Forecasting Chilean Inflation with a new method

Javier Finot

May 2017

Abstract In this paper we present a new method called concatenated forecast to build forecast combining the law of movement of one forecast with the forecast of another, we show a theoretical analysis about when gains by lowering the MSPE could be achieved, then we present an empirical example of the method by forecasting inflation in Chile concatenating the Economic Expectation Survey with the law of movement of an SARIMA forecast showing that this method could improve the forecast and beat an univariate benchmark in a pseudo-out-of-sample scheme.

.

JEL Codes:

Keywords:

I. Introduction

In an environment of uncertainty, it is common that multiple information sets and forecast are available for distinct individuals. The exist of distinct forecast reflex not only de different information set available for distinct forecaster, it also reflexes that there is not a unique way to use the information even when two equally trained forecasters have the same information set. In this context is that the combination of forecast arises. The gains of doing so could arise from many ways like diversification, different reactions of the forecasts to structural brakes, misspecification bias that could affect individual models, etc. these benefits were pointed by several authors and its well documented in Timmermann (2006).

Nevertheless, empirical results show that the discussion about how to combine is far from over, especially because optimal weights are often beaten against the alternative of equally weighted forecast in a pseudo out of sample test (Stock & Watson (2004)). This suggest that information could be use in a different way and improve somehow the forecast.

In this work, we present a new method to combinate forecasts that we call "concatenation", the main idea is that one can take advantage of the accurateness of one forecast in some horizon of time and combine it with the movement law of a forecast function from another forecast. This idea was motivated by two reasons. The first is that like Faust and Wright (2013) point out, in some context judgmental forecast are remarkably hard to beat, this and the fact that judgmental forecast are not always available for all horizon suggest that another form to combine this information could be useful. The second reason is that univariate forecast with drift has proven useful in the short term but does not in the long term because the tendency added by the constant term and, by other hand, drift less univariate forecast perform well in long term but may not be the better for the short terms, in this context one can ask if the accurateness of the first could improve the performance of the second in the long term by providing better forecast that feed the forecast function in the first horizons of time.

The rest of the paper is organized as follows. In section 2 we present a brief literature review. Section 3 contains a simple mathematical development to present when concatenation could be useful. Section 4 describe the data used in our empirical example. Section 5 display an in-sample analysis about the data. Section 6 describe our prediction evaluation strategy. Section 7 present the empirical results. Finally, section 8 concludes.

II. Literature Review

As far as we know, we are the first to formally purpose this method as an alternative in combination of forecast. So, there is no previous works about this method. Nevertheless, there are results in topics close to the method that motivate this alternative, in this literature review we want to show some of the results that motivates the theoretical and the empirical result. It is important to say that we keep some distance in the traditional combination scheme because this work is not focused on deepen the discussion about combination forecast but to open a new alternative to do so.

Since Bates and Granger (1969) do his seminal work about combination of forecast there are several papers showing the gains in accurateness in out of sample schemes by combinate forecast, we stand in the work of Timmermann (2006) that provide the most recent summarize in previous literature and a good point to start in the actual state of the art about these methods. In his work, he points that (among other things) the empirical evidence show that simple combination schemes are hard to beat and combination often dominates the best individual forecast in out-of-sample forecasting experiments. Also, research like Stock & Watson (2004) and Elliot and Timmermann (2005) points that optimal weights are often beaten against equally weighted forecast and Pincheira (2012) shows that averaging forecast could produce inefficiencies. Our interpretation about these results is that combine is a good option to forecast but there is not a unique way to combine information and improvements can still be made by deepen in the search of good weights or broaden the discussion in to new methods.

To forecast inflation in Chile, and provide an empirical example of the method, we stand in one of the conclusions of Faust and Wright (2013) that judgmental forecasts are remarkably hard to beat and use the of economics expectations survey (EES) taken by the Central bank and concatenate with a univariate model that performs well in an in-sample scheme. The model was chosen by taking in account the work of Pincheira & García (2012), that pointed that specific members of the SARIMA family perform well to forecast inflation in Chile and that the EES has a remarkably performance in forecast inflation and following the Box & Jenkins (1976) strategy

III. Simple mathematical development

In this section, we develop a simple example for a concatenation using a h-step-ahead forecast, namely $\pi_t^{fA}(h)$, with another forecast for the same process, namely $\pi_t^{fB}(h)$, to show under which conditions are possible the gains in accuracy for this method.

To do so let

$$\pi_t^{fA}(h) = \rho_1 * \pi_t^{fA}(h-1)$$

Be a AR(1) h-step ahead forecast for π_t using the information available in t.

So, the first step ahead forecast for the AR(1) is

$$\pi_t^{fA}(1) = \rho_1 * \pi_t^{fA}(0)$$

$$\pi_t^{fA}(1) = \rho_1 * \pi_t$$

The second step ahead forecast is

$$\pi_t^{fA}(2) = \rho_1 * \pi_t^{fA}(1)$$

$$\pi_t^{fA}(2) = \rho_1^2 * \pi_t$$

And the h-step ahead forecast is

$$\pi_t^{fA}(h) = \rho_1^h * \pi_t$$

Then the forecast error for the one-step-ahead forecast is

$$e_t^A(1) = \pi_{t+1} - \pi_t^{fA}(1)$$

$$e_t^A(1) = \pi_{t+1} - \rho_1 * \pi_t$$

For the two-step-ahead forecast the forecast error is

$$e_t^A(2) = \pi_{t+2} - \pi_t^{fA}(2)$$

$$e_t^A(2) = \pi_{t+2} - \rho_1^2 * \pi_t$$

$$e_t^A(2) = \pi_{t+2} - \rho * \pi_{t+1} + \rho_1 * \pi_{t+1} - \rho_1^2 * \pi_t$$

$$e_t^A(2) = e_{t+1}^A(1) + \rho_1 * e_t^A(1)$$

And for the h-step-ahead forecast, the forecast error is

$$e_t^A(h) = \sum_{i=1}^{h} \rho^i \, e_{t+h-i}^A(1)$$

Analogously, let

$$\pi_t^{fB}(h)$$

Be another (e.g. survey from experts) h-step ahead forecast for π_t using the information available in t.

The one-step-ahead forecast is

$$\pi_t^{fB}(1)$$

The second step ahead forecast is

$$\pi_t^{fB}(2)$$

And the h-step-ahead is

$$\pi_t^{fB}(h)$$

Then the forecast error in the one-step-ahead is

$$e_t^B(1) = \pi_{t+1} - \pi_t^{fB}(1)$$

For the second is

$$e_t^B(2) = \pi_{t+2} - \pi_t^{fB}(2)$$

And for the h-step ahead is

$$e_t^B(h) = \pi_{t+h} - \pi_t^{fB}(h)$$

Now we can define the concatenate method of forecast by

$$\pi_t^{fC}(h)$$

Where

$$\pi_t^{fC}(h) = \pi_t^{fB}(h)$$

If the concatenation is with the forecast method B in the h-step-ahead.

And

$$\pi_t^{fC}(h) = \rho_1 * \pi_t^{fC}(h-1)$$

Otherwise, where ρ_1 is the same parameter as in the forecast method A.

So, if the concatenation of the forecast B with A is only in the first forecast then

$$\pi_t^{fC}(1) = \pi_t^{fB}(1)$$

and

$$\pi_t^{fC}(h) = \rho_1 * \pi_t^{fC}(h-1)$$

For all $h \geq 2$.

Then the forecast error for the one-step-ahead is

$$e_t^C(1) = \pi_{t+1} - \pi_t^{fC}(1)$$

$$e_t^C(1) = \pi_{t+1} - \pi_t^{fB}(1)$$

For the two-step-ahead is

$$e_t^C(2) = \pi_{t+2} - \pi_t^{fC}(2)$$

$$e_t^C(2) = \pi_{t+2} - \rho_1 * \pi_t^{fB}(1)$$

$$e_t^C(2) = \pi_{t+2} - \rho_1 \pi_t^{fA}(1) + \rho_1 \pi_t^{fA}(1) - \rho_1 * \pi_t^{fB}(1)$$

$$e_t^C(2) = e_t^A(2) + \rho_1(\pi_t^{fA}(1) - \pi_t^{fB}(1))$$

$$e_t^C(2) = e_t^A(2) + \rho_1(\pi_t^{fA}(1) - \pi_{t+1} + \pi_{t+1} - \pi_t^{fB}(1))$$

$$e_t^C(2) = e_t^A(2) + \rho_1(e_t^B(1) - e_t^A(1))$$

And using the fact that

$$e_t^A(2) = e_{t+1}^A(1) + \rho_1 * e_t^A(1)$$

In the las equation

$$e_t^C(2) = e_{t+1}^A(1) + \rho_1 * e_t^A(1) + \rho_1(e_t^B(1) - e_t^A(1))$$

$$e_t^C(2) = e_{t+1}^A(1) + \rho_1 * e_t^B(1)$$

Analogously

$$e_t^C(h) = \sum_{i=0}^{h-2} \rho_1^i * e_{t+h-1+i} + \rho_1^{h-1} * e_t^B(1)$$

Then, gains from concatenation exist if

$$E\left(e_t^A(h)\right)^2 - E\left(e_t^C(h)\right)^2 > 0$$

If h=1 the last expression is

$$E\left(e_t^A(1)\right)^2 - E\left(e_t^C(1)\right)^2 > 0$$

Which is the same as

$$E\left(e_t^A(1)\right)^2 - E\left(e_t^B(1)\right)^2 > 0$$

So, for the first step ahead, the concatenation method show gains only if the Mean Square Prediction Error is lower for the method B.

For the two-step-ahead forecast, gains for concatenation exist if

$$E\left(e_t^A(2)\right)^2 - E\left(e_t^C(2)\right)^2 > 0$$

Replacing this last for the expression for $e_t^A(2)$ and $e_t^C(2)$,

$$E\left(e_{t+1}^A(1) + \rho_1 * e_t^A(1)\right)^2 - E\left(e_{t+1}^A(1) + \rho_1 * e_t^B(1)\right)^2 > 0$$

And using simple algebra and rearranging terms this can be expressed as

$$E(\rho_1^2 * (e_t^A(1)^2 - e_t^B(1)^2)) + 2 * E\left(\rho_1 * e_{t+1}^A(1) * \left(e_t^A(1) - e_t^B(1)\right)\right) > 0$$

This expression can be interpreted as the one-step-ahead differential between MSPE plus a correlation between forecast error for the one-step-ahead forecast A in t+1 with the differential of forecast errors one-step-ahead between the two models in t.

This expression could be positive for several reasons, for example, if we use a more accurate forecast to concatenate the forecast A, then the statement before only depends in the correlations to be true, so if the correlation between the error in t+1 for the method A with the differential of forecast errors in t is less than the differential of MSPE between methods, then gain in accuracy is possible. Also, the last part of the expression show that even a concatenation with a method with lower accuracy could generate gains, the necessary condition is that the correlation be positive and greater than the differential of MSPE between the two methods.

Generalizing the last expression for horizon h (proof in the appendix), we get

$$E\left(\rho_1^{2h-2} * \left(e_t^A(1)^2 - e_t^B(1)^2\right)\right) + 2E\left(\sum_{i=1}^{h-1} \rho_1^{h+i-2} * e_{t+h-i}^A(1) * \left(e_t^A(1) - e_t^B(1)\right)\right) > 0$$

Analyzing this expression, we can see that the gain driven by the difference in MSPE are less important for long horizons and the result depend heavily in the correlation explain before. This implies that the gains of using a more accurate forecast could be less important if we want to forecast longer horizons and its far more relevant to choose a method with the last part of the expression positive.

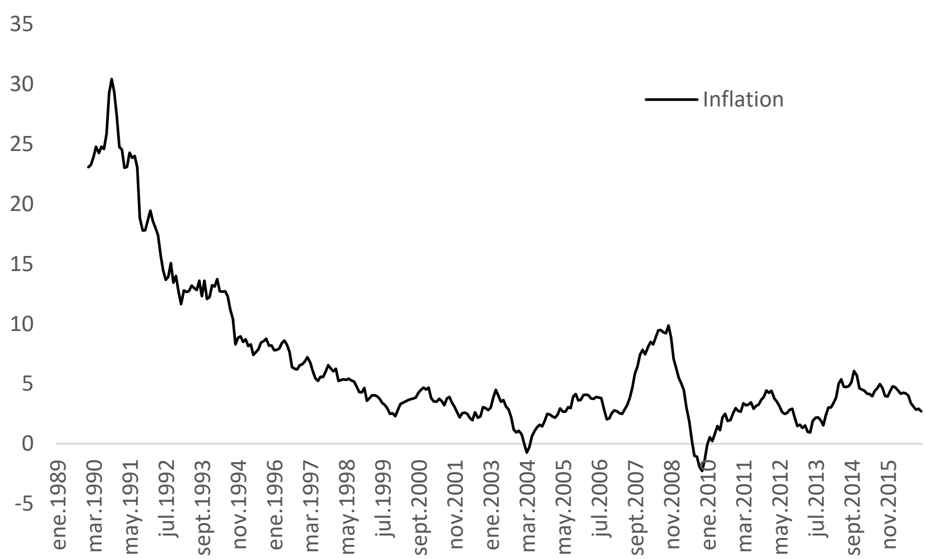
IV. Data

We use monthly data obtained from the Central Bank of Chile to obtain time series on inflation and time series from the Economic Expectation Survey (EES). For inflation, we construct the index using year-on-year variations of two-historic series for Consumer Price Index (CPI) with base period 2008 and 2013. These series have been spliced to get a final series with observations from January of 1990 to December of 2016.

The EES is a survey made by the Central Bank of Chile month by month since February of 2000 to the present, his objective is asking to a group of 50 professionals (that includes academics, manager and consultant for the private sector) his forecast about several macroeconomics variables including inflation. Until 2010, since answer the survey is voluntary, the mean of answers each month was 35, his max was 50, the min was 20, also the group is renewal each year and the survey need not to be answer completely, that generate that the answers for immediate months were more abundant than the answers for several months later, an introductory note to this survey are summarized in Pedersen (2010). From this survey, we use 2 series for different time horizons, the first forecast inflation one month ahead and the second eleven months ahead. The first is in monthly variations so we transform it to annual variations. The second is already in yearly variation so no transformation is needed. It is important to mention that the forecast in 11 months are released before the actual data on inflation is released, so it's 12-steps-ahead if we count the time since the last data is available (i.e. if the announced by the EES for 11 month it is on February, then the realization for inflation in that month it is not available yet and the forecast is 12-step-ahead from the last data available on January).

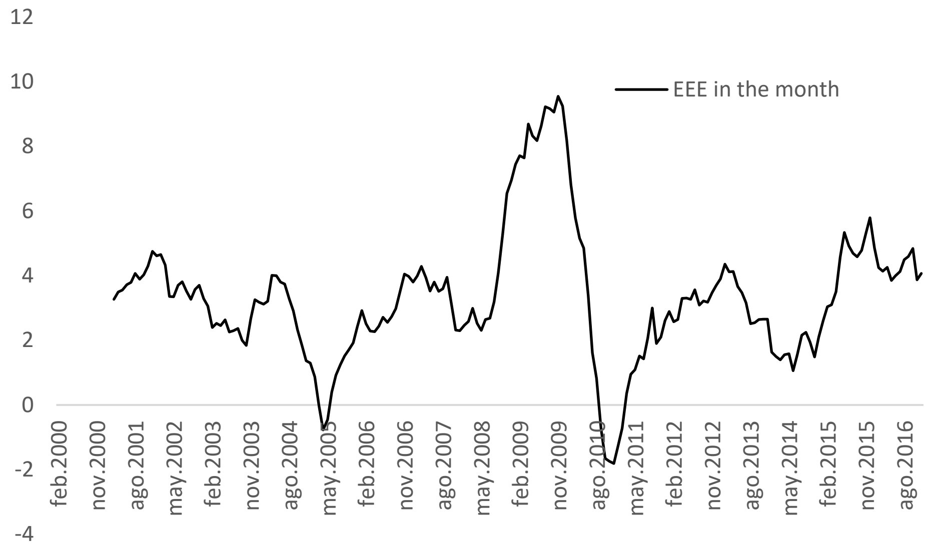
The graphs below show the time series from the year-on-year variation and the EES for the month and eleven months.

Graph x: Year on year variation of the CPI.



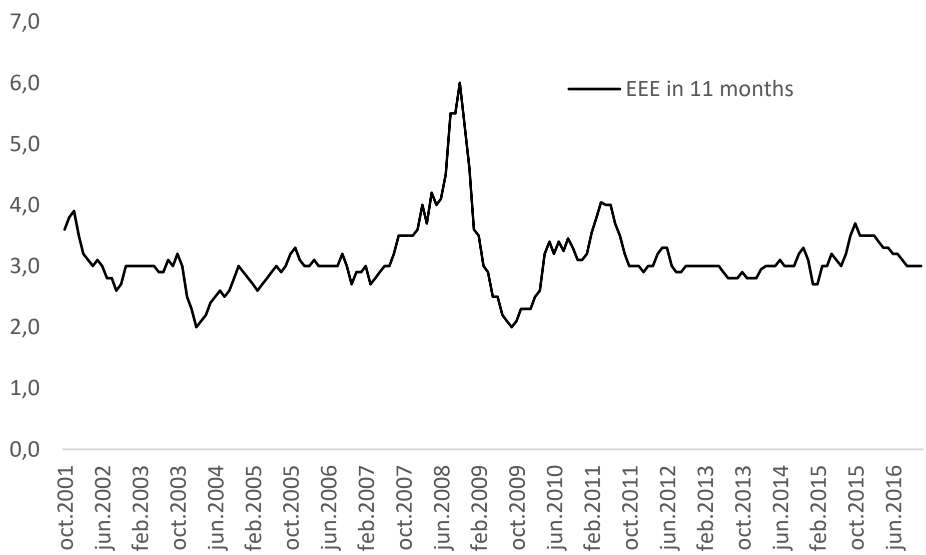
Source: Own calculations based on Central Bank of Chile data.

Graph x: Year on year variation of the EEE forecast for the inflation one month ahead.



Source: Own calculations based on Central Bank of Chile data.

Graph x: EEE forecast for the inflation eleven months ahead.



Source: Own calculations based on Central Bank of Chile data.

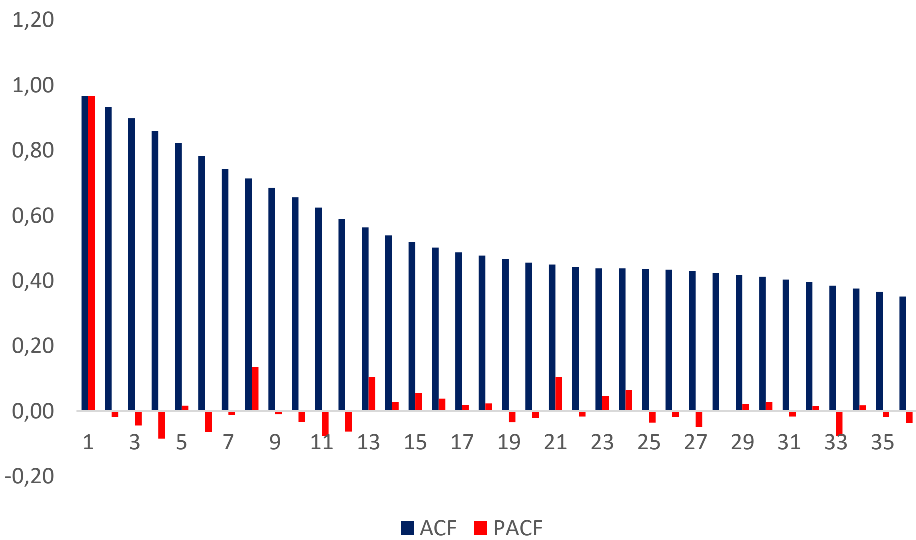
Our pseudo-out-of-sample analysis uses a rolling window approach, to maximize the use of the EES we use an initial estimation window length of 109 observations such that our first one-step-ahead forecast is made for February 2000 (the date of the first EES available), while the last one is made for December 2016. Additional to this forecast we construct 3 types of Concatenation, the first is using the EES for the month, the second and third use the EES for both the month and 11 months and, in order to construct a fair comparison because the EES for 11 months start in September 2001, we change the window length to start with 128 observations, the first one-step-ahead forecast is made for September 2001 and the last is made for December 2016. All the equations were estimated with HAC Standard Errors using Newey West (1987, 1994).

V. In-Sample analysis

For the in-sample analysis we use the Box & Jenkins (1976) strategy to get possible functions to forecast inflation, we choose the candidate with lower Akaike to perfom the out-of-sample part.

In the next graph, we show the Autocorrelation Function (ACF) for the process and the Partial Autocorrelation Function (PACF).

Graph x: ACF and PACF for the Year on Year inflation in Chile.



Source: Own calculations based on Central Bank of Chile data.

The possible methods to forecast the inflation are presented in the next table.

Table n: Possible methods to forecast inflation.

N° Method	Representation
1:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_{12} * \pi_t^f(h-12)$
2:	$\Delta\pi_t^f(h) = \rho_1 * \Delta\pi_t^f(h-1) + \theta_1 * \epsilon_{t+h-1} + \theta_{12} * \epsilon_{t+h-12}$
3:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_2 * \pi_t^f(h-2) + \theta_1 * \epsilon_{t+h-1} + \theta_{12} * \epsilon_{t+h-12}$
4:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_2 * \pi_t^f(h-2) + \theta_{12} * \epsilon_{t+h-12}$
5:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_2 * \pi_t^f(h-2)$
6:	$\Delta\pi_t^f(h) = \rho_1 * \Delta\pi_t^f(h-1) + \rho_{12} * \Delta\pi_t^f(h-12)$

7:	$\Delta\pi_t^f(h) = \rho_1 * \Delta\pi_t^f(h-1) + \rho_{12} * \Delta\pi_t^f(h-12) + \theta_1 * \epsilon_{t+h-1} + \theta_{12} * \epsilon_{t+h-12} + \theta_{13} * \epsilon_{t+h-13}$
8:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_2 * \pi_t^f(h-2) + \theta_{12} * \epsilon_{t+h-12} + \theta_{13} * \epsilon_{t+h-13}$

For these methods, we present the results for estimation, to provide a fair comparison we use the same amount of data for every possible method. The results are showed in table XX.

Table n: In-Sample Analysis 8 possible models.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
constant	0.11*		0.06	0.06***	0.1*			0.06***
AR(1)	1.01***	0.54***	1.28	1.27***	1.17***	0.23***	0.97***	1.27***
AR(2)			-0.30	-0.29***	-0.19**			-0.29***
SAR(12)	-0.04**					-0.47***		
MA(1)		0	0				-0.88***	
SMA(12)		-0.97***	-0.97***	-0.97***			-0.97***	-0.97***
SMA(13)							0.86***	0
R-squared	0.99	0.50	0.99	0.99	0.99	0.32	0.54	0.99
N	311	311	311	311	311	311	311	311
Akaike Criterion	1.80	1.21	1.02	1.01	1.79	1.51	1.12	1.02
Schwarz Criterion	1.84	1.25	1.08	1.06	1.83	1.53	1.17	1.08
Durbin-Watson	1.64	1.23	2.04	2.04	1.98	1.52	1.14	2.04

Source: Own calculations based on Central Bank of Chile data. * p<10%, ** p<5%, *** p<1%. HAC standard errors according to Newey & West (1987) in parentheses. Each parameter corresponds to his analog of table n.

The model we use in the out of sample analysis is the model 1 and 3. The first one is chosen because it is similar to our mathematical analysis (and the SAR term allow to concatenate again in this period with the EEE.) and the model 3 was chosen on basis of Akaike criterion.

VI. Predictive Evaluation Strategy

In our baseline exercise, we evaluate the out of sample predictive ability of the standard models against their concatenated versions. we construct 3 types of concatenated forecasts. The first is based on a concatenation in the first-step-ahead with the second forecast method (i.e. Economic surveys), so the forecast is given by

$$\pi_t^f(1) = \pi_t^E(1)$$

$$\pi_t^f(2) = \alpha + \rho_1 * \pi_t^f(1) + \rho_{12} * \pi_{t-10}$$

$$\vdots$$

$$\pi_t^f(12) = \alpha + \rho_1 * \pi_t^f(11) + \rho_{12} * \pi_t$$

$$\pi_t^f(13) = \alpha + \rho_1 * \pi_t^f(12) + \rho_{12} * \pi_t^f(1)$$

$$\vdots$$

$$\pi_t^f(24) = \alpha + \rho_1 * \pi_t^f(23) + \rho_{12} * \pi_t^f(12)$$

where $\pi_t^f(h)$ it's the h-step-ahead forecast using information available at time t and $\pi_t^E(h)$ it's the h-step-ahead forecast with the second method using information available at time t.

The second forecast uses not only the first-step-ahead but another exogenous forecast for the 12-steps-ahead concatenation with the second forecast method. So, we get

$$\pi_t^f(1) = \pi_t^E(1)$$

$$\pi_t^f(2) = \alpha + \rho_1 * \pi_t^f(1) + \rho_{12} * \pi_{t-10}$$

$$\vdots$$

$$\pi_t^f(12) = \alpha + \rho_1 * \pi_t^f(11) + \rho_{12} * \pi_t$$

$$\pi_t^f(13) = \alpha + \rho_1 * \pi_t^E(12) + \rho_{12} * \pi_t^f(1)$$

$$\vdots$$

$$\pi_t^f(24) = \alpha + \rho_1 * \pi_t^f(23) + \rho_{12} * \pi_t^f(12)$$

And the third forecast, as the second, uses the first and the 12-steps ahead from the second method, but the difference between both is that in the third method we use the 12-step-ahead prediction in the same way that we use the first one. So

$$\pi_t^f(1) = \pi_t^E(1)$$

$$\pi_t^f(2) = \alpha + \rho_1 * \pi_t^f(1) + \rho_{12} * \pi_{t-10}$$

$$\vdots$$

$$\pi_t^f(12) = \pi_t^E(12)$$

$$\pi_t^f(13) = \alpha + \rho_1 * \pi_t^f(12) + \rho_{12} * \pi_t^f(1)$$

$$\vdots$$

$$\pi_t^f(24) = \alpha + \rho_1 * \pi_t^f(23) + \rho_{12} * \pi_t^f(12)$$

To evaluate the predictive ability, we perform the test attributed to Diebold & Mariano (1995) and West (1996) (hereafter DMW) to do a horse race between the standard model and the concatenated version. To do so we split the database and estimate the models in rolling windows of fixed size, then generate the first h-step-ahead forecast and get the sample RMSPE (SRMSPE) and the sample MSPE.

To describe this test, let us suppose we have a data set of $\pi_t$with T+1 observations and that we want to determine if the model 2 has better forecasting performance than model 1. we can use the first R observations of the data set to estimate both models. Then, with this estimation, we can do a multi-step ahead forecast from R+1 to R+24 and obtain the forecast error by doing:

$$\pi_R^f(h) - \pi_{R+h} = e_R(h)$$

Where $h = 1, \ldots, 24$, and as before $\pi_R^f(h)$ represent the h-step-ahead forecast using information available in time R and π_{R+h} is the actual value of inflation in the period R+h.

Next we use the observations between the second through the R+1 and repeat the process by continuing with this fixed window scheme, reach the last observations between T+1-R and T+1. By doing so, we generate a total of $P(h)$ forecast for each horizon with $P(h)$ satisfying $R + (P(h) - 1) + h = T + 1$. So

$$P(h) = T + 2 - h - R$$

To measure the forecast accuracy with the RMSPE we use the sample analog of this population moment. This is:

$$SRMSPE = \sqrt{\frac{1}{P(h)} \sum_{t=R}^{T+1-h} e_R(h)^2}$$

For the DMW test, we use the difference in sample MSPE for the models $i = 1,2$. This is

$$\bar{d} = SMSPE_i - SMSPE_j$$

with

$$SMSPE_i = \frac{1}{P(h)} \sum_{t=R}^{T+1-h} e_{Ri}(h)^2$$

Then the DMW t-statistic test is:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}^*}}$$

Where $\hat{V}^*$is an estimator from the long run variance of $\bar{d}$.

And the null hypothesis is

H_0: $MSPE_1 - MSPE_2 \leq 0$

Against the alternative

H_A: $MSPE_1 - MSPE_2 > 0$

As a practical matter, a simple way to proceed is to regress d_i on a constant and use a t-test with Newey and West (1987, 1994) Standard Errors to determine whether the constant is statistically positive.

VII. Empirical Results

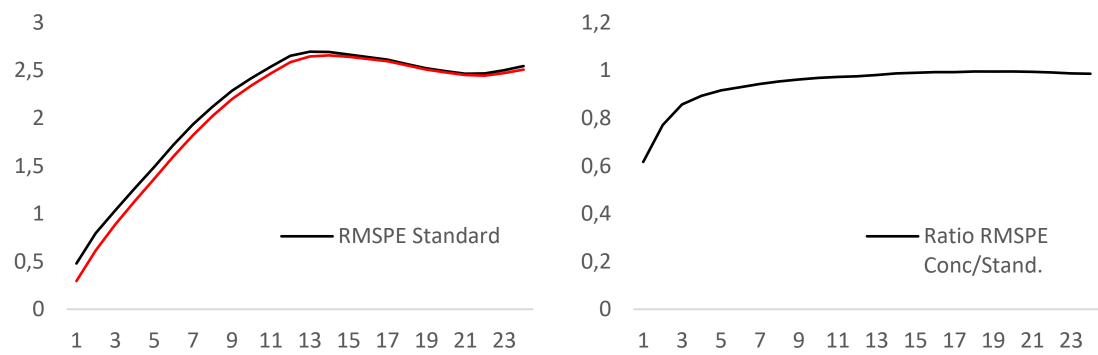
In this section, we present the result from the DMW test for two forecast models, each model has been concatenated with the Survey of Economic expectation using the 3 different types of concatenation presented in the previous section.

The forecast models were chosen using the Box & Jenkins (1976) strategy, and the DMW test was used to compare forecast accuracy off the two prediction models (standard models) and three possible concatenations for these models

Table n: Standard Forecast models

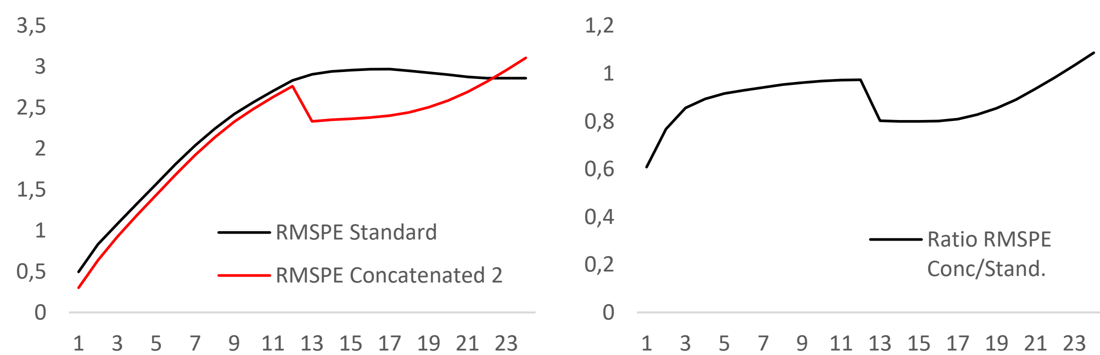
1:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_{12} * \pi_t^f(h-12)$
2:	$\pi_t^f(h) = \alpha + \rho_1 * \pi_t^f(h-1) + \rho_2 * \pi_t^f(h-2) + \theta_1 * \epsilon_{t+h-1} + \theta_{12} * \epsilon_{t+h-12}$

Graph x: RMSPE and the ratio from forecast model 1 and first concatenated forecast.



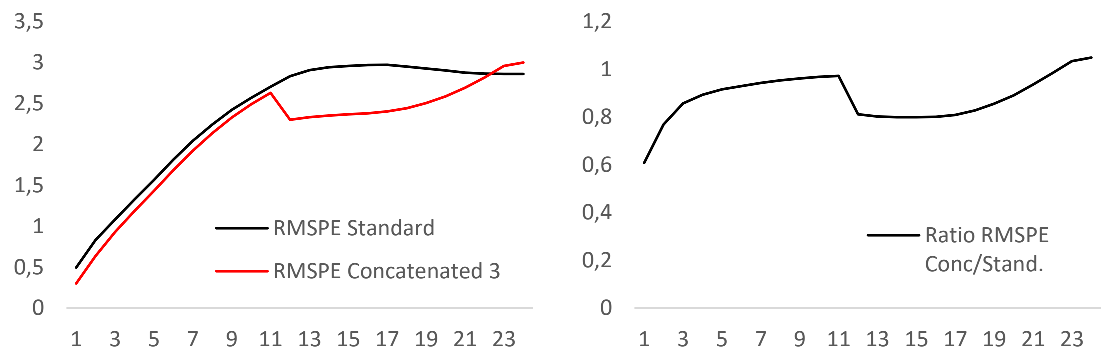
Source: Own calculations based on Central Bank of Chile data.

Graph x: RMSPE and the ratio from forecast model 1 and second concatenated forecast.



Source: Own calculations based on Central Bank of Chile data.

Graph x: RMSPE and the ratio from forecast model 1 and third concatenated forecast.



Source: Own calculations based on Central Bank of Chile data.

Table n: DMW test for compare accuracy in forecast for prediction model 1, horizons 1 to 6.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DMW and RMSPE							
		Forecast					
Model	DMW/RMSPE	h=1	h=2	h=3	h=4	h=5	h=6
Model 1	DMW	0,14***	0,26***	0,28***	0,32***	0,36***	0,4***
	RMSPE	0,30	0,62	0,89	1,13	1,37	1,60
	RMSPE benchmark	0,48	0,80	1,04	1,27	1,49	1,72
Model 2	DMW	0,15***	0,28***	0,31***	0,36***	0,39***	0,45***
	RMSPE	0,30	0,64	0,93	1,19	1,43	1,68
	RMSPE benchmark	0,50	0,83	1,08	1,33	1,57	1,81
Model 3	DMW	0,15***	0,28***	0,31***	0,36***	0,39***	0,45***
	RMSPE	0,30	0,64	0,93	1,19	1,43	1,68
	RMSPE benchmark	0,50	0,83	1,08	1,33	1,57	1,81

Source: Own calculations based on Central Bank of Chile data. * p<10%, ** p<5%, *** p<1%. HAC standard errors according to Newey and West (1987) in parentheses.

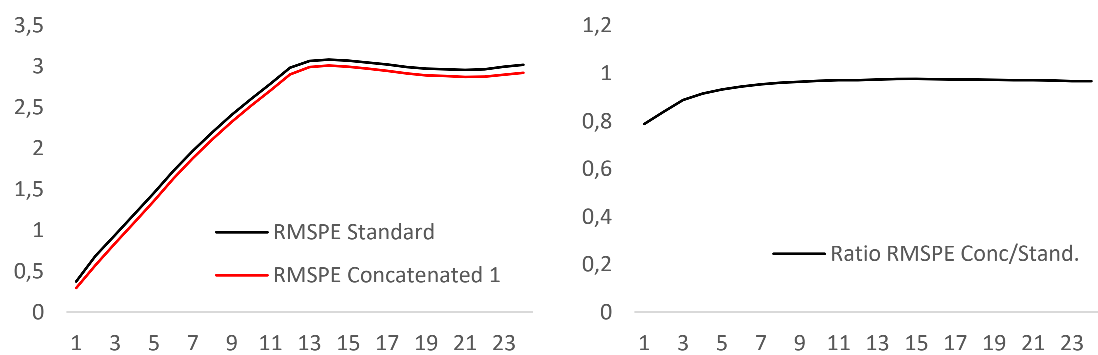
Table n: DMW test for compare accuracy in forecast for prediction model 1, horizons 12 to 17.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DMW and RMSPE							
Forecast							
Model	DMW/RMSPE	h=12	h=13	h=14	h=15	h=16	h=17
Model 1	DMW	0,35**	0,27**	0,19*	0,13	0,11	0,09
	RMSPE	2,58	2,65	2,66	2,64	2,62	2,60

	RMSPE benchmark	2,65	2,70	2,69	2,67	2,64	2,61
Model 2	DMW	0,4**	3,01**	3,12***	3,16**	3,15**	3,05**
	RMSPE	2,76	2,33	2,35	2,37	2,38	2,40
	RMSPE benchmark	2,83	2,91	2,94	2,96	2,97	2,97
Model 3	DMW	2,74**	3,01**	3,12***	3,16**	3,15**	3,05**
	RMSPE	2,30	2,33	2,35	2,37	2,38	2,40
	RMSPE benchmark	2,83	2,91	2,94	2,96	2,97	2,97

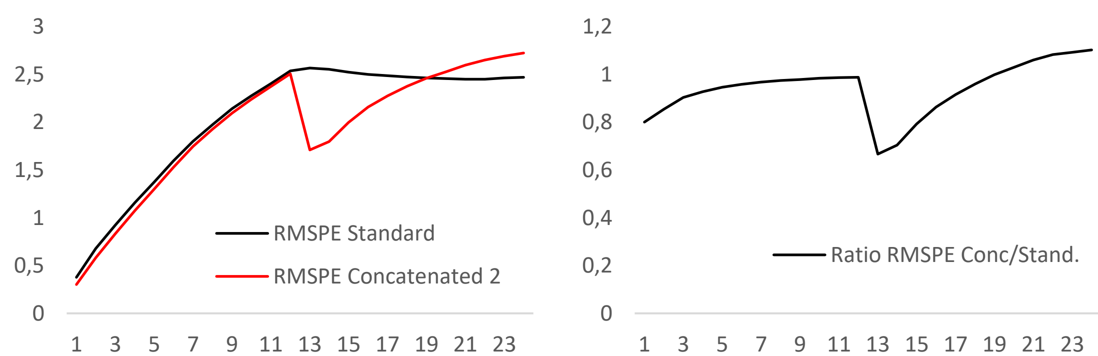
Source: Own calculations based on Central Bank of Chile data. * p<10%, ** p<5%, *** p<1%. HAC standard errors according to Newey and West (1987) in parentheses.

Graph x: RMSPE and the ratio from forecast model 2 and first concatenated forecast.



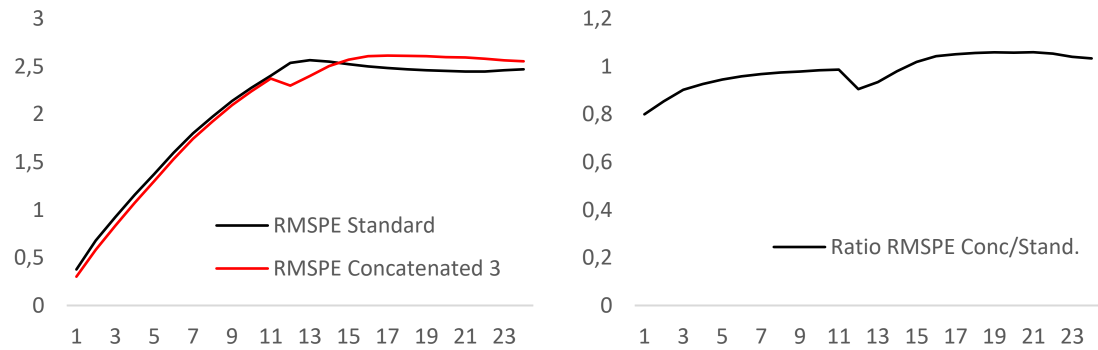
Source: Own calculations based on Central Bank of Chile data.

Graph x: RMSPE and the ratio from forecast model 2 and second concatenated forecast.



Source: Own calculations based on Central Bank of Chile data.

Graph x: RMSPE and the ratio from forecast model 2 and third concatenated forecast.



Source: Own calculations based on Central Bank of Chile data.

Table n: DMW test for compare accuracy in forecast for prediction model 2, horizons 1 to 6.

(1) DMW and RMSPE	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Forecast							
Model	DMW/RMSPE	h=1	h=2	h=3	h=4	h=5	h=6
Model 1	DMW	0,05***	0,14***	0,19***	0,23***	0,28***	0,32***
	RMSPE	0,30	0,62	0,89	1,13	1,37	1,60
	RMSPE benchmark	0,48	0,80	1,04	1,27	1,49	1,72
Model 2	DMW	0,05***	0,12***	0,16***	0,19***	0,2***	0,21***
	RMSPE	0,30	0,64	0,93	1,19	1,43	1,68
	RMSPE benchmark	0,50	0,83	1,08	1,33	1,57	1,81
Model 3	DMW	0,05***	0,12***	0,16***	0,19***	0,2***	0,21***
	RMSPE	0,30	0,64	0,93	1,19	1,43	1,68
	RMSPE benchmark	0,50	0,83	1,08	1,33	1,57	1,81

Source: Own calculations based on Central Bank of Chile data. * p<10%, ** p<5%, *** p<1%. HAC standard errors according to Newey and West (1987) in parentheses.

Table n: DMW test for compare accuracy in forecast for prediction model 2, horizons 12 to 17.

(1) DMW and RMSPE	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Forecast							
Model	DMW/RMSPE	h=12	h=13	h=14	h=15	h=16	h=17
Model 1	DMW	0,5**	0,47*	0,44*	0,44*	0,45*	0,47*
	RMSPE	2,58	2,65	2,66	2,64	2,62	2,60
	RMSPE benchmark	2,65	2,70	2,69	2,67	2,64	2,61

Model 2	DMW	0,15***	3,66***	3,29**	2,37*	1,59	1
	RMSPE	2,76	2,33	2,35	2,37	2,38	2,40
	RMSPE benchmark	2,83	2,91	2,94	2,96	2,97	2,97
Model 3	DMW	1,15**	0,83	0,25	-0,25	-0,55	-0,66
	RMSPE	2,30	2,33	2,35	2,37	2,38	2,40
	RMSPE benchmark	2,83	2,91	2,94	2,96	2,97	2,97

Source: Own calculations based on Central Bank of Chile data. * p<10%, ** p<5%, *** p<1%. HAC standard errors according to Newey and West (1987) in parentheses.

It is worth noting that, for the prediction model 1, the first type of concatenation show gains in accuracy vs his standard model until 7-steps-ahead with 1 percent statistical significance, until 13-steps-ahead with 5 percent and until 14-steps-ahead with 10 percent, and with the second and third types of concatenation show gains in accuracy until 7-steps-ahead and in 14-steps-ahead with 1 percent of tolerance, and until 18-steps-ahead with 5 percent of tolerance, the most important difference between the concatenation 2 and 3 is that for the 12-steps-ahead forecast, the third concatenation show a greater difference between RMSPE with his standard counterpart, but the loss of RMSPE is accompanied with more variance. In summary, by concatenating the firs forecast model it is possible to obtain significance improvement in terms of accuracy measured by DMW test against the same models without the concatenation.

For the prediction model 2, the first type of concatenation show gains in accuracy vs the standard model until 7-steps-ahead with 1 percent of tolerance, until 12-steps-ahead with 5 percent, and until 18-steps-ahead with 10 percent, the second type of concatenation show gains in accuracy until 9-steps-ahead, in 12 and 13-steps-ahead with 1 percent of tolerance and until 14-steps-ahead 5 percent and until 15-steps-ahead with 10 percent, and the third type of concatenation show gains in accuracy until 9-steps-ahead with 1 percent of tolerance and until 12-steps-ahead with 5 percent of tolerance. In summary, by concatenating the second forecast model it is also possible to obtain significance improvement in terms of accuracy measured by DMW test against the same models without the concatenation.

VIII. Concluding Remarks

In this paper, we present a new method for combination of forecast that we called "Concatenation of forecast", this method takes advantage of the law of movement of one forecast and the accurateness of another.

In section 3 we show that gains by concatenating in an out-of-sample scheme are driven by two factors, the first is the existence of a method with better performance and the second are gains by correlations between both methods. The first is important because by concatenation, the accurateness of one forecast could be used to forecast longer horizons even when the method is only available to forecast specific time horizons (e.g. judgmental forecast). The second factor take more importance in long run horizons and shows that gains could be achieved taking advantage of the positive correlations between the difference in forecast errors in horizons when concatenation are made with longer horizon errors of the method with the law of movement, this is important because shows that this method could be useful even when there is not clear winner to make forecast for some horizon.

Also, we presented an empirical example to improve the forecast of inflation in Chile by concatenating the Economic Expectations Survey (EES) and univariate method in two different ways. First, we concatenate the first step ahead forecast of the EES with different univariate forecasts and then we concatenate forecasts in the first and 12 steps ahead with the same univariate methods. We show that gains could be achieved for long horizons in the two options (at least for 12-steps-ahead) and these gains are statistically significant in a pseudo-out-of-sample scheme. By concatenating the one-step-ahead forecast made by EES, the gains are fading away in longer horizons but last longer, by other hand, concatenating both (the first and twelve steps ahead), the gains are greater after the second concatenation, but show a lower performance than the univariate counterpart in the long run.

Our finding is that Concatenation is a good alternative to make forecast of the inflation in Chile. In further research, we hope to found the same gains to forecast other relevant variables in an international context.

References

Newey, W., & West, K. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica, 55*(3), 703-708.

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 253-263.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day.

Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, *1*, 135-196.

Faust, J., & Wright, J. H. (2013). Forecasting inflation. *Handbook of economic forecasting*, *2*(Part A), 3-56.

Pedersen, M. (2010). Una nota introductoria a la Encuesta de Expectativas Económicas. *Estudios Económicos Estadísticos, Banco central de Chile.*

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Or*, 451-468.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, *23*(6), 405-430.

Pincheira, P. M. (2012). Are Forecast Combinations Efficient?.

Pincheira, P., & García, Á. (2012). En busca de un buen marco de referencia predictivo para la inflación en Chile. *El Trimestre Económico*, *79*(313), 85-123.

Hansen, B. E. (2008). Least-squares forecast averaging. *Journal of Econometrics*, *146*(2), 342-350.

Appendix

Prof for the section 3.

$$E\left(\rho_1^{2h-2} * (\, e_t^A(1)^2 - e_t^B(1)^2)\right) + 2E\left(\sum_{i=1}^{h-1} \rho_1^{h+i-2} * e_{t+h-i}^A(1) * (\, e_t^A(1) - e_t^B(1))\right) > 0$$

To solve this expression, we use the iterative method to get the general expression

For h=1

$$E\left(e_t^A(1)\right)^2 - E\left(e_t^C(1)\right)^2 > 0$$

And because we are concatenating for the first-step-ahead this is equal to

$$E\left(e_t^A(1)\right)^2 - E\left(e_t^B(1)\right)^2 > 0$$

For h=2

$$E\left(e_t^A(2)\right)^2 - E\left(e_t^C(2)\right)^2 > 0$$

Replacing this last for the expression for $e_t^A(2)$ and $e_t^C(2)$,

$$E\left(e_{t+1}^A(1) + \rho_1 * e_t^A(1)\right)^2 - E\left(e_{t+1}^A(1) + \rho_1 * e_t^B(1)\right)^2 > 0$$

$$E(e_{t+1}^A(1)^2 + 2 * e_{t+1}^A(1) * \rho_1 * e_t^A(1) + \rho_1^2 * e_t^A(1)^2) \\ - E\left(e_{t+1}^A(1)^2 + 2 * e_{t+1}^A(1) * \rho_1 * e_t^B(1) + \rho_1^2 * e_t^B(1)\right) > 0$$

The factorizing and eliminating repeated terms we get

$$E\left(\rho_1^2 * (e_t^A(1)^2 - e_t^B(1)^2)\right) + 2 * E\left(\rho_1 * e_{t+1}^A(1) * \left(e_t^A(1) - e_t^B(1)\right)\right) > 0$$

For h=3

$$E\left(e_t^A(3)\right)^2 - E\left(e_t^C(3)\right)^2 > 0$$

Replacing

$$E\left(e_{t+2}^A(1) + \rho_1 * e_{t+1}^A(1) + \rho_1^2 * e_t^A(1)\right)^2 - E\left(e_{t+2}^A(1) + \rho_1 * e_{t+1}^A(1) + \rho_1^2 * e_t^B(1)\right)^2 > 0$$

Operating the polynomial we get

$$E\left(e_{t+2}^{A}(1)^2 + \rho_1^2 * e_{t+1}^{A}(1)^2 + \rho_1^4 * e_t^{A}(1)^2 + 2 * e_{t+2}^{A}(1) * \rho_1 * e_{t+1}^{A}(1) + 2 * e_{t+2}^{A}(1) * \rho_1^2 * e_t^{A}(1) + 2 * \rho_1^3 * e_{t+1}^{A}(1) * e_t^{A}(1)\right)^2 - E\left(e_{t+2}^{A}(1)^2 + \rho_1^2 * e_{t+1}^{A}(1)^2 + \rho_1^4 * e_t^{B}(1)^2 + 2 * e_{t+2}^{A}(1) * \rho_1 * e_{t+1}^{A}(1) + 2 * e_{t+2}^{A}(1) * \rho_1^2 * e_t^{B}(1) + 2 * \rho_1^3 * e_{t+1}^{A}(1) * e_t^{B}(1)\right)^2 > 0$$

Then re arranging terms.

$$E\left(\rho_1^4 * (e_t^{A}(1)^2 - e_t^{B}(1)^2)\right) + 2 * E\left(\rho_1^2 * e_{t+2}^{A}(1) * \left(e_t^{A}(1) - e_t^{B}(1)\right) + \rho_1^3 * e_{t+1}^{A}(1) * \left(e_t^{A}(1) - e_t^{B}(1)\right)\right) > 0$$

$$E\left(\rho_1^4 * (e_t^{A}(1)^2 - e_t^{B}(1)^2)\right) + 2 * E\left(\left(\rho_1^2 * e_{t+2}^{A}(1) + \rho_1^3 * e_{t+1}^{A}(1)\right) * \left(e_t^{A}(1) - e_t^{B}(1)\right)\right) > 0$$

Then the solution for any h-step-ahead is:

$$E\left(\rho_1^{2h-2} * (\, e_t^{A}(1)^2 - e_t^{B}(1)^2)\right) + 2E\left(\sum_{i=1}^{h-1} \rho_1^{h+i-2} * e_{t+h-i}^{A}(1) * (\, e_t^{A}(1) - e_t^{B}(1))\right) > 0$$