Documentos de Trabajo

ISSN 2409-1863
DT 060-Diciembre 2017
Banco Central de Nicaragua

THE NICARAGUAN INFLATION COMBINATION ASSESSMENT (NICA): A forecast combination system through an efficient forecast path

Carlos E. Castillo-Maldonado
Edson R. Ortiz-Cardona

Banco Central de Nicaragua
Emitiendo confianza y estabilidad

Banco Central de Nicaragua

THE NICARAGUAN INFLATION COMBINATION ASSESSMENT (NICA): A forecast combination system through an efficient forecast path

Carlos E. Castillo-Maldonado - Edson R. Ortiz-Cardona

DT-060-2017

THE NICARAGUAN INFLATION COMBINATION ASSESSMENT (NICA): A forecast combination system through an efficient forecast path[1]

Carlos E. Castillo-Maldonado[2] Edson R. Ortiz-Cardona[3]

Summary

A forecast combination algorithm called NICA is developed and tested to forecast Nicaraguan inflation. This procedure aims to build horizon-specific weights, based on a model's historical performance under a rolling window estimation setting. It considers an endogenous trimming procedure based on statistical distributions for forecast weights, which are constructed for every forecast path period. Inflation forecasts are performed based on ARMA, OLS, SWLS, VAR, and VEC models using quarterly data from 2001Q4 to 2017Q1. It is found statistical support for NICA over more common benchmark combination methods.

JEL Classification codes: C5, C6,
Key words: Forecast Combination, Inflation, Nicaragua

---

[1] The views expressed in this document are the unique responsibility of the authors.

[2] Economic Research Department, Bank of Guatemala. 7a. avenida 22-01 Zona 1, Ciudad de Guatemala. Tel: + (502) 2238-1552, Email: cecm@banguat.gob.gt.

[3] Economic Research Department, Bank of Guatemala. 7a. avenida 22-01 Zona 1, Ciudad de Guatemala. Tel: + (502) 2238-1552, Email: eroc@banguat.gob.gt.

INDEX

I. INTRODUCTION

Combining forecasts to obtain more efficient projections is a method employed in multidisciplinary fields (Clemen, 1989). Since the seminal works of Reid (1968 and 1969) and Bates and Granger (1969), forecast combination has become commonly used among researchers, private firms, and government institutions to improve the accuracy level of their projections. Central banks in particular, have performed research, and developed internal procedures based on forecast combination methods, to provide policymakers more accurate insight about short term expected fluctuations of main variables that influence monetary policy actions, particularly of inflation (Kapetanios et al., 2005 and 2008; Adolfson et al., 2007; Coletti and Murchison, 2002, Samuels and Sekkel, 2013; Bjornland et al., 2012; Aiolfi, Capistran and Timmermann, 2010; González, 2010; Bello, 2009; and Hubrich and Skudelny, 2016).

The reason why forecast combination provides more robust estimates than a single optimal forecast is still a puzzle. Nevertheless, some justifications indicate that this method provides an insurance advantage against possible model misspecification or omitted variable biases (Baumeister, et al., 2015; Bjornland et al., 2012); it is a useful hedging strategy against structural breaks in data (Hendry and Clements, 2004; Diebold and Pauly, 1987; Makridakis, 1989); it is a reasonable approximation for underlying non-linearities (Pesaran and Timmermann, 2005; Marcellino, 2004; Hubrich and Skudelny, 2016); or for simple portfolio diversification arguments (Bates and Granger, 1969).

Despite its well-recognized and demonstrated advantages, there is no common ground about the method to combine forecasts. At this respect it is important to mention that the method to be chosen has to deal with two simultaneous and interrelated issues (Samuels and Sekkel, 2013): i) a weighting scheme; and ii) a preselection of forecasts to combine (also called trimming). On the first case, Timmermann (2010) provides a complete survey of recent methods employed, emphasizing relative performance weights, and equal weighting schemes, as the two methods most commonly used given their simplicity and intuitive properties. On the other issue, forecast trimming methods could be classified into two main categories: exogenous and endogenous. However, both type of procedures appear to provide mixed results. While some studies have found that forecast trimming (particularly hard trimming) provides better estimates (Makridakis and Winkler, 1983), some other studies argue in favor of crowd wisdom: no trimming at all (Stock and Watson, 2002).

This document attempts to add to the current forecast combination literature by applying and comparing different combination methodologies to obtain robust estimates for Nicaraguan inflation. We develop an algorithm, called the Nicaraguan Inflation Combination Assessment (NICA), which consists of an horizon-specific weighting scheme that favors forecasts according to their historical performance. In particular, we construct statistical weight distributions for every forecasted period, based on rolling window estimates, and select just those forecasts whose weights fall on the upper side of such distributions, particularly, those statistically significant at the 5% level. Forecast combination results generated by this methodology are compared with forecast combination outcomes produced by an equal weighting scheme, and by the top performance model for each period predicted. Projections are generated by five kind of models: ARMA, OLS, SWLS, VAR and VEC, which are estimated using quarterly data for the period 2001Q4-2017Q1. Four inflation fundamentals are considered for estimation and forecasting: US inflation, domestic money supply, banking credit to the Nicaraguan private sector, and the Cordoba-US$ nominal exchange rate. Through different variable combinations, and lag structures, we ended up with 309 models and forecasts to combine. Our results favor NICA over benchmark combination methods. In fact, we strongly recommend that NICA's algorithm be incorporated to complement the set of tools used by the Central Bank of Nicaragua to generate macroeconomic forecasts used as basis for monetary policy decisions. The remaining sections of this document are organized as follows. Section II describes our forecast combination methodology; Section III depicts the set of forecast performing models and data span; Section IV provides our empirical results; and Section V Concludes.

II. FORECAST COMBINATIONS

In this section we describe the methods used in our empirical analysis to combine forecasts, and to compare such combinations among different criteria.

A. COMBINATION METHOD

A forecast combination method involves two simultaneous and interrelated issues to deal with (Samuels and Sekkel, 2013): i) a weighting scheme; and ii) a preselection of forecasts to combine. Since Bates and Granger (1969) original work, many weighting schemes have been proposed, some more complex than others: relative performance weights (Granger and Newbold, 1974; Stock and Watson, 2002), equal weighting (Stock and Watson, 2004), least squares estimation (Granger and Ramanathan, 1984), and moment estimators (Elliot and Timmermann, 2004), to mention some of them. The most recent and complete survey of forecast combination methods in economic literature is listed in Timmermann (2010). From all these criteria, the first two of them continue to receive a great deal of attention in the literature because of their intuitive nature and their simplicity. They are also two of the main methodologies used in our empirical analysis, as it will be mentioned below.

The second issue, which is interrelated to the previous one, refers to the number and the kind of forecasts to combine. Should we combine all the forecasts available? Or should we select among them, and combine just the most efficient forecasts? And if so, how to proceed about it? Economic literature classifies forecast trimming into two main categories: exogenous or predetermined, and endogenous to past forecast performance. Exogenous trimming is a selection procedure where the number of forecasts to combine (independent of the method to be used), or a given percentage of the total available forecast set, are previously defined by the researcher (for instance: using just 5 models, or employing 5% of the total number of models available according to their R squared estimated value). On the other hand, endogenous trimming is a selection procedure where the number or forecasts combined will depend upon a relative comparison of past performance (usually through a Root Mean Squared Error, RMSE, criteria) among all available forecasts. In practice, both type of procedures appear to provide mixed results. While some studies have found that forecast trimming provides better results (Makridakis and Winkler, 1983) [4], some other studies argue in favor of crowd wisdom, which implies no trimming at all (Stock and Watson, 2002).

The method employed for empirical analysis in this document is based on forecast past performance efficiency at each of eight quarters ahead. We are interested in a forecast scope of this length, since a central bank's definition of short run tends to be equal or lower than 2 years. The empirical analysis performed was based on an algorithm called the Nicaraguan Inflation Combination Assessment (NICA), which consists of a five-step method to construct a weighting scheme for model combination that includes an ex-ante endogenous procedure for model trimming. The Annex of this document presents a more detailed derivation of it. The first step

[4] Makridakis and Winkler (1983) argue that there is a decreasing marginal benefit from adding forecasts to the combination pool, and that such marginal benefit significantly decreases after considering five-eight forecasts.

of such algorithm consists in performing rolling window estimations for all *K* models of different types considered in this document (ARMA, OLS, SWLS, VAR, and VEC), and generating pseudo out-of-sample inflation forecasts from every model, for each period *S+h* belonging to the established forecast horizon. A second step consists of computing window and model specific forecast weights, $q^{w}_{S+h,k}$, defined as the ratio of model *k* point forecast error's inverse, in absolute terms, $\widehat{\boldsymbol{fi}}^{w}_{S+h,k}$, to the sum overall similar values for all *K* models, $\widehat{FI}^{w}_{S+h}$ ($\widehat{FI}^{w}_{S+h} = \sum_{k=1}^{K} \widehat{fi}^{w}_{S+h,k}$).[5] The third step consists of getting the horizon-specific average weight for each model, $\widehat{\boldsymbol{q}}_{t,k}$, equivalent to the average of model *k* forecast weights computed in each rolling window *w*. In the fourth step we compute the mean and standard deviation of average weights for each period *S+h* that belongs to the forecast horizon. Then an endogenous trimming procedure was performed by selecting just those models whose average weights were above two standard deviations from the mean value. Hence, we renormalized weights by setting to zero those who belong to models that did not made the cut, while making the chosen ones sum up to one. Under this criteria the set of models to be combined were those whose performance was statistically significant at 95% confidence. It is important to mention that the final model set, as well as the best performance model, varied depending on the *S+h* period forecasted. The fifth and final step was multiplying each model out-of-sample forecast, $\hat{y}_{N+h,k}$, times its final horizon-specific normalized average weight, $\widehat{Q}^{n}_{S+h,k}$. The resulting inflation forecast path is what we called NICA.

In addition, to measure NICA's reliability, we compare its RMSE with those resulting from two hard-to-beat benchmarks: i) forecast combination using equal weights for all sample models; and ii) the best performing model for each forecast horizon. Our method for comparing combinations criteria is presented next.

B. COMPARING FORECAST COMBINATION CRITERIA

Accuracy gains from forecast combination methods are usually illustrated by comparing RMSEs or FEs computed from an optimal/proposed combination method, and those

[5] Window size depends on data availability. We considered a 35 observation window (around 9 years) given the short span of available information. Nevertheless, as described in Section IV, we performed a sensibility analysis or our results based on smaller and larger sized windows, and found that our results are robust to different window size specifications.

computed from one or several benchmarks. Nevertheless, when comparisons are made in relative terms, there might be cases when conclusions taken from RMSEs could differ from those taken from FEs. Consider the example illustrated in Table 1, where forecasts from four different models (A, B, C, and D) are combined to forecast inflation for T+1. Each model forecast is weighted according to an equal weighting (EW), and to other (unknown) weighting criteria (OW), and both methods would be compared to determine the most efficient (lower error) inflation forecast combination. Column [2] in panel (a) describes FEs produced by each model; columns [3] and [4] depict the weights assigned to each forecast, according to both criteria; columns [5] and [6] present the calculations required to get the combined FE for each method; while columns [7] and [8] describe the calculations required to obtain the RMSE statistic according to both criteria.

Table 1. Comparing forecast combination criteria

(a)

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
MODEL	FE	EW	OW	FE*EW	FE*OW	$(FE)^2$ *EW	$(FE)^2$ *OW
A	0.16	0.25	1	0.040	0.160	0.006	0.026
B	-0.10	0.25	0	-0.025	0.000	0.003	0.000
C	0.50	0.25	0	0.125	0.000	0.063	0.000
D	-0.04	0.25	0	-0.010	0.000	0.000	0.000
TOTAL				0.130	0.160	0.072	0.026

(b)

STATISTIC	OW/EW	EW Method	OW Method
ABS(FE)		0.130	0.160
RMSE		0.268	0.160
FE Ratio (OW/EW)	1.231		
RMSE Ratio (OW/EW)	0.597		
Z = ABS(FE)*RMSE		**0.035**	**0.026**
Efficiency Gains (OW/EW)	**26.5**		

FE: Forecast Error; EW: Equal Weighting Scheme; OW: Other Weighting Scheme; RMSE: Root Mean Squared Error.

Panel (b) of Table 1 shows the results obtained. The first pair of rows depict FE and RMSE values for period T+1 resulting from each combination method, where the former is in absolute terms, hence ABS(FE). The following pair of rows describe ratios (OW/EW) for both statistics. In each case, a ratio lower than one would imply that the OW method provides a more efficient combination criteria. Likewise, a ratio greater than one would imply a preference to combine model forecasts according to the EW criteria. However, a problem could arise when each ratio provides a different signal.[6] In the example illustrated this is exactly the case. The ABS(FE) ratio (OW/EW) is greater than one (1.231), which indicates that forecasts errors are lower under the EW method. However, the RMSE ratio (OW/EW) is less than one (0.597), indicating a forecast efficiency advantage favoring the OW method.

To avoid getting ambiguous conclusions, as those illustrated above, we propose an alternative statistic, called Z. Consider the last two rows of panel (b) in Table 1. The first of them depicts the value of Z, which is the product of the former statistics ABS(FE) and RMSE. The computed value is a single number. Therefore, the lower value (resulting from the OW method in this example) would indicate the more efficient criteria.[7] The latter row computes efficiency gains from using OW as a forecast combination method with respect to EW. They are equivalent to the percentage difference between both Z values (0.026 vs 0.035), which for this example resulted in 26.5%.

Our proposed methodology to compare forecast combination procedures becomes more useful with a longer forecast horizon. In this case we would have different values for Z, not only for every method, but also for each period forecasted. Therefore, by adding them up, and comparing their cumulative values we would get a more robust conclusion about the most efficient forecast combination method. Hence, efficiency gains would be computed as a percentage difference from such a cumulative sum of Z values. A description of the number and type of models, as well as of all data series employed in our empirical analysis is given in the following section.

---

[6] The same intuitive result is obtained if we compute absolute or percentage differences between both statistics.

[7] Hypothesis testing could be performed under this scenario. In particular a null hypothesis, H_0, could maintain that the difference between both values of Z is not different from zero. To our knowledge there is no statistical distribution identified for Z; this will be subject to further research. Nevertheless, in the meantime, we would argue in favor of choosing the lower value.

III. FORECASTING MODELS AND DATA

A. MODELS

In order to proceed with forecast combination we created a set of models to forecast Nicaraguan inflation. In particular, we employed five kind of models: ARMA(*p,q*), Ordinary Least Squares (OLS), Stepwise Least Squares (SWLS), Vector Autoregression (VAR), and Vector Error Correction (VEC) models. Different lag and variable combinations were performed for each kind of model. All variables were transformed to their logarithmic form, and rolling estimation, through moving windows, was performed for each model specification. Table 2 describes the number of models estimated, the period of data available for estimation, as well as the window size established to produce our results. A brief description of empirical estimations performed under each kind of model is provided below.

1. Autoregressive Moving Average (ARMA) Models

ARMA(*p, q*) models are univariate representations that express a variable (y_t) as a function of its own lags (*p*), and lags (*q*) of the error term (ε_t). Let $\Omega_{ARMA} = ARMA(1,0), ARMA(2,0), \ldots ARMA(p,0), ARMA(1,1), \ldots ARMA(p,q),$ be the set of estimated models. Therefore, the total number of models contained in such a set is equal to $\Omega_{ARMA} = 2p(1+q)$. We let $p = 13$, and $q = 8$. Hence, Ω_{ARMA} is composed of 234 models.[8]

2. Ordinary Least Squares (OLS) Models

The OLS models that we employed are classical econometric representations of variables as a function of their past values, and one or more independent variables and their lags. Let's assume that Ω_{OLS} is the set of all OLS models estimated, which contains all possible combinations of multivariable models, along with all possible combinations of lags. In this case, the total number of OLS models contained in Ω_{OLS} equals to $5p_0 + 4p_1 + 6p_2 + 4p_3 + p_4$. The sub-index "*i*" in the lag expression p_i denotes the number of inflation fundamentals included in each estimation. For instance, p_2 denotes an OLS expression where the Nicaraguan Consumer Price Index was estimated using its own lag(s), and two of its fundamentals. We let

[8] The values for p and q were the maximum values allowed, given data availability.

$p_0 = 2$; and $p_1 = p_2 = p_3 = p_4 = 1$, which implies that Ω_{OLS} contains 25 (=10+4+6+4+1) models.[9]

Table 2. Data span and model sets

Data Span	Window Size	Model Sets (Ω)					
		Ω_{ARMA}	Ω_{OLS}	Ω_{SWLS}	Ω_{VAR}	Ω_{VEC}	Ω_{TOTAL}
2001Q4-2017Q1	35	234	25	20	15	15	309

Source: Own elaboration.

3. Stepwise Least Squares (SWLS) Models

SWLS is an iterative algorithm proposed by Efroymson (1960) to automatically obtain OLS regressions' best fit.[10] Each model representation is identical to those of the previous section, but the final results differ, since the SWLS algorithm was established to select just those regressors whose *p-value* was lower or equal to 0.05. We let $p_0 = p_1 = p_2 = p_3 = p_4 = 1$. Hence, it would imply that Ω_{SWLS} contains 20 (=5+4+6+4+1) models.

4. Vector Autoregression (VAR) Models

Unrestricted VAR models are systems of equations that express each variable as a function of its own past values, and lags of the remaining variables in the system. We tried all combinations from two to five variables (domestic IPC and its four fundamentals), and different lag specifications. The set of all VAR models estimated, Ω_{VAR}, contained $4P_2 + 6P_3 + 4P_4 + P_5$ models. As before, the sub-index in the lag expression denotes the number of inflation fundamentals included in each VAR model. By letting $P_2 = P_3 = P_4 = P_5 = 1$, it would imply that Ω_{VAR} is composed of 15 (=4+6+4+1) models (see Table 2).

5. Vector Error Correction (VEC) Models

We estimated equilibrium VEC models for one cointegrating relationship through the Johansen procedure. As with the previous case, combinations for all possible variable and lag

[9] We made exercises (not reported) with higher lag values for each fundamental varialbe, but they didn't add significantly to our final results. This was also the case for the rest of models presented in this section.
[10] Derksen and Keselman (1992), and Burnham and Anderson (1998) provide a description of the algorithm, and describe some advantages and disadvantages of this methodology.

specifications were performed. Therefore, the number of lags considered to estimate each VEC model is the same as in the VAR case. Hence, Ω_{VEC} is composed of 15 models.

B. DATA

Nicaraguan inflation was estimated and forecasted based on the models described before, using quarterly data from 2001Q4 to 2017Q1.[11] The data set was obtained from the Central Bank of Nicaragua's website, and included information for four well-identified inflation fundamentals: i) US Inflation; ii) Nominal Exchange Rate Cordoba-US Dollar (C/US$); iii) Real Money Supply; and iv) Real Banking Credit. The US represents Nicaragua's main trading partner. Hence, US price fluctuations, along with C/US$ variations are transferred fast, and almost full to domestic prices (Treminio, 2014). Moreover, changes in domestic money supply and banking credit affect domestic inflation indirectly, through their effect on domestic GDP.[12] Following Clements and Hendry (1999), data were not deseasonalized, nor detrended to avoid missing important forecasting information. This is also why we estimated ARMA instead ARIMA models. Finally, for empirical purposes all variables were transformed to their log form. The results obtained are described in the following section.

IV. EMPIRICAL RESULTS

Our main findings are presented in this section. First we present our best performing models according to their past performance. Then forecast accuracy gains from NICA are portrayed with respect to the benchmarks used to compare our results. Finally, we describe results from a sensitivity analysis under different window sizes for rolling regressions.[13]

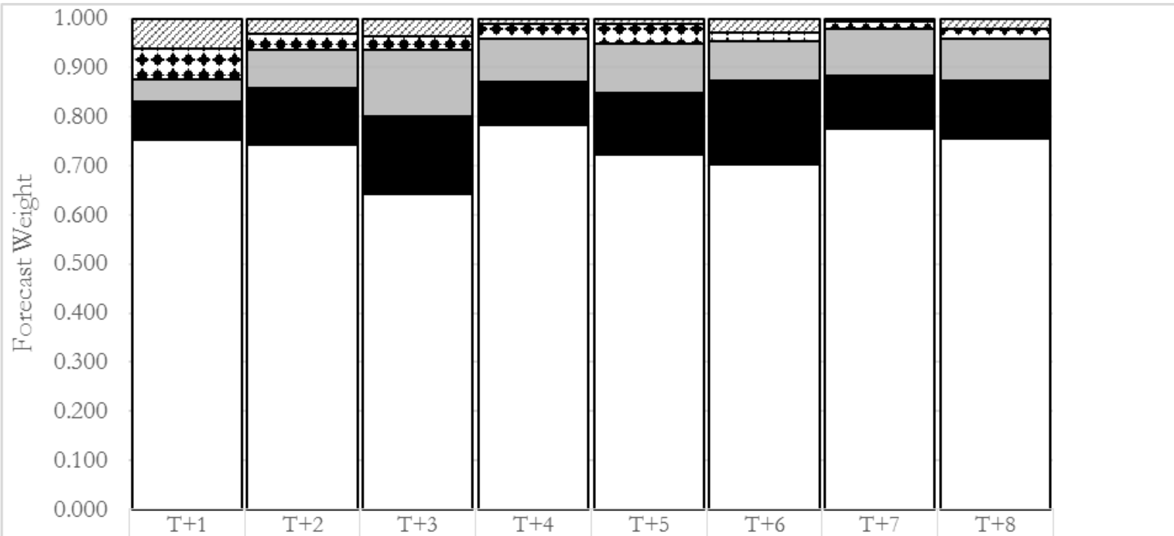
[11] Data is constrained to start in 2001Q4, since we were not able to obtain information for all series before that period.

[12] It is important to mention that we used money and credit as a proxy for Nicaraguan GDP, since a quarterly series for such a variable is not available for the whole period under consideration.

[13] Our empirical analysis and outcome were based on computer programs made in Eviews. Those programs are available upon request for further replication of our results. In addition, suggestions from the authors can be provided if NICA is to be implemented as a complementary tool at the Central Bank of Nicaragua.

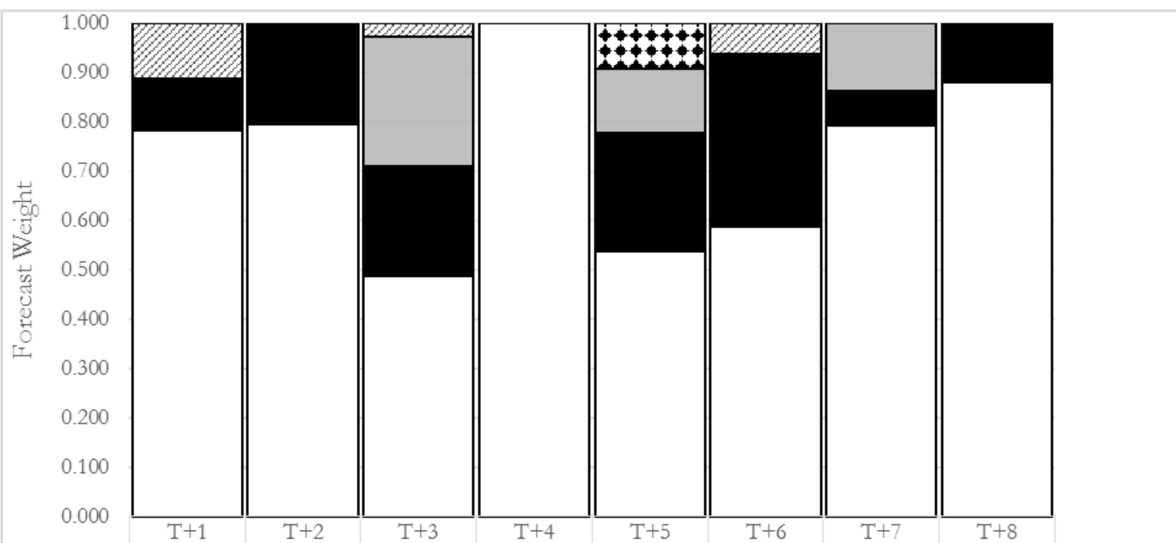
Figure 1. Model Weight Composition

(a) Untrimmed estimation



	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8
VEC	0.060	0.030	0.036	0.010	0.011	0.029	0.007	0.020
VAR	0.065	0.033	0.029	0.030	0.040	0.018	0.014	0.022
SWLS	0.045	0.077	0.136	0.090	0.100	0.079	0.096	0.085
OLS	0.078	0.117	0.158	0.087	0.127	0.171	0.107	0.116
ARMA	0.752	0.742	0.642	0.783	0.722	0.703	0.777	0.757

(b) Endogenous Trimming (Weights above Mean + 2 SD)



	T+1	T+2	T+3	T+4	T+5	T+6	T+7	T+8
VEC	0.111	0.000	0.027	0.000	0.000	0.062	0.000	0.000
VAR	0.000	0.000	0.000	0.000	0.092	0.000	0.000	0.000
SWLS	0.000	0.000	0.263	0.000	0.130	0.000	0.138	0.000
OLS	0.106	0.205	0.222	0.000	0.238	0.350	0.069	0.119
ARMA	0.782	0.795	0.488	1.000	0.539	0.588	0.793	0.881

A. TOP PERFORMING FORECASTS

Figure 1 depicts forecasts weights grouped by model type (ARMA, OLS, SWLS, VAR, and VEC). Panel (a) presents the results from an unrestricted (untrimmed) estimation, while panel (b) presents the results from an endogenous trimming method, where the final forecasts combination was made from those forecasts whose weights were two standard deviations above the mean, at each forecasts horizon (from T+1 to T+8). Both graphs present model types ordered by past performance, that is, by their cumulative sum of forecast weights. From the untrimmed estimation we found that inflation forecasts from ARMA models outperform the remaining ones at each forecast period projected, since their cumulative sum of weights ranged from 64.2% to 78.3%. Moreover, OLS, and SWLS forecast importance is quite similar, and ranged from 4.5% to 17.1%. Forecasts from VAR and VEC models lagged much behind, since their cumulative sum of weights do not surpass 6.5% (in T+1). Nonetheless, such results could also be affected by the number of models used to produce forecast in the first place. In fact, when we trimmed the models to account just for the top performers (those models whose weights are over two standard deviations from the mean), we found that top performer type models are still in the same order as above (see panel b). In particular, forecast weights for ARMA models represented in average over 73% of the total. Moreover, the average cumulative sum of weights for OLS models was about 17%, while the weight sum of the remaining models added up to about 10%.[14]

B. FORECAST EFFICIENCY GAINS

Our main results are presented in Table 3. Panels (a) and (b) depict Z values for NICA, as well as for both comparison methods used as benchmarks: the Equal Weight (EW) and the Top Model (T) criteria. Results presented in panel (a) include all 309 model forecasts estimated, while results in panel (b) just include the top performing models. Cumulative sums for Z values were performed for a short run horizon (from T+1 to T+2), for a medium term span (from T+1 to T+4), and for the whole forecast path (from T+1 to T+8). Two main conclusions can be derived from this information. First, according to both untrimmed and endogenously trimmed estimations, NICA is the most efficient procedure to combine Nicaraguan inflation forecasts.

[14] The only exception is T+8, where SWLS models' importance represented 29.3% of the total.

Table 3. Forecast Combination results

(a) Untrimmed estimation

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.156	0.188	1.118
T + 2	0.272	0.330	0.462
T + 3	0.413	0.513	2.879
T + 4	0.435	0.898	0.264
T + 5	0.491	1.044	2.409
T + 6	0.517	0.868	0.864
T + 7	0.191	0.866	0.651
T + 8	0.507	1.037	2.780
Cumulative Sum (Z_{t+1} to Z_{t+2})	0.428	0.518	1.580
Cumulative Sum (Z_{t+1} to Z_{t+4})	1.276	1.928	4.724
Cumulative Sum (Z_{t+1} to Z_{t+8})	2.982	5.743	11.427

(b) Endogenous Trimming (Weights above Mean + 2SD)

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.112	0.188	1.118
T + 2	0.305	0.330	0.462
T + 3	0.572	0.513	2.879
T + 4	0.432	0.898	0.264
T + 5	0.292	1.044	2.409
T + 6	0.774	0.868	0.864
T + 7	0.229	0.866	0.651
T + 8	0.100	1.037	2.780
Cumulative Sum (Z_{t+1} to Z_{t+2})	0.417	0.518	1.580
Cumulative Sum (Z_{t+1} to Z_{t+4})	1.421	1.928	4.724
Cumulative Sum (Z_{t+1} to Z_{t+8})	2.816	5.743	11.427

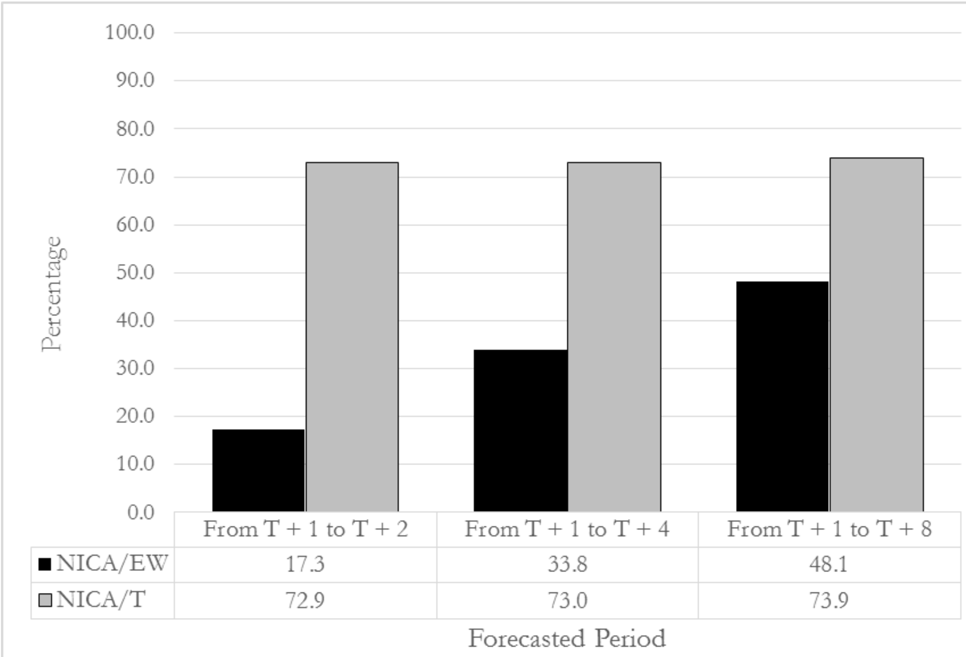
NICA: Nicaraguan Inflation Combination Assessment; EW: Equal weights; T: Top weighted model;

FE: Forecast Error; ABS: Absolute value; RMSE: Root Mean Squared Error;

Z = ABS(FE)*RMSE.

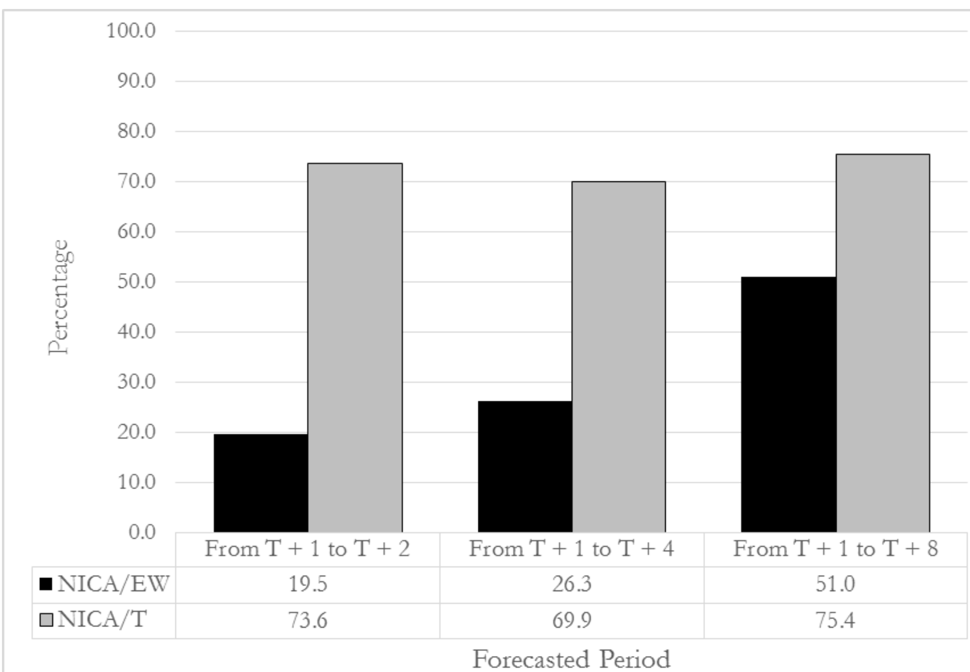
Figure 2. Efficiency Gains from NICA

(1) Untrimmed estimation



NICA: Nicaraguan Inflation Combination Assessment; EW: Equal Weights; T: Top Weighted model

(b) Endogenous Trimming (Weights above Mean + 2 SD)

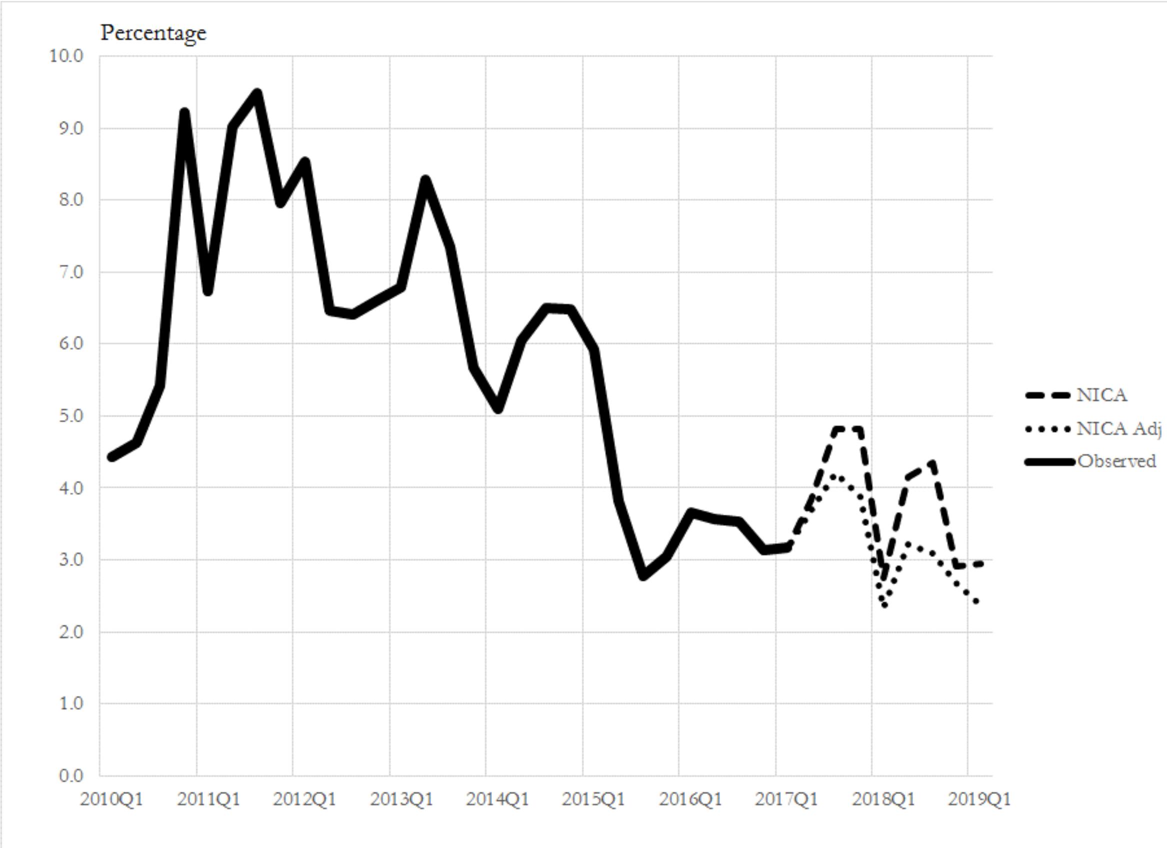


NICA: Nicaraguan Inflation Combination Assessment; EW: Equal Weights; T: Top Weighted

Cumulative sums for Z are lower under the NICA criterion, particularly in the short run. An interesting result is that both NICA and the EW criteria provide more efficient short run forecasts than the T method (a hard-to-beat method in practice). The second important conclusion, which particularly holds for the NICA criteria is that endogenous trimming performs more efficiently in the short term (from T+1 to T+2), while in larger horizons untrimmed forecast combinations are more efficient. This result might be due to model forecast inertia or to the fact that the Nicaraguan inflation series might be subject to nonlinearities or to continuous structural breaks, so it could be better forecasted through a more ample set of models.

Efficiency gains at each forecast horizon are depicted in Figure 2. The vertical axis shows Z values' percentage variations computed between the NICA combination method and each of the remaining criteria, at each forecast horizon established. The darker column represents NICA's forecast efficiency gains respect to the EW method, while the lighter column denotes its efficiency gains respect to the T criteria. Panel (a) presents results from the untrimmed estimation. For this case, forecast efficiency gains are produced with respect to both benchmark methodologies, and the range from 17.3% to 48.1% for the EW combination method, and from 72.9% to 73.9% for the T method. While NICA's efficiency gains seem to remain stable (around 73%), they are increasing along the forecast path with respect to the EW method. Panel (b) depicts NICA's efficiency gains at the three forecast horizons established under our endogenously trimmed method. Efficiency gains are stronger, particularly for longer horizons with respect to the other two forecast combination criteria. In summary, NICA's efficiency gains are strongly predominant with respect to both comparison benchmarks (EW and T) for either trimmed or untrimmed estimations. Furthermore, endogenous trimming provides the highest efficiency gains, particularly for the latter forecast path periods, which makes NICA a suitable methodology to generate inflation forecasts for monetary policy considerations. Figure 3 shows Nicaraguan inflation forecasts using NICA. The darker line shows observed year to year inflation from 2010Q1 to 2017Q1. The wide dotted line from 2017Q2 to 2019Q1 shows forecasts under NICA. In fact, the series forecasted was the unseasoned Nicaraguan IPC; hence, internual inflation was calculated on a posterior basis. In addition, we computed inflation forecasts through the same methodology from 2014Q1 to 2017Q1 (not shown), in order to calculate an average forecast error, which we latter added to the original NICA forecast. Such adjusted forecast (NICA Adj) is shown by the narrow dotted line in Figure 3.

Figure 3. Nicaraguan Inflation Forecasts under NICA



C. SENSITIVITY ANALYSIS

We tested our results' robustness through sensitivity analysis employing lower and larger size rolling windows. Tables 4 and 5 present the results obtained using window sizes of 40 and 30 observations, respectively.[15] In both tables, panel (a) shows results for the untrimmed case, while panel (b) describe those for the endogenous trimming methodology. Our main conclusions still hold, particularly with respect to the EW method: NICA leads to lower Z values along the whole forecast path for both trimmed and untrimmed estimations. We attribute such a robustness to model forecast inertia: the fact that in-sample forecast combination arrangements are still hold when forecasting a series out of the sample.

[15] We also tested the methodology using different window sizes, and the conclusions still hold.

Table 4. Sensitivity Analysis (Window Size = 40 observations)

(a) Untrimmed estimation

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.330	0.527	0.155
T + 2	0.582	0.840	0.883
T + 3	1.154	1.463	0.321
T + 4	1.271	2.204	2.185
T + 5	0.530	1.357	0.158
T + 6	0.842	1.381	0.575
T + 7	0.081	1.149	1.668
T + 8	0.222	1.259	5.774
Cumulative Sum (Z_{t+1} to Z_{t+2})	0.912	1.367	1.038
Cumulative Sum (Z_{t+1} to Z_{t+4})	3.337	5.033	3.544
Cumulative Sum (Z_{t+1} to Z_{t+8})	5.012	10.180	11.719

(b) Endogenous Trimming (Weights above Mean + 2SD)

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.594	0.527	0.155
T + 2	0.650	0.840	0.883
T + 3	1.473	1.463	0.321
T + 4	0.695	2.204	2.185
T + 5	0.284	1.357	0.158
T + 6	1.020	1.381	0.575
T + 7	0.528	1.149	1.668
T + 8	0.000	1.259	5.774
Cumulative Sum (Z_{t+1} to Z_{t+2})	1.244	1.367	1.038
Cumulative Sum (Z_{t+1} to Z_{t+4})	3.412	5.033	3.544
Cumulative Sum (Z_{t+1} to Z_{t+8})	5.244	10.180	11.719

NICA: Nicaraguan Inflation Combination Assessment; EW: Equal weights; T: Top weighted model;
FE: Forecast Error; ABS: Absolute value; RMSE: Root Mean Squared Error;
Z = ABS(FE)*RMSE.

Table 5. Sensitivity Analysis (Window Size = 30 observations)

(a) Untrimmed estimation

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.483	0.595	0.394
T + 2	0.485	0.614	0.632
T + 3	0.645	0.806	0.062
T + 4	0.859	0.973	0.259
T + 5	0.351	0.841	1.017
T + 6	0.506	0.806	1.061
T + 7	0.225	0.700	1.818
T + 8	0.383	0.925	0.248
Cumulative Sum (Z_{t+1} to Z_{t+2})	0.968	1.209	1.026
Cumulative Sum (Z_{t+1} to Z_{t+4})	2.472	2.988	1.347
Cumulative Sum (Z_{t+1} to Z_{t+8})	3.936	6.260	5.491

(b) Endogenous Trimming (Weights above Mean + 2SD)

FORECASTED PERIOD	Z(NICA)	Z(EW)	Z(T)
T + 1	0.567	0.595	0.394
T + 2	0.220	0.614	0.632
T + 3	0.855	0.806	0.062
T + 4	1.311	0.973	0.259
T + 5	0.218	0.841	1.017
T + 6	1.463	0.806	1.061
T + 7	0.067	0.700	1.818
T + 8	0.123	0.925	0.248
Cumulative Sum (Z_{t+1} to Z_{t+2})	0.787	1.209	1.026
Cumulative Sum (Z_{t+1} to Z_{t+4})	2.953	2.988	1.347
Cumulative Sum (Z_{t+1} to Z_{t+8})	4.824	6.260	5.491

NICA: Nicaraguan Inflation Combination Assessment; EW: Equal weights; T: Top weighted model;
FE: Forecast Error; ABS: Absolute value; RMSE: Root Mean Squared Error;

Z = ABS(FE)*RMSE.

It is important to mention that even when NICA's algorithm was originally built to produce inflation forecasts, it can also be adapted to generate forecasts for other series, particularly monetary policy sensitive variables such as economic growth, nominal and real exchange rates, or foreign inflation and output, among others. Hence, we strongly recommend that NICA's algorithm be incorporated within the set of tools used by the Central Bank of Nicaragua to generate macroeconomic forecasts that are used as basis for monetary policy decisions.

V. CONCLUSIONS

We attempt to add to the current forecast combination literature by developing a methodology to combine forecasts, and to evaluate their outcomes, in order to forecast Nicaraguan inflation. This methodology, called the Nicaraguan Inflation Combination Assessment (NICA), is an horizon-specific weighting scheme, where forecasts are selected (trimmed) according to their historical statistical significance. Projections were generated by five type of models: ARMA, OLS, SWLS, VAR and VEC, based on quarterly data for the period 2001Q4-2017Q1. Four inflation fundamentals are considered for estimation and forecasting: US inflation, domestic money supply, banking credit to the Nicaraguan private sector, and the Cordoba-US$ nominal exchange rate. Forecast combination results generated by NICA are found to over perform two high standard benchmarks, which are difficult to beat in empirical analysis: Equal Weighting (EW) and Top Model (T) criteria. Our results and sensitivity analysis favor endogenous trimming both for the short and long term forecast horizons.

BIBLIOGRAPHY

Adolfson, M., Andersson, M.K., Linde, J., Villani, M., and Vredin, A. (2007). Bayesian forecast combination for VAR models. International Journal of Central Banking, 3, 111-144.

Aiolfi, M., Capistrán, C., Timmerman, A. (2010). Forecast combinations. Banco de México Working Papers No. 2010-04.

Bates J.M., Granger, C.W.J. (1969). The combination of forecasts. Operations Research Quarterly 20, 451-468.

Baumeister, C., Kilian, L. (2015). Forecasting the real price of oil in a changing world: a forecast combination approach. Journal of Business and Economic Statistics, 33:3 338-351.

Bello, O. (2009). Combinación de pronósticos de inflación en Nicaragua. Documento de Trabajo DT-011, Banco Central de Nicaragua.

Burnham, K.P., and Anderson D.R. (1998). Model Selection and Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Bjornland, H., Gerdrup, K., Jore, A.S., Smith, C., Thorsrud, L.A. (2012). Does Forecast Combination Improve Norges Bank Inflation Forecasts? Oxford Bulletin of Economics and Statistics, 74, 2, 163-179.

Clemen, R.T. (1989). Combining forecasts: A review and annotates bibliography. International Journal of Forecasting, 5, 559-583.

Clements, M.P, and Hendry, D. (1999). Forecasting Non-stationary Economic Time Series. Cambridge, MIT Press.

Coletti, D. and Murchison, S. (2002). Models in policy-making. Bank of Canada Review, Summer, 19-26.

Derksen, S., and Keselman, H. J. (1992). Backward, Forward and Stepwise Automated Subset Selection Algorithms: Frequency of Obtaining Authentic and Noise Variables. British Journal of Mathematical and Statistical Psychology, 45, 265–282.

________________________ (1987). Structural change and the combination of forecasts. Journal of Forecasting 6, 21-40.

Efroymson, M.A. (1960). Multiple Regression Analysis. In *Mathematical Methods for Digital Computers*, by A. Ralston and H.S.Wilf, Wiley & Sons, New York.

Elliot, G., Timmermann, A. (2004). Optimal forecast combinations under general loss functions and forecast error distributions. Journal of Econometrics, 122 (1), 47-79.

González, E. (2010). Bayesian model averaging: an application to forecast inflation in Colombia. Borradores de Economía, No. 604-2010. Banco de la República.

Granger, C., and Newbold, P. Spurious regressions in econometrics. Journal of Econometrics, 2:2, 111-120.

Granger, C., and Ramanathan, R. (1984). Improved methods of combining forecasts. Journal of Forecasting, 3(2), 197-204.

Hendry, D., and Clements, M. (2004). Pooling of forecasts. Econometrics Journal, 7, 1-31.

Hubrich, K., Skudelny, F. (2016). Forecast combination for Euroarea inflation: a cure in times of crisis? Finance and Economics Discussion Series 2016-104. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2016.104.

Kapetanios, G., Labhard, V., Price, S. (2005). Forecast using Bayesian and Information Theoretic Averaging: An Application to UK Inflation. Discussion Paper 268, Bank of England.

___________________________________(2008). Forecast combination and the Bank of England's suite of statistical forecasting models. Economic Modeling, 25, 772-792.

Makridakis, S. and Winkler, R. (1983). Average of forecasts: some empirical results. Management Science, 29, 987-996.

Makridakis, S. (1989). Why combining works? International Journal of Forecasting, 5, 601-603.

Marcellino, M. (2004). Forecast pooling for short time series of macroeconomic variables. Oxford Bulletin of Economics and Statistics, 66, 91-112.

Pesaran, M.H. and Timmerman, A. (2005). Selection of estimation window in the presence of breakes. Mimeo, Cambridge University and Department of Economics, University of California, San Diego.

Reid, D.J. (1968). Combining three estimates of gross domestic product. Economica, 35, 431-444.

________ (1969). A comparative study of time series prediction techniques on economic data. PhD thesis, University of Nottingham.

Samuels, J., Sekkel, R, (2013). Forecasting with many models: model confidence sets and forecast combination. Bank of Canada, Working Paper 2013-11.

Stock, J. and Watson, M. (2002). Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics, 20, 147-162.

Stock, J. and Watson, M. (2004). Combination forecasts of output growth in a seven-country data set. Journal of Forecasting, 23, 405-430.

Tamayo, E. and Lalil, M. Pass-through of International Food Prices to Domestic Inflation During and After the Great Recession: Evidence from a Set of Latin American Economies. Intern-American Development Bank. 167-170.

Timmermann, A. (2010). Forecast combinations. In: *Handbook of Economic Forecasting* I-4, Elsevier, 136-196.

Treminio, J.C. (2014). Factores macroeconómicos que afectan el Traspaso Cambiario a Precios. Documento de Trabajo DT-048, Banco Central de Nicaragua.

ANNEX: NICA'S ALGORITHM

The Nicaraguan Inflation Combination Assessment (NICA) is designed to generate an efficient out-of-sample forecast path for a finite time series, based on a variety of forecasting models. A forecast path is the number of periods that such a time series will be forecasted, which for this case it will be restricted to be lower or equal than the number of observations in a sample.

Let's assume that $\boldsymbol{y}_t = (y_{1,} \, y_{2,} \, y_{3,} \, \dots . \, y_N)'$ is an N-size vector containing a finite time series, which we are interested to forecast for *H* periods ahead of the sample size (where $H = 8 < N$). To perform this task we can use *K* models of the following types: ARMA, OLS, SWLS, VAR, and VEC. Hence, the specific out-of-sample forecast path for $\boldsymbol{y}_t$ could be represented by $\hat{\boldsymbol{y}}_t = (\hat{y}_{N+1,} \, \hat{y}_{N+2,} \, \dots . \, \hat{y}_{N+H})'$. In addition, let $\boldsymbol{y}_t^w = (y_1^w, y_2^w, y_3^w, \dots . \, y_S^w)'$ be an *S*-size vector containing an ordered subsample of the original series (where $S < N$), such that $\boldsymbol{y}_t^w$ is contained *W* times within $\boldsymbol{y}_t$. In other words, there could be a total of *W* continuous-size-*S* windows that can be derived from the original sample. Based on this information, NICA can be computed through the following steps.

STEP 1. ROLLING REGRESSIONS

The first step consists of estimating rolling regressions of sample size *S* for each model *k*, so to generate in-sample forecasts for *h* periods ahead, for *h = 1, 2,…H*. Those forecasts can be arranged in the *H x W* matrix depicted in expression (A.1).

$$\hat{\boldsymbol{y}}_{t,k}^w = \begin{pmatrix} \hat{y}_{S+1,k}^1 & \hat{y}_{S+1,k}^2 & \cdots & \hat{y}_{S+1,k}^W \\ \hat{y}_{S+2,k}^1 & \hat{y}_{S+2,k}^2 & \cdots & \hat{y}_{S+2,k}^W \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{S+H,k}^1 & \hat{y}_{S+H,k}^2 & \cdots & \hat{y}_{S+H,k}^W \end{pmatrix}_{HxW} \qquad \text{(A.1)}$$

Where the columns of $\hat{\boldsymbol{y}}_{t,k}^w$ represent *H*-size forecast vectors performed by model *k*. The number of column vectors is equivalent to *W*, which is the number of size-*S* sample windows that could be estimated from the whole sample size *N*.

STEP 2. ABSOLUTE FORECAST ERRORS AND MODEL WEIGHTS

A different weight, $q^{w}_{S+h,k}$, is assigned to each model *k*, for *k = 1, 2,…K*, according to its performance (or forecast accuracy) to project each period *S+h*, for *h = 1, 2,…H*. Consider expression (A.2), which is a *H x W* matrix containing the inverse of in-sample absolute forecast errors generated by model *k*.

$$\boldsymbol{\widehat{fi}}^{w}_{t,k} = \begin{pmatrix} \widehat{fi}^{1}_{S+1,k} & \widehat{fi}^{2}_{S+1,k} & \cdots & \widehat{fi}^{W}_{S+1,k} \\ \widehat{fi}^{1}_{S+2,k} & \widehat{fi}^{2}_{S+2,k} & \cdots & \widehat{fi}^{W}_{S+2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{fi}^{1}_{S+H,k} & \widehat{fi}^{2}_{S+H,k} & \cdots & \widehat{fi}^{W}_{S+H,k} \end{pmatrix}_{HxW} \tag{A.2}$$

An absolute forecast error represents the distance, from its observed value, of an *S+h* forecast generated by model *k* based on a sample window *w*. Therefore, the smaller its value, the more accurate the forecast. Now, its inverse value could be interpreted as model *k*'s importance to forecast *S+h*, since the lower the absolute forecast error, the higher its inverse. For instance, $\widehat{fi}^{3}_{S+2,5}$ represents the absolute forecast error inverse for the *S+2* in-sample forecast performed by the 5th model (from a total of *K* models) in the third window estimated (out of *W*); the greater its value (the lower its absolute forecast error), the more accurate is model 5th to forecast the second period ahead of the sample size, so the greater should be the weight assigned to its *S+2* forecast. To get such weights, lets firs add up (A.2) over all *k* models, and represent that summation in the *H x W* matrix $\boldsymbol{\widehat{FI}}^{w}_{t}$, as illustrated in (A.3).

$$\boldsymbol{\widehat{FI}}^{w}_{t} = \sum_{k=1}^{K} \begin{pmatrix} \widehat{fi}^{1}_{S+1,k} & \widehat{fi}^{2}_{S+1,k} & \cdots & \widehat{fi}^{W}_{S+1,k} \\ \widehat{fi}^{1}_{S+2,k} & \widehat{fi}^{2}_{S+2,k} & \cdots & \widehat{fi}^{W}_{S+2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{fi}^{1}_{S+H,k} & \widehat{fi}^{2}_{S+H,k} & \cdots & \widehat{fi}^{W}_{S+H,k} \end{pmatrix} = \begin{pmatrix} \widehat{FI}^{1}_{S+1} & \widehat{FI}^{2}_{S+1} & \cdots & \widehat{FI}^{W}_{S+1} \\ \widehat{FI}^{1}_{S+2} & \widehat{FI}^{2}_{S+2} & \cdots & \widehat{FI}^{W}_{S+2} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{FI}^{1}_{S+H} & \widehat{FI}^{2}_{S+H} & \cdots & \widehat{FI}^{W}_{S+H} \end{pmatrix} \tag{A.3}$$

Where each value on the right hand side matrix represents the summation overall absolute forecast error inverse values generated by every model *k* when forecasting each *S+h* period at the specific window *w*. In other words, $\widehat{FI}^{3}_{S+2}$ stands for the summation of absolute forecast error inverse values that result from forecasting *S+2* by all *K* models during the third rolling regression window ($\widehat{FI}^{3}_{S+2} = \sum_{k=1}^{K} \widehat{fi}^{3}_{S+2,k}$). Therefore, model *k*'s weight or significance for each rolling window estimation, is obtained by dividing each element of (A.2) by its corresponding element in (A.3), as indicated below.

$$\hat{\boldsymbol{q}}_{t,k}^{w} = \begin{pmatrix} \frac{\widehat{fi}_{S+1,k}^{1}}{\widehat{FI}_{S+1}^{1}} & \frac{\widehat{fi}_{S+1,k}^{2}}{\widehat{FI}_{S+1}^{2}} & \cdots & \frac{\widehat{fi}_{S+1,k}^{W}}{\widehat{FI}_{S+1}^{W}} \\ \frac{\widehat{fi}_{S+2,k}^{1}}{\widehat{FI}_{S+2}^{1}} & \frac{\widehat{fi}_{S+2,k}^{2}}{\widehat{FI}_{S+2}^{2}} & \cdots & \frac{\widehat{fi}_{S+2,k}^{W}}{\widehat{FI}_{S+2}^{W}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\widehat{fi}_{S+H,k}^{1}}{\widehat{FI}_{S+H}^{1}} & \frac{\widehat{fi}_{S+H,k}^{2}}{\widehat{FI}_{S+H}^{2}} & \cdots & \frac{\widehat{fi}_{S+H,k}^{W}}{\widehat{FI}_{S+H}^{W}} \end{pmatrix} = \begin{pmatrix} \hat{q}_{S+1,k}^{1} & \hat{q}_{S+1,k}^{2} & \cdots & \hat{q}_{S+1,k}^{W} \\ \hat{q}_{S+2,k}^{1} & \hat{q}_{S+2,k}^{2} & \cdots & \hat{q}_{S+2,k}^{W} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{q}_{S+H,k}^{1} & \hat{q}_{S+H,k}^{2} & \cdots & q_{S+H,k}^{W} \end{pmatrix} \tag{A.4}$$

Each element of matrix $\hat{\boldsymbol{q}}_{t,k}^{w}$ represents the weight or importance of model *k* when forecasting period *S+h* at rolling window *w*. Notice that each element of $\hat{\boldsymbol{q}}_{t,k}^{w}$ has a value between zero and one. The higher its value, or the closer is it to one, the more accurate its forecast of *S+h*, and the more important would be model *k* (or the higher weight will have model *k*) when forecasting *N+h*. Moreover, the summation of all *k* weights for the same rolling window *w*, and for each *S+h* period forecasted, is equal to one ($\sum_{k=1}^{K} \hat{q}_{S+h,k}^{w} = 1$; for *w = 1,2,…W* and *h = 1,2,…H*).

STEP 3. AVERAGE WEIGHT FOR EACH MODEL

This step aims to reduce the *H x W* matrix $\hat{\boldsymbol{q}}_{t,k}^{w}$, depicted in expression (A.4), to a column vector *H* that contains an optimal weight for each *S+h* period forecasted by model *k*. This is done by taking a simple average of weights, overall rolling regression windows, computed for the same forecasting period *S+h*,; that is, to compute an average of all values that belong to the same row of matrix $\hat{\boldsymbol{q}}_{t,k}^{w}$. In matrix notation, this operations is performed by multiplying (A.4) by a *W x 1* vector containing a constant value equivalent to the inverse of the number of windows (1/*W*), as described below.

$$\hat{\boldsymbol{q}}_{t,k} = \begin{pmatrix} \hat{q}_{S+1,k}^{1} & \hat{q}_{S+1,k}^{2} & \cdots & \hat{q}_{S+1,k}^{W} \\ \hat{q}_{S+2,k}^{1} & \hat{q}_{S+2,k}^{2} & \cdots & \hat{q}_{S+2,k}^{W} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{q}_{S+H,k}^{1} & \hat{q}_{S+H,k}^{2} & \cdots & q_{S+H,k}^{W} \end{pmatrix}_{HxW} \begin{pmatrix} \frac{1}{W} \\ \vdots \\ \vdots \\ \frac{1}{W} \end{pmatrix}_{Wx1} = \begin{pmatrix} \hat{q}_{S+1,k} \\ \hat{q}_{S+2,k} \\ \vdots \\ \hat{q}_{S+H,k} \end{pmatrix}_{Hx1} \tag{A.5}$$

Where each term of (A.5) is given by $\hat{q}_{S+h,k} = \frac{1}{W}\hat{q}^{1}_{S+h,k} + \frac{1}{W}\hat{q}^{2}_{S+h,k} + \cdots + \frac{1}{W}\hat{q}^{W}_{S+h,k}$. This term represents model *k*'s pseudo out-of-sample forecasting efficiency to predict *S+h*, and its value fluctuates between zero and one. Therefore, the greater its value (the closer it is to one), the more accurate is model *k*'s *S+h* forecast. Notice that model *k*'s forecast efficiency does not necessarily have to hold among all forecast path periods; in other words, model *k* could be an efficient predictor of *S+1*, but it is, not necessarily, an efficient predictor of periods *S+2, S+3,..S+H.*

STEP 4. AVERAGE WEIGHT DISTRIBUTIONS AND ENDOGENOUS TRIMMING

The last step involves getting rid of the models whose forecast efficiency, denoted by its average weight $\hat{q}_{S+h,k}$, is too low relatively to the rest of models. Consider first an expansion of expression (A.5), into an *H x K* matrix containing average weights for all *K* models used to generate pseudo out-of-sample forecasts:

$$\widehat{\boldsymbol{Q}}_{t,k} = \begin{pmatrix} \hat{q}_{S+1,1} & \hat{q}_{S+1,2} & \cdots & \hat{q}_{S+1,K} \\ \hat{q}_{S+2,1} & \hat{q}_{S+2,2} & \cdots & \hat{q}_{S+2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{q}_{S+H,1} & \hat{q}_{S+H,2} & \cdots & \hat{q}_{S+H,K} \end{pmatrix}_{HxK} \qquad \text{(A.6)}$$

Each column of matrix $\widehat{\boldsymbol{Q}}_{t,k}$ contains model *k*'s average weights (for *k=1,2,…K*) that result from forecasting $\boldsymbol{y}_t$ for each period of the forecast horizon: *S+1, S+2,…S+H*, represented by each row of (A.6). Notice that average weights from a single model *k* are not necessarily equally efficient for all periods forecasted, since it could be a good short-term predictor, but a poor long term forecaster. As mentioned above, matrix $\widehat{\boldsymbol{Q}}_{t,k}$ row values contain each model's average accuracy to predict a specific period of the forecast path. Therefore, the summation overall same row values is equal to one $\left(\sum_{k=1}^{K}\hat{q}_{S+h,k} = 1\right)$ for each *h=1,2,…H*. Furthermore, we computed the mean and standard deviation for all elements in a row, so to obtain an average weight distribution, which allowed us to select models based on above average performance. Through this endogenous trimming method we could get rid of (set to zero) those average weight values at the lower extreme of each distribution. The proportion of models whose average weights were set to zero, represented those who were below two standard deviations above the mean, so we ended up with the top 5% performers. Hence, we also renormalized each row, so that their elements were still add up to one. The resulting matrix will be similar in dimensions to (A.6), but it will contain the final normalized weights for each model *k*, and for each period *S+h*, where *h=1,2,…H*.

$$\hat{Q}^n_{t,k} = \begin{pmatrix} \hat{Q}^n_{S+1,1} & \hat{Q}^n_{S+1,2} & \cdots & \hat{Q}^n_{S+1,K} \\ \hat{Q}^n_{S+2,1} & \hat{Q}^n_{S+2,2} & \cdots & \hat{Q}^n_{S+2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{Q}^n_{S+H,1} & \hat{Q}^n_{S+H,2} & \cdots & \hat{Q}^n_{S+H,K} \end{pmatrix}_{HxK} \tag{A.7}$$

Where $\hat{Q}^n_{S+h,k}$ represents model *k* renormalized average weight to forecast *S+h*, for *k=1,2,…K*, and *h=1,2,…H*. Those weights are understood as the importance of each model in predicting *S+h*. As before, it follows that the summation of weights along the same row is equal to one, $\sum_{k=1}^{K} \hat{Q}^n_{S+h,k} = 1$.

STEP 5. EFFICIENT FORECAST PATH

The final step in the algorithm is to generate out-of-sample forecasts for *h* periods ahead for each model *k*, and to weight each forecast by the normalized weight matrix $\hat{Q}^n_{t,k}$, expression (A.7). Let's express these out-of-sample forecasts by the *H x K* matrix $\hat{\boldsymbol{y}}_{k,t}$, as depicted in expression (A.8).

$$\hat{\boldsymbol{y}}_{t,k} = \begin{pmatrix} \hat{y}_{N+1,1} & \hat{y}_{N+1,2} & \cdots & \hat{y}_{N+1,K} \\ \hat{y}_{N+2,1} & \hat{y}_{N+2,2} & \cdots & \hat{y}_{N+2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{N+H,1} & \hat{y}_{N+H,2} & \cdots & \hat{y}_{N+H,K} \end{pmatrix}_{HxK} \tag{A.8}$$

Where $\hat{y}_{N+h,k}$ represents variable y_t's forecast generated by model *k* for period *N+h*. Now, by multiplying each element of (A.8) by its corresponding element of (A.7), the following expression is obtained:

$$\hat{Q}^n_{t,k}\hat{\boldsymbol{y}}_{t,k} = \begin{pmatrix} \hat{Q}^n_{S+1,1}\hat{y}_{N+1,1} & \hat{Q}^n_{S+1,2}\hat{y}_{N+1,2} & \cdots & \hat{Q}^n_{S+1,K}\hat{y}_{N+1,K} \\ \hat{Q}^n_{S+2,1}\hat{y}_{N+2,1} & \hat{Q}^n_{S+2,2}\hat{y}_{N+2,2} & \cdots & \hat{Q}^n_{S+2,K}\hat{y}_{N+2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{Q}^n_{S+H,1}\hat{y}_{N+H,1} & \hat{Q}^n_{S+H,2}\hat{y}_{N+H,2} & \cdots & \hat{Q}^n_{S+H,K}\hat{y}_{N+H,K} \end{pmatrix}_{HxK} \quad (A.9)$$

Each column in (A.9) represents the weighted out-of-sample forecast of model *k* for each period *N+h*, for *h=1,2,…H*. Normalized weights $\hat{Q}^n_{S+h,k}$ are understood as model *k*'s forecast efficiency of pseudo out-of-sample predictions of y_t; their value falls between zero and one, and their summation overall *k* models is equal to one. Therefore, each element of (A.9), say $\hat{Q}^n_{S+h,k}\hat{y}_{N+h,k}$, can also be interpreted as the contribution of model *k* to y_t's forecast of period $N+h$. Hence, the summation of each row (or period *N+h*) of matrix $\hat{Q}^n_{t,k}\hat{\boldsymbol{y}}_{t,k}$ overall columns (or *k* models), is equivalent to a single forecast for y_t. In other words, such a summation represents the total weighted contribution of all *K* models to y_t forecast in period *N+h*., as represented in (A.10).

$$\hat{\boldsymbol{y}}_t = \begin{pmatrix} \hat{Q}^n_{S+1,1}\hat{y}_{N+1,1} + \hat{Q}^n_{S+1,2}\hat{y}_{N+1,2} + \cdots + \hat{Q}^n_{S+1,K}\hat{y}_{N+1,K} \\ \hat{Q}^n_{S+2,1}\hat{y}_{N+2,1} + \hat{Q}^n_{S+2,2}\hat{y}_{N+2,2} + \cdots + \hat{Q}^n_{S+2,K}\hat{y}_{N+2,K} \\ \vdots \\ \hat{Q}^n_{S+H,1}\hat{y}_{N+H,1} + \hat{Q}^n_{S+H,2}\hat{y}_{N+H,2} + \cdots + \hat{Q}^n_{S+H,K}\hat{y}_{N+H,K} \end{pmatrix} = \begin{pmatrix} \hat{y}_{N+1} \\ \hat{y}_{N+2} \\ \vdots \\ \hat{y}_{N+H} \end{pmatrix}$$

(A.10)

Where vector $\hat{\boldsymbol{y}}_t$ represents the NICA's estimate of y_t for period *N+h*, for *h=1,2,…H.*

Finally, it is important to mention that NICA's algorithm presented in this section was programmed in Eviews to produce our empirical outcome. Those programs are available upon request for further replication of our results. Additional suggestions from the authors can also be provided if NICA is to be implemented as a complementary tool at the Central Bank of Nicaragua.