Comment on: Taylor Rule Exchange Rate Forecasting During the Financial Crisis, by T. Molodtsova and D. Papell

Barbara Rossi

ICREA, UPF, CREI, BGSE

September 3, 2012

Abstract

This comment analyzes the importance of time variation in the forecasting performance of Taylor rule models of exchange rate determination as well as the robustness of the results in Molodtsova and Papell (2012).

Acknowledgements: I thank T. Molodtsova and D. Papell for sharing their data, and to K. West for organizing the session at 2012 ISoM in Oslo.

Keywords: Forecasting, Exchange Rates, Taylor Rules.

J.E.L. Codes: C22, C52, C53

1 Introduction

Molodtsova and Papell (2012, MP hereafter) investigate the out-of-sample predictability of exchange rates during the recent financial crisis. Several recent works, including Engel and West (2006), Molodtsova and Papell (2009) and Molodtsova et al. (2010) among others, have emphasized the usefulness of Taylor rules for forecasting exchange rates. Let me first provide some background on how Taylor rules can be used to forecast exchange rates. To fix ideas, let the traditional Taylor rule be: $i_t = \alpha + \gamma(\pi_t - \overline{\pi}) + \phi y_t^{gap}$, where the interest rate of the "home" country (i_t) is decided according to a monetary policy reaction function that reacts to the inflation gap (that is, how much the inflation rate π_t exceeds the inflation target $\overline{\pi}$) and to the output gap (y_t^{gap}). Let asterisks denote the foreign country's variables and assume that the foreign country's monetary policy follows a similar Taylor rule, with the same coefficients: $i_t^* = \alpha^* + \gamma(\pi_t^* - \overline{\pi}) + \phi y_t^{gap,*}$. The difference of the Taylor rules for the two countries, together with the uncovered interest parity condition (UIRP: $E_t \Delta e_{t+1} = i_t - i_t^*$, where e_t is the nominal exchange rate between the two countries and $\Delta e_{t+1} \equiv e_{t+1} - e_t$), gives:

$$E_t \Delta e_{t+1} = \kappa + \gamma(\pi_t - \pi_t^*) + \phi(y_t^{gap} - y_t^{gap*}), \tag{1}$$

where $\kappa = \alpha - \alpha^*$ and E_t is the conditional expectation at time t. Among the papers that investigated whether eq. (1) can forecast future exchange rate fluctuations out-of-sample, Engel and West (2006) proposed using Taylor rules with fixed coefficients, whereas Molodtsova and Papell (2009) proposed using Taylor rules with estimated coefficients. The contribution of MP is to augment the Taylor rule with indicators of financial stress:

$$E_t \Delta e_{t+1} = \kappa + \gamma(\pi_t - \pi_t^*) + \phi(y_t^{gap} - y_t^{gap*}) + \delta(s_t - s_t^*), \tag{2}$$

where s_t, s_t^* are the financial stress indicators for the home and the foreign country, respectively. MP consider several empirical specifications of eqs. (1) and (2), including different measures of output gap and financial stress. In this comment we will focus on two of their specifications: (i) the traditional Taylor rule fundamental model without financial conditions indices, eq. (1), where the output gap is measured by OECD estimates; and (ii) the Taylor rule fundamental model (2) augmented with the TED spread differential as the measure of financial stress.

There are two features of MP's work that I find especially interesting and overall very important for the debate on exchange rate predictability. The first is the attempt to use real-time data (when available), that is data that were actually available to forecasters at the point in time in which the pseudo out-of-sample forecast is generated. This practice avoids finding predictive ability due to either subsequent data revisions or, in general, to any information that became available to researchers ex-post and which may potentially bias the results towards predictive ability.[1] The second important and very interesting feature of their work is the fact that they recognize the importance of instabilities in the forecasting performance of the predictors. As pointed out in Giacomini and Rossi (2010), failure to do so may result in an incorrect evaluation of the predictors' forecasting ability.

This comment has two goals. First, we highlight the importance of taking into account instabilities in the predictors' forecasting ability; we also highlight the dangers of repeatedly testing predictive ability over time without appropriately correcting the critical values. The latter procedure may spuriously find predictive ability even when there is none in the data. This problem is especially important in the exchange rate literature where a predictor is successful if it forecasts better than the random walk. Fortunately, there is a very simple way to correct the problem, and we revisit the empirical evidence in MP accordingly.[2] Second, we perform additional analyses to evaluate the robustness of the results in MP.[3]

2 The Danger of Data Mining Over Time

Monetary policy does change over time (see Clarida, Gali and Gertler, 2000, Curdia and Woodford, 2010, and Gertler and Karadi, 2011, among others). Thus, presumably, the predictive ability of Taylor rules may change over time as well. In particular, one might expect that a model including traditional Taylor rule fundamentals may work well before the 2007-2009 crisis, but may work poorly subsequently; conversely, one might expect that indices of financial stress may have become important at the time of the crisis. It is therefore very

[1] Note that, however, one finding of the literature is that the use of real-time data actually improves the forecasting ability of exchange rate models with traditional monetary fundamentals out-of-sample – see Faust, Rogers and Wright (2003).

[2] The analysis is done using MP's (2012) data, although based on an independent replication of their results. All tests are implemented with Newey and West's (1987) heteroskedasticity and serial correlation robust variance estimator with a bandwidth equal to two.

[3] Several additional robustness analyses would include: (i) the use of other test statistics for predictive ability; (ii) the extension to other countries where, unlike in Europe, data were not backcasted; (iii) the use of other financial condition indices with longer time series, such as housing prices or indicators extracted from factor models; (iv) the extension to other Taylor rules (e.g. with time-varying inflation targets); (v) an analysis of the theoretical properties of the Taylor rule model, including signs and significance of the coefficients – see Chinn (2008). We will not discuss these additional issues due to space constraints.

important to evaluate the predictive ability of the models *over time*, as their (relative) forecasting performance might have changed. This is true in general when forecasting exchange rates with Taylor rule and UIRP fundamentals, as shown in Giacomini and Rossi (2010), and in particular when forecasting exchange rates using Taylor rules that are either traditional or augmented with financial indicators, as in this paper.

MP realize the importance of instabilities and report the Clark and West (2006, CW hereafter) test recursively calculated over time. Figure 1 shows the importance of taking into account instabilities when forecasting exchange rates using Taylor rules. The solid line in Figure 1, Panel A, reports CW's (2006) test statistic over time for the fundamental Taylor rule model without financial indices, eq. (1), against the random walk. Large values of the statistics are evidence in favor of the Taylor rule model. The figure shows that a researcher who evaluates the predictive ability of Taylor rules in 2012 would estimate a CW (2006) test statistic equal to 0.8; since the critical value at the 5% level is 1.645, the researcher would conclude that Taylor rules have no predictive ability. However, he/she would miss that if he/she were to evaluate the predictive ability of Taylor rules in 2008, the CW (2006) test statistic would have been 1.9, and he/she would have found the opposite result. So it is very important to use techniques that follow models' relative performance over time when the environment is unstable and when the predictive ability of the fundamentals may change over time.

However, unfortunately, testing for predictive ability over time is not as simple as in MP: while CW's (2006) critical values are valid point-wise, they are not designed for multiple testing (over time) the way the authors implement it. In other words, the original CW's (2006) critical values were calculated for a "one-shot" test, not for "repeated" (or multiple) tests. Their application of the test is problematic since repeating the test using the usual critical values may lead to over-rejections. This means that the empirical evidence in favor of predictive ability might be spurious. To understand the intuition why the test cannot be implemented without adjusting the critical values, think about the simple example of flipping a coin. When one flips a fair coin, there is a 50% probability of head (tail). However, if the researcher keeps flipping the same coin, the probability of finding head in at least one of the draws will be much higher. Similarly, if one implements CW's (2006) test (or any test) once using their 5% critical values, he(she) will have a 5% chance of incorrectly rejecting the null of no predictability in favor of predictability; however, if one implements CW's (2006) test multiple times using CW's (2006) 5% critical values for each outcome, the probability of incorrectly rejecting the null of no predictability in favor of predictability will be higher than

5%. The probability of spuriously finding predictive ability will increase with the number of times the test is repeated.

A simple test that is available to control for potential over-rejections is Giacomini and Rossi's (2010) Fluctuation test statistics. Their procedure works as follows: CW's (2006) test is repeated in rolling windows over the out-of-sample period, and the critical values are adjusted to take into account that the test is repeated multiple times. For details on the implementation of the test see Giacomini and Rossi (2010, p. 601). There is one minor difference between the way Giacomini and Rossi (2010) propose to implement the repeated CW's (2006) test statistic and the way MP implement it: the latter recursively report CW's (2006) test using all the forecasts available at the evaluation time, whereas Giacomini and Rossi (2010) re-calculate CW's (2006) test in a rolling fashion, using only the most recent forecasts in a window of fixed size. Therefore, Giacomini and Rossi's (2010) procedure updates more quickly to the instabilities in the data.

Going back to Figure 1, Panel A not only depicts the CW (2006) test statistic (solid line) but also the "one-shot" critical values that MP use (dotted line). According to these critical values, when the CW (2006) test statistic is above the dotted line, we conclude that the Taylor rule model forecasts significantly better than the random walk. The solid line in Panel B of Figure 1 shows instead CW's (2006) statistic calculated over rolling windows.[4] Note that the largest value is comparable to that in Panel A; however, the value of the statistic in Panel B decreases more rapidly after 2008 than that in Panel A; this suggests that the predictive ability was a transitory phenomenon. Either way, using the "one-shot" critical value (1.645) would have still led to significant predictive ability around the end of 2007-beginning of 2008. However, as previously explained, the critical values are only valid if the test were applied once. When we correct the critical values to take into account the repeated tests (dotted line in Panel B), the critical values become much higher and there is no evidence of superior predictive ability anymore. Figure 2 shows that the results are similar for the Taylor rule model augmented with the TED spread, eq. (2).

Overall, Figures 1 and 2 show that correctly taking into account the repeated nature of the test is crucial to avoid incorrectly concluding that a model forecasts better than the benchmark. However, these are only two of the cases considered in MP: it is possible that, in some of their other specifications, the empirical evidence in favor of the model is so strong that it remains significant even after correcting for the repeated nature of the test.

[4] The size of the window used for CW is 30 observations, the same as in MP. Thus, the first point on the solid line is the same in both panels.

3 Robustness Analysis

The window size used for estimation of the Taylor rule parameters plays two roles: first, it determines the out-of-sample forecasting period; second, it controls the amount of "smoothing" used to re-estimate the parameters over time. Thus, one might potentially obtain different empirical results using different window sizes. So, how should the window size be chosen? And how do different values of the window size affect MP's finding of predictive ability?

Figure 3 depicts CW's (2006) test statistic calculated for various window sizes, labeled "R". MP use $R = 26$.[5] Panel A in the figure reports results for the traditional Taylor rule fundamental model, and Panel B reports results for the Taylor rule fundamental model augmented with the TED differential. Clearly, the figures show that varying the window size does not alter the results much. Thus, either a larger or a smaller window size than the one used by MP only slightly affects predictive ability. Thus, their results are very robust to changes in the window size.

But how should the window size be chosen? Ideally, researchers are not interested in the window size per se; instead, they would like their results to be robust to the window size choice. As discussed in Inoue and Rossi (2012), when evaluating the robustness of the results to the choice of the window size, one runs into the same problem as those described in Section 2, namely the fact that the test is repeated across multiple window sizes. Again, repeated tests may result in spurious evidence in favor of predictive ability. On the other hand, it is also the case that smaller or larger window sizes may find more or less evidence of predictive ability depending on the degree of instability in the data.[6] Thus, varying the window size may improve the evidence in favor of predictive ability too. To correct for these issues, we report Inoue and Rossi's (2012) test statistic, which is robust to these problems. In the traditional Taylor rule model, Inoue and Rossi's (2012) CW_T test statistic is 5.0802, with a p-value of 0.6502; in the Taylor rule model with financial condition indices, the test statistic is 12.7049, with a p-value of 0.0796. Thus, in neither case the empirical evidence is favorable to the model at conventional (5%) significance levels, although there is marginal evidence in favor of the second model at the 10% significance level.

[5] When the window size is smaller than 26, we simply use the latest observations to estimate the parameters, and then forecast over the same out-of-sample period as MP. When the window size is larger, we start forecasting later in the sample.

[6] Smaller window sizes may be more robust to instabilities; larger window sizes may provide more precise estimates when there is no instability.

4 Conclusions

Understanding the factors that drive exchange rate dynamics is a crucial question for several reasons, including predicting the transmission effects of policies in open economies, and assessing the benefits and risks faced by international businesses. Recent research has suggested that Taylor rules may be potentially important predictors of future exchange rate fluctuations. The work by MP is an interesting example of using these alternative predictors for exchange rates in a framework where instabilities are important, as well as in a framework that recognizes the importance of real-time data. Both the latter aspects are very important, and make the paper of MP a worthwhile enterprise.

Regarding the first aspect, MP show that, when using predictors such as the Taylor rule, it becomes crucial to contemplate the possibility that the performance of the predictor may be time-varying. As shown in MP, during the latest financial crisis the forecasting ability of the Taylor rule model has worsened significantly. Failing to acknowledge the possibility that the models' relative performance may change over time would incorrectly lead the researcher to conclude that the Taylor rule model does not forecast well, when, in reality, this conclusion is heavily influenced by the financial crisis period. However, examining the evolution of predictive ability over time is not simple. In fact, as we highlighted in this comment, simply utilizing existing tools with the existing critical values may lead to spurious evidence of predictive ability. Fortunately, tools that are designed to evaluate the out-of-sample predictive content over time are available and can make a difference; in particular, we showed that they do in some of the cases analyzed in MP. We also argued that the choice of the window size might also be potentially important, especially in the presence of instabilities, although the results in MP's analysis are overall robust to the latter.

5 References

Chinn, M. (2008), "Nonlinearities, Business Cycles and Exchange Rates," *Economic Notes* 37(3), 219-239.

Clarida, R., J. Galí and M. Gertler (2000), "Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory," *Quarterly Journal of Economics* 115(1), 147-180.

Clark, T.E. and K.D. West (2006), "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics* 135, 155-186.

Curdia, V. and M. Woodford (2010), "Credit Spreads and Monetary Policy", *Journal of Money, Credit and Banking* 42, 3-35.

Engel, C. and K.D. West (2006), "Taylor Rules and the Deutschemark-Dollar Real Exchange Rate", *Journal of Money, Credit and Banking* 38, 1175-1194.

Faust, J., J.H. Rogers and J.H. Wright (2003), "Exchange Rate Forecasting: the Errors We've Really Made," *Journal of International Economics* 60(1), 35-59.

Gertler, M. and P. Karadi (2011), "A Model of Unconventional Monetary Policy", *Journal of Monetary Economics* 58, 17-34.

Giacomini, R. and B. Rossi (2010), "Forecast Comparisons in Unstable Environments", *Journal of Applied Econometrics* 25(4), 595-620.

Inoue, A. and B. Rossi (2011), "Out-of-Sample Forecast Tests Robust to the Choice of Window Size," *Journal of Business and Economic Statistics* 30(3), 432-453.

Molodtsova, T. and D.H. Papell (2009), "Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals," *Journal of International Economics*, 77(2), 167-180.

Molodtsova, T., A. Nikolsko-Rzhevskyy and D.H. Papell (2010), "Taylor Rules and the Euro", *Journal of Money, Credit and Banking*, 535-552.

Molodtsova, T. and D.H. Papell (2012), "Taylor Rule Exchange Rate Forecasting During the Financial Crisis," *NBER International Seminar on Macroeconomics*, this volume.

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

Figure 1. Panel A

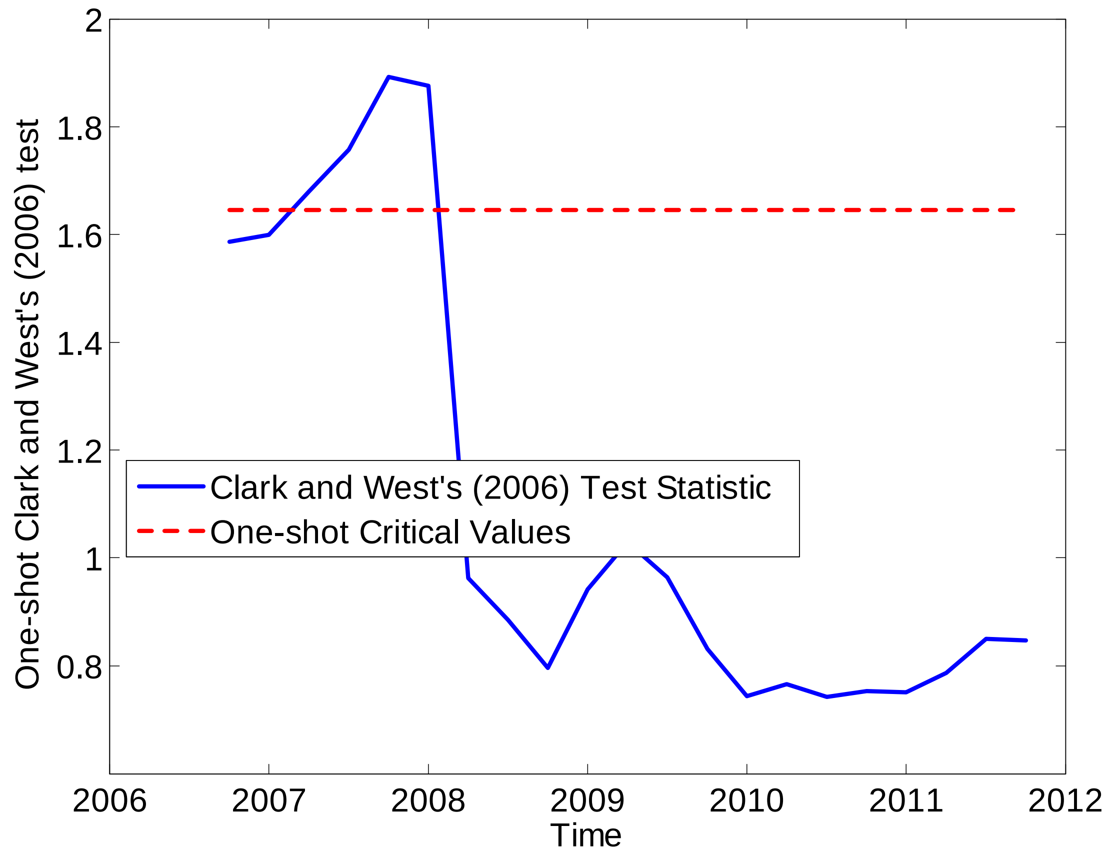

Figure 1. Panel B

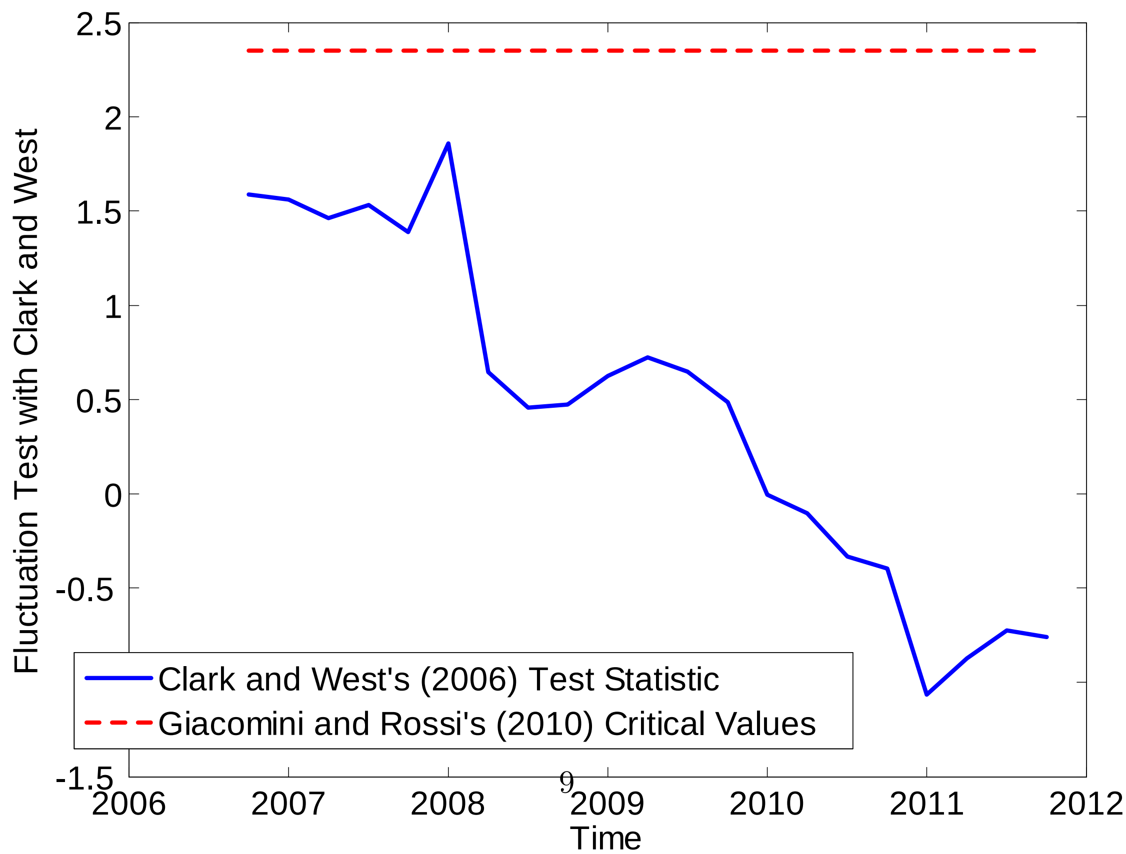

Figure 2. Panel A

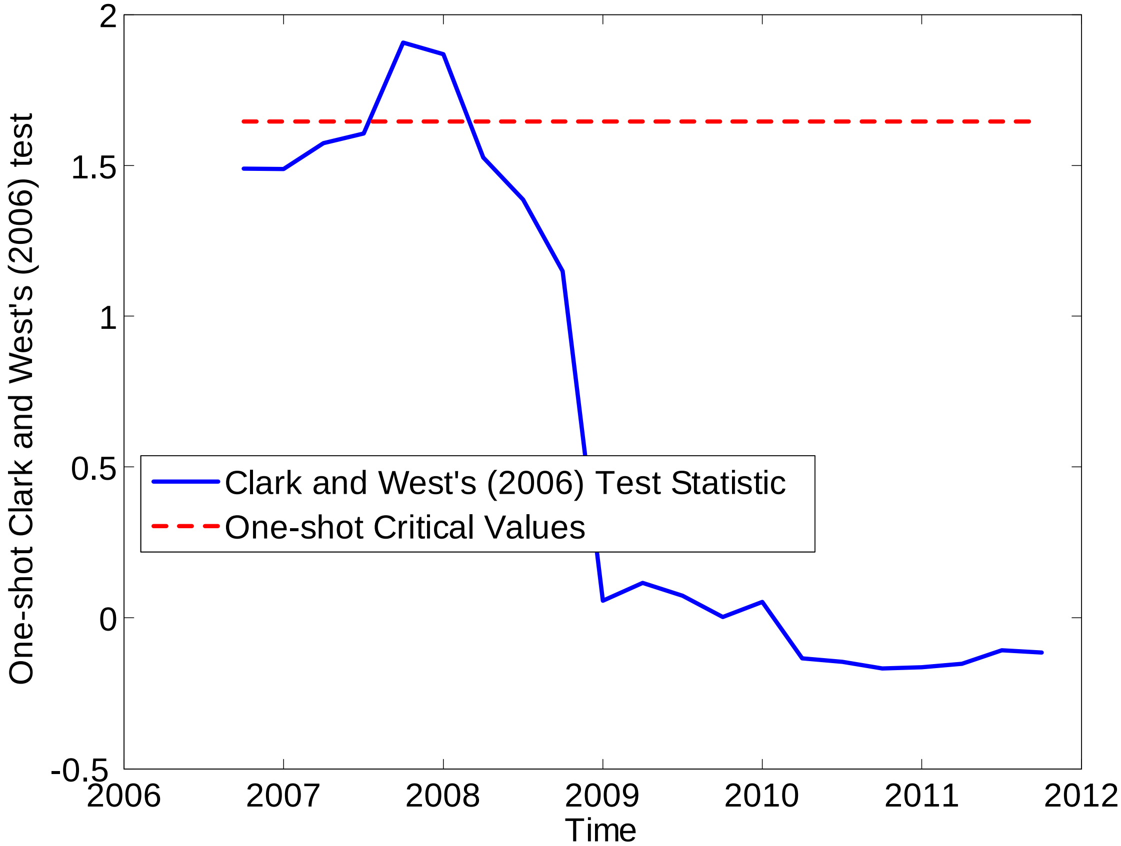

Figure 2. Panel B

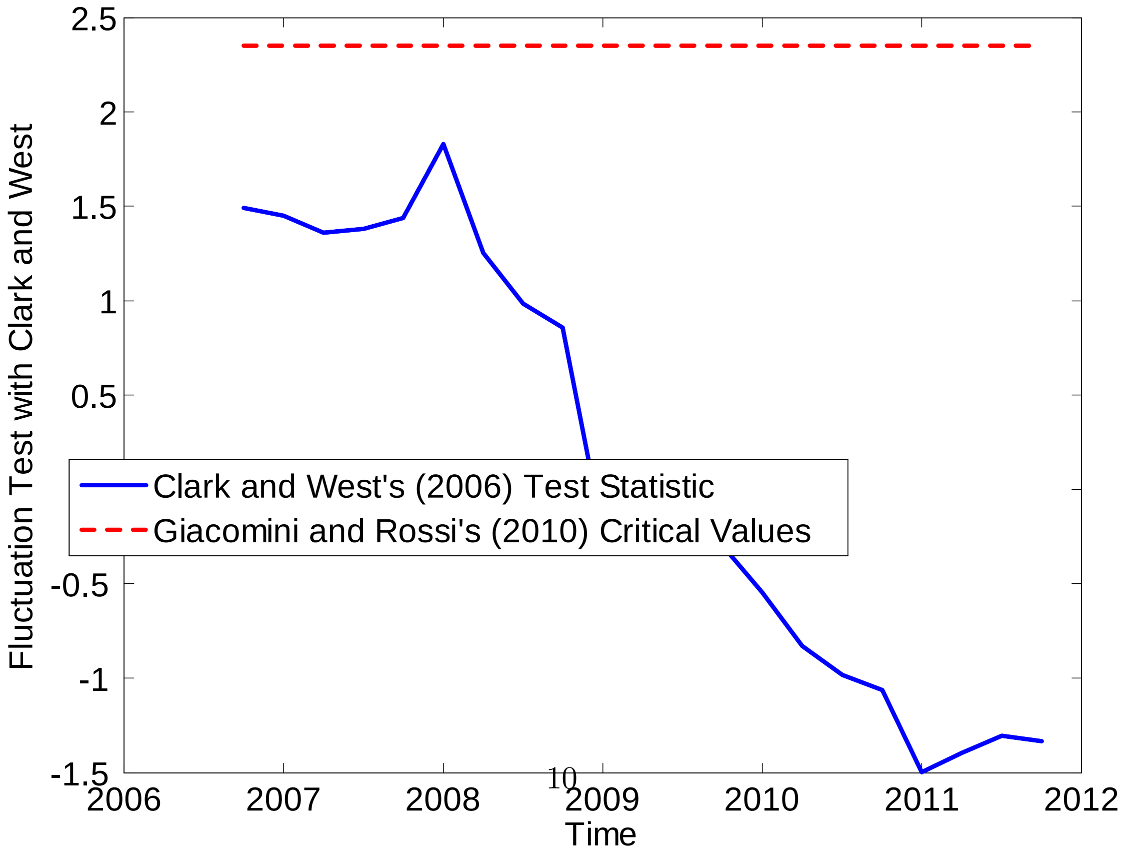

Figure 3. Panel A

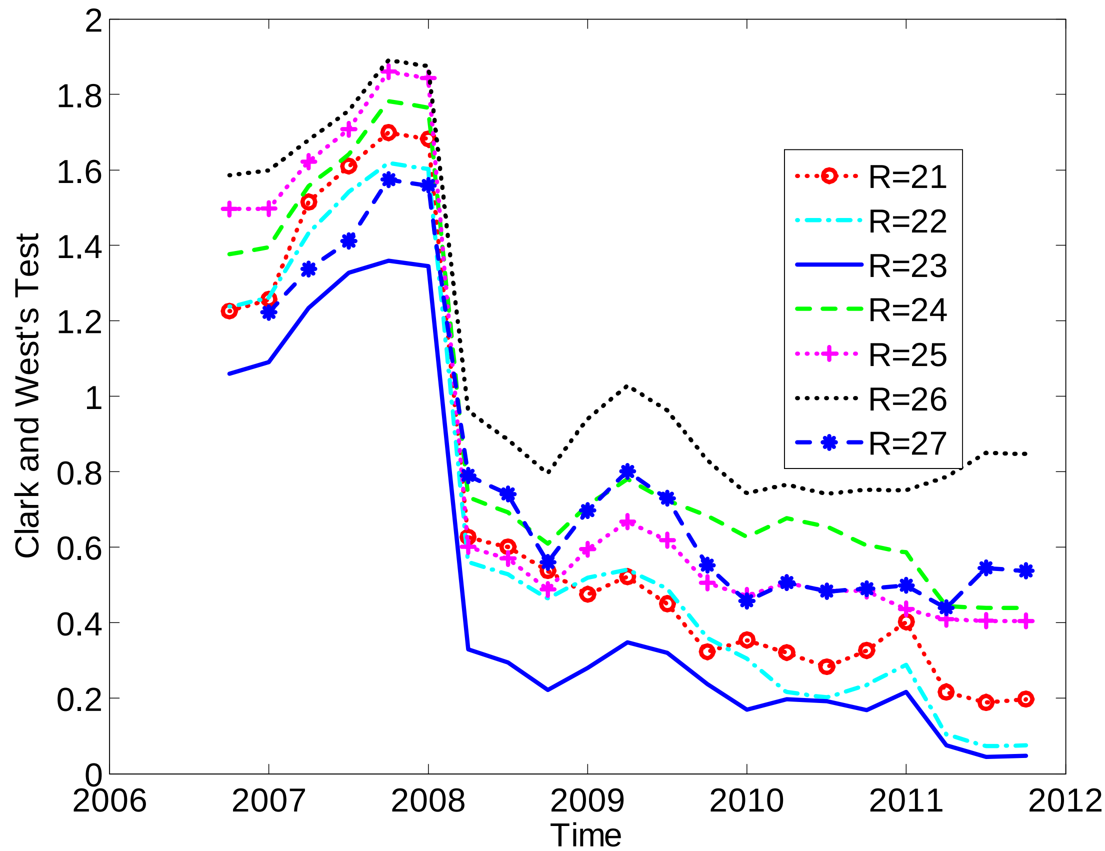

Figure 3. Panel B