EVALUATION OF A NEW OUT-OF-SAMPLE TEST OF PREDICTABILITY AND ITS COMPARISON WITH PREVIOUS TESTS IN ITERATED MULTI-STEP-AHEAD FORECASTS

TESIS PARA OPTAR AL GRADO DE MAGISTER EN ECONOMIA

17-06-2020

FELIPE IGNACIO MUÑOZ LAMILLA

PROFESOR GUIA: PABLO PINCHEIRA, Ph.D

PROFESORES CORRECTORES: VIVIANA FERNANDEZ, Ph.D; JUAN PABLO MEDINA, Ph.D

Evaluation of a new out-of-sample test of predictability and its comparison with previous tests in iterated multi-step-ahead forecasts.

by Felipe Muñoz

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Economics

at the

Universidad Adolfo Ibáñez

Author: Felipe Muñoz

Thesis Advisor: Pablo Pincheira, Ph.D

Thesis Committee: Viviana Fernández, Ph.D; Juan Pablo Medina, Ph.D.

2020

Acknowledgement

First and foremost, I would like to express my deep and sincere gratitude to my thesis advisor, Prof. Pablo Pincheira, for giving me the opportunity to do research and providing invaluable guidance throughout this work. His vision, sincerity and motivation have deeply inspired me. It was a great privilege and honor to work and study under his guidance.

Besides my advisor, I would like to thank my thesis committee: Prof. Viviana Fernández and Prof. Juan Pablo Medina.

My sincere thanks also goes to Prof. Andrea Repetto for her enthusiasm and encouragement.

I cannot forget to thank my parents for all the unconditional support during this process.

Last but not least, gratitude goes to my friend Fernando who directly or indirectly helped me during this process.

To Francisca, thank you for your love, advice, patience and faith.

Contents

1 Introduction

Predicting the behavior of variables is an important activity for making decisions. Economic forecasting describes the process of making an attempt to predict the future condition of an economy of interest using a set of predefined indicators. This is a key area in economics as it allows policymakers to make informed decisions. For example, if we forecast the economy will enter into a recession, the government could consider the implementation of an expansionary fiscal policy (increasing spending) to make sure demand in the economy is sustained and to prevent a significant downturn in the economy.

Economists usually deal with forecasts for economic variables such as: inflation, population growth, exchange rates, etc. Forecasts are subject to error and this is because different variables are simultaneously affecting the economy. For example, what will happen to oil prices? how will consumers react to a catastrophic event?.

One key aspect is to test models' forecasting ability. On the one hand, there are "tests of absolute forecasting performance", which are used to evaluate a single model. Some key performance indicators might include lack of serial correlation in one step ahead prediction errors, bias in predictions or correlation between prediction and realization. Examples of these tests are available in Mincer and Zarnowitz (1969) and West and McCracken (1998). These tests are used to evaluate one specific forecasting model. On the other hand, there are "tests of relative forecast comparisons", which involve comparing between two or more competing models with different loss functions at the population level to determine which one of them predicts the best. In this case, measures of relative model quality include ratios or differences of mean, mean-squared or mean-absolute prediction errors. The most prominent is the Mean Squared Prediction Error (MSPE) and the most prominent of these tests is the Diebold and Mariano (1995) and West (1996) test (hereafter DMW). While DMW test is asymptotically normal when comparing non-nested models, West (1996) shows that when the models are nested, the traditional DMW test fails to have an asymptotically normal distribution, i.e. when the null model is a random walk or random walk with drift and is compared to other models including additional predictors. There is a vast literature available on tests of equal forecast accuracy in non-nested models: Diebold and Mariano(1995), West (1996) and Clark and McCracken (2001). Other authors have focused on encompassing tests, as described in Ericsson (1992), West (1996, 2001a, b), West and McCracken (1998), Harvey et al. (1998), McCracken (2000), Clark and McCracken (2001), Chao et al. (2001), Corradi et al. (2001) and Clark and West (2006, 2007).

Ashley et al. (1980) introduced the method of comparing out-of-sample forecasts from nested models stating that it would be best to use post-sample forecast tests. Additionally, since the work of Meese and Rogoff (1983), out-of-sample forecasts have been used to determine predictive ability of one variable for another. On the one hand, in-sample analyses estimate the model using all the available data, which introduces the risk of overfitting. On the other hand, out-of-sample analyses split the data into a purely estimation window which is used for parameter estimation and a forecasting evaluation and parameter update subset. Out-of-sample analysis is usually preferred as some literature shows its ability to reduce overfitting (see Clark (2004)). Out-of-sample analyses are also preferred because they better reflect the information available to the forecaster in real time. It is then crucial to develop better technology to perform evaluation of nested model hypothesis using out-of-sample techniques

After West (1996) discovery, there has been a significant effort to have an equal MSPE test under nested environments. In an effort to solve the problem identified for DMW test, different authors have proposed encompassing tests or modifications to previous tests. The problem with the currently available tests is that usually they are not asymptotically normal, depend on nuisance parameters, require boostrapping, are ill-sized for longer horizons or there is no proof of their asymptotically normal distribution when using multi-step-ahead forecasts. In a number of different issues, these tests do not behave well and it is harder to work on predictive analysis.

In the forecasting econometrics literature, there is an increasing demand for the availability of asymptotically normal tests of predictability for nested models. One problem that arises when doing forecast evaluation with encompassing tests is that central limit theorem holds only under special conditions. West and McCracken (1998) have shown that Chong and Hendry (1986) is asymptotically normal when applied to nested models in some cases. Clark and McCracken (2001) and Clark and West (2006) have also approached this problem in the past with limited results. West (2006) points out that it is required to work in the development of an asymptotically normal test for nested models, as currently only special case results are available in the literature. This is particularly important as asymptotically normal critical values not only allow us to build confidence intervals in an easy way, but also allow us to rely on both rolling and recursive windows when dealing with out-of-sample analysis.

Our work will focus in an attempt to fill this gap. We propose an asymptotically normal test for nested models for multi-step ahead forecast horizons in a wide setting that does not depend on nuisance parameters. Although we will proceed here without theoretical proof, we will rely on a vast set of simulations to theck the adequacy of our approach.

Our main approach here will be to compare how a modified ENC-T test performs against Clark and West (2007). The main goal of this thesis is to use simulations to see how this new test behaves relative to other available competitors.

Size and power of a modified ENC-T test (CWN), in comparison with CW, will be presented as part of the simulations results. This approach is based on the work done by Pincheira and West (2016), where the authors compare CW and DMW tests.

Simulation results presented one important key finding when using persistent Data Generating Processes (DGPs), which shows that while CW test tends to be oversized, the proposed modification leads to a near-nominal size test. We set a 10% significance threshold for all our tests. The main findings of our asymptotic simulations show that CW is slightly undersized but tolerable at short horizons ($h = \{1, 2, 3\}$) with an average size of 8% (all simulations have been performed using a nominal size of 10% and additional simulations are available in the appendix with a nominal size of 5%). At long horizons, CW is slightly oversized with an average size of 10.6%. In the case of persistent DGPs, CW gets oversized quickly with an average of 20%. In comparison, CWN is always slightly oversized and stable across all horizons and DGPs. In terms of power, CW is always better than CWN, sometimes dramatically and sometimes moderately. Similar results were obtained for finite sample simulations.

This work is organized as follows. Section 2 outlines ENC-T test and presents a new test called CWN while presenting the general econometric framework that will be used to run the simulations. Section 3 describes our DGPs and the simulation setup. Section 4 presents simulation evidence showing size, power and size adjusted power of both tests and results for the normality tests conducted after the simulations in order to conclude that the proposed test tends to be asymptotically normal under the null. Section 5 illustrates the use of both tests in an empirical application. Section 6 concludes.

2 Econometric Setup and Evaluation Framework

The linear econometric setup considers nested specifications for a scalar dependent variable y_{t+1} as follows:

$$\begin{aligned} y_{t+1} &= X_t'\beta_1 + e_{1t+1} && \text{(model 1: null model)} \\ y_{t+1} &= X_t'\beta_2 + Z_t'\gamma + e_{2t+1} && \text{(model 2: alternative model)} \end{aligned}$$

where e_{1t+1} and e_{2t+1} are both zero mean martingale difference processes, meaning that $E(e_{it+1}|F_t) = 0$ for $i = 1, 2$. F_t represents the sigma-field generated by current and past values of X_t, Z_t and e_{it} for $i = 1, 2$.

Under the null:

$$\begin{aligned} \gamma &= 0 \\ e_{1t+1} &= e_{2t+1} \\ \beta_1 &= \beta_2 \end{aligned}$$

In this case, the alternative model reduces to the null model, and it means Z_t' in the alternative model does not help prediction.
Under the alternative, $\gamma \neq 0$ and $e_{1t+1} \neq e_{2t+1}$. β_1 might be different from β_2. If Z_t is orthogonal to X_t, then $\beta_1 = \beta_2$ even if $\gamma \neq 0$. In this case, Z_t' does help prediction.

Clark and McCracken (2001) propose a test of encompassing called ENC-T which takes this form:

$$ENC-T = \sqrt{P}\frac{P^{-1}\sum_{t=R}^{T}\widehat{e}_{1t+1}(\widehat{e}_{1t+1} - \widehat{e}_{2t+1})}{\sqrt{\widehat{\sigma}^2}}$$

Where $\widehat{\sigma}^2$ is the usual variance estimator for $(\widehat{e}_{1t+1}(\widehat{e}_{1t+1} - \widehat{e}_{2t+1}))$ and σ^2 is the population variance for $(e_{1t+1}(e_{1t+1} - e_{2t+1}))$.

The asymptotic distribution of the ENC-T test is asymptotically normal in non-nested environments. This conclusion is similar as the one shown by West (1996). However, this proof does not hold for nested models. West (1996) results require the long run variance $\sigma^{*2} > 0$. This is violated with nested models as the null hypothesis implies the population errors of the forecasting models are identically the same. This implies that σ^2 is equal to 0. The problem is also explained by West (2006) as under the null hypothesis $e_{1t+1} = e_{2t+1}$, which impacts normality because the full rank condition is not met as it requires $(e_{1t+1}(e_{1t+1} - e_{2t+1}))$ to be non zero.

Under the null:

$$\begin{aligned} e_{1t+1} &= e_{2t+1} \\ e_{1t+1} - e_{2t+1} &= 0 \\ e_{1t+1}(e_{1t+1} - e_{2t+1}) &= 0 \end{aligned}$$

Immediately, we can see that rank condition is not met and $\sigma^2 = 0$.

Under the alternative:

$$\begin{aligned} e_{1t+1} &= y_{t+1} - X^{'}\beta_1 \\ e_{2t+1} &= y_{t+1} - X^{'}\beta_2 - Z^{'}\gamma \\ e_{1t+1} - e_{2t+1} &= Z^{'}\gamma + X^{'}(\beta_2 - \beta_1) \\ e_{1t+1} &= e_{2t+1} + Z^{'}\gamma + X^{'}(\beta_2 - \beta_1) \end{aligned}$$

Therefore:

$$\mathbb{E}[e_{1t+1}(e_{1t+1} - e_{2t+1})] = \mathbb{E}[Z^{'}\gamma + X^{'}(\beta_2 - \beta_1)]^2 > 0$$

Notice that this justifies the use of a one-sided test.

The main idea in this case is to modify the test to make it asymptotically normal under the null. This requires the introduction of a new random variable and to use the long-run variance instead, i.e.:

$$\theta_t \;\; i.i.d.$$

It is important to note that θ_t is also independent fromX, Z and e_{it} for $i = 1, 2$.

In this case, we will use the following specification for the new random variable:

$$\theta_t \sim N(1, \phi^2) \;\; i.i.d.$$

We can modify ENC-T now to follow the new specification:

$$CWN - T = \sqrt{P}\frac{P^{-1}\sum_{t=R}^{T}\widehat{e}_{1t+1}(\widehat{e}_{1t+1} - \widehat{e}_{2t+1}\theta_t)}{\sqrt{\widehat{V}^2}}$$

We will now assume $\widehat{V}^2$ to be an estimate of the long run variance of $(\widehat{e}_{1t+1}(\widehat{e}_{1t+1} - \widehat{e}_{2t+1}\theta_t))$.

Under the null we have:

$$e_{1t+1} = e_{2t+1}$$

Therefore:

$$
\begin{aligned}
\mathbb{E}[e_{1t+1}(e_{1t+1} - e_{2t+1}\theta_t)] &= \mathbb{E}[e_{1t+1}(e_{1t+1} - e_{1t+1}\theta_t)] \\
&= \mathbb{E}[e_{1t+1}^2(1 - \theta_t)] \\
&= \mathbb{E}[e_{1t+1}^2]\mathbb{E}[(1 - \theta_t)] \\
&= \mathbb{E}[e_{1t+1}^2] * 0 \\
&= 0
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
Var[e_{1t+1}(e_{1t+1} - e_{2t+1}\theta_t)] &= \mathbb{E}[e_{1t+1}^2(1 - \theta_t)^2] \\
&= \mathbb{E}[e_{1t+1}^2]\mathbb{E}[1 - \theta_t]^2 \\
&= \mathbb{E}[e_{1t+1}^2]\mathbb{E}[\mathbb{E}[\theta_t] - \theta_t]^2 \\
&= \mathbb{E}[e_{1t+1}^2]Var[\theta_t] \\
&= \mathbb{E}[e_{1t+1}^2]\phi^2 > 0
\end{aligned}
$$

This transformation is very important as now under the null, the core statistic retains its random behavior and the full rank condition is met.

Now, under the alternative we have:

$$
\begin{aligned}
e_{1t+1} &= y_{t+1} - X^{'}\beta_1 \\
e_{1t+1} &= y_{t+1} - X^{'}\beta_2 - Z^{'}\gamma \\
e_{1t+1} - e_{2t+1} &= Z^{'}\gamma + X^{'}(\beta_2 - \beta_1) \\
e_{2t+1} &= e_{1t+1} - Z^{'}\gamma - X^{'}(\beta_2 - \beta_1)
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
\mathbb{E}[e_{1t+1}(e_{1t+1} - e_{2t+1}\theta)] &= \mathbb{E}[e_{1t+1}(e_{1t+1} - (e_{1t+1} - Z^{'}\gamma - X^{'}(\beta_2 - \beta_1))\theta)] \\
&= \mathbb{E}[e_{1t+1}^2 - e_{1t+1}^2\theta + e_{1t+1}(Z^{'}\gamma + X^{'}(\beta_2 - \beta_1))\theta] \\
&= \mathbb{E}[e_{1t+1}^2(1 - \theta)] + \mathbb{E}[e_{1t+1}(Z^{'}\gamma + X^{'}(\beta_2 - \beta_1))\theta] \\
&= \mathbb{E}[e_{1t+1}(Z^{'}\gamma + X^{'}(\beta_2 - \beta_1))\theta] \\
&= \mathbb{E}[e_{1t+1}(Z^{'}\gamma + X^{'}(\beta_2 - \beta_1))] \\
&= \mathbb{E}[Z^{'}\gamma + X^{'}(\beta_2 - \beta_1)]^2 > 0
\end{aligned}
$$

Now as before, under the alternative, the test will still be a one-sided test.

With this result, we can see that in this new scenario, the central limit theorem should hold under stationarity and ergodicity. However, while we do not offer a proof that the new test satisfies all the conditions required for the proof in West (1996), we will rely on simulations to check the adequacy of our test under different setups in the next sections. Simulations will include both asymptotic analysis and finite sample analysis.

3 Monte Carlo simulations

Six different DGPs were included in the simulations based on empirical work in asset pricing (DGPs 1 and 2), macroeconomics data (DGPs 3 and 4) using DGP 3 as a benchmark and DGP 4 as a comparison when the model is persistent, a model estimated using non-linear least squares estimation (DGP 5), and core inflation forecasting using monthly CPI data (DGP 6).

In all cases, DGPs 1 to 5 are based on work from Pincheira and West (2016) and DGP 6 is based on work from Pincheira (2017). All driving shocks from DGPs are i.i.d. normal.

Simulations were done using both rolling and recursive estimation windows, using a single value of initial regression sample size R and four values of P. Addtionally, for CWN test, 5 different variance values were used for $V(\theta) = \{0.01; 0.1; 0.25; 0.5; 1\}$. 2000 independent replications are being used for testing purposes.

In the case of asymptotic analysis simulations, an initial estimation window of $R = 500$ is used and we report results for $P = 400$ predictions. For finite sample analysis simulations, an initial estimation window of $R = 80$ is used and we report results for $P = 80$ predictions.

Results are being reported for a single value of P and a single value of $V(\theta)$ ($V(\theta) = 0.1$). Additional results are available in the appendix.

3.1 Experimental Design

DGPs 1-2 describe a case where the null is a white noise, we consider DGPs such as the ones used in Pincheira and West (2016), Clark and West (2006), Campbell (2001), Stambaugh (1999), Nelson and Kim (1993) and Mankiw and Shapiro (1986). The general setup is the following:

$$\begin{aligned}
\text{Null Model} &: \quad y_{t+1} = \varepsilon_{t+1} && (2)\\
\text{Alernative Model} &: \quad y_{t+1} = \alpha_y + \gamma r_t + \varepsilon_{t+1} && (3)\\
&: \quad r_{t+1} = \alpha_r + \varphi_1 r_t + \varphi_2 r_{t-1} + ... + \varphi_p r_{t-p} + v_{t+1}
\end{aligned}$$

In all our simulations, we use the following values:

$$\begin{aligned}
\alpha_y &= \alpha_r = \varphi_3 = ... = \varphi_p = 0\\
\mathsf{var}(\varepsilon_{t+1}) &= \sigma_e^2\\
\mathsf{var}(v_{t+1}) &= \sigma_v^2\\
corr(\varepsilon_{t+1}, v_{t+1}) &= \rho
\end{aligned}$$

Parameters are set as follows:

	φ_1	φ_2	σ_e^2	σ_v^2	ρ	γ, under H_0	γ, under H_A
DGP 1	1.19	-0.25	$(1.75)^2$	$(0.075)^2$	0	0	-2
DGP 2	0.5	0	$(0.06)^2$	$(0.06)^2$	-0.4	0	-0.9

In both DGPs, the null yields the "no change" forecast of 0 for all t and all forecasting horizons. Additionally, the alternative forecast for multistep horizons is obtained from a regression of r_{t+1} on its own lags and a constant.

The first parametrization, labeled DGP 1, is the same DGP1 used in Pincheira and West (2016) and it is based on estimates done against the exchange rate application considered in the empirical work available in Clark and West (2006). Therefore, y_{t+1} represents the monthly percentage change in a US dollar bilateral exchange rate and r_t is the corresponding interest differential. The model was calibrated from monthly data.

The second parametrization, labeled DGP 2, is the same DGP2 used in Pincheira and West (2016) and it is based on monthly returns in the copper price y_{t+1} and the Chilean Peso-Dollar exchange rate r_t.

DGPs 3-4: for DGPs calibrated to macroeconomic data, we consider two DGPs. DGP 3 is the same DGP 4 in Pincheira and West (2016) and the same DGP 2 in Clark and West (2007). This DGP process is based on models exploring the relationship between US GDP growth and the Federal Reserve Bank of Chicago's factor index of economic activity. DGP 4 is the same as DGP 3 but with strong persistence, but it needs to be reviewed with caution, as DGP4 has not been calibrated and it's just a modification performed against DGP 3 to illustrate how CWN test compares against CW test using multi-step ahead forecasts when the DGP is highly persistent. These DGPs take the following form:

$$\begin{aligned} \text{Null Model} &: \ y_{t+1} = \alpha_y + \delta y_t + \varepsilon_{t+1} && (4) \\ \text{Alernative Model} &: \ y_{t+1} = \alpha_y + \delta y_t + \gamma_1 r_t + \gamma_2 r_{t-1} + \ldots \gamma_p r_{t-p} + \varepsilon_{t+1} && (5) \\ &: \ r_{t+1} = \alpha_r + 0.804 r_t - 0.221 r_{t-1} + 0.226 r_{t-2} - 0.205 r_{t-3} + v_{t+1} \end{aligned}$$

In all our simulations, we use the following values:

$$\begin{aligned} \alpha_y &= 2.237 \\ \alpha_r &= \gamma_5 = \ldots = \gamma_p = 0 \\ \mathsf{var}(\varepsilon_{t+1}) &= \sigma_e^2 \\ \mathsf{var}(v_{t+1}) &= \sigma_v^2 \\ corr(\varepsilon_{t+1}, v_{t+1}) &= \rho \end{aligned}$$

γ_1, under H_0	γ_2, under H_0	γ_3, under H_0	γ_4, under H_0
0	0	0	0
γ_1, under H_A	γ_2, under H_A	γ_3, under H_A	γ_4, under H_A
3.363	−0.633	−0.377	−0.529

Parameters are set as follows:

	δ	σ_e^2	σ_v^2	ρ
DGP 3	0.261	10.505	0.366	0.528
DGP 4	0.95	10.505	0.366	0.528

DGP 5 is as follows:

$$
\begin{aligned}
\text{Null Model} \quad &: \quad \pi_{t+1} = \alpha + \varphi_\pi \pi_t + \varepsilon_{t+1} && (6)\\
&: \quad \varepsilon_{t+1} = u_{t+1} - \theta u_t - \tau u_{t-11} + \tau\theta u_{t-12} && (7)\\
\text{Alernative Model} \quad &: \quad \pi_{t+1} = \alpha + \varphi_\pi \pi_t + \gamma \pi_t^{core} + \varepsilon_{t+1} && (8)\\
&: \quad \varepsilon_{t+1} = u_{t+1} - \theta u_t - \tau u_{t-11} + \tau\theta u_{t-12} &&\\
&: \quad \pi_{t+1}^{core} = \delta + \omega_{t+1} && (9)\\
&: \quad \omega_{t+1} = \varphi_\omega \omega_t + \mu_{t+1} - b\mu_{t-11} && (10)
\end{aligned}
$$

This DGP needs to be estimated using nonlinear least squares as there is serial correlation between ε_{t+1} and ω_{t+1}. This DGP is calibrated as in Pincheira and West (2016) and Pincheira, Selaive and Nolazco (2016) , i.e.:

$$
\begin{aligned}
\alpha &= 0\\
\varphi_\pi &= 0.96\\
\theta &= -0.45\\
\tau &= 0.93\\
\delta &= 0\\
\varphi_\omega &= 0.99\\
b &= 0.93\\
\sigma_u^2 &= 0.075\\
\sigma_\mu^2 &= 0.011\\
cov(u_t, \mu_t) &= 0.007
\end{aligned}
$$

γ, under H_0	γ, under H_A
0	0.05

Finally, DGP 6 is DGP 3 from Pincheira (2017), based on recent work on predictive linkages between domestic and international inflation:

$$\text{Null Model} \quad : \quad \pi_{t+1}^{core} = \alpha_\pi + \varphi_\pi \pi_t^{core} + \varepsilon_{t+1} \tag{11}$$

$$\text{Alternative Model} \quad : \quad \pi_{t+1}^{core} = \alpha_\pi + \varphi_\pi \pi_t^{core} + \gamma_1 \pi_t^{CIIF} + \gamma_2 \pi_{t-1}^{CIIF} + \varepsilon_{t+1}$$

$$: \quad \pi_{t+1}^{CIIF} = \alpha_r + \varphi_1 \pi_t^{CIIF} + \varphi_2 \pi_{t-1}^{CIIF} + v_{t+1} \tag{12}$$

This DGP is calibrated to match in-sample estimates for monthly core inflation, i.e.:

$$\begin{aligned}
\alpha_\pi &= 0.15 \\
\alpha_r &= 0.05 \\
\varphi_\pi &= 0.9 \\
\varphi_1 &= 1.27 \\
\varphi_2 &= -0.3 \\
\sigma_e^2 &= 0.0625 \\
\sigma_v^2 &= 0.01 \\
cov(e_{t+1}, v_{t+1}) &= 0.005
\end{aligned}$$

γ_1, under H_0	γ_2, under H_0	γ_1, under H_A	γ_2, under H_A
0	0	0.51	−0.50

This DGP is highly persistent.

4 Simulation results

Results are presented for 9 horizons x 4 values of P x 5 variances for CWN test x 6 DGPs = 1080 simulations. As previously stated, results are being reported for a single value of the number of predictions P, one-sided tests of nominal size 10% and a single value of the parameter ϕ for the CWN test. In other exercises, we carried out simulations with different values of the variance ϕ to explore the behavior of the CWN test. In summary it seems to us that the CWN test is better sized but there is a decline in power. Tables 1-8 contain the results for asymptotic analyses while tables 9-16 contain the results for finite sample analyses.

Table 1 (Asymptotic Simulations)

Size, nominal 10% tests, DGPs 1 and 2

Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.066500	0.062500	0.108000	0.111500
h=2	0.063500	0.078500	0.111000	0.106500
h=3	0.056000	0.084500	0.109500	0.109000
h=6	0.086000	0.099000	0.111500	0.107000
h=9	0.081000	0.099500	0.108500	0.107000
h=12	0.088000	0.098500	0.110000	0.109500
h=18	0.097000	0.097000	0.112000	0.110500
h=24	0.087000	0.119000	0.105500	0.105500
h=36	0.095500	0.106500	0.111000	0.111000
Panel B: Recursive Regressions				
h=1	0.060000	0.069000	0.104000	0.107000
h=2	0.059000	0.073000	0.105000	0.105500
h=3	0.055500	0.091000	0.106500	0.106000
h=6	0.078500	0.114500	0.107500	0.106500
h=9	0.079000	0.102000	0.105500	0.106500
h=12	0.085500	0.105500	0.108000	0.107500
h=18	0.097000	0.104500	0.108000	0.108000
h=24	0.100000	0.103000	0.108500	0.106500
h=36	0.109000	0.111000	0.109500	0.109000

Note: The table presents size for both CW and CWN tests using asymptotic simulations. The following parameters were used to run the simulations: R=500,P=400,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 2 (Asymptotic Simulations)				
Size, nominal 10% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	0.081000	0.078500	0.102000	0.106000
h=2	0.070000	0.088500	0.096500	0.106500
h=3	0.068000	0.091000	0.095000	0.096000
h=6	0.087500	0.106500	0.100000	0.091000
h=9	0.106000	0.116000	0.101500	0.105000
h=12	0.111000	0.133000	0.099000	0.104500
h=18	0.108500	0.172500	0.100500	0.103000
h=24	0.113500	0.204500	0.097500	0.099000
h=36	0.115000	0.238000	0.099500	0.106500
Panel B: Recursive Regressions				
h=1	0.078000	0.086500	0.100500	0.102500
h=2	0.086500	0.085500	0.095000	0.109000
h=3	0.078500	0.085500	0.096500	0.100500
h=6	0.089500	0.105000	0.101500	0.086000
h=9	0.112500	0.113000	0.100500	0.101000
h=12	0.122000	0.128500	0.099500	0.106000
h=18	0.106000	0.177000	0.099000	0.101000
h=24	0.108000	0.209000	0.094500	0.097000
h=36	0.102500	0.251000	0.100000	0.103000

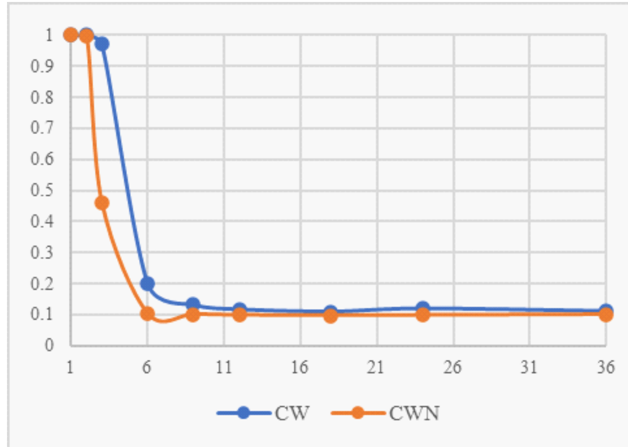
Note: The table presents size for both CW and CWN tests using asymptotic simulations. The following parameters were used to run the simulations: R=500,P=400,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 3 (Asymptotic Simulations)				
Size, nominal 10% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.087500	0.074500	0.111500	0.107500
h=2	0.093500	0.080000	0.105500	0.108000
h=3	0.104500	0.082000	0.094000	0.101500
h=6	0.207500	0.101500	0.102500	0.103000
h=9	0.306000	0.132000	0.104000	0.118500
h=12	0.338500	0.147500	0.106500	0.119000
h=18	0.307500	0.169000	0.105500	0.117000
h=24	0.265500	0.181500	0.106500	0.110000
h=36	0.199500	0.213500	0.097000	0.107000
Panel B: Recursive Regressions				
h=1	0.093500	0.062000	0.112000	0.104500
h=2	0.093500	0.074500	0.103500	0.107000
h=3	0.106500	0.074000	0.095000	0.105000
h=6	0.213500	0.103000	0.102000	0.094000
h=9	0.322000	0.131500	0.101000	0.116500
h=12	0.352000	0.150000	0.103500	0.115500
h=18	0.335000	0.168000	0.104000	0.112000
h=24	0.270500	0.180000	0.103000	0.111500
h=36	0.185500	0.212500	0.098500	0.109000

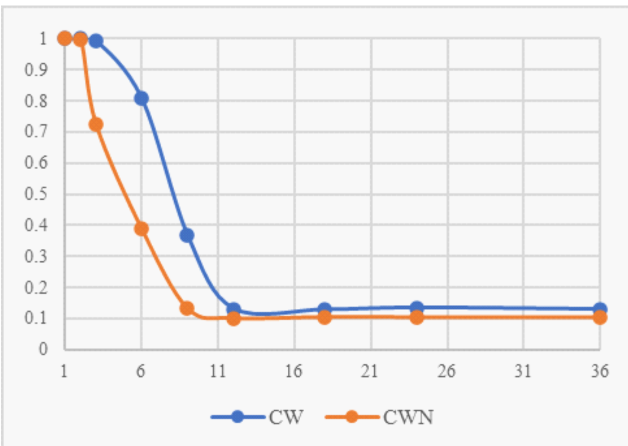
Note: The table presents size for both CW and CWN tests using asymptotic simulations. The following parameters were used to run the simulations: R=500,P=400,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Raw Power, Nominal 10% Tests - Rolling Regressions

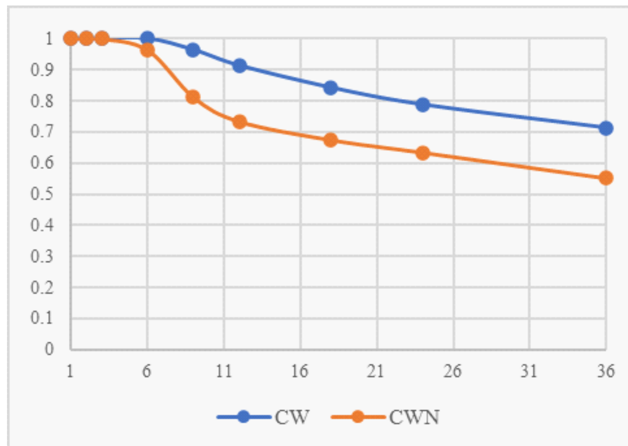
DGP 1



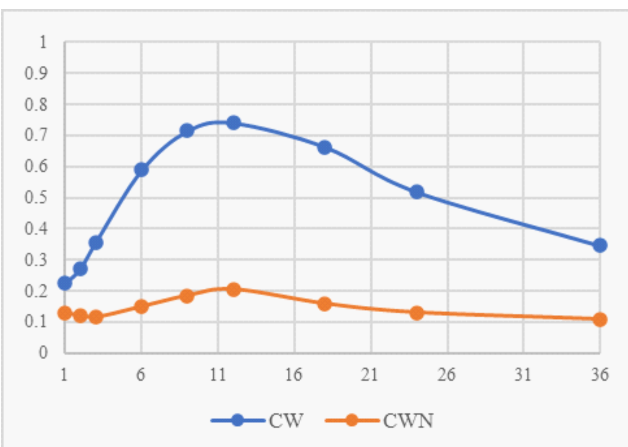
DGP 2



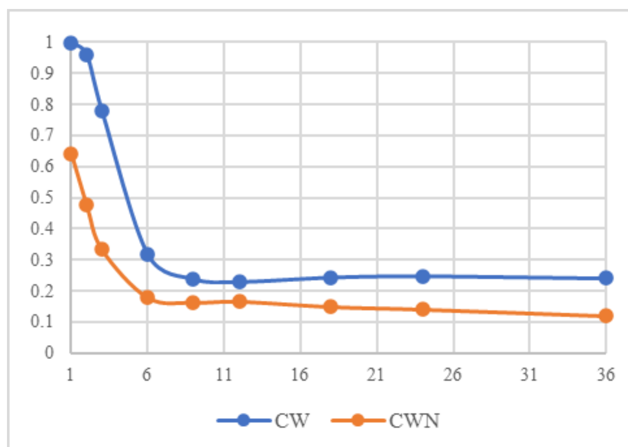
DGP 3



DGP 4



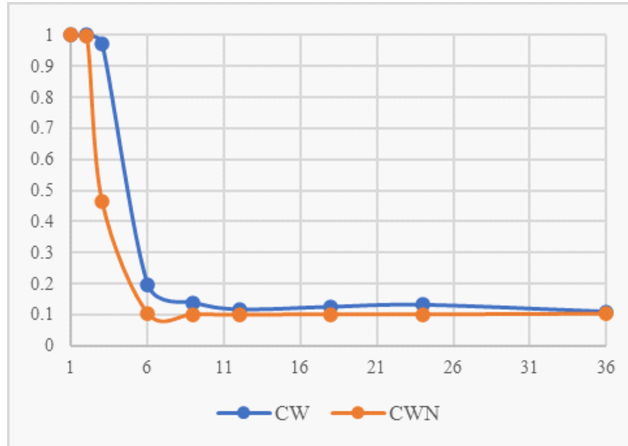
DGP 5



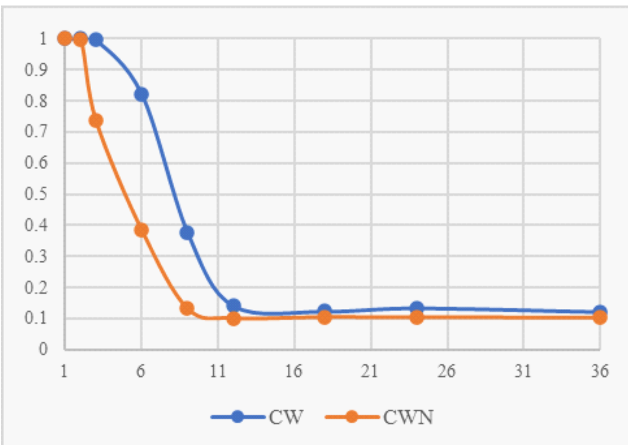
DGP 6

Raw Power, Nominal 10% Tests - Recursive Regressions

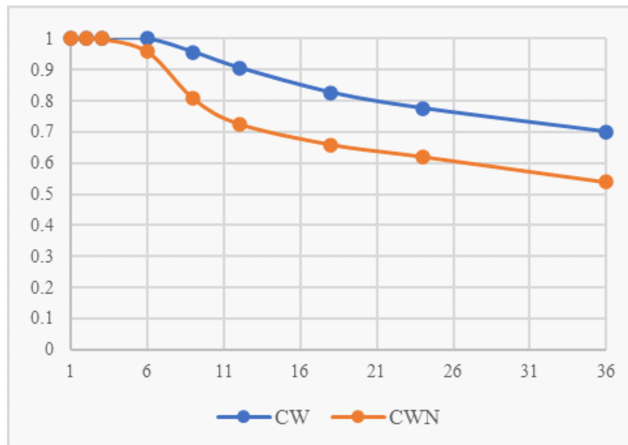
DGP 1



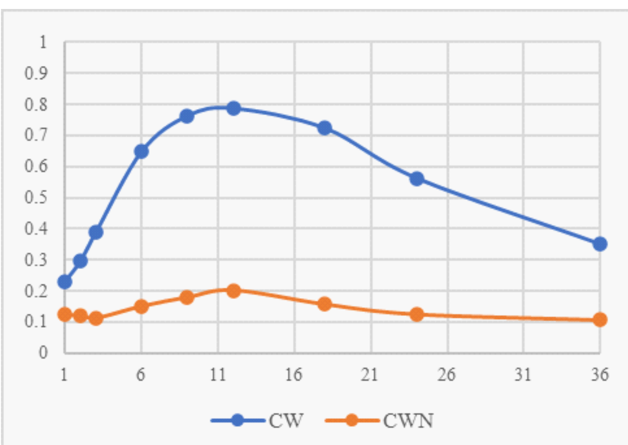
DGP 2



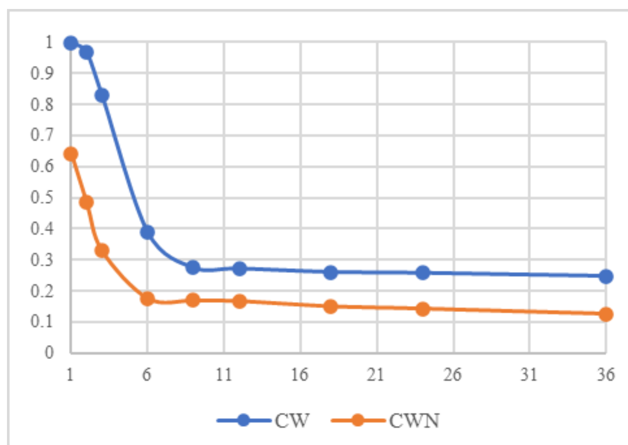
DGP 3



DGP 4



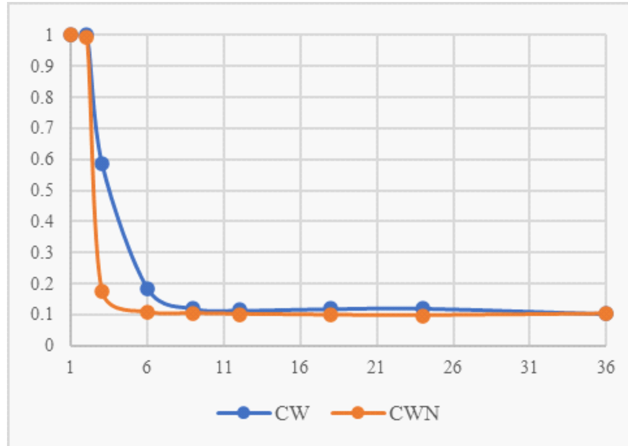
DGP 5



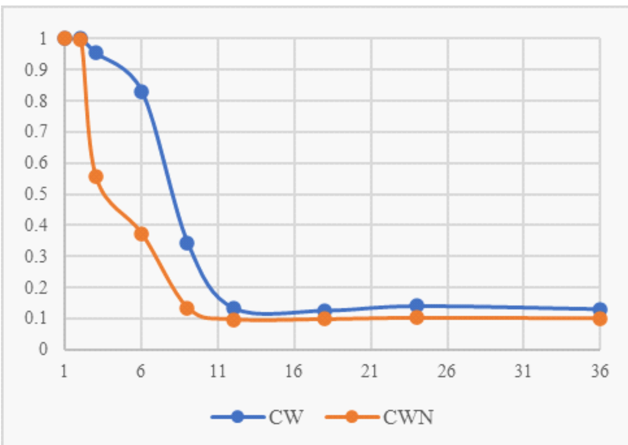
DGP 6

Size-Adjusted-Power, Nominal 10% Tests - Rolling Regressions

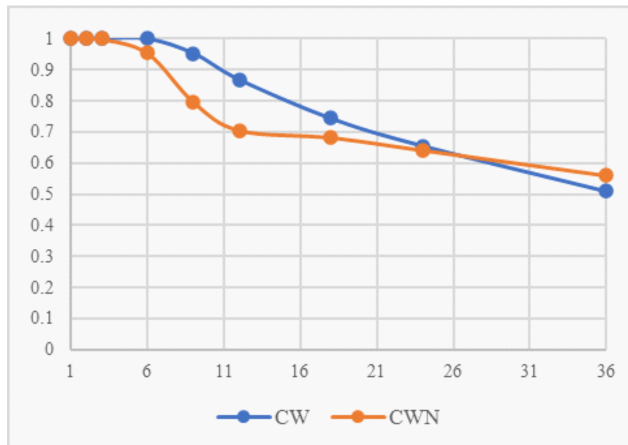
DGP 1



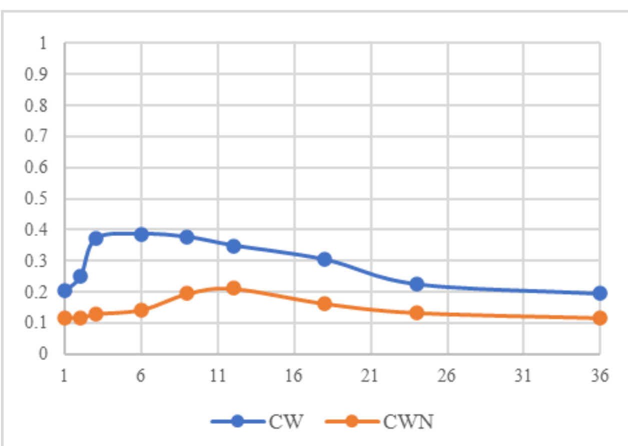
DGP 2



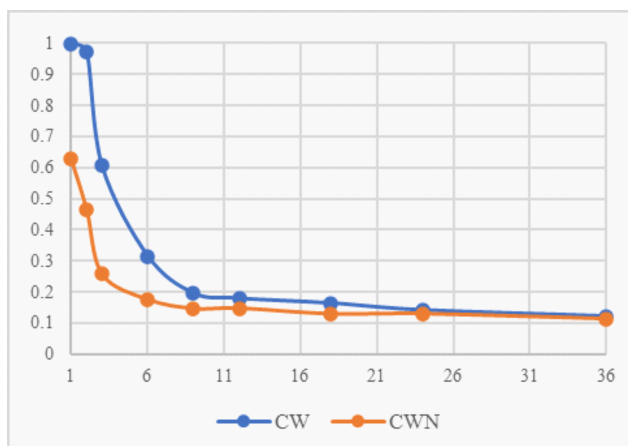
DGP 3



DGP 4

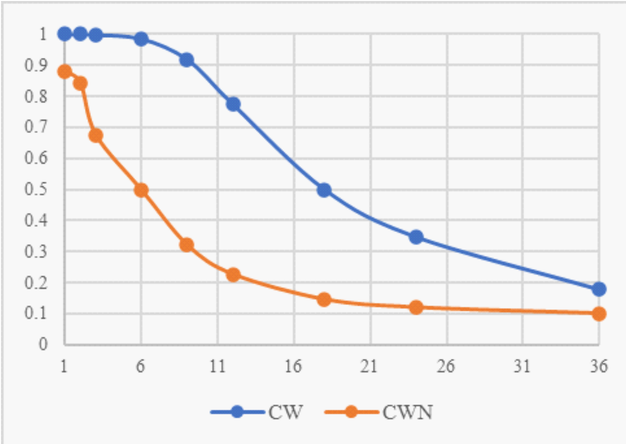


DGP 5

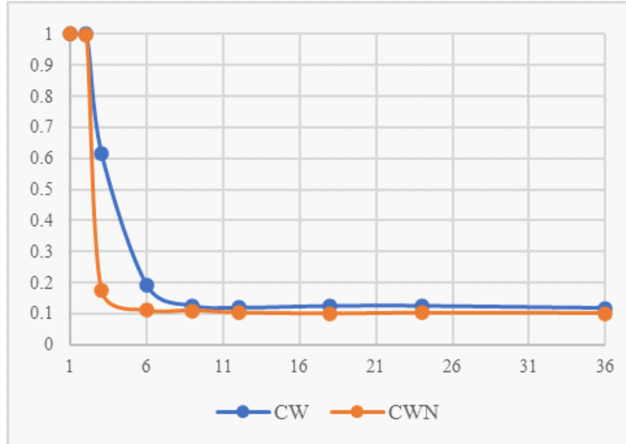


DGP 6

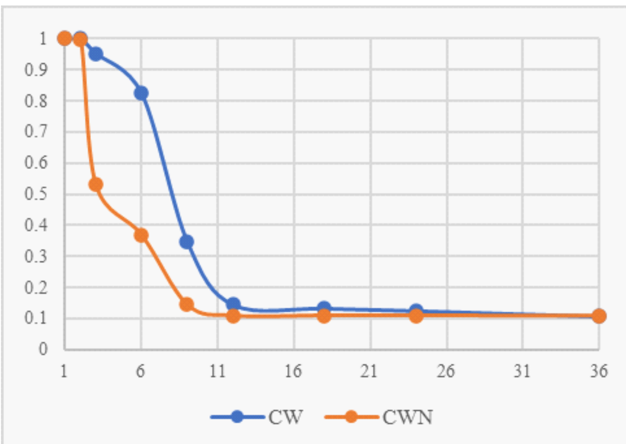
Size Adjusted Power, Nominal 10% Tests - Recursive Regressions



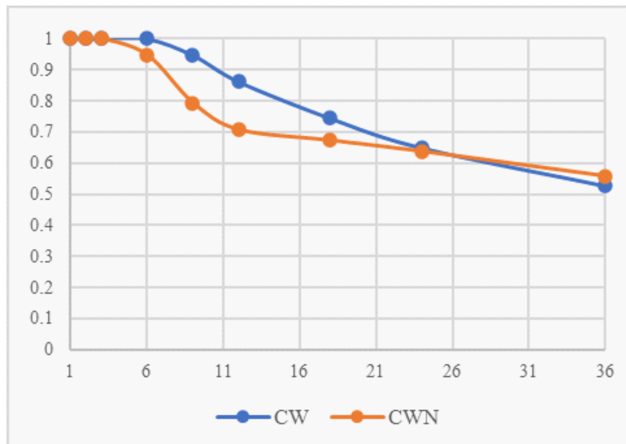
DGP 1



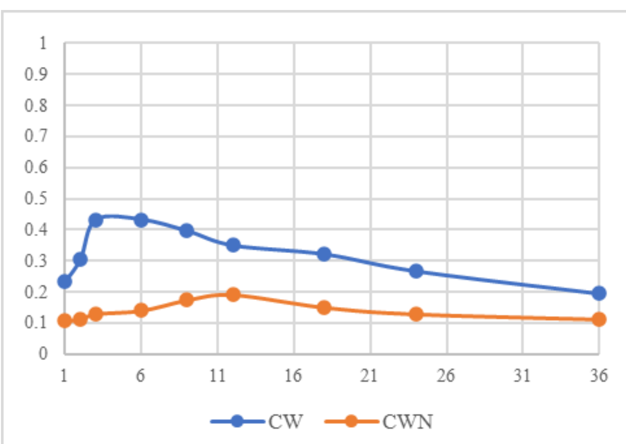
DGP 2



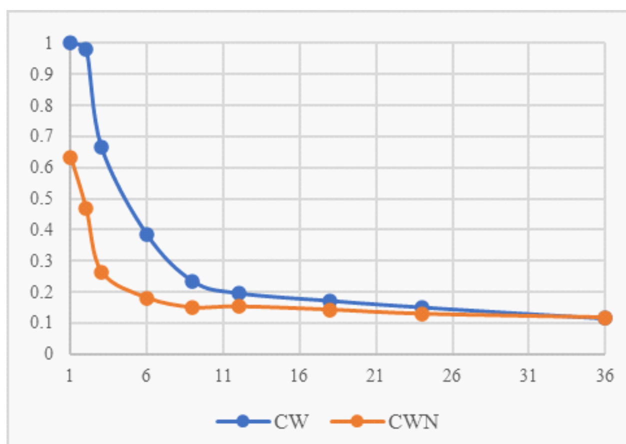
DGP 3



DGP 4



DGP 5



DGP 6

In summary, the main findings of our asymptotic simulations are that CW is slightly undersized but tolerable at short horizons ($h = \{1, 2, 3\}$) with an average size of 8%. At long horizons, CW is slightly oversized with an average size of 10.6%. In the case of persistent DGPs at long horizons, CW gets oversized quickly with an average of 20%. In comparison, CWN is always slightly oversized and stable across all horizons and DGPs. In terms of power, CW is always better than CWN, sometimes dramatically and sometimes moderately. Given CW tendency to be undersized and CWN tendency to be oversized, we see from Size-Adjusted-Power simulations that there are bigger differences in terms of power when it comes to CW, which is expected.

In order to compare different values of $V(\theta)$ and how its change impacts size and power for short, mid and long term horizons, we present the following tables for DGPs 3 and 4 to account for low and high persistence:

Table 4 (Asymptotic Simulations)

Size, nominal 10% tests, DGPs 3 and 4

Clark-West-Normal

(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	DGP 3			DGP 4			
Panel A: Rolling Regressions	h=1	h=12	h=24	h=1	h=12	h=24	Average
$V(\theta)=0$	0.081000	0.111000	0.113500	0.078500	0.133000	0.204500	0.120250
$V(\theta)=0.01$	0.096500	0.099000	0.097000	0.097000	0.106000	0.126000	0.103583
$V(\theta)=0.1$	0.102000	0.099000	0.097500	0.106000	0.104500	0.099000	0.101333
$V(\theta)=0.25$	0.103500	0.098500	0.097500	0.106500	0.104500	0.096500	0.101167
$V(\theta)=0.5$	0.101500	0.099000	0.097500	0.104000	0.104000	0.097500	0.100583
$V(\theta)=1$	0.102000	0.099000	0.097500	0.102500	0.104000	0.097000	0.100333
Average	0.097750	0.100917	0.100083	0.099083	0.109333	0.120083	
Panel B: Recursive Regressions							
$V(\theta)=0$	0.078000	0.122000	0.108000	0.086500	0.128500	0.209000	0.122000
$V(\theta)=0.01$	0.103000	0.100000	0.094500	0.100000	0.115000	0.119000	0.105250
$V(\theta)=0.1$	0.100500	0.095000	0.094500	0.102500	0.106000	0.097000	0.099250
$V(\theta)=0.25$	0.100000	0.099500	0.094500	0.101000	0.104000	0.098000	0.099500
$V(\theta)=0.5$	0.101500	0.100000	0.094500	0.103000	0.105500	0.096000	0.100083
$V(\theta)=1$	0.103000	0.100000	0.094500	0.102500	0.105500	0.097000	0.100417
Average	0.097667	0.252750	0.096750	0.099250	0.110750	0.119333	

Note: The table presents size variation for CWN test using different values for ϕ^2 with asymptotic simulations. DGPs 3 and 4 are compared to account for low and high persistence. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 5 (Asymptotic Simulations)							
Raw Power, nominal 10% tests, DGPs 3 and 4							
Clark-West-Normal							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	DGP 3			DGP 4			
Panel A: Rolling Regressions	h=1	h=12	h=24	h=1	h=12	h=24	Average
$V(\theta) = 0$	1.000000	0.141500	0.134500	1.000000	0.912500	0.787500	0.662667
$V(\theta) = 0.01$	1.000000	0.098500	0.104500	1.000000	0.892500	0.759000	0.642417
$V(\theta) = 0.1$	1.000000	0.101500	0.104000	1.000000	0.726000	0.620000	0.591917
$V(\theta) = 0.25$	1.000000	0.101500	0.104000	1.000000	0.559000	0.498500	0.543833
$V(\theta) = 0.5$	0.999500	0.101000	0.104000	1.000000	0.423500	0.384500	0.502083
$V(\theta) = 1$	0.985500	0.101000	0.104000	1.000000	0.314500	0.301500	0.467750
Average	0.997500	0.107500	0.109167	1.000000	0.638000	0.558500	
Panel B: Recursive Regressions							
$V(\theta) = 0$	1.000000	0.141500	0.134500	1.000000	0.906000	0.775500	0.659583
$V(\theta) = 0.01$	1.000000	0.098500	0.104500	1.000000	0.892500	0.759000	0.642417
$V(\theta) = 0.1$	1.000000	0.101500	0.104000	1.000000	0.726000	0.620000	0.591917
$V(\theta) = 0.25$	1.000000	0.101500	0.104000	1.000000	0.559000	0.498500	0.543833
$V(\theta) = 0.5$	0.999500	0.101000	0.104000	1.000000	0.423500	0.384500	0.502083
$V(\theta) = 1$	0.985500	0.101000	0.104000	1.000000	0.314000	0.292500	0.466167
Average	0.997500	0.107500	0.109167	1.000000	0.636833	0.555000	

Note: The table presents raw power variation for CWN test using different values for ϕ^2 with asymptotic simulations. DGPs 3 and 4 are compared to account for low and high persistence. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

In addition to the asymptotic analyses results, we present the finite sample analyses results:

Table 6 (Finite Sample Simulations)				
Size, nominal 10% tests, DGPs 1 and 2				
Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.081500	0.073000	0.122500	0.110000
h=2	0.078000	0.102500	0.119500	0.107500
h=3	0.081000	0.096500	0.118500	0.109000
h=6	0.096000	0.119000	0.112000	0.104500
h=9	0.112000	0.113000	0.116000	0.107000
h=12	0.101000	0.112000	0.118000	0.109000
h=18	0.115000	0.101000	0.109000	0.108000
h=24	0.111000	0.092000	0.114500	0.106000
h=36	0.118000	0.112000	0.109500	0.106000
Panel B: Recursive Regressions				
h=1	0.066500	0.075000	0.106500	0.105000
h=2	0.061500	0.089500	0.111000	0.104000
h=3	0.065500	0.095500	0.114500	0.103500
h=6	0.076500	0.115000	0.110000	0.102500
h=9	0.100000	0.115000	0.108500	0.104500
h=12	0.099500	0.108500	0.110000	0.109500
h=18	0.119500	0.111000	0.107000	0.106500
h=24	0.115000	0.115500	0.103000	0.104000
h=36	0.124000	0.114500	0.112000	0.110000

Note: The table presents size for both CW and CWN tests using finite sample simulations. The following parameters were used to run the simulations: R=80,P=80,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 7 (Finite Sample Simulations)				
Size, nominal 10% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	0.085500	0.086500	0.119000	0.120500
h=2	0.094000	0.089500	0.119500	0.120000
h=3	0.090500	0.097000	0.126000	0.119000
h=6	0.107500	0.111000	0.117000	0.119500
h=9	0.107500	0.134500	0.110000	0.123000
h=12	0.107000	0.159000	0.114000	0.123000
h=18	0.115000	0.223000	0.115500	0.134500
h=24	0.110500	0.230000	0.113000	0.145500
h=36	0.129500	0.243000	0.107000	0.148000
Panel B: Recursive Regressions				
h=1	0.092000	0.082500	0.098500	0.102000
h=2	0.087500	0.089500	0.106500	0.112500
h=3	0.088500	0.096000	0.106000	0.110500
h=6	0.088500	0.095000	0.106500	0.109500
h=9	0.115500	0.116500	0.111000	0.116500
h=12	0.115500	0.148500	0.108000	0.120000
h=18	0.128000	0.188500	0.105500	0.139000
h=24	0.141500	0.230000	0.112500	0.143500
h=36	0.144500	0.268000	0.113500	0.142500

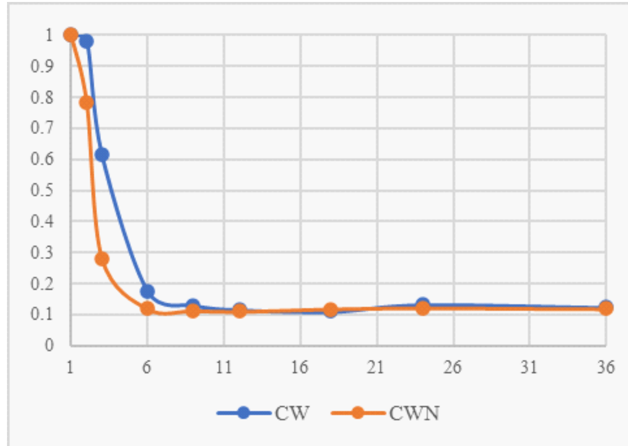
Note: The table presents size for both CW and CWN tests using finite sample simulations. The following parameters were used to run the simulations: R=80,P=80,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 8 (Finite Sample Simulations)				
Size, nominal 10% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ^2=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.111000	0.097000	0.127500	0.127000
h=2	0.128000	0.112500	0.128000	0.130500
h=3	0.144000	0.123000	0.130500	0.134000
h=6	0.212000	0.157000	0.151000	0.152000
h=9	0.303500	0.185500	0.168000	0.174000
h=12	0.335500	0.201500	0.172500	0.173000
h=18	0.300000	0.247500	0.151500	0.180000
h=24	0.286000	0.264000	0.152500	0.176000
h=36	0.277000	0.256500	0.145000	0.170500
Panel B: Recursive Regressions				
h=1	0.102000	0.085500	0.117000	0.120000
h=2	0.113000	0.105000	0.120500	0.127500
h=3	0.129000	0.113000	0.119500	0.121500
h=6	0.209000	0.140000	0.135500	0.140000
h=9	0.302000	0.170000	0.148000	0.156000
h=12	0.361500	0.192000	0.152500	0.162500
h=18	0.330000	0.242000	0.133500	0.181500
h=24	0.300500	0.256500	0.136500	0.174000
h=36	0.279500	0.259500	0.130500	0.156000

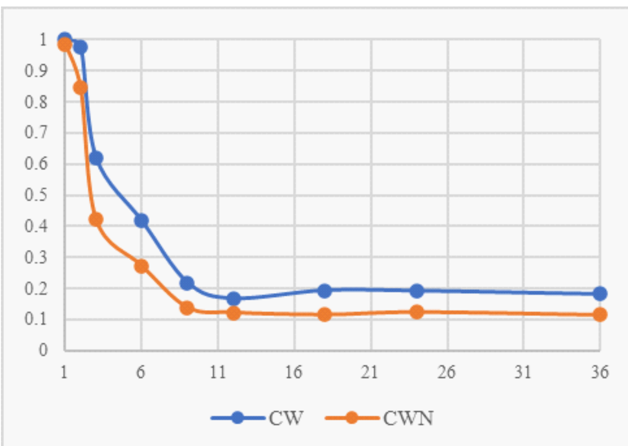
Note: The table presents size for both CW and CWN tests using finite sample simulations. The following parameters were used to run the simulations: R=80,P=80,ϕ^2=0.1. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Raw Power, Nominal 10% Tests - Rolling Regressions

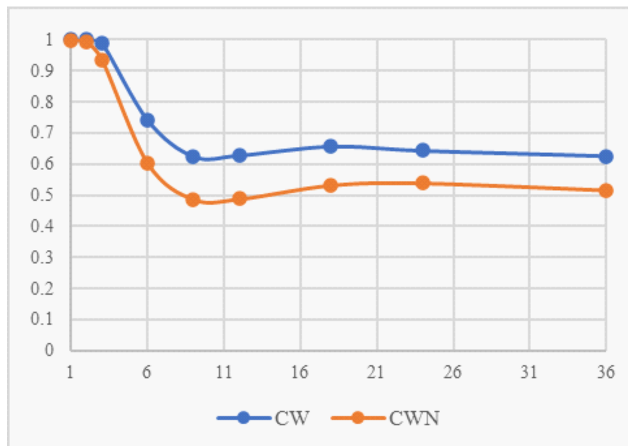
DGP 1



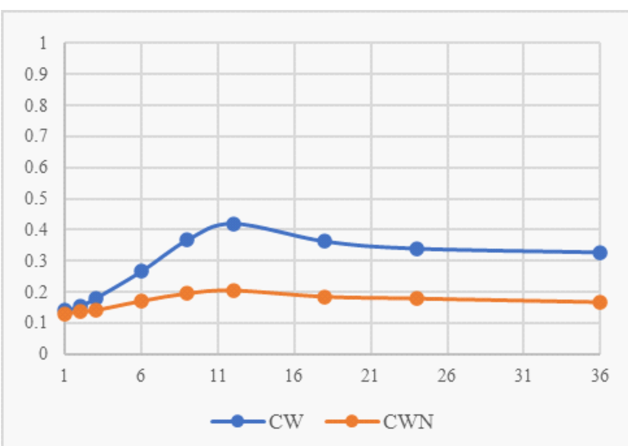
DGP 2



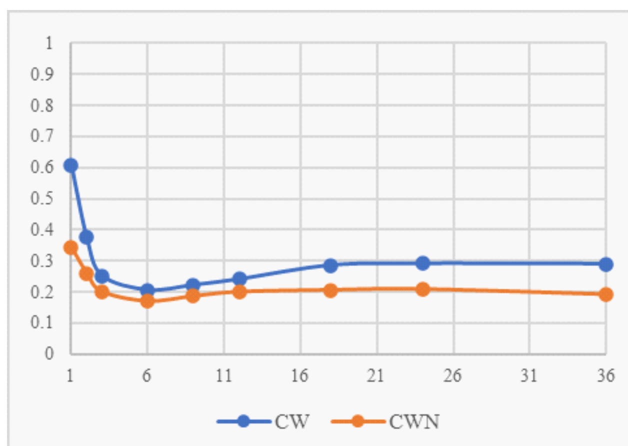
DGP 3



DGP 4



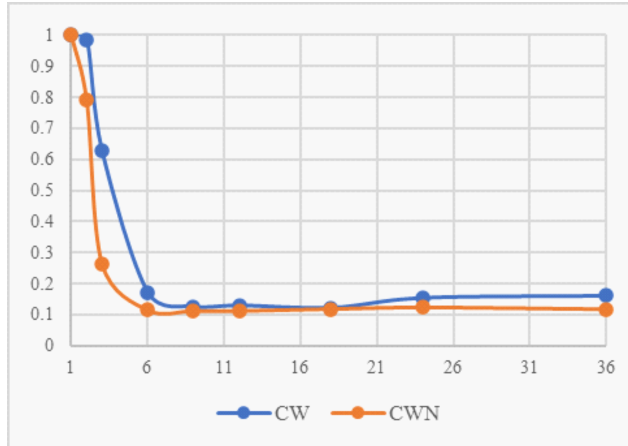
DGP 5



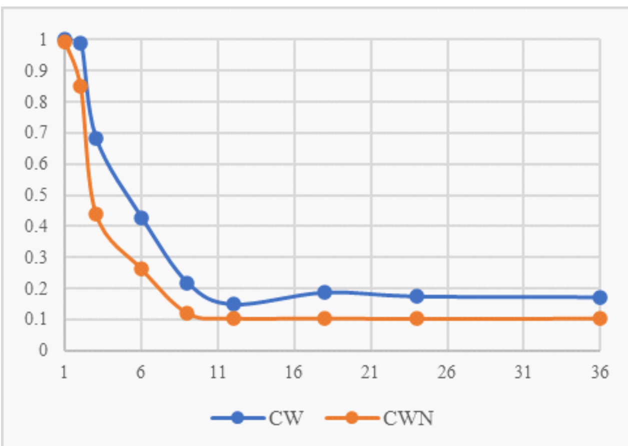
DGP 6

Raw Power, Nominal 10% Tests - Recursive Regressions

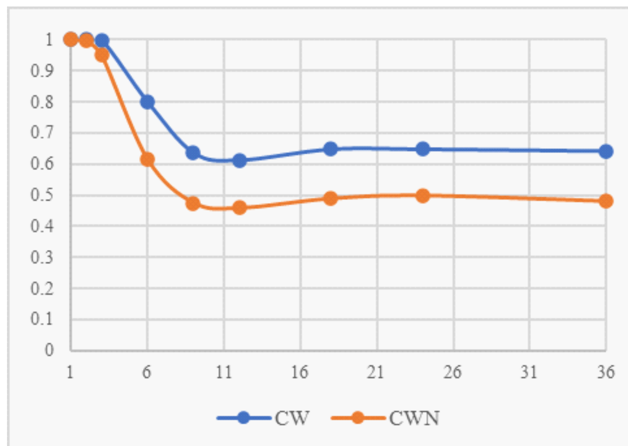
DGP 1



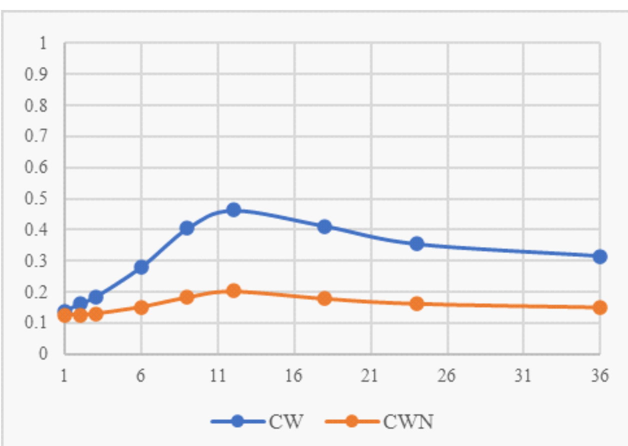
DGP 2



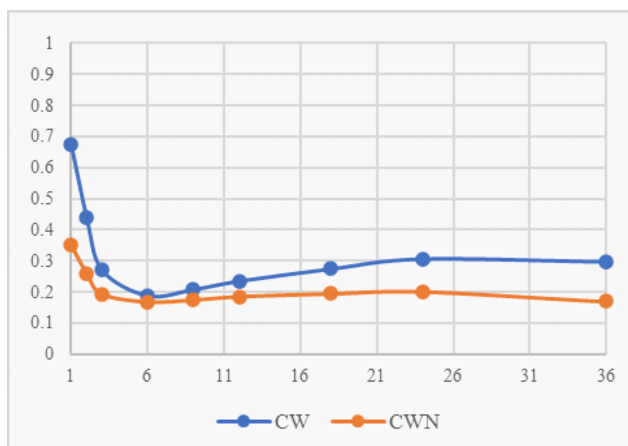
DGP 3



DGP 4



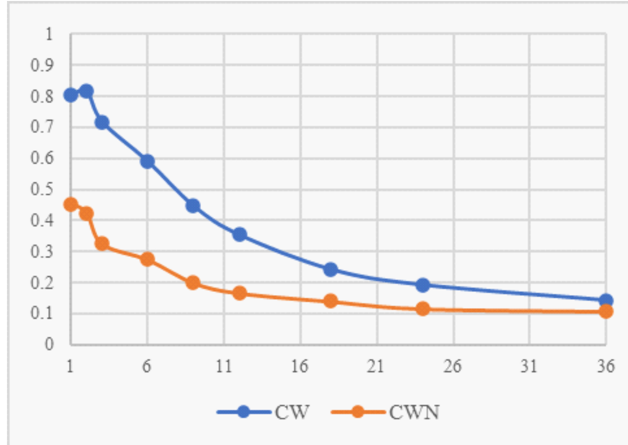
DGP 5



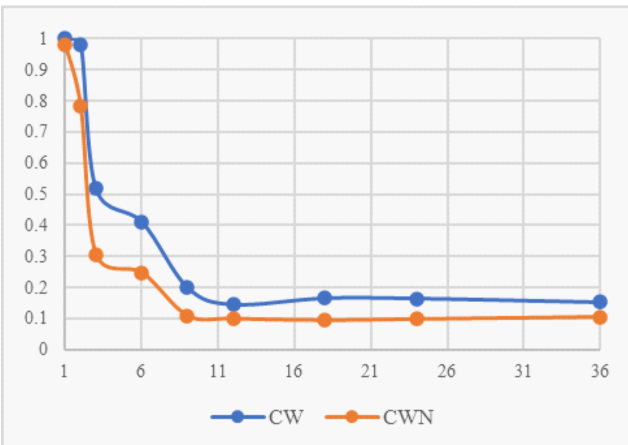
DGP 6

Size-Adjusted-Power, Nominal 10% Tests - Rolling Regressions

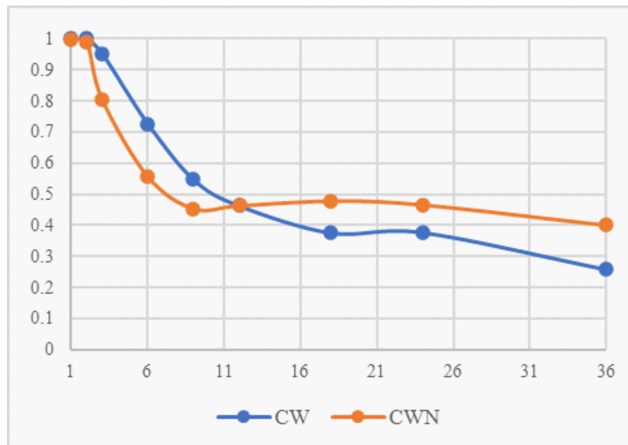
DGP 1



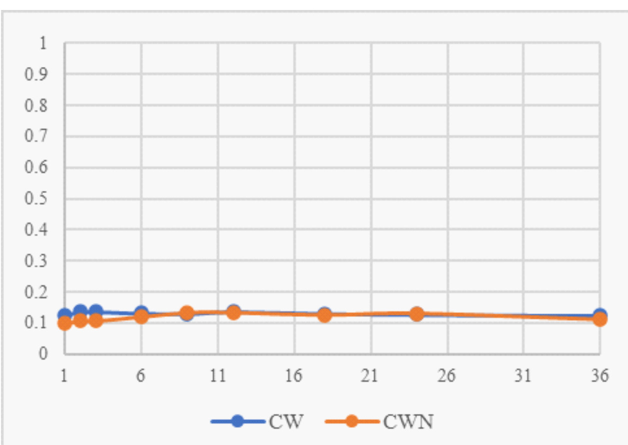
DGP 2



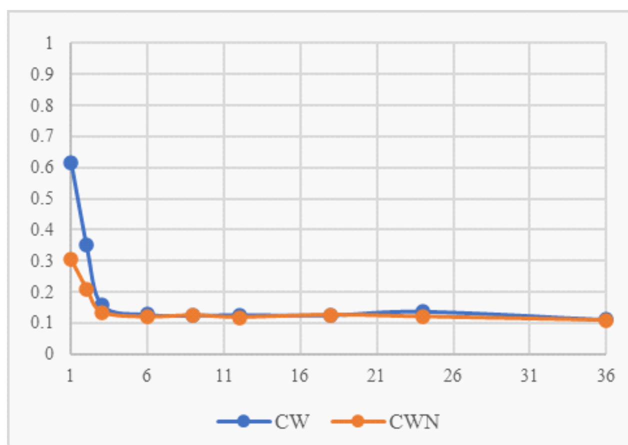
DGP 3



DGP 4

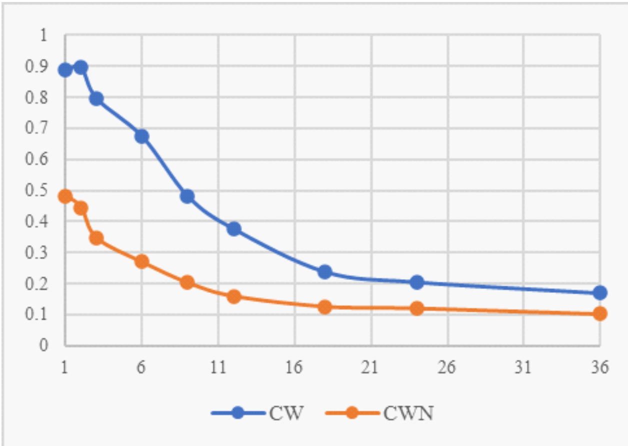


DGP 5

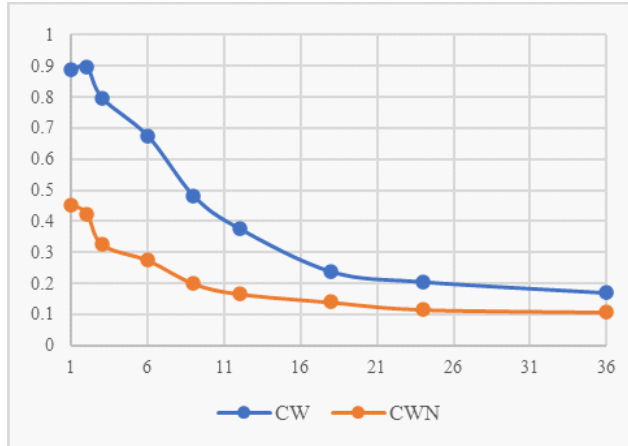


DGP 6

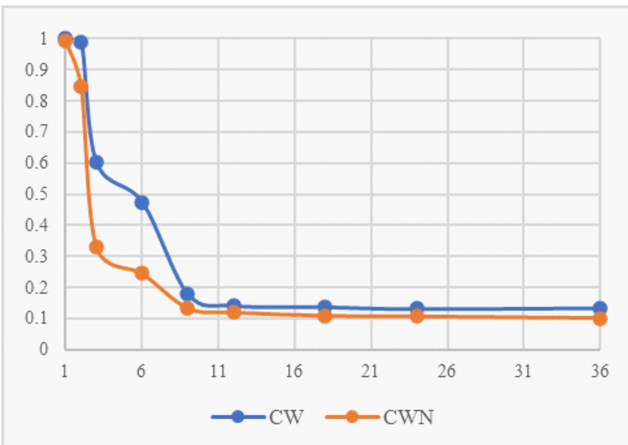
Size Adjusted Power, Nominal 10% Tests - Recursive Regressions



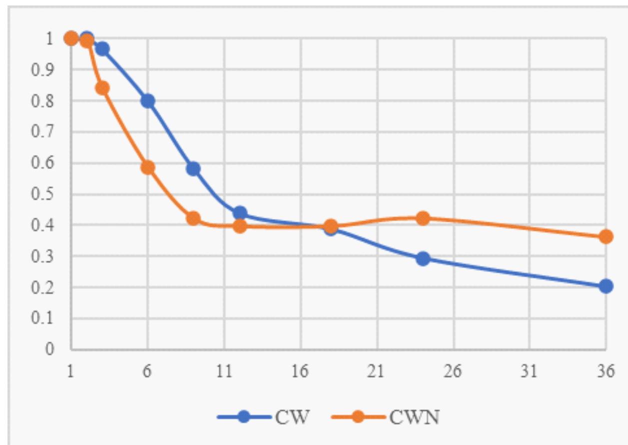
DGP 1



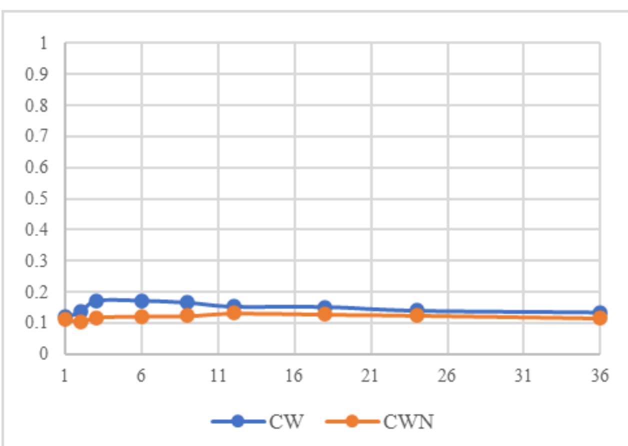
DGP 2



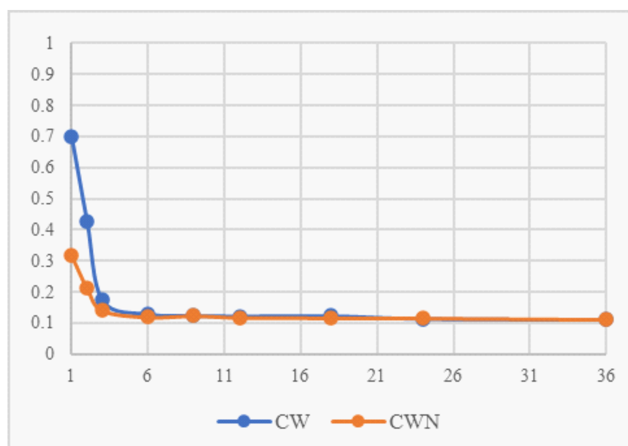
DGP 3



DGP 4



DGP 5



DGP 6

In summary, the main findings of our finite sample simulations are that CW is slightly undersized but tolerable at short horizons ($h = \{1, 2, 3\}$) with an average size of 8.5%. At long horizons, CW is slightly oversized with an average size of 11%. In the case of persistent DGPs with long horizons, CW gets oversized quickly with an average of 23%. In comparison, CWN is always slightly oversized and stable across all horizons and DGPs. In terms of power, CW is always better than CWN, sometimes dramatically and sometimes moderately. Given CW tendency to be undersized and CWN tendency to be oversized, we see from Size-Adjusted-Power simulations that there are bigger differences in terms of power when it comes to CW, which is expected.

In order to compare different values of $V(\theta)$ and how its change impacts size and power for short, mid and long term horizons, we present the following tables for DGPs 3 and 4 to account for low and high persistence:

Table 9 (Finite Sample Simulations)

Size, nominal 10% tests, DGPs 3 and 4

Clark-West-Normal

(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	DGP 3			DGP 4			
Panel A: Rolling Regressions	h=1	h=12	h=24	h=1	h=12	h=24	Average
$V(\theta) = 0$	0.085500	0.107000	0.110500	0.086500	0.159000	0.230000	0.129750
$V(\theta) = 0.01$	0.099500	0.104500	0.114000	0.096000	0.146000	0.177500	0.122917
$V(\theta) = 0.1$	0.119000	0.114000	0.113000	0.120500	0.123000	0.145500	0.122500
$V(\theta) = 0.25$	0.116500	0.112500	0.113000	0.116500	0.118500	0.129500	0.117750
$V(\theta) = 0.5$	0.116500	0.113000	0.112000	0.117500	0.117500	0.125000	0.116917
$V(\theta) = 1$	0.118500	0.113000	0.112000	0.116000	0.114500	0.118500	0.115417
Average	0.109250	0.110667	0.112417	0.108833	0.129750	0.154333	
Panel B: Recursive Regressions							
$V(\theta) = 0$	0.092000	0.115500	0.141500	0.082500	0.148500	0.230000	0.135000
$V(\theta) = 0.01$	0.091500	0.111000	0.111000	0.090000	0.133000	0.182000	0.119750
$V(\theta) = 0.1$	0.098500	0.108000	0.112500	0.102000	0.116500	0.143500	0.113500
$V(\theta) = 0.25$	0.098000	0.107500	0.112500	0.106000	0.115500	0.132500	0.112000
$V(\theta) = 0.5$	0.102000	0.107000	0.112500	0.104500	0.112000	0.122000	0.110000
$V(\theta) = 1$	0.100000	0.107500	0.112500	0.102000	0.105000	0.118000	0.107500
Average	0.097000	0.109417	0.117083	0.097833	0.121750	0.154667	

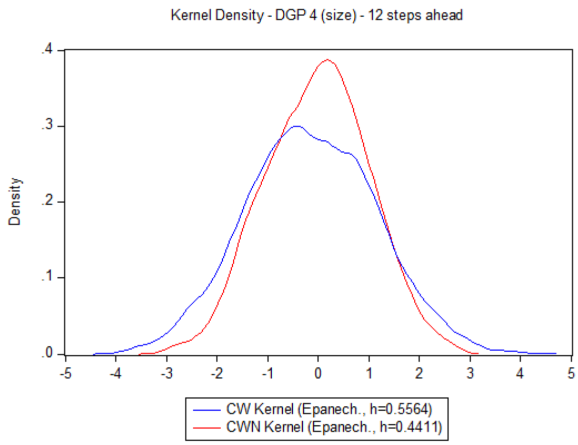
Note: The table presents size variation for CWN test using different values for ϕ^2 with finite sample simulations. DGPs 3 and 4 are compared to account for low and high persistence. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Table 10 (Finite Sample Simulations)							
Raw Power, nominal 10% tests, DGPs 3 and 4							
Clark-West-Normal							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	DGP 3			DGP 4			
Panel A: Rolling Regressions	h=1	h=12	h=24	h=1	h=12	h=24	Average
$V(\theta) = 0$	1.000000	0.168000	0.192500	1.000000	0.627500	0.643000	0.605167
$V(\theta) = 0.01$	1.000000	0.129500	0.133000	1.000000	0.609000	0.624500	0.582667
$V(\theta) = 0.1$	0.986500	0.122500	0.123500	0.999000	0.488000	0.538500	0.543000
$V(\theta) = 0.25$	0.941000	0.121000	0.120500	0.991000	0.398500	0.465000	0.506167
$V(\theta) = 0.5$	0.856500	0.120000	0.120000	0.973000	0.339000	0.402000	0.468417
$V(\theta) = 1$	0.715500	0.118500	0.120500	0.919500	0.275000	0.343000	0.415333
Average	0.916583	0.129917	0.135000	0.980417	0.456167	0.502667	
Panel B: Recursive Regressions							
$V(\theta) = 0$	1.000000	0.150000	0.175500	1.000000	0.613000	0.648000	0.597750
$V(\theta) = 0.01$	0.999500	0.115000	0.107000	1.000000	0.586500	0.617500	0.570917
$V(\theta) = 0.1$	0.991500	0.105500	0.104000	0.999500	0.458500	0.498500	0.526250
$V(\theta) = 0.25$	0.952000	0.104000	0.104500	0.995500	0.376000	0.422500	0.492417
$V(\theta) = 0.5$	0.859500	0.103000	0.104500	0.980000	0.312500	0.354000	0.452250
$V(\theta) = 1$	0.728500	0.103000	0.104500	0.905500	0.254000	0.294500	0.398333
Average	0.921833	0.113417	0.116667	0.980083	0.433417	0.472500	

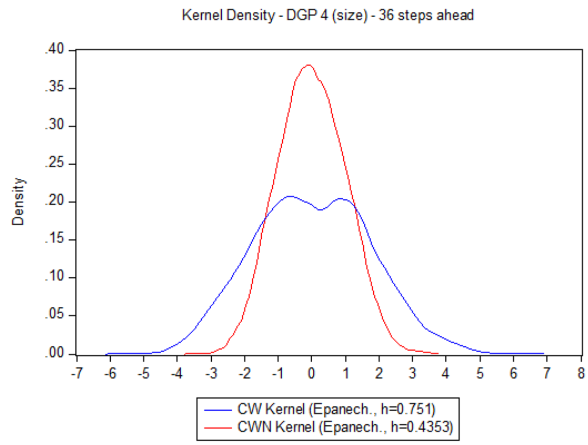
Note: The table presents raw power variation for CWN test using different values for ϕ^2 with finite sample simulations. DGPs 3 and 4 are compared to account for low and high persistence. Results are based on 2000 replications. Results for nominal 5% tests are available in the appendix.

Additionally, we present some findings regarding kernel density for asymptotic simulations. This is particulary important as Kolmogorov normality tests show that CWN-based results tend to be asymptotically normal while CW-based results are not. Some of the most important results are presented in the next charts and the rest of the results are available upon request. Tables 11 and 12 show Kolmogorov-Smirnov Statistic and estimated values for both mean and variance using $N(0,1)$ as the reference distribution for three selected horizons ($h = \{1, 12, 36\}$) to illustrate the results obtained when comparing CWN-based results with CW-based results. Finite sample results are available upon request.

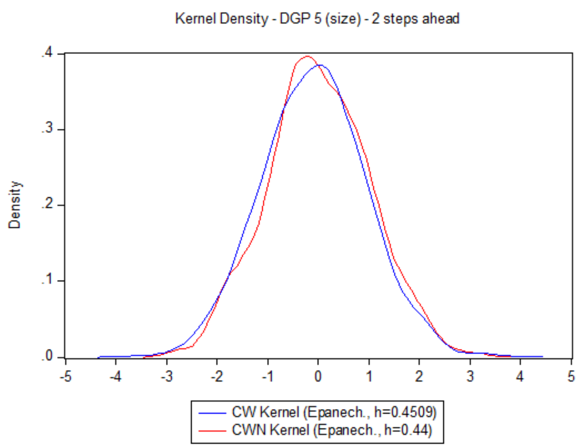
Kernel Density



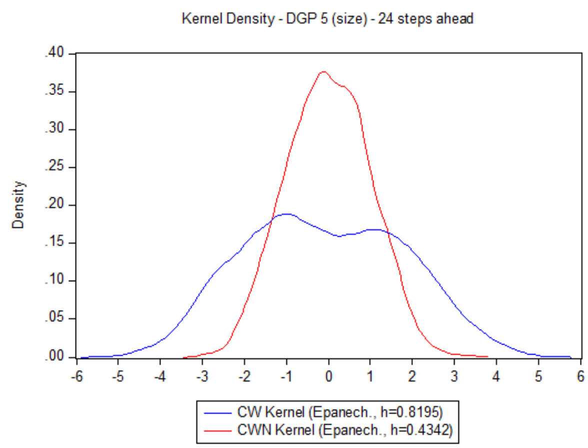
DGP 4 (12 steps ahead)



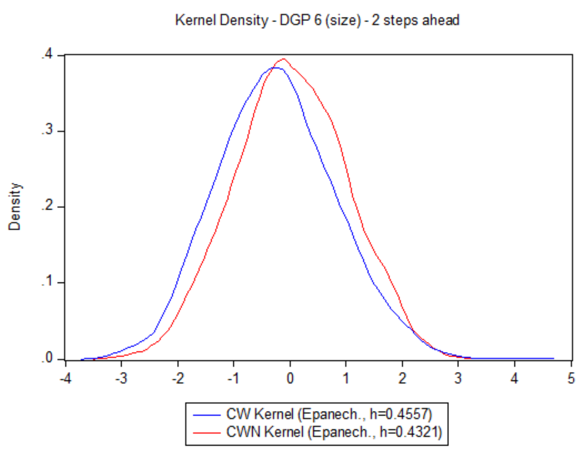
DGP 4 (36 steps ahead)



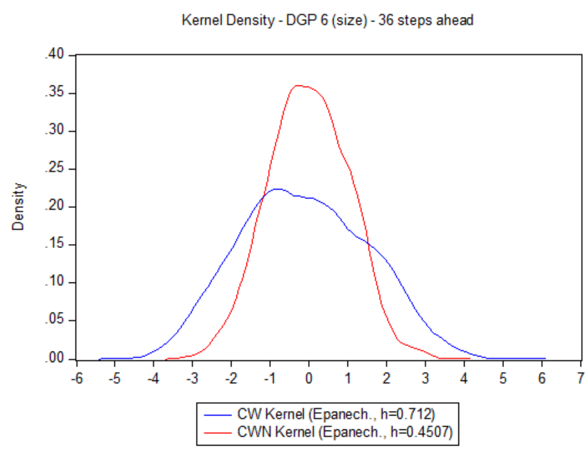
DGP 5 (2 steps ahead)



DGP 5 (24 steps ahead)



DGP 6 (2 steps ahead)



DGP 6 (36 steps ahead)

Table 11 (Asymptotic Simulations)

Kolmogorov Normality Test

	DGP 1			DGP2			DGP3		
Rolling Regressions	h=1	h=12	h=36	h=1	h=12	h=36	h=1	h=12	h=36
CW									
Kolmogorov (D)	0.09***	0.06***	0.02	0.14***	0.02	0.02	0.08***	0.03**	0.03*
Mean	-0.22	-0.12	-0.02	-0.29	-0.02	0.02	-0.18	0.03	0.00
Std	1.01	1.01	1.01	0.99	1.01	1.01	1.04	1.05	1.06
CWN (ϕ^2=0.1)									
Kolmogorov (D)	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.02	0.02
Mean	0.02	0.02	0.03	0.02	0.02	0.03	0.01	0.02	0.03
Std	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99

Note: * Means statistically significant at the 10% significant level. ** Means statistically significant at the 5% level and *** denotes statistically significance at the1% level

Table 12 (Asymptotic Simulations)

Kolmogorov Normality Test

	DGP 4			DGP5			DGP6		
Rolling Regressions	h=1	h=12	h=36	h=1	h=12	h=36	h=1	h=12	h=36
CW									
Kolmogorov (D)	0.08***	0.10***	0.15***	0.03**	0.37***	0.14***	0.10***	0.18***	0.17***
Mean	-0.18	-0.14	0.01	-0.05	-0.34	-0.09	-0.21	-0.29	-0.12
Std	1.04	1.28	1.72	1.00	2.89	1.51	1.02	1.40	1.63
CWN (ϕ^2=0.1)									
Kolmogorov (D)	0.01	0.02	0.02	0.02	0.03*	0.01	0.01	0.04***	0.02
Mean	0.02	-0.01	0.00	0.03	-0.02	-0.01	0.02	-0.03	0.00
Std	1.00	1.01	1.00	1.04	1.04	0.98	0.99	1.07	1.03

Note: * Means statistically significant at the 10% significant level. ** Means statistically significant at the 5% level and *** denotes statistically significance at the1% level

5 Empirical Illustration

We will explore predicting core inflation with an international inflation factor using nested models as described in DGP 6. There is recent literature available that has reviewed predictive linkages between domestic and international inflation. See Cicarelli and Mojon (2010), Morales-Arias and Moura (2013), Hakkio (2009), Pincheira and Gatty (2016), Pincheira and West (2016) and Medel et al (2016).
Let π_{it} be year-on-year domestic inflation rates in country i. Following the literature cited in the previous paragraph, we build an international inflation factor (IIF) as the single average of π_{it} measured using monthly CPI data, with i ranging over the current 31 OECD countries:

$$\pi_t^{IIF} = \frac{1}{31}\sum_{i=1}^{31} \pi_{it}$$

The following model is used:

$$\text{Null Model} \quad : \quad \pi_{t+1}^{core} = \alpha_\pi + \varphi_\pi \pi_t^{core} + \varepsilon_{t+1} \tag{13}$$

$$\begin{aligned} \text{Alternative Model} \quad : \quad & \pi_{t+1}^{core} = \alpha_\pi + \varphi_\pi \pi_t^{core} + \gamma_1 \pi_t^{IIF} + \gamma_2 \pi_{t-1}^{IIF} + \varepsilon_{t+1} \\ : \quad & \pi_{t+1}^{IIF} = \alpha_r + \varphi_1 \pi_t^{IIF} + \varphi_2 \pi_{t-1}^{IIF} + \upsilon_{t+1} \end{aligned} \tag{14}$$

We use data ranging from January 1995 to September 2019 (297 observations). For the out-of-sample analysis we estimate our models by OLS in recursive windows with an initial window length of 100 observations ($R = 100$, from January 1995 to April 2003). This means that our first one-step-ahead forecast is made for May 2003, while the last one is made for December 2019. We consider forecasts for the following horizons: $h = 1, 2, 3$ and 12 months ahead. We analyze if the IIF has the ability to predict inflation for different OCDE countries.

Table 13								
Forecasts on CPI core inflation								
Clark-West-Normal (ϕ=0.01)	CW				CWN			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Country	h=1	h=2	h=3	h=12	h=1	h=2	h=3	h=12
Austria	2.34***	1.84**	1.22	0.54	1.86**	1.71**	1.21	0.42
Belgium	2.48***	2.07**	1.83**	1.87**	1.82**	2.14**	1.66**	1.73**
Canada	1.92**	1.41*	0.81	1.55*	1.35*	0.78	0.21	0.51
Chile	2.59***	2.07**	1.99**	1.77**	2.66***	2.02**	1.86**	1.81**
Czech Republic	1.66**	1.32*	1.11	0.84	1.86**	1.42*	1.18	0.81
Germany	2.29**	1.63*	0.92	1.26	1.91**	1.56*	0.73	1.41*
Denmark	1.56*	0.77	0.23	-0.98	1.06	0.71	0.2	-1.25
Finland	2.56***	2.41***	2.31**	-0.03	1.5*	1.76**	1.4*	-0.58
France	2.09**	1.6*	1.06	0.83	1.96**	1.6*	0.91	0.74
Hungary	1.45*	0.93	0.7	-1.45	0.41	0.48	0.31	-1.64
Ireland	1.87**	1.66**	1.58*	1.27	1.86**	1.66**	1.51*	1.077
Iceland	3.31***	3.23***	3.1***	2.37***	3.38***	3.09***	2.96***	2.44***
Italy	2.41***	1.72**	1.27	0.88	2.29**	1.62*	1.25	0.75
Mexico	1.95**	1.86**	1.86**	0.04	1.99**	1.87**	1.86**	0.03
USA	2.26**	1.63*	1.3*	1.86**	2.13**	1.67**	1.27	1.82**
Switzerland	2.24**	1.80**	1.60*	0.83	2.22**	1.82**	1.61*	0.84
Spain	2.09**	1.36*	0.83	0.67	1.61*	1.15	0.65	0.56
United Kingdom	2.72***	2.55***	2.14**	2.07**	2.67***	2.73***	2.33***	2.27**
Greece	2.41***	1.53*	1.11	0.87	1.77**	1.54*	0.79	0.80
Israel	1.80**	0.89	0.31	-1.42	1.12	0.25	-0.17	-1.47
Japan	1.93**	1.30*	0.81	-0.71	1.72**	1.08	0.59	-0.87
Korea	3.94***	4.01***	3.32***	0.90	3.75***	3.71***	3.33***	0.77
Luxembourg	2.29**	1.45*	1.03	1.05	2.22**	1.34*	0.88	0.99
Netherlands	1.33*	1.23	1.13	0.67	0.33	1.12	1.06	0.33
Norway	2.12**	1.70**	1.02	-0.18	1.67**	1.62*	1.42*	-0.30
Poland	1.96**	1.44*	1.06	-0.13	1.89**	1.36*	0.98	-0.06
Portugal	2.03**	1.39*	1.05	0.98	0.40	1.03	0.80	0.71
Slovak Republic	2.21**	1.25	0.69	-1.09	1.40*	0.59	0.22	-1.23
Slovenia	2.20**	0.76	-0.24	-0.43	1.74**	0.25	-0.50	-0.75
Sweden	1.67**	1.48*	1.27	0.99	1.48*	1.22	1.07	0.59
Turkey	3.43***	3.00***	2.56***	1.57*	3.42***	2.96***	2.55***	1.54*

Note: * Means statistically significant at the 10% significant level. ** Means statistically significant at the 5% level and *** denotes statistically significance at the1% level.

Consistent with the simulation results in size and power, CWN rejects less frequently than CW does. In columns 6-9 in Table 13 there are 19 rejections at the 10% level, 31 rejections at the 5% level and 14 rejections at the 1% level using CWN. In columns 2-5 of that table there are 20 rejections at the 10% level, 35 rejections at the 5% level and 19 rejections at the 1% level using CW. It is also noticeable that for short horizons CW is able to detect predictability while CWN is unable, i.e. Hungary, Israel, Netherlands and Portulgal. However, the case of Canada is interesting, as CW rejects the null for $h = 12$ while CWN does not, and it would be convenient to incline for CWN result as CW is oversized for long horizons which increases the possibility of a false rejection, however, we need to remember there are two different effects here: the decline in power from CWN and that CW is oversized for longer horizons.

6 Concluding remarks

In this paper, we reviewed the behavior of one of the most widely and commonly used test (CW) for comparing forecasts from competing nested models. Additionally, we evaluated a new asymptotically-normal test and performed comparisons against this test using CW as a benchmark. The focus of interest was multistep ahead forecasts computed using the iterated method. Simulations were computed for 9 horizons x 4 values of P x 5 variances for CWN test x 6 DGPs = 1080 simulations

Our Monte Carlo simulations for both asymptotic and finite sample analyses show that CW is slightly undersized but tolerable at short horizons while it gets slightly oversized for long horizons. In the case of persistent DGPs, CW gets oversized quickly while CWN is usually well contained, slightly oversized in some cases but under tolerable levels and near-nominal size when using large sample size. While reviewing DGP 4 and DGP 6 results (highly persistent models), CW gets increasingly oversized as the horizon becomes larger while CWN remains near nominal-size for all horizons. Size Adjusted Power simulations show results consistent with the behavior of CW vs CWN, where we see bigger differences in favor of CW.

In terms of power, CW is always better than CWN, sometimes dramatically and sometimes moderately and we also see that power decreases as horizons become longer.

For multistep ahead forecasts in persistent series there is an important finding and benefit from using CWN test. While CWN is not always better in terms of size, it remains near nominal-size when CW is oversized. This is key when doing analyses for long horizons ($h = \{12, 24, 36\}$) in highly persistent models (DGP 6), as with CWN test the null hypothesis is better protected. In cases where CW is undersized, CWN remains near nominal-size but its power decreases.

Our results are relevant in light of existing literature, as we presented an asymptotically normal test for nested models that performs well for both persistent and non-persistent data generating processes, while previous work from different authors have focused in special cases only.

With CWN having a higher guarantee of normality (although we did not provide formal proof), it could result useful for additional applications or other tests. For example, Hubrich and West (2010) propose a joint test of predictability where they compare several models with a nested benchmark. This joint test is built with a vector of CW tests, and they assume each component is asymptotically normal, which might not be true as this assumption does not hold under general conditions. This test may benefit from using CWN test, as assuming normality using CW can lead to misleading results while switching to CWN test should address those issues with the power tradeoff already discussed.

An application in the context of inflation forecasts is consistent with our simulation results.

Future research could explore direct multi-step ahead forecasts using CWN test. Additional work could review a more detailed behavior of the CWN test and whether West (1996) conditions hold or not.

Appendix

Tables 1-9 illustrate size, raw power and size adjusted power for asymptotic nominal 5% tests.

Table 1				
Size, nominal 5% tests, DGPs 1 and 2				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.034000	0.062500	0.054000	0.051000
h=2	0.027500	0.078500	0.053000	0.052000
h=3	0.026000	0.084500	0.051500	0.051000
h=6	0.033000	0.099000	0.054500	0.054000
h=9	0.044000	0.099500	0.052500	0.052500
h=12	0.039000	0.098500	0.053000	0.054000
h=18	0.049000	0.097000	0.054000	0.055500
h=24	0.041500	0.119000	0.053000	0.053000
h=36	0.051000	0.106500	0.054500	0.054500
Panel B: Recursive Regressions				
h=1	0.031500	0.033500	0.051500	0.050000
h=2	0.027000	0.036000	0.053000	0.052000
h=3	0.025000	0.034500	0.053500	0.052000
h=6	0.035500	0.056500	0.055000	0.054500
h=9	0.037500	0.051000	0.053000	0.053000
h=12	0.037000	0.055000	0.054000	0.053500
h=18	0.050000	0.056000	0.052000	0.053500
h=24	0.042500	0.058000	0.055000	0.053500
h=36	0.053500	0.056000	0.053500	0.054000

Table 2				
Size, nominal 5% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	0.035500	0.040000	0.052000	0.050000
h=2	0.034500	0.044000	0.052000	0.052000
h=3	0.031000	0.045500	0.050000	0.042000
h=6	0.046000	0.056500	0.051000	0.045500
h=9	0.063000	0.067000	0.051000	0.051500
h=12	0.057000	0.079500	0.051500	0.048500
h=18	0.054000	0.121000	0.050500	0.046500
h=24	0.058500	0.148500	0.052500	0.045000
h=36	0.062000	0.174500	0.053000	0.052000
Panel B: Recursive Regressions				
h=1	0.044500	0.046500	0.049000	0.050000
h=2	0.041500	0.045500	0.048000	0.053000
h=3	0.039000	0.046500	0.046500	0.045500
h=6	0.042000	0.060500	0.050000	0.045500
h=9	0.058000	0.064500	0.051000	0.048000
h=12	0.066500	0.080500	0.050000	0.045500
h=18	0.056500	0.122000	0.049500	0.044500
h=24	0.056000	0.151000	0.052500	0.041500
h=36	0.056000	0.181000	0.053500	0.041000

Table 3

Size, nominal 5% tests, DGPs 5 and 6

Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.044500	0.033500	0.064500	0.054000
h=2	0.049500	0.041500	0.059500	0.056500
h=3	0.060000	0.046500	0.055000	0.045000
h=6	0.155500	0.065000	0.048000	0.054000
h=9	0.272000	0.081000	0.046500	0.058000
h=12	0.313000	0.095500	0.058000	0.061500
h=18	0.258000	0.123500	0.049000	0.059000
h=24	0.204500	0.129000	0.047000	0.057000
h=36	0.135500	0.163500	0.044000	0.049000
Panel B: Recursive Regressions				
h=1	0.044500	0.032000	0.066000	0.051500
h=2	0.048500	0.039000	0.059000	0.054500
h=3	0.061000	0.039500	0.051000	0.046000
h=6	0.162500	0.060500	0.049000	0.052000
h=9	0.293500	0.082500	0.048000	0.050500
h=12	0.332500	0.097500	0.053500	0.055000
h=18	0.302000	0.114500	0.043500	0.058500
h=24	0.219500	0.131500	0.046000	0.056500
h=36	0.130000	0.157000	0.041000	0.053000

Table 4				
Raw Power, nominal 5% tests, DGPs 1 and 2				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.999500	1.000000	0.804500	1.000000
h=2	0.998500	1.000000	0.754500	0.989500
h=3	0.997500	0.930500	0.668500	0.318000
h=6	0.961500	0.115000	0.376500	0.047500
h=9	0.823000	0.072500	0.210500	0.043500
h=12	0.611000	0.053500	0.141500	0.045000
h=18	0.318500	0.055500	0.082500	0.048500
h=24	0.200000	0.061500	0.061500	0.047000
h=36	0.108500	0.053500	0.052500	0.048500
Panel B: Recursive Regressions				
h=1	0.999000	1.000000	0.810000	1.000000
h=2	0.999000	1.000000	0.758000	0.992000
h=3	0.999000	0.933500	0.676000	0.316500
h=6	0.968000	0.113000	0.376000	0.047000
h=9	0.825500	0.076500	0.209500	0.042500
h=12	0.626500	0.056500	0.138500	0.044500
h=18	0.321000	0.065500	0.080000	0.047000
h=24	0.206500	0.074000	0.060500	0.046000
h=36	0.103500	0.058000	0.055000	0.047500

Table 5				
Raw Power, nominal 5% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	1.000000	1.000000	1.000000	1.000000
h=2	1.000000	1.000000	0.998500	1.000000
h=3	0.982000	1.000000	0.606500	0.999500
h=6	0.704000	0.998500	0.266500	0.921000
h=9	0.257500	0.923000	0.066500	0.695500
h=12	0.075000	0.850500	0.051000	0.590000
h=18	0.074000	0.755500	0.049500	0.540000
h=24	0.079000	0.704500	0.051000	0.507000
h=36	0.079500	0.637000	0.050000	0.435000
Panel B: Recursive Regressions				
h=1	1.000000	1.000000	1.000000	1.000000
h=2	1.000000	1.000000	0.998000	1.000000
h=3	0.985500	1.000000	0.612500	0.999500
h=6	0.723500	0.997500	0.266000	0.911500
h=9	0.268500	0.911000	0.067500	0.691500
h=12	0.073500	0.836000	0.051000	0.594500
h=18	0.075000	0.745000	0.049000	0.542500
h=24	0.075000	0.698000	0.050000	0.497000
h=36	0.063000	0.627500	0.051500	0.425000

Table 6				
Raw Power, nominal 5% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.119000	0.992000	0.072500	0.488500
h=2	0.184500	0.913000	0.066000	0.344500
h=3	0.258000	0.661500	0.063000	0.220500
h=6	0.517500	0.223000	0.074000	0.094500
h=9	0.671000	0.174500	0.100000	0.093000
h=12	0.710500	0.170000	0.113000	0.092000
h=18	0.592000	0.184500	0.076500	0.093000
h=24	0.433500	0.178000	0.066000	0.082000
h=36	0.251500	0.181500	0.050500	0.066000
Panel B: Recursive Regressions				
h=1	0.141000	0.995500	0.073500	0.487000
h=2	0.201000	0.935000	0.065000	0.349500
h=3	0.291500	0.728500	0.063000	0.222000
h=6	0.582000	0.284500	0.074500	0.102000
h=9	0.741000	0.208000	0.101000	0.099000
h=12	0.774500	0.191500	0.110000	0.093500
h=18	0.667500	0.202500	0.074000	0.085500
h=24	0.468500	0.199500	0.063500	0.079500
h=36	0.252500	0.183500	0.050000	0.066500

Table 7				
Size Adjusted Power, nominal 5% tests, DGPs 1 and 2				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.999500	1.000000	0.806000	1.000000
h=2	1.000000	1.000000	0.751500	0.987500
h=3	0.993000	0.441500	0.584500	0.100500
h=6	0.959500	0.108000	0.415500	0.058500
h=9	0.843500	0.069000	0.250000	0.055000
h=12	0.679000	0.050500	0.150000	0.054000
h=18	0.346000	0.074000	0.082500	0.055000
h=24	0.224000	0.057000	0.064000	0.052000
h=36	0.126500	0.059000	0.053000	0.055000
Panel B: Recursive Regressions				
h=1	1.000000	1.000000	0.805000	1.000000
h=2	0.999500	1.000000	0.754000	0.991500
h=3	0.994500	0.478500	0.564500	0.104000
h=6	0.968500	0.115500	0.392500	0.059000
h=9	0.864000	0.067000	0.239500	0.054000
h=12	0.660500	0.063500	0.144000	0.056000
h=18	0.393500	0.071000	0.084500	0.060000
h=24	0.258000	0.072000	0.060000	0.056500
h=36	0.120500	0.068000	0.057500	0.053500

Table 8				
Size Adjusted Power, nominal 5% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	1.000000	1.000000	1.000000	1.000000
h=2	1.000000	1.000000	0.999000	1.000000
h=3	0.901000	1.000000	0.447000	0.999500
h=6	0.709000	0.999500	0.268000	0.907500
h=9	0.226000	0.880500	0.070000	0.694500
h=12	0.068000	0.749000	0.046500	0.612500
h=18	0.074000	0.568500	0.049500	0.563500
h=24	0.063000	0.468500	0.043000	0.535000
h=36	0.067000	0.346500	0.050000	0.452000
Panel B: Recursive Regressions				
h=1	1.000000	1.000000	1.000000	1.000000
h=2	1.000000	1.000000	0.998000	1.000000
h=3	0.902000	1.000000	0.427500	0.998500
h=6	0.709500	0.997000	0.257000	0.909000
h=9	0.236000	0.882000	0.080500	0.670500
h=12	0.066500	0.732000	0.057000	0.586000
h=18	0.060500	0.565000	0.059500	0.573500
h=24	0.065000	0.462500	0.056000	0.532000
h=36	0.046500	0.386000	0.056500	0.456000

Table 9				
Size Adjusted Power, nominal 5% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.141000	0.995500	0.060500	0.477000
h=2	0.185500	0.933000	0.063000	0.325500
h=3	0.261000	0.408500	0.069000	0.152000
h=6	0.256500	0.187000	0.075500	0.091000
h=9	0.210500	0.117500	0.099000	0.086000
h=12	0.184500	0.103500	0.113000	0.073500
h=18	0.188000	0.095500	0.084500	0.078500
h=24	0.128500	0.079500	0.075500	0.065500
h=36	0.102000	0.064000	0.058000	0.067000
Panel B: Recursive Regressions				
h=1	0.146500	0.998000	0.107500	0.485000
h=2	0.205000	0.951500	0.113000	0.342000
h=3	0.290500	0.523000	0.129500	0.155000
h=6	0.278000	0.257500	0.140000	0.100000
h=9	0.229000	0.142000	0.173500	0.099000
h=12	0.208000	0.125500	0.190000	0.088500
h=18	0.168500	0.092000	0.150000	0.079500
h=24	0.143500	0.073500	0.129000	0.077000
h=36	0.105000	0.054000	0.112500	0.064000

Tables 10-18 illustrate size, raw power and size adjusted power for finite sample nominal 5% tests.

Table 10				
Size, nominal 5% tests, DGPs 1 and 2				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.038500	0.036500	0.065000	0.055000
h=2	0.039500	0.048000	0.063000	0.054500
h=3	0.040000	0.045000	0.063000	0.053500
h=6	0.047500	0.056000	0.056500	0.052500
h=9	0.053000	0.058500	0.059500	0.052000
h=12	0.047500	0.053000	0.058500	0.053500
h=18	0.047000	0.051000	0.053500	0.051500
h=24	0.049000	0.048000	0.054500	0.051000
h=36	0.061000	0.049500	0.054500	0.059500
Panel B: Recursive Regressions				
h=1	0.037000	0.037000	0.050500	0.048500
h=2	0.027000	0.044500	0.052000	0.052000
h=3	0.035000	0.045500	0.051000	0.052500
h=6	0.037500	0.056000	0.050500	0.050000
h=9	0.046000	0.050000	0.056500	0.052000
h=12	0.047000	0.054000	0.052000	0.047500
h=18	0.060000	0.055500	0.053000	0.050500
h=24	0.058000	0.064500	0.051000	0.050500
h=36	0.053000	0.059500	0.057000	0.057000

Table 11				
Size, nominal 5% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	0.044000	0.042000	0.067000	0.065500
h=2	0.050500	0.048500	0.067000	0.068500
h=3	0.047000	0.051000	0.060000	0.068500
h=6	0.051500	0.064500	0.061000	0.068000
h=9	0.053000	0.076500	0.056500	0.068000
h=12	0.053000	0.094500	0.056000	0.065500
h=18	0.055000	0.144000	0.055000	0.071500
h=24	0.060000	0.144000	0.059500	0.076000
h=36	0.059000	0.168000	0.057000	0.083500
Panel B: Recursive Regressions				
h=1	0.047500	0.046000	0.055000	0.056500
h=2	0.045500	0.050500	0.054000	0.061000
h=3	0.046000	0.055500	0.057000	0.067000
h=6	0.044500	0.057000	0.058500	0.054000
h=9	0.058500	0.063000	0.057000	0.058000
h=12	0.062500	0.098500	0.055000	0.057000
h=18	0.070500	0.133000	0.053000	0.073000
h=24	0.079000	0.167500	0.055000	0.073000
h=36	0.081500	0.198500	0.059000	0.077500

Table 12				
Size, nominal 5% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.055000	0.057000	0.065500	0.061500
h=2	0.070500	0.070500	0.072500	0.072500
h=3	0.084500	0.081500	0.071500	0.074000
h=6	0.131500	0.102000	0.075000	0.090000
h=9	0.202000	0.134000	0.084000	0.108500
h=12	0.240500	0.145500	0.098000	0.112000
h=18	0.208500	0.184000	0.089000	0.119500
h=24	0.192500	0.202500	0.079000	0.110500
h=36	0.184500	0.194000	0.075500	0.105500
Panel B: Recursive Regressions			-	
h=1	0.055000	0.047500	0.064000	0.062500
h=2	0.065000	0.060500	0.065000	0.072000
h=3	0.080000	0.072500	0.065000	0.070000
h=6	0.134500	0.095500	0.073000	0.083500
h=9	0.235500	0.123000	0.072500	0.100500
h=12	0.288500	0.136000	0.076000	0.106500
h=18	0.247500	0.185000	0.079500	0.113000
h=24	0.220500	0.203000	0.072000	0.109000
h=36	0.191000	0.208500	0.061000	0.096500

Table 13

Raw Power, nominal 5% tests, DGPs 1 and 2

Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.660500	1.000000	0.355000	1.000000
h=2	0.665000	0.959000	0.325000	0.676000
h=3	0.605500	0.459000	0.282500	0.170500
h=6	0.446500	0.102000	0.181500	0.062000
h=9	0.306000	0.072500	0.132500	0.060000
h=12	0.224000	0.060500	0.097500	0.060000
h=18	0.152500	0.062000	0.066500	0.058000
h=24	0.132000	0.072000	0.060000	0.061500
h=36	0.090500	0.071500	0.055000	0.052500
Panel B: Recursive Regressions				
h=1	0.762500	1.000000	0.406000	1.000000
h=2	0.751500	0.966500	0.362000	0.684500
h=3	0.687000	0.471500	0.322500	0.165000
h=6	0.517000	0.099500	0.197500	0.063500
h=9	0.356000	0.077500	0.137500	0.062500
h=12	0.264000	0.077500	0.104500	0.060000
h=18	0.168000	0.073000	0.084000	0.054500
h=24	0.145500	0.098500	0.074000	0.058000
h=36	0.123000	0.095500	0.058500	0.055500

Table 14				
Raw Power, nominal 5% tests, DGPs 3 and 4				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	1.000000	1.000000	0.971000	0.996500
h=2	0.952000	0.999500	0.755500	0.983500
h=3	0.498000	0.973500	0.302000	0.875000
h=6	0.301000	0.576500	0.178500	0.448500
h=9	0.120500	0.416500	0.071000	0.322500
h=12	0.099500	0.420000	0.065500	0.307500
h=18	0.119000	0.473500	0.058000	0.365000
h=24	0.113500	0.470000	0.057500	0.365000
h=36	0.109000	0.422500	0.062500	0.330500
Panel B: Recursive Regressions				
h=1	1.000000	1.000000	0.982000	0.998500
h=2	0.972500	0.999500	0.759500	0.993000
h=3	0.568000	0.986500	0.314500	0.904500
h=6	0.318500	0.663500	0.161500	0.470000
h=9	0.183000	0.466000	0.053500	0.317500
h=12	0.083000	0.457500	0.046000	0.310500
h=18	0.099500	0.512000	0.046000	0.351500
h=24	0.101000	0.518500	0.045000	0.352500
h=36	0.104000	0.503000	0.052000	0.348000

Table 15				
Raw Power, nominal 5% tests, DGPs 5 and 6				
Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.074000	0.474500	0.0710000	0.250000
h=2	0.098500	0.255500	0.072000	0.164000
h=3	0.106000	0.160000	0.076500	0.124500
h=6	0.172000	0.130000	0.084000	0.108500
h=9	0.268000	0.150500	0.115500	0.132500
h=12	0.310500	0.174000	0.126500	0.138000
h=18	0.276000	0.223500	0.113000	0.146500
h=24	0.226000	0.221500	0.103500	0.138000
h=36	0.222500	0.210000	0.088500	0.118500
Panel B: Recursive Regressions				
h=1	0.075000	0.536000	0.065500	0.248500
h=2	0.094500	0.293500	0.069500	0.163500
h=3	0.116000	0.184000	0.069500	0.118000
h=6	0.206000	0.120000	0.081000	0.103500
h=9	0.321000	0.150500	0.104500	0.121000
h=12	0.378500	0.173500	0.118500	0.123000
h=18	0.320500	0.214500	0.094000	0.129500
h=24	0.263000	0.238000	0.086500	0.124500
h=36	0.220000	0.234000	0.074000	0.106000

Table 16

Size Adjusted Power, nominal 5% tests, DGPs 1 and 2

Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 1	DGP 2	DGP 1	DGP 2
Panel A: Rolling Regressions				
h=1	0.711000	0.711000	0.330000	0.333000
h=2	0.702000	0.702000	0.318000	0.318000
h=3	0.596000	0.596000	0.233500	0.233500
h=6	0.446000	0.446000	0.188000	0.188000
h=9	0.319500	0.319500	0.123500	0.123500
h=12	0.264500	0.264500	0.103000	0.103000
h=18	0.158000	0.158000	0.077000	0.077000
h=24	0.135500	0.135500	0.068000	0.068000
h=36	0.083000	0.083000	0.060500	0.060500
Panel B: Recursive Regressions				
h=1	0.800000	0.800000	0.374000	0.374000
h=2	0.825000	0.825000	0.338000	0.338000
h=3	0.689000	0.689000	0.249500	0.249500
h=6	0.549500	0.549500	0.184000	0.184000
h=9	0.352500	0.352500	0.125000	0.125000
h=12	0.263500	0.263500	0.099000	0.099000
h=18	0.148000	0.148000	0.066000	0.066000
h=24	0.126500	0.126500	0.065000	0.065000
h=36	0.119000	0.119000	0.050000	0.050000

Table 17

Size Adjusted Power, nominal 5% tests, DGPs 3 and 4

Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 3	DGP 4	DGP 3	DGP 4
Panel A: Rolling Regressions				
h=1	0.999500	1.000000	0.955500	0.994500
h=2	0.955000	0.998500	0.680000	0.978000
h=3	0.390500	0.883000	0.223000	0.687500
h=6	0.301500	0.535500	0.171500	0.391000
h=9	0.122000	0.342000	0.064500	0.278000
h=12	0.082500	0.292500	0.056000	0.291000
h=18	0.096000	0.225500	0.053000	0.321000
h=24	0.098000	0.193500	0.048500	0.307000
h=36	0.089500	0.129000	0.049000	0.254000
Panel B: Recursive Regressions				
h=1	1.000000	1.000000	0.975500	0.998500
h=2	0.976000	1.000000	0.759500	0.987000
h=3	0.445000	0.910000	0.214500	0.765500
h=6	0.347500	0.587000	0.149500	0.447000
h=9	0.119500	0.388500	0.060500	0.273000
h=12	0.073000	0.250000	0.053500	0.267000
h=18	0.072000	0.198500	0.050500	0.280500
h=24	0.069000	0.139000	0.051000	0.287500
h=36	0.072500	0.082000	0.046000	0.267000

Table 18

Size Adjusted Power, nominal 5% tests, DGPs 5 and 6

Clark-West			Clark-West-Normal (ϕ=0.1)	
(1)	(2)	(3)	(4)	(5)
Horizon	DGP 5	DGP 6	DGP 5	DGP 6
Panel A: Rolling Regressions				
h=1	0.660000	0.615500	0.052500	0.303500
h=2	0.071500	0.350000	0.056000	0.210000
h=3	0.073000	0.158000	0.059500	0.134500
h=6	0.063000	0.127000	0.063000	0.119500
h=9	0.065500	0.124500	0.073500	0.124500
h=12	0.075000	0.126000	0.076500	0.118500
h=18	0.062000	0.125500	0.070000	0.125500
h=24	0.061500	0.137000	0.066500	0.120000
h=36	0.056500	0.110500	0.060500	0.108500
Panel B: Recursive Regressions				
h=1	0.064500	0.699000	0.052000	0.317500
h=2	0.064500	0.427500	0.058000	0.212000
h=3	0.081500	0.173500	0.061000	0.140500
h=6	0.094500	0.129000	0.062500	0.119000
h=9	0.097000	0.124000	0.072000	0.123500
h=12	0.084500	0.122500	0.074500	0.118000
h=18	0.092000	0.123000	0.068000	0.116500
h=24	0.069500	0.114000	0.067000	0.115500
h=36	0.067500	0.111000	0.062000	0.111500

References

1. Ashley, R., Granger, C.W.J., Schmalensee, R., 1980. Advertising and aggregate consumption: an analysis of causality. Econometrica 48, 1149—1168.

2. Ciccarelli, M., Mojon, B., 2010. Global inflation. Rev. Econ. Stat. 92 (3), 524–535

3. Chong, Y.Y., Hendry, D.F., 1986. Econometric evaluation of linear macroeconomic models. Review of Economic Studies 53, 671—690.

4. Chao, J., Corradi, V., Swanson, N., 2001. An out of sample test for Granger causality. Macroeconomic Dynamics, forthcoming.

5. Clark, T.E., 1999. Finite-sample properties of tests for equal forecast accuracy. Journal of Forecasting 18, 489—504.

6. Clark, T.,McCracken,M.,2001.Tests of equal forecast accuracy and encompassing for nested models. J.Econom.105,85—110.

7. Clark, T.,West,K.D.,2006. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis.J.Econom.135(1–2), 155—186.

8. Clark, T.,West,K.D.,2007. Approximately normal tests for equal predictive accuracy in nested models. J.Econom.138,291—311.

9. Corradi, V., Swanson, N.R., Olivetti, C., 2001. Predictive ability with cointegrated variables. Journal of Econometrics, forthcoming.

10. Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253—263

11. Ericsson, N.R., 1992. Parameter constancy, mean square forecast errors, and measuring forecast performance: an exposition, extensions, andillustration. Journal of Policy Modeling 14, 465—495.

12. Hakkio, C., 2009. Global inflation dynamics. Research Working Paper 09-01, Federal Reserve Bank of Kansas City.

13. Harvey, D.I., Leybourne, S.J., Newbold, P., 1998. Tests for forecast encompassing. Journal of Business and Economic Statistics 16, 254—259

14. Hubrich, K., West, K., 2010. Forecast evaluation of small nested model sets. Journal of Applied Econometrics 25, 574—594.

15. Mankiw, N.G.,Shapiro,M.D.,1986.Do we reject too often? Small sample properties of tests of rational expectations models. Econ.Lett.20,139—14

16. McCracken, M.W., 2000. Robust out of sample inference. Journal of Econometrics 99, 195—223.

17. Medel, C., Pedersen, M., Pincheira, P., 2016. The elusive predictive ability of global inflation. Int. Financ. (forthcoming)

18. Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: do they . . . fit out of sample? J. Int. Econ. 14, 3-24.

19. Mincer, J., Zarnowitz, V., 1969. "The Evaluation of Economic Forecasts," NBER Chapters, in: Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance, pages 3-46, National Bureau of Economic Research, Inc.

20. Morales-Arias, L., Moura, G.V., 2013. A conditional heteroskedastic global inflation model. J. Econ. Stud. 40 (4), 572–596

21. Pincheira, P., Gatty, A., 2016. Forecasting chilean inflation with international factors. Empir. Econ.http://dx.doi.org/10.1007/s00181

22. Pincheira, P., West, K.D., 2016. A comparison of some out-of-sample tests of predictability in iterated multi-step-ahead forecasts. Research in Economics 70, Elsevier, pp 304-319.

23. Pincheira, P, J Selaive, and J L Nolazco (2016), "The Evasive Ability of Core Inflation", BBVA Working Paper 15/34, January.

24. Pincheira, P., 2017, "A Power Booster Factor for Out-of-Sample Tests of Predictability", MPRA Paper No. 77027

25. Tauchen,G.,2001.The bias of tests for a risk premium in forward exchange rates.J.Empir.Financ.8, 695—704.

26. West,K.D.,1996.Asymptotic inference about predictive ability. Econometrica 64(5), 1067—1084.

27. West, K.D., McCracken, M.W., 1998. Regression-basedtests of predictive ability. International Economic Review 39, 817—840

28. West, K.D., 2001a. Tests for forecast encompassing when forecasts depend on estimated regression parameters. Journal of Business and Economic Statistics 19, 29—33.

29. West, K.D., 2001b. Encompassing tests when no model is encompassing. Journal of Econometrics 105.

30. West,K.D.,2006.Forecast evaluation. Elliott,G.,Granger,C.,Timmerman,A.(Eds.), Handbook of Economic Forecasting,vol.1.,Elsevier,Amsterdam, pp. 100—134