NOWCASTING MEXICO'S QUARTERLY GDP USING FACTOR MODELS AND BRIDGE EQUATIONS

NOWCASTING DEL PIB DE MÉXICO USANDO MODELOS DE FACTORES Y ECUACIONES PUENTE

Oscar de J. Gálvez-Soriano

Banco de México/University of Houston

Resumen: Se evalúan cinco modelos de Nowcasting: un modelo de factores dinámicos (MFD), dos ecuaciones puente (BE) y dos basados en componentes principales (PCA). Los resultados indican que el promedio de los pronósticos de las BE es estadísticamente mejor que el del resto de los modelos considerados, de acuerdo con la prueba de precisión de pronósticos de Diebold-Mariano. Utilizando información en tiempo real, se encuentra que el promedio de las BE es más preciso que la mediana de los pronósticos de los analistas encuestados por Bloomberg, que la mediana de los especialistas que responden la encuesta de expectativas del Banco de México y que la estimación oportuna del PIB publicada por el INEGI.

Abstract: I evaluate five nowcasting models that I used to forecast Mexico's quarterly GDP in the short run: a dynamic factor model (DFM), two bridge equation (BE) models and two models based on principal components analysis (PCA). The results indicate that the average of the two BE forecasts is statistically better than the rest of the models under consideration, according to the Diebold-Mariano accuracy test. Using real-time information, I show that the average of the BE models is also more accurate than the median of the forecasts provided by the analysts surveyed by Bloomberg, the median of the experts who answer Banco de México's Survey of Professional Forecasters and the rapid GDP estimate released by INEGI.

Clasificación JEL/JEL Classification: C32, C38, C53, E52.

Palabras clave/keywords: forecasting; state space model; principal component analysis; monetary policy; Kalman filter; Diebold-Mariano test; pronósticos, modelos de estado espacio; análisis de componentes principales; política monetari; filtro de Kalman; prueba de Diebold-Mariano

Fecha de recepción: 22 I 2019 *Fecha de aceptación: 29 VII 2019*

https://doi.org/10.24201/ee.v35i2.402

1. Introduction

Information on the current state of the economy is a crucial aspect in decision making for policymakers. Nonetheless, key statistics on the evolution of the economy are available only with a certain delay, which is why we rely on forecasting procedures in order to get timely estimations of those key figures. This is the case of series that are calculated on a quarterly basis, such as the Gross Domestic Product (GDP). Indeed, it is of a particular importance for central banks to use precise short-term GDP estimates in guiding monetary policies that will affect the long run; in the words of Lucas (1976:22), "...forecasting accuracy in the short-run implies reliability of long-term policy..."

Indeed, an increasingly common forecasting practice among central banks is nowcasting, which has been broadly studied in developed countries, such as China, France, Germany, Ireland, New Zealand, Norway, Spain, Switzerland, UK, United States (US), among others (and whose practice is much less generalized in developing economies, where it is mainly used by the IMF) with the purpose of obtaining timely GDP estimations. In particular, Mexico's National Institute of Statistics (INEGI) publishes its estimate of Mexico's GDP and the official measure of National Accounts four and seven weeks after the end of the reference quarter, respectively. And although forecasts from Bloomberg and from Banco de México's Survey of Professional Forecasters (SPF) are updated on a regular basis, a more precise estimate of GDP would be helpful for policymakers.

For example, during the third quarter of 2019 (July-September), policymakers would prefer to take decisions based on that quarter's data and on the short-term forecasts of economic activity. However, in Mexico, the firstGDP estimated figures were not released until October 2019, which means that policymakers had to wait about 120 days for needed GDP data and at least 30 days after the end of the quarter, in order to have the first reliable estimate of economic activity for that quarter (the rapid GDP estimate is conducted by INEGI). Furthermore, official GDP statistics for the third quarter are not published until the end of November, which means a larger delay in their availability and, hence, reducing its usefulness for decision making purposes.

My goal in this paper is to find a nowcasting model that is more accurate than the consensus GDP estimates of professional forecasters and the GDP estimations released by INEGI. I propose five nowcasting models that forecast quarterly GDP using monthly data (which are inspired by the work of Rünstler and Sédillot, 2003; Baffigi, Golinelli, and Parigi, 2004; Giannone, Reichlin and Small, 2008). These include a dynamic factor model (DFM), two bridge equation (BE) models and two principal components (PCA) models, which are the most common methods used for nowcasting. All of them use high-frequency variables (monthly data) to predict a lower frequency variable (quarterly GDP). The high frequency variables are related to economic activity and include data on sales, production, employment, and foreign trade as well as financial variables.

Although previous research has already proposed nowcasting models for Mexican GDP (Caruso, 2018; Dahlhaus, Guénette, and Vasishtha, 2017), none compare nowcasting models, nor do they include BE or PCA models in their analysis. Rather, they compare their forecasts with those of the SPF. In fact, my results suggest that the BE models produce Mexican quarterly GDP forecasts which are more accurate than both the DFM and those reported by the SPF (and are even more accurate than the preliminary GDP estimations made by INEGI), which opens a new discussion about the convenience of using a more complicated model, such as the DFM, *versus* a "simpler" approach (i.e. the BE model), when forecasting the GDP growth rate of a developing economy.

Furthermore, the aforementioned authors have only evaluated their models within their data sample, which reduces their robustness for practical applications because both the GDP and the monthly series are constantly revised. In an attempt to deal with those revisions, Delajara, Hernández and Rodríguez (2016) retrieved data series originally published for the five variables of their DFM with which they were able to perform a pseudo real-time analysis; however, they do not consider BE models in their analysis.

In my research I evaluate the BE model forecasts in real time, which has never been done before. This evaluation was possible because I kept a record of the forecasts of all the proposed models during 12 consecutive quarters (from the second quarter of 2014, henceforth 2014-II, to the first quarter of 2017, henceforth 2017-I). Based on these records and using the Diebold-Mariano test, I find that the BE models generate more accurate predictions than the median forecasts of the analysts surveyed by Bloomberg, the median of the forecasts provided by the specialists who answer the SPF and the rapid GDP estimate released by INEGI.

Moreover, the analysis of the BE model's forecast errors suggests that their variance decreases consistently with the inclusion of more information as new observed data are available. Indeed, for the period from 2014-II to 2017-I, more information led to a significant reduction in the variance of errors from forecasts made one month before INEGI published the official GDP growth, so that 75 percent of the time the margin of error of the BE is, in absolute terms, less than 0.1 percentage points of the observed quarterly GDP growth, which is a quite low forecast error, one hardly ever reached by professional forecasters or by the INEGI's timely GDP estimate in the same period of study.

The structure of this document is as follows: after introduction, section two presents a review of the literature that has proposed nowcasting models; in section three the BE, the DFM and the PCA models are theoretically described; section four shows the data that will be used to apply the models of section three to the case of Mexico, while section five shows the main results and section six presents the discussion and conclusions.

2. Literature review

The first researches that used high frequency variables to predict the quarterly GDP were based on BE models (Rünstler and Sédillot, 2003; Baffigi, Golinelli, and Parigi, 2004). The BE method consists of using dynamic and linear equations where the explanatory variables are formed with the quarterly aggregates of daily or monthly series. However, the BE models are not precisely parsimonious due to the large number of explanatory variables included. In order to reduce the number of independent variables, Klein and Sojo (1989) use the PCA model and, years later, Stock and Watson (2002a,b) confirmed the efficiency of the forecasts provided with this method.

Recently, Giannone, Reichlin and Small (2008) developed a method to obtain forecasts of the GDP growth rates using the factors of a state-space representation whose coefficients are estimated with the filter developed by Kalman (1960). This method is known in the literature as DFM and has been widely used to forecast the GDP of developed countries (Rünstler *et al.*, 2009; Banbura and Modugno, 2014; Angelini *et al.*, 2011; Yiu and Chow, 2011; and de Winter, 2011, are some examples). However, most of the research using DFM is based on large information sets that, according to Álvarez, Camacho and Perez-Quiros (2012), imply a strong assumption about the orthogonality of the factors obtained. A large number of series will show at least some degree of correlation, which suggests that this assumption of orthogonality does not always hold.[1] The empirical

[1] For example, Giannone, Reichlin and Small (2008) use 200 monthly indicators of US economic activity, while Álvarez, Camacho and Perez-Quiros (2012)

findings of Alvarez, Camacho and Perez-Quiros (2012) indicate that, although neither of their two DFM (with large and with small information sets) had consistently superior results over the other, the accuracy of the forecasts generated by the model with the small information set was equal to or greater than the one of the model with the large information set. Recently, other authors (Camacho and Doménech, 2012; Barnett, Chauvet and Leiva-Leon, 2016; Delajara, Hernandez and Rodriguez, 2016; Dahlhaus, Guénette and Vasishtha, 2017; and Caruso, 2018) have chosen to use small-scale models. Thus, based on the literature described, in this document I only consider small information sets in the proposed models.

The first research suggesting a nowcasting model for Mexico was conducted by Liu, Matheson and Romeu (2012), who compared a nowcast and the forecast of the GDP growth rate using five models: an autoregressive model (AR), BE, VAR bivariate, Bayesian VAR and DFM, for 10 Latin American countries.[2] Their results indicate that, for most of the countries considered, the monthly data flow helps to improve the accuracy of the estimates and that the DFM produces, in general, more precise nowcasts and forecasts relative to other model specifications. However, one of the exceptions was the case of Mexico, where better results were achieved with the Bayesian VAR.

Likewise, the first antecedent of the timely estimate published by INEGI was proposed by Guerrero, García and Sainz (2013), who suggested a procedure to make timely estimates of Mexico's quarterly GDP using bridge equations based on vector autoregressive (VAR) models. Guerrero, García and Sainz (2013) structure the forecast by economic sectors and then by activity, analogously to how INEGI presents the official data. Their results suggest that their estimates have relatively small forecast errors, so they recommend using their model to estimate Mexico's quarterly GDP. However, Caruso (2018) does not consider this proposal as a nowcast, but catalogs it as a backcast since, along with the model of Guerrero, García and Sainz (2013), the estimate of GDP growth is not available until 15 days after the conclusion of the reference quarter.

Due to this lag, Caruso (2018) prefers the use of a DFM based on Doz, Giannone and Reichlin (2012), and Banbura and Modugno (2014). Using this model, the author forecasts Mexico's GDP growth

---

show the convenience of using small sets of 12 indicators over large sets of 146 US indicators.

[2] Argentina, Brazil, Chile, Colombia, the Dominican Republic, Ecuador, Mexico, Peru, Uruguay and Venezuela.

using monthly series from Mexico and the United States. His results indicate that the DFM generates more precise forecasts than those offered by the IMF, the OECD, the forecasts of the SPF and the forecasts of the analysts surveyed by Bloomberg. However, the comparisons made by Caruso (2018) between the forecasts of his DFM and those of the specialists are not necessarily the most appropriate, since the latter are published in real time, while the DFM he estimates include data revisions.

Similarly, Dahlhaus, Guénette and Vasishtha (2017) use a DFM based on Giannone, Reichlin and Small (2008) in order to model and forecast the GDP of Brazil, Russia, India, China and Mexico (BRIC-M). The DFM that the authors use for Mexico includes variables similar to those of the DFM that I propose in this research, except for the price indicators that I do not consider and the Global Indicator of Economic Activity (IGAE, for its initials in Spanish),[3] which is not included by the authors. Dahlhaus, Guénette and Vasishtha. (2017) compare the forecasts of their DFM with those generated by an $AR(2)$ and a $MA(4)$; their results suggest that the DFM produces better forecasts than the reference models.

In another research similar to that of Caruso (2018), Delajara, Hernandez and Rodriguez (2016) use a DFM to forecast Mexico's GDP, but, unlike the former, the authors test their model in pseudo real-time. Delajara, Hernandez and Rodriguez (2016) use five variables of the economic activity in Mexico and compare the forecasts of their model with those offered by the SPF. Their results show that their DFM produces more accurate forecasts than those of the SPF. However, with the exception of Liu, Matheson and Romeu (2012), none of the aforementioned researches consider BE models in their comparisons. In this sense, the present document provides new evidence about the convenience of the use of BE models to make nowcasting of Mexico's GDP growth.

3. Nowcasting

Nowcasting can be defined as a forecast of economic activity of the recent past, the present and the near future. These forecasts are calculated as the linear projection of quarterly (contemporary) GDP given

[3] The IGAE is an economic indicator published monthly by INEGI, approximately eight weeks after the end of the reference month and which represents 93.9% of GDP in the base year, 2008=100.

a dataset that consists of greater-frequency (usually monthly) figures. Intuitively, specifications are estimated through ordinary least squares (OLS) in which the GDP is a function of its own lags, as well as of the contemporaneous and lagging values of the independent variables that are constructed from a set of monthly indicators.

Formally, let us denote quarterly GDP growth as ${y_t}^Q$, and the monthly information set as $\mathbf{X_t}$, where the superscript Q refers to quarterly variables and the subscript t refers to time (months or quarters). We want to estimate the GDP of the current quarter, so we calculate the linear projection of GDP given the information set $\mathbf{X_t^Q}$:

$$Proy\left[y_t^Q|\mathbf{X_t^Q}\right]$$

We start from the fact that our information set is composed of n variables, $\mathbf{X^Q_{it|v_j}}$, where $i = 1, ..., n$ identifies the individual series and $t = 1, ..., T_{v_j}$ denotes the time of publication, which varies between series v_j according to its publication schedule. The differences among the publication schedules of the different indicators produce a problem known in the literature *as jagged edges or ragged edges*. In this sense, the first forecasts offered by nowcasting (at the beginning of the reference quarter) are made despite missing observations from the end (edges) of the series.

The nowcast is calculated as the expected value of GDP given the available information and the underlying model, $\mathcal{M}$, under which a conditional expectation is calculated:

$$\hat{y}_t^Q = E\left[y_t^Q|\mathbf{X^Q_{v_j}};\mathcal{M}\right]$$

Usually, a linear model is used, where the regressors are the variables of the information set (or the factors) and the dependent variable is quarterly GDP growth. The uncertainty (variance) associated with this projection is:

$$V_{y^Q_{t|v_j}} = E\left[\left(\hat{y}^Q_{t|v_j} - y_t^Q\right)^2|\mathbf{X^Q_{v_j}};\mathcal{M}\right]$$

Because the number of observed data grows over time, the variance of the error decreases, that is:

$$V_{y^Q_{t|v_j}} \leq V_{y^Q_{t|v_{j-1}}}$$

3.1. *Bridge equation models*

In bridge equation models, factors are not calculated. Instead, the same monthly indicators are used as explanatory variables. Let us denote the vector of n monthly indicators as $X_t = (X_{1,t}, \ldots, X_{n,t})$, for $t = 1, \ldots, T$. The bridge equation is estimated with quarterly aggregates, $X^Q_{i,t}$, of the three corresponding monthly data.

$$X^Q_{i,t} = \frac{1}{3}\left(X_{i,1} + X_{i,2} + X_{i,3}\right)$$

These quarterly aggregates are used as regressors in the bridge equation models to obtain a quarterly GDP growth forecast:

$$y^Q_t = \mu + \psi\left(L\right)\mathbf{X^Q_t} + \varepsilon^Q_t$$

where μ is the coefficient of the constant, $\psi\left(L\right) = \psi_0 + \psi_1 L^1 + \ldots + \psi_p L^p$ denotes the lag polynomial, and ε^Q_t is the error term, which is assumed white noise with normal distribution.

3.2. *Dynamic factor models*

The DFM were developed and applied for the first time by Giannone, Reichlin and Small (2008) to forecast the quarterly GDP growth of United States. However, the idea of using state space models (SSM) in order to obtain coincident US indicators was originally proposed and studied by Stock and Watson (1988, 1989), based on Geweke's original proposal (1977).

Consider the vector of n monthly series $\mathbf{X_t} = \left(X_{1,t}, \ldots, X_{n,t}\right)'$, for $t = 1, ..., T$. The dynamics of the factors considered by Giannone, Reichlin and Small (2008) is given by the following state space representation:

$$\mathbf{X_t} = \mathbf{\Lambda f_t} + \xi_\mathbf{t} \qquad \xi_\mathbf{t} \sim \mathbb{N}\left(\mathbf{0}, \mathbf{\Sigma}_\xi\right) \tag{1}$$

$$\mathbf{f_t} = \sum\nolimits_{\mathbf{i=1}}^{\mathbf{p}} \mathbf{A_i f_{t-i}} + \zeta_\mathbf{t} \tag{2}$$

$$\zeta_\mathbf{t} = \mathbf{B}\eta_\mathbf{t} \qquad \eta_\mathbf{t} \sim \mathbb{N}\left(\mathbf{0}, \mathbb{I}_\mathrm{q}\right) \tag{3}$$

where $\mathbf{\Lambda}$ is an $n \times r$ matrix of weights, which implies that equation (1) relates the monthly series X_t to an $r \times 1$ vector of latent factors $\mathbf{f_t} = (f_{1,t}, \ldots, f_{r,t})^{'}$ plus an idiosyncratic component $\xi_\mathbf{t} = (\xi_{1,t}, \ldots, \xi_{n,t})^{'}$. It is assumed that the latter is white noise with a diagonal covariance matrix $\mathbf{\Sigma}_\xi$. Equation (2) describes the law of movement of latent factors $\mathbf{f_t}$, which are driven by an autoregressive process of order p, plus a *q-dimensional* white noise component, where $\mathbf{B}$ is an $n \times q$ matrix, and where $q \leq r$. Thus, the number of common shocks, q, is less than or equal to the number of common factors, r. Hence, $\xi_\mathbf{t} \sim \mathbb{N}(0, B, B^{'})$. Finally, $\mathbf{A_1}, ..., \mathbf{A_p}$ are $r \times r$ matrices of coefficients and, in addition, it is assumed that the stochastic process of $\mathbf{f_t}$ is stationary.[4]

3.3. *Principal components analysis models*

The PCA method is a statistical technique that is typically used for data reduction.[5] This implies that from a large information set, eigenvectors are obtained from the decomposition of the covariance matrix of the original series. These eigenvectors describe series of uncorrelated linear combinations of the variables that contain most of the variance of the entire information set. In my research I use this technique to make predictions with those eigenvectors, generating more parsimonious models.

---

[4] The complete development of the SSM proposed by Giannone, Reichlin and Small (2008) is found in Forni *et al.* (2009).

[5] The PCA method can be attributed to the work of Pearson (1901) and Hotelling (1933). For a useful introductory explanation see Afifi, May and Clark (2012).

Starting with the information set $\mathbf{X_t}$ of n monthly series, let us define the $n \times n$ covariance matrix of the information set as $\mathbf{\Sigma}_{X_t}$. Where $\mathbf{\Phi}$ is an $n \times n$ orthogonal matrix, whose columns are the $\mathbf{c}$ eigenvectors of $\mathbf{\Sigma}_{X_t}$, and a diagonal matrix, $\mathbf{\Psi}$, where the elements of its main diagonal are the eigenvalues of $\mathbf{\Sigma}_{X_t}$, such that,

$$\mathbf{\Phi}'\mathbf{\Sigma}_{X_t}\mathbf{\Phi} = \mathbf{\Psi}$$

The n vectors C_t are orthogonal and are arranged according to the proportion of the variance they represent of the set $\mathbf{X_t}$.

4. Data

In this paper, I use quarterly series of Mexico's GDP at constant prices, from the first quarter of 1993 (1993-I) to the first quarter of 2017 (2017-I). I consider three information sets as explanatory variables. The first one (CI-1) includes 25 monthly indicators that, when converted to quarterly indicators as explained above, have a correlation with GDP greater than 0.30 (this correlation is calculated with respect to the quarterly variations of seasonally adjusted series). However, if the indicator is published in the first week after the reference month, I keep it in the information set, even if correlation is less than 0.30. An additional criterion is that I only use monthly series that are available since 1993, in order to have explanatory variables whose observation period corresponds that of the Mexican GDP figures.

The second information set (CI-2) consists of eight variables, some of which are included in CI-1 set but have a rate of correlation with GDP growth of at least 0.40 (instead of .30 as above). I no longer consider the initial data availability date as a criterion, so now there are indicators that were not included in the CI-1. The third set (CI-3) is exclusive for the DFM estimation and in it I use 11 variables that I chose arbitrarily from CI-1 and CI-2 sets because they represent different and representative sectors of the Mexican economy (see appendix A1 for a detailed list of variables included in each information set).[6]

The three information sets can be described as being formed by "hard" variables and "soft" variables. The former, offer timely and

[6] A special information set is needed for the DFM due to the way it is estimated, and due to the fact that including variables with few observations (such as the Monthly Survey on Commercial Companies EMEC- that begins in 2008) makes it more difficult to reach a recursive solution to the model.

coincident information on the economic activity, while the latter, although more timely and better able to anticipate economic activity, come from perception surveys, and are therefore more likely to be inaccurate. Indeed, the hard indicators are very important for the estimation of quarterly GDP, since they have a relatively greater weight in the estimated factors, while the soft indicators have a lower impact, which reflects the fact that most of their contribution is mainly due to their timely availability. Moreover, the literature has shown that the variables that provide the most timely information contribute to an improvement in the estimation only at the beginning of the quarter and that once the updated data of the hard indicators is included, their contribution fades (Banbura *et al.*, 2013).

Regarding the use of the data, I seasonally adjust all the variables included in the information set with the X-12-ARIMA program,[7] except those that are already seasonally adjusted by INEGI before publication, and those that come from the perception surveys (because they do not present a seasonal pattern). In addition, I only use stationary series; thus I transform non- stationary series by means of a logarithmic difference, based on unit root tests (see appendix A2, table A2). Finally, following a convention in the literature, I standardize all the series before applying the methodologies of nowcasting.

5. Results

To deal with the jagged edges problem, I elaborate ARIMA models for each monthly variable, in order to forecast the missing observations at the end of the series. In this way, to generate the quarterly GDP growth forecast,[8] the BE, the DFM and the PCA models are estimated from previously completed information sets with ARIMA equations. This allows me to compare the predictive power of each model regardless of how it deals with incomplete information sets. All this

[7] With the change of the base year (in October 31st, 2017), INEGI started using the X-13ARIMA-SEATS program in order to seasonally adjust most of the Mexican time series. Nonetheless, I worked with the previous base year (2008=100). Thus, all the seasonal adjustment models I developed were based on the document "Procedure for obtaining seasonal adjustment models with the X12-ARIMA program" of the Specialized Group on Seasonal Adjustment (GED) of the Specialized Committee of Macroeconomic Statistics and National Accounts of Mexico.

[8] I define growth here as the rate of variation of GDP of a quarter with respect to the previous one, using seasonally adjusted series.

despite the fact that both the PCA and the DFM models could make forecasts of their own factors. It is important to mention that all results from this section were obtained with GDP data published until 2014-II, except those of subsection 5.6, which were conducted in real time (from 2014-II to 2017-I).

5.1. *BE model estimation*

I used the CI-1 and CI-2 data sets to obtain the BE1 and BE2 models, respectively. Theoretically, a BE model uses an OLS method for its estimation with lags of the variables included in the model; however, most of the aforementioned research proposes ARIMA models with exogenous variables to improve the accuracy of the estimates. Consequently, I estimate the following equation:

$$\phi(L)\,\mathrm{y}_t^Q = \theta(L)\,\varepsilon_t^Q + \psi(L)\,\mathbf{X_t^Q} \tag{4}$$

where all the variables were treated with a logarithmic difference to approximate a growth rate.

Table 1
Bridge equation estimation models

Variables	*BE1*		*BE2*	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*
IGAE_t	.713	(.029)	.746	(.048)
Consumption_{t-2}	-.080	(.018)		
$\text{Industrial Activity}_{t-2}$	.161	(.021)	.059	(.036)
$\text{Manufacturing}_{t-1}$	.076	(.023)		
Import_{st}	.055	(.008)		
IndustryUS_t			.083	(.030)
$\text{Construction}_{t-2}$	-.041	(.009)		
ANTAD_t	.083	(.012)	.049	(.029)
$\text{ExportNoPetrolManu}_t$	-.059	(.008)	.048	(.010)
$\text{ForwardIndicator}_{t-1}$	-.270	(.047)		

Table 1
(continued)

Variables	*BE1*		*BE2*	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*
BMV_t	.249	(.049)		
$Cement_{t-3}$	-.035	(.006)	.020	(.010)
$AMIA_{t-1}$	-.012	(.003)	-.008	(.003)
$M4_{t-3}$	.030	(.010)		
$EMEC_t$			.023	(.012)
$AutoParts_t$	.003	(.001)		
$Aluminium_{t-3}$	.013	.002		
$Export_{t-1}$	.053	(.009)		
$Hotel_t$	.034	(.007)		
$Electricity_{t-1}$	.041	(.020)		
$IndustrialGas_{t-1}$	-.004	(.002)		
$Rail_{t-4}$	-.017	(.005)		
$Tires_{t-2}$	-.006	(.002)		
$Movie_{t-2}$	-.004	(.002)		
$Fuel_{t-2}$	-.044	(.016)		
$TIIE_{t-3}$	-.072	(.028)		
ER_{t-3}	.065	(.042)		
AR(1)	-.647	(.100)	-.434	(.123)
Adjusted R-squared	.986		.942	
S.E. of regression	.002		.002	
Durbin-Watson stat	2.088		2.190	
Akaike info. criterion	-9.823		-9.267	
Schwarz criterion	-8.967		-8.905	
Hannan-Quinn criter.	-9.477		-9.124	

Note: Models shown with data available until July 31st, 2017 in order to forecast GDP growth rate of 2017-II, which was published in August 22nd, 2017.

We have that $\phi(L)$, $\theta(L)$ and $\psi(L)$ are lag polynomials whose order was determined based on the error autocorrelation function, the Q statistic of Ljung-Box, statistical significance tests of estimated coefficients, and the conventional information criterions (AIC, BIC and HQC). Finally, ε_t^Q is assumed white noise with normal distribution.

Note that, during the nowcast of the previous section, the BE models were updated according to the data revisions as well as the seasonal adjustments. This means that models are changing as needed. As an example, in table 1 I show the model estimation for the BE models with data available until July 2017, which is the latest available model from estimations made in real time (The autocorrelation analysis and the normality test are shown in appendix A3.1). Table 1 also shows how some variables could lose their significant levels (see ER_{t-3}, $Industrial\ Activity_{t-2}$, $ANTAD_t$, and $EMEC_t$ in table 1) due to data revisions and due to changes in the seasonally adjustment models, but I included those variables nevertheless, in order to keep track of them and to have comparable forecast among quarters, despite data revisions.

5.2. *DFM estimation*

To estimate the coefficients of the DFM, I use the 11 variables of the CI-3 data set, and the maximum likelihood method (ML). In turn, the parameters of the likelihood function are estimated with the Kalman filter.[9] This requires initial values for the state variables, as well as a covariance matrix to begin the recursive process. For this, I use the method suggested in Hamilton (1994b).[10] The estimated state space model (the state and observation equations, respectively) is shown by equations (6) and (7), where $\eta_\mathbf{t} \sim \mathbb{N}(\mathbf{0}, \mathbf{1})$ and $\xi_\mathbf{t} \sim \mathbb{N}\left(\mathbf{0}, \sigma_\mathbf{i}^\mathbf{2}\right)$.

To use the factor(s) obtained from the estimation of the DFM (figure 2) it is necessary to make it quarterly. This is done by taking the average of the three monthly observations corresponding to each quarter $f_t^Q = \frac{1}{3}(f_1 + f_2 + f_3)$. In this way, GDP growth is estimated with the following equation.

$$y_t^Q = \alpha + \beta f_t^Q + \theta(L)\,\varepsilon_t^Q \tag{5}$$

[9] For further information on SSM and the Kalman filter see Hamilton (1994a, 1994b), Harvey (1989), and Brockwell and Davis (1991).

[10] That is, the initial values are obtained with the estimated coefficients with a linear regression of X_t on f_t, since the latter has an autoregressive structure.

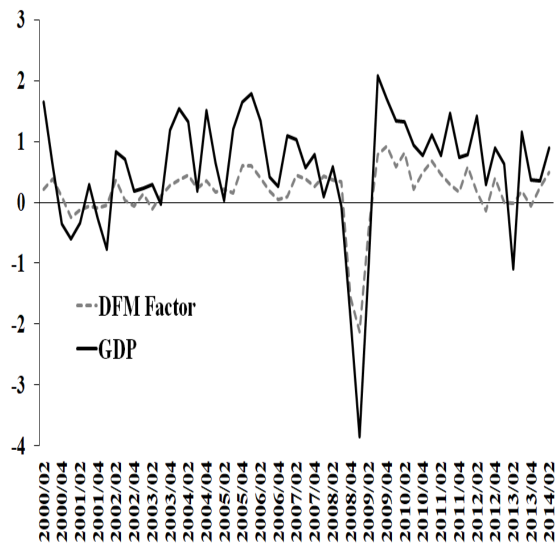
$$
\begin{pmatrix} f_t \\ \xi_{1,t} \\ \xi_{2,t} \\ \xi_{3,t} \\ \xi_{4,t} \\ \xi_{5,t} \\ \xi_{6,t} \\ \xi_{7,t} \\ \xi_{8,t} \\ \xi_{9,t} \\ \xi_{10,t} \\ \xi_{11,t} \end{pmatrix}
=
\begin{pmatrix}
\varphi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \delta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \delta_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \delta_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \delta_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \delta_5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \delta_6 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_7 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_8 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_9 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_{10} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_{11}
\end{pmatrix}
\begin{pmatrix} f_{t-1} \\ \xi_{1,t-1} \\ \xi_{2,t-1} \\ \xi_{3,t-1} \\ \xi_{4,t-1} \\ \xi_{5,t-1} \\ \xi_{6,t-1} \\ \xi_{7,t-1} \\ \xi_{8,t-1} \\ \xi_{9,t-1} \\ \xi_{10,t-1} \\ \xi_{11,t-1} \end{pmatrix}
+
\begin{pmatrix} \eta_t \\ \varsigma_{1,t} \\ \varsigma_{2,t} \\ \varsigma_{3,t} \\ \varsigma_{4,t} \\ \varsigma_{5,t} \\ \varsigma_{6,t} \\ \varsigma_{7,t} \\ \varsigma_{8,t} \\ \varsigma_{9,t} \\ \varsigma_{10,t} \\ \varsigma_{11,t} \end{pmatrix}
\quad (6)
$$

$$
\begin{pmatrix} IGAE_t \\ Consumption_t \\ IMAI_t \\ M_t \\ ANTAD_t \\ Cement_t \\ M4_t \\ EMEC_t \\ X_t \\ CFE_t \\ PEMEX_t \end{pmatrix}
=
\begin{pmatrix}
\lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\lambda_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\lambda_3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\lambda_4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\lambda_5 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\lambda_6 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\lambda_7 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\lambda_8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\lambda_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\lambda_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\lambda_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix} f_t \\ \xi_{1,t} \\ \xi_{2,t} \\ \xi_{3,t} \\ \xi_{4,t} \\ \xi_{5,t} \\ \xi_{6,t} \\ \xi_{7,t} \\ \xi_{8,t} \\ \xi_{9,t} \\ \xi_{10,t} \\ \xi_{11,t} \end{pmatrix}
\quad (7)
$$

That is, the nowcast of the quarterly GDP is a linear function of the factor. The estimation procedure of this linear regression of equation (5) uses the method presented in Cochrane and Orcutt (1949) to obtain robust estimators in the presence of residual autocorrelation.

Figure 1

Quarterly GDP growth vs. DFM factor



An example for the DFM estimation is shown in table 2, where data is available until July 2017. This model is the latest available from estimations made in real time (The autocorrelation analysis and the normality test are shown in appendix A3.2).

Table 2

Dynamic factor model estimation

Variables	*Coefficient*	*Std. Error*
Factor_t	.042	(.002)
Constant	.001	(.000)
MA(1)	-.590	(.113)
Adjusted R-squared	.871	
S.E. of regression	.003	
Durbin-Watson stat	2.040	

Table 2
(*continued*)

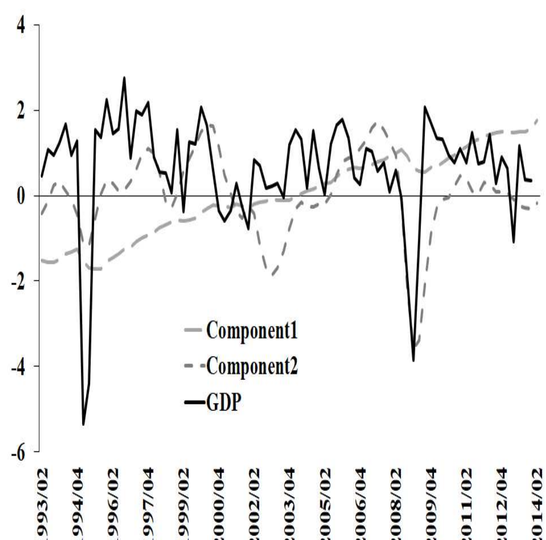
Variables	*Coefficient*	*Std. Error*
Akaike info criterion	-8.546	
Schwarz criterion	-8.416	
Hannan-Quinn criter.	-8.494	

Note: Models shown with data available until July 31st, 2017 in order to forecast GDP growth rate of 2017-II, which was published in August 22nd, 2017.

5.3. *PCA estimation*

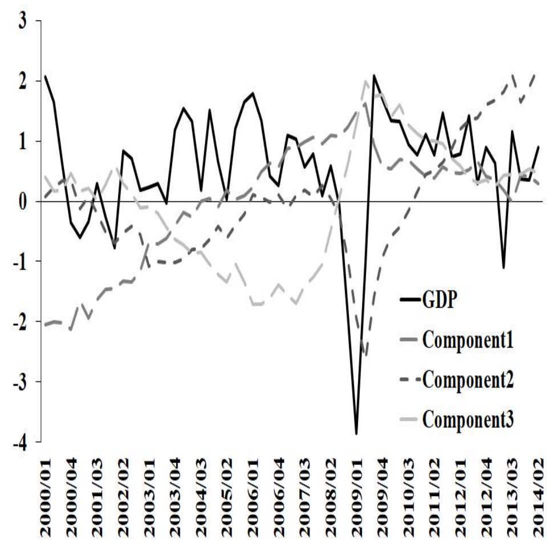
Finally, I used principle components analysis, PCA, and CI-1 and CI-2 data sets, obtaining PCA1 and PCA2 models, respectively. In the analysis I obtain all the principal components (eigenvectors), c_t, that arise from each information set. However, I only consider the k components ($k < n$) whose eigenvalue is greater than or equal to unity, according to the criterion developed in Kaiser (1958).

Figure 2
Quarterly GDP growth vs. components of PCA1



After obtaining these components I rotate them in order to distribute the variance explained by each one, so I use the *varimax method*, which also allows me to maintain the property of orthogonality between the components even after having distributed their variance. In turn, maintaining the property of orthogonality implies a significant reduction in the number of components obtained from the high correlation between the variables considered. This helps maintain parsimony in the model that forecasts GDP growth.

Figure 3
Quarterly GDP growth vs. components of PCA2



Consequently, and analogously to the methods described above, in the PCA approach, I use the k quarterly principal components (figures 2 and 3) as regressors in the following linear equation to forecast quarterly GDP growth:

$$y_t^Q = \gamma + \delta \mathbf{c_t^Q} + \varepsilon_t^Q \tag{8}$$

where δ is an $n \times k$ matrix of coefficients, $\mathbf{c_t^Q}$ is the vector of k quarterly principal components, and ε_t^Q is the error term, which I assume to be white noise with normal distribution. The estimation of equation (8) is held by the generalized least squares (GLS) technique by using the Cochrane-Orcutt method.

The practical example for the PCA estimation is shown in table 3, where data is available until July 2017. This model is the latest available using estimations made in real time (the autocorrelation analysis and the normality test are shown in appendix A3.3).

Table 3
Principal components analysis model estimation

Variables	*PCA1*		*PCA2*	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*
$Factor1_t$	.079	(.012)	.050	(.006)
$Factor1_{t-1}$	-.082	(.012)	-.033	(.010)
$Factor1_{t-2}$			-.016	(.005)
$Factor2_t$	.008	(.001)		
$Factor2_{t-1}$	-.008	(.001)		
$Factor3_t$	.007	.003		
$Factor3_{t-1}$	-.008	(.003)		
Constant	.005	(.001)	.002	(.001)
Adjusted R-squared		.836		.795
S.E. of regression		.005		.004
Durbin-Watson stat		2.010		2.295
Akaike info criterion		-7.597		-8.063
Schwarz criterion		-7.330		-7.866
Hannan-Quinn criter.		-7.489		-7.985

Note: Models shown with data available until July 31st, 2017 in order to forecast GDP growth rate of 2017-II, which was published in August 22nd, 2017.

5.4. *Diebold-Mariano tests for series within sample*

Under the hypothesis that the accuracy of the forecasts can be improved by taking the average or the median of the five models, I consider both of these approaches as additional forecasts. I also consider the average of BE forecasts, as well as a univariate model, AR(1), which is included as a reference. This AR model uses the quarterly GDP series as input, treating it with a logarithmic difference to induce stationarity and to approximate the GDP growth rate.

In order to evaluate the predictive power of each model and thus discern which is more appropriate to perform nowcasting, I use the modified Diebold-Mariano (DM) test proposed by Harvey, Leybourne, and Newbold (1997), hereafter referred to as the HLN-modified DM test, for small samples.[11] Table 4 summarizes this test comparing each model with the rest. In the main diagonal, the mean square error (MSE) of each model is indicated in bold and the columns show which model is more accurate, according to this test. I indicate the statistical significance of error differences in each pair of compared models with asterisks. The test considers the quarters that go from 2009-I to 2016-II, the financial crisis of 2009 is included.

The results of the HLN-modified DM test suggest that the forecasts generated with the DFM and with the BE models were more accurate than those obtained with the PCA2 model, but with inconclusive results with respect to the AR and the PCA1 models. Although there are no statistically significant differences between the forecast errors of the BE models and those of the DFM, there are significant differences between the forecasts of the average of BE models and the DFM.

Moreover, I find that the forecasts using this average of BE models are more accurate than those obtained with the mean or median of all the models. Indeed, those forecasts obtained the smallest MSE (MSE=0.026), which implies an error of 14 hundredths compared to the observed seasonally adjusted quarterly GDP growth (table 5). Based on these results, I conclude that the average of the BE models is the best predictor of quarterly GDP of the models analyzed.

In order to analyze the robustness of my results, I performed the HLN-modified DM test under a different loss criterion. To do this, I use the Mean Absolute Forecast Error (MAE) as the loss criterion and a Bartlett kernel to compute the long-term variance of the differences series. The results show that the accuracy of the forecasts generated with the DFM is better than the univariate model and the PCA models, with statistically significant differences. This result is consistent with the findings of Giannone, Reichlin, and Small (2008), Rünstler *et al.* (2009) and Banbura and Modugno (2014), who have proposed the use of DFM when nowcasting. However, this new test strengthens my previous conclusion that the forecasts generated by all the models (including the DFM) are surpassed by the average of BE, with statistically significant differences (see appendix A5, table A3).

[11] The details of this test are presented in appendix A4.

Table 4

HLN-modified DM test (MSE loss criterion)

Models	*AR*	*PCA1*	*PCA2*	*DFM*	*BE1*	*BE2*	*Mean*	*Median*	*Mean BE*
AR	**.808**								
PCA1	PCA1	**.665**							
PCA2	PCA2	PCA2	**.542**						
DFM	DFM	DFM	DFM**	**.120**					
BE1	BE1	BE1*	BE1***	BE1	**.045**				
BE2	BE2	BE2*	BE2***	BE2	BE1	**.056**			
Mean (a)	Mean	Mean	Mean***	DFM	BE1***	BE2*	**.124**		
Median (a)	Median	Median	Median***	Median	BE1	BE2	Median*	**.058**	
Mean (BE)	MeanBE*	MeanBE*	MeanBE***	MeanBE*	MeanBE***	MeanBE***	MeanBE**	MeanBE**	**.026**

Notes: (a) = all models; p-value for statistically significance differences in MSE between compared models ***$p<0.01$, **$p<0.05$, *$p<0.1$ The sample includes forecasts from 2009-I to 2016-II. The mean squared error (MSE) is used as loss criterion and the uniform kernel distribution is used to compute the long-term variance.

Table 5

Forecast errors (from 2009-I to 2016-II)

Criterion	*AR*	*PCA1*	*PCA2*	*DFM*	*BE1*	*BE2*	*Mean*	*Median*	*Mean BE*
BIAS	.001	.295	-.146	.035	-.003	.000	.036	.015	-.001
MAE	.550	.546	.555	.243	.168	.174	.248	.165	.136
MSE	.808	.665	.542	.120	.045	.056	.124	.058	.026
RMSE	.899	.816	.736	.346	.212	.237	.353	.241	.162

Note: Forecast errors are calculated as the difference between the observed and the predicted value. The criteria shown in this table are detailed in section 5.5.

Furthermore, I also performed the modified DM tests on a smaller sample, one which excludes the period of the 2008-2009 financial crisis. Thus, I evaluate the forecasts from 2011-I until the end of the sample (see appendix A5, table A.4). The results favor the average of BE over any other model, with statistically significant differences (except when compared with the BE1 model, where my result is not conclusive). Again, the DFM offers more accurate forecasts than the AR and the PCA models, but the BE models produce more accurate forecasts than the former.

5.5. *BE forecasts within sample*

To evaluate the efficiency of the BE, I perform an analysis of the forecast errors using the following criteria, where k refers to the number of predicted periods.

Forecast Bias (BIAS):

$$BIAS = \frac{1}{k}\sum_{i=1}^{k}(\hat{y}_i - y_i)$$

Mean Squared Error (MSE):

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(\hat{y}_i - y_i)^2$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\hat{y}_i - y_i)^2}$$

I calculated these three previously described equations using two approaches, one of *rolling window* and another of *expanded window*. In the first I estimate equation (4) with data from 1993-I to 2006-IV (X_1^{56}, subscript 1 refers to the first observed data [1993-I] and the superscript 56 refers to 2006-IV, that is 56 quarters from the first observed data). I used the resulting equation to forecast four quarters

($k = 4$) forward. Analogously, I re-estimated equation (4) with data from 1993-II to 2007-I (X_2^{57}), and so on in the following way,

$$\hat{y}^Q_{j,T+i+j} = E\left[y^Q_{j,T+j}|X_j^{T+j}\right] \qquad \forall\, i = (1,4)\,,\; j = (1,28)$$

where $\hat{y}^Q_{j,T+i+j}$ refers to the forecast i quarters after the estimation made from j to $(T+J)$ quarters after; X_j^{T+j} refers to the information set that begins j quarters after the first observation and ends $T+j$ after this. Thus, the estimation window always has the same length, i.e. $T = 56$ quarters. The estimation windows were rotated 28 times until 2013-IV, so the last rotation makes forecasts until 2014-IV.

In the second approach I estimate equation (4) with X_1^{56} and I forecast four quarters ($k = 4$) forward. Next, I re-estimate equation (4) with X_2^{57}, that is, with an additional quarter, and so on in the following way,

$$\hat{y}^Q_{T+i+j} = E\left[y^Q_{T+j}|X^{T+j}\right] \qquad \forall\, i = 1,\ldots,4 \; and \; j = 1,\ldots,28$$

This implies that in the last window I perform an estimation from 1993-I to 2013-IV, with which I forecast four quarters (until 2014-IV).

Results of *rolling window* analysis show that the forecast errors generated with BE1 show an ascending behavior as k grows. That is, errors become larger when the forecast horizon is longer. However, this is not true for the BE2 model, where the smallest error was obtained by forecasting three forward periods. Similarly, the *expanded window* does not show an increase in the forecast error in any of the models (table 6).

Table 6

Forecast errors analysis in pseudo real time

	Rolling window				*Expanding window*			
Error	*Forecast horizon*							
measure	*k=1*	*k=2*	*k=3*	*k=4*	*k=1*	*k=2*	*k=3*	*k=4*
BEI								
BIAS	.004	-.059	-.053	-.029	.017	-.042	-.040	-.022
MSE	.118	.133	.149	.159	.117	.115	.138	.159
RMSE	.344	.365	.386	.399	.342	.339	.372	.399

Table 6
(*continued*)

	Rolling window				*Expanding window*			
Error	*Forecast horizon*							
measure	*k=1*	*k=2*	*k=3*	*k=4*	*k=1*	*k=2*	*k=3*	*k=4*
BE2								
BIAS	.021	.033	.013	.048	.076	.074	.026	.077
MSE	.275	.229	.183	.353	.220	.185	.134	.229
RMSE	.524	.479	.428	.594	.469	.430	.366	.478

Note: Table shows the average statistics obtained from an estimate with a rolling window and an expanded one. The size of the first window is 56 quarters in BE1 and 28 quarters in BE2; from 1993-I to 2006-IV and from 2000-I to 2006-IV, respectively.

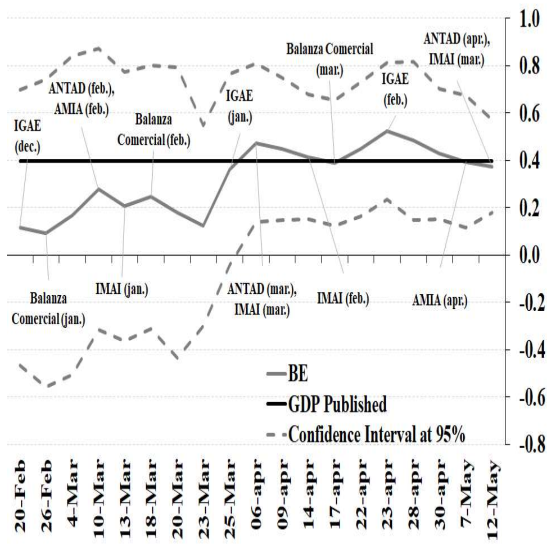
The bias of BE1 model shows an underestimation of GDP growth as the forecast horizon grows (except in $k = 4$, where the bias is reduced). On the other hand, in the BE2 model with an expanded window, the bias remains relatively constant and even decreases in $k = 3$. These results are consistent with the findings of Giannone, Reichlin y Small (2008), who suggest the use of nowcasting to forecast one step ahead and advise against using it for future quarters.

5.6. *Real-time forecasts (out of sample)*

Because the BE average provides more accurate GDP forecasts than the rest of the models considered, I used it to perform a number of real-time tests in order to analyze the evolution and sensitivity of its forecast before the publication of new information corresponding to each series that belongs to the information set.

The period of study includes quarters from 2014-II to 2017-I, and the analysis consists of observing the evolution of the forecast, updating it based on the monthly release of variables that "complete" the information set. This allows us to evaluate variables to which the forecast is more sensitive and to identify the moment at which it improves its accuracy until a reliable estimate of GDP growth is obtained. Figure 4 shows an example of how the nowcasting update behaves as each of the indicators that make up the information set are published. The forecast begins with the IGAE release of data for the third month of the quarter prior to the reference one, that is, it begins during the current quarter.

Figure 4
Nowcasting evolution in real time vs. quarterly variation of GDP (2015-I)



I recorded my GDP growth forecasts during twelve quarters, and from these forecasts I obtained the forecast errors that were grouped by "moments". I identified 12 moments of particular relevance that make up the real-time GDP forecast evolution for any quarter. These moments include indicators of interest that have important effects on nowcasting:

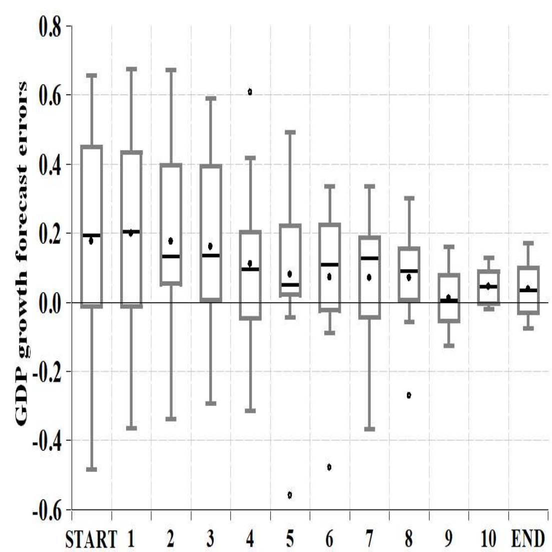
START. Publication of IGAE data; last month of previous quarter.

1. Publication of balance of trade data; first month of reference quarter.
2. Publication of car sales (AMIA) data; second month of reference quarter.
3. Publication of industrial activity (IMAI) data; first month of reference quarter.
4. Publication of balance of trade data; second month of reference quarter.
5. Publication of IGAE data; first month of reference quarter.
6. Publication of car sales (AMIA) data; third month of reference quarter.
7. Publication of industrial activity (IMAI) data; second month of reference quarter.

8. Publication of balance of trade data; third month of reference quarter.
9. Publication of IGAE data; second month of reference quarter.
10. Publication of car sales (AMIA) data; first month of following quarter.
END. Publication of industrial activity (IMAI) data; third month of reference quarter.

With this information I constructed a boxplot to evaluate the speed and efficiency of nowcasting to improve its accuracy until I obtained a reliable estimate of GDP growth. Figure 5 shows the average of the forecast errors in period t(represented by a solid point) and the median of the errors (black horizontal line). The boxes in the diagram represent the dispersion limits of the forecast errors in the central quartiles and the "arms" show the forecast errors in the first and last quartiles (the hollow circles represent extreme values). The figure shows that incorporating the balance of trade data from the second month of reference quarter (moment 4) into the data set improves the accuracy of the forecast compared to that obtained with the accumulated information until the publication of the monthly indicator of industrial activity (IMAI) from the first month of the reference quarter.

Figure 5
Nowcasting forecast errors in percentage points on quarterly GDP growth (2014-II to 2017-I)



With the release of IGAE data for the second month of reference quarter (moment 9), the forecast not only approximates the true value of GDP growth, but also reduces the variance of the forecast error considerably. This means that the model I propose can offer an accurate forecast of Mexico's GDP growth one month before INEGI publishes the official GDP data. Hence, once the IGAE data for the second month of reference quarter are included in the information set, the forecast is, on average, equal to the observed quarterly GDP growth, and 75 percent of time the margin of error is, in absolute terms, less than 0.1 percentage points of the aforementioned quarterly variation.

5.7. *Bridge equations vs. specialists*

As in the case of Caruso (2018), in this research I compare the forecasts of the "preferred" BE model against the INEGI rapid GDP estimations and the forecasts of the analysts surveyed by Bloomberg, as well as those of the SPF.[12] However, the forecasts introduced in Caruso's (2018) analysis are not comparable because the estimates in his DFM are not made in real time, while those of the specialists are and, moreover, he does not include the rapid GDP estimate published by INEGI.

To address the problem of data revisions, Delajara, Hernández y Rodríguez (2016) recover the historical series of GDP and those of the five indicators they included in their DFM to simulate the generation of forecasts in real time and, thus, improve the comparability with those offered by specialists.

In my case, I have a record from 2014-II to 2017-I of forecasts generated in real time with the five models that I propose in this research. As a result, I was able to compare the BE average records with the rapid GDP estimations, the median of the analysts surveyed by Bloomberg and the median of those registered in the SPF.[13]

[12] GDP estimations were recovered from INEGI press releases (for the period 2015-III to 2017-I) and from the Technical Note published in 2015 (for the period 2014-II to 2015-II). The forecasts of the analysts surveyed by Bloomberg were obtained from their platform, while those of the SPF were obtained from the database published by Banco de México on its website.

[13] In the results of the SPF published by Banco de México, the forecast for annual GDP growth is reported only with original series, which is why I translated these annual rates into their corresponding quarterly variations with seasonally adjusted series. To do this, I seasonally adjust the respective historical series of

Table 7
HLN-modified DM test

Models	*MeanBE*	*Bloom*	*SPF*	*INEGI*
Mean BE	**.004**			
Bloomberg	MeanBE*	**.019**		
Survey of professional forecasters	MeanBE***	Bloom	**.051**	
INEGI rapid estimation	MeanBE*	INEGI	INEGI*	**.015**

Notes: p-value for the significance of differences in MSE between compared models ***p<0.01, **p<0.05, *p<0.1. The sample includes forecasts from 2014-II to 2017-I. The Mean Squared Error (MSE) is used as loss criterion and the uniform kernel distribution is used to calculate the long-term variance. The MSE of each model is in the main diagonal in bold.

Table 8
Real-time forecast errors (from 2014-II to 2017-I)

Criterion	*Mean BE*	*Bloomberg*	*SPF*	*INEGI*
BIAS	.022	.023	.052	.077
MAE	.051	.117	.197	.098
MSE	.004	.019	.051	.015
RMSE	.065	.138	.227	.123

Note: Forecast errors are obtained as the difference between the observed and the predicted value.

To carry out the comparison I used the HLN-modified DM test for small samples. The results of this test show that the BE model's MSE is lower than that of Bloomberg's forecasts, as well as the SPF's and the rapid GDP estimations released by INEGI, with statistically

GDP using the annual variation predicted by the analysts and with the INEGI's official models.

significant differences. The BE model's MSE, 0.004, indicates that during the analysis in real time, GDP growth forecast has differed, on average, 5 hundredths of the seasonally adjusted quarterly GDP variation observed (table 8), which means it offers a timely and relatively precise forecast of Mexican GDP growth rate.

6. Discussion and conclusions

In this paper, I propose a set of models to nowcast the seasonally adjusted quarterly growth of Mexico's GDP, updating the forecasts when new information is released in the reference quarter. The forecast models that I consider are one DFM, two BE and two PCA models. I use the HLN-modified DM tests in order to evaluate the forecast errors of each model. First, the evaluation is done within sample, during the period 2009-I to 2016-II. As a reference, I include in the analysis the predictions of a univariate model (AR).

The results of the DM tests suggest that the average of the two BE models is a better predictor of quarterly Mexican GDP growth than the AR model, the DFM or the PCA models. Even compared to the mean and the median forecasts of all models (without considering the AR), the BE average is more accurate. These results were consistent under robustness checks in which I changed the loss criterion and the period of analysis. My findings contrast with those of Liu, Matheson, and Romeu (2012), who suggest the use of DFM to forecast GDP growth of emerging economies, with the exception of Mexico, where they opt for a Bayesian VAR model. However, the information set they use is substantially different from the one I propose in this document. As a preliminary explanation I suggest that the information set has such a wide variance among and within the economic variables that it is quite difficult to condense the whole information into one or a few factors. This was already noted by Gálvez-Soriano (2018) when forecasting agricultural sector growth in Mexico. This leads me to conclude that BE models are more appropriate than factor models when the dependent variable and/or the explanatory variables have relatively high variances.

In addition, I provide an analysis for predictions in real time. This was possible because I recorded the nowcasts for twelve consecutive quarters (from 2014-II to 2017-I) as new information was incorporated in each model. From this tracking I obtained the forecast errors from BE model average. My results show that the error variance declines as more information is released from the reference quarter. I also find that the model I propose in this research can offer an accurate GDP forecast a month before INEGI publishes the official National Account GDP and a week before it publishes its timely GDP estimate. Indeed, once the IGAE data for the second month of the reference quarter are included in the information set, the forecast is, on average, equal to the observed quarterly GDP variation, and 75 percent of the time the margin of error is, in absolute terms, less than 0.1 percentage points of the aforementioned quarterly variation.

Finally, I compared the nowcast with the rapid GDP estimate (INEGI), the median forecasts of Bloomberg's analysts and with the median of the SPF, using the HLN-modified DM test. The results of the test show that the BE's MSE is smaller than the MSE obtained from the median of the forecasts provided by the analysts surveyed by Bloomberg, and that it is also smaller than the MSE obtained from the median of the forecasts provided by the specialists who answer the SPF and the rapid GDP estimate released by INEGI. In all three cases the difference in MSE was statistically significant.

Acknowledgements

I am grateful for the valuable comments of Alejandrina Salcedo, Aldo Heffner, David Papell, Rodolfo Ostolaza, two anonymous reviewers from Banco de México, and two anonymous reviewers from Estudios Económicos, as well as those provided by participants in the seminars of Banco de México, the Statistics Department at ITAM and the Faculty of Sciences at UNAM. The main results of this paper were developed when I worked at Banco de México. This research was also supported by the Mexican National Council for Science and Technology (Conacyt) and the University of Houston. The views in this article correspond to the author and do not necessarily reflect those of Banco de México. ogalvezs@cougarnet.uh.edu

References

Afifi, A., S. May, and V.A. Clark, 2012. *Practical Multivariate Analysis*, fifth edition, CRC Press.

Álvarez, R., M. Camacho, and G. Perez-Quiros. 2012. Finite sample performance of small versus large scale dynamic factor models, Documento de Trabajo, núm. 1204, Banco de España.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler. 2011. Short-term forecasts of euro area GDP growth, *The Econometrics Journal*, 14(1): C25-C44.

Baffigi, A., R. Golinelli, and G. Parigi. 2004. Bridge models to forecast the euro area GDP, *International Journal of Forecasting*, 20(3): 447-460.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin. 2013. Now-casting and the real-time data flow, Working Paper Series, no. 1564, European Central Bank.

Banbura, M., and M. Modugno. 2014. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data, *Journal of Applied Econometrics*, 29(1): 133-160.

Barnett, W., M. Chauvet, and D. Leiva-Leon. 2016. Real-time nowcasting of nominal GDP with structural breaks, *Journal of Econometrics*, 191(2): 312-324.

Brockwell, P.J., and R.A. Davis. 1991. *Time Series: Theory and Methods*, second ed. New York: Springer.

Camacho, M., and R. Doménech. 2012. MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting, *Journal of the Spanish Economic Association*, 3(4): 475-497.

Caruso, A. 2018. Nowcasting with the help of foreign indicators: The case of Mexico, *Economic Modelling*, 69(C): 160-168.

Cochrane, D., and G.H. Orcutt. 1949. Application of least squares regression to relationships containing autocorrelated error terms, *Journal of the American Statistical Association*, 44(245): 32-61.

de Winter, J. 2011. Forecasting GDP growth in times of crisis: private sector forecasts *versus* statistical models, DNB Working Paper no. 320.

Dahlhaus, T., J.D. Guénette, and G. Vasishtha. 2017. Nowcasting BRIC+ M in real time, *International Journal of Forecasting*, 33(4): 915-935.

Delajara, M., F. Hernández and A. Rodríguez. 2016. Nowcasting Mexico's short-term GDP growth in real-time: A factor model *versus* professional forecasters, *Economía*, 17(1): 167-182.

Diebold, F., and R. Mariano. 1995. Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13(3): 253-263.

Doz, C., D. Giannone and L. Reichlin. 2012. A quasimaximum likelihood approach for large, approximate dynamic factor models, *Review of Economics and Statistics*, 94(4): 1014-1024.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin. 2009. Opening the black box: Structural factor models with large cross sections, *Econometric Theory*, 25(05): 1319-1347.

Gálvez-Soriano, O.J. 2018. Forecasting the agricultural sector of Mexico, in F. Pérez *et al.* (eds.), *Economía, Finanzas y Desarrollo Social en México*, ASMIIA, pp. 42-58.

Guerrero, V., A. García, and E. Sainz. 2013. Rapid estimates of Mexico's quarterly GDP, *Journal of Official Statistics*, 29(3): 397-423.

Geweke, J. 1977. The dynamic factor analysis of economic time series, in D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socio-Economic Models*, Amsterdam: North Holland.

Giannone, D., L. Reichlin, and D. Small. 2008. Nowcasting: The real-time informational content of macroeconomic data, *Journal of Monetary Economics*, 55: 665-676.

Hamilton, J.D. 1994a. State-space models, in R.F. Engel and D.L. McFadden (eds.), *Handbook of Econometrics*, pp. 3039-3080. Amsterdam: Elsevier.

Hamilton, J.D. 1994b. *Time Series Analysis*, Princeton University Press.

Harvey, A.C. 1989. *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Harvey, D., S. Leybourne, and P. Newbold. 1997. Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 13(2): 281-291.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, 24(6): 417-441.

Kaiser, H.F. 1958. The varimax criterion for analytic rotation in factor analysis, *Psychometrika*, 23(3): 187-200.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, 82(1): 35-45.

Klein, L.R., and Sojo, E. 1989. Combinations of high and low frequency data in macroeconometric models, in L.R. Klein and J. Marquez (eds.), *Economics in Theory and Practice: An Eclectic Approach*, Norwell, MA: Kluwer Academic Publishers, pp. 3-16.

Liu, P., T. Matheson, and R. Romeu. 2012. Real-time forecasts of economic activity for Latin American economies, *Economic Modelling*, 29(4): 1090-1098.

Lucas, R.E., Jr. 1976. Econometric policy evaluation: A critique, *Carnegie-Rochester Conference Series on Public Policy*, 1: 19-46.

Maddala, G.S., and I. Kim. 1998. *Unit Roots, Cointegration, and Structural Change*, Cambridge: Cambridge University Press.

Pearson, K. 1901. On lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11): 559-572.

Rünstler, G., and F. Sédillot. 2003. Short-term estimates of euro area real GDP by means of monthly data, Working Paper Series, 276, European Central Bank.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Reijer, A. Jakaitiene, P. Jelonk, A. Rua, K. Ruth, and C. van Nieuwenhuyze. 2009. Short-term forecasting of GDP using large datasets: a real-time forecast evaluation exercise, *Journal of Forecasting*, 28(7): 595-611.

Sánchez, I., and D. Peña. 2001. Properties of predictors in overdifferenced nearly nonstationary autoregression,*Journal of Time Series Analysis*, 22(1): 45-66.

Stock, J.H., and M.W. Watson. 1988. A probability model of the coincident economic indicators, NBER Working Paper Series, no. 2772.

Stock, J.H., and M.W. Watson. 1989. New indexes of coincident and leading economic indicators, in O.J. Blanchard and S. Fischer (eds.), NBER *Macroeconomics Annual 1989, Volume* 4, Cambridge: MIT Press, pp. 351-394.

Stock, J.H., and M.W. Watson. 2002a. Macroeconomic forecasting using diffusion indices, *Journal of Business and Economics Statistics*, 20(2): 147-162.

Stock, J.H., and M.W. Watson. 2002b. Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association*, 97(460): 1167-1179.

Yiu, M.S., and K.K. Chow. 2011. Nowcasting ChineseGDP: Information content of economic and financial data, *China Economic Journal*, 3(3): 223-240

Data sources

AMIA. Statistic Information. Retrieved from: http://www.amia.com.mx/prodtot.html.

ANTAD. Information Release (Comunicado de Prensa ANTAD). Retrieved from: https://antad.net/indicadores/comunicado-de-prensa/.

Banco de México. Economic Information System (Sistema de Información Económica, SIE). Retrieved from: https://www.banxico.org.mx/SieInternet/.

Banco de México. Survey of Professional Forecasters (Encuestas sobre las expectativas de los especialistas en economía del sector privado). Retrieved from (Encuestas): https://www.banxico.org.mx/SieInternet/.

Bloomberg. Forecasts of quarterly Mexican GDP. Retrieved from (Bloomberg Terminal): https://www.bloomberg.com/professional/solution/bloomberg-terminal/.

CFE. Statistic Information. Retrieved from: https://datos.gob.mx.

Federal Reserve Bank. Data. Retrieved from: https://www.federalreserve.gov/data.

INEGI. Economic Information Repository (Banco de Información Económica). Retrieved from: https://www.inegi.org.mx/sistemas/bie/.

Pemex. Statistic Information. Retrieved from: https://datos.gob.mx.

Sectur. Statistic Information (DataTur). Retrieved from: https://datatur.sectur.gob.mx/.

Appendix

A1. *Databases*

Table A.1 summarizes the variables contained in each of the three information sets (CI-1, CI-2 and CI-3), indicating the source of those variables and the correlation with quarterly GDP at a quarterly seasonally adjusted variation level for both sides; the explanatory variables and the GDP. Note that the quarterly variation induces stationarity for all series analyzed.

A2. *Variables with logarithmic differences and seasonal adjustment*

In table A.2 I list all the variables that are part of the three information sets. I also show the results of the unit root tests. Finally, I indicate which variables have already been seasonally adjustment by INEGI (marked with a check, $\surd$). The previously unadjusted variables (marked with a cross, X) were seasonally adjusted with the X12 ARIMA program.

I performed a correlation analysis between the quarterly GDP growth and the growth rate of each quarterly variable for the period 1993-2016. Then I applied unit root tests for all the variables, and found that four of the 27 variables were stationary, so I did not transform them with logarithmic differences.

It is important to mention that I applied the unit root tests to seasonally adjusted series, even knowing that these tests could be biased. In fact, according to Maddala and Kim (1998), in finite samples, the ADF and Phillips-Perron unit-root test statistics may be biased against rejecting the null hypothesis when the data are seasonally adjusted. If this were the case for the tests previously done, there would be some over-differentiated variables, which does not present a problem according to Sánchez and Peña (2001), who argue that it is better over-differentiate than to under-differentiate when using autoregressive models to generate forecasts.

Table A1

Variables of the information sets

Indicator name	*Source*	*Coefficient of correlation with GDP1*[1/] (1993-I to 2016-I)	*CI-1*	*CI-2*	*CI-3*
IGAE	INEGI	.958	√	√	√
Private consumption	INEGI	.850	√		√
Industrial activity	INEGI	.849	√	√	√
Manufactures	INEGI	.811	√		
Manufactures	INEGI	.811	√		√
Mexican imports	Banco de México	.802			
Industrial production in US	Federal Reserve System	.746		√	
Construction	INEGI	.699	√		
ANTAD sales	ANTAD	.626	√	√	√
No-oil manufacture exports	Banco de México	.617	√	√	
Forward indicator	INEGI	.559	√		
BMV	INEGI	.533	√		
Cement production	Banco de México	.478	√	√	√
AMIA vehicle production	AMIA	.470	√	√	
M4 monetary aggregate	Banco de México	.443	√		√
EMEC wholesales*	INEGI	.439		√	√
Vehicle parts	Banco de México	.387	√		
Aluminum production	Banco de México	.373	√		

Table A1
(*continued*)

Indicator name	*Source*	*Coefficient of correlation with GDP1*[1] (1993-I to 2016-I)	*CI-1*	*CI-2*	*CI-3*
Mexican exports	Banco de México	.348	√		√
Hotels (occupancy)	Sectur	.348	√		
Electricity sales	CFE	.322	√		√
Industrial gases	Banco de México	.319	√		
Train people flow	Banco de México	.288	√		
Tires production	Banco de México	.249	√		
Movie theaters occupancy	Banco de México	.181	√		
Fuel sales	PEMEX	.176	√		√
TIIE (interest rate)	INEGI	-.341	√		
Real exchange rate (Mex-US)	Banco de México	-.544	√		

Notes: The correlations of the indicators with (*) are presented for shorter periods because the series do not start in 1993. 1/The correlation coefficient was obtained with respect to the seasonally adjusted quarterly GDP variations and those of the indicator selected.

Table A2

Unit root tests

Indicator name	*Ho: The series has a unit root*						*Ho: The series is stationary*		*Logarithmic difference*	*Seasonal adjustment of INEGI*
	Augmented Dickey-Fuller Test			*Phillips-Perron Test*			*Kwiatkowski-Phillips-Schmidt-Shin Test*			
	None	*Intercept*	*Intercept and Trend*	*None*	*Intercept*	*Intercept and Trend*	*Intercept*	*Intercept and Trend*		
IGAE	1.000	.948	.395	1.000	.935	.159	[p<.01]	[p<.01]	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Private consumption	1.000	.965	.107	1.000	.955	.213	[p<.01]	[p<.01]	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Industrial activity	.995	.664	.620	.991	.659	.460	[p<.01]	[p<.01]	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Manufactures	.997	.773	.693	.990	.742	.500	[p<.01]	[p<.01]	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Mexican imports	.968	.912	.021	.991	.939	.124	p<.01	.01p<p<.05	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Industrial production in US	.664	.096	.209	.759	.364	.644	p<.01	.01<p<.05	✓	✓
	[.001]	[.010]	[.044]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
Construction	.946	.718	.497	.939	.709	.332	p<.01	.05<p<.10	✓	✓
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>.10]		
ANTAD sales	1.000	1.000	.611	1.000	.999	.615	p<.01	p<.01	✓	X
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[p>.10]	[p>10]		

Table A2
(*continued*)

Indicator name	*Ho: The series has a unit root*						*Ho: The series is stationary*		*Logarithmic difference*	*Seasonal adjustment of INEGI*
	Augmented Dickey-Fuller Test			*Phillips-Perron Test*			*Kwiatkowski-Phillips-Schmidt-Shin Test*			
	None	*Intercept*	*Intercept and Trend*	*None*	*Intercept*	*Intercept and Trend*	*Intercept*	*Intercept and Trend*		
No-oil manufacture exports	.990 [.000]	.952 [.000]	.503 [.000]	.985 [.000]	.941 [.000]	.364 [.000]	p<.01 [p>.10]	.01<p<.05 [p>.10]	√	√
Forward indicator	.657 [.000]	.000 [.000]	.001 [.000]	.678 [.000]	.012 [.001]	.047 [.003]	p>.10 [p>.10]	p>.10 [p>.10]	X	X
BMV	.651 [.000]	.000 [.000]	.000 [.000]	.652 [.000]	.015 [.000]	.066 [.003]	p>.10 [p>.10]	p>.10 [p>.10]	X	X
Cement production	.855 [.000]	.747 [.000]	.057 [.000]	.879 [.000]	.421 [.000]	.007 [.000]	p<.01 [p>.10]	.05<p<.10 [p>.10]	√	X
AMIA production	.944 [.000]	.895 [.000]	.430 [.000]	.918 [.000]	.801 [.000]	.066 [.000]	p>.01 [p>.10]	.01<p<.05 [p>.10]	√	X
M4 monetary aggregate	1.000 [.000]	1.000 [.000]	.994 [.000]	1.000 [.000]	1.000 [.000]	.997 [.000]	p<.01 [p>.10]	p<.01 [p>.10]	√	X
EMEC wholesales	.967 [.000]	.239 [.000]	.474 [.000]	.954 [.000]	.252 [.000]	.342 [.000]	p<.01 [p>.10]	p<.01 [p>.10]	√	√
Vehicle parts	.943 [.000]	.975 [.000]	.887 [.000]	.658 [.000]	.775 [.000]	.155 [.000]	p<.01 [p>.10]	p<.01 [p>.10]	√	X

Table A2
(*continued*)

Indicator name	*Ho: The series has a unit root*						*Ho: The series is stationary*		*Logarithmic difference*	*Seasonal adjustment of INEGI*
	Augmented Dickey-Fuller Test			*Phillips-Perron Test*			*Kwiatkowski-Phillips-Schmidt-Shin Test*			
	None	*Intercept*	*Intercept and Trend*	*None*	*Intercept*	*Intercept and Trend*	*Intercept*	*Intercept and Trend*		
Aluminum production	.988 [.000]	.991 [.000]	.338 [.000]	.933 [.000]	.906 [.000]	.001 [.000]	$p<.01$ $[p>.10]$	$p<.01$ $[p>.10]$	√	X
Mexican exports	.964 [.000]	.885 [.000]	.031 [.000]	.991 [.000]	.923 [.000]	.161 [.000]	$p<.01$ $[p>.10]$	$.01<p<.05$ $[p>.10]$	√	√
Hotels (occupancy)	.951 [.000]	.696 [.000]	.004 [.000]	.972 [.000]	.574 [.000]	.008 [.000]	$p<.01$ $[p>.10]$	$p<.01$ $[p>.10]$	√	X
Electricity sales	1.000 [.000]	.799 [.000]	.181 [.000]	1.000 [.000]	.780 [.000]	.037 [.000]	$p<.01$ $[p>.10]$	$p<.01$ $[p>.10]$	√	X
Industrial gases	.582 [.000]	.001 [.000]	.000 [.000]	.551 [.000]	.004 [.000]	.000 [.000]	$p<.01$ $[p>.10]$	$.01<p<.05$ $[p>.10]$	√	X
Train (flow of people)	.958 [.000]	.730 [.000]	.382 [.000]	.957 [.000]	.576 [.000]	.022 [.000]	$p<.01$ $[p>.10]$	$p<.01$ $[p>.10]$	√	X
Tires production	.353 [.001]	.438 [.010]	.148 [.046]	.080 [.000]	.000 [.000]	.000 [.000]	$p<.01$ $[p>.10]$	$.01<p<.05$ $[p>.10]$	√	X
Movie theaters occupancy	.942 [.000]	.899 [.000]	.297 [.000]	.789 [.000]	.048 [.000]	.000 [.000]	$p<.01$ $[p>.10]$	$p<.01$ $[p>.10]$	√	X

Table A2
(*continued*)

Indicator name	*Ho: The series has a unit root*						*Ho: The series is stationary*		*Logarithmic difference*	*Seasonal adjustment of INEGI*
	Augmented Dickey-Fuller Test			*Phillips-Perron Test*			*Kwiatkowski-Phillips-Schmidt-Shin Test*			
	None	*Intercept*	*Intercept and Trend*	*None*	*Intercept*	*Intercept and Trend*	*Intercept*	*Intercept and Trend*		
Fuel sales	1.000 [.000]	.951 [.000]	.899 [.000]	.999 [.000]	.818 [.000]	.955 [.000]	$p<.01$ [$p>.10$]	$p<.01$ [$p>.10$]	$\checkmark$	X
TIIE (interest rate)	.737 [.000]	.002 [.000]	.012 [.000]	.668 [.000]	.012 [.000]	.053 [.002]	$p>.10$ [$p>.10$]	$p>.10$ [$p>.10$]	X	X
Real exchange rate (Mex-US)	.736 [.000]	.000 [.000]	.001 [.000]	.718 [.000]	.029 [.000]	.118 [.001]	$p>.10$ [$p>.10$]	$p>.10$ [$p>.10$]	X	X

Notes: unit root tests were done for the period 1993-2013. The p-values are shown to reject the Ho. In blue, the tests in which the series has a unit root are highlighted. The p-value in brackets refers to the unit root tests with the differences of the original series.

A3. **Residual diagnostics**

A3.1. *Bridge equation residuals from model shown in table 1*

From the BE models I obtained residuals that fulfil the required assumptions, namely, that the residuals from BE1 do not show significant autocorrelation in the first four lags (see figure A1) and that they are normally distributed, according to the Jarque-Bera test (see figure A2).

Likewise, residuals from BE2 are uncorrelated with their own first four lags (see figure A3) and are normally distributed, as the Jarque-Bera test suggests (see figure A4).

A3.2. *Dynamic factor model residuals from estimations shown in table 2*

Residual analysis from DFM shows that errors are uncorrelated with their own first four lags (see figure A5) and that they are normally distributed, as the Jarque-Bera test suggests (see figure A6).

A3.3. *Principal components model residuals from estimations shown in table 3*

From the PCA models I obtained residuals that fulfil the required assumptions, namely, residuals from PCA1 that do not show significant autocorrelation in the first four lags (see figure A7) and that are normally distributed, according to the Jarque-Bera test (see figure A8).

Finally, residual diagnostics from PCA2 show that errors are uncorrelated with their own first four lags (see figure A9) and that they are normally distributed, according to the Jarque-Bera test (see figure A10).

Figure A1
Residual autocorrelation, BE1

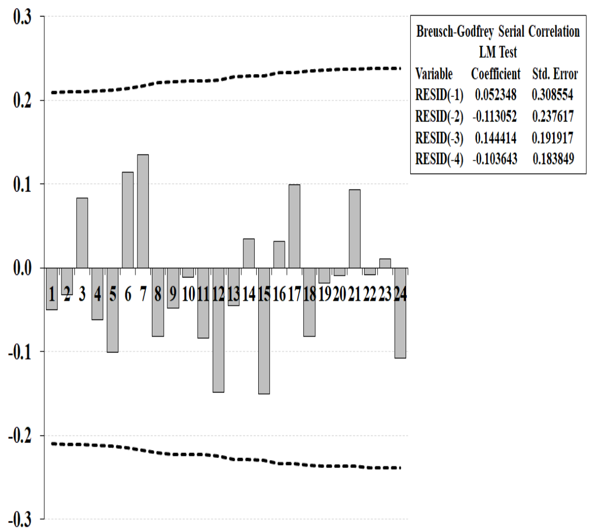


Figure A2
Residual normality test, BE1

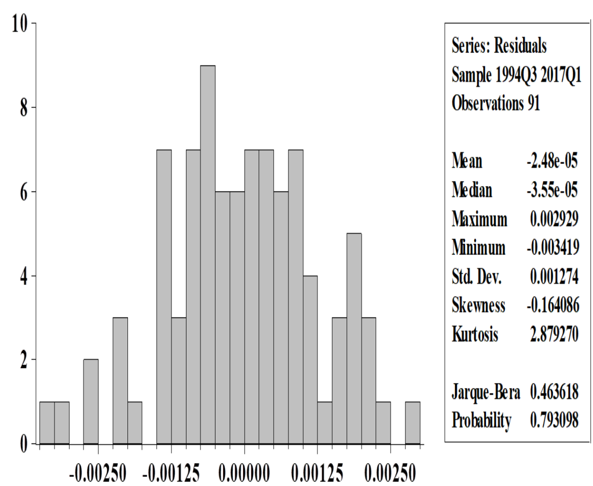


Figure A3
Residual autocorrelation, BE2

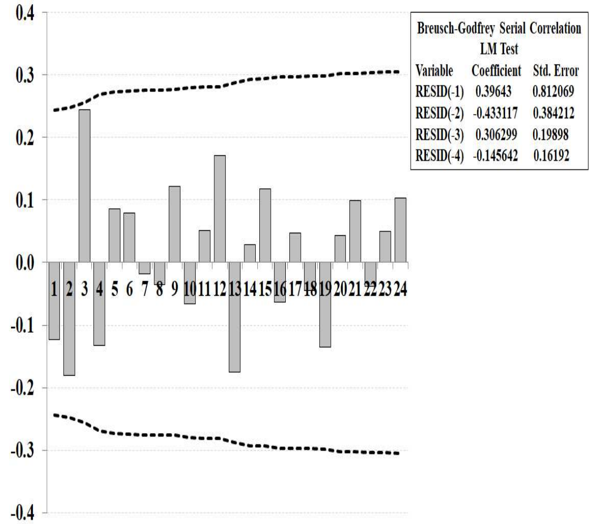


Figure A4
Residual normality test, BE2

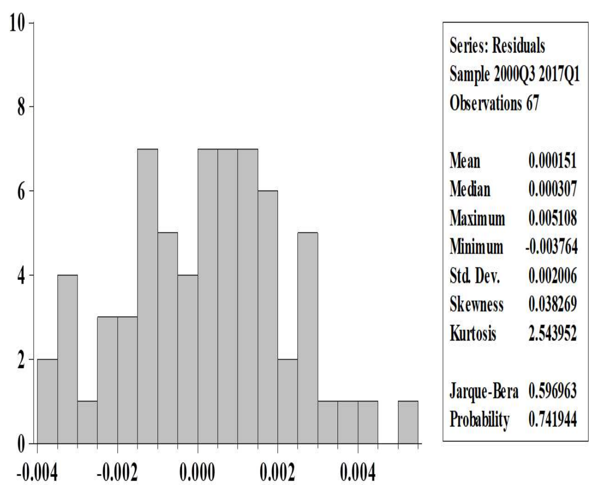


Figure A5
Residual autocorrelation, DFM

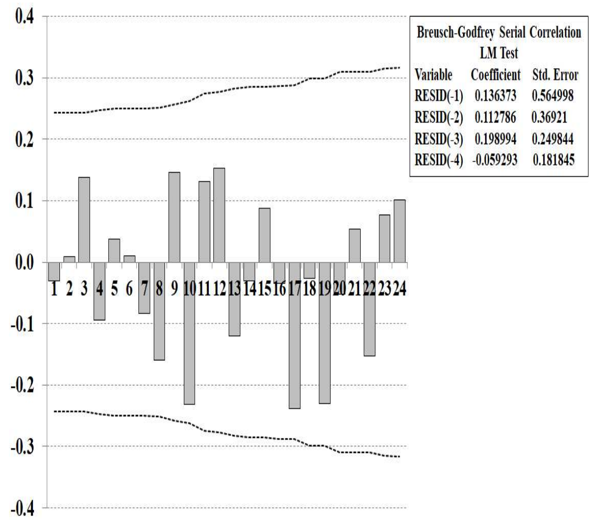


Figure A6
Residual normality test, DFM

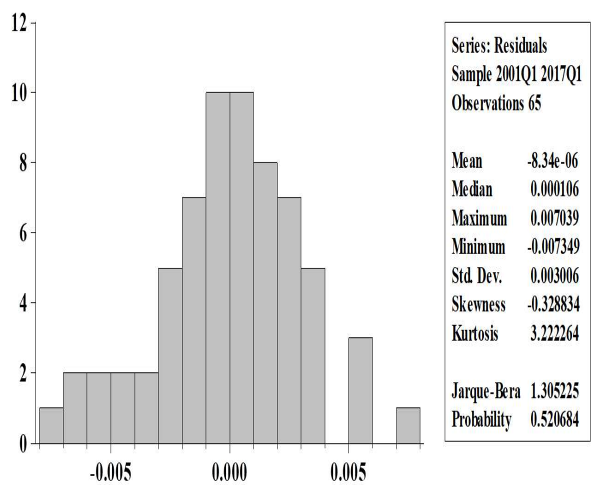


Figure A7

Residual autocorrelation, PCA1

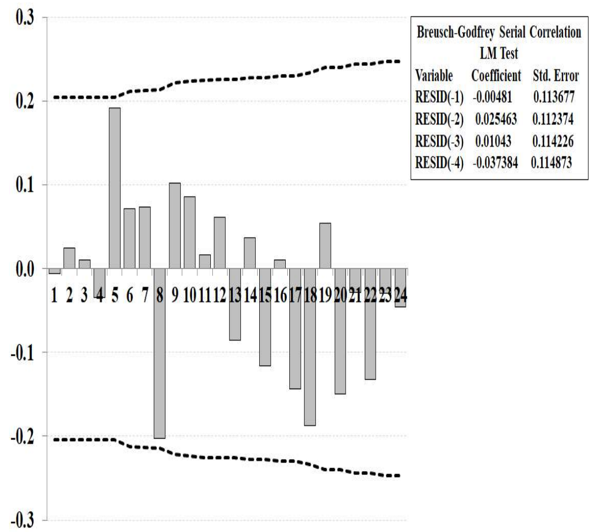


Figure A8

Residual normality test, PCA1

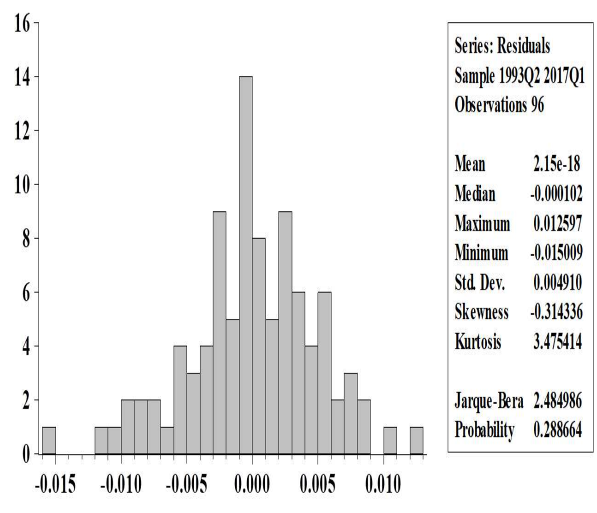


Figure A9
Residual autocorrelation, PCA2

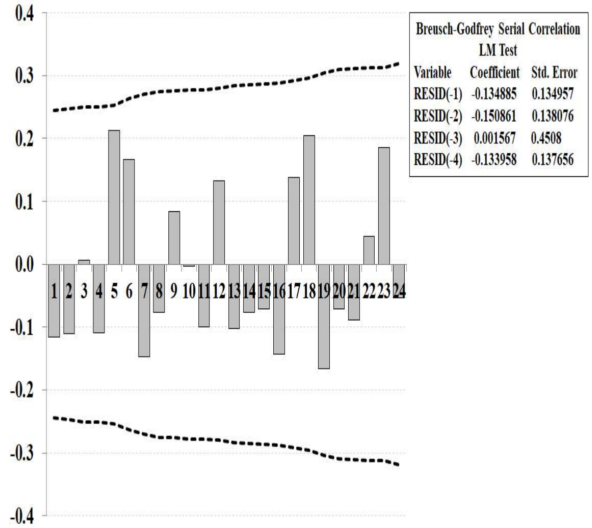
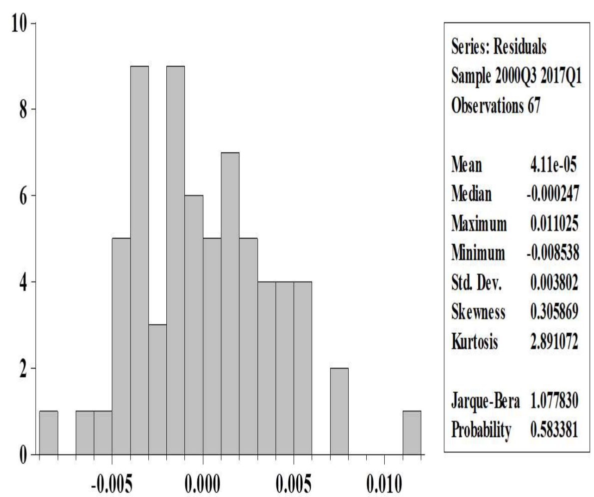


Figure A10
Residual normality test, PCA2



A4. Diebold-Mariano test

Forecast errors are defined as:

$$\varepsilon_{it} = \hat{y}_{it} - y_t, \qquad i = 1, 2$$

It is assumed that the loss function associated with the forecast i is a function of the forecast error, ε_{it}, and it is denoted by $g(\varepsilon_{it})$. The function $g(\bullet)$ is a loss function, such that: it takes the value of zero when no mistake is made, it is never negative, and it is increasing as errors become larger in magnitude. Typically, $g(\varepsilon_{it})$ is the squared-error loss or the absolute error loss of ε_{it}.

$$g(\varepsilon_{it}) = {\varepsilon_{it}}^2$$

$$g(\varepsilon_{it}) = |\varepsilon_{it}|$$

A problem with these loss functions is that they are symmetric. In fact, in some cases, the symmetry between forecast errors, positive and negative, may be inappropriate. The loss difference between two forecasts is defined as:

$$d_t = g(\varepsilon_{1t}) - g(\varepsilon_{2t})$$

The two forecasts have equal accuracy if and only if the loss difference has an expectation of zero for all t. Next I apply this to test the null hypothesis,

$$H_0 : E(d_t) = 0 \qquad \forall t$$

against the alternative hypothesis,

$$H_1 : \quad E(d_t) \neq 0$$

The null hypothesis is that the two forecasts have the same precision. The alternative hypothesis is that the two forecasts have different levels of precision. Consider the statistic:

$$\sqrt{T}\left(\bar{d}-\mu\right)$$

where $\bar{d}=\sum_{t=1}^{T}d_t$ is the sample mean of the difference between loss functions, $\mu=E\left(d_t\right)$ is the population mean of the difference between loss functions, $f_d\left(0\right)=\frac{1}{2\pi}\left(\sum_{k=-\infty}^{\infty}\gamma_d\left(k\right)\right)$ is the spectral density of the loss difference at the frequency 0, and $\gamma_d\left(k\right)$ is the autocovariance of the loss difference to the lag k.

It is possible to show that if the series generated by the difference between loss functions $\{d_t; t=1,\ldots,T\}$ is stationary and short-memory, then:

$$\sqrt{T}\left(\bar{d}-\mu\right)\xrightarrow{d}N\left(0,2\pi f_d\left(0\right)\right)$$

In the following I assume that the series generated by the loss difference are stationary and short-memory. Suppose the forecasts are $h(>1)-periods$ forward. To test the null hypothesis that the two forecasts have the same accuracy, Diebold and Mariano (1995) use the following statistic:

$$DM=\frac{\bar{d}}{\sqrt{\frac{2\pi\widehat{f_d}(0)}{T}}}$$

Where $\widehat{f_d}\left(0\right)$ is a consistent estimator of $f_d\left(0\right)$ defined by

$$\widehat{f_d}\left(0\right)=\frac{1}{2\pi}\sum_{k=-(t-1)}^{T-1}I\left(\frac{k}{h-1}\right)\hat{\gamma}_d\left(k\right)$$

where $\hat{\gamma}_d\left(k\right)=\frac{1}{T}\sum_{t=|k|+1}^{T}\left(d_t-\bar{d}\right)\left(d_{t-|k|}-\bar{d}\right)$ and $I\left(\frac{k}{h-1}\right)=$ $\begin{cases}1 & for\ \left|\frac{k}{h-1}\right|\leq 1\\ 0 & otherwise\end{cases}$. Under the null hypothesis, the DM test statistic is asymptotically $N(0)$ distributed. The null hypothesis of no difference will be rejected if the computed DM statistic falls outside the range of $-z_{\alpha/2}$ and $z_{\alpha/2}$.

Note that the normal distribution can be a poor approximation of the DM test's finite-sample null distribution for small samples. Thus, depending on the degree of serial correlation among the forecast errors and the sample size T, the results of the DM test can cause the null hypothesis to be rejected too often. Harvey, Leybourne, and Newbold (97) suggest that improved small-sample properties can be obtained by: making a bias correction to the DM test statistic and comparing the corrected statistic with a Student-t distribution with $(T-1)$ degrees of freedom, rather than comparing it to the standard normal statistic.

The corrected DM test statistic is obtained as:

$$DM^{*} = \left[\frac{T+1-2h+h\,(h-1)}{T}\right]^{1/2} DM$$

where, T is the size of the sample and h is the forecast horizon, which in our case it is equal to one, $h=1$, so our HLN-modified DM test is:

$$DM^{*} = \sqrt{\frac{T-1}{T}} DM$$

A5. Robustness check for DM test

As a robustness check I performed the HLN-modified DM test with a different loss criterion than the MSE, so I used the mean absolute error (MAE) as a loss criterion instead. I proceed with a Bartlett kernel to compute the long-term variance (table A3). I also analyze a different period in order to determine if the BE average is consistently better than the rest of the models (table A4).

The first robustness check is consistent with the previous findings, and the DFM is shown to be even more accurate than the AR, PCA1 and PCA2 models, but it is not more accurate than the BE-average model. The BE average provides more precise forecasts with respect to the rest of the models over the period analyzed, and the difference in accuracy is statistically significant (table A3), except when it is compared with the BE1 model, and with the BE-median model, where I found inconclusive results.

In addition, I performed the HLN-modified DM test for a different period of time. Since the original test includes the 2008-2009 financial crisis, in this additional test I omit the aforementioned period, in such a way that the analysis is carried out from 2011-I to 2016-II. In this test I use the MSE as a loss criterion and the uniform kernel distribution to compute the long-term variance.

The HLN-DM tests made for the period 2011-I to 2016-II confirms that the BE's forecasts are more accurate than those of the rest of the models. However, the tests also highlight the fact that the DFM performed better than the PCA and the AR models. Furthermore, note that for this period, the results of the BE1 model are indeed better than those of the rest of the models, except for the "mean BE", for which results were not conclusive result. In summary, the "mean BE" model provides the most accurate forecast of Mexican GDP, even in the most recent period of time for which I conducted the corresponding analysis (table A4).

Table A3

HLN-modified Diebold-Mariano test (using loss criterion MAE)
forecasts from 2009-I to 2016-II

Models	*AR*	*PCA1*	*PCA2*	*DFM*	*BE1*	*BE2*	*Mean*	*Median*	*Mean BE*
AR	**.550**								
PCA1	PCA1	**.546**							
PCA2	AR	PCA1	**.555**						
DFM	DFM*	DFM*	DFM***	**.243**					
BE1	BE1***	BE1**	BE1***	BE1	**.168**				
BE2	BE2**	BE2**	BE2***	BE2*	BE1	**.174**			
Mean (a)	Mean**	Mean**	Mean***	DFM	BE1*	BE2	**.248**		
Median (a)	Median***	Median**	Median***	Median**	Median	Median	Median**	**.165**	
Mean (BE)	MeanBE***	MeanBE**	MeanBE***	MeanBE**	MeanBE	MeanBE**	MeanBE***	MeanBE	**.136**

Notes: (a) = all models; p-value for the significance of differences in MSE between compared models ***p<0.01, **p<0.05, *p<0.1 The sample includes forecasts from 2009-I to 2016-II. The mean absolute error (MAE) is used as loss criterion and a Bartlett kernel is used to compute the long-term variance. The main diagonal shows the MAE of each model.

Table A4

HLN-modified Diebold-Mariano test (loss criterion MSE)
forecasts from 2011-I to 2016-II

Models	*AR*	*PCA1*	*PCA2*	*DFM*	*BE1*	*BE2*	*Mean*	*Median*	*Mean BE*
AR	**.293**								
PCA1	PCA1***	**.234**							
PCA2	AR	PCA1**	**.378**						
MDF	DFM*	DFM*	DFM**	**.143**					
BE1	BE1***	BE1***	BE1***	BE1	**.026**				
BE2	BE2***	BE2***	BE2***	BE2	BE1	**.048**			
Mean (a)	Mean***	Mean***	Mean***	Mean	BE1	BE2**	**.82**		
Median (a)	Median***	Median***	Median***	Median*	BE1	BE2	Median***	**.059**	
Mean (BE)	MeanBE***	MeanBE***	MeanBE***	MeanBE*	MeanBE	MeanBE***	MeanBE**	MeanBE*	**.020**

Notes: (a) = all models; p-value for the significance of differences in MSE between compared models ***p<0.01, **p<0.05, *p<0.1 The sample includes forecasts from 2011-I to 2016-II. The mean squared error (MAE) is used as loss criterion and the uniform kernel distribution is used to compute the long-term variance. The main diagonal shows the MSE of each model.