DOCUMENTOS DE TRABAJO

A Simple Out-of-Sample Test for the Martingale Difference Hypothesis

Pablo Pincheira

N.º 698 Julio 2013

BANCO CENTRAL DE CHILE

DOCUMENTOS DE TRABAJO

A Simple Out-of-Sample Test for the Martingale Difference Hypothesis

Pablo Pincheira

N.º 698 Julio 2013

BANCO CENTRAL DE CHILE

BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

La serie Documentos de Trabajo es una publicación del Banco Central de Chile que divulga los trabajos de investigación económica realizados por profesionales de esta institución o encargados por ella a terceros. El objetivo de la serie es aportar al debate temas relevantes y presentar nuevos enfoques en el análisis de los mismos. La difusión de los Documentos de Trabajo sólo intenta facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de los miembros del Consejo del Banco Central de Chile. Tanto el contenido de los Documentos de Trabajo como también los análisis y conclusiones que de ellos se deriven, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión del Banco Central de Chile o de sus Consejeros.

The Working Papers series of the Central Bank of Chile disseminates economic research conducted by Central Bank staff or third parties under the sponsorship of the Bank. The purpose of the series is to contribute to the discussion of relevant issues and develop new analytical or empirical approaches in their analyses. The only aim of the Working Papers is to disseminate preliminary research for its discussion and comments.

Publication of Working Papers is not subject to previous approval by the members of the Board of the Central Bank. The views and conclusions presented in the papers are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 3882231

A SIMPLE OUT-OF-SAMPLE TEST FOR THE MARTINGALE DIFFERENCE HYPOTHESIS*

Pablo Pincheira
Gerencia de Investigación Económica
Banco Central de Chile

Abstract

We show that a straightforward modification of a trading based test for predictability displays interesting advantages over the Excess Profitability (EP) test proposed by Anatolyev and Gerco when testing the Martingale Difference Hypothesis. Our statistic is called Straightforward Excess Profitability (SEP) and avoids the calculation of a term that under the null of no predictability should be zero, but in practice may be sizable. In addition, our test does not require the strong assumption of independency used to derive the EP test. We claim that dependency is the rule and not the exception. We show via Monte Carlo simulations that the SEP test outperforms the EP test in terms of size and power. Finally, we illustrate the use of these tests in an empirical application within the context of the exchange rate literature.

Resumen

En este artículo mostramos que una modificación directa de un test de predictibilidad basado en reglas de transacción de activos financieros posee ventajas interesantes sobre el Test de Exceso de Utilidades (EU) propuesto por Anatolyev y Gerco cuando se evalúa la hipótesis de una martingala en diferencias. Nuestro estadístico se denomina Exceso de Utilidades Directo (EUD) y evita el cálculo de un término que bajo la hipótesis nula de ausencia de predictibilidad debería ser cero, pero que en la práctica tiene un tamaño considerable. Adicionalmente, nuestro test no requiere los fuertes supuestos de independencia requeridos para derivar el test EU. Nosotros planteamos que la dependencia es la regla y no la excepción. Mostramos vía simulaciones de Monte Carlo que el test EUD tiene un mejor comportamiento que el test EU en términos de error de tipo I y error de tipo II. Finalmente ilustramos el uso de estos tests en una aplicación empírica dentro del contexto de la literatura de tipos de cambio.

* Acknowledgements: I am grateful to Todd Clark for valuable comments and for sharing their database for the empirical application. Alfonso Barrero has provided superb assistance. The paper has also benefited from the opinions expressed at the Forecasting Financial Markets 2013 Conference carried out in Hannover, Germany and from the opinions expressed at the Brown Bag Meetings of the Research Department of the Central Bank of Chile. The views expressed in this paper do not necessarily represent those of the Central Bank of Chile or its Board members. All errors are mine. Address: Agustinas 1180, Santiago, Chile. Phone: (562) 6702874. E-mail: ppinchei@bcentral.cl.

1. Introduction

The objective of this paper is to provide a new out-of-sample test for the Martingale Difference Hypothesis (MDH). This is important because the economic literature has usually used a martingale model as a benchmark to test for predictability. In the context of asset prices, for instance, the martingale model posits that the best forecast of tomorrow's price is today's price. In other words, the martingale model assumes that future expected returns, given information available today, are zero.

We propose the construction of the Straightforward Excess Profitability (SEP) test. This test is constructed as a simple modification of the Excess Profitability (EP) test originally developed by Anatolyev and Gerco (2005). The SEP and EP test statistics differ because the SEP avoids the calculation of a term that under the null of no predictability should be zero, but in practice is sizable. In addition, our test does not require the strong assumption of independency used to derive the EP test. We claim that dependency is the rule and not the exception. We also show via Monte Carlo simulations that the size and power of the SEP test are in general better than those of the EP test. Despite of these distinctions, the SEP and EP tests share the same rich economic interpretation. The finding of predictability with these tests indicates that a given trading strategy provides positive returns on average. Therefore these two tests depart from the vast literature focused on statistical measures of predictability as Mean Squared Prediction Errors (MSPE), Least Absolute Deviations, Linex loss functions and Direction of Change, among others. See McCracken and West (2002), Chinn and Meese (1995), Cheung, Chinn and Garcia (2005), Patton and Timmermann (2007) and Giacomini and White (2006).

While the simple MDH is generally rejected when the econometrician engages in conventional in-sample analysis, it is indeed a difficult benchmark to beat when an out-of-sample approach is followed. The seminal paper of Meese and Rogoff (1983) is a classical example of this problem in the context of the exchange rate literature. This is sometimes interpreted as an indication that in-sample analysis is affected by overfitting or data mining problems and therefore should be disregarded. While the conflicting results from the in-sample and out-of-sample approaches are not entirely clear, Inoue and Kilian (2005) emphasize the higher power of in-sample strategies over out-of-sample analysis. According to this argument, out-of-sample tests of the MDH would fail to reject the null of no predictability mainly due to the low power of these tests. Therefore, it is essential to derive out-of-sample tests of the MDH that display power improvements with respect to their competitors.

A key point to consider when testing the MDH is that traditional statistical methods of comparing predictive accuracy, as those presented by West (1996) and Diebold and Mariano (1995), are not adequate. The reason for this is that the models under analysis are nested, see West (1996). McCracken (2007) and Clark and McCracken (2005) show that when comparing nested models, direct application of traditional methods may result in tests of inadequate size. In particular, Clark and West (2006) show that traditional comparisons of MSPE render a test with low power and questionable size.

Two interesting alternatives to construct proper statistical methods of comparing predictive accuracy when models are nested are presented by McCracken (2007) and Clark and West (2006). In the first paper, the asymptotic distribution of a t-type statistic, comparable to that suggested by West (1996) and Diebold and Mariano (1995), is derived. While the distribution is non standard, tables of asymptotically correct critical values are provided. In

the second paper, Clark and West (2006) make an important observation about the null hypothesis of equal MSPE when models are nested. They show that under the null the sample distribution of the difference in MSPE is not zero. Instead, the sample MSPE from the null model is expected to be smaller than that of the alternative model. They propose an adjusted procedure that takes into account this finding and demonstrate via simulations that this adjusted procedure is well-sized and relatively powerful in small and moderate samples. A third alternative to properly test for model adequacy when models are nested is provided by Chao, Corradi and Swanson (2000). They provide a test of Granger causality that displays good size and power properties in small samples when applied to the MDH.

Despite these interesting alternatives to properly test the MDH, this field of research is still not closed. From a theoretical point of view, we have already mentioned that it is desirable to derive more powerful tests in small and moderate samples and, more generally, to derive tests with better small sample properties. We also would like to have a test with a direct economic interpretation that allow us, for example, to obtain positive returns when using a given trading rule. That is exactly what we do in this paper. We introduce the SEP test which has a direct economic interpretation and displays better size and power than the EP test in our simulations.

The rest of the paper is organized as follows: Section 2 introduces the econometric environment; Section 3 shows that the original EP test may be biased under the null hypothesis of a martingale model; Section 4 introduces the new SEP test. Section 5 describes the experimental design and delivers the simulation results; Section 6 presents an empirical application within the exchange rate literature and Section 7 concludes.

2. Econometric Environment

Let us consider the simple setup introduced by Clark and West (2006) in which two simple models for a scalar stationary and ergodic time series y_{t+1} are defined as follows:

$$\text{Model 1 (null)} : y_{t+1} = e_{t+1} \quad (1)$$

$$\text{Model 2 (alternative)} : y_{t+1} = X'_{t+1}\beta + e_{t+1} \quad (2)$$

where X_{t+1} is a vector of ergodic and stationary exogenous random variables and e_{t+1} is a zero mean martingale difference meaning that $\mathbb{E}(e_{t+1}|\mathfrak{F}_t) = 0$, where $\{\mathfrak{F}_t\}$ represents a filtration such that $\mathfrak{F}_t$ is the sigma-field generated by current and past X 's and e 's

$$\mathfrak{F}_t = \sigma\{X_{t+1}, e_t, X_t, e_{t-1}, X_{t-1}, e_{t-2} \dots\}$$

Notice that we are using the index $t + 1$ to denote exogenous variables known at time t . Thus X_{t+1} is a vector containing known variables at time t . Notice also that we will use the notation $\mathbb{E}(e_{t+1}|\mathfrak{F}_t)$ or $\mathbb{E}_t(e_{t+1})$. The alternative model posits that the conditional expectation of y_{t+1} with respect to the filtration $\mathfrak{F}_t$ only depends linearly on the vector X_{t+1} and an unknown parameter β :

$$\mathbb{E}(y_{t+1}|\mathfrak{F}_t) = X'_{t+1}\beta \quad (3)$$

According to this condition, the null hypothesis may be written in terms of a restriction over the vector of parameters: $\beta = 0$ against the alternative $\beta \neq 0$.

Anatolyev and Gerco (2005) provide a very attractive out-of-sample test in a slightly different econometric context, but that in principle might be used to test $\beta = 0$ against $\beta \neq 0$. They introduce the Excess Profitability (EP) test that is based on the idea that y_{t+1}

corresponds to an asset return. Given a forecast $y_{t,t+1}^f$ built on information available up until time t for the return y_{t+1} they consider the following trading strategy:

1. To buy shares of the asset if the forecast $y_{t,t+1}^f$ predicts nonnegative returns in the following period, that is to say: $y_{t,t+1}^f \geq 0$.
2. To sell shares of the asset if the forecast $y_{t,t+1}^f$ predicts negative returns in the following period, that is to say: $y_{t,t+1}^f < 0$.

The one-period-return of this trading strategy is given by $r_{t+1} = \text{sign}(y_{t,t+1}^f)y_{t+1}$, where

$$\text{sign}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

Based upon this one-period-return, Anatolyev and Gerco (2005) construct an asymptotically normal statistic under the null hypothesis of conditional mean independence

$$A - G \text{ null hypothesis: } \mathbb{E}(y_{t+1}|\mathfrak{I}_t) = \mathbb{E}(y_{t+1}) = \mu \quad (4)$$

where $\mathfrak{I}_t$ represents the information set available to the forecaster at time t . This asymptotically normal statistic is given by

$$EP \equiv \frac{A_T - B_T}{\sqrt{\widehat{V}(A_T - B_T)}} \quad (5)$$

where

$$A_T = \frac{1}{T} \sum_t r_{t+1} \quad (6)$$

$$B_T = \left(\frac{1}{T} \sum_t \text{sign}(y_{t,t+1}^f) \right) \left(\frac{1}{T} \sum_t y_{t+1} \right) \quad (7)$$

and $\widehat{V}(A_T - B_T)$ represents a consistent estimator of the Variance of $A_T - B_T$ that is computed according to the following expressions:

$$\widehat{V}_{EP} \equiv \widehat{V}(A_T - B_T) = \frac{4}{T^2} \widehat{p}_{y^f} (1 - \widehat{p}_{y^f}) \sum_t (y_{t+1} - \bar{y})^2 \quad (8)$$

$$\widehat{p}_{y^f} = \frac{1}{2} \left[1 + \left(\frac{1}{T} \sum_t \text{sign}(y_{t,t+1}^f) \right) \right] \quad (9)$$

The rationale for using the two expressions above comes from the fact that under the null hypothesis (4) and some strong assumptions of independence (see Anatolyev and Gerco (2005) for details) we have

$$V(A_T - B_T) = \frac{4(T-1)}{T^2} p_{y^f} (1 - p_{y^f}) \text{var}(y_{t+1})$$

$$p_{y^f} = \Pr(\text{sign}(y_{t,t+1}^f) = 1)$$

so $\widehat{V}_{EP}$ is a “degrees of freedom” corrected variance estimator.

The intuition behind the definition of the EP test relies on the fact that under the null of no predictability both A_T and B_T converge in probability to the same value, so they are asymptotically equal.

It is tempting to export the EP test to our econometric framework. In fact our null hypothesis coincides with (4) when we consider the same information sets ($\mathfrak{F}_t = \mathfrak{I}_t$) and $\mu = 0$. Nevertheless, when $\mu = 0$ and no additional assumptions of independence are imposed we will show that the core EP statistic may be biased under the null hypothesis when some usual out-of-sample methods are implemented. We see this next.

3. Bias in the core EP statistic

Traditional out-of-sample analyses must split the available sample into two windows: an estimation window and a predictive window. For simplicity we will focus on one step ahead forecasts and we will use the notation in West (1996) and Clark and West (2006). If we have a total of $T + 1$ observations, then the last P will be used for predictions and $R = T + 1 - P$ will be used for the initial estimation of the parameters. $\hat{\beta}_t$ denotes a generic estimate of β with information available until time t . Usually, the updating scheme may be either fixed, rolling or recursive. The fixed scheme is one in which $\hat{\beta}_t$ is estimated only once using the first R observations. So, no updating occurs here. The rolling scheme updates the estimate of $\hat{\beta}_t$ using always the last R available observations. The recursive scheme also updates the estimate of $\hat{\beta}_t$, but this time using all available information until time t . That is to say, in the recursive scheme the estimation sample increases with t . We will work with the rolling scheme, partly because it is appropriate when working with time series that may have experienced breaks, and partly because we do not want parameter uncertainty to vanish asymptotically¹.

¹Working with the fixed scheme is also consistent with not vanishing parameter uncertainty, but it is probably inefficient. The recursive scheme, however, is conflicting with not vanishing uncertainty. The in-

Let us analyze the behavior of the core EP statistic under the null hypothesis. Under our notation, we will now have a total of P forecasts for evaluation. The EP test should be written according to this notation as follows:

$$EP \equiv \frac{A_P - B_P}{\sqrt{\widehat{V}(A_P - B_P)}} \quad (10)$$

where

$$A_P = \frac{1}{P} \sum_{t=R+1}^{T+1} r_t \quad (11)$$

$$B_P = \left(\frac{1}{P} \sum_{t=R+1}^{T+1} \text{sign}(y_{t-1,t}^f) \right) \left(\frac{1}{P} \sum_{t=R+1}^{T+1} y_t \right) \quad (12)$$

The next proposition shows the behavior of both terms, A_P and B_P , of the core EP statistic under the null hypothesis.

Proposition 1 *In the econometric framework characterized by models (1) and (2) we have that under the null hypothesis*

$$\begin{aligned} \mathbb{E}(A_P) &= 0 \\ \mathbb{E}(B_P) &= \frac{1}{P^2} \left\{ \sum_{t>j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) y_j) \right\} \end{aligned}$$

This last term may or may not be zero, depending on the updating scheme used in the out-of-sample exercise.

creasing precision of the recursive estimation due to the use of additional data might reduce the variance of the estimate, making it converge to the population parameter. This intuition is given by West (2006).

Proof. See the appendix. ■

Proposition 1 tells us that under the null hypothesis the core statistic from the EP test might not be zero when there is dependence between future values of the forecasts and past values of the target variable. A simple example in which $\mathbb{E}(B_P) \neq 0$ is given next.

Let us consider the simple case in which there is only one predictor under the alternative: a constant different from zero. Let us also assume that e_{t+1} is not only a zero mean martingale difference sequence but also is an independent sequence. Accordingly, here we assume that the alternative model looks as follows:

$$\text{Model 2 (alternative)} : y_{t+1} = c + e_{t+1} \quad (13)$$

Under this assumption, the best forecast is given by $y_{t-1,t}^f = c \neq 0$. If this constant c is unknown, it needs to be estimated. Let us assume that we use a rolling scheme to estimate c by OLS. Therefore we have that our feasible forecast corresponds to

$$\hat{y}_{t-1,t}^f = \hat{c}_t = \frac{1}{R} \sum_{i=t-R}^{t-1} y_i ; t = R+1, \dots, T+1$$

In the very particular case when $R = 1$ we have that

$$\hat{y}_{t-1,t}^f = \hat{c}_t = y_{t-1} ; t = 2, \dots, T+1$$

therefore

$$\begin{aligned}
\mathbb{E}(B_P) &= \frac{1}{P^2} \left\{ \sum_{t>j}^{T+1} \mathbb{E}(\text{sign}(\hat{y}_{t-1,t}^f) y_j) \right\} = \frac{1}{P^2} \left\{ \sum_{j=R+1}^{T+1} \sum_{t=j+1}^{T+1} \mathbb{E}(\text{sign}(\hat{y}_{t-1,t}^f) y_j) \right\} \\
&= \frac{1}{P^2} \left\{ \sum_{j=R+1}^{T+1} \sum_{t=j+1}^{T+1} \mathbb{E}(\text{sign}(y_{t-1}) y_j) \right\} \stackrel{H_0}{=} \frac{1}{P^2} \left\{ \sum_{j=R+1}^{T+1} \sum_{t=j+1}^{T+1} \mathbb{E}(\text{sign}(e_{t-1}) e_j) \right\} \\
&= \frac{1}{P^2} \sum_{j=R+1}^{T+1} \mathbb{E}(e_j \text{sign}(e_j)) = \frac{\mathbb{E}|e_j|}{P} > 0
\end{aligned}$$

In this particular case, the term under consideration is strictly positive, but it goes to zero as the number of forecasts increases.

This is just an example to show that this $\mathbb{E}(B_P)$ term might not be zero when the null hypothesis is true. We have used an extreme simplification in the context of a rolling estimation scheme to prove our point. Nevertheless we should also mention that the term $\mathbb{E}(B_P)$ will be zero in some circumstances. For instance, this happens when using the fixed scheme and our alternative model is given by (13). Under these conditions, the forecast $\hat{y}_{t-1,t}^f$ is constructed using information contained in $\mathfrak{F}_R$ and is never updated again. In this case we have that the forecasts could be expressed as

$$\hat{y}_{t-1,t}^f = \hat{c}_t = \widehat{c}_R \text{ for all } t = R+1, \dots, T+1$$

where $\widehat{c}_R$ is a function of the data $y_1, y_2, \dots, y_R$. A natural candidate for $\widehat{c}_R$ is the simple average:

$$\overline{c}_R = \frac{1}{R} \sum_{j=1}^R y_j$$

therefore

$$\begin{aligned}
\mathbb{E}(B_P) &= \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{y}_{t-1,t}^f) y_j) \right\} = \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{c}_R) y_j) \right\} \\
&= \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E} \{ \mathbb{E}(\text{sign}(\widehat{c}_R) y_j) | \mathfrak{F}_R \} \right\} = \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{c}_R) \mathbb{E}\{y_j | \mathfrak{F}_R\}) \right\}
\end{aligned}$$

but under the null hypothesis we have

$$\begin{aligned}
\mathbb{E}(B_P) &\stackrel{H_0}{=} \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{c}_R) \mathbb{E}\{y_j | \mathfrak{F}_R\}) \right\} = \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{c}_R) \mathbb{E}\{e_j | \mathfrak{F}_R\}) \right\} \\
&= \frac{1}{P^2} \left\{ \sum_{\substack{t>j \\ j=R+1}}^{T+1} \mathbb{E}(\text{sign}(\widehat{c}_R) 0) \right\} = 0
\end{aligned}$$

and the core test EP statistic is unbiased under the null.

In the next sections we will see that the bias of the test statistic may be important in more general examples than the extremely simple one we have described so far. Previous to that, we show next that for the null hypothesis in (1) we could construct an adjusted EP test which is unbiased under the null. The idea is simply to remove the B_P term from the construction of the test. The new test is called the Straightforward Excess Profitability (SEP) test.

4. The Straightforward Excess Profitability Test (SEP)

Here we propose a simple modification of the EP test by removing the B_P term from the EP test. In other words we propose the construction of a test for the martingale difference hypothesis based upon the following statistic:

$$A_P = \frac{1}{P} \sum_{t=R+1}^{T+1} r_t \quad (14)$$

$$r_{t+1} = \text{sign}(y_{t,t+1}^f) y_{t+1} \quad (15)$$

Assuming stationarity and ergodicity for r_t we have that under the null

$$H_0 : \mathbb{E}[r_{t+1}] = \mathbb{E}[\mathbb{E}[r_{t+1}|\mathfrak{F}_t]] = \mathbb{E}\left[\mathbb{E}\left[\text{sign}(y_{t,t+1}^f)e_t|\mathfrak{F}_t\right]\right] = \mathbb{E}\left[\text{sign}(y_{t,t+1}^f)\mathbb{E}[e_{t+1}|\mathfrak{F}_t]\right] = 0$$

this is because

$$\text{sign}(y_{t,t+1}^f) \in \mathfrak{F}_t$$

Under the alternative we have that

$$\begin{aligned} H_A : \quad \mathbb{E}[r_{t+1}] &= \mathbb{E}[\mathbb{E}[r_{t+1}|\mathfrak{F}_t]] = \mathbb{E}\left[\mathbb{E}\left[\text{sign}(y_{t,t+1}^f)y_{t+1}|\mathfrak{F}_t\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\text{sign}(y_{t,t+1}^f)[X'_{t+1}\beta + e_{t+1}]|\mathfrak{F}_t\right]\right] = \\ &= \mathbb{E}\left[\mathbb{E}\left[\text{sign}(y_{t,t+1}^f)[X'_{t+1}\beta]|\mathfrak{F}_t\right]\right] + \mathbb{E}\left[\text{sign}(y_{t,t+1}^f)\mathbb{E}[e_{t+1}|\mathfrak{F}_t]\right] \\ &= \mathbb{E}\left[\text{sign}(y_{t,t+1}^f)[X'_{t+1}\beta]\right] = \mathbb{E}\left[\text{sign}(X'_{t+1}\hat{\beta}_t)[X'_{t+1}\beta]\right] \end{aligned}$$

As long as $\widehat{\beta}_t$ is an adequate estimate of β , then we will have that

$$H_A : \mathbb{E} \left[\text{sign}(X'_{t+1} \widehat{\beta}_t) [X'_{t+1} \beta] \right] \approx \mathbb{E} [|X'_{t+1} \beta|] > 0 \quad (16)$$

under mild assumptions we will have that under the null²

$$\sqrt{P} A_P = \frac{\sqrt{P}}{P} \sum_{t=R+1}^{T+1} r_t \rightsquigarrow N(0, V)$$

We could then construct a consistent estimate of V that we will denote $\widehat{V}$ to build the following statistic:

$$SEP \equiv \frac{\frac{\sqrt{P}}{P} \sum_{t=R+1}^{T+1} r_t}{\sqrt{\widehat{V}}} \rightsquigarrow N(0, 1)$$

given (16) we will construct a one sided test, as under the alternative we expect a positive value for our statistic.

In the next section we compare size and power properties of the EP and SEP test via simulations. We also include in the comparison a widely cited test developed by Clark and West (2006) just to have an external and relatively recent benchmark.

5. Monte Carlo Simulations

We use Monte Carlo simulations for three multivariate data-generating processes (DGPs) to evaluate the finite-sample properties of the SEP test. The idea is to compare small sample

²Under the null, r_t is a martingale difference sequence, so we invoke the assumptions required for the central limit theorem for a martingale difference sequence to apply. See Hamilton (1994).

properties of this test with those of some benchmark tests already available in the literature. We consider two benchmark tests: the traditional EP test and the MSPE-Adjusted test introduced by Clark and West (2006). This last test is derived from typical comparisons of MSPE between a linear model and the null of a martingale difference sequence. Clark and West claim that the MSPE-Adjusted test has better size than traditional tests of MSPE comparisons when the models under evaluation are nested. Intuitively this test shows good size because it does not take into account a term that introduces noise into its forecasts by estimating a parameter vector that under the null should be zero. To see this we notice that the sample analog of the difference in MSPE between the two models considered in (1) and (2) is given by

$$\begin{aligned}
\widehat{\Delta MSPE} &= \widehat{MSPE}_1 - \widehat{MSPE}_2 \\
&= \frac{1}{P} \sum_{t=R}^T (y_{t+1})^2 - \frac{1}{P} \sum_{t=R}^T (y_{t+1} - X'_{t+1} \hat{\beta}_t)^2 \\
&= \frac{2}{P} \sum_{t=R}^T y_{t+1} X'_{t+1} \hat{\beta}_t - \frac{1}{P} \sum_{t=R}^T (X'_{t+1} \hat{\beta}_t)^2
\end{aligned}$$

Clark and West notice that the second term in the right hand side introduces a bias that does not vanish as P goes to ∞ . They propose to build a test based upon the first term in the right hand side. Their test is given by

$$MSPE - Adjusted : P^{1/2} \frac{\frac{2}{P} \sum_{t=R}^T y_{t+1} X'_{t+1} \hat{\beta}_t}{\sqrt{4\widehat{V}(y_{t+1} X'_{t+1} \hat{\beta}_t)}}$$

So, we see that under the null hypothesis and mild assumptions, the three tests under considerations are asymptotically normal.

We use DGPs based on Clark and West (2006) and Anatolyev and Gerko (2005). These DGPs are calibrated to match features of exchange rate and asset returns for which the martingale difference is a plausible null model. Specifically the first DGP (DGP1) is given by the following process:

DGP1:

$$\begin{aligned} y_{t+1} &= \beta x_t + e_{t+1} \\ x_t &= 0,95x_{t-1} + u_t; \\ e_{t+1} &= N(0, \sigma_e^2); \quad u_{t+1} = N(0, \sigma_u^2) \end{aligned}$$

with $\mathbb{E}(e_{t+1}|\mathfrak{F}_t) = 0$, $\mathbb{E}(u_{t+1}|\mathfrak{F}_t) = 0$ and $\sigma_e^2 = 1$.

DGP1 is calibrated to match exchange rate features based on interest parity so we will have $var(u_t) = \sigma_u^2$ (with $\sigma_u = 0,025$) and $corr(e_t, u_t) = 0$. We set $\beta = -2$ in experiments evaluating power and $\beta = 0$ in experiments evaluating size. We assume data to be generated from homoskedastic draws of their respective distributions³.

The next two DGPs are simple processes calibrated to match S&P 500 returns and are identical to those used by Anatolyev and Gerko (2005):

DGP2: This is a simple first order autoregressive model for the S&P 500 returns:

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim N(0, 0,000249) \end{aligned}$$

We set $\rho = 0,1256$ for power results and $\rho = 0$ for size results.

³We are aware that $\beta = 1$ is the theoretical implication. Empirical estimations for a number of industrial countries, however, provide estimates around -2. See Clark and West (2006) for further details.

DGP3: This DGP is similar to DGP2 but now this process includes a threshold parameter to simulate possible breaks in the data.

$$y_t = \begin{cases} c_1 + \rho_1 y_{t-1} + \varepsilon_t \rightarrow |y_{t-1}| \leq 0,01848 \\ c_2 + \rho_2 y_{t-1} + \varepsilon_t \rightarrow |y_{t-1}| > 0,01848 \end{cases}$$

$$\varepsilon_t \sim N(0, 0,000245)$$

We set $c_1 = \rho_1 = c_2 = \rho_2 = 0$ when analyzing the size of the tests. For power, and following Anatolyev and Gerco (2005), we set

$$c_1 = 0,000844; \rho_1 = 0,2453; c_2 = 0,002679; \rho_2 = 0,0664$$

For all our simulations we focus, for simplicity, only on one step ahead forecasts. Estimation always includes a constant term in each regression. We explore the performance of rolling schemes for a number of sample sizes $(T + 1)$ and decompositions of the sample in the estimation window (size R) and the prediction window (size $P, (T + 1 = R + P)$). We run simulations for the following sample sizes: $R = 50$ and 100 ; $P = 25, 100$ and 750 . The total number of simulations is 5000 . We use OLS to get estimates of the parameters. The main features arising from the simulations with DGP1 and DGP 2 are analyzed in the following subsection. Results for DGP 3 are mostly used as a robustness check. We do this because for this DGP we estimate an AR(1) model with drift, so the threshold parameter is omitted in the estimation and a misspecification error is present.

5.1. Simulation Results

Tables 1-2 display results of empirical size and power for the three tests under consideration when DGP1 is simulated. Table 1 shows results when the rolling estimation window has only 50 observations. Table 2 shows results when 100 observations are used for the rolling windows.

**Table 1: Empirical Power and Size for DGP1 (Exchange Rate Model)
Estimation Window of 50 Observations (R=50), One tail test**

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.113	0.367	0.608	0.190	0.490	0.723
EP	0.039	0.275	0.545	0.077	0.384	0.660
MSPE-Adjusted	0.094	0.413	0.709	0.170	0.531	0.799
	<i>Empirical Size</i>					
SEP	0.049	0.046	0.044	0.097	0.098	0.097
EP	0.021	0.029	0.033	0.045	0.063	0.072
MSPE-Adjusted	0.035	0.038	0.042	0.072	0.080	0.084

**Table 2: Empirical Power and Size for DGP1 (Exchange Rate Model)
Estimation Window of 100 Observations (R=100) One tail test**

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.129	0.482	0.745	0.223	0.596	0.829
EP	0.058	0.365	0.675	0.107	0.475	0.767
MSPE-Adjusted	0.118	0.538	0.851	0.201	0.649	0.905
	<i>Empirical Size</i>					
SEP	0.049	0.051	0.057	0.105	0.100	0.108
EP	0.022	0.020	0.036	0.052	0.049	0.074
MSPE-Adjusted	0.036	0.033	0.041	0.082	0.071	0.085

Tables 1 and 2 show similar results. In general the original EP seems to be undersized, while the SEP has better size properties than both the EP and MSPE-Adjusted tests. These

results does not change much with the size of the estimation window. In terms of power we see that for small P the SEP test has much more power than the EP test and a little more power than the MSPE-Adjusted test. As the number of forecasts gets larger, we still see a better performance of the SEP test over the EP test. Nevertheless, the MSPE-Adjusted test dominates the rest of the tests. The size of the estimation window does not seem to have an impact on the empirical size of the tests. Nevertheless, it does have an impact in their power. We can see that with a larger estimation window, all three tests display higher power.

**Table 3: Empirical Power and Size for DGP2 (AR(1) for S&P 500)
Estimation Window of 50 Observations (R=50), One tail test**

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.092	0.228	0.389	0.159	0.340	0.516
EP	0.046	0.173	0.338	0.095	0.265	0.461
MSPE-Adjusted	0.078	0.297	0.535	0.150	0.407	0.649
	<i>Empirical Size</i>					
SEP	0.049	0.055	0.052	0.105	0.101	0.097
EP	0.032	0.036	0.036	0.067	0.067	0.077
MSPE-Adjusted	0.044	0.040	0.049	0.085	0.083	0.093

**Table 4: Empirical Power and Size for DGP2 (AR(1) for S&P 500)
Estimation Window of 100 Observations (R=100) One tail test**

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.094	0.313	0.501	0.175	0.434	0.637
EP	0.060	0.240	0.434	0.110	0.346	0.569
MSPE-Adjusted	0.090	0.383	0.648	0.172	0.500	0.760
	<i>Empirical Size</i>					
SEP	0.049	0.048	0.047	0.102	0.099	0.091
EP	0.031	0.023	0.030	0.070	0.051	0.060
MSPE-Adjusted	0.039	0.031	0.038	0.087	0.072	0.076

Tables 3 and 4 show power and size results when DGP2 is simulated. These results corroborate the findings shown in tables 1-2: First, the SEP test dominates the EP test both

in terms of size and power. Second, the MSPE-Adjusted test has no better empirical size compared to the SEP test. Third, the MSPE-Adjusted test is the best in terms of power when the number of forecasts is moderate or large and fourth, with a larger estimation window, all three tests display higher power.

Table 5: Empirical Power and Size for DGP3 (SETAR Model estimated as an AR(1))
Estimation Window of 50 Observations (R=50), One tail test

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.165	0.647	0.916	0.263	0.751	0.954
EP	0.091	0.542	0.879	0.167	0.656	0.931
MSPE-Adjusted	0.166	0.747	0.973	0.273	0.830	0.987
	<i>Empirical Size</i>					
SEP	0.051	0.048	0.050	0.106	0.099	0.104
EP	0.032	0.028	0.037	0.066	0.060	0.078
MSPE-Adjusted	0.040	0.034	0.040	0.084	0.073	0.082

Table 6: Empirical Power and Size for DGP3 (SETAR Model estimated as an AR(1))
Estimation Window of 100 Observations (R=100) One tail test

	5%			10%		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>Empirical Power</i>					
SEP	0.185	0.780	0.969	0.301	0.858	0.985
EP	0.118	0.692	0.950	0.210	0.782	0.973
MSPE-Adjusted	0.195	0.851	0.993	0.319	0.912	0.998
	<i>Empirical Size</i>					
SEP	0.048	0.046	0.054	0.098	0.100	0.106
EP	0.029	0.020	0.034	0.064	0.048	0.072
MSPE-Adjusted	0.038	0.035	0.040	0.084	0.072	0.086

Tables 5-6 present results when DGP3 is estimated as a simple first order autoregressive process. Let us recall that under the null hypothesis DGP3 is simple a white noise, but under the alternative is a SETAR model. Results in tables 5-6 corroborate those shown in tables 1-4 with only one difference: now the MSPE-Adjusted test dominates the rest of the tests in

terms of power even when the number of forecasts is small. Aside from this salient feature, the rest of the findings coincide with those in tables 1-4.

5.2. Statistics Bias Results

In previous sections we proved that under the null hypothesis the core EP statistic may have an expected value different from zero. Tables 7-8 show descriptive statistics (Mean and Median) for the SEP, EP and MSPE-Adjusted statistics and it's differences. The idea is to explore whether this bias is present in our simulations. From tables 7-8 we see that it is indeed present. In fact we detect a negative bias for the EP test in all our simulations. Interestingly this bias decreases (in absolute terms) as the number of forecasts grows larger. The MSPE-Adjusted statistic follows the same pattern. As mentioned by Clark and West (2006) and Pincheira (2013), this adjusted statistic has also a negative bias under the null hypothesis that decreases as the number of forecasts grows larger. Finally, we see that our proposed SEP test presents a negligible bias with no clear sign, which suggests that the modification we propose here is unbiased under the null hypothesis.

Table 7: Mean and Median of the Tests Statistics under the Null DGP1 (Exchange Rate Model)

	Mean			Median		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>R=50</i>					
SEP	0.014	-0.007	-0.019	0.018	-0.020	-0.040
EP	-0.440	-0.236	-0.166	-0.762	-0.253	-0.176
Difference (SEP minus EP)	0.426	0.229	0.147	0.744	0.234	0.137
MSPE-Adjusted	-0.187	-0.136	-0.099	-0.236	-0.144	-0.124
Difference (SEP minus MSPE-Adjusted)	0.172	0.129	0.080	0.219	0.124	0.084
	<i>R=100</i>					
SEP	-0.009	-0.001	-0.010	-0.009	0.007	-0.017
EP	-0.444	-0.327	-0.225	-1.105	-0.318	-0.230
Difference (SEP minus EP)	0.435	0.326	0.215	1.097	0.311	0.213
MSPE-Adjusted	-0.199	-0.178	-0.127	-0.250	-0.183	-0.140
Difference (SEP minus MSPE-Adjusted)	0.190	0.177	0.117	0.242	0.176	0.122

**Table 8: Mean and Median of the Tests Statistics under the Null
DGP2 (AR(1) for S&P 500)**

	Mean			Median		
	P=25	P=300	P=750	P=25	P=300	P=750
	<i>R=50</i>					
SEP	<i>0.029</i>	<i>-0.014</i>	<i>-0.013</i>	<i>0.018</i>	<i>-0.001</i>	<i>-0.016</i>
EP	<i>-0.299</i>	<i>-0.346</i>	<i>-0.176</i>	<i>-0.485</i>	<i>-0.261</i>	<i>-0.175</i>
Difference (SEP minus EP)	<i>0.270</i>	<i>0.333</i>	<i>0.164</i>	<i>0.467</i>	<i>0.260</i>	<i>0.159</i>
MSPE-Adjusted	<i>-0.132</i>	<i>-0.185</i>	<i>-0.100</i>	<i>-0.164</i>	<i>-0.145</i>	<i>-0.126</i>
Difference (SEP minus MSPE-Adjusted)	<i>0.103</i>	<i>0.172</i>	<i>0.087</i>	<i>0.146</i>	<i>0.144</i>	<i>0.110</i>
	<i>R=100</i>					
SEP	<i>-0.005</i>	<i>-0.009</i>	<i>0.014</i>	<i>-0.008</i>	<i>-0.009</i>	<i>0.008</i>
EP	<i>-0.247</i>	<i>-0.260</i>	<i>-0.206</i>	<i>-0.533</i>	<i>-0.359</i>	<i>-0.215</i>
Difference (SEP minus EP)	<i>0.243</i>	<i>0.251</i>	<i>0.192</i>	<i>0.526</i>	<i>0.350</i>	<i>0.206</i>
MSPE-Adjusted	<i>-0.141</i>	<i>-0.130</i>	<i>-0.116</i>	<i>-0.169</i>	<i>-0.176</i>	<i>-0.108</i>
Difference (SEP minus MSPE-Adjusted)	<i>0.136</i>	<i>0.121</i>	<i>0.102</i>	<i>0.161</i>	<i>0.168</i>	<i>0.099</i>

In summary our simulations show that both the EP test and MSPE-Adjusted test are undersized, while the SEP test presents good size properties. In addition, the SEP test presents lower bias under the null and more power than the EP test, although less power than the MSPE-Adjusted test.

6. Empirical Illustration

In this section we study the behavior of the SEP test using monthly forecasts of four US dollar bilateral exchange rates. We analyze the cases of Canada, Japan, Switzerland, and the U.K. While the null model corresponds to a zero mean martingale difference for the percentage change in exchange rates, the alternative model posits that this percentage change is explained by two regressors: a constant and the one-month interest differential. The data were generously provided by Todd Clark and correspond to the same database used in Clark and West (2006). Interest rates correspond to 1-month euro-currency deposit rates, taking an average of bid and ask rates at London close. Monthly time series are formed as the last daily rate of each month. Data was obtained from Global Insight's FACS database.

Using rolling regressions estimated by OLS we engage in two empirical exercises. We assume that the number of observations used for the first estimation (R) as well as the number of predictions (P) are fixed. We follow Clark and West (2006) to choose R relatively small with respect to P . In the first empirical exercise we set $R = 120$ and $P = 166$. In the second we choose $R = 50$ and $P = 236$. Then we compute the EP, SEP and MSPE-Adjusted statistics to analyze whether the tests are able to reject the null of a martingale difference hypothesis.

Results from the two empirical exercises are displayed in Tables 9-10. Robust evidence of predictability is only found for Canada. For this country, all three tests reject the martingale difference hypothesis at the 10% significance level irrespective of the values for R and P . For the rest of the countries the evidence is mixed in the sense that some tests, for some choices of R and P , detect predictability, but some others do not. When comparing the performance of the EP and SEP tests we observe that the SEP test statistic is always higher than the EP test statistic with the only exception of the UK when $R = 50$. When compared to the MSPE-Adjusted statistic, no clear pattern is found and a mixture of results is the rule. For instance, in table 9 both tests statistics tend to reject the null for the same countries at similar significance levels with, maybe, the subtle exception of Japan. Nevertheless, in table 10 we detect two important differences. First, the SEP test rejects the null for Japan at the 5% significance level but the MSPE-Adjusted test is unable to reject it even at the 10% significance level. A similar discrepancy happens for Switzerland where now the SEP test is unable to reject the null but the MSPE-Adjusted test rejects at the 10% level.

Table 9: Forecasts of Monthly Changes in U.S. Dollar Exchange Rates

(1)			(2)	(3)	(4)
Country	Prediction Sample	R, P	MSPE Adjusted	EP	SEP
Canada	1990:1-2003:10	R=120, P=166	0.134 (0.076)	0.206 (0.109)	0.238 (0.119)
			1.760	1.894	1.995
Pvalue			0.039	0.029	0.023
Japan	1990:1-2003:10	R=120, P=166	0.529 (0.430)	0.415 (0.258)	0.441 (0.261)
			1.230	1.612	1.687
Pvalue			0.109	0.053	0.046
Switzerland	1985:01-2003:10	R=120, P=226	0.899 (0.482)	0.475 (0.205)	0.613 (0.233)
			1.867	2.315	2.629
Pvalue			0.031	0.010	0.004
U.K.	1985:01-2003:10	R=120, P=226	0.011 (0.334)	-0.071 (0.188)	0.000 (0.208)
			0.034	-0.375	-0.002
Pvalue			0.487	0.646	0.501

Notes: For each country and test statistic table 9 display the corresponding core statistic (above), standard error (middle), t-statistic (below) and p-value (bold).

Table 10: Forecasts of Monthly Changes in U.S. Dollar Exchange Rates

(1)			(2)	(3)	(4)
Country	Prediction Sample	R, P	MSPE Adjusted	EP	SEP
Canada	1990:1-2003:10	R=50, P=236	0.179 (0.095)	0.133 (0.093)	0.137 (0.095)
			1.884	1.430	1.448
Pvalue			0.03	0.08	0.07
Japan	1990:1-2003:10	R=50, P=236	0.355 (0.476)	0.277 (0.201)	0.422 (0.227)
			0.746	1.376	1.860
Pvalue			0.23	0.08	0.03
Switzerland	1985:01-2003:10	R=50, P=296	0.646 (0.442)	0.163 (0.194)	0.186 (0.204)
			1.463	0.839	0.911
Pvalue			0.07	0.20	0.18
U.K.	1985:01-2003:10	R=50, P=296	0.686 (0.334)	0.328 (0.176)	0.313 (0.181)
			2.051	1.868	1.726
Pvalue			0.02	0.03	0.04

Notes: For each country and test statistic table 10 display the corresponding core statistic (above), standard error (middle), t-statistic (below) and p-value (bold).

7. Conclusion

In this paper we propose a new out-of sample test of predictability that is constructed as a simple modification of the Excess Profitability (EP) test developed by Anatolyev and Gerco (2005). Our proposed test is called Straightforward Excess Profitability (SEP) test, and avoids the calculation of a term in the EP test that under the null of no predictability should be zero, but in practice may be sizable. In addition, our test does not require the strong assumption of independency used to derive the EP test. We claim that dependency is the rule and not the exception.

We use Monte Carlo simulations to explore the small and moderate sample behavior of the SEP test, comparing it with the EP and MSPE-Adjusted test derived by Clark and West (2006). Our simulations make use of three different data generating processes designed to match features of exchange rates and S&P 500 monthly returns as shown by Clark and West (2006) and Anatolyev and Gerco (2005). Simulation results show that the SEP test displays better size and more power than the EP test. We link this better behavior to an important bias that is present in the EP test under the null of no predictability. This bias generates an undersized test which may also explain the lower power of the test. By removing the term that generates the bias, the SEP test displays a superior performance.

Our SEP test also fares well when compared to the MSPE-Adjusted test in terms of empirical size. Nevertheless, the MSPE-Adjusted test tends to show more power than the SEP test, specially when the number of forecasts is large.

We illustrate the use of the SEP test in an empirical application within the exchange rate literature. Using monthly series of bilateral exchange rates and an alternative model based

on interest parity, we find robust evidence of predictability for Canada. For this country, all three tests (EP, SEP and MSPE-Adjusted) reject the martingale difference hypothesis at the 10 % significance level. For the rest of the countries (Japan, Switzerland and the UK) the evidence is mixed. When comparing the performance of the EP and SEP tests we observe that the SEP test statistic is always higher than the EP test statistic with only one exception.

As a final remark let us notice that the SEP test could also be considered as a “shrinkage-based-test” according to the terminology used in Pincheira (2013). In this regard, the estimation of an optimal shrinkage factor to improve forecast accuracy is an open question that might be addressed as the object of future research.

References

1. Anatolyev S. and A. Gerco (2005): “A Trading Approach to Testing for Predictability”. *Journal of Business & Economic Statistics* **4**(2): 455-461.
2. Chao J, Corradi V. and N. Swanson (2000): “An Out of Sample Test for Granger Causality”. *Macroeconomic Dynamics* **5**: 598-620.
3. Cheung Y, Chinn M and A. Garcia Pascual (2005): “Empirical Exchange Rate Models of the 1990’s: Are Any Fit to Survive?”. *Journal of International Money and Finance* **24**:1150-1175.
4. Chinn, M. and R. Meese (1995): “Banking on currency forecasts: How predictable is change in money?”. *Journal of International Economics* **38**(1-2):161-178. February.
5. Clark T. and M. McCracken (2005): “Evaluating Direct Multistep Forecasts”. *Econometric Reviews* **24**(4): 369-404.
6. Clark T and K. West (2006): “Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis”. *Journal of Econometrics* **135**: 155-186.
7. Diebold, F. and R. Mariano (1995): “Comparing Predictive Accuracy”. *Journal of Business & Economic Statistics* **13**: 253-263.

8. Giacomini, R. and H. White (2006): “Tests of Conditional Predictive Ability”. *Econometrica* **74**(6): 1545-1578.
9. Hamilton J. (1994): “Time Series Analysis”, Princeton University Press.
10. Inoue A. and L. Kilian (2005): “In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?”. *Econometric Reviews* **23**(4): 371-402.
11. McCracken M. (2007): “Asymptotics for Out of Sample Tests of Granger Causality”. *Journal of Econometrics* **140**: 719-752.
12. McCracken M and K. West. (2002): “Inference About Predictive Ability”. In M. Clements and D. Hendry (eds) *Companion to Economic Forecasting*, Oxford: Blackwell. 299-321.
13. Meese R and K. Rogoff (1983): “Empirical Exchange Rate Models of the Seventies. Do They Fit Out-of-Sample?”. *Journal of International Economics* **14**: 3-24.
14. Patton A. and A. Timmermann (2007). “Properties of optimal forecasts under asymmetric loss and nonlinearity”. *Journal of Econometrics* **140**(2): 884-918, October.
15. Pincheira P. (2013): “Shrinkage Based Tests of Predictability”. *Journal of Forecasting* **32**(4): 289-384.
16. West K.(1996): “Asymptotic Inference About Predictive Ability”. *Econometrica* **64**: 1067-1084.
17. West K. (2006): “Forecast Evaluation” *Handbook of Economic Forecasting*, Elsevier.

8. Appendix

Proposition 1: In the econometric framework characterized by models (1) and (2) we have that under the null hypothesis

$$\begin{aligned}\mathbb{E}(A_P) &= 0 \\ \mathbb{E}(B_P) &= \frac{1}{P^2} \left\{ \sum_{t>j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) y_j) \right\}\end{aligned}$$

This last term may or may not be zero, depending on the updating scheme used in the out-of-sample exercise.

Proof: Let us take expectations on both A_P and B_P :

$$\mathbb{E}(A_P) = \mathbb{E}\left(\frac{1}{P} \sum_{t=R+1}^{T+1} r_t\right) = \frac{1}{P} \sum_{t=R+1}^{T+1} \mathbb{E}(r_t) \quad (17)$$

$$\mathbb{E}(r_t) = \mathbb{E}(\mathbb{E}(r_t|\mathfrak{F}_{t-1})) = \mathbb{E}(\mathbb{E}(\text{sign}(y_{t-1,t}^f)y_t|\mathfrak{F}_{t-1})) \quad (18)$$

$$\mathbb{E}(r_t) = \mathbb{E}(\text{sign}(y_{t-1,t}^f)\mathbb{E}(y_t|\mathfrak{F}_{t-1})) \quad (19)$$

$$\mathbb{E}(r_t) = \mathbb{E}(\text{sign}(y_{t-1,t}^f)\mathbb{E}(e_t|\mathfrak{F}_{t-1})) \quad (20)$$

$$\mathbb{E}(r_t) = \mathbb{E}(\text{sign}(\hat{y}_t)0) = 0 \quad (21)$$

$$\mathbb{E}(B_P) = \mathbb{E}\left\{\left(\frac{1}{P} \sum_{t=R+1}^{T+1} \text{sign}(y_{t-1,t}^f)\right)\left(\frac{1}{P} \sum_{t=R+1}^{T+1} y_t\right)\right\} \quad (22)$$

$$\mathbb{E}(B_P) = \frac{1}{P^2} \sum_{t,j=R+1}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f)y_j) \quad (23)$$

$$\mathbb{E}(B_P) = \frac{1}{P^2} \left\{ \sum_{t=j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f)y_j) + \sum_{t \neq j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f)y_j) \right\} \quad (24)$$

therefore

$$\mathbb{E}(B_P) = \frac{1}{P^2} \left\{ \sum_{t=R+1}^{T+1} \mathbb{E}(r_t) + \sum_{t < j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f)y_j) + \sum_{t > j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f)y_j) \right\} \quad (25)$$

Under the null hypothesis we have:

$$\mathbb{E}(B_P) \stackrel{H_0}{=} \frac{1}{P^2} \left\{ \sum_{t=R+1}^{T+1} 0 + \sum_{t < j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) \mathbb{E}(y_j | \mathfrak{F}_{j-1})) + \sum_{t > j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) y_j) \right\} \quad (26)$$

$$\mathbb{E}(B_P) \stackrel{H_0}{=} \frac{1}{P^2} \left\{ \sum_{t=R+1}^{T+1} 0 + \sum_{t < j}^{T+1} 0 + \sum_{t > j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) y_j) \right\} \quad (27)$$

$$\mathbb{E}(B_P) \stackrel{H_0}{=} \frac{1}{P^2} \left\{ \sum_{t > j}^{T+1} \mathbb{E}(\text{sign}(y_{t-1,t}^f) y_j) \right\} \quad (28)$$

which is the desired result.

Documentos de Trabajo Banco Central de Chile	Working Papers Central Bank of Chile
<p>NÚMEROS ANTERIORES</p> <p>La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica:</p> <p>www.bcentral.cl/esp/estpub/estudios/dtbc.</p> <p>Existe la posibilidad de solicitar una copia impresa con un costo de Ch\$500 si es dentro de Chile y US\$12 si es fuera de Chile. Las solicitudes se pueden hacer por fax: +56 2 26702231 o a través del correo electrónico: bcch@bcentral.cl.</p>	<p>PAST ISSUES</p> <p>Working Papers in PDF format can be downloaded free of charge from:</p> <p>www.bcentral.cl/eng/stdpub/studies/workingpaper.</p> <p>Printed versions can be ordered individually for US\$12 per copy (for order inside Chile the charge is Ch\$500.) Orders can be placed by fax: +56 2 26702231 or by email: bcch@bcentral.cl.</p>

DTBC – 697

Vulnerability, Crises and Debt Maturity: Do IMF Interventions Shorten the Length of Borrowing?

Diego Saravia

DTBC – 696

Are All Capital Inflows Associated with *Booms* in House Prices? An Empirical Evaluation

Alejandro Jara y Eduardo Olaberria

DTBC – 695

Household Debt During the Financial Crisis: Micro-Evidence from Chile

Roberto Álvarez y Luis Opazo

DTBC – 694

Efficient CPI-Based Taylor Rules in Small Open Economies

Rodrigo Caputo y Luis Oscar Herrera

DTBC – 693

Interventions and Inflation Expectations in an Inflation Targeting Economy

Pablo Pincheira

DTBC – 692

The Response Speed of the International Monetary Fund

Ashoka Mody y Diego Saravia

DTBC – 691

Efectos de *Shocks* al Precio del Petróleo sobre la Economía en Chile y sus Socios Comerciales

Michael Pedersen y Miguel Ricaurte

DTBC – 690

Aumento del Salario Mínimo y sus Efectos sobre el Mercado Laboral

Gonzalo Castex

DTBC – 689

The Impact of Persistence in Volatility over the Probability of Default

Rodrigo Alfaro y Natán Goldberger

DTBC – 688

Rebellions, Technical Change, and the Early Development of Political Institutions in Latin America

Álvaro Aguirre

DTBC – 687

Are International Market Linkages Stronger? Comparison between 1990s and 2000s

Francisca Pérez

DTBC – 686

What Affects the Predictions of Private Forecasters? The Role of Central Bank Forecasts

Michael Pedersen

DTBC – 685

Pronósticos con Métodos Shrinkage utilizando una Gran Base de Datos

Wildo González y Hernán Rubio

DTBC – 684

Precio de Materias Primas y Spread Soberano en Economías Emergentes ¿Importa la Concentración de las Exportaciones?

Ercio Muñoz

DTBC – 683

Evolution of a Small Open Emerging Economy's External Vulnerability: Evidence for Chile

Gastón Chaumont y Markus Kirchner

BANCO CENTRAL
DE CHILE