Forecasting Macroeconomic Variables in Colombia: A Methodological Comparison Approach

Santiago Díaz Espitia[1]

Pontificia Universidad Javeriana

Julio 2023

[1] The complete dataset in Excel and all the code used can be found in My Github Page, some ChatGPT suggestions were incorporated in the proofreading of the document and the making of this presentation.

"Because of the things we don't know we don't know, the future is largely unpredictable."

Maxine Singer (1997) 'Thoughts of a nonmillennarian', *Bulletin of the American Academy of Arts and Sciences*, 51,2,p.39

Literature Review

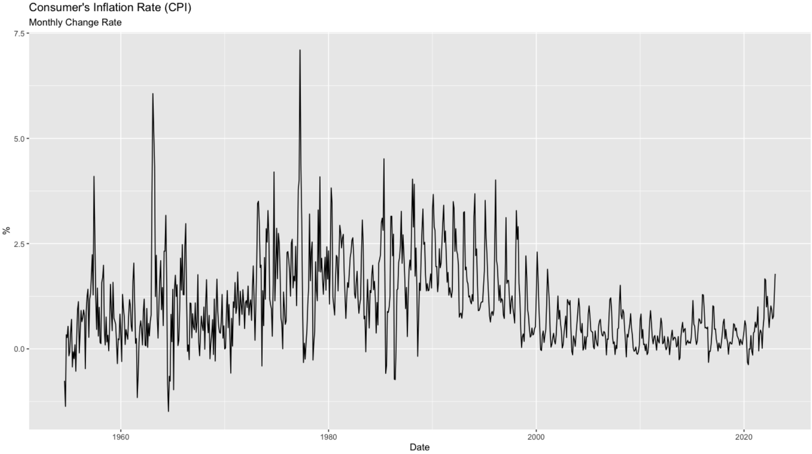
- Presenting the taxonomy of forecast errors, Clements and Hendry (2001) show the pernicious effects of unanticipated location shifts (changes in the unconditional mean of the DGP) (Castle et al., 2017).
- Model uncertainty and parameter instability are crucial aspects in macroeconomic forecasting (Rapach and Zhou, 2013).
- Following Castle et al. (2019), augmenting the information set does not imply better forecasts.
- In Colombia, Carmona Restrepo (2022) uses regularized neural networks to forecast inflation and Cárdenas-Cárdenas et al. (2023) used a LSTM neural network to forecast inflation; both finding results in favour of ML models.
- In Faust and Wright (2013), it is shown the importance of taking into account locations shifts when forecasting inflation in the United States and model selection algorithms are used for forecasting UK's inflation in Castle (2006).
- Since the findings of Meese and Rogoff (1983), Petropoulos et al. (2022) points to the "exchange rate disconnect puzzle".
- Petropoulos et al. (2022) give a state-of-the-art summary analyzing traditional statistical approaches, "Big Data" methodologies, and forecast combinations, among others.

Data

- The two target variables are the monthly inflation rate (CPI) and the depreciation rate (TRM). Even though approximate publication delays are taken into account, revisions are not (pseudo real-time evaluation).
- Variables were divided into 8 blocks: 1) Monetary and credit aggregates, 2) Activity indicators, 3) Prices, 4) Interest rates, 5) External, 6) Surveys, 7) Global, and 8) Google trends data. None of the series was seasonally adjusted.
- For I(1)/I(0) representations variables were specified in a "theory-consistent" way[2] and for I(0) models the variables were differenced until the ADF test was rejected.

[2] For example, the Phillips Curve denotes a relationship between inflation and unemployment rates and not between the unemployment rate and the CPI level.

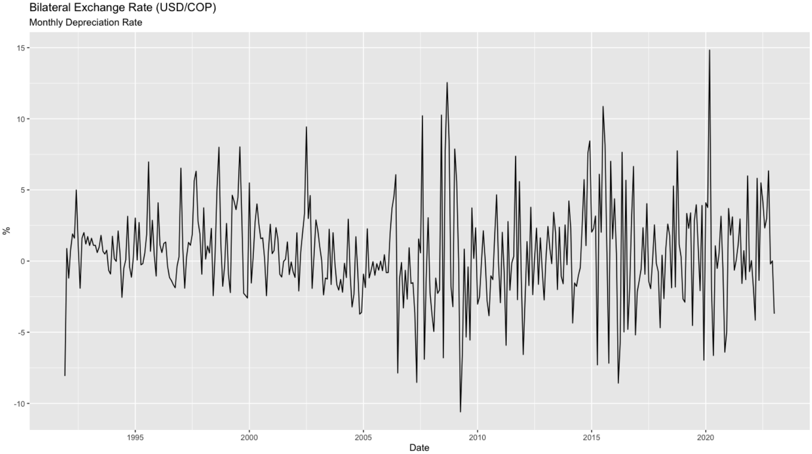
Figure: Inflation Rate



Source: BANREP

Descriptive Statistics

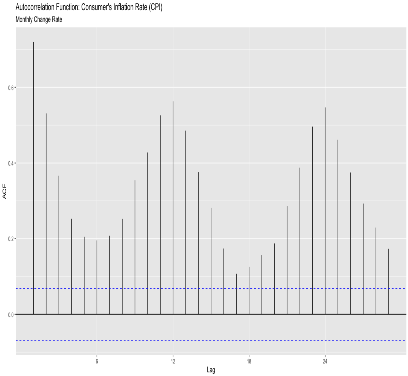
Figure: Depreciation Rate



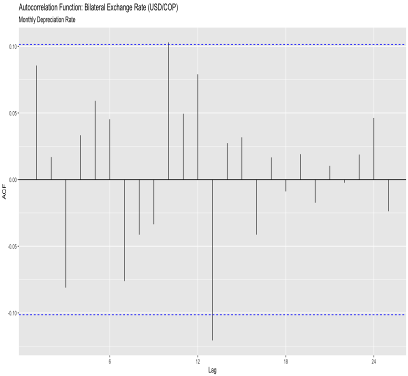
Source: BANREP

Figure: Autocorrelation Functions

(a) Autocorrelation Function (Inflation Rate)



(b) Autocorrelation Function (Depreciation Rate)



Source: Author's calculations

- For inflation the information set limited by t-1 may matter, but for the depreciation rate......

- Approaches are compared by their RMSFE and the evaluated horizons were: 1) One-month, 2) One-year and, 3) Two-years ahead.
  1. **Univariate Time Series Models:** ARIMA, RW, ETS, MEAN and DDD.
  2. **Multivariate Time Series Models:** VAR, VARdiff, GETSISAT, GETSIC, DFM.
  3. **Machine Learning Models:** LASSO, RIDGE, RF, SVM, ENet, KNN[3]. As some allow for high-dimensionality, and/or non-linearities and/or cross-validation (Goulet Coulombe et al., 2022).
  4. **Forecast Combinations:** SA, OLS,TA, among others (To avoid "poisonous" methods in the combinations (Castle and Hendry, 2022) , only the 4 best individual models were used in combinations).
- For both target variables the main test period used for the pseudo out-of-sample evaluation is 2019(1)-2023(1) (49 periods).
- For the Machine Learning models, hyperparameters were tuned using Time Series Cross-Validation (TSCV).
- For one-month ahead forecasts the "Encuesta Mensual de Expectativas Económicas" EMEE from Colombia's Central Bank was used as a reference regarding survey-based forecasts[4].
- Each forecast was done *as if* it was 12^{th} of every month to favour comparability to EMEE for one-month ahead forecasts.

[3] Default settings of the *caret* package (Kuhn, 2008) were used, it is important to denote that pre-processing input variables was not done if it was not the default option, but an additional experiment was made.

[4] Although the information set was not completely comparable, it allows to assess the plausible dominance of survey expectations when forecasting (Faust and Wright, 2013) with an information set of lagged variables.

Figure: Forecasting Results by Horizon (Inflation)

Monthly Inflation Rate (CPI): Ratios to RMSFE of a RW Model						
	One-month	RMSFE Ratio	One-year	RMSFE Ratio	Two-years	RMSFE Ratio
Best Model	ETS	0.7813	GETSISAT*	0.8738	RF	0.8599
Best Combination	MED	0.7693	EIG1	0.8504	MED	0.8518

* Variables were consistent with a non-unit root representation.
Shaded values are significant at 5% with respect to the RW model using the Diebold-Mariano test.
For one-month ahead forecasts, the RMSFE ratio of the EMEE was 0.7274.

Source: Author's calculations

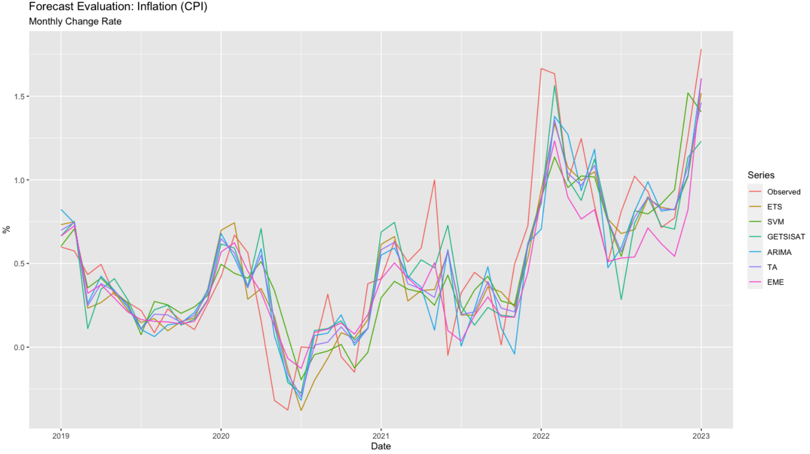
- Without predictor standardization, KNN was one of the worst performing but with it, the model became the best for inflation at longer horizons, for one-year ahead the ratio of the RMSFE was 0.8707 and for two-years ahead was 0.7278 (for a non-unit root representation).
- Actually, only using inflation's autoregressive lags and seasonal dummies lead to a RMSFE's ratio of 0.8169 for one-month ahead forecasts and for two-years ahead this model lead to an RMSFE's ratio of 0.7672[5].

[5] This was an additional exercise done for this presentation with some ChatGPT suggestions.

Forecast Evaluation: Inflation Rate

Figure: Forecasting Results (Inflation)

Figure: One-month Ahead

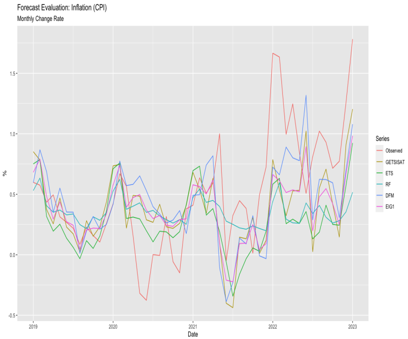


Source: Author's calculations

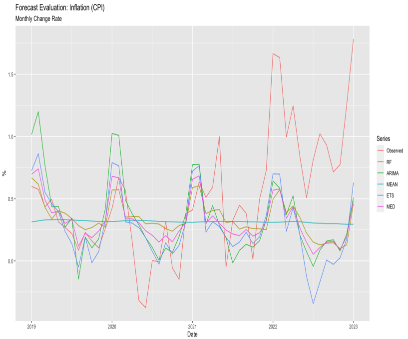
Forecast Evaluation: Inflation Rate

Figure: Forecasting Results by Horizon (One-year and Two-years ahead)

(a) One-year Ahead



(b) Two-years Ahead



Source: Author's calculations

Forecast Evaluation: Depreciation Rate

Figure: Forecasting Results by Horizon (Depreciation Rate)

Monthly Depreciation Rate (USD/COP): Ratios to RMSFE of a RW						
	One-month	RMSFE Ratio	One-year	RMSFE Ratio	Two-years	RMSFE Ratio
Best Model	LASSO	0.7022	LASSO	0.6869	SVM*	0.6910
Best Combination	CLS	0.7037	EIG1	0.6879	EIG1	0.6858

* Variables were consistent with a non-unit root representation.

Shaded values are significant at 5% with respect to the RW model using the Diebold-Mariano test.

For one-month ahead forecasts, the RMSFE ratio of the EMEE was 0.5708.

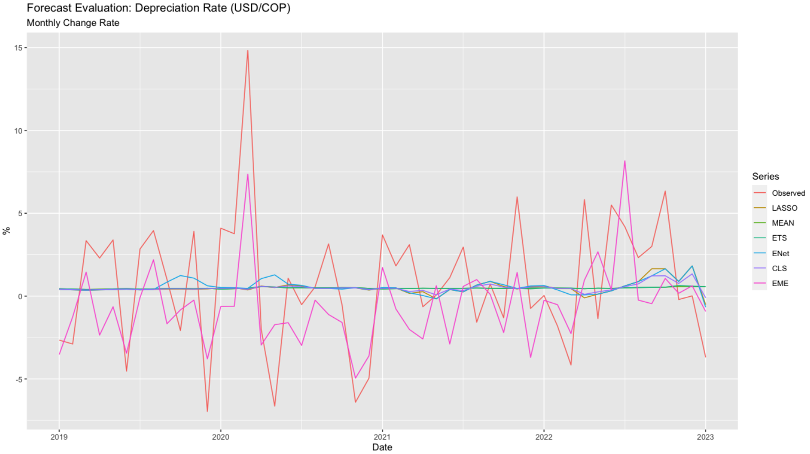
Source: Author's calculations

- For the depreciation rate, the improvement after scaling the "non-default scale" models was not appreciable when compared to the one seen for inflation. Now, although multivariate models seem to improve the RW, when compared to the MEAN forecasts, the "information benefits" are not important [6].

[6] The relative RMSFEs of the MEAN forecast for the one-month, one-year and two-years ahead horizons are: 0.708, 0.692, and 0.713, respectively.

Figure: Forecasting Results by Horizon (Depreciation Rate)

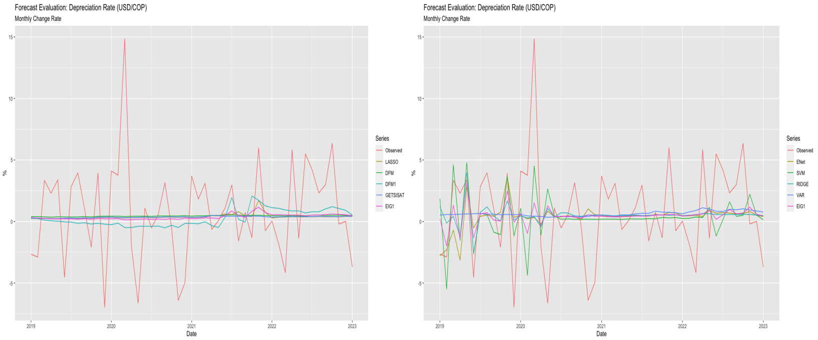
Figure: One-month Ahead



Source: Author's calculations

Figure: Forecasting Results by Horizon (Depreciation Rate)

(a) One-year Ahead (b) Two-years Ahead



Source: Author's calculations

Figure: Results from January 2017 to January 2020

(a) Inflation (CPI)

Monthly Inflation Rate (CPI)			
	Forecast Horizon		
Models	One-month	One-year	Two-years
ARIMA	0.77	1.15	0.98
KNN*	0.72	1.25	2.14
RW	0.214	0.297	0.391

* Variables were consistent with a non-unit root representation.
Shaded values are significant at 5% with respect to the RW model using the Diebold-Mariano test.
For one-month ahead forecasts, the RMSFE ratio of the EMEE was 0.4557.

(b) Depreciation Rate (USD/COP)

Monthly Depreciation Rate (USD/COP)			
	Forecast Horizon		
Models	One-month	One-year	Two-years
MEAN	0.63	0.67	0.56
LASSO	0.63	0.85	0.57
RW	4.886	4.610	5.481

* Variables were consistent with a non-unit root representation.
Shaded values are significant at 5% with respect to the RW model using the Diebold-Mariano test.
For one-month ahead forecasts, the RMSFE ratio of the EMEE was 0.5448.

Source: Author's calculations

- At last, using daily information for forecasting the depreciation rate lead to an economically significant improvement in predictive accuracy (RMSFE of 3.1369 versus 3.1787 for the EMEE for 2019-2023), suggesting that higher frequency data could be of key importance[7].

[7] For the writing of the code used in this exercise, some suggestions from ChatGPT and other similar AI tools were incorporated.

Limitations and Future Work

- Other transformations could be analyzed (e.g., annual instead of monthly rates).
- Only lagged information was employed.
- Models are sensitive to hyperparametrization, estimation and evaluation periods.
- Other models such as Neural Networks, as at least in-sample and with high non-linearities may prove fruitful (Cárdenas-Cárdenas et al., 2023).
- The use of alternative datasets and higher-frequency information (e.g., online prices as in Cavallo and Rigobon (2016)) and real-time datasets are promising for future research. Macias et al. (2023) used online prices for nowcasting inflation.
- Other loss functions that inform decision-making (Blaskowitz and Herwartz, 2011). As costs could be asymmetrical (Montenegro, 2010) or the emphasis could be on directional accuracy (Duncan and Martínez-García, 2019).

Main Takeaways: To Forecasters

- In this case parsimony pays-off! Consistent with: Nelson (1972), Atkeson et al. (2001), Duncan and Martínez-García (2019), and Leal et al. (2020).
- Survey-based (consensus) forecasts are at the "frontier of our forecasting ability" (Faust and Wright, 2013).
- Differentiation may help, but the key aspect is whether the model captures the wide-sense non-stationarities in a variable selection context (Hendry and Doornik, 2014).
- Combinations may help, but here they were not a game changer.
- Try different initial transformations.
- Higher frequency information could be crucial for competing with survey-based approaches, as lagged additional information (besides on outcome's history and "special features" such as seasonality, as pointed in Castle et al. (2022)) may be superfluous.
- Focus on: 1) Extending simple univariate forecast to higher-frequency information (Martinez et al., 2022), 2) Incorporating relevant events (Faust and Wright, 2013) at the forecast origin and 3) Nowcasting (Bańbura et al., 2013).

Atkeson, A., Ohanian, L. E., et al. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis quarterly review*, 25(1):2–11.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Now-casting and the real-time data flow. In *Handbook of economic forecasting*, volume 2, pages 195–237. Elsevier.

Blaskowitz, O. and Herwartz, H. (2011). On economic evaluation of directional forecasts. *International journal of forecasting*, 27(4):1058–1065.

Cárdenas-Cárdenas, J.-A., Cristiano-Botia, D. J., and Martínez-Cortés, N. (2023). Colombian inflation forecast using long short-term memory approach. *Borradores de Economía; No. 1241*.

Carmona Restrepo, N. (2022). Colombian inflation forecasting by regularized neural networks. Master's thesis, Universidad Nacional de Colombia.

Castle, J., Hendry, D., and Kitov, O. (2017). *Forecasting and nowcasting macroeconomic variables: A methodological overview*. Eurostat.

Castle, J., Hendry, D. F., and Clements, M. P. (2019). *Forecasting*. Yale University Press.

Castle, J. L. (2006). Empirical modelling and model selection for forecasting inflation in a non-stationary world. Phd thesis, University of Oxford; Nuffield College.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2022). Forecasting facing economic shifts, climate change and evolving pandemics. *Econometrics*, 10(1).

Castle, J. L. and Hendry, D. F. (2022). Econometrics for modelling climate change. In *Oxford Research Encyclopedia of Economics and Finance*.

Cavallo, A. and Rigobon, R. (2016). The billion prices project: Using online prices for measurement and research. *Journal of Economic Perspectives*, 30(2):151–178.

Clements, M. P. and Hendry, D. F. (2001). *Forecasting Non-stationary Economic Time Series*. MIT Press.

Duncan, R. and Martínez-García, E. (2019). New perspectives on forecasting inflation in emerging market economies: An empirical assessment. *International Journal of Forecasting*, 35(3):1008–1031.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In *Handbook of economic forecasting*, volume 2, pages 2–56. Elsevier.

Goulet Coulombe, P., Leroux, M., Stevanovic, D., and Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5):920–964.

Hendry, D. F. and Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: automatic selection methods in econometrics*. MIT Press.

Kuhn, M. (2008). Building predictive models in r using the caret package. *Journal of statistical software*, 28:1–26.

Leal, F., Molina, C., Zilberman, E., et al. (2020). *Proyección de la inflación en chile con métodos de machine learning*. Banco Central de Chile.

Macias, P., Stelmasiak, D., and Szafranek, K. (2023). Nowcasting food inflation with a massive amount of online prices. *International Journal of Forecasting*, 39(2):809–826.

Martinez, A. B., Castle, J. L., and Hendry, D. F. (2022). Smooth robust multi-horizon forecasts. In *Essays in Honor of M. Hashem Pesaran: Prediction and Macro Modeling*, volume 43, pages 143–165. Emerald Publishing Limited.

Meese, R. A. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14(1-2):3–24.

Montenegro, A. (2010). Análisis de series de tiempo. *Bogotá: Pontificia Universidad Javeriana*.

Nelson, C. R. (1972). The prediction performance of the frb-mit-penn model of the us economy. *The American Economic Review*, 62(5):902–917.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., et al. (2022). Forecasting: theory and practice. *International Journal of Forecasting*.

Rapach, D. and Zhou, G. (2013). Forecasting stock returns. In *Handbook of economic forecasting*, volume 2, pages 328–383. Elsevier.