Box 1
Evaluation of the Predictive Capacity of Expected Inflation Measures

César Anzola-Bravo
Anderson Grajales-Olarte
Alexander Guarín-López
Julián Camilo Mateus-Gamboa
Jonathan Alexander Muñoz-Martínez
Carlos Andrés Quicazán-Moreno
Juan Sebastián Rojas-Moreno
Cristhian Hernando Ruiz-Cardozo*

Economic theory suggests that inflation expectations are a key driver of decision-making among economic agents. Alongside nominal interest rates, expected inflation helps determine the real interest rate, which is fundamental to decisions regarding consumption, production, and savings. Given the relationship between inflation and economic activity, expected inflation can also be an indicator of future behavior related to production and employment[1]. It serves a central role in determining prices and salaries. As businesses look to set salary increases for an upcoming year, for example, they may incorporate inflation expectations into their decision-making processes, compensating employees for an anticipated increase in the cost of living and making determinations over increases in labor costs.

Given its importance in informing such decisions, this supplement examines which of the numerous measures of expected inflation are most predictive of future behavior, and under what circumstances[2].

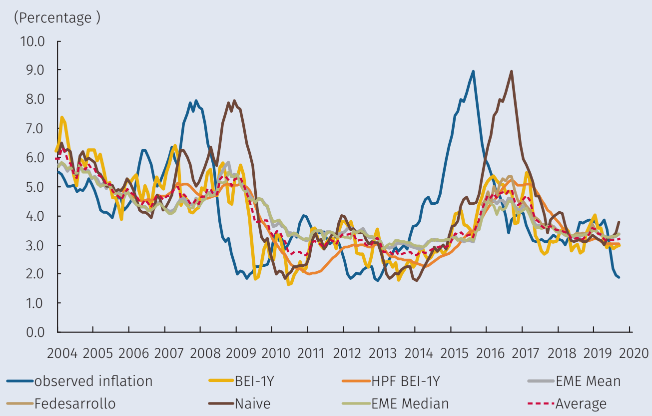
The measures used to gauge expected inflation in Colombia are varied, and aim to capture a diverse range of perspectives, time horizons (short, medium, long) and frequencies (monthly, quarterly). Although these measures generally reveal similar dynamics, their results tend to be more mixed in periods of uncertainty regarding the nature and duration of economic shocks.

Diagram B1.1 summarizes the measures analyzed in this supplement. Those included consider annual inflation expectations on three different forecast horizons (one, two, and five years) and using different general sources of information: surveys, financial market instruments, and macroeconomic models. The surveys include both the Central Bank's monthly (EME in Spanish) and quarterly (ETE in Spanish) surveys of analyst expectations, the survey of financial opinion conducted by Colombian think-tank Fedesarrollo, and a survey from Focus Economics[3].

The market instruments used include breakeven inflation (BEI)[4] and forward breakeven inflation (FBEI), derived from public debt securities, and a smoothed measure of BEI using a recursive Hodrick-Prescott filter (HPF BEI)[5].

The macroeconomic modeling measures of expected inflation are based on the Bank's central forecasting models: Patacon and 4GM. They also include averages of the measures in question and a naive expectation measure, the value of which corresponds to the most recent observable data point. Graph B1.1 compares monthly annual inflation expectations on a 12-month forecast horizon with the corresponding observed inflation.

Graph B1.1
Annual Inflation and Measures of Expected Inflation



Sources: DANE, *Banco de la República*, Fedesarrollo; calculations by the authors.

* The authors are members of the Central Bank's Department of Macroeconomic Modeling and Department of Operations and Market Analysis. The opinions contained herein are theirs alone, and do not necessarily reflect those of the Bank or its Board of Directors.

1 Inflation expectations in this sense serve as a market proxy of monetary policy credibility.

2 This analysis does not attempt to measure the relative degree of divergence in inflation expectations, which would be of value but falls outside the bounds of this supplement. This supplement is a purely retrospective exercise and its results can not necessarily be extrapolated to future behavior. Given the difficulty in anticipating supply and demand shocks (e.g. El Niño or Covid-19), measures of expected inflation show high levels of forecast error. This limited forecasting capacity in absolute terms does not diminish the importance of expected inflation measures in decision-making on behalf of diverse economic agents.

3 Expected inflation measures at 12 and 24 months from the EME begin in September 2003 and January 2015, respectively, while those for the ETE at four and eight quarters are available from 2000 and 2015. Expectations at five years from Focus Economics and at 12 months from Fedesarrollo are available beginning in February 2004 and October 2015, respectively.

4 BEI expectations are constructed based on nominal government bonds denominated in pesos and UVR.

5 Smoothed expectations aim to eliminate short-term movement derived from government bonds that are not necessarily related to changes in market expectations. Conducted with a recursive Hodrick-Prescott filter.

This supplement presents three exercises used to examine the predictive capacity of expected inflation measures:

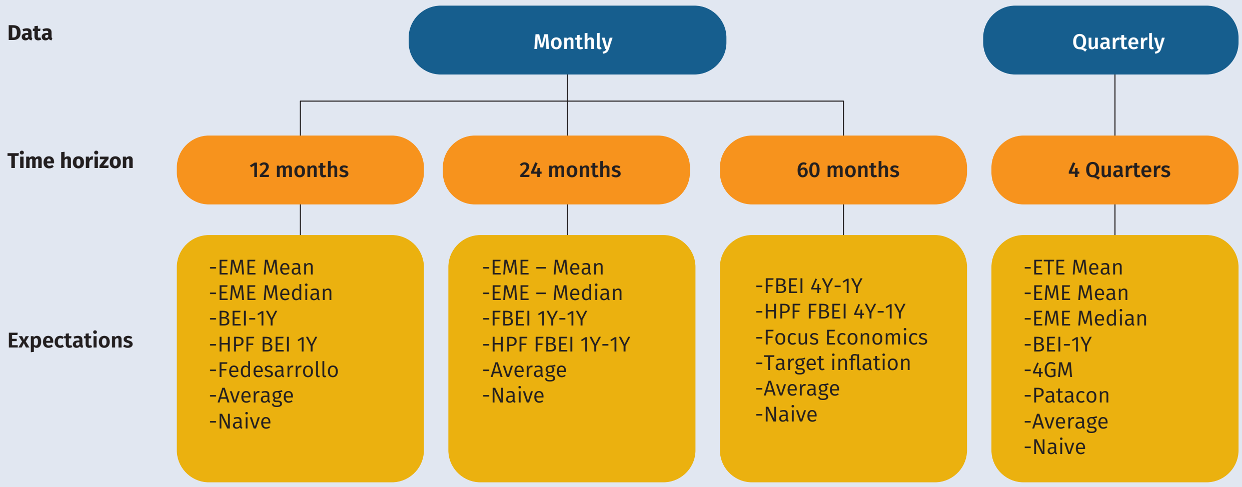
a. Traditional evaluation of statistics such as mean absolute error (MAE) and root-mean-squared error (RMSE).
b. Forecast error distribution and likelihood analysis
c. Statistical tests comparing expected values of forecast errors (see Giacomini and White, 2006; Giacomini and Rossi, 2010).

These three exercises were performed for the various measures of inflation, frequencies, and time horizons presented in Diagram B1.1 for three different sample periods: the full period[6], the last three years (2017-2020), and periods of high inflation (2008 and 2015-2016)[7]. This supplement only shows exercises related to monthly expected inflation at 12 months. A comparative analysis of all the measures considered for the three sample periods is presented in the final section of this supplement.

1. Traditional Evaluation

MAE and RMSE[8] were calculated for the various measures of expected inflation. These statistics reflect the average magnitudes of forecast errors for each measure with respect to observed inflation at the future date. Chart B1.1 reflects these statistics for expected inflation at 12 months

Diagram B1.1
Measures of Expected Inflation: Source, Frequency, and Time Horizon



Sources: DANE, *Banco de la República* and Fedesarrollo; calculations by the authors

Chart B1.1
MAE and RMSE for Expected Inflation at 12 Months

	MAE (Percentage Points)			RMSE (percentage points)		
	Full period	High-inflation periods	Last three years	Full period	High-inflation periods	Last three years
BEI-1Y	1.2	3.6	0.6	1.7	3.7	0.7
HPF BEI-1Y	1.4	3.8	0.9	1.8	3.9	1.1
EME Mean	1.3	3.9	0.5	1.8	4.1	0.7
EME Median	1.3	4.0	0.5	1.8	4.1	0.7
Fedesarrollo	-	-	0.6	-	-	0.8
Average	1.3	3.8	0.6	1.7	3.9	0.7
Naive	1.7	2.9	1.5	2.1	3.1	2.0

Sources: DANE, *Banco de la República* and Fedesarrollo; calculations by the authors

6 Available information is considered for each measure, and the most recent data is considered the start of the period, such that all measures use the same set of information and are comparable among themselves.

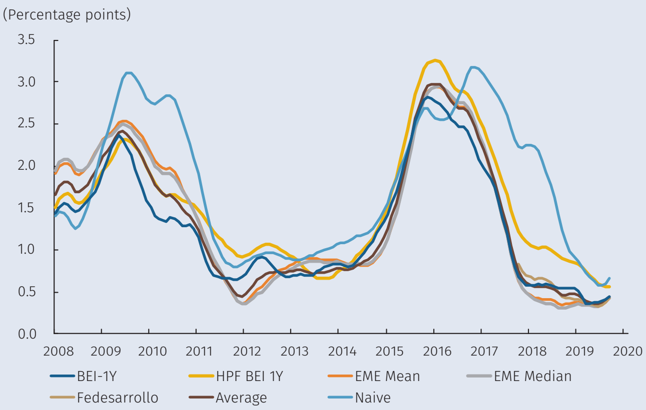
7 High-inflation periods are defined as those in which the annual observed change in CPI is above the constructed range of 1.5 times greater than its standard variation around the mean.

8 $MAE = \frac{1}{T}\sum_{t=1}^{T}|e_t^m|$ and $RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(e_t^m)^2}$, where $e_t^m = E_t[\pi_{t+m}] - \pi_{t+m}$ is the forecast error $e_t^m = E_t[\pi_{t+m}]$ and the expected value t of inflation $t+m$; π_{t+m} is the value observed at that date; t is time indexed; m time horizon, and T is the number of periods.

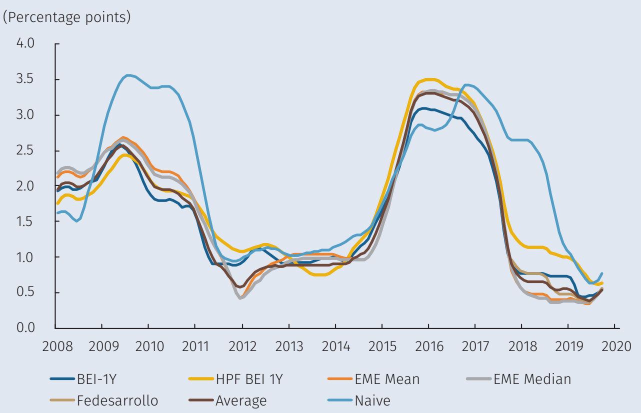
for the three sample periods considered, while Graph B1.2 shows their historical behavior using a centered rolling window of 24 months.

Graph B1.2
MAE and RMSE for Expected Inflation at 12 Months

A. MAE



B. RMSE



Sources: *Banco de la República* and Fedesarrollo; calculations by the authors.

Broadly speaking, the results of this exercise show that the predictive ability of distinct measures of expected inflation changes over time. Both the MAE and RMSE suggest a relatively high degree of prediction error from 2009-2011 and from 2015-2017. In the first case, these errors can be associated with a rapid decline in observed inflation resulting from the global financial crisis. For the second period in question, inflation rose as the result of nominal depreciation as well as from supply pressures created by an El Niño weather pattern from 2015 to 2016. By contrast, from 2012 to 2014 and from the end of 2017 to the first half of 2020, prediction errors were relatively low, and in the latter case were trending downward.

For the full sample period, prediction errors for the various measures of expected inflation oscillated between 1.2 and 1.8 percentage points (pp). For the sample period of the last three years, those figures were reduced by half, and even more in some cases. In periods of high inflation, however, the MAE and RMSE rose to values between 3.6 and 4.1 pp, the result of supply or demand shocks the nature and duration of which were difficult for economic agents to determine.

A comparison of the MAE and RMSE results for the different measures of expected inflation was also conducted. The results suggested that one-year BEI (BEI-1Y) performed better relative to the other measures for the sample period and in particular before 2012 and from 2016 to the beginning of 2018. These results are illustrated in Graph B1.2.

2. Distribution of the Probability of Forecast Errors and Log-Score

This exercise directly compared the probability distribution of forecast errors reflected in the various measures of expected inflation. Distributions were estimated using parametric methods for each measure on each of the study's sample periods, using the same centered rolling window of 24 months defined above.

A log-score indicator was calculated for each of these probability distributions, defined as the natural logarithm of the relative probability of observed forecast errors equal to zero. The higher the value of this indicator, the better the predictive capacity of the measure being considered (Geweke and Amisano, 2010). Graph B1.3, Panel A shows the log-score over time for expected inflation at 12 months, while Panel B illustrates the probability distribution of forecasting errors for the entire period.

As with the results described in the previous section, the predictive capacity of the various measures of expected inflation depended on the period of analysis. According to the log-score, BEI-1Y showed the best predictive capacity before 2012 and was among the best-performing measures after 2017. The mean and median of analysts' EME responses were the most predictive after 2015. Similarly, the distribution of forecast errors indicates that BEI-1Y was the most predictive over the full study period, followed by statistics from the EME.

3. Statistical Tests

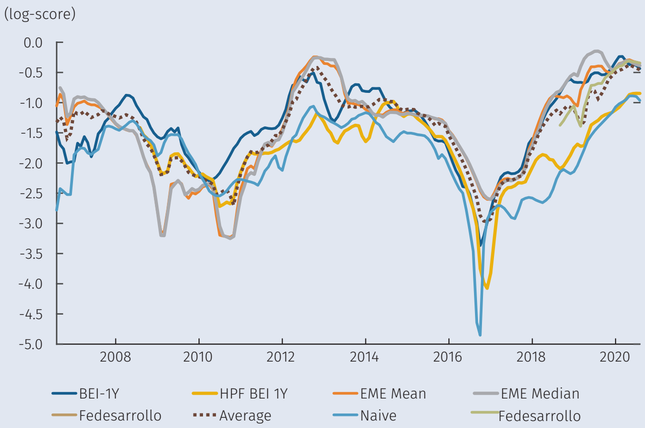
Finally, the predictive capacity of the various measures of expected inflation was evaluated using Giacomini and White (GW, 2006)[9] and Giacomini and Rossi (GR, 2010)[10] tests. These tests are used to evaluate the statistical significance of the expected difference between forecast

9 The unconditional (conditional) GW test evaluates $H_0{=}E[\Delta\hat{L}^{S1,S2}_{t+h}] = 0$ $(H_0{=}E[\Delta\hat{L}^{S1,S2}_{t+h}]|G_t]{=}0)$ equivalent predictive capacity among measures of expected inflation s_1 and s_2 (conditioned on the set of information G_t).

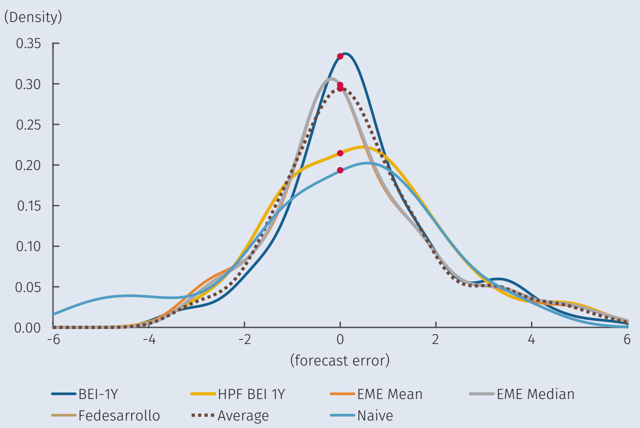
10 The *GR* fluctuation test evaluates $H_0{=}E[\Delta\hat{L}^{S1,S2}_{t+h}] = 0$, equivalent predictive capacity among measures of expected inflation s_1 and s_2 on a rolling time horizon with set magnitude v.

Graph B1.3
Expected Inflation at 12 Months: distribution of Probability in Forecast Errors and Log-Score

A. Log-score



B. Probability distribution of forecast errors (full period)



Note: in Panel B, positive values on the x-axis indicate that inflation expectations underestimated observed annual inflation
Source: calculations by the authors.

errors of two different measures of expected inflation. The conditional GW test also examines whether the predictive capacity of the measures can be differentiated given prior conditions. The GR test captures the variation over time in the relative performance of the forecast.

The results of the unconditional and conditional GW tests at 12 months are reflected in Panel A and Panel B, respectively, of Chart B1.2. The values in the first line correspond to the p-values of each test. The values in parenthesis in the second line represent the ratio between the RMSE of the measure of expected inflation on the upper part of the box with respect to the measure on the left. The blue (orange) shading indicates that the test rejects an equivalent predictive capacity among the two measures at a 10% level of significance, and that the measure on the left has a lower (higher) RMSE than the measure above.

Graph B1.4 illustrates the results of a GR fluctuation test for the same measures of expected inflation at 12 months considered throughout this supplement. This is an unconditional test of predictive equivalency among two measures for each of the moving averages defined above. The result of the test for each measure is presented relative to a benchmark rate, in this case the BEI-1Y. Positive (negative) values in the test correspond to measures that were less (more) predictive than the benchmark. The dotted lines denote critical values at a 5% level of significance.

Chart B1.2 shows statistically significant differences in the predictive capacity of various expected inflation measures and suggests that BEI-1Y was the most predictive on the 12-month horizon. The results shaded in blue on both the conditional and unconditional GW tests show a greater predictive capacity in the BEI-1Y compared to the other measures analyzed. The results of the GR test displayed in Graph B1.4 do not reject the hypothesis of equivalent predictive ability among the different time frames considered. Nevertheless, the BEI-1Y showed lower expected forecast errors between 2010 and 2011, as well as between mid-2015 and the beginning of 2017, and is in line with the EME expectations.

4. Results and Conclusions

Chart B1.3 shows the best-performing measures of expected inflation in each of the four evaluation exercises for each frequency and time horizon and highlights the measure that had the best results on average.

BEI-1Y showed the best predictive capacity at 12 months for the full sample period and when using monthly data, while for 24 and 60 months the average of the various measures of expected inflation would have performed best. The 4GM monetary policy model performed best for the quarterly frequency.

For the sample period of the last three years, the median of analyst responses in the EME was the most predictive at 12 and 24 months. For periods of high inflation at 12 months and four quarters, again the BEI-1Y and 4GM were most predictive.

In conclusion, the results of the evaluation suggest that 1) measures of expected inflation are imprecise, showing high levels of forecast error in absolute terms; that 2) the measures' predictive capacity depends significantly on the time horizon considered (12, 24, or 60 months); and that 3) their predictive capacity changes over time, depending on the existence, nature, and duration of economic shocks.

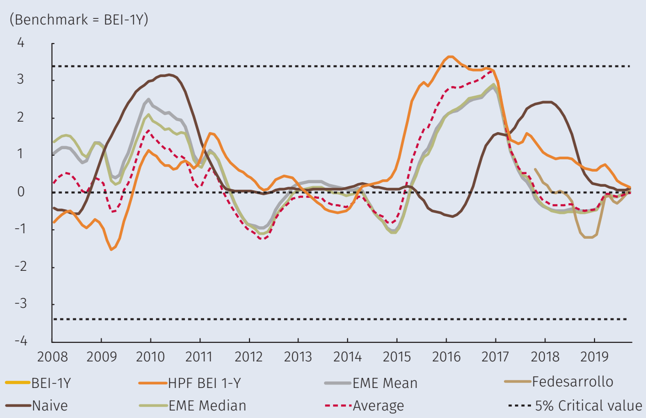
Chart B1.2
Giacomini & White (2006) Statistical Tests: expected Inflation at 12 Months

A. Unconditional test	HPF BEI-1Y	EME Mean	EME Median	Average	Naive
BEI-1Y	0.03 (1.18)	0.05 (1.16)	0.07 (1.15)	0.14 (1.07)	0.03 (1.63)
HPF BEI-1Y		0.41 (0.98)	0.35 (0.97)	0.03 (0.90)	0.11 (1.37)
EME Mean			0.21 (0.98)	0.01 (0.91)	0.10 (1.39)
EME Median				0.03 (0.92)	0.10 (1.41)
Average					0.06 (1.52)

B. Conditional test	HPF BEI-1Y	EME Mean	EME Median	Average	Naive
BEI-1Y	0.06 (1.18)	0.00 (1.16)	0.00 (1.15)	0.01 (1.07)	0.08 (1.63)
HPF BEI-1Y		0.00 (0.98)	0.00 (0.97)	0.00 (0.90)	0.04 (1.37)
EME Mean			0.14 (0.98)	0.00 (0.91)	0.04 (1.39)
EME Median				0.00 (0.92)	0.03 (1.41)
Average					0.06 (1.52)

Note: In Chart B1.2 Panels A and B, the values on the first line correspond to the p-values for each test, the numbers in parenthesis are ratios between RMSE for the expected inflation measure on the upper part of the table compared to the measure on the left. Shading in blue (orange) indicates that the test rejects equivalency in the predictive capacity between measures at a level of significance α=10%, and that the measure on the left has a smaller (larger) RMSE than the measure above.
Source: calculations by the authors.

Graph B1.4
Giacomini & Rossi (2010) Statistical Tests: Expected Inflation at 12 Months



Note: each line evaluates the relative performance of a measure of expected inflation against the benchmark. The dotted line denotes a critical value in the Giacomini & Rossi (2010) test at a significance level of α=5%. Positive (negative) values in the test corres-pond to a measure of expected inflation performing below (above) the benchmark.
Source: calculations by the authors

Chart B1.3
Measures of Expected Inflation: Summary of Comparative Analysis

A. Full period				
Frequency	Monthly			Quarterly
Time horizon	12 months	24 months	60 months	4 quarters
Traditional Statistics	BEI-1Y	Average	Average	4GM
Giacomini & White (2006)	BEI-1Y	Average	Average	4GM
Giacomini & Rossi (2010)	BEI-1Y	Average	HPF FBEI 4Y-1Y	Patacon
Log-score	BEI-1Y	Average	Target inflation	4GM
Summary	BEI-1Y	Average	Average	4GM
B. Last three years				
Frequency	Monthly			Quarterly
Time horizon	12 months	24 months	60 months	4 quarters
Traditional Statistics	EME (Median)	EME (Median)	HPF FBEI 4Y-1Y	Average
Giacomini & White (2006)	Inconclusive	EME (Median)	Target inflation	Inconclusive
Giacomini & Rossi (2010)	Fedesarrollo	EME (Median)	HPF FBEI 4Y-1Y	BEI-1Y
Log-score	EME (Median)	EME (Median)	HPF FBEI 4Y-1Y	Average
Summary	EME (Median)	EME (Median)	HPF FBEI 4Y-1Y	Average
C. High-inflation periods				
Frequency	Monthly			Quarterly
Time horizon	12 months	24 months	60 months	4 quarters
Traditional Statistics	BEI-1Y	Average	HPF FBEI 4Y-1Y	4GM
Giacomini & Rossi (2010)	BEI-1	FBEI 1Y-1Y	HPF FBEI 4Y-1Y	BEI-1Y
Log-score	BEI-1Y	FBEI 1Y-1Y	HPF FBEI 4Y-1Y	4GM
Summary	BEI-1Y	FBEI 1Y-1Y	HPF FBEI 4Y-1Y	4GM

Source: Calculations by the authors.

References

Geweke, J.; Amisano, G. (2010). "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns", *International Journal of Forecasting*, pp. 216-230.

Giacomini, R.; Rossi, B. (2010). "Forecast Comparisons in Unstable Environments", *Journal of Applied Econometrics*, vol. 25, no. 4, pp. 595-620.

Giacomini, R.; White, H. (2006). "Test of Conditional Predictive Ability", *Econométrica*, vol. 74, no. 6, pp. 1545-1578.

González, A.; Guarín, A.; Rodríguez, D.; Vargas, H. (2020). "4GM: A New Model for Monetary Policy in Colombia", Borradores de Economía, No. 1106, *Banco de la República*.

González, A.; Mahadeva, L.; Prada, J.; Rodriguez, D. (2011). "Policy Analysis Tool Applied to Colombian Needs: Patacon model description", *Ensayos sobre Política Económica*, vol. 29, no. 66, pp. 222-245.