

Libonatti, Luis

**Working Paper**

## MIDAS modeling for core inflation forecasting

Economic Research Working Papers, No. 72

**Provided in Cooperation with:**

Economic Research Department (ie), Central Bank of Argentina

*Suggested Citation:* Libonatti, Luis (2017) : MIDAS modeling for core inflation forecasting, Economic Research Working Papers, No. 72, Banco Central de la República Argentina (BCRA), Investigaciones Económicas (ie), Buenos Aires

This Version is available at:

<https://hdl.handle.net/10419/202482>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Economic Research Working Papers | 2017 N° 72



*ie* | BCRA  
INVESTIGACIONES ECONÓMICAS

Economic Research  
Working Papers | 2017  
N° 72

**MIDAS Modeling for Core Inflation  
Forecasting**

**Luis Libonatti**

*Banco Central de la República Argentina*

December 2017



*ie* | BCRA  
INVESTIGACIONES ECONÓMICAS

Working Papers, N° 72

**MIDAS Modeling for Core Inflation  
Forecasting**

**Luis Libonatti**

*Banco Central de la República Argentina*

December 2017

December 2017

ISSN 1850-3977

*Electronic Edition*

Reconquista 266, C1003ABF

Ciudad Autónoma de Buenos Aires, Argentina

Phone | 54 11 4348-3582

Email | [investig@bcra.gob.ar](mailto:investig@bcra.gob.ar)

Website | [www.bcra.gob.ar](http://www.bcra.gob.ar)

The opinions expressed in this paper are the sole responsibility of its authors and do not necessarily represent the position of the Central Bank of Argentina. The **Working Papers** series is comprised of preliminary material intended to stimulate academic debate and receive comments. This paper may not be referenced without authorization from the authors.

# MIDAS Modeling for Core Inflation Forecasting

Luis Libonatti\*

First draft: August 2017

This draft: December 2017

## Abstract

A forecasting exercise is presented to assess the predictive potential of a daily price index based on online prices, compiled by web scrapping by the private company PriceStats in cooperation with a finance research corporation, State Street Global Markets, as a predictor for a measure of the monthly core inflation rate in Argentina, known as “resto IPCBA” and published by the Statistics Office of the Government of the City of Buenos Aires. Mixed frequency regression models offer a convenient arrangement to accommodate variables sampled at different frequencies and hence many specifications are tested. Various classes of MIDAS models are found to produce a slight boost in terms of out-of-sample predictive performance at immediate horizons when compared to benchmark naïve models and estimators. Additionally, an analysis of intraperiod forecasts, reveals a slight trend towards increased forecast accuracy as the daily variable approaches a full month for certain horizons.

**Keywords:** MIDAS, distributed lags, core inflation, forecasting

**JEL Classification:** C22, C53, E37

---

\*Central Bank of Argentina, Economic Research Division. E-mail: [luis.libonatti@bcra.gob.ar](mailto:luis.libonatti@bcra.gob.ar).

# 1 Introduction

Forecasting inflation as frequently as possible has become increasingly important in Argentina as it is essential for economic agents to adjust wages and prices. In regard to this matter, having timely updates about the future trajectory of the inflation rate is essential for conducting monetary policy, specially, since the Central Bank is transitioning towards an inflation targeting regime. Recent developments in the use of “big data” has greatly facilitated tracking macroeconomics variables in real-time. A remarkable example is the construction of online price indexes that are sampled daily, at a much higher frequency when compared to standard price indexes from statistical offices that are often sampled monthly. A natural question arises of whether this information can help anticipate the lower frequency inflation (in terms of which inflation targets are usually defined). [Ghysels et al. \(2004\)](#) introduced a regression framework that allows the exploitation of times series sampled at different frequencies, known in the literature as *Mixed Data Sampling* (MIDAS) regression models. The methodology reduces to fitting a regression model to some low frequency variable using high frequency data as regressors. As it will be shown later, this technique closely resembles distributed lag models. In this paper, I will employ this methodology to assess whether the combination of price series sampled at different frequencies is an effective tool to improve forecast accuracy compared to naïve models, in particular using the online price index constructed by PriceStats in cooperation with State Street Global Markets.

The rest of the paper is organized as follows, in **Section (2)** I present a brief introduction to MIDAS models; in **Section (3)** I go through existing research on applications of MIDAS regressions to forecasting inflation as well as the derivation of some useful models and results; in **Section (4)** I describe the forecasting exercise and comment on the results and the accuracy of the models in terms of root mean square error (RMSFE) and in **Section (5)** I present a short conclusion; and motivate further research in this area.

## 2 MIDAS Regression Models

MIDAS regression models propose a data driven method to aggregate high frequency variables into lower frequency predictors and provides an alternative to the well known “bridge” approach ([Schumacher, 2016](#)) where the high frequency variables are aggregated with equal weights (flat aggregation).<sup>1</sup> [Ghysels et al. \(2004\)](#) suggest to combine  $y_t$ , a low frequency process, and  $x_\tau$  a high frequency process that is observed a discrete and fixed number of times  $m$  each time a new value of  $y_t$  is observed, in a plain regression equation,

$$y_t = \sum_{j=0}^{m-1} \theta_j x_{t-j/m} + u_t, \quad (1)$$

or more compactly,

$$y_t = \mathbf{x}_t \boldsymbol{\theta} + u_t, \quad (2)$$

where  $\mathbf{x}_t \equiv [x_t \dots x_{t-(m-1)/m}]$  is a  $1 \times m$  row vector that collects all the  $x_\tau$  corresponding to period  $t$  and  $\boldsymbol{\theta} \equiv [\theta_0 \dots \theta_{m-1}]'$  is the  $m \times 1$  vector of weight coefficients.<sup>2</sup> Each  $j$  high frequency observation  $x_{t-j/m}$  within the low frequency period  $t$  enters the model linearly as a variable accompanied by its specific weight,  $\theta_j$ , totaling  $m$  explanatory variables and  $m$  weights, plus an error term. The high frequency subindex  $\tau$  needs to

---

<sup>1</sup>In fact, this can be considered as a special case of a MIDAS regression.

<sup>2</sup>This equation may also include constants, trends, seasonal terms or other low frequency explanatory variables.

be represented in terms of the low frequency index  $t$  by noting that  $\tau = t - 1 + j/m$  for  $j = 1, \dots, m$  since  $m$  is fixed, where  $x_{t-0/m}$  would be the most recent observation. This structure, in fact, conceals a high frequency lag polynomial  $\theta(L^{1/m}) \equiv \sum_{j=0}^{m-1} \theta_j L^{j/m} x_t$  so that  $L^{j/m} x_t = x_{t-j/m}$  in a similar fashion to a distributed lags model.

In order to provide a clearer perspective, it is perhaps easier to introduce matrix notation. Defining  $\mathbf{X} \equiv [\mathbf{x}'_1 \dots \mathbf{x}'_T]'$  as the  $T \times m$  matrix that groups all the  $\mathbf{x}_t$  vectors together;  $\mathbf{y} \equiv [y_1 \dots y_T]'$ , the collection of the low frequency observations of size  $T \times 1$ ; and  $\mathbf{u} \equiv [u_0 \dots u_T]'$ , the residuals of the same length as  $\mathbf{y}$ , it is possible to unveil a simple multiple regression equation,

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{T-1} \\ y_T \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} x_1 & \dots & x_{1-(m-1)/m} \\ x_2 & \dots & x_{2-(m-1)/m} \\ \vdots & \ddots & \vdots \\ x_{T-1} & \dots & x_{T-1-(m-1)/m} \\ x_T & \dots & x_{T-(m-1)/m} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \theta_0 \\ \vdots \\ \theta_{m-1} \end{bmatrix}}_{\boldsymbol{\theta}} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{T-1} \\ u_T \end{bmatrix}}_{\mathbf{u}}. \quad (3)$$

Indeed this problem can be solved by ordinary least squares (OLS) and this method will produce consistent coefficient estimates. **Equation (1)** is usually referred to as the unrestricted MIDAS regression model (U-MIDAS).<sup>3</sup> However, an inconvenience arises when  $m$ , the length of the vector  $\boldsymbol{\theta}$ , is large relative to the sample size  $T$ , which is often the case with MIDAS regressions. When this occurs, the models suffers from parameter proliferation and OLS estimation produces poor estimates and consequently, poor forecasts. A simple way to overcome this deficiency is to impose restrictions on the coefficients of the high frequency lag polynomial and restate each  $\theta_j$  as function of some  $q$  hyperparameters and its subindex  $j$  (its position within the low frequency lag polynomial) in a such a way that  $q \gg m$ . Each  $\theta_j$  is redefined as  $\theta_j \equiv w_j(\boldsymbol{\gamma}; j)$  where  $\boldsymbol{\gamma}$  is the collection of  $q$  hyperparameters that characterize the weight function  $w_j(\cdot)$ . **Equation (1)** is then transformed to,

$$y_t = \lambda \sum_{j=0}^{m-1} \left( \frac{w_j(\boldsymbol{\gamma}; j)}{\sum_{j=0}^{m-1} w_j(\boldsymbol{\gamma}; j)} \right) x_{t-j/m} + u_t. \quad (4)$$

where  $\lambda$  is an impact parameter and the weights are normalized so that they sum up to unity. Ghysels et al. (2004) initially recommended what is known as the exponential Almon polynomial as a candidate for weight function as it allows for many different shapes and depends only on very few parameters. This is an exponentiated version of an Almon lag polynomial well known in the distributed lags literature,<sup>4</sup>

$$\theta_j(\gamma_1, \dots, \gamma_q; j) = e^{\sum_{s=1}^q \gamma_s j^s}. \quad (a)$$

Another conventional candidate is the beta probability density (also normalized),

$$\theta_j(\gamma_1, \gamma_2; j) = z_j^{\gamma_1 - 1} (1 - z_j)^{\gamma_2 - 1}, \quad (b)$$

with  $z_j \equiv j/(m-1)$ .

Parameterization as in **Equation (a)** has proved to be quite popular and has become the standard among researchers, particularly when  $q = 2$ .

The introduction of constrained coefficients like the parameterizations above has many far-reaching implications. The model turns nonlinear and lacks a closed form solution. It is necessary to resort to nonlinear least

<sup>3</sup>Froni et al. (2015) present a detailed assessment of this model.

<sup>4</sup>See for example the book by Judge et al. (1985).

squares and approximate the solution by numerical optimization routines. Additionally, the constraints are highly likely to introduce a bias in each  $\theta_j$ . However, based on Monte Carlo simulations, when the sample size is small relative to the number of parameters, [Ghysels et al. \(2016\)](#) argue that both parameter estimation precision and out-of-sample forecast accuracy gained by the increase in degrees of freedom, far offset the effects of the bias generated by misspecified constraints.

MIDAS models are intended as a direct forecasting tool as this can be more robust against misspecification ([Marcellino et al., 2006](#)) and are horizon specific in opposition to iterative forecasting in bridge equations; in other words, parameter estimates will differ based on the desired forecast horizon  $h$ , since the projection will be done on a different information set.<sup>5</sup> Setting  $W(L^{1/m}; \gamma) \equiv \sum_{j=0}^{m-1} w_j(\gamma; j)L^{1/m}$ , the  $h$  period ahead forecast will be given by

$$\hat{y}_{T+h|T} = \hat{\lambda}_h W(L^{1/m}; \hat{\gamma}_h) x_T. \quad (5)$$

Note that MIDAS regressions allow for noninteger forecast horizons, i.e. intraperiod forecasts. One simply needs to project on  $x_{t+j/m}$  for desired  $t$  and  $j$ .

To get to **Equation (5)**, it is first necessary to estimate,

$$y_t = \lambda W(L^{1/m}; \gamma) x_{t-h} + u_t, \quad (6)$$

and then compute  $\hat{y}_{T+h|T}$  with the estimated parameters,  $\hat{\lambda}_h$  and  $\hat{\gamma}_h$ , and  $\{x_t, \dots, x_{t-(m-1)/m}\}$ .

An immediate extension of the MIDAS regression model that is going to be employed throughout this paper is to allow for more than  $m$  high frequency regressors, for example by including  $p_x$  lags of the vector  $x_t$ , totaling  $m \times L_x$  high frequency variates where  $L_x = p_x + 1$ , as sort of a distributed lags model that is known as MIDAS-DL,

$$y_t = \sum_{r=0}^{p_x} (\theta_r x_{t-r})' + u_t, \quad (7)$$

or in matrix form,

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{T-1} \\ y_T \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} x_1 & \dots & x_{1-(m-1)/m} & \dots & x_{1-p_x} & \dots & x_{1-p_x-(m-1)/m} \\ x_2 & \dots & x_{2-(m-1)/m} & \dots & x_{2-p_x} & \dots & x_{2-p_x-(m-1)/m} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ x_{T-1} & \dots & x_{T-1-(m-1)/m} & \dots & x_{T-1-p_x} & \dots & x_{T-1-p_x-(m-1)/m} \\ x_T & \dots & x_{T-(m-1)/m} & \dots & x_{T-p_x} & \dots & x_{T-p_x-(m-1)/m} \end{bmatrix}}_{\mathbf{X}_{DL}} \underbrace{\begin{bmatrix} \theta_{0,0} \\ \vdots \\ \theta_{0,m-1} \\ \vdots \\ \theta_{p_x,0} \\ \vdots \\ \theta_{p_x,m-1} \end{bmatrix}}_{\boldsymbol{\theta}_{DL}} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{T-1} \\ u_T \end{bmatrix}}_{\mathbf{u}}. \quad (8)$$

Estimation of different weight functions for each  $\theta$  in **Equation (7)** constitutes the multiplicative or aggregates based MIDAS model. On the contrary, employing a single weight function for all  $m \times L_x$  coefficients  $\theta_r$  is also possible, and this last version will be employed in the paper.

Other possible extensions include adding high frequency leads ([Andreou et al., 2013](#)), building high frequency factors ([Marcellino and Schumacher, 2010](#)), incorporating cointegration relations ([Miller, 2013](#)), developing a bayesian framework ([Rodriguez and Puggioni, 2010](#)), integrating markov switching ([Guérin and Marcellino, 2013](#)), estimating multivariate models ([Ghysels et al., 2007](#)), using infinite polynomials ([Ghysels et al., 2007](#)) or adding low frequency autoregressive augmentations ([Ghysels et al., 2007](#); [Clements and Galvão, 2008](#); [Duarte, 2014](#)), for example. [Forni and Marcellino \(2013\)](#) provide a comprehensive survey of possible extensions in a recent survey about mixed frequency models.

<sup>5</sup>[Marcellino et al. \(2006\)](#) go through an extensive comparison of direct and iterated approaches to forecasting.



### 3 Literature Review

[Clements and Galvão \(2008\)](#) were among the first to study applications of MIDAS regressions to macroeconomic variables. In their paper, they forecast U.S. real quarterly output growth in combination with three different monthly variables: industrial production, employment growth and capacity utilization. They find a slight increase in out-of-sample forecast accuracy in terms of RMSFE both with vintage and revised data compared to two benchmarks (an autoregression and an ADL model), particularly, for short term horizons. They also derive and assess a model with autoregressive dynamics introduced as a common factor shared by the low and the high frequency lag polynomials. Based on [Ghysels et al. \(2007\)](#) comments, they argue that including an autoregressive term in a standard MIDAS model, as in the next equation,

$$y_t = \phi y_{t-1} + \lambda W(L^{1/m})x_t + u_t, \quad (9)$$

induces a seasonal response from  $y_t$  to  $x_t$  irrespective of whether  $x_t$  exhibits a seasonal pattern. They suggest to further restrict the model, by adding a common lag polynomial shared between  $y_t$  and  $x_t$ ,

$$(1 - \phi L)y_t = \lambda(1 - \phi L)W(L^{1/m})x_t + u_t, \quad (10)$$

so that when writing the model in distributed lag representation, the polynomial in  $L$  cancels out, eliminating the spurious seasonal response. A multi-step generalization of [Equation \(10\)](#) for  $h$  step ahead forecasts would be,

$$(1 - \phi L^{f+1})y_t = \lambda(1 - \phi L^{f+1})W(L^{1/m})x_t + u_t, \quad (11)$$

where  $f$  is just the integer part of the forecast horizon  $h$ .

[Armesto et al. \(2010\)](#) analyze the performance of MIDAS models for the U.S. economy regarding four different variable combinations: i) quarterly GDP growth and monthly employment growth; ii) monthly CPI inflation and daily fed funds rate; iii) monthly industrial production growth and a measure of term spread; and iv) employment growth and again a measure of term spread. They contrast the results of flat aggregation, the exponential Almon polynomial and a step weight function, but are unable to find a dominant model specification. They test configurations which include both leads and lags and provide detailed results for intraperiod forecasting performance of the models, computed by accumulating lead variables as the leads approach a full low frequency period. They find an erratic pattern for the RMSFE of the models as a function of the leads included in them. This implies that, in a real time setting, it is not trivial which intraperiod forecasts are the most accurate.

[Monteforte and Moretti \(2013\)](#) develop MIDAS models to forecast the euro area harmonized price index inflation. They put forward a two step approach involving low and high frequency variables. In the first place, they estimate a generalized dynamic factor model ([Forni et al., 2000](#)) for the inflation rate based on a set of variables, then extract a common component and separate that into a long-run and a cyclical or short run component. The second step consists in fitting the model of [Clements and Galvão \(2008\)](#) to capture short term dynamics and use financial time series as high frequency regressors, in addition to the long-run component previously estimated as well as other low frequency variables. They design three MIDAS models, M1, M2 and M3, each with different high frequency regressors: i) M1 includes the short term interest rate and changes in interest rate spread and oil future prices; ii) M2 uses changes in the wheat price, oil future quotes and the exchange rate; and finally, iii) M3 consists of long-term rates and changes in the interest rates spreads and in the short-term rate. They contrast the out-of-sample performance in terms of RSMFE of these models against the equation for the inflation rate of two different low frequency VARs, a random walk, an autoregressive moving average model and an autoregression. They compute all the intraperiod forecasts for the MIDAS models, the monthly average of these daily forecasts and compare this mean to all the low frequency models. All the analysis is conducted for one-month-ahead and two-month-ahead forecasts. They find on average a 20% reduction in forecast error dispersion. The authors also provide a last empirical exercise, by using forecast combinations with the MIDAS models and the inflation rate implied by financial derivatives, but are not able to find any significant gains this way.

Duarte (2014) discusses in detail the implications of autoregressive augmentations in MIDAS regression models and different ways to incorporate them. The author explores the out-of-sample performance of MIDAS models with autoregressive augmentations with no restrictions, models with an autoregressive augmentation with a common factor restriction and models with autoregressive augmentations with no restrictions and a multiplicative scheme to aggregation, and compares those models to the same versions but excluding the autoregressive component and to two benchmarks, a low frequency autoregression and a low frequency multiple regression model. She computes forecasts for quarterly euro area GDP growth based on three different series: i) industrial production, ii) an economic sentiment indicator and iii) the Dow Jones Euro Stoxx index. She disregards the seasonal spikes impulse responses as the relevant impulse responses, as she argues that it is not possible to single out a particularly relevant impulse response for a mixed frequency process since they vary depending on when the shocks occur within the low frequency process. Although there is no superior model among all of those tested, the author finds sizable gains compared to the naïve benchmarks at all horizons.

Breitung and Roling (2015) propose a “nonparametric” MIDAS model to forecast monthly inflation rates using a daily predictor. Instead of imposing any particular polynomial parametrization, the nonparametric approach consists on enforcing some degree of smoothness to the lag distribution by minimizing a penalized least squares cost function,

$$S(\theta) = (\mathbf{y} - \mathbf{X}\theta)'(\mathbf{y} - \mathbf{X}\theta) + \eta\theta' \mathbf{D}' \mathbf{D}\theta, \quad (12)$$

where  $D$  is a  $(m - 1) \times (m + 1)$  matrix such that,

$$D = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & -2 & 1 \end{bmatrix}, \quad (13)$$

and  $\eta$  is a pre-specified smoothing parameter. They call this estimator the Smoothed Least Squares (SLS) estimator and its structure closely resembles the well known Hodrick-Prescott filter. If  $\eta$  is not known, they suggest to solve for the  $\eta$  that minimizes the Akaike Information Criterion. Their target variable is the harmonized index of consumer prices for the euro area and they use a commodity price index as a high frequency regressor. They compare their model against the recursive unconditional mean and the parametric MIDAS model (exponential Almon weights) for two different forecast horizons. They conclude that the commodity index paired with the nonparametric MIDAS results in a reasonably good one-month-ahead forecast. Additionally, the authors conduct a Monte Carlo experiment and compare their model to four parametric MIDAS alternatives: the exponential Almon polynomial, a hump shaped function, a declining linear function and sinusoidal type. They find that the nonparametric method attains similar performance compared to the parametric competitors.

## 4 Data, Exercise and Results

In this paper, the out-of-sample predictive performance of a daily inflation rate series computed from a daily price index based on online prices is analyzed in order to forecast a monthly price index, employing many different MIDAS regression model specifications mentioned in the previous two sections.

### 4.1 Data

The online index is compiled by the company PriceStats in cooperation with State Street Global Markets, a leading financial research corporation. PriceStates is a spin-off company that emerged from the Billions Prices Project at MIT, founded by professors Alberto Cavallo and Roberto Rigobon. It is the first company, institution or organization to apply a big data

approach to produce real-time (daily) price indexes to track general price inflation and other related metrics. Essentially, they collect daily data of prices from a multiplicity of online retailers by “web scraping” (i.e. recording price information contained inside specific HyperText Markup Language (HTML) tags in the retailers’ websites) and aggregate the data by replicating the methodology of standard consumer price indexes, as is done by National Statistics Offices with offline prices. Cavallo (2013) analyses in detail the methodology and provides comparisons between online and offline price indexes for Argentina, Brazil, Chile, Colombia and Venezuela. He concludes that online prices are able to track well the dynamic behavior of inflation rates over time with the exception of Argentina. In fact, the construction of online price indexes was initially motivated in order to provide the public with an alternate measure of the inflation rate in Argentina, since that from the years 2007 to 2015, there were large discrepancies between the official price indexes compiled by the National Institute of Statistics and Census (INDEC) and price indexes compiled by provincial statistics offices or those compiled by private consultants. Throughout the rest of the paper, this price index will be referred to as the State Street PriceStats Index (SSPS). Data for Argentina is available since November 1, 2007 with a 3 day publication lag.

A provincial price index that raised itself to prominence in recent years is the consumer price index compiled by the General Department of Statistics and Censuses of the Government of the Autonomous City of Buenos Aires, known as IPCBA. Although this index only contemplates the territory of the City of Buenos Aires (with a population close to 3 million), it should be reasonable to expect that price dynamics in the Buenos Aires Metropolitan Area (which encompasses a much larger population, close to 14 million or 1/3 of the total population of Argentina) share most of its features with the pricing structure of the City Buenos Aires, due to arbitrage by reason of geographical proximity, would prevent large distortions, at least in nonregulated markets. A more restricted version of the index is also published, called “resto IPCBA” (rIPCBA) and it serves as a measure of core inflation. Compared to the headline version, it excludes products with strong seasonal patterns and regulated prices (public utility services), and represents 78.15% of the headline index. rIPCBA is available since July 2012 and is released monthly, with approximately a two week publication lag.

The above indexes and as well as other provincial private and public price indexes are closely monitored by the monetary authorities, as well as the general public, particularly the recently released National Price Consumer Index by INDEC, that as the name implies is the only index with full national coverage. However, this index still consists of less than a year of data points and this severely limits the possibilities of drawing any relevant inference as of today.

## 4.2 Forecasting Exercise

By combining the high frequency daily inflation rate of SSPS,  $\pi_{\tau}^{SSPS} \equiv 100\Delta \ln P_{\tau}^{SSPS}$ , and the low frequency monthly core inflation rate of rIPCBA, the target variable,  $\pi_t^{rIPCBA} \equiv 100\Delta \ln P_t^{rIPCBA}$ , I will estimate different MIDAS models and examine if these specifications can outperform single frequency naïve models. Work done by Atkeson and Ohanian (2001), Stock and Watson (2007) and Faust and Wright (2009) found evidence for the US that naïve models, are sometimes not easily beaten by more sophisticated models.

Inflation in Argentina in recent years has been high, unstable and volatile, particularly from 2012 to most of 2016 when Argentina experienced high monetization of fiscal deficits, strict capital controls and two major devaluations of the currency.<sup>6</sup> The average monthly core inflation rate according to rIPCBA has been fluctuating around 2.2% and 2.1% for the monthly aggregated SSPS series, with coefficients of variation at 34% and 55%, respectively. This should pose a significant challenge to formulate accurate forecasts. **Figure 1** illustrates the comparison between these two indexes and provides a quick glimpse at the potential predictive power of the high frequency index. Overall and for the scope of this work, rIPCBA is available from July 2012 to June 2017 (60 data points) while SSPS ranges from November 1, 2007 to June 30, 2017 (3530 data points).

---

<sup>6</sup>The last one coinciding with the lifting of the majority of the capital controls in December 2015 and a subsequent transition to a flexible exchange rate regime and inflation targeting.

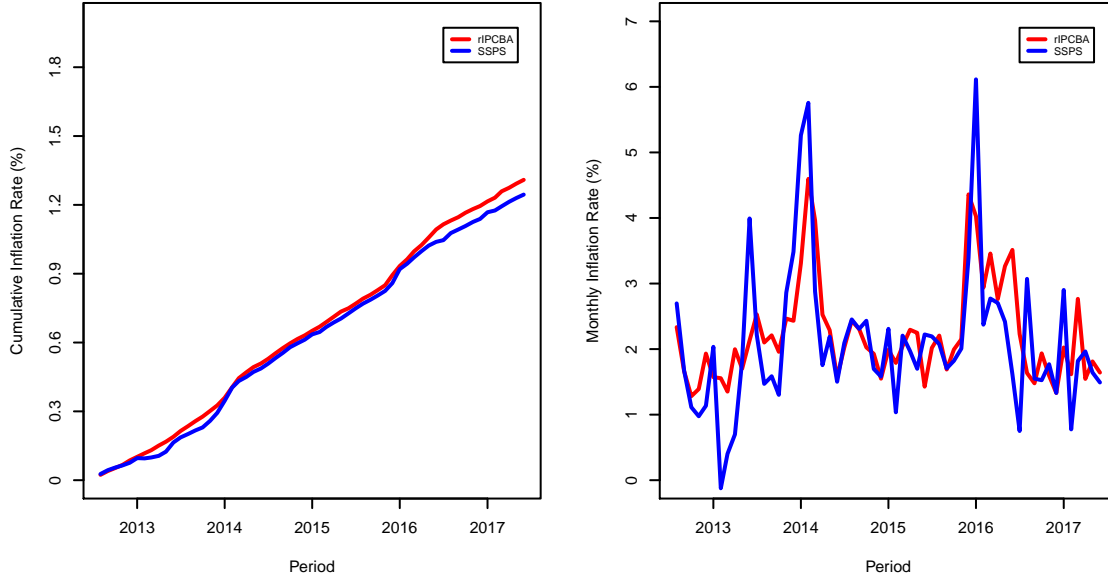


Figure 1: Comparison between rIPCBA inflation and SPSS inflation aggregated to monthly frequency

The MIDAS specifications to be estimated are the MIDAS-DL, the unrestricted autoregressive MIDAS-DL (MIDAS-ADL) and the autoregressive MIDAS-DL with the common factor restriction (MIDAS-ADL-CF). All MIDAS specifications are tested with a number of high frequency variables equal to  $m \times L_x$ ,<sup>7</sup> with  $L_x \in \{1, 2, 3\}$ , and forecasts are computed for periods  $T + 1$  ( $h = 0$ ),  $T + 2$  ( $h = 1$ ) and  $T + 3$  ( $h = 2$ ) over a 24 observation evaluation sample, spanning from 2015.07 to 2017.06, and a 12 observation subsample from 2016.07 to 2017.06 (a period with a more stable inflation rate), using recursive (expanding) windows. MIDAS-ADL-CF models include quadratic and cubic variations of the standard Almon polynomial and the exponential version, as well as the Beta probability density function. MIDAS-ADL models further add flat aggregation (equal weights); and, moreover, MIDAS-DL add the nonparametric (NP) model described in **Section 3**. In addition, all these models are compared to two benchmarks: i) the low frequency recursive unconditional mean and ii) a low frequency first order autoregression.

In a first stage, the models are estimated with a balanced dataset. There is exact frequency matching;  $m$  daily observations from the same month or  $L_x$  groups of  $m$  daily observations from the same months correspond to a specific low frequency monthly observation of the dependent variable. In total, two sets of 168 RMSFE are computed, one corresponding to the large sample and the other to a reduced subsample.

A second stage involves estimating intraperiod forecasts for the best selected  $L_x$  for each period based on the results from the large sample of the first stage and briefly analyzing the stability of the forecasts as more recent information is incorporated in the models. Forecasts from the autoregression and the unconditional mean remain the same throughout the month.

To account for the fact that SPSS is an irregularly spaced series, the frequency is assumed fixed at  $m = 28$ , and so days 29, 30 and 31 of each month are discarded.<sup>8</sup>

Estimation is conducted in **R** with the **midasr** package developed by [Ghysels et al. \(2016\)](#) and optimization is performed with three routines included in **optimx**<sup>9</sup> for models without a closed form solution or with the **lm** function from the **stats** package for linear models. Models that require **optimx** are solved sequentially, they are first solved with **ucminf**, then

<sup>7</sup>MIDAS-ADL-CF includes at least  $m \times L_x$ , high frequency regressors or more since the common factor restriction may increase the number of variables depending on the forecast horizon and the number of high frequency lags.

<sup>8</sup>This is not contemplated in **Figure 1**. Daily observations are aggregated considering full data.

<sup>9</sup>Details about this package can be found in [Nash and Varadhan \(2011\)](#).

with **nlsminb** and finally with **Nelder-Mead**; inputting as initial conditions the solution from the previous algorithm. The first algorithm is initialized taking as starting conditions the hypothesis of equal weights and a null impact parameter. This strategy delivered reasonable results empirically and serves as a rough check on whether the high frequency regressors are actually relevant.

### 4.3 Empirical Results

**Tables (1) and (2)** summarize the main results of the first stage. In general, for  $h = 0$  (nowcasts), larger values of  $L_x$  produce better results while this tends to reverse when forecasting further into the future, i.e.  $h = 1$  or  $h = 2$ . For  $h = 1$ , the result is ambiguous for the larger sample but the smaller sample indicates that a result in the middle like  $L_x = 2$  performs best. MIDAS-ADL specifications exhibit slightly worse predictive accuracy compared to MIDAS-DL and MIDAS-ADL-CF for  $h = 0$ , the latter two also being able to produce a small gain at around 10% when compared to the autoregression and a larger 25% against the recursive unconditional mean.<sup>10</sup> The smaller sample greatly amplifies these results. For  $h = 1$ , the results are again ambiguous but for  $h = 2$ , there seems to be an indication that MIDAS-DL models outperform the rest. Note that for each  $h$ , there is a flat aggregation model that preformed very well, even better at times than standard MIDAS models but overall, there is not a single MIDAS model that systematically outperforms the rest.

**Figures (2) to (4)** condense the main findings of the second stage. Forecasts for period  $T + 1$  display a clear trend towards better accuracy as the high frequency variable reaches a full low frequency period. In day 1 to day 28 point to point comparison, the RMSFE is reduced by approximately 20% and particularly, in the second half of the month, the models start to surpass the accuracy of the autoregression by a 15% at most for some days. This trend is less evident for the MIDAS-ADL models, except in the case of flat aggregation. The improved performance when evaluated in the subsample suggests that it is even possible to obtain better results as the inflation rate stabilizes. Similar behavior, although less evident, is observed for forecasts for period  $T + 3$  in the case of MIDAS-DL models. Forecasts for period  $T + 2$  display a rather erratic pattern with the exception of the flat aggregation MIDAS-DL and MIDAS-ADL models.

**Figures (5)** zooms in on the evolution of all intraperiod forecasts for models that performed best for each specific period, either  $T + 1$ ,  $T + 2$  or  $T + 3$ . In spite of the intraperiod forecasts evidencing some volatility within the month, this does not seem to be a major concern as inflation stabilizes at the end of the sample. Additionally, note that forecasting further into the future yields a dynamic closer to the unconditional mean of the whole process.

Even though the results look promising, they should be interpreted with caution since that even the “large” out-of-sample evaluation set, actually, constitutes a very small sample by literature standards and this prevents any meaningful inference from conventional predictive accuracy tests such as those proposed by [Giacomini and White \(2006\)](#), for example. In the near future, as more observations accumulate, these results could be used as a training sample from which to compute inverse mean square error weights and perform forecast combinations, which could prove to be effective in mitigating intraperiod forecast volatility.

---

<sup>10</sup>Tables with RMSE ratios are presented in the **Appendix**.

Table 1: Out-of-sample predictive performance, RMSFE

	$h = 0$			$h = 1$			$h = 2$		
	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$
<i>Almon (d = 2)</i>									
MIDAS-DL	0.701	0.703	<b>0.633</b>	0.889	0.804	<b>0.803</b>	<b>0.794</b>	<i>0.804</i>	<i>0.824</i>
MIDAS-ADL	<i>0.660</i>	<i>0.668</i>	<b>0.636</b>	<i>0.882</i>	<i>0.817</i>	<b>0.811</b>	<i>0.825</i>	<b>0.817</b>	<i>0.840</i>
MIDAS-ADL-CF	0.678	0.701	<b>0.626</b>	0.887	<i>0.802</i>	<b>0.797</b>	0.811	<b>0.807</b>	0.853
<i>Almon (d = 3)</i>									
MIDAS-DL	0.730	0.702	<b>0.641</b>	0.919	0.833	<b>0.813</b>	<i>0.807</i>	<b>0.796</b>	<i>0.822</i>
MIDAS-ADL	<i>0.687</i>	<i>0.659</i>	<b>0.646</b>	<i>0.914</i>	<i>0.853</i>	<b>0.817</b>	<i>0.835</i>	<b>0.804</b>	<i>0.838</i>
MIDAS-ADL-CF	0.710	0.721	<b>0.635</b>	0.936	<i>0.823</i>	<b>0.805</b>	0.816	<b>0.805</b>	0.861
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	0.749	0.753	<b>0.635</b>	0.974	0.873	<b>0.833</b>	0.818	0.821	<b>0.816</b>
MIDAS-ADL	<i>0.673</i>	<b>0.656</b>	0.673	<i>0.901</i>	<b>0.839</b>	0.843	0.840	<b>0.805</b>	<b>0.805</b>
MIDAS-ADL-CF	0.699	0.707	<b>0.623</b>	0.909	<i>0.832</i>	<b>0.830</b>	<i>0.802</i>	<b>0.787</b>	0.826
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	0.803	0.727	<b>0.636</b>	<i>0.856</i>	0.859	<b>0.834</b>	<b>0.819</b>	<i>0.821</i>	<i>0.822</i>
MIDAS-ADL	0.689	<b>0.649</b>	0.670	0.902	<b>0.844</b>	0.850	0.835	0.835	<b>0.834</b>
MIDAS-ADL-CF	<i>0.670</i>	0.724	<b>0.629</b>	1.020	<i>0.844</i>	<b>0.835</b>	<b>0.828</b>	0.833	0.834
<i>Beta</i>									
MIDAS-DL	0.712	0.710	<b>0.646</b>	0.833	<b>0.773</b>	<i>0.785</i>	<b>0.765</b>	<i>0.783</i>	<i>0.781</i>
MIDAS-ADL	<b>0.635</b>	0.711	0.690	<i>0.821</i>	<b>0.791</b>	0.797	<b>0.807</b>	0.817	0.815
MIDAS-ADL-CF	0.654	<i>0.705</i>	<b>0.629</b>	0.827	<b>0.768</b>	0.778	0.810	<b>0.807</b>	0.811
<i>Flat</i>									
MIDAS-DL	1.347	0.694	<b>0.637</b>	0.830	<b>0.749</b>	<i>0.769</i>	<b>0.784</b>	<i>0.815</i>	<i>0.827</i>
MIDAS-ADL	<i>1.052</i>	<i>0.671</i>	<b>0.636</b>	<i>0.821</i>	<b>0.768</b>	0.814	<b>0.798</b>	0.839	0.853
<i>Nonparametric</i>									
MIDAS-DL	0.691	0.704	<b>0.637</b>	0.869	<b>0.804</b>	0.806	<b>0.787</b>	0.799	0.825
<i>Autoregressive</i>									
$p = 1$	0.695	0.695	0.695	0.794	0.794	0.794	0.848	0.848	0.848
<i>Unconditional Mean</i>									
$\bar{y}$	0.860	0.860	0.860	0.872	0.872	0.872	0.880	0.880	0.880

Notes: The evaluation sample comprises 24 data points, from 2015.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .

Table 2: Out-of-sample predictive performance, RMSE

	<i>h</i> = 0			<i>h</i> = 1			<i>h</i> = 2		
	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3
<i>Almon (d = 2)</i>									
MIDAS-DL	0.573	0.556	<b>0.459</b>	0.668	<b>0.525</b>	0.552	<b>0.502</b>	0.512	0.509
MIDAS-ADL	0.517	0.535	<b>0.478</b>	0.644	<b>0.543</b>	0.560	0.560	<b>0.550</b>	0.560
MIDAS-ADL-CF	0.562	0.586	<b>0.471</b>	0.657	<b>0.560</b>	0.572	<b>0.572</b>	0.583	0.585
<i>Almon (d = 3)</i>									
MIDAS-DL	0.607	0.591	<b>0.464</b>	0.705	<b>0.548</b>	0.565	0.513	<b>0.494</b>	0.500
MIDAS-ADL	0.549	0.556	<b>0.490</b>	0.683	0.592	<b>0.574</b>	0.579	<b>0.544</b>	0.564
MIDAS-ADL-CF	0.628	0.629	<b>0.484</b>	0.703	0.613	<b>0.583</b>	0.581	<b>0.576</b>	0.617
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	0.623	0.536	<b>0.466</b>	0.807	0.586	<b>0.549</b>	<b>0.519</b>	<b>0.519</b>	0.533
MIDAS-ADL	0.552	<b>0.509</b>	<b>0.509</b>	0.652	<b>0.535</b>	0.555	0.560	<b>0.554</b>	<b>0.554</b>
MIDAS-ADL-CF	0.653	0.590	<b>0.474</b>	0.654	<b>0.551</b>	0.574	<b>0.564</b>	0.570	0.570
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	0.635	0.594	<b>0.462</b>	0.546	<b>0.538</b>	0.543	<b>0.519</b>	<b>0.519</b>	0.524
MIDAS-ADL	0.538	<b>0.503</b>	0.514	0.717	<b>0.536</b>	0.547	<b>0.552</b>	<b>0.552</b>	0.546
MIDAS-ADL-CF	0.617	0.612	<b>0.473</b>	0.997	0.551	<b>0.564</b>	<b>0.562</b>	0.588	0.587
<i>Beta</i>									
MIDAS-DL	0.631	0.526	<b>0.482</b>	0.594	<b>0.473</b>	0.521	<b>0.480</b>	0.508	0.522
MIDAS-ADL	0.563	0.526	<b>0.502</b>	0.567	<b>0.490</b>	0.527	<b>0.540</b>	0.573	0.564
MIDAS-ADL-CF	0.612	0.543	<b>0.485</b>	0.580	<b>0.505</b>	0.541	<b>0.546</b>	0.569	0.576
<i>Flat</i>									
MIDAS-DL	0.615	0.544	<b>0.428</b>	0.614	<b>0.476</b>	0.505	<b>0.479</b>	0.532	0.548
MIDAS-ADL	0.574	0.521	<b>0.426</b>	0.598	<b>0.499</b>	0.530	<b>0.512</b>	0.547	0.580
<i>Nonparametric</i>									
MIDAS-DL	0.569	0.569	<b>0.465</b>	0.651	<b>0.526</b>	0.550	<b>0.488</b>	0.504	0.508
<i>Autoregressive</i>									
<i>p</i> = 1	0.604	0.604	0.604	0.609	0.609	0.609	0.674	0.674	0.674
<i>Unconditional Mean</i>									
$\bar{y}$	0.607	0.607	0.607	0.612	0.612	0.612	0.616	0.616	0.616

Notes: The evaluation sample comprises 12 data points, from 2016.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .

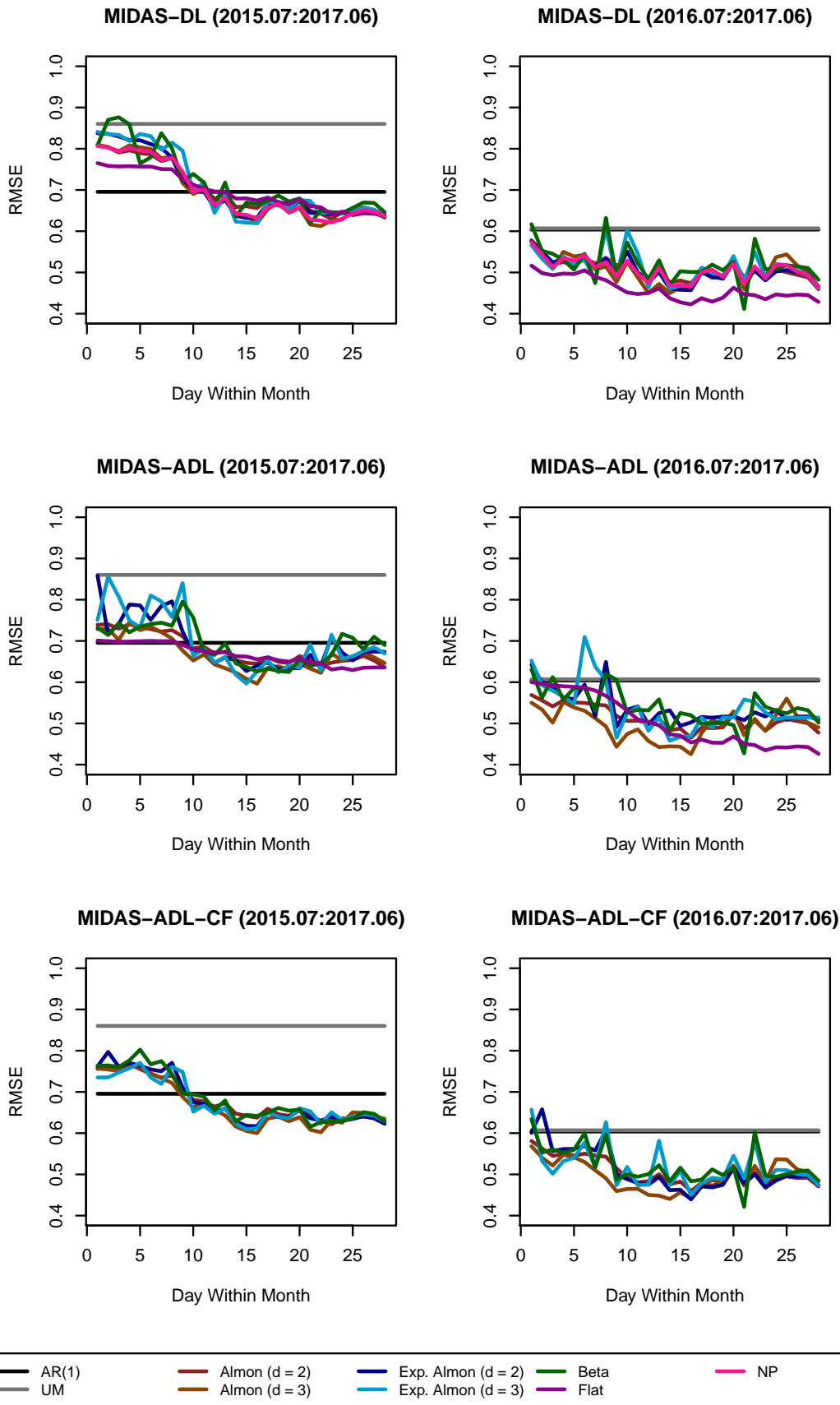


Figure 2: Evolution of RMSFE of period  $T + 1$  forecasts within a month for selected models with  $L_x = 3$



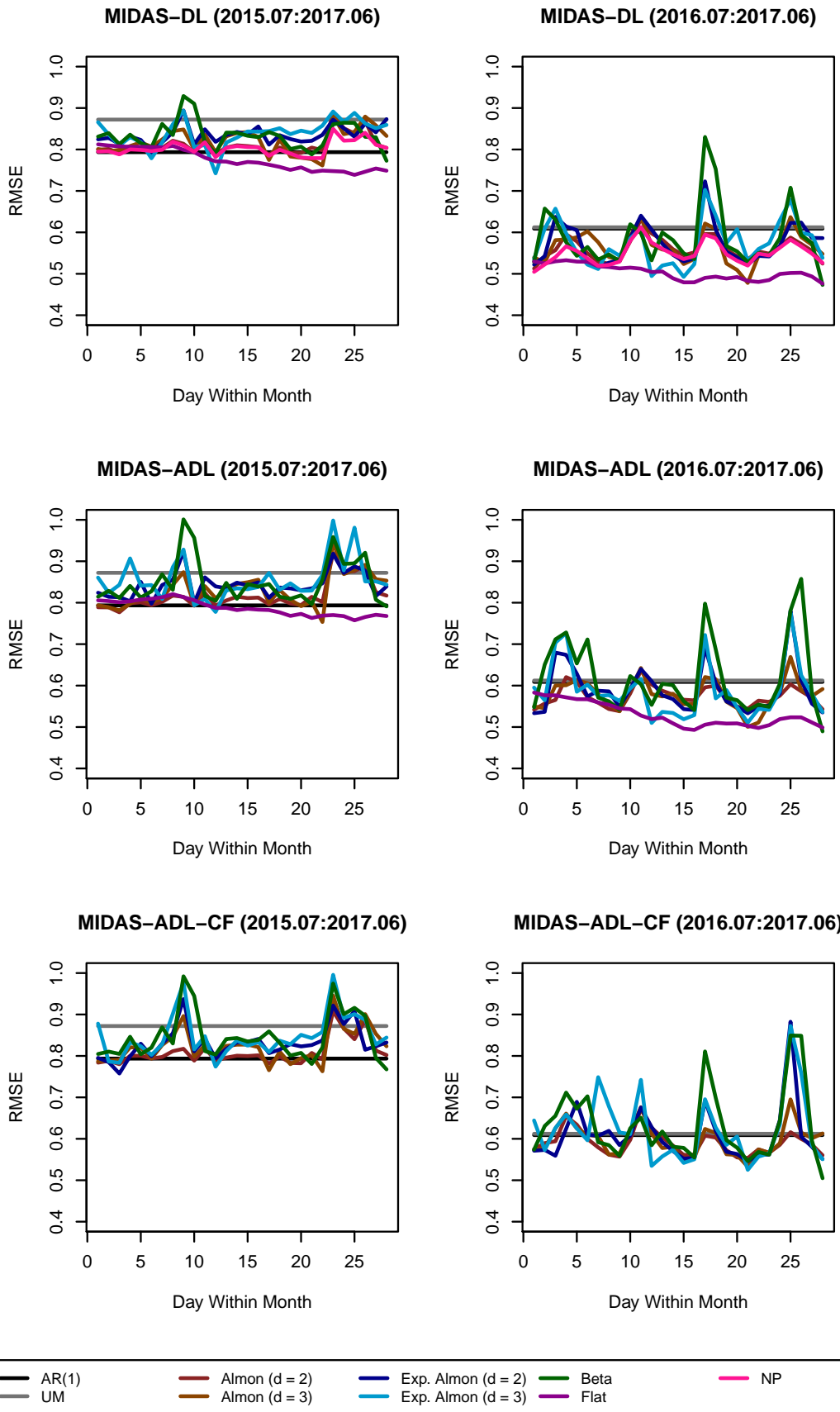


Figure 3: Evolution of RMSFE of period  $T + 2$  forecasts within a month for selected models with  $L_x = 2$

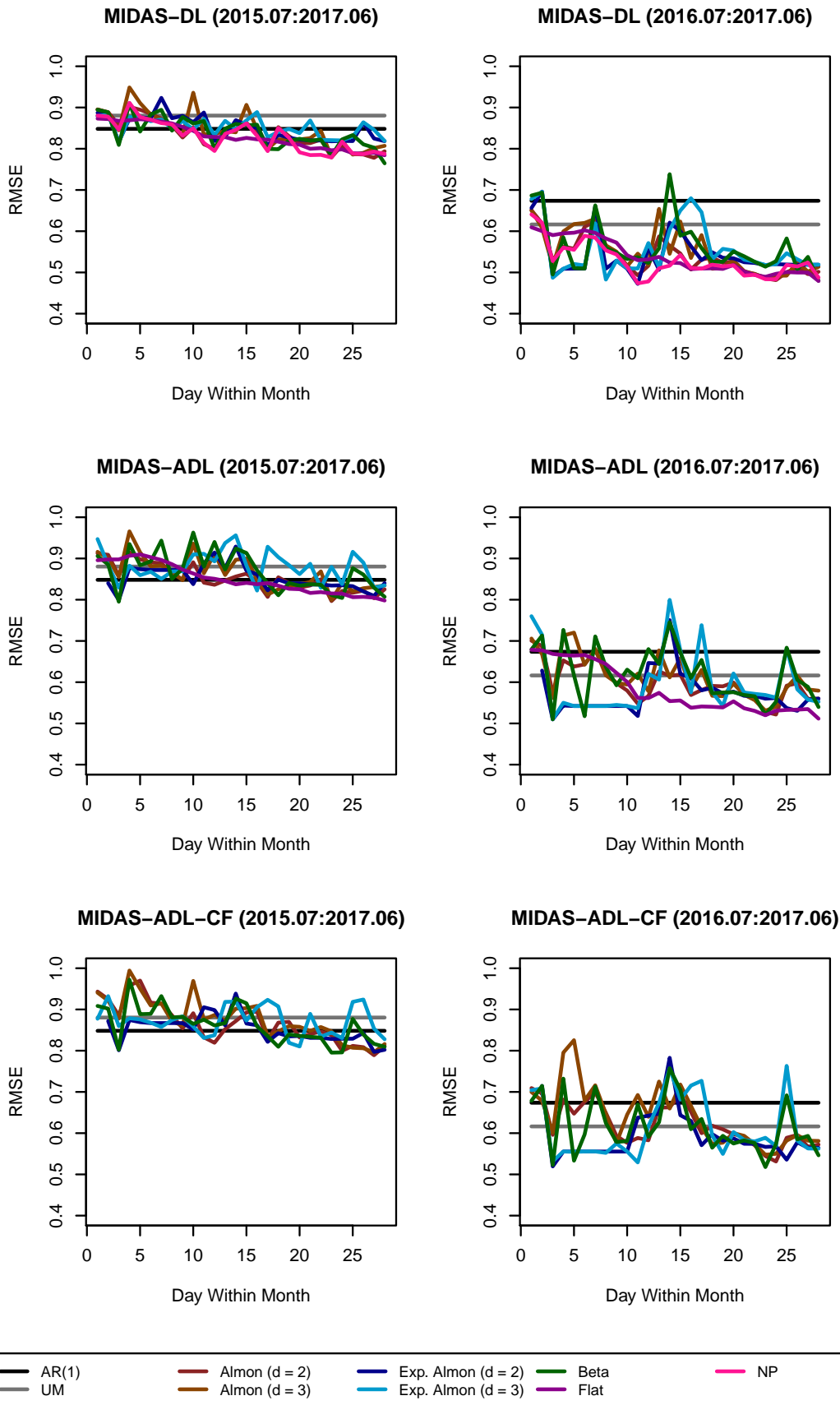
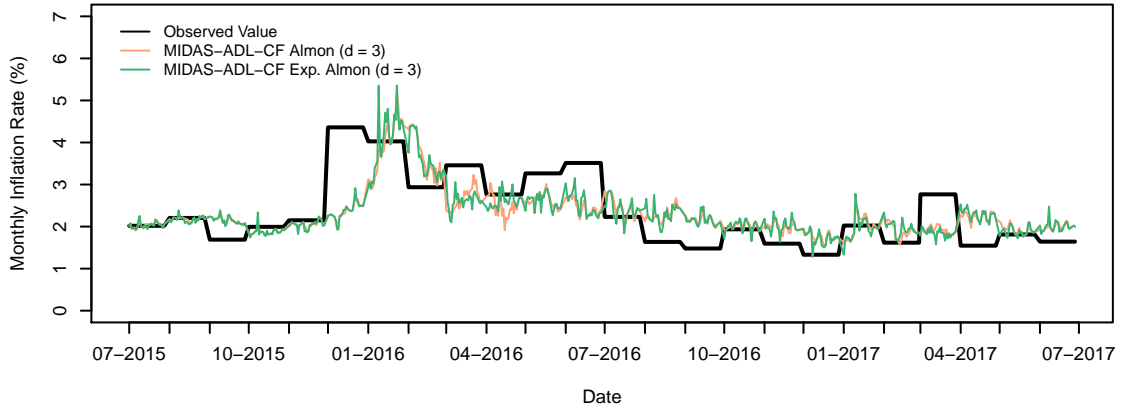
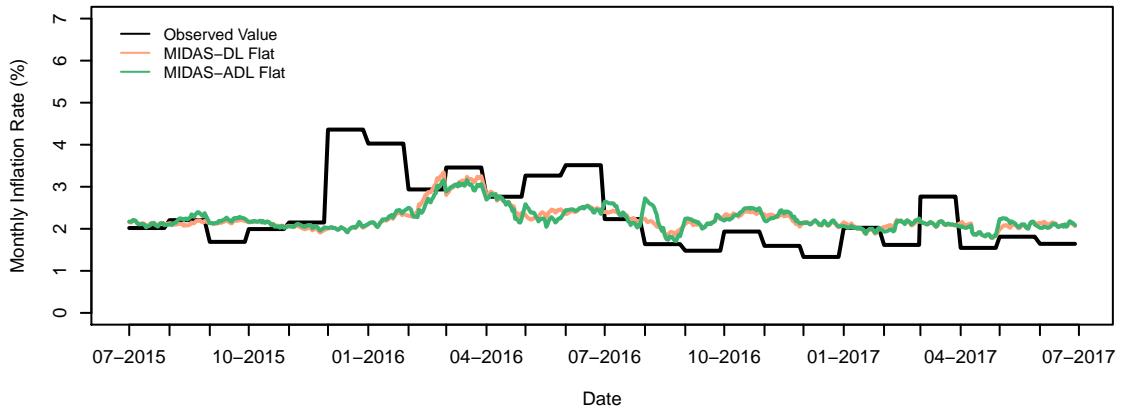


Figure 4: Evolution of RMSFE of period  $T + 3$  forecasts within a month for selected models with  $L_x = 1$

(a) Intraproduct forecasts for period T+1



(b) Intraproduct forecasts for period T+2



(c) Intraproduct forecasts for period T+3

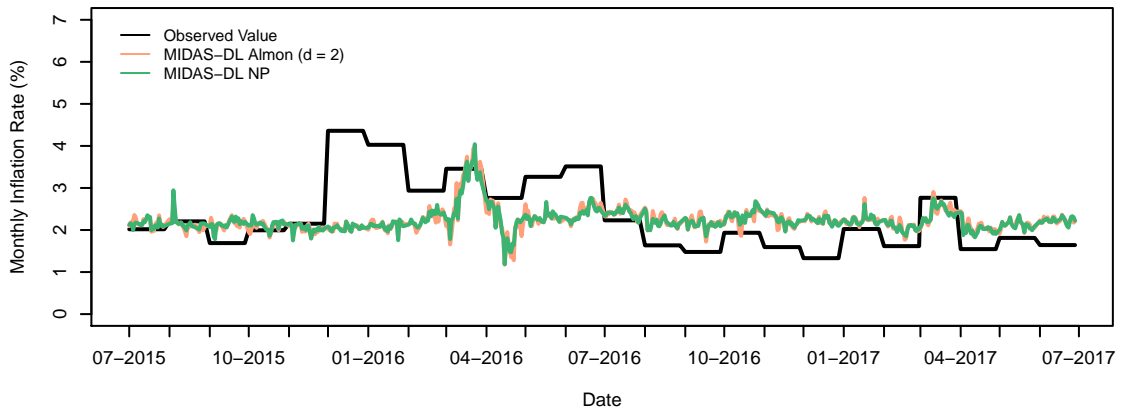


Figure 5: Evolution of intraproduct forecasts for selected models and forecast horizons

## 5 Conclusion

For some particular MIDAS specifications, there is a slight improvement compared to the low frequency benchmark autoregression and the recursive unconditional mean. In principle, this would imply that high frequency online price indices have a good potential to forecast future behavior of consumer inflation for immediate horizons in Argentina. This could serve as a useful complementary tool to assess the out-of-sample performance of perhaps more sophisticated models. Future research could focus on building an alternative variable such as a daily financial factor as suggested by [Monteforte and Moretti \(2013\)](#) or comparing against measures of market expectations in order to further validate the findings of this paper.

## References

- Andreou, E., Ghysels, E., and Kourtellis, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2):240–251.
- Armesto, M. T., Engemann, K. M., and Owyang, M. T. (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review*, 92(6):521–36.
- Atkeson, A. and Ohanian, L. E. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis*, 25(1):2.
- Breitung, J. and Roling, C. (2015). Forecasting inflation rates using daily data: A nonparametric MIDAS approach. *Journal of Forecasting*, 34(7):588–603.
- Cavallo, A. (2013). Online and official price indexes: Measuring Argentina’s inflation. *Journal of Monetary Economics*, 60(2):152–165.
- Clements, M. P. and Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. *Journal of Business & Economic Statistics*, 26(4):546–554.
- Duarte, C. (2014). Autoregressive augmentation of MIDAS regressions. Banco de Portugal Working Paper Series.
- Faust, J. and Wright, J. H. (2009). Comparing Greenbook and reduced form forecasts using a large realtime dataset. *Journal of Business & Economic Statistics*, 27(4):468–479.
- Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82(4):540–554.
- Forni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1):57–82.
- Forni, C. and Marcellino, M. G. (2013). A survey of econometric methods for mixed-frequency data. Norges Bank Working Paper Series.
- Ghysels, E., Kvedaras, V., and Zemlys, V. (2016). Mixed frequency data sampling regression models: The R package midasr. *Journal of Statistical Software*, pages 1–35.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models. CIRANO 2004s-20 Working Papers Series.
- Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.
- Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.
- Guérin, P. and Marcellino, M. (2013). Markov-switching MIDAS models. *Journal of Business & Economic Statistics*, 31(1):45–56.
- Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H., and Lee, T. C. (1985). *Theory and Practice of Econometrics*. John Wiley & Sons, New York NY, 2nd edition.
- Marcellino, M. and Schumacher, C. (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for german GDP. *Oxford Bulletin of Economics and Statistics*, 72(4):518–550.

- Marcellino, M., Stock, J. H., and Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of econometrics*, 135(1):499–526.
- Miller, J. I. (2013). Mixed-frequency cointegrating regressions with parsimonious distributed lag structures. *Journal of Financial Econometrics*, 12(3):584–614.
- Monteforte, L. and Moretti, G. (2013). Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, 32(1):51–61.
- Nash, J. C. and Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, 43(9):1–14.
- Rodriguez, A. and Puggioni, G. (2010). Mixed frequency models: Bayesian approaches to estimation and prediction. *International Journal of Forecasting*, 26(2):293–311.
- Schumacher, C. (2016). A comparison of MIDAS and bridge equations. *International Journal of Forecasting*, 32(2):257–270.
- Stock, J. H. and Watson, M. W. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39(s1):3–33.

## Appendix

Table A.1: Out-of-sample predictive performance, ratio to RMSE of autoregressive model  $\times 100$

	$h = 0$			$h = 1$			$h = 2$		
	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$
<i>Almon (d = 2)</i>									
MIDAS-DL	100.8	101.1	<b>91.0</b>	112.0	101.3	<b>101.2</b>	<b>93.6</b>	94.8	97.2
MIDAS-ADL	94.9	96.1	<b>91.4</b>	<i>111.1</i>	103.0	<b>102.2</b>	97.3	<b>96.3</b>	99.0
MIDAS-ADL-CF	97.5	100.8	<b>90.1</b>	111.8	<i>101.1</i>	<b>100.5</b>	95.7	<b>95.1</b>	100.5
<i>Almon (d = 3)</i>									
MIDAS-DL	105.0	100.9	<b>92.2</b>	115.9	104.9	<b>102.4</b>	95.2	<b>93.9</b>	96.9
MIDAS-ADL	98.8	94.8	<b>92.9</b>	<i>115.2</i>	107.5	<b>102.9</b>	98.4	<b>94.8</b>	98.8
MIDAS-ADL-CF	102.2	103.7	<b>91.3</b>	118.0	<i>103.7</i>	<b>101.5</b>	96.2	<b>94.9</b>	101.5
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	107.7	108.3	<b>91.3</b>	122.8	110.0	<b>105.0</b>	96.5	96.8	<b>96.2</b>
MIDAS-ADL	96.8	<b>94.3</b>	96.8	<i>113.5</i>	<b>105.8</b>	106.2	99.1	<b>94.9</b>	<b>94.9</b>
MIDAS-ADL-CF	100.5	101.7	<b>89.6</b>	114.5	<i>104.9</i>	<b>104.6</b>	<i>94.6</i>	<b>92.7</b>	97.4
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	115.5	104.5	<b>91.5</b>	<i>107.8</i>	108.3	<b>105.1</b>	<b>96.5</b>	96.8	96.9
MIDAS-ADL	99.0	<b>93.4</b>	96.3	113.7	<b>106.3</b>	107.1	98.4	98.4	<b>98.3</b>
MIDAS-ADL-CF	96.3	104.2	<b>90.4</b>	128.5	106.4	<b>105.3</b>	<b>97.6</b>	98.2	98.3
<i>Beta</i>									
MIDAS-DL	102.4	102.0	<b>92.9</b>	104.9	<b>97.4</b>	98.9	<b>90.1</b>	92.3	92.1
MIDAS-ADL	<b>91.3</b>	102.2	99.2	103.5	<b>99.7</b>	100.4	<b>95.1</b>	96.3	96.1
MIDAS-ADL-CF	94.0	101.4	<b>90.4</b>	104.2	<b>96.7</b>	98.1	<i>95.4</i>	<b>95.2</b>	95.6
<i>Flat</i>									
MIDAS-DL	193.7	99.8	<b>91.6</b>	104.6	<b>94.4</b>	96.9	<b>92.4</b>	96.1	97.6
MIDAS-ADL	<i>151.2</i>	96.5	<b>91.4</b>	<i>103.4</i>	<b>96.8</b>	102.5	<b>94.1</b>	98.9	100.6
<i>Nonparametric</i>									
MIDAS-DL	99.3	101.2	<b>91.5</b>	109.4	<b>101.3</b>	101.5	<b>92.8</b>	94.2	97.3
<i>Unconditional Mean</i>									
$\bar{y}$	123.7	123.7	123.7	109.9	109.9	109.9	103.8	103.8	103.8

Notes: The evaluation sample comprises 24 data points, from 2015.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .

Table A.2: Out-of-sample predictive performance, ratio to RMSFE of unconditional mean  $\times 100$ 

	<i>h</i> = 0			<i>h</i> = 1			<i>h</i> = 2		
	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3
<i>Almon (d = 2)</i>									
MIDAS-DL	81.5	81.8	<b>73.6</b>	101.9	92.2	<b>92.1</b>	<b>90.2</b>	91.3	93.6
MIDAS-ADL	76.7	77.7	<b>73.9</b>	<i>101.1</i>	93.7	<b>93.0</b>	93.7	<b>92.8</b>	95.4
MIDAS-ADL-CF	78.9	81.5	<b>72.8</b>	101.7	92.0	<b>91.4</b>	92.2	<b>91.6</b>	96.9
<i>Almon (d = 3)</i>									
MIDAS-DL	84.9	81.6	<b>74.5</b>	105.4	95.5	<b>93.2</b>	<i>91.7</i>	<b>90.5</b>	93.4
MIDAS-ADL	79.9	76.7	<b>75.1</b>	<i>104.8</i>	97.8	<b>93.6</b>	94.8	<b>91.3</b>	95.2
MIDAS-ADL-CF	82.6	83.8	<b>73.8</b>	107.4	<i>94.4</i>	<b>92.4</b>	92.7	<b>91.4</b>	97.8
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	87.1	87.5	<b>73.9</b>	111.7	100.1	<b>95.5</b>	<b>93.0</b>	93.3	92.7
MIDAS-ADL	78.2	<b>76.3</b>	78.3	<i>103.2</i>	<b>96.2</b>	96.6	95.4	<b>91.4</b>	<b>91.4</b>
MIDAS-ADL-CF	81.2	82.2	<b>72.4</b>	104.2	<i>95.4</i>	<b>95.2</b>	<i>91.1</i>	<b>89.3</b>	93.8
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	93.4	84.5	<b>73.9</b>	<i>98.1</i>	98.5	<b>95.6</b>	<b>93.0</b>	93.3	93.3
MIDAS-ADL	80.1	<b>75.5</b>	77.9	103.4	<b>96.7</b>	97.4	94.8	94.8	<b>94.7</b>
MIDAS-ADL-CF	77.8	84.2	<b>73.1</b>	116.9	96.8	<b>95.8</b>	<b>94.0</b>	94.6	94.7
<i>Beta</i>									
MIDAS-DL	82.8	82.5	<b>75.1</b>	95.5	<b>88.6</b>	90.0	<b>86.8</b>	<i>89.0</i>	88.7
MIDAS-ADL	<b>73.8</b>	82.6	80.2	<i>94.1</i>	<b>90.7</b>	91.4	<b>91.7</b>	92.8	92.6
MIDAS-ADL-CF	76.0	<i>81.9</i>	<b>73.1</b>	94.8	<b>88.0</b>	89.2	92.0	91.7	<b>92.1</b>
<i>Flat</i>									
MIDAS-DL	156.6	80.7	<b>74.1</b>	95.1	<b>85.8</b>	88.2	<b>89.1</b>	92.6	<i>94.0</i>
MIDAS-ADL	<i>122.3</i>	<i>78.0</i>	<b>73.9</b>	<i>94.1</i>	<b>88.0</b>	93.3	<b>90.7</b>	95.3	96.9
<i>Nonparametric</i>									
MIDAS-DL	80.3	81.8	<b>74.0</b>	99.6	<b>92.2</b>	92.4	<b>89.4</b>	90.8	93.7
<i>Autoregressive</i>									
<i>p</i> = 1	80.9	80.9	80.9	91.0	91.0	91.0	96.3	96.3	96.3

Notes: The evaluation sample comprises 24 data points, from 2015.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .



Table A.3: Out-of-sample predictive performance, ratio to RMSFE of autoregressive model  $\times 100$ 

	<i>h</i> = 0			<i>h</i> = 1			<i>h</i> = 2		
	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$	$L_x = 1$	$L_x = 2$	$L_x = 3$
<i>Almon (d = 2)</i>									
MIDAS-DL	94.8	92.0	<b>76.0</b>	109.7	<b>86.1</b>	90.6	<b>74.4</b>	75.9	75.5
MIDAS-ADL	85.5	88.6	<b>79.1</b>	105.7	<b>89.1</b>	91.9	83.1	<b>81.7</b>	83.1
MIDAS-ADL-CF	93.0	97.1	<b>77.9</b>	107.8	<b>92.0</b>	93.9	<b>85.0</b>	86.5	86.8
<i>Almon (d = 3)</i>									
MIDAS-DL	100.5	97.8	<b>76.9</b>	115.8	<b>90.0</b>	92.7	76.2	<b>73.3</b>	74.2
MIDAS-ADL	90.8	92.1	<b>81.2</b>	112.1	97.2	<b>94.2</b>	86.0	<b>80.7</b>	83.7
MIDAS-ADL-CF	104.0	104.1	<b>80.1</b>	115.5	100.6	<b>95.8</b>	86.2	<b>85.5</b>	91.7
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	103.2	88.7	<b>77.1</b>	132.5	96.2	<b>90.2</b>	<b>77.0</b>	77.1	79.0
MIDAS-ADL	91.4	84.3	<b>84.2</b>	107.1	<b>87.9</b>	91.2	83.1	<b>82.2</b>	<b>82.2</b>
MIDAS-ADL-CF	108.2	97.7	<b>78.4</b>	107.4	<b>90.5</b>	94.2	<b>83.8</b>	84.6	84.6
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	105.1	98.4	<b>76.5</b>	89.7	<b>88.4</b>	89.1	<b>77.0</b>	<b>77.0</b>	77.8
MIDAS-ADL	89.1	83.3	<b>85.1</b>	117.8	<b>88.1</b>	89.9	82.0	82.0	<b>81.0</b>
MIDAS-ADL-CF	102.1	101.3	<b>78.4</b>	163.7	<b>90.4</b>	92.6	<b>83.5</b>	87.3	87.2
<i>Beta</i>									
MIDAS-DL	104.5	87.1	<b>79.8</b>	97.5	77.7	85.5	<b>71.2</b>	75.4	77.4
MIDAS-ADL	93.3	87.0	<b>83.1</b>	93.1	<b>80.4</b>	86.6	<b>80.1</b>	85.0	83.8
MIDAS-ADL-CF	101.4	89.8	<b>80.2</b>	95.3	<b>82.9</b>	88.8	<b>81.1</b>	84.4	85.5
<i>Flat</i>									
MIDAS-DL	101.8	90.0	<b>70.9</b>	100.8	<b>78.2</b>	82.9	<b>71.1</b>	79.0	81.4
MIDAS-ADL	95.0	86.3	<b>70.6</b>	98.3	<b>81.9</b>	87.1	<b>75.9</b>	81.2	86.1
<i>Nonparametric</i>									
MIDAS-DL	94.2	94.2	<b>77.0</b>	106.8	<b>86.4</b>	90.4	<b>72.4</b>	74.8	75.5
<i>Unconditional Mean</i>									
$\bar{y}$	100.5	100.5	100.5	100.5	100.5	100.5	91.5	91.5	91.5

Notes: The evaluation sample comprises 12 data points, from 2016.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .

Table A.4: Out-of-sample predictive performance, ratio to RMSFE of unconditional mean  $\times 100$ 

	<i>h</i> = 0			<i>h</i> = 1			<i>h</i> = 2		
	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3	<i>L<sub>x</sub></i> = 1	<i>L<sub>x</sub></i> = 2	<i>L<sub>x</sub></i> = 3
<i>Almon (d = 2)</i>									
MIDAS-DL	94.4	91.6	<b>75.7</b>	109.2	<b>85.7</b>	90.1	<b>81.4</b>	83.0	82.6
MIDAS-ADL	85.1	88.2	<b>78.7</b>	105.1	<b>88.7</b>	91.4	90.9	<b>89.3</b>	90.8
MIDAS-ADL-CF	92.6	96.6	<b>77.6</b>	107.3	91.5	93.4	<b>92.9</b>	94.6	94.8
<i>Almon (d = 3)</i>									
MIDAS-DL	100.0	97.4	<b>76.5</b>	115.2	<b>89.5</b>	92.3	83.3	<b>80.1</b>	<i>81.1</i>
MIDAS-ADL	90.4	91.7	<b>80.8</b>	111.5	96.7	<b>93.7</b>	94.0	<b>88.3</b>	91.5
MIDAS-ADL-CF	103.5	103.7	<b>79.7</b>	114.9	100.1	<b>95.3</b>	94.3	<b>93.4</b>	100.2
<i>Exp. Almon (d = 2)</i>									
MIDAS-DL	102.7	88.3	<b>76.7</b>	131.8	95.7	<b>89.7</b>	<b>84.1</b>	84.2	86.4
MIDAS-ADL	90.9	84.0	<b>83.8</b>	106.6	<b>87.4</b>	90.7	90.9	<b>89.8</b>	89.8
MIDAS-ADL-CF	107.7	97.3	<b>78.1</b>	106.9	<b>90.0</b>	93.7	<b>91.6</b>	92.5	92.5
<i>Exp. Almon (d = 3)</i>									
MIDAS-DL	104.6	97.9	<b>76.1</b>	89.2	<b>87.9</b>	88.7	<b>84.2</b>	<b>84.2</b>	85.1
MIDAS-ADL	88.7	82.9	<b>84.7</b>	117.2	<b>87.6</b>	89.4	89.6	89.6	<b>88.5</b>
MIDAS-ADL-CF	101.6	100.8	<b>78.0</b>	162.8	<b>90.0</b>	92.2	<b>91.2</b>	95.5	95.3
<i>Beta</i>									
MIDAS-DL	104.0	86.7	<b>79.5</b>	97.0	<b>77.3</b>	85.1	<b>77.8</b>	82.4	84.7
MIDAS-ADL	92.8	86.6	<b>82.7</b>	92.6	<b>80.0</b>	86.1	<b>87.6</b>	92.9	91.6
MIDAS-ADL-CF	100.9	89.4	<b>79.9</b>	94.8	<b>82.5</b>	88.3	<b>88.6</b>	92.3	93.4
<i>Flat</i>									
MIDAS-DL	101.3	89.6	<b>70.6</b>	100.3	<b>77.8</b>	82.5	<b>77.7</b>	86.3	88.9
MIDAS-ADL	94.6	85.9	<b>70.2</b>	97.8	<b>81.5</b>	86.6	<b>83.0</b>	88.7	94.1
<i>Nonparametric</i>									
MIDAS-DL	93.7	93.7	<b>76.6</b>	106.3	<b>85.9</b>	89.9	<b>79.1</b>	81.8	82.5
<i>Autoregressive</i>									
<i>p</i> = 1	99.5	99.5	99.5	99.5	99.5	99.5	109.3	109.3	109.3

Notes: The evaluation sample comprises 12 data points, from 2016.07 to 2017.06. Characters in **bold** indicate the best number of variables,  $L_x$ , for each model and forecast horizon,  $h$ . Characters in *italics* indicate the best model for each number of variables,  $L_x$ , and forecast horizon,  $h$ .

Reconquista 266, C1003ABF  
Ciudad Autónoma de Buenos Aires, Argentina  
Phone | 54 11 4348-3582  
E-mail | [investig@bcra.gob.ar](mailto:investig@bcra.gob.ar)  
Web | [www.bcra.gob.ar](http://www.bcra.gob.ar)



*ie* | BCRA  
INVESTIGACIONES ECONÓMICAS