Data Description: Prices

The column tcode denotes the following data transformation for a series x : (1) no transformation; (2) Δx_t ; (3) $2\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t)$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GS.

Group 8: Stock Market						
id	tcode	fred	description	gsi	gsi:description	
1	80*	5	S&P 500	S&P's Common Stock Price Index: Composite	M_110155044	S&P 500
2	81*	5	S&P: indust	S&P's Common Stock Price Index: Industrials	M_110155047	S&P: indust
3	82*	2	S&P div yield	S&P's Composite Common Stock: Dividend Yield		S&P div yield
4	83*	5	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio		S&P PE ratio
5	135*	1	VXOCLSx	VXO		

Table 6.20: Data Description: Stock Market

Appendix Models

For all models, with the exception of the RW and UCSV specifications, we include a dummy for the November 2008, when a huge deflation was observed. This appendix was removed because all the information concerning models was discussed in chapter 2.

Appendix Additional Results

Variable Selection: Word Clouds

This Appendix presents the variable selection for several models as *word clouds*. In the present context, a word cloud is an image composed of the names of variables selected by a specific model across the estimation windows in which the size of each word indicates its frequency or importance. The names displayed in the clouds are as defined in the third column of Tables 6.13–6.20. These names represent FRED mnemonics. The clouds also indicate the degree of sparsity of each model. For instance, the smaller the cloud is, the more sparse the model is.

Figures 6.8 and 6.9 display the word clouds for the linear model estimated with the adaLASSO method for the first and second subsamples, respectively. In each figure, panel (a) shows the cloud for one-month-ahead models ($h = 1$), panel (b) presents the results for the three-month horizon ($h = 3$), and panels (c) and (d) consider the cases for $h = 6$ and $h = 12$, respectively. A number of findings emerge from the word clouds. First, as expected, the adaLASSO method delivers very sparse methods, and this did not change much according to the subsample considered. Second, the models across different horizons, as shown before, are quite different. For example, in the first subsample and for $h = 1$, the three variables that stand out from the cloud are CUSR0000SAOL5 (CPI: all items less medical care), WPSFD49207

(PPI: finished goods), and DSERRG3M086SBEA (PCE: Services). However, for $h = 12$, the most important variables are AMDMUOx (unfilled orders for durable goods) and HOUSTMW (Housing starts, Midwest). Finally, the pool of selected variables also changes from the first to the second sample, specially for $h = 1$. In this case, oil prices turn out to be one of the most relevant variables.

Figures 6.10 and 6.11 shows the word clouds for the RF model. From the pictures it is clear that the number of important variables are much higher. As in the adaLASSO case, the variable composition changes from the first to the second subsample.

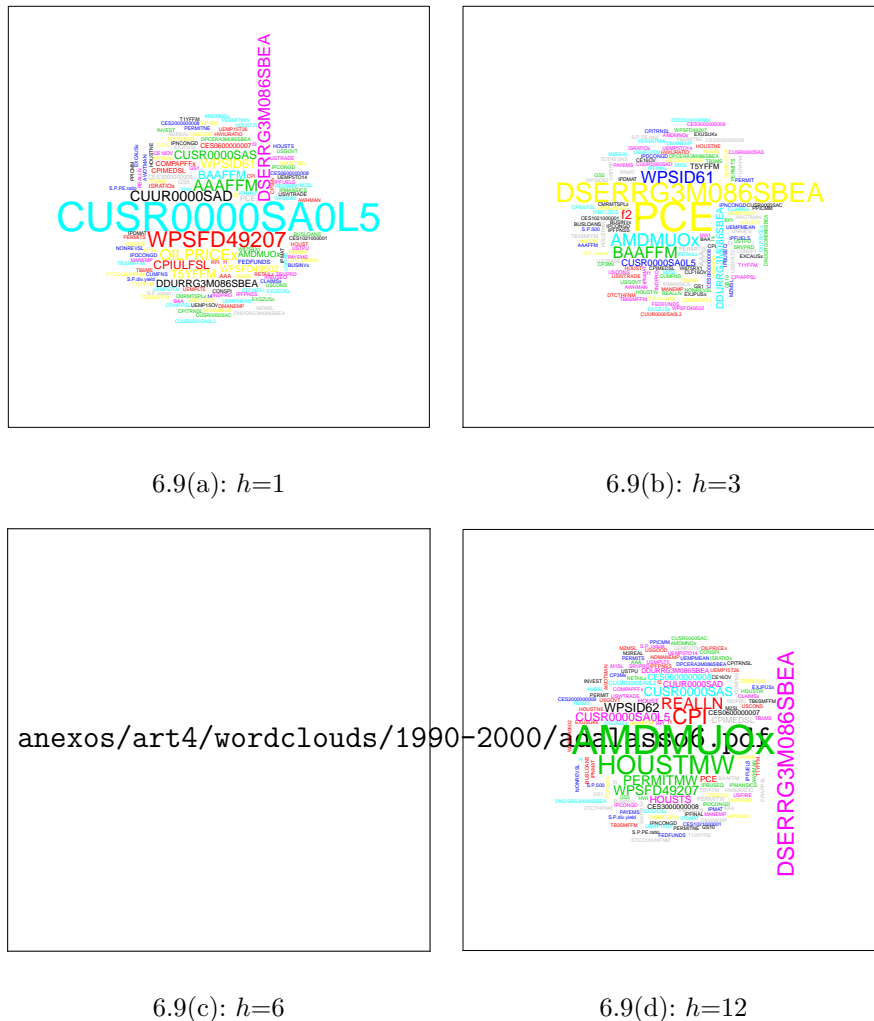
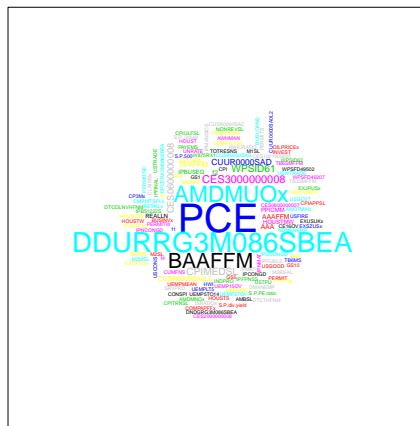


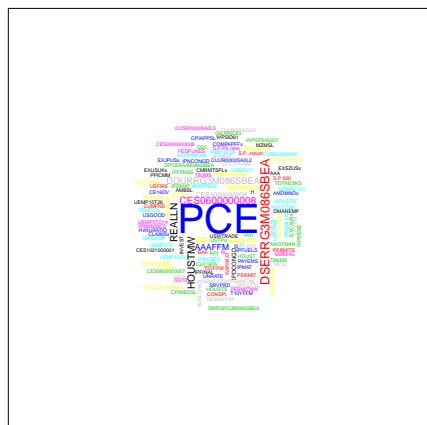
Figure 6.8: Word clouds for the adaLASSO method (1990–2000).



6.10(a): $h=1$



6.10(b): $h=3$

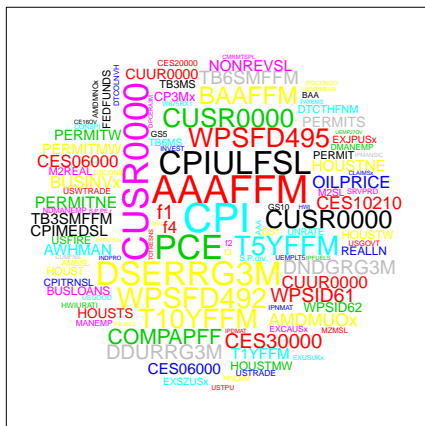


6.10(c): $h=6$

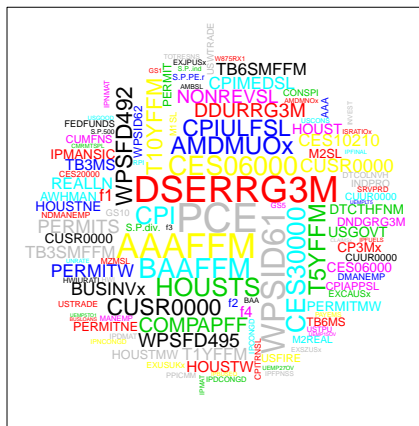


6.10(d): $h=12$

Figure 6.9: Word clouds for the adaLASSO method (2001–2015).



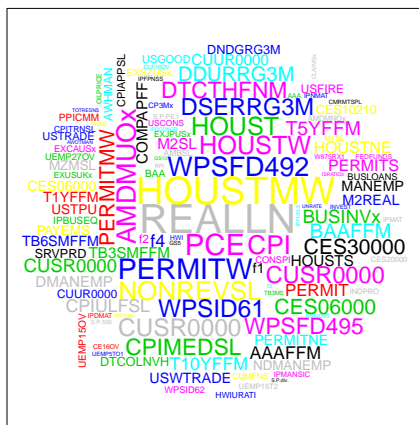
6.11(a): $h=1$



6.11(b): $h=3$



6.11(c): $h=6$



6.11(d): $h=12$

Figure 6.10: Word clouds for the Random Forest model (1990-2000).



Figure 6.11: Word clouds for the Random Forest model (2001—2015).

Additional Results: Personal Consumption Expenditure (PCE)

In this section, we report forecasting results for PCE. The main message is similar to the one described in the main text: RF models outperform traditional benchmarks as well as other linear ML methods.

In Tables 6.21–6.23, we report for each model a number of different summary statistics across all the forecasting horizons, including the accumulated twelve-month horizon for the full out-of-sample period (1990–2015) as well as for the two subsamples considered, namely, 1990–2000 and 2001–2015. Columns (1), (2) and (3) report the RMSE, the MAE and the MAD, respectively. In columns (4), (5) and (6) we report the number of

times (across horizons) each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the **range** or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the SPA test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84). The superiority of the RF models is clear from the tables.

Tables 6.24–6.26 show the RMSE and, in parenthesis, the MAE for all models relative to the RW. The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% MCS using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum t statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table report how many models were included in the MCS for square and absolute losses. Again, the performance of the RF model is remarkable.

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.063	0.002	0.095	0.012	0.008	0.000	0.000	0.000
AR	0.835	0.825	0.802	0	0	0	0.305	0.193	0.483	0.361	0.082	0.042	0.000	0.000
UCSV	0.848	0.853	0.873	0	0	0	0.132	0.071	0.335	0.232	0.019	0.010	0.000	0.000
LASSO	0.792	0.787	0.758	0	0	1	0.254	0.286	0.603	0.570	0.153	0.188	0.623	0.623
adaLASSO	0.800	0.789	0.759	0	0	0	0.265	0.337	0.577	0.507	0.125	0.212	0.033	0.033
ElNet	0.791	0.791	0.752	0	0	1	0.266	0.264	0.603	0.554	0.163	0.174	0.997	0.997
adaElnet	0.800	0.792	0.761	0	0	0	0.250	0.226	0.581	0.480	0.106	0.118	0.001	0.001
Ridge	0.776	0.774	0.784	0	0	1	0.460	0.512	0.678	0.632	0.302	0.406	0.869	0.869
BVAR	0.837	0.852	0.813	0	0	0	0.244	0.216	0.432	0.334	0.062	0.048	0.000	0.000
Bagging	0.810	0.831	0.847	0	0	0	0.300	0.194	0.560	0.390	0.112	0.085	0.054	0.054
CSR	0.813	0.798	0.781	0	0	1	0.360	0.474	0.602	0.573	0.200	0.310	0.000	0.000
JMA	0.870	0.913	0.950	0	0	0	0.070	0.010	0.238	0.092	0.010	0.000	0.000	0.000
Factor	0.857	0.861	0.822	0	0	1	0.103	0.014	0.245	0.128	0.016	0.003	0.000	0.000
T. Factor	0.834	0.842	0.849	0	0	0	0.071	0.059	0.302	0.183	0.020	0.031	0.000	0.000
Boosting	0.859	0.894	0.926	0	0	0	0.158	0.077	0.495	0.230	0.007	0.003	0.000	0.000
RF	0.745	0.736	0.740	11	9	3	0.589	0.664	0.960	0.968	0.928	0.946	1	1
Mean	0.777	0.767	0.761	0	0	0	0.546	0.589	0.682	0.680	0.397	0.559	0.995	0.997
T. Mean	0.778	0.766	0.753	0	1	1	0.392	0.550	0.680	0.684	0.286	0.581	0.985	0.997
Median	0.779	0.766	0.749	0	0	1	0.417	0.541	0.682	0.684	0.308	0.562	0.969	0.997
RF/OLS	0.780	0.779	0.783	1	1	1	0.492	0.647	0.702	0.655	0.444	0.519	0.993	0.997
adaLASSO/RF	0.756	0.750	0.742	1	2	2	0.658	0.790	0.746	0.766	0.528	0.598	0.905	0.997

Table 6.21: Forecasting Results PCE: Summary statistics for the out-of-sample period from 1990–2015

PUC-Rio - Certificação Digital N° 1421640/CA

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.095	0.028	0.335	0.346	0.008	0.003	0.000	0.000
AR	0.849	0.871	0.782	0	0	1	0.366	0.351	0.715	0.716	0.302	0.258	0.199	0.199
UCSV	0.882	0.892	0.902	0	0	0	0.407	0.495	0.740	0.773	0.204	0.377	0.228	0.228
LASSO	0.875	0.908	0.812	0	0	1	0.122	0.161	0.534	0.482	0.032	0.083	0.008	0.008
adaLASSO	0.856	0.870	0.818	0	0	0	0.173	0.350	0.650	0.730	0.080	0.211	0.008	0.008
EINet	0.887	0.928	0.803	1	0	1	0.127	0.115	0.477	0.345	0.079	0.080	0.777	0.863
adaElnet	0.861	0.883	0.817	0	0	0	0.119	0.183	0.583	0.585	0.058	0.086	0.001	0.001
Ridge	0.828	0.850	0.827	0	1	0	0.481	0.525	0.845	0.832	0.326	0.376	0.863	0.863
BVAR	1.011	1.078	0.853	0	0	1	0.268	0.165	0.383	0.263	0.054	0.044	0.005	0.005
Bagging	0.892	0.920	0.936	0	0	0	0.448	0.634	0.773	0.755	0.215	0.293	0.177	0.177
CSR	0.825	0.841	0.839	3	4	0	0.589	0.640	0.863	0.890	0.526	0.600	0.993	0.993
JMA	1.025	1.044	1.052	0	0	0	0.107	0.131	0.271	0.268	0.016	0.029	0.000	0.000
Factor	0.949	0.985	0.866	0	0	2	0.059	0.068	0.277	0.252	0.006	0.010	0.000	0.000
T. Factor	0.929	0.959	0.876	0	0	1	0.083	0.102	0.363	0.325	0.029	0.071	0.001	0.001
Boosting	0.987	1.038	1.062	0	0	0	0.293	0.274	0.497	0.452	0.024	0.042	0.000	0.000
RF	0.799	0.816	0.725	7	5	5	0.676	0.625	0.953	0.912	0.889	0.748	1	1
Mean	0.817	0.834	0.789	0	0	0	0.658	0.668	0.863	0.882	0.638	0.642	0.961	0.993
T. Mean	0.827	0.843	0.788	0	0	0	0.303	0.380	0.836	0.876	0.250	0.284	0.067	0.067
Median	0.829	0.846	0.788	0	0	0	0.342	0.405	0.815	0.848	0.266	0.311	0.150	0.150
RF/OLS	0.815	0.837	0.845	2	3	0	0.794	0.742	0.888	0.910	0.794	0.750	0.943	0.993
adaLASSO/RF	0.835	0.852	0.790	0	0	1	0.453	0.457	0.754	0.736	0.373	0.362	0.795	0.863

Table 6.22: Forecasting Results PCE: Summary statistics for the out-of-sample period from 1990–2000

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	0	0	0	0.090	0.002	0.090	0.004	0.011	0.000	0.000	0.000
AR	0.832	0.805	0.788	0	0	0	0.332	0.165	0.543	0.274	0.082	0.032	0.000	0.000
UCSV	0.841	0.837	0.842	0	0	0	0.138	0.041	0.316	0.144	0.024	0.007	0.000	0.000
LASSO	0.771	0.733	0.708	0	1	0	0.470	0.547	0.802	0.680	0.309	0.451	0.887	0.935
adaLASSO	0.786	0.752	0.708	0	0	3	0.423	0.430	0.756	0.616	0.234	0.245	0.209	0.209
ElNet	0.767	0.730	0.708	0	0	0	0.449	0.541	0.812	0.699	0.292	0.485	0.824	0.935
adaElnet	0.785	0.751	0.715	0	0	0	0.446	0.436	0.760	0.615	0.236	0.220	0.149	0.149
Ridge	0.763	0.750	0.706	0	0	0	0.529	0.584	0.818	0.645	0.380	0.417	0.408	0.408
BVAR	0.790	0.750	0.701	0	0	3	0.405	0.538	0.767	0.679	0.209	0.317	0.200	0.200
Bagging	0.789	0.794	0.782	0	0	2	0.416	0.278	0.661	0.359	0.223	0.105	0.176	0.176
CSR	0.810	0.778	0.758	0	0	0	0.431	0.399	0.657	0.518	0.188	0.188	0.013	0.013
JMA	0.830	0.856	0.883	0	0	0	0.288	0.074	0.480	0.153	0.062	0.006	0.000	0.000
Factor	0.834	0.805	0.765	0	0	0	0.296	0.103	0.445	0.234	0.070	0.024	0.000	0.000
T. Factor	0.809	0.790	0.771	0	0	0	0.319	0.129	0.543	0.262	0.114	0.050	0.000	0.000
Boosting	0.826	0.832	0.829	0	0	0	0.334	0.228	0.690	0.407	0.040	0.029	0.000	0.000
RF	0.733	0.700	0.688	8	4	2	0.768	0.823	0.934	0.887	0.863	0.875	1	1
Mean	0.768	0.736	0.710	0	0	0	0.591	0.482	0.818	0.671	0.424	0.402	0.876	0.876
T.Mean	0.766	0.731	0.714	0	0	0	0.482	0.616	0.811	0.693	0.348	0.622	0.794	0.935
Median	0.766	0.731	0.711	0	0	0	0.507	0.640	0.816	0.704	0.397	0.640	0.868	0.935
RF/OLS	0.772	0.754	0.705	1	0	2	0.540	0.550	0.826	0.648	0.343	0.330	0.935	0.935
adaLASSO/RF	0.736	0.705	0.680	4	8	1	0.706	0.741	0.878	0.893	0.774	0.746	0.910	0.935

Table 6.23: Forecasting Results PCE: Summary statistics for the out-of-sample period from 2001–2015

Additional Results: CPI-Core

In this section, we report forecasting results for the Core of the Consumer Price Index. The CPI-Core series exhibits a dynamics quite different from the other two inflation indexes reported before. More specifically there is a clear seasonal pattern in the series.

In Tables 6.27–6.29, we report for each model a number of different summary statistics across all the forecasting horizons, including the accumulated twelve-month horizon for the full out-of-sample period (1990–2015) as well as for the two subsamples considered, namely, 1990–2000 and 2001–2015. Columns (1), (2) and (3) report the RMSE, the MAE and the MAD, respectively. In columns (4), (5) and (6) we report the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD, respectively. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the **range** or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the SPA test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Tables 6.30–6.32 show the RMSE and, in parenthesis, the MAE for all models relative to the RW. The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% MCS using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum t statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table report how many models were included in the MCS for square and absolute losses.

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Personal Consumer Expenditure 1990–2015														
RMSE/(MAE)	1	2	3	4	Forecasting Horizon				9	10	11	12	Acc.	RMSE count (MAE count)
					6	7	8							
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0 (0)
AR	0.89 (0.87)	0.83 (0.81)	0.82 (0.79)	0.84 (0.83)	0.82 (0.82)	0.80 (0.81)	0.79 (0.79)	0.78 (0.75)	0.82 (0.82)	0.84 (0.84)	0.87 (0.89)	0.80 (0.82)	0.96 (0.91)	7 (3)
UCSV	0.94 (0.92)	0.86 (0.86)	0.84 (0.82)	0.86 (0.85)	0.84 (0.84)	0.83 (0.83)	0.83 (0.84)	0.82 (0.82)	0.83 (0.84)	0.84 (0.84)	0.86 (0.90)	0.83 (0.83)	0.85 (0.90)	5 (0)
LASSO	0.83 (0.80)	0.78 (0.77)	0.75 (0.73)	0.79 (0.80)	0.78 (0.81)	0.78 (0.79)	0.78 (0.78)	0.77 (0.73)	0.81 (0.80)	0.82 (0.81)	0.84 (0.86)	0.76 (0.77)	0.79 (0.77)	9 (7)
adaLASSO	0.84 (0.82)	0.79 (0.78)	0.77 (0.74)	0.80 (0.80)	0.79 (0.79)	0.78 (0.77)	0.78 (0.77)	0.79 (0.74)	0.82 (0.80)	0.82 (0.80)	0.84 (0.85)	0.76 (0.77)	0.83 (0.82)	9 (6)
ElNet	0.83 (0.80)	0.78 (0.77)	0.75 (0.73)	0.79 (0.81)	0.78 (0.81)	0.78 (0.80)	0.77 (0.78)	0.77 (0.74)	0.81 (0.81)	0.82 (0.81)	0.85 (0.87)	0.76 (0.78)	0.79 (0.78)	9 (6)
adaElnet	0.84 (0.83)	0.79 (0.79)	0.76 (0.74)	0.80 (0.80)	0.79 (0.80)	0.78 (0.78)	0.78 (0.77)	0.78 (0.74)	0.82 (0.80)	0.82 (0.81)	0.84 (0.86)	0.76 (0.77)	0.83 (0.81)	9 (5)
Ridge	0.85 (0.83)	0.76 (0.75)	0.76 (0.73)	0.79 (0.78)	0.76 (0.79)	0.77 (0.78)	0.75 (0.75)	0.75 (0.72)	0.79 (0.79)	0.78 (0.78)	0.81 (0.83)	0.74 (0.75)	0.77 (0.77)	13 (9)
BVAR	0.90 (0.89)	0.80 (0.78)	0.79 (0.79)	0.81 (0.83)	0.79 (0.83)	0.80 (0.84)	0.81 (0.83)	0.82 (0.82)	0.86 (0.89)	0.87 (0.89)	0.90 (0.94)	0.83 (0.87)	0.90 (0.89)	5 (3)
Bagging	0.87 (0.86)	0.78 (0.76)	0.78 (0.76)	0.83 (0.86)	0.82 (0.90)	0.81 (0.86)	0.82 (0.83)	0.80 (0.80)	0.83 (0.86)	0.81 (0.83)	0.84 (0.88)	0.76 (0.78)	0.76 (0.82)	8 (4)
CSR	0.85 (0.82)	0.78 (0.78)	0.77 (0.74)	0.80 (0.80)	0.79 (0.79)	0.78 (0.78)	0.79 (0.76)	0.78 (0.73)	0.82 (0.80)	0.84 (0.82)	0.87 (0.87)	0.81 (0.81)	0.90 (0.87)	11 (8)
JMA	0.95 (0.94)	0.88 (0.91)	0.84 (0.85)	0.89 (0.97)	0.87 (0.97)	0.84 (0.91)	0.88 (0.91)	0.85 (0.86)	0.92 (0.97)	0.87 (0.92)	0.91 (0.95)	0.82 (0.88)	0.79 (0.83)	2 (1)
Factor	0.89 (0.87)	0.84 (0.84)	0.83 (0.83)	0.84 (0.84)	0.83 (0.85)	0.84 (0.85)	0.83 (0.84)	0.83 (0.81)	0.87 (0.89)	0.87 (0.88)	0.89 (0.90)	0.85 (0.88)	0.95 (0.92)	0 (0)
T. Factor	0.90 (0.91)	0.83 (0.85)	0.81 (0.82)	0.81 (0.80)	0.81 (0.83)	0.81 (0.83)	0.80 (0.81)	0.80 (0.77)	0.83 (0.84)	0.84 (0.86)	0.85 (0.87)	0.82 (0.84)	0.92 (0.90)	2 (1)
Boosting	0.99 (1.00)	0.83 (0.83)	0.82 (0.83)	0.84 (0.89)	0.82 (0.89)	0.84 (0.89)	0.84 (0.88)	0.84 (0.85)	0.87 (0.91)	0.86 (0.88)	0.88 (0.92)	0.81 (0.83)	0.91 (1.02)	7 (1)
RF	0.86 (0.82)	0.76 (0.74)	0.74 (0.73)	0.76 (0.77)	0.73 (0.77)	0.73 (0.75)	0.72 (0.72)	0.72 (0.68)	0.75 (0.73)	0.75 (0.74)	0.78 (0.78)	0.71 (0.70)	0.67 (0.63)	13 (13)
Mean	0.84 (0.81)	0.77 (0.77)	0.76 (0.73)	0.78 (0.78)	0.77 (0.78)	0.77 (0.77)	0.76 (0.76)	0.75 (0.71)	0.78 (0.77)	0.78 (0.77)	0.80 (0.81)	0.75 (0.75)	0.80 (0.77)	13 (10)
T.Mean	0.84 (0.81)	0.77 (0.76)	0.75 (0.73)	0.78 (0.78)	0.77 (0.78)	0.76 (0.77)	0.76 (0.75)	0.75 (0.71)	0.78 (0.77)	0.79 (0.77)	0.81 (0.82)	0.75 (0.75)	0.80 (0.77)	13 (10)
Median	0.83 (0.81)	0.77 (0.76)	0.75 (0.73)	0.78 (0.78)	0.77 (0.78)	0.77 (0.77)	0.76 (0.75)	0.76 (0.71)	0.78 (0.77)	0.79 (0.77)	0.81 (0.82)	0.75 (0.75)	0.79 (0.77)	13 (10)
RF/OLS	0.82 (0.80)	0.76 (0.76)	0.75 (0.73)	0.78 (0.78)	0.78 (0.78)	0.76 (0.78)	0.76 (0.75)	0.76 (0.72)	0.79 (0.79)	0.79 (0.85)	0.82 (0.85)	0.76 (0.78)	0.82 (0.82)	13 (9)
adaLASSO/RF	0.85 (0.81)	0.79 (0.78)	0.75 (0.73)	0.76 (0.78)	0.74 (0.77)	0.74 (0.75)	0.72 (0.72)	0.72 (0.68)	0.76 (0.74)	0.76 (0.76)	0.80 (0.81)	0.74 (0.73)	0.70 (0.68)	13 (11)
RMSE count	13	16	13	16	16	17	8	8	16	18	7	8	18	
MAE count	(10)	(16)	(13)	(15)	(14)	(11)	(7)	(8)	(1)	(3)	(1)	(8)	(10)	

Table 6.24: Forecasting Errors for the PCE from 1990 to 2015

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Personal Consumer Expenditure 1990-2000														
RMSE/(MAE)	Forecasting Horizon											Acc.	RMSE count (MAE count)	
	1	2	3	4	5	6	7	8	9	10	11			12
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	3 (2)
AR	0.84 (0.86)	0.80 (0.79)	0.86 (0.88)	0.82 (0.85)	0.79 (0.79)	0.80 (0.84)	0.83 (0.89)	0.84 (0.80)	0.90 (0.89)	0.86 (0.90)	0.95 (1.00)	0.92 (0.98)	0.82 (0.84)	12 (11)
UCSV	0.89 (0.90)	0.85 (0.84)	0.86 (0.87)	0.88 (0.88)	0.85 (0.84)	0.84 (0.85)	0.86 (0.89)	0.87 (0.86)	0.89 (0.86)	0.87 (0.87)	0.90 (0.93)	0.89 (0.89)	1.04 (1.15)	12 (11)
LASSO	0.83 (0.83)	0.83 (0.82)	0.86 (0.87)	0.89 (0.94)	0.85 (0.89)	0.88 (0.93)	0.89 (1.00)	0.89 (0.86)	0.96 (0.96)	0.88 (0.91)	0.95 (1.02)	0.88 (0.95)	0.78 (0.81)	8 (6)
adaLASSO	0.84 (0.84)	0.84 (0.83)	0.86 (0.87)	0.85 (0.87)	0.81 (0.82)	0.83 (0.85)	0.86 (0.92)	0.88 (0.83)	0.93 (0.91)	0.83 (0.85)	0.90 (0.94)	0.87 (0.92)	0.82 (0.85)	10 (11)
ElNet	0.80 (0.81)	0.83 (0.83)	0.87 (0.89)	0.90 (0.97)	0.86 (0.92)	0.89 (0.96)	0.92 (1.02)	0.91 (0.88)	0.95 (0.96)	0.88 (0.92)	1.00 (1.08)	0.91 (0.98)	0.80 (0.83)	7 (2)
adaElnet	0.85 (0.86)	0.84 (0.84)	0.86 (0.87)	0.86 (0.90)	0.80 (0.82)	0.84 (0.87)	0.86 (0.93)	0.88 (0.84)	0.95 (0.93)	0.84 (0.87)	0.92 (0.97)	0.88 (0.94)	0.80 (0.83)	8 (6)
Ridge	0.82 (0.82)	0.77 (0.75)	0.85 (0.86)	0.83 (0.85)	0.78 (0.82)	0.82 (0.87)	0.83 (0.90)	0.84 (0.80)	0.90 (0.90)	0.83 (0.87)	0.91 (0.96)	0.84 (0.87)	0.76 (0.80)	13 (13)
BVAR	0.99 (1.00)	0.83 (0.80)	0.95 (1.00)	0.91 (0.99)	0.87 (0.95)	0.94 (1.05)	1.02 (1.14)	1.07 (1.08)	1.15 (1.22)	1.05 (1.15)	1.16 (1.29)	1.10 (1.22)	1.11 (1.17)	6 (2)
Bagging	0.85 (0.85)	0.82 (0.80)	0.94 (0.95)	0.90 (0.94)	0.86 (0.93)	0.86 (0.92)	0.84 (0.91)	0.83 (0.80)	0.91 (0.90)	0.86 (0.88)	0.97 (1.00)	0.89 (0.91)	1.09 (1.19)	12 (9)
CSR	0.83 (0.83)	0.83 (0.81)	0.86 (0.87)	0.81 (0.83)	0.76 (0.78)	0.80 (0.80)	0.84 (0.85)	0.81 (0.76)	0.87 (0.86)	0.83 (0.86)	0.90 (0.93)	0.85 (0.89)	0.81 (0.87)	13 (13)
JMA	0.94 (0.97)	1.00 (1.03)	1.06 (1.10)	1.04 (1.06)	1.00 (1.06)	0.92 (0.95)	1.02 (1.11)	1.01 (0.92)	1.17 (1.15)	1.00 (1.03)	1.08 (1.06)	1.00 (0.99)	1.09 (1.16)	4 (2)
Factor	0.89 (0.91)	0.89 (0.89)	0.96 (1.01)	0.88 (0.89)	0.86 (0.87)	0.92 (0.95)	0.91 (0.98)	0.92 (0.89)	1.05 (1.09)	0.98 (1.04)	1.05 (1.10)	1.04 (1.13)	0.96 (1.04)	3 (1)
T. Factor	0.95 (0.96)	0.91 (0.91)	1.01 (1.07)	0.84 (0.85)	0.83 (0.84)	0.88 (0.91)	0.89 (0.98)	0.88 (0.86)	1.00 (1.01)	0.96 (1.01)	0.99 (1.03)	0.97 (1.02)	0.94 (1.02)	6 (3)
Boosting	0.99 (1.03)	0.90 (0.91)	1.01 (1.09)	0.96 (1.02)	0.91 (0.97)	0.98 (1.05)	1.00 (1.09)	1.04 (1.01)	1.06 (1.06)	0.94 (0.96)	0.99 (1.03)	0.92 (0.94)	1.09 (1.30)	6 (4)
RF	0.82 (0.82)	0.77 (0.77)	0.86 (0.90)	0.83 (0.87)	0.78 (0.81)	0.80 (0.85)	0.80 (0.85)	0.81 (0.76)	0.84 (0.82)	0.78 (0.80)	0.85 (0.90)	0.79 (0.82)	0.67 (0.63)	13 (12)
Mean	0.82 (0.83)	0.79 (0.79)	0.84 (0.85)	0.82 (0.84)	0.78 (0.79)	0.80 (0.83)	0.81 (0.88)	0.81 (0.77)	0.88 (0.86)	0.82 (0.84)	0.87 (0.92)	0.84 (0.88)	0.74 (0.74)	13 (13)
T.Mean	0.82 (0.83)	0.80 (0.79)	0.85 (0.86)	0.82 (0.84)	0.79 (0.80)	0.80 (0.83)	0.82 (0.89)	0.83 (0.79)	0.90 (0.88)	0.82 (0.85)	0.89 (0.94)	0.85 (0.90)	0.75 (0.76)	13 (13)
Median	0.82 (0.83)	0.80 (0.79)	0.84 (0.85)	0.82 (0.85)	0.79 (0.81)	0.81 (0.83)	0.83 (0.89)	0.84 (0.79)	0.90 (0.89)	0.82 (0.85)	0.89 (0.94)	0.85 (0.90)	0.75 (0.76)	13 (13)
RF/OLS	0.81 (0.81)	0.77 (0.76)	0.83 (0.84)	0.80 (0.82)	0.76 (0.79)	0.79 (0.84)	0.80 (0.87)	0.83 (0.78)	0.87 (0.86)	0.80 (0.84)	0.87 (0.93)	0.83 (0.87)	0.81 (0.89)	13 (13)
adaLASSO/RF	0.81 (0.82)	0.89 (0.87)	0.91 (0.94)	0.87 (0.90)	0.81 (0.83)	0.82 (0.87)	0.84 (0.89)	0.81 (0.77)	0.85 (0.84)	0.79 (0.83)	0.90 (0.94)	0.84 (0.87)	0.71 (0.71)	11 (10)
RMSE count	13	15	14	16	17	20	11	12	11	15	18	16	21	
MAE count	(12)	(14)	(11)	(11)	(13)	(13)	(11)	(11)	(12)	(15)	(18)	(13)	(16)	

Table 6.25: Forecasting Errors for the PCE from 1990 to 2000

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 2001-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Personal Consumer Expenditure 1990–2015														
RMSE/(MAE)	Forecasting Horizon											Acc.	RMSE count (MAE count)	
	1	2	3	4	5	6	7	8	9	10	11			12
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0 (0)
AR	0.91 (0.88)	0.84 (0.82)	0.81 (0.75)	0.84 (0.82)	0.82 (0.83)	0.80 (0.79)	0.79 (0.74)	0.77 (0.72)	0.80 (0.78)	0.83 (0.81)	0.85 (0.83)	0.77 (0.74)	0.99 (0.93)	8 (0)
UCSV	0.96 (0.94)	0.86 (0.87)	0.84 (0.81)	0.86 (0.84)	0.83 (0.84)	0.83 (0.83)	0.82 (0.82)	0.81 (0.81)	0.82 (0.83)	0.83 (0.83)	0.85 (0.88)	0.81 (0.80)	0.81 (0.81)	3 (1)
LASSO	0.83 (0.79)	0.77 (0.74)	0.73 (0.67)	0.77 (0.74)	0.76 (0.77)	0.76 (0.72)	0.75 (0.69)	0.75 (0.68)	0.78 (0.73)	0.80 (0.76)	0.81 (0.78)	0.73 (0.68)	0.79 (0.76)	13 (11)
adaLASSO	0.84 (0.80)	0.77 (0.76)	0.74 (0.69)	0.78 (0.77)	0.78 (0.78)	0.77 (0.74)	0.76 (0.71)	0.77 (0.70)	0.79 (0.74)	0.81 (0.78)	0.83 (0.81)	0.73 (0.70)	0.83 (0.80)	13 (10)
ElNet	0.84 (0.80)	0.76 (0.74)	0.72 (0.67)	0.76 (0.73)	0.76 (0.76)	0.75 (0.72)	0.74 (0.68)	0.74 (0.67)	0.78 (0.74)	0.80 (0.76)	0.81 (0.78)	0.72 (0.68)	0.78 (0.76)	13 (12)
adaElnet	0.84 (0.81)	0.78 (0.76)	0.74 (0.68)	0.78 (0.76)	0.79 (0.78)	0.77 (0.74)	0.76 (0.71)	0.76 (0.69)	0.79 (0.75)	0.81 (0.78)	0.82 (0.81)	0.73 (0.70)	0.83 (0.80)	13 (9)
Ridge	0.87 (0.84)	0.76 (0.74)	0.73 (0.67)	0.77 (0.75)	0.77 (0.78)	0.75 (0.75)	0.74 (0.69)	0.73 (0.68)	0.76 (0.74)	0.77 (0.74)	0.78 (0.76)	0.71 (0.70)	0.77 (0.76)	13 (11)
BVAR	0.85 (0.83)	0.79 (0.77)	0.75 (0.70)	0.78 (0.76)	0.77 (0.77)	0.77 (0.74)	0.76 (0.69)	0.76 (0.70)	0.79 (0.75)	0.81 (0.76)	0.82 (0.78)	0.76 (0.71)	0.85 (0.80)	12 (12)
Bagging	0.88 (0.87)	0.76 (0.75)	0.74 (0.69)	0.81 (0.82)	0.81 (0.88)	0.79 (0.83)	0.82 (0.80)	0.79 (0.80)	0.81 (0.84)	0.80 (0.80)	0.81 (0.82)	0.73 (0.72)	0.68 (0.70)	11 (5)
CSR	0.86 (0.81)	0.77 (0.76)	0.75 (0.68)	0.79 (0.79)	0.79 (0.80)	0.79 (0.77)	0.79 (0.73)	0.77 (0.71)	0.81 (0.77)	0.84 (0.81)	0.86 (0.84)	0.80 (0.77)	0.92 (0.87)	10 (6)
JMA	0.96 (0.92)	0.84 (0.84)	0.78 (0.75)	0.85 (0.92)	0.83 (0.93)	0.82 (0.89)	0.84 (0.82)	0.82 (0.83)	0.86 (0.89)	0.82 (0.87)	0.82 (0.91)	0.77 (0.83)	0.72 (0.73)	6 (1)
Factor	0.89 (0.85)	0.83 (0.82)	0.80 (0.75)	0.83 (0.82)	0.82 (0.83)	0.81 (0.80)	0.81 (0.78)	0.81 (0.77)	0.83 (0.80)	0.83 (0.80)	0.84 (0.81)	0.80 (0.77)	0.95 (0.88)	4 (1)
T. Factor	0.88 (0.88)	0.80 (0.81)	0.76 (0.72)	0.80 (0.78)	0.80 (0.82)	0.79 (0.80)	0.78 (0.74)	0.78 (0.74)	0.79 (0.77)	0.80 (0.78)	0.82 (0.80)	0.78 (0.76)	0.91 (0.86)	8 (2)
Boosting	1.00 (0.98)	0.80 (0.79)	0.77 (0.73)	0.81 (0.82)	0.79 (0.84)	0.80 (0.82)	0.80 (0.79)	0.79 (0.78)	0.82 (0.85)	0.84 (0.84)	0.85 (0.87)	0.79 (0.78)	0.87 (0.93)	11 (3)
RF	0.88 (0.82)	0.76 (0.73)	0.71 (0.66)	0.74 (0.73)	0.72 (0.74)	0.71 (0.71)	0.70 (0.67)	0.70 (0.65)	0.73 (0.70)	0.74 (0.71)	0.76 (0.72)	0.70 (0.64)	0.67 (0.63)	13 (13)
Mean	0.85 (0.81)	0.76 (0.75)	0.74 (0.68)	0.77 (0.75)	0.77 (0.77)	0.76 (0.74)	0.75 (0.71)	0.74 (0.68)	0.76 (0.73)	0.77 (0.73)	0.78 (0.76)	0.72 (0.69)	0.81 (0.78)	13 (11)
T. Mean	0.84 (0.80)	0.76 (0.75)	0.73 (0.67)	0.77 (0.75)	0.76 (0.77)	0.75 (0.73)	0.74 (0.69)	0.74 (0.67)	0.76 (0.72)	0.78 (0.74)	0.79 (0.76)	0.72 (0.68)	0.80 (0.77)	13 (12)
Median	0.84 (0.80)	0.76 (0.75)	0.73 (0.68)	0.77 (0.75)	0.76 (0.76)	0.76 (0.74)	0.74 (0.69)	0.74 (0.67)	0.76 (0.72)	0.78 (0.74)	0.79 (0.76)	0.72 (0.68)	0.80 (0.77)	13 (13)
RF/OLS	0.83 (0.80)	0.76 (0.75)	0.74 (0.68)	0.78 (0.76)	0.76 (0.78)	0.75 (0.75)	0.74 (0.70)	0.77 (0.69)	0.78 (0.76)	0.81 (0.77)	0.74 (0.81)	0.74 (0.74)	0.82 (0.80)	13 (11)
adaLASSO/RF	0.86 (0.81)	0.75 (0.73)	0.71 (0.65)	0.73 (0.73)	0.73 (0.74)	0.72 (0.70)	0.69 (0.66)	0.70 (0.64)	0.74 (0.70)	0.76 (0.72)	0.78 (0.75)	0.71 (0.67)	0.69 (0.67)	13 (13)
RMSE count	15	16	18	17	20	16	15	15	18	20	17	13	16	
MAE count	(12)	(15)	(15)	(9)	(12)	(12)	(11)	(13)	(3)	(16)	(10)	(13)	(16)	

Table 6.26: Forecasting Errors for the PCE from 2001 to 2015

PUC-Rio - Certificação Digital N° 1421640/CA

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	2	2	2	0.046	0.065	0.159	0.155	0.141	0.140	1	1
AR	0.835	0.847	0.747	0	0	0	0.177	0.120	0.405	0.324	0.153	0.090	0.211	0.211
UCSV	0.891	0.887	0.882	0	0	0	0.080	0.081	0.100	0.104	0.001	0.001	0	0
LASSO	0.882	0.892	0.780	0	0	0	0.000	0.000	0.031	0.018	0.000	0.000	0	0
adaLASSO	0.826	0.817	0.766	0	0	0	0.037	0.027	0.192	0.209	0.020	0.015	0	0
EINet	0.895	0.905	0.792	0	0	0	0.000	0.000	0.030	0.015	0.000	0.000	0	0
adaElnet	0.833	0.830	0.778	0	0	0	0.008	0.008	0.141	0.150	0.003	0.003	0	0
Ridge	0.881	0.890	0.835	0	0	0	0.002	0.002	0.068	0.094	0.000	0.000	0	0
BVAR	1.126	1.143	0.937	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000	0	0
Bagging	0.787	0.765	0.746	3	3	2	0.387	0.354	0.607	0.631	0.386	0.411	1	1
CSR	0.800	0.797	0.732	1	0	1	0.481	0.469	0.578	0.523	0.493	0.414	1	1
JMA	0.902	0.871	0.814	0	0	0	0.013	0.042	0.107	0.178	0.002	0.010	0	0
Factor	0.877	0.878	0.805	0	0	0	0.034	0.050	0.219	0.215	0.008	0.015	0	0
T. Factor	0.895	0.894	0.842	0	0	0	0.067	0.092	0.268	0.237	0.011	0.013	0	0
Boosting	1.080	1.099	1.071	0	0	0	0.001	0.000	0.011	0.005	0.000	0.000	0	0
RF	0.749	0.751	0.681	6	5	6	0.598	0.605	0.739	0.725	0.695	0.665	1	1
Mean	0.792	0.795	0.742	0	0	0	0.340	0.228	0.436	0.391	0.363	0.293	0.999	1
T.Mean	0.801	0.806	0.737	0	0	0	0.116	0.062	0.295	0.307	0.122	0.049	1	1
Median	0.803	0.807	0.734	0	0	0	0.125	0.083	0.293	0.314	0.123	0.080	0.984	0.984
RF/OLS	0.777	0.771	0.746	0	0	0	0.347	0.335	0.539	0.526	0.402	0.377	1	1
adaLASSO/RF	0.771	0.758	0.700	1	3	2	0.547	0.588	0.531	0.610	0.378	0.464	0.965	0.965

Table 6.27: Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 1990–2015

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	2	2	2	0.076	0.079	0.188	0.186	0.149	0.144	1	1
AR	0.872	0.906	0.774	0	0	0	0.150	0.109	0.379	0.318	0.103	0.062	0	0.001
UCSV	0.881	0.888	0.938	0	0	0	0.086	0.087	0.156	0.202	0.002	0.004	0.0	0
LASSO	0.906	0.937	0.804	0	0	0	0.002	0.002	0.108	0.120	0.001	0.003	0	0
adaLASSO	0.826	0.835	0.777	0	0	0	0.109	0.077	0.324	0.303	0.086	0.085	0.022	0.022
EINet	0.926	0.959	0.830	0	0	0	0.001	0.001	0.120	0.123	0.000	0.001	0.001	0.001
adaElnet	0.838	0.857	0.796	0	0	0	0.066	0.035	0.244	0.237	0.039	0.038	0.005	0.005
Ridge	0.871	0.894	0.871	0	0	0	0.046	0.044	0.200	0.234	0.048	0.028	0.002	0.002
BVAR	1.252	1.327	0.867	0	0	0	0.000	0.000	0.001	0.001	0.000	0.000	0	0
Bagging	0.768	0.792	0.737	4	6	5	0.317	0.325	0.681	0.648	0.457	0.493	1	1
CSR	0.782	0.792	0.697	4	1	4	0.749	0.745	0.697	0.687	0.723	0.652	1	1
JMA	0.888	0.867	0.845	0	0	0	0.091	0.215	0.294	0.359	0.043	0.070	0.020	0.020
Factor	0.954	0.971	0.869	0	0	0	0.069	0.102	0.225	0.225	0.010	0.028	0	0
T. Factor	1.019	1.055	0.841	0	0	0	0.044	0.056	0.183	0.148	0.006	0.006	0	0
Boosting	1.075	1.093	1.014	0	0	0	0.031	0.060	0.095	0.138	0.001	0.004	0	0
RF	0.767	0.782	0.709	3	2	1	0.521	0.478	0.687	0.617	0.594	0.542	0.993	1
Mean	0.822	0.842	0.753	0	0	0	0.165	0.142	0.396	0.309	0.210	0.171	0.833	0.837
T. Mean	0.829	0.852	0.754	0	0	0	0.082	0.071	0.383	0.271	0.056	0.066	0.327	0.327
Median	0.828	0.850	0.751	0	0	0	0.079	0.126	0.363	0.266	0.064	0.111	0.263	0.263
RF/OLS	0.777	0.784	0.742	0	0	0	0.301	0.309	0.543	0.479	0.342	0.370	0.837	0.837
adaLASSO/RF	0.785	0.775	0.733	0	2	1	0.492	0.576	0.594	0.635	0.424	0.517	0.766	0.837

Table 6.28: Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 1990–2000

The table reports, for each model, a number of different summary statistics across all the forecasting horizons, including as well the accumulated twelve-month horizon. Columns (1), (2) and (3) report the average root mean square error (RMSE), the average mean absolute error (MAE) and the average median absolute deviation (MAD). Columns (4), (5) and (6) report, respectively, the number of times (across horizons) each model achieved the lowest RMSE, MAE, and MAD. Columns (7)–(10) present, for square and absolute losses, the average p -values based either on the range or the t_{\max} statistics as described in (55). Columns (11) and (12) show the average p -values of the Superior Predictive Ability (SPA) test proposed by (83). Finally, columns (13) and (14) display the p -value of the multi-horizon test for superior predictive ability proposed by (84).

Model	Forecasting Precision						Model Confidence Set				Sup. Pred. Ability		Multi-Horizon SPA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	ave. RMSE	ave. MAE	ave. MAD	# min. RMSE	# min. MAE	# min. MAD	ave. p.v. range sq	ave. p.v. range abs	ave. p.v. Tmax sq	ave p.v. Tmax abs	ave. p.v. sq	ave. p.v. abs	p.v. model	p.v. test
RW	1.000	1.000	1.000	1	1	1	0.100	0.117	0.139	0.117	0.139	0.139	1	1
AR	0.804	0.803	0.712	0	0	1	0.347	0.389	0.615	0.591	0.284	0.291	0.348	0.368
UCSV	0.889	0.888	0.889	1	1	1	0.138	0.171	0.267	0.291	0.079	0.102	0.999	0.999
LASSO	0.860	0.858	0.761	0	0	0	0.002	0.006	0.100	0.096	0.001	0.000	0	0
adaLASSO	0.822	0.803	0.746	0	0	0	0.199	0.345	0.385	0.487	0.097	0.178	0.222	0.222
ElNet	0.870	0.865	0.756	0	0	0	0.000	0.002	0.085	0.064	0.000	0.000	0	0
adaElnet	0.824	0.808	0.738	0	0	0	0.102	0.288	0.328	0.448	0.049	0.145	0.124	0.124
Ridge	0.890	0.888	0.831	0	0	0	0.007	0.011	0.104	0.098	0.000	0.001	0	0
BVAR	1.017	1.008	0.828	0	0	0	0.002	0.000	0.032	0.014	0.001	0.000	0	0
Bagging	0.802	0.776	0.747	2	3	1	0.572	0.535	0.702	0.710	0.410	0.418	1	1
CSR	0.814	0.802	0.753	0	0	0	0.504	0.508	0.580	0.626	0.383	0.358	0.947	0.947
JMA	0.914	0.875	0.785	0	0	0	0.110	0.139	0.220	0.260	0.019	0.027	0	0
Factor	0.812	0.809	0.761	0	0	0	0.235	0.324	0.483	0.542	0.157	0.200	0.014	0.014
T. Factor	0.781	0.774	0.757	0	1	0	0.554	0.534	0.725	0.730	0.435	0.404	0.368	0.368
Boosting	1.087	1.105	1.127	0	0	0	0.006	0.003	0.040	0.017	0.000	0.000	0	0
RF	0.736	0.728	0.676	5	6	7	0.677	0.758	0.851	0.872	0.765	0.827	1	1
Mean	0.768	0.761	0.710	2	0	0	0.518	0.569	0.652	0.670	0.539	0.571	1	1
T. Mean	0.778	0.773	0.719	0	0	0	0.227	0.264	0.563	0.544	0.209	0.249	0.999	1
Median	0.782	0.776	0.719	0	0	0	0.241	0.297	0.511	0.552	0.192	0.266	0.508	0.508
RF/OLS	0.775	0.762	0.740	0	0	0	0.653	0.630	0.760	0.759	0.591	0.543	1	1
adaLASSO/RF	0.759	0.744	0.683	2	1	2	0.624	0.713	0.810	0.801	0.498	0.600	1	1

Table 6.29: Forecasting Results CPI-Core: Summary statistics for the out-of-sample period from 2001–2015

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Consumer Price Index (Core) 1990-2015															
RMSE/(MAE)	Forecasting Horizon											Acc.	RMSE count (MAE count)		
	1	2	3	4	5	6	7	8	9	10	11			12	
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	2
AR	0.87 (0.92)	0.69 (0.70)	0.59 (0.59)	0.62 (0.62)	0.79 (0.80)	0.93 (0.91)	0.90 (0.85)	0.68 (0.63)	0.47 (0.45)	0.49 (0.49)	0.67 (0.70)	1.21 (1.27)	1.96 (2.09)	4	
UCSV	0.96 (1.03)	0.80 (0.80)	0.76 (0.74)	0.77 (0.76)	0.86 (0.87)	0.96 (0.93)	0.93 (0.90)	0.84 (0.80)	0.76 (0.74)	0.73 (0.74)	0.82 (0.84)	1.38 (1.38)	1.02 (1.01)	1	
LASSO	0.85 (0.92)	0.66 (0.66)	0.62 (0.60)	0.69 (0.68)	0.86 (0.85)	0.97 (0.93)	0.90 (0.86)	0.72 (0.68)	0.54 (0.52)	0.58 (0.57)	0.80 (0.83)	1.39 (1.43)	1.90 (2.06)	0	
adaLASSO	0.84 (0.89)	0.65 (0.65)	0.59 (0.57)	0.66 (0.65)	0.84 (0.82)	0.94 (0.90)	0.89 (0.85)	0.73 (0.68)	0.51 (0.48)	0.52 (0.51)	0.72 (0.74)	1.24 (1.27)	1.64 (1.62)	0	
ElNet	0.86 (0.92)	0.67 (0.67)	0.61 (0.60)	0.68 (0.67)	0.86 (0.85)	0.98 (0.94)	0.89 (0.85)	0.72 (0.68)	0.55 (0.53)	0.59 (0.58)	0.83 (0.85)	1.46 (1.49)	1.94 (2.12)	0	
adaElnet	0.85 (0.90)	0.66 (0.66)	0.60 (0.58)	0.68 (0.67)	0.83 (0.82)	0.94 (0.90)	0.90 (0.85)	0.72 (0.68)	0.50 (0.48)	0.52 (0.52)	0.73 (0.75)	1.26 (1.30)	1.66 (1.69)	0	
Ridge	0.92 (0.99)	0.66 (0.67)	0.61 (0.60)	0.67 (0.66)	0.85 (0.85)	0.95 (0.93)	0.84 (0.81)	0.68 (0.65)	0.59 (0.58)	0.62 (0.63)	0.87 (0.92)	1.56 (1.62)	1.63 (1.69)	0	
BVAR	0.91 (0.97)	0.71 (0.71)	0.70 (0.67)	0.79 (0.76)	1.01 (0.99)	1.13 (1.08)	1.01 (0.96)	0.83 (0.79)	0.75 (0.73)	0.80 (0.80)	1.10 (1.15)	1.93 (1.99)	2.97 (3.26)	0	
Bagging	0.81 (0.81)	0.57 (0.56)	0.55 (0.52)	0.67 (0.65)	0.91 (0.87)	0.96 (0.88)	0.81 (0.77)	0.65 (0.62)	0.48 (0.46)	0.50 (0.49)	0.71 (0.73)	1.28 (1.28)	1.33 (1.31)	7	
CSR	0.82 (0.86)	0.65 (0.65)	0.57 (0.56)	0.62 (0.60)	0.78 (0.77)	0.89 (0.87)	0.85 (0.82)	0.66 (0.61)	0.46 (0.45)	0.48 (0.47)	0.65 (0.68)	1.21 (1.25)	1.77 (1.78)	8	
JMA	0.91 (0.92)	0.68 (0.65)	0.68 (0.65)	0.75 (0.74)	1.05 (0.98)	1.06 (0.99)	0.96 (0.91)	0.74 (0.68)	0.57 (0.53)	0.55 (0.53)	0.82 (0.81)	1.45 (1.46)	1.50 (1.48)	0	
Factor	0.90 (0.95)	0.71 (0.71)	0.62 (0.61)	0.68 (0.67)	0.86 (0.87)	0.96 (0.93)	0.90 (0.86)	0.73 (0.68)	0.47 (0.46)	0.52 (0.51)	0.73 (0.75)	1.26 (1.29)	2.07 (2.13)	1	
T. Factor	0.90 (0.94)	0.72 (0.72)	0.64 (0.62)	0.69 (0.68)	0.86 (0.86)	0.97 (0.95)	0.92 (0.88)	0.75 (0.69)	0.47 (0.47)	0.51 (0.51)	0.71 (0.75)	1.25 (1.31)	2.25 (2.25)	1	
Boosting	0.94 (1.00)	0.70 (0.70)	0.69 (0.67)	0.79 (0.78)	1.02 (1.02)	1.10 (1.07)	0.98 (0.94)	0.77 (0.75)	0.64 (0.64)	0.72 (0.73)	0.99 (1.05)	1.67 (1.74)	3.03 (3.19)	0	
RF	0.84 (0.89)	0.61 (0.61)	0.56 (0.54)	0.62 (0.60)	0.78 (0.78)	0.87 (0.84)	0.80 (0.77)	0.64 (0.61)	0.43 (0.42)	0.45 (0.45)	0.63 (0.66)	1.14 (1.16)	1.36 (1.45)	9	
Mean	0.81 (0.86)	0.64 (0.63)	0.58 (0.57)	0.64 (0.63)	0.79 (0.79)	0.88 (0.86)	0.81 (0.78)	0.66 (0.62)	0.50 (0.49)	0.52 (0.52)	0.69 (0.72)	1.19 (1.22)	1.59 (1.66)	5	
T.Mean	0.83 (0.88)	0.64 (0.65)	0.58 (0.57)	0.63 (0.63)	0.79 (0.79)	0.89 (0.87)	0.83 (0.80)	0.67 (0.63)	0.49 (0.47)	0.51 (0.51)	0.70 (0.72)	1.21 (1.24)	1.64 (1.73)	2	
Median	0.83 (0.88)	0.65 (0.65)	0.58 (0.57)	0.64 (0.63)	0.79 (0.79)	0.89 (0.87)	0.84 (0.80)	0.68 (0.64)	0.49 (0.47)	0.51 (0.51)	0.70 (0.72)	1.22 (1.25)	1.64 (1.73)	2	
RF/OLS	0.82 (0.88)	0.63 (0.63)	0.57 (0.55)	0.64 (0.62)	0.80 (0.80)	0.89 (0.86)	0.81 (0.77)	0.65 (0.61)	0.46 (0.44)	0.48 (0.48)	0.66 (0.69)	1.19 (1.22)	1.51 (1.48)	7	
adaLASSO/RF	0.84 (0.89)	0.62 (0.62)	0.56 (0.54)	0.61 (0.60)	0.78 (0.76)	0.91 (0.86)	0.89 (0.82)	0.69 (0.64)	0.45 (0.43)	0.46 (0.45)	0.65 (0.66)	1.16 (1.15)	1.40 (1.43)	7	
RMSE count	3	1	5	4	8	8	4	5	8	2	5	1	2		
MAE count	(10)	(1)	(3)	(3)	(3)	(8)	(4)	(12)	(8)	(2)	(3)	(1)	(2)		

Table 6.30: Forecasting Errors for the CPI-Core from 1990 to 2015

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Consumer Price Index (Core) 1990-2000															
RMSE/(MAE)	Forecasting Horizon											Acc.	RMSE count (MAE count)		
	1	2	3	4	5	6	7	8	9	10	11			12	
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	2
AR	0.88 (0.92)	0.69 (0.70)	0.59 (0.58)	0.62 (0.62)	0.79 (0.82)	0.96 (0.96)	0.91 (0.90)	0.71 (0.65)	0.50 (0.48)	0.54 (0.53)	0.71 (0.76)	1.26 (1.43)	2.16 (2.38)	5	
UCSV	0.91 (0.98)	0.68 (0.67)	0.64 (0.60)	0.72 (0.69)	0.90 (0.93)	1.04 (1.02)	0.91 (0.91)	0.75 (0.70)	0.66 (0.61)	0.68 (0.70)	0.90 (0.95)	1.61 (1.75)	1.06 (1.03)	1	
LASSO	0.85 (0.91)	0.65 (0.64)	0.61 (0.58)	0.67 (0.65)	0.67 (0.88)	0.87 (0.98)	1.01 (0.90)	0.91 (0.68)	0.73 (0.56)	0.59 (0.63)	0.63 (0.92)	0.87 (1.59)	1.44 (2.18)	1.91 (1)	
adaLASSO	0.85 (0.90)	0.63 (0.63)	0.56 (0.53)	0.63 (0.61)	0.83 (0.83)	0.97 (0.95)	0.90 (0.90)	0.75 (0.70)	0.56 (0.53)	0.56 (0.56)	0.77 (0.80)	1.27 (1.38)	1.39 (1.47)	3	
ElNet	0.85 (0.91)	0.66 (0.65)	0.61 (0.58)	0.68 (0.66)	0.87 (0.88)	1.02 (0.99)	0.90 (0.90)	0.73 (0.67)	0.59 (0.56)	0.64 (0.65)	0.89 (0.94)	1.55 (1.70)	2.00 (2.30)	0	
adaElnet	0.85 (0.91)	0.64 (0.64)	0.58 (0.54)	0.65 (0.63)	0.82 (0.83)	0.97 (0.95)	0.90 (0.90)	0.74 (0.69)	0.55 (0.52)	0.57 (0.58)	0.79 (0.83)	1.32 (1.45)	1.47 (1.61)	1	
Ridge	0.88 (0.95)	0.65 (0.65)	0.60 (0.57)	0.66 (0.63)	0.84 (0.85)	0.99 (0.97)	0.85 (0.85)	0.69 (0.65)	0.59 (0.58)	0.63 (0.65)	0.86 (0.92)	1.51 (1.62)	1.52 (1.66)	1	
BVAR	0.91 (0.97)	0.72 (0.71)	0.72 (0.68)	0.84 (0.80)	1.09 (1.10)	1.28 (1.26)	1.10 (1.10)	0.90 (0.85)	0.81 (0.79)	0.88 (0.90)	1.19 (1.27)	2.04 (2.26)	3.73 (4.49)	0	
Bagging	0.77 (0.79)	0.57 (0.55)	0.55 (0.50)	0.65 (0.62)	0.82 (0.76)	0.93 (0.87)	0.79 (0.77)	0.65 (0.59)	0.52 (0.47)	0.56 (0.54)	0.74 (0.74)	1.19 (1.25)	1.22 (1.19)	8	
CSR	0.81 (0.85)	0.65 (0.64)	0.55 (0.53)	0.59 (0.57)	0.76 (0.77)	0.90 (0.89)	0.83 (0.82)	0.66 (0.61)	0.47 (0.45)	0.49 (0.48)	0.66 (0.67)	1.15 (1.22)	1.56 (1.68)	10	
JMA	0.85 (0.86)	0.72 (0.68)	0.62 (0.58)	0.73 (0.71)	1.00 (0.95)	1.11 (1.04)	0.96 (0.92)	0.70 (0.63)	0.58 (0.54)	0.57 (0.56)	0.78 (0.77)	1.44 (1.54)	1.40 (1.41)	2	
Factor	0.88 (0.91)	0.75 (0.74)	0.65 (0.61)	0.70 (0.68)	0.89 (0.91)	1.04 (1.03)	0.93 (0.92)	0.82 (0.77)	0.52 (0.50)	0.57 (0.57)	0.78 (0.82)	1.36 (1.48)	2.49 (2.63)	1	
T. Factor	0.92 (0.96)	0.81 (0.80)	0.72 (0.69)	0.74 (0.72)	0.91 (0.94)	1.07 (1.08)	0.99 (0.99)	0.86 (0.81)	0.52 (0.51)	0.58 (0.59)	0.78 (0.83)	1.35 (1.53)	2.91 (3.14)	1	
Boosting	0.92 (0.98)	0.69 (0.69)	0.68 (0.66)	0.80 (0.77)	1.03 (1.05)	1.15 (1.13)	0.97 (0.96)	0.77 (0.72)	0.63 (0.62)	0.73 (0.74)	0.96 (1.03)	1.55 (1.70)	2.96 (2.93)	0	
RF	0.83 (0.90)	0.61 (0.61)	0.56 (0.54)	0.62 (0.60)	0.79 (0.79)	0.92 (0.90)	0.82 (0.81)	0.67 (0.62)	0.46 (0.43)	0.48 (0.47)	0.64 (0.67)	1.14 (1.22)	1.40 (1.57)	10	
Mean	0.81 (0.86)	0.64 (0.63)	0.58 (0.55)	0.63 (0.61)	0.80 (0.82)	0.93 (0.92)	0.83 (0.83)	0.69 (0.64)	0.53 (0.51)	0.56 (0.56)	0.74 (0.77)	1.24 (1.35)	1.65 (1.81)	4	
T.Mean	0.82 (0.87)	0.65 (0.64)	0.58 (0.55)	0.63 (0.62)	0.80 (0.82)	0.94 (0.93)	0.84 (0.84)	0.69 (0.64)	0.52 (0.50)	0.55 (0.56)	0.74 (0.78)	1.26 (1.38)	1.68 (1.86)	4	
Median	0.83 (0.88)	0.65 (0.64)	0.58 (0.56)	0.63 (0.62)	0.80 (0.81)	0.93 (0.92)	0.85 (0.84)	0.69 (0.64)	0.52 (0.50)	0.56 (0.56)	0.74 (0.78)	1.26 (1.39)	1.66 (1.83)	4	
RF/OLS	0.83 (0.89)	0.63 (0.63)	0.56 (0.53)	0.62 (0.59)	0.79 (0.80)	0.93 (0.90)	0.83 (0.83)	0.68 (0.64)	0.49 (0.46)	0.50 (0.50)	0.68 (0.69)	1.18 (1.24)	1.35 (1.42)	7	
adaLASSO/RF	0.85 (0.93)	0.64 (0.64)	0.55 (0.51)	0.61 (0.58)	0.77 (0.77)	0.93 (0.89)	0.91 (0.87)	0.72 (0.65)	0.49 (0.43)	0.49 (0.47)	0.66 (0.66)	1.17 (1.21)	1.37 (1.39)	7	
RMSE count	9	1	8	6	9	8	3	10	8	3	3	1	2		
MAE count	(3)	(13)	(5)	(7)	(5)	(5)	(1)	(9)	(6)	(3)	(3)	(1)	(2)		

Table 6.31: Forecasting Errors for the CPI-Core from 1990 to 2000

The table shows the root mean squared error (RMSE) and, between parenthesis, the mean absolute errors (MAE) for all models relative to the Random Walk (RW). The error measures were calculated from 132 rolling windows covering the 1990-2015 period and 180 rolling windows covering the 2001-2015 period. Values in bold show the most accurate model in each horizon. Cells in gray (blue) show the models included in the 50% model confidence set (MCS) using the squared error (absolute error) as loss function. The MCSs were constructed based on the maximum *t*-statistic. The last column in the table reports in how many horizons the row model was included in the MCS for square (absolute) loss. The last two rows in the table reports how many models were included in the MCS for square and absolute losses.

Consumer Price Index (Core) 1990-2015															
RMSE/(MAE)	Forecasting Horizon											Acc.	RMSE count (MAE count)		
	1	2	3	4	5	6	7	8	9	10	11			12	
RW	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	2
AR	0.86 (0.91)	0.68 (0.69)	0.60 (0.59)	0.62 (0.63)	0.79 (0.78)	0.90 (0.87)	0.89 (0.81)	0.66 (0.61)	0.43 (0.43)	0.45 (0.45)	0.62 (0.66)	1.16 (1.15)	1.80 (1.90)	1.80 (1.90)	8
UCSV	1.00 (1.07)	0.90 (0.91)	0.85 (0.85)	0.81 (0.82)	0.83 (0.82)	0.89 (0.87)	0.94 (0.90)	0.90 (0.87)	0.84 (0.84)	0.78 (0.77)	0.74 (0.75)	1.09 (1.09)	0.99 (1.00)	1.80 (1.85)	4
LASSO	0.86 (0.92)	0.66 (0.68)	0.62 (0.62)	0.71 (0.71)	0.85 (0.84)	0.94 (0.90)	0.89 (0.82)	0.71 (0.69)	0.49 (0.49)	0.53 (0.52)	0.73 (0.75)	1.34 (1.31)	1.85 (1.98)	1.85 (1.98)	0
adaLASSO	0.84 (0.88)	0.66 (0.66)	0.60 (0.60)	0.69 (0.69)	0.85 (0.82)	0.91 (0.87)	0.88 (0.81)	0.71 (0.67)	0.45 (0.44)	0.48 (0.46)	0.66 (0.69)	1.20 (1.18)	1.74 (1.72)	1.74 (1.72)	4
ElNet	0.87 (0.94)	0.68 (0.69)	0.62 (0.62)	0.69 (0.69)	0.85 (0.83)	0.95 (0.91)	0.88 (0.82)	0.72 (0.68)	0.52 (0.51)	0.54 (0.52)	0.76 (0.78)	1.37 (1.32)	1.86 (2.00)	1.86 (2.00)	0
adaElnet	0.84 (0.89)	0.67 (0.68)	0.62 (0.61)	0.70 (0.70)	0.84 (0.81)	0.91 (0.87)	0.89 (0.81)	0.71 (0.67)	0.45 (0.44)	0.48 (0.47)	0.67 (0.69)	1.20 (1.17)	1.74 (1.74)	1.74 (1.74)	3
Ridge	0.95 (1.01)	0.68 (0.68)	0.62 (0.62)	0.68 (0.68)	0.86 (0.85)	0.92 (0.90)	0.84 (0.78)	0.68 (0.66)	0.58 (0.58)	0.61 (0.61)	0.87 (0.92)	1.61 (1.61)	1.67 (1.70)	1.67 (1.70)	0
BVAR	0.91 (0.97)	0.71 (0.70)	0.68 (0.66)	0.75 (0.73)	0.93 (0.90)	1.00 (0.95)	0.93 (0.86)	0.76 (0.74)	0.69 (0.69)	0.73 (0.73)	1.02 (1.05)	1.81 (1.78)	2.29 (2.49)	2.29 (2.49)	0
Bagging	0.84 (0.83)	0.58 (0.57)	0.55 (0.54)	0.69 (0.67)	0.98 (0.92)	0.99 (0.88)	0.82 (0.77)	0.66 (0.64)	0.45 (0.44)	0.46 (0.45)	0.68 (0.71)	1.36 (1.30)	1.38 (1.39)	1.38 (1.39)	10
CSR	0.82 (0.87)	0.65 (0.65)	0.59 (0.59)	0.64 (0.63)	0.80 (0.78)	0.88 (0.85)	0.87 (0.81)	0.66 (0.62)	0.45 (0.44)	0.47 (0.46)	0.65 (0.68)	1.27 (1.27)	1.84 (1.85)	1.84 (1.85)	8
JMA	0.97 (0.97)	0.65 (0.63)	0.72 (0.70)	0.77 (0.77)	1.09 (1.00)	1.02 (0.95)	0.96 (0.91)	0.77 (0.72)	0.57 (0.52)	0.53 (0.50)	0.86 (0.85)	1.46 (1.40)	1.52 (1.53)	1.52 (1.53)	1
Factor	0.91 (0.98)	0.68 (0.70)	0.60 (0.60)	0.65 (0.66)	0.84 (0.83)	0.90 (0.87)	0.88 (0.82)	0.65 (0.60)	0.43 (0.43)	0.48 (0.47)	0.68 (0.70)	1.15 (1.14)	1.71 (1.81)	1.71 (1.81)	6
T. Factor	0.88 (0.92)	0.63 (0.65)	0.57 (0.57)	0.64 (0.64)	0.81 (0.80)	0.89 (0.85)	0.86 (0.80)	0.63 (0.59)	0.42 (0.43)	0.45 (0.45)	0.63 (0.67)	1.13 (1.13)	1.61 (1.70)	1.61 (1.70)	10
Boosting	0.96 (1.01)	0.71 (0.71)	0.69 (0.68)	0.79 (0.79)	1.01 (1.00)	1.07 (1.03)	0.99 (0.93)	0.78 (0.78)	0.65 (0.66)	0.72 (0.73)	1.00 (1.07)	1.78 (1.78)	2.99 (3.35)	2.99 (3.35)	0
RF	0.86 (0.88)	0.61 (0.61)	0.55 (0.54)	0.62 (0.61)	0.78 (0.76)	0.83 (0.79)	0.78 (0.73)	0.62 (0.60)	0.41 (0.41)	0.43 (0.43)	0.62 (0.65)	1.13 (1.10)	1.32 (1.38)	1.32 (1.38)	11
Mean	0.81 (0.86)	0.63 (0.64)	0.58 (0.58)	0.64 (0.64)	0.78 (0.77)	0.84 (0.81)	0.80 (0.74)	0.64 (0.61)	0.47 (0.46)	0.49 (0.48)	0.65 (0.67)	1.14 (1.12)	1.51 (1.56)	1.51 (1.56)	8
T.Mean	0.83 (0.88)	0.64 (0.65)	0.58 (0.58)	0.64 (0.64)	0.78 (0.77)	0.86 (0.83)	0.82 (0.76)	0.66 (0.63)	0.46 (0.45)	0.48 (0.47)	0.65 (0.67)	1.16 (1.13)	1.58 (1.64)	1.58 (1.64)	7
Median	0.83 (0.88)	0.64 (0.66)	0.58 (0.58)	0.64 (0.64)	0.79 (0.77)	0.86 (0.83)	0.83 (0.77)	0.66 (0.63)	0.46 (0.45)	0.48 (0.47)	0.65 (0.67)	1.16 (1.14)	1.60 (1.66)	1.60 (1.66)	5
RF/OLS	0.82 (0.86)	0.63 (0.63)	0.57 (0.57)	0.65 (0.65)	0.81 (0.80)	0.86 (0.83)	0.79 (0.73)	0.63 (0.60)	0.44 (0.42)	0.46 (0.46)	0.65 (0.68)	1.20 (1.20)	1.57 (1.51)	1.57 (1.51)	11
adaLASSO/RF	0.83 (0.86)	0.61 (0.61)	0.57 (0.56)	0.62 (0.61)	0.80 (0.76)	0.89 (0.84)	0.86 (0.79)	0.66 (0.63)	0.42 (0.42)	0.43 (0.42)	0.64 (0.66)	1.16 (1.11)	1.39 (1.45)	1.39 (1.45)	11
RMSE count	9	6	5	11	10	14	4	10	9	7	13	9	2	2	
MAE count	(9)	(4)	(5)	(11)	(12)	(14)	(8)	(9)	(10)	(14)	(13)	(12)	(2)	(2)	

Table 6.32: Forecasting Errors for the CPI-Core from 2001 to 2015

References

- [1] HASTIE, T.; TIBSHIRAMI, R.; FRIEDMAN, J. *The elements of statistical learning; data mining, inference and prediction*. Springer, 2001.
- [2] TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 267–288, 1996.
- [3] ZOU, H. The adaptive lasso and its oracle properties. *Journal of the American statistical association*, v. 101, p. 1418–1429, 2006.
- [4] BAI, J.; NG, S. Determining the number of factors in approximate factor models. *Econometrica*, v. 70, p. 191–221, 2002.
- [5] BAI, J.; NG, S. Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, v. 74, n. 4, p. 1133–1150, 2006.
- [6] BAI, J.; NG, S. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, v. 146, n. 2, p. 304–317, 2008.
- [7] BAI, J.; NG, S. Boosting diffusion indices. *Journal of Applied Econometrics*, v. 24, n. 4, p. 607–629, 2009.
- [8] BREIMAN, L. Bagging predictors. *Machine learning*, v. 24, n. 2, p. 123–140, 1996.
- [9] BREIMAN, L. Random forests. *Machine learning*, v. 45, n. 1, p. 5–32, 2001.
- [10] BAÑBURA, M.; GIANNONE, D.; REICHLIN, L. Large bayesian vector auto regressions. *Journal of Applied Econometrics*, v. 25, n. 1, p. 71–92, 2010.

- [11] MEDEIROS, M. C.; VASCONCELOS, G.; FREITAS, E. Forecasting brazilian inflation with high-dimensional models. *Brazilian Review of Econometrics*, v. 36, n. 2, p. 223–254, 2016.
- [12] GARCIA, M. G.; MEDEIROS, M. C.; VASCONCELOS, G. F. Real-time inflation forecast with high-dimensional models: the case of brazil. *International Journal of Forecasting - Accepted for publication*.
- [13] MEDEIROS, M. C.; VASCONCELOS, G. F.; VEIGA, Á.; ZILBERMAN, E. Forecasting inflation with high-dimensional data. *Working Paper*.
- [14] HOERL, A. E.; KENNARD, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, v. 12, n. 1, p. 55–67, 1970.
- [15] HOERL, A. E.; KENNARD, R. W. Ridge regression: applications to nonorthogonal problems. *Technometrics*, v. 12, n. 1, p. 69–82, 1970.
- [16] ZHAO, P.; YU, B. On model selection consistency of lasso. *The Journal of Machine Learning Research*, v. 7, p. 2541–2563, 2006.
- [17] MEDEIROS, M.; MENDES, E. ℓ_1 -regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors. *Journal of Econometrics*, v. 191, p. 255–271, 2016.
- [18] ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 67, n. 2, p. 301–320, 2005.
- [19] ELLIOTT, G.; GARGANO, A.; TIMMERMANN, A. Complete subset regressions. *Journal of Econometrics*, v. 177, n. 2, p. 357–373, 2013.
- [20] ELLIOTT, G.; GARGANO, A.; TIMMERMANN, A. Complete subset regressions with large-dimensional sets of predictors. *Journal of Economic Dynamics and Control*, v. 54, p. 86–110, 2015.

- [21] HANSEN, B. E.; RACINE, J. S. Jackknife model averaging. *Journal of Econometrics*, v. 167, n. 1, p. 38–46, 2012.
- [22] ZHANG, X.; WAN, A. T.; ZOU, G. Model averaging by jackknife criterion in models with dependent data. *Journal of Econometrics*, v. 174, n. 2, p. 82–94, 2013.
- [23] MINELLA, A.; DE FREITAS, P. S.; GOLDFAJN, I.; MUINHOS, M. K. Inflation targeting in brazil: constructing credibility under exchange rate volatility. *Journal of international Money and Finance*, v. 22, p. 1015–1040, 2003.
- [24] BOGDANSKI, J.; TOMBINI, A. A.; WERLANG, S. R. D. C. Implementing inflation targeting in Brazil. *Banco Central do Brasil Working Paper*, , n. 1, 2000.
- [25] CARRARAA, A. F.; CORREAB, A. L. O regime de metas de inflação no brasil: uma análise empírica do ipca. *Revista de Economia Contemporânea*, v. 16, p. 441–462, 2012.
- [26] DE POOTER, M.; ROBITAILLE, P.; WALKER, I.; ZDINAK, M. Are long-term inflation expectations well anchored in brazil, chile and mexico? *International Journal of Central Banking*, v. 10, p. 337–400, 2014.
- [27] NUNES, M. S.; DA COSTA JR, N. C.; MEURER, R. A relação entre o mercado de ações e as variáveis macroeconômicas: uma análise econométrica para o brasil. *Revista Brasileira de Economia*, v. 59, p. 585–607, 2005.
- [28] MENDONÇA, M. J. C. D.; SACHSIDA, A.; MEDRANO, L. A. T. Inflação versus desemprego: novas evidências para o brasil. *Economia Aplicada*, v. 16, p. 475–500, 2012.
- [29] SACHSIDA, A. Inflação, desemprego e choques cambiais: Uma revisão da literatura sobre a curva de phillips no brasil. *Revista Brasileira de Economia*, v. 67, p. 549–559, 2013.

- [30] ARRUDA, E. F.; FERREIRA, R. T.; CASTELAR, I. Modelos lineares e não lineares da curva de phillips para previsão da taxa de inflação no brasil. *Revista Brasileira de Economia*, v. 65, p. 237–252, 2011.
- [31] FIGUEIREDO, E. A. D.; MARQUES, A. M. Inflação inercial como um processo de longa memória: análise a partir de um modelo arfima-figarch. *Estudos Econômicos (São Paulo)*, v. 39, p. 437–458, 2009.
- [32] KOHLSCHEEN, E. Uma nota sobre erros de previsão da inflação de curto-prazo. *Revista Brasileira de Economia*, v. 66, p. 289–297, 2012.
- [33] STOCK, J. H.; WATSON, M. W. Forecasting inflation. *Journal of Monetary Economics*, v. 44, p. 293–335, 1999.
- [34] ATKESON, A.; OHANIAN, L. E. Are phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis quarterly review*, v. 25, p. 2–11, 2001.
- [35] CHEN, Y.-C.; TURNOVSKY, S. J.; ZIVOT, E. Forecasting inflation using commodity price aggregates. *Journal of Econometrics*, v. 183, p. 117–134, 2014.
- [36] FORNI, M.; HALLIN, M.; LIPPI, M.; REICHLIN, L. Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, v. 50, p. 1243–1255, 2003.
- [37] GROEN, J. J.; PAAP, R.; RAVAZZOLO, F. Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics*, v. 31, p. 29–44, 2013.
- [38] MEINSHAUSEN, N.; YU, B. Lasso-type recovery of sparse representations for high-dimensional data. *The Annals of Statistics*, p. 246–270, 2009.

- [39] FAN, J.; LI, R. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, v. 96, p. 1348–1360, 2001.
- [40] WANG, H.; LI, G.; TSAI, C.-L. Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 69, p. 63–78, 2007.
- [41] ZOU, H.; HASTIE, T.; TIBSHIRANI, R. et al. On the “degrees of freedom” of the lasso. *The Annals of Statistics*, v. 35, p. 2173–2192, 2007.
- [42] GIACOMINI, R.; WHITE, H. Tests of conditional predictive ability. *Econometrica*, v. 74, p. 1545–1578, 2006.
- [43] FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, v. 33, n. 1, p. 1–22, 2010.
- [44] INOUE, A.; KILIAN, L. How useful is bagging in forecasting economic time series? a case study of us consumer price inflation. *Journal of the American Statistical Association*, v. 103, n. 482, p. 511–522, 2008.
- [45] ELLIOTT, G.; GARGANO, A.; TIMMERMANN, A. Complete subset regressions. *Journal of Econometrics*, v. 177, n. 2, p. 357–373, 2013.
- [46] JURADO, K.; LUDVIGSON, S. C.; NG, S. Measuring uncertainty. Technical report, National Bureau of Economic Research, 2013.
- [47] FISHER, I. *The theory of interest*. Macmillan, New York, 1930.
- [48] IVERSEN, J.; LASÉEN, S.; LUNDVALL, H.; SÖDERSTRÖM, U. Real-time forecasting for monetary policy analysis: The case of Sveriges Riksbank. Working Paper 318, Sveriges Riksbank Working Paper Series, 2016.
- [49] MONTEFORTE, L.; MORETTI, G. Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, v. 32, p. 51–61, 2013.

- [50] CLEMENTS, M.; GALVÃO, A. Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. *Journal of Applied Econometrics*, v. 28, p. 458–477, 2013.
- [51] BELLONI, A.; CHERNOZHUKOV, V. Least squares after model selection in high-dimensional sparse models. *Bernoulli*, v. 19, n. 2, p. 521–547, 2013.
- [52] STOCK, J.; WATSON, M. Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, v. 39, p. 3–33, 2007.
- [53] MARQUES, A. Central Bank of Brazil's market expectations system: a tool for monetary policy. In: *IFC Bulletins*. Bank for International Settlements, 2013. v. 36, p. 304–324.
- [54] SAMUELS, J.; SEKKEL, R. Model confidence sets and forecast combination. *International Journal of Forecasting*, v. 33, p. 48–60, 2017.
- [55] HANSEN, P.; LUNDE, A.; NASON, J. The model confidence set. *Econometrica*, v. 79, p. 453–497, 2011.
- [56] BERNANKE, B.; BOIVIN, J.; ELIASZ, P. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, v. 120, p. 387–422, 2005.
- [57] HAN, X. Tests for overidentifying restrictions in factor-augmented var models. *Journal of Econometrics*, v. 184, p. 394–419, 2015.
- [58] MEDEIROS, M. C.; VASCONCELOS, G. F. Forecasting macroeconomic variables in data-rich environments. *Economics Letters*, v. 138, p. 50–52, 2016.
- [59] EFRON, B.; HASTIE, T.; JOHNSTONE, I.; TIBSHIRANI, R. Least angle regression. *The Annals of Statistics*, v. 32, p. 407–499, 2004.

- [60] AMISANO, G.; GIACOMINI, R. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, v. 25, n. 2, p. 177–190, 2007.
- [61] STOCK, J. H.; WATSON, M. W. Modeling inflation after the crisis. Technical report, National Bureau of Economic Research, 2010.
- [62] FAUST, J.; WRIGHT, J. Forecasting inflation. In: ELLIOTT, G.; TIMMERMANN, A. (Eds.) *Handbook of Economic Forecasting*. Elsevier, 2013. v. 2A.
- [63] VARIAN, H. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, v. 28, p. 3–28, 2014.
- [64] MULLAINATHAN, S.; SPIESS, J. Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, v. 31, p. 87–106, 2017.
- [65] STOCK, J.; WATSON, M. Macroeconomic forecasting with diffusion indexes. *Journal of Business and Economic Statistics*, v. 20, p. 147–162, 2002.
- [66] GU, S.; KELLY, B.; XIU, D. Empirical asset pricing with machine learning. Working paper, University of Chicago, 2018.
- [67] MCCRACKEN, M.; NG, S. FRED-MD: A monthly database for macroeconomic research. *Journal of Business and Economic Statistics*, v. 34, p. 574–589, 2016.
- [68] GIANNONE, D.; LENZA, M.; PRIMICERI, G. Economic predictions with big data: The illusion of sparsity. Working paper, Northwestern University, 2018.
- [69] SCORNET, E.; BIAU, G.; VERT, J.-P. Consistency of random forests. *Annals of Statistics*, v. 43, p. 1716–1741, 2015.
- [70] WAGNER, I.; ATHEY, S. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, forthcoming.

- [71] BLOOM, N. The impact of uncertainty shocks. *Econometrica*, v. 77, n. 3, p. 623–685, 2009.
- [72] KRUGMAN, P. It's baaack: Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity*, , n. 2, p. 137–205, 1998.
- [73] EGGERTSSON, G. B.; WOODFORD, M. Zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, , n. 1, p. 139–233, 2003.
- [74] IACOVIELLO, M. House prices, borrowing constraints, and monetary policy in the business cycle. *American Economic Review*, v. 95, n. 3, p. 739–764, June 2005.
- [75] MIAN, A.; SUFI, A. The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis*. *Quarterly Journal of Economics*, v. 124, n. 4, p. 1449–1496, 2009.
- [76] JUSTINIANO, A.; PRIMICERI, G.; TAMBALOTTI, A. Credit supply and the housing boom. *Journal of Political Economy*, forthcoming.
- [77] GEANAKOPOLOS, J. The leverage cycle. *NBER Macroeconomics Annual*, v. 24, n. 1, p. 1–66, 2010.
- [78] SHILLER, R. J. Speculative asset prices. *American Economic Review*, v. 104, n. 6, p. 1486–1517, June 2014.
- [79] REICHLIN, L.; FORNI, M.; HALLIN, M.; LIPPI, M. The generalised dynamic factor model: Identification and estimation. *Review of Economics and Statistics*, v. 82, p. 540–554, 2000.
- [80] REICHLIN, L.; FORNI, M.; HALLIN, M.; LIPPI, M. The generalised dynamic factor model consistency and rates. *Journal of Econometrics*, v. 119, p. 231–255, 2004.

- [81] TERÄSVIRTA, T.; VAN DIJK, D.; MEDEIROS, M. Linear models, smooth transition autoregressions and neural networks for forecasting macroeconomic time series: A reexamination (with discussion). *International Journal of Forecasting*, v. 21, p. 755–774, 2005.
- [82] NAKAMURA, E. Inflation forecasting using a neural network. *Economics Letters*, v. 86, p. 373–378, 2005.
- [83] HANSEN, P. A test for superior predictive ability. *Journal of Business and Economic Statistics*, v. 23, p. 365–380, 2005.
- [84] QUAEDVLIEG, R. Multi-horizon forecast comparison. Working paper, Erasmus School of Economics, 2017.