

Bond University
Research Repository



A response surface analysis of critical values for the lead-lag ratio with application to high frequency and non-synchronous financial data

O'Neill, Michael; Rajaguru, Gulasekaran

Published in:
Accounting and Finance

DOI:
[10.1111/acfi.12546](https://doi.org/10.1111/acfi.12546)

Licence:
Other

[Link to output in Bond University research repository.](#)

Recommended citation(APA):
O'Neill, M., & Rajaguru, G. (2020). A response surface analysis of critical values for the lead-lag ratio with application to high frequency and non-synchronous financial data. *Accounting and Finance*, 60(4), 3979-3990. <https://doi.org/10.1111/acfi.12546>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

A Response Surface Analysis of Critical Values for the Lead-Lag Ratio with Application to High Frequency and Non-Synchronous Financial Data

Michael O'NEILL^a, and Gulasekaran RAJAGURU^{a*}

^a *Bond Business School, Bond University, Gold Coast, Australia – 4229*

Abstract

Granger causality tests are being supplanted by new methods such as the Lead-Lag Ratio, particularly in finance where data arrives at random times and systematic sampling often produces spurious results. Existing approaches are insufficient; outside of block-sampling using a bootstrap, the lead-lag ratio has generally been assessed against a benchmark of 1 without regard for statistical significance. We use simulations to generate a response surface for the Lead-Lag Ratio. Our modelled critical values are applied to reassess the findings of three previous studies of lead/lag relations between financial return series with high frequency data. Our response surface method proves to be a convenient and efficient alternative to using a bootstrap.

JEL classification: G10, G12, G13

Keywords: Lead-lag; Response Surface; High frequency

*Corresponding author: Gulasekaran Rajaguru, Email: rgulasek@bond.edu.au, Tel: +61 7 55952049.

Acknowledgements: We thank the Editor and an anonymous referee for their useful comments and suggestions. We thank James Todd for excellent research assistance.

1. Introduction

Most studies on financial data of causality rely on Granger causality tests using systematically sampled data. However, such studies can result in spurious results if the Granger causality runs from a non-stationary variable to either stationary or non-stationary variables (Rajaguru et al, 2018). On the other hand, genuine causal inferences are established when the cause variable is stationary in nature. These results are restricted based on synchronized data, and new methods have been introduced to overcome the issues with non-synchronous data where information arrives at random times (Hayashi and Yoshida, 2005).

The standard approach to calculating correlations would involve censoring the two series at arbitrary times, resulting in the duplication of some observations and elimination of others. In contrast, the method of Hayashi and Yoshida (2005) estimates the correlation of two series based on overlapping pairs of observations; that is, observations enter the calculation wherever the time period between observations overlaps. The cross-correlation estimate can then be recalculated with positive and negative lags between the two series.

A popular measure of whether one variable leads or lags another is the lead/lag ratio (*LLR*) which involves taking the sum of squared correlations across all positive lags divided by the sum of squared correlations across all negative lags (Huth and Abergel, 2014; Fung et al, 2015). Hoffmann et al (2013) use the contrast function defined by covariances to reach similar conclusions relating to lead/lag relations. The *LLR* measure is of increasing importance in analyzing flow of information and supply and demand in financial markets

which are traded at high frequency (Bollen et al, 2016). For example, if the sum of squared correlations across all lags of variable X against Y is greater than the converse, then X leads Y and the $LLR > 1$.

The problem with the LLR as a measure of leads and lags is that prior studies benchmark the LLR to a 50% critical value of 1 or simply comment on the magnitude of the LLR (Huth et al, 2014; De Jong and Nijman, 1997). The approach classifies causality from either X to Y or Y to X . But if the underlying data generating process is either bi-directional or no-causality between X and Y then the benchmark value of 1 will not be able to identify these two cases. Although confidence intervals are available for the Hiyashi–Yoshida covariance matrix (Bibinger and Mykland, 2013), it is not clear how to translate these into confidence intervals for the LLR . One solution has been to bootstrap the Hiyashi–Yoshida covariance matrix under the null using the empirical distribution, but this can be troublesome when dealing with high frequency time series data which arrives at random times. Resampling needs to be performed in blocks in order to preserve the correlation structure (Bollen et al, 2016).

Recent studies have focused on lead/lag relations between markets which trade at high frequency, with the aim of understanding volatility transmission between financial markets and providing better volatility forecasts.¹ For example, Bollen et al (2016) analyse the lead/lag relation between the CBOE VIX and the S&P 500 VIX short-term total return index (SPVXSTR).² They bootstrap the time series data to create 2,000 block resamples,

¹ Volatility forecasts are of increasing importance to investors who manage portfolio risk and may look to time the market. See, for example, Benson et al (2010), Do (2002), Gallagher et al (2014a,b), In et al (2014), Kim et al (2014), Jain and Jiang (2014), Jun et al (2014), Lee et al (2015), Nartea and Wu (2013), Nicholls and Tonuri (2006), O’Neill and Liu (2015), Sault (2005), Sinclair (2009) and Watson et al (2015).

² VIX is not a unique measure, nor is it necessarily the best predictor of future volatility. A number of studies have used VIX and other volatility indices to examine volatility transmission for other financial markets. See Jubinski and Lipton (2013), Liu and O’Neill (2016), Le (2017), O’Neill and Liu (2010), Wang

with 25 blocks of 20 minutes each per day, during the period December 20, 2005 through April 30, 2013. The authors highlight the issues with the bootstrap, suggesting that 2,000 replicates are likely to be sufficient, and blocks of 20 minutes are likely to be long enough to overcome the loss of dependence at the end of each block and to preserve correlation structures. This procedure provides 2,000 estimates for the *LLR* per day with associated confidence intervals.

Bollen et al (2016) focus on “Phase 4” of their study, between February 29, 2012 and April 30, 2013. Their central conclusion is that VIX futures trading raced ahead of VIX options trading during this period, and that the VIX futures index (SPVXSTR) began to lead the VIX itself, and to predict its direction (“the tail wagging the dog”). The bootstrap results show P-values less than or equal to 0.05 in 31% of days in this period, suggesting that the lead/lag relation with the SPVXSTR leading the VIX is significant at the 95% confidence level for 31% of days in Phase 4. In the same phase, VIX leads SPVXSTR with 95% confidence in none of the days.

The *LLR* is of increasing importance in determining lead/lag ratios in high frequency data with random arrival of information. Relying on an arbitrary benchmark such as 1 may be misleading. On the other hand, a two-tailed bootstrap confidence interval can be mis-leading since it identifies presence rather than direction of causality. We propose a simulation method to produce a response surface for the *LLR* in order to determine whether predictive lead/lag relations are statistically different from 1 with the

(2010) and Wang and Guo (2010). Implied volatility measures generally have better information content than other measures of historic volatility (Brace and Hodgson, 2009). In an Australian context, Hathaway (2009) finds that share price volatility is not stationary, while Yang and Allen (2005) investigate the optimisation of GARCH hedge ratios in futures markets. Brailsford and Faff (1996) apply existing models finding that ARCH and simple regression models provide better forecasts of volatility, but caution on the use of complex models.

one-tailed alternative. We then use these modelled values to reassess the statistical significance of lead/lag relations highlighted in two previous studies (Huth et al, 2014; De Jong and Nijman, 1997), and also compare our results to those obtained using the bootstrap method by Bollen et al (2016).

This study does not consider the theoretical validity of the bootstrap versus simulation methods. Further investigation could involve assessment of the circumstances under which assumptions for block-resampling differ from our simulated critical values. In any case, the simulated values represent a very convenient benchmark, particularly for high frequency analysis where a bootstrap might not be possible due to data or computing constraints.

2. Simulation Experiments and Response Surface Analysis

2.1. LLR Critical Values and Response Surface Function

The underlying data generating process considered is $y_t = \varepsilon_{1t}$, $x_t = \varepsilon_{2t}$, $\varepsilon_{it} \sim iidN(0,1)$, $t = 1, 2, \dots, T$; $i = 1, 2$. We estimate the *LLR* for the sample size of T at various lag lengths. The critical values are computed via stochastic simulation of 100,000 replications. We construct upper-tail critical values for *LLR* at the 1 percent, 5 percent and 10 percent levels of significance for sample sizes varying from $T = 100$ to $T = 10,000$.³ In traditional time series analysis critical values are purely a function of sample size T . Here, it is observed that the critical values for *LLR* are sensitive to both effective sample size T and lag p length. We construct *LLR* critical values at the three levels of significance for the

³ The *LLR* critical values are invariant to the mean and variances of the underlying normally distributed data generating process.

various combinations of sample size T and lag length p . For a given sample size T , the lag length p is expected to vary from 1 to p_{max} , where p_{max} is approximately equal to 10% of the sample size. We consider all combinations of $p \in \{1, 2, \dots, p_{max}\}$ and $T \in \{100, 110, 120, \dots, 10000\}$. For each combination of T and p , the LLR is calculated using

$$LLR = \frac{\sum_{i=1}^p \rho^2(l_i)}{\sum_{i=1}^p \rho^2(-l_i)}$$

where

$$\hat{\rho} = \frac{\sum_{i,j} r_i^X r_j^Y 1_{\{0, i, j \neq \emptyset\}}}{\sqrt{\sum_i (r_i^X)^2 \sum_j (r_j^Y)^2}}$$

In order to generalize the estimators of the critical values for any combination of sample size T and lag length p at a given level of significance, we use the response surface regression techniques proposed by MacKinnon (1994, 1996). Suppose that we are interested in $q^\alpha(T, p)$, i.e., α quantile of the distribution, where $\alpha = 90\%$, 95% and 99% . In each case, three response surfaces are estimated based on the 90th, 95th and 99th quantiles. Hence, a total of three response surface regressions are estimated.

In contrast to response surface regressions based on pure time series studies, in which the regression equation is a function of sample size T , we construct the response surfaces equation as a function of T and p and the response surface equation for the LLR :

$$q^\alpha(p, T) = \beta_0 + \beta_1 T^{-1} + \beta_2 T^{-2} + \beta_3 T^{-3} + \beta_4 p^{-1} + \beta_5 p^{-2} + \beta_6 p^{-3} + e \quad (1)$$

We also consider the possible interaction between T and p . Results have improved at the fourth decimal, and hence interaction terms are omitted from the analysis.⁴ We present results for the most parsimonious representation, and results for those models where an inclusion of interaction term can be obtained from authors upon request. In all three cases, the asymptotic critical value is approximately equal to the intercept β_0 , as lag length approaches infinity. The response surface at the 10%, 5% and 1% level of significance are reported in Table 1. It can be inferred from Table 1 that the asymptotic critical values for LLR at the 10%, 5% and 1% level of significance are 1.10, 1.13 and 1.18, respectively. This clearly indicates that the critical values obtained through stochastic simulation are far from the benchmark value of LLR equal to one, suggesting that findings from earlier literature could be misleading.

The independent variables in the response surface function are expected to be collinear. Since our objective is to obtain unbiased estimates to generate critical values at various levels of effective sample size and lag lengths, violation of this property will not be problematic.

Critical values at the 10%, 5% and 1% are plotted against effective sample size (T) and lag length (p) in Figure 1, panels a, b, and c, respectively. The vertical axis in these three figures is the simulated critical value.

Insert Table 1 HERE

Insert Figure 1 HERE

⁴ For example, the critical value at the 5% level of significance based on the response surface function represented by equation (1) is 1.6607 when $T=100$ and $p = 50$. On the other hand, the critical value for the same scenario with the interaction terms in the response surface regression is 1.6602.

2.2. Non-stationary data generating process

The *LLR* can be applied to examine the causal direction between stationary series. If the variables are non-stationary in nature, then we take an appropriate difference to convert them to stationary series. The unit root property is invariant to temporal aggregation and systematic sampling. Thus, the unit root property of the time series can be examined at equal intervals even if observations arrive at random time.⁵ In order to examine the validity of *LLR* for the non-stationary variables, we consider the following data generating process:

$$\begin{aligned}y_t &= y_s + \varepsilon_{1t} \\x_{t'} &= x_{s'} + \varepsilon_{2t'}\end{aligned}$$

Where t and t' are randomly chosen sequentially available non-synchronous time periods. y_s is previously available information at time t , $x_{s'}$ is previously available information at time t' and $\varepsilon_{1t}, \varepsilon_{2t'} \sim iidN(0,1)$. The *LLR* critical values are obtained for a given sample size T and the lag length p under two scenarios: (i) *LLR* between y_t and $x_{t'}$ and (ii) *LLR* between Δy_t and $\Delta x_{t'}$, where Δ denotes the difference between subsequent observations. The *LLR* critical values based on scenario (ii) are identical to the stationary case. The *LLR* based on scenario (i) will be spurious if both x and y are not cointegrated. We have also analysed the case where one of the variables is stationary and the other is non-stationary. The *LLR* is computed by taking appropriate difference of the non-stationary variable and the critical values are found to be same as the stationary case.

⁵ Marcellino (1999) shows that unit root property is invariant to temporal aggregation and systematic sampling.

2.3. Power of LLR

In order to analyse the usefulness of the response surface function under the case where the variables are causally related (i.e., the variables are coupled with each other), we construct the following data generating process. Without loss of generality, let the unidirectional causality run from the variable x to y , but not vice versa. The data generating process is

$$y_t = \varphi x_s + \varepsilon_{1t}$$

$$x_{t'} = \varepsilon_{2t'}$$

Where t and t' are randomly chosen sequentially available non-synchronous time periods. x_s is the previously available information at time t .

We compute the *LLR* for the various combinations of the parameter φ , sample size T and lag length p . We consider all combinations of $\varphi \in \{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, $p \in \{1, 2, 5, 10, 20, 30, 40, 50\}$ and $T \in \{100, 1000\}$. The LLR are computed via stochastic simulation of 100,000 replications. Table 2 summarizes the percentage of rejection frequencies of Granger non-causality from x to y at the 5% level of significance. The critical values at the 5% level of significance are benchmarked using the response surface function represented by equation (1). It can be observed from Table 2 that the genuine causal relationship from x to y is captured in almost all cases. The causal distortion is observed up to 6% of the cases when the causal parameter is close to 0 (i.e., $\varphi = 0.01$ and 0.05) and the lag length is high. We observe 100% accuracy in terms of power of the LLR test in establishing the genuine causal inferences for the case of parsimonious lag lengths.

Insert Table 2 HERE

3. Empirical Results

In this section we use our modelled critical values for the *LLR* to reassess the statistical significance of lead/lag relations highlighted in three previous studies (Huth et al, 2014; De Jong and Nijman, 1997; Bollen et al, 2016). Results are reported in Table 3. The methods used are robust to different time series dynamics in the datasets.

Beginning with Bollen et al (2016), we can apply the *LLR* confidence limits modelled in this study to Phase 4 to assess the significance of results and to compare results with the bootstrap approach used by the authors. There are 1,608 overlapping observations per day in Phase 4 and +/- 60 lags. Using the response surface functions, the critical values are 1.48 at 10%, 1.66 at 5% and 2.05 at 1%, being near the asymptotic limits. Applying these limits to the Bollen et al (2016) results leads to very similar conclusions. SPVXSTR leads VIX at the 95% significance level for 38% of days, and VIX leads SPVXSTR in 5% of days. This compares to 31% and 0% in Bollen et al (2016), respectively. It may be that the bootstrap is slightly more sensitive to the fatter tails of the empirical distribution, or that there is autocorrelation or cointegration between variables. It is not clear whether the errors in block sampling or the assumptions underlying the simulations are more limiting to the sensitivity of the test. We will endeavor to assess the relative validity of the two approaches in a further theoretical study. The response surface results are consistent with the bootstrap results of Bollen et al (2016), supporting the robustness of our method to different time series processes. Further, the power test in section 2 above demonstrates the robustness of methods to coupling or causality between variables.

Next we reassess the results of Huth et al (2014) for statistical significance. This study analyses the lead/lag relations between 42 stocks and 4 market indices. The number

of observations for different series can be determined from the trading hours summarized in Table 1 and the average time between observations in Table 3 shown in Huth et al (2014). Fifty four lags are listed at irregular intervals from 0, 0.01,...,300 seconds, allowing us to calculate the degrees of freedom for each pair as a range. For the stocks studied there are 8.5 hours of trading and 3-22 seconds average duration between trades, meaning 1,384 and 10,146 trades per day, with 54 lags on the grid. For futures, there are between 14 and 19.5 hrs trading with the maximum trading for Footsie 100 future (FFI) at around 51,166 trades per day with average spacing 1.372 seconds across 19.5 trading hours. Thus, the range of critical values can be determined for all pairs summarized, with the asymptotic levels being generally applicable. A ratio above 2.15 is significant at the 99% level, above 1.71 at the 95% level, and above 1.52 at the 90% level.

For example, for the pair CAC40 future (fce) and Total (totf.pa), although we are not told the number of overlapping observations entering the calculation of the *LLR*, we can take the limits as 9,321-35,644 with 54 lags in calculating a range of critical values.⁶ On this basis, the *LLR* of 1.85 is significant at the 5% significance level. Similarly, for the pair Renault (rena.pa) and peupa.pa, the *LLR* is 1.15 which is clearly not statistically significant with 3,026 to 5,280 overlapping observations and 54 lags.

De Jong and Nijman (1997) provide further empirical data to test our simulated *LLR* confidence limits. They analyze the S&P 500 index and its futures at 1, 5 and 10 minute intervals over the last quarter of 1993. This is similar to Bollen et al (2016) in that the futures on S&P 500 index lead the index itself. They censor data and give the number

⁶ There are 14 hours of trading in fce and an average of 1.414 seconds between trades meaning 35,644 observations per day. Similarly, with 8.5 trading hours and 3.283 average time between trades there are 9,321 observations per day in totf.pa.

of missing observations in Tables 1 and 2. Using data in their Table 3, we can sum the squared correlations at negative lags and divide by the squared correlations at positive lags giving an *LLR* of 3.825. We have 7,989 observations for the index and 6,807 for the futures (with 13% missing). With between 6,807 and 7,989 overlapping observations and +/-15 lags we have an effective sample size of between 6,792 and 7,974. Thus, we can test the *LLR* against a 90% critical value of 2.41, 95% of 2.92 and 99% of 4.54. We conclude that the results are again significant at the 95% level.

In Bollen et al (2016), Huth et al (2014) and De Jong and Nijman (1997) the method is applied to return series. The potential issue of non stationarity of time series processes is therefore avoided through differencing, and the method remains robust to different time series dynamics in these prior studies.

Insert Table 3 HERE

4. Conclusion

The Lead-Lag Ratio is a new measure of lead/lag relations that is becoming an important alternative to Granger causality analysis with high frequency and non-synchronous data. However, benchmarking the *LLR* to 1 can be misleading. This study conducts extensive Monte Carlo simulations to generate a response surface for the *LLR*. We generate critical values via stochastic simulation and construct the upper-tail critical values for *LLR* at the 1 percent, 5 percent and 10 percent levels of significance for sample sizes varying from $T = 100$ to $T = 10,000$. Critical values for *LLR* are sensitive to both effective sample size and lag length when we use lag lengths of up to 10% of the sample

size. The estimators are generalized using response surface regression techniques, with all combinations of lags and sample sizes within these limits. We validate these methods through application of modelled critical values to three prior studies. We also demonstrate the similarity of our results with an empirical bootstrap. Simulated values prove to be a more efficient and convenient benchmark than existing measures. There are also advantages of the simulation approach over an empirical bootstrap method, particularly in terms of computational intensity for high frequency data. We expect the *LLR* response surface to be of increasing importance as trading frequency between markets increases, and synchronicity of markets and information flow continues to diverge.

References

Benson, K. L., R. W. Faff, and T. Smith, 2010, The simultaneous relation between fund flows and returns, *Australian Journal of Management* 35, 51–68.

Bollen, N.P., M. O'Neill, , and R.E. Whaley, 2016, Tail Wags Dog: Intraday Price Discovery in VIX Markets, *Journal of Futures Markets* 37(5), 431-451.

Bibinger, M., and P.A. Mykland, 2013, Inference for multi-dimensional high-frequency data: Equivalence of methods, central limit theorems, and an application to conditional independence testing, online at <http://arxiv.org/pdf/1301.2074.pdf>.

Brace, A., and A. Hodgson, 2009, Index futures options in Australia – an empirical focus on volatility, *Accounting and Finance* 31, 13–30.

Brailsford, T., and R. Faff, 1996, An evaluation of volatility forecasting techniques, *Journal of Banking and Finance* 20, 419–438.

De Jong, F., and T. Nijman, 1997, High frequency analysis of lead-lag relationships between financial markets, *Journal of Empirical Finance* 4(2–3), 259–277.

Do, H. D., 2002, Relative performance of dynamic portfolio insurance strategies: Australian evidence, *Accounting and Finance* 42, 279–296.

Doung, L., 2017, The Information Content of Implied Volatility in the Crude Oil and Natural Gas Markets, *Accounting and Finance* 17(9), 15-30.

Fung, J.K.W, F. Lau, and Y. Tse, 2015, The impact of sampling frequency on intraday correlation and lead/lag relations between index futures and individual stocks, *Journal of Futures Markets* 35, 939–652.

Hathaway, N., 2009, The non-stationarity of share price volatility, *Accounting and Finance* 26, 35–54.

Hayashi, T., and N. Yoshida, 2005, On covariance estimation of non-synchronously observed diffusion processes, *Bernoulli* 11, 359–379.

Hoffmann, M., M. Rosenbaum, and N. Yoshida, 2013, Estimation of the lead-lag parameter from non-synchronous data, *Bernoulli* 19, 426–461.

Huth, N., and F. Abergel, 2014, High frequency lead/lag relationships—Empirical facts, *Journal of Empirical Finance* 26, 41–58.

In, F., S. Kim, and P.I. Ji, 2014, On timing ability in Australian managed funds, *Australian Journal of Management* 39, 93–106.

Jain, P., and C. Jiang, 2014, Predicting future price volatility: empirical evidence from an emerging limit order market, *Pacific-Basin Finance Journal* 27, 72–93.

Jubinski, D. and A.F. Lipton, 2013, VIX, Gold, Silver, and Oil: How do Commodities React to Financial Market Volatility? *Accounting and Finance* 13(1), 70-88.

Jun, X., M. Li, and J. Shi, 2014, Volatile market condition and investor clientele effects on mutual fund flow performance relationship, *Pacific-Basin Finance Journal* 29, 310–334.

Kim, S. W., B. S. Lee, and Y. M. Kim, 2014, Who mimics whom in the equity fund market? Evidence from the Korean equity fund market, *Pacific-Basin Finance Journal* 29, 199–218.

Liu, Z. and M.J. O’Neill, 2016, Partial moment volatility indices, *Accounting and Finance* 58(1), 195-215.

MacKinnon, J.G., 1994, Approximate asymptotic distribution function for unit-root and cointegration tests, *Journal of Business and Economic Statistics* 167-176.

MacKinnon, J.G., 1996, Numerical distribution functions for unit root and cointegration tests, *Journal of Applied Econometrics* 11, 601-618.

Marcellino, M., 1999, Some consequences of temporal aggregation in empirical analysis, *Journal of Business and Economic Statistics*, 17(1), 129-136.

O'Neill, M., and Z. Liu, 2015, Fund volatility index using equity market state prices. *Accounting and Finance* 57(3), 837-853.

O'Neill, M.J., and Z. Liu, 2016, Tail risk hedging for mutual funds using equity market state prices. *Australian Journal of Management* 41(4), 687-698.

Rajaguru, G., M. O'Neill, and T. Abeysinghe, 2017, Does Systematic Sampling Preserve Granger Causality with an Application to High Frequency Financial Data? *Econometrics* 6(2), 1-17.

Sault, S., 2005, Movements in Australian stock volatility: a disaggregated approach, *Australian Journal of Management* 30, 303–320.

Sinclair, N.A., 2009, Market timing ability of pooled superannuation funds January 1981 to December 1987, *Accounting and Finance* 30, 51–65.

Wang, K., 2010, Forecasting volatilities in equity, bond and money markets: A market-based approach, *Australian Journal of Management* 35(2), 165-180.

Wang, K., and Y. Guo, 2013, Predictability of time-varying jump premiums: Evidence based on calibration, *Australian Journal of Management* 39(3), 369-394.

Watson, J., J. Delaney, M. Dempsey, and J. Wickramanayake, 2015, Australian superannuation, pension, fund product ratings and performance: a guide for fund managers, *Australian Journal of Management* 41(2), 189-211.

Yang, W., and D. Allen, 2005, Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets, *Accounting and Finance* 45, 301–321.

Table 1: Estimates of Response Surface Functions

Variable	Response Surface Function 10% - significance level		Response Surface Function 5% -significance level		Response Surface Function 1% - significance level	
	Coefficient	Se [#]	Coefficient	Se [#]	Coefficient	Se [#]
C	1.10***	0.00	1.13***	0.01	1.18***	0.01
1/T	-4.67*	2.76	-8.22*	4.32	-24.64*	13.75
1/T²	-1809.88*	953.78	-2249.86*	1252.97	281.29**	102.97
1/T³	174147.20**	80329.11	212417.00**	100053.1	-45305.34	298567.12
1/P	29.86***	0.92	39.92***	1.39	62.34***	3.78
1/P²	-461.21***	56.00	-533.77***	80.75	-652.78***	209.47
1/P³	4633.66***	999.58	5071.46***	1279.25	7111.98**	3314.30
R²	0.9987		0.9943		0.9891	

*, ** and *** indicates significance at the 10%, 5% and 1% level, respectively.

White heteroscedasticity consistent standard errors.

Note: Table summarizes the estimates of response surface functions used to generate LLR critical values at the 10, 5 and 1% significance levels. The estimated intercepts 1.10, 1.13 and 1.18 indicate asymptotic critical values at the 10%, 5% and 1% levels, respectively. The inverse of the effective sample size and lag lengths are used to ensure convergence.

Table 2: Power of LLR test at the 5% level of significance

Panel A: T=100											
φ	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>p</i>											
1	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
10	98%	98%	100%	100%	100%	100%	100%	100%	100%	100%	100%
20	95%	95%	100%	100%	100%	100%	100%	100%	100%	100%	100%
30	95%	95%	96%	98%	99%	100%	100%	100%	100%	100%	100%
40	94%	94%	95%	97%	99%	100%	100%	100%	100%	100%	100%
50	94%	94%	94%	96%	98%	99%	100%	100%	99%	100%	100%
Panel B: T=1000											
φ	0.01	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>p</i>											
1	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
10	99%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%
20	95%	97%	100%	100%	100%	100%	100%	100%	100%	100%	100%
30	95%	96%	99%	100%	100%	100%	100%	100%	100%	100%	100%
40	94%	96%	98%	100%	100%	100%	100%	100%	100%	100%	100%
50	94%	95%	99%	100%	100%	100%	100%	100%	100%	100%	100%

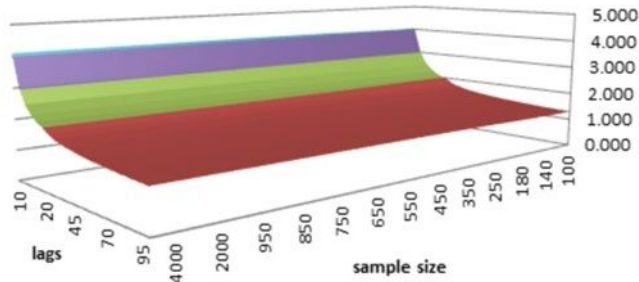
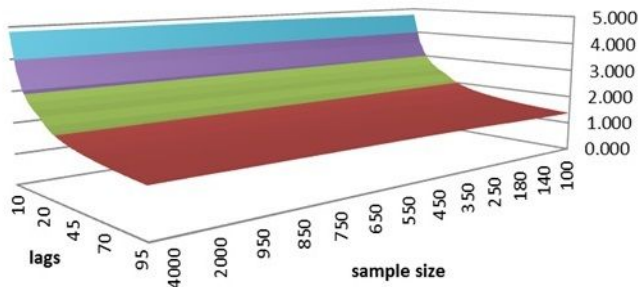
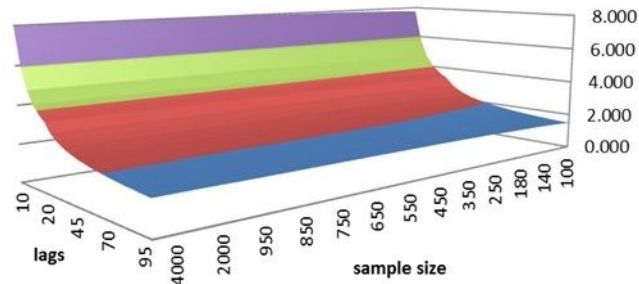
Note: table summarizes the percentage rejection frequencies of Granger non-causality from x to y at the 5% level of significance. The underlying data generating process assumes unidirectional causality from x to y.

Table 3 – Validation of Prior Studies using *LLR*

Example	Lags	Observations	Critical Value*		
			90%	95%	99%
Bollen et al (2016)	+/-60	1,608	1.48	1.66	2.05
Huth et al (2014)	54 (0, 0.01,...,300s)	1,384-51,166	1.52	1.71	2.15
De Jong and Nijman (1997)	+/-15	6,807-7,989	2.41	2.92	4.54

* critical values are based on response surface functions.

Note: table summarizes the modelled critical values for three prior studies. The number of lags used in these studies is combined with the range of overlapping observations allowing us to estimate response surface functions and to generate the 90, 95 and 99% critical values, to reassess the statistical significance of prior results.

Figure 1: Response Surface for *LLR***a. 10% Critical Values****b. 5% Critical Values****c. 1% Critical Values**

Note: Figure shows the modelled critical values at the 10, 5 and 1% levels in panels a, b and c, respectively.