Long Run Confidence: Estimating Confidence Intervals when using Long Run Multipliers

Mark David Nieman* and David A.M. Peterson†

Abstract

The recent exchange on Error Correction Models in *Political Analysis* and elsewhere dealt with several important issues involved in time series analysis. While there was much disagreement in the symposium, one common theme was the lack of power due to the few number of observations for much of this work. In this paper we highlight two well known but rarely discussed problems this has for inferences from standard time series techniques. First, one result of low power is inflated standard errors. One issue low powered time series can face is that the confidence interval on a lagged dependent variable, even when the series is stationary, includes values ≥ 1. This is particularly problematic when calculating the confidence interval of the long run multiplier. If the confidence interval of the lagged dependent variable includes 1, the standard error of the long run multiplier will be explosive. Second, the calculation of the long run multiplier is the ratio of coefficients, which makes the calculation of the uncertainty slightly more complicated. Unfortunately, the two standard approaches to calculating the uncertainty in the long run multiplier, the delta method and the Bewley transformation, are asymptotically accurate, but may have difficulties in small samples. As a solution, we suggest using a Bayesian approach. For autoregressive distributed lag models, the Bayesian approach can formalizes the stationarity assumption by using a beta prior that is strictly less than 1. With error correction models, the researcher can easily calculate the credible region of the long run multiplier from the posterior distribution of the ratio of the coefficients. As a result, we obtain theoretically informed estimates of the confidence regions for the distribution of the long run multiplier.

*Assistant Professor of Political Science, Iowa State University.
†Professor of Political Science, Iowa State University.

1. INTRODUCTION

The recent exchange on the utility of and concerns with error correction models did not generate much consensus on the appropriateness of those models (e.g., Enns et al, 2016; Grant and Lebo, 2016; Keele, Linn, and Webb, 2016). What is notable, however, that there were at last two places of agreement between the authors that have implications for users of time series models. First, as most clearly articulated by DeBoef and Keele (2008), it is important for practitioners to calculate the long run multiplier (LRM) and its standard error to present a complete picture of the effect of an independent variable. Second, many of our time series applications are short and the small number of observations limits the power of the estimation of specification tests and important parameters.[1]

This creates potential difficulties for two of the predominant approaches to estimating models of stationary time series. The standard set up for an ADL(1,1;1) is:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t. \tag{1}$$

In the case of an autoregressive distributed lag (ADL) model, one potential problem is that the lack of power in the estimation may lead to the estimation of the coefficient for the lagged dependent variable, even if the series is stationary, where the confidence interval includes 1.[2] The calculation of the LRM in this model is $lrm_{\text{ADL}} = \frac{\beta_0+\beta_1}{1-\alpha_1}$ where the α_1 term is the coefficient on the lagged dependent variable. If the confidence interval for that term includes 1.0, then denominator of that LRM will include zero. In that case, the estimation of the variance of the LRM will be "mildly explosive"[3] (which is bad).

A second approach, the error correction model (ECM) does not tend have the same problem. Existing measures of uncertainty for the LRM are typically calculated from an ECM, which DeBoef and Keele (2008: 189-190) have shown can be written as mathematically equivalent to the ADL. The basic form of the ECM, analogous to the ADL model in equation 1 is:

$$\Delta Y_t = \alpha_0 + (\alpha_{1-1}) Y_{t-1} + \beta_0 \Delta X_t + (\beta_0 + \beta_1) X_{t-1} + \epsilon_t \tag{2}$$

and the LRM is calculated as: $LRM_{ECM} = \frac{\beta_0+\beta_1}{\alpha-1}$.

The "mildly explosive" LRM problem for the coefficient estimate for lagged value of Y_t would be for it to include 1.0 in its confidence interval. Based on our examination of the empirical time series literature in political science this does not seem to be a problem for ECMs. These models, however, often face a different potential problem. Because the LRM is calculated as the ratio of coefficients, the estimation of the error variances is more complicated. The two common solutions are either the Bewley transformation or the delta

[1] Freeman (2016) provides a counter to this ubiquity of small T time series.

[2] One may suggest that if the confidence interval for α_1 contains 1, then the researcher should not treat the series as stationary. While this might seem like a reasonable suggestion, there may be cases where stationarity is still appropriate for small samples. Undoubtedly, if a the coefficient on the lagged dependent variable is not significantly different than 1, that series is unlikely to reject the null of a unit root in stationarity tests. Researchers, however, often have more knowledge about the properties of their dependent variable beyond the small set of observations used. Researchers who are limited by the timespan of their independent variables, for instance, may have a series that is stationary with larger T than they can use in their analyses. The small T also means that that the stationarity tests themselves are often weak.

[3] We borrow this phrase from Hill and Peng (2014, 293) and Hill, Li, and Peng (2016, 126).

method. Both of these are easy to implement and provide asymptotically accurate estimates of the variance in the LRM. But the small sample properties of these estimates are less clear and may lead to inaccurate estimates of the variance of the LRM.

In this paper, we examine the effects of having a short series on the estimation of the LRM in particular and suggest a simple fix that will provide more accurate estimates of the uncertainty in the LRM. Specifically, we propose a straightforward Bayesian approach to solve both of these potential difficulties. For the ADL models, this involves placing a prior on the lagged dependent variable in an ADL model that constrains the coefficient of the lagged dependent variable to be strictly between zero and one. This prior simply incorporates the decision to treat the series as stationary and, as we will show, leads to much more accurate estimates of the uncertainty in the LRM and the specification of the ADL. ECMs do not seem to need this prior. But estimating the ECM via MCMC makes the calculation of the variance of the ratio of coefficients straightforward and does not require appeals to asymptotic theory.

In the next section, we discuss the general problem, outlining the main methods for calculating the LRM and its uncertainty. We then present a series of Monte Carlos that illustrate the problems with each of theses approaches and the improvements due to the inclusion of a Beta prior on the lagged dependent variable in an ADL. Finally, we provide two replications to illustrate how this approach can help with the estimates of the uncertainty in the LRM.

2. UNCERTAINTY AND THE LRM

The notion that we should provide uncertainty estimates for estimated quantities of interest should be obvious. This is standard practice for reporting results from statistical models. For most cross sectional linear regression models, this is quite easy. The main quantity of interest is captured by the coefficient estimates and their standard errors. Models with limited dependent variables or interaction terms can be more complicated, but it is still rare to see the reporting of some quantity of interest without the corresponding uncertainty estimate.

In time series work, this is less common. If a long run relationship exists between two variables, the regression coefficients do not directly capture the extent of how differences in the independent variable correspond to differences in the dependent variable. Because the effects are likely to accumulate over time, the better way to calculate the full effect of X_t on Y_t is to calculate the long run multiplier. In most time series specifications, the point estimate is quite easy. It is simply the ratio of two coefficients. The more difficult part is the calculation of the uncertainty around this point estimate.

In the literature there are three main approaches to calculating the calculating variance of the LRM. As DeBoef and Keele (2008) note, neither the ADL nor the ECM provide a direct estimate of the variance of LRM. Since the LRM is a ratio of coefficients, the calculation of the variance of the ratio of coefficients with known variances can be used. The formula is:

$$Var(\frac{a}{b}) = (\frac{1}{b^2})Var(a) + (\frac{a^2}{b^4})Var(b) - 2(\frac{a}{b^3})Cov(a,b). \tag{3}$$

There are two common alternatives to this calculation. The first alternative calculation of the LRM from an ECM is to use the Bewley (1979) transformation, which estimates the variance of the LRM directly.[4] The Bewley transformation is:

$$Y_t = \alpha_0\phi - \alpha_1\phi\Delta Y_t + \phi(\beta_0 + \beta_1)X_t + \phi\beta_1\Delta X_t + \phi\epsilon_t \tag{4}$$

where $\phi = (\frac{1}{\alpha_1 - 1})$ and an instrument for ΔY_t is calculated as the predicted values from the equation $\Delta Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \gamma_2 X_t + \gamma_3 \Delta X_t + \epsilon_t$. The LRM is the coefficient on X_t from equation 3. The second approach is to use the delta method. The delta method relies on expanding a random variable (in this case the LRM) via a Taylor series and calculating the resulting asymptotic variance of this estimate.

In small samples, however, both of these techniques face difficulties. When T is small, the estimate of α_1 can be imprecise. Normally, this lack of power will simple lead to inflated standard errors and not many other issues. In this case, however, there is an additional potential pitfall. If the confidence interval of α_1 includes 1, the variance calculation will be "slightly explosive". This can be seen in the calculation of the LRM. If the value of α_1 has probability mass at 1 the LRM is undefined for that point. If there is mass where $\alpha_1 > 1$, the denominator will be negative. In each case, the calculation of the variance breaks down.

Our intuition is that the difficulty with the small sample variance estimates for the Bewley method is partially responsible for Webb, Linn, and Lebo (2019) (hereafter, WLL) need for stochastic simulations to calculate the bounds on their proposed LRM test for the presence of the long run relationship. WLL demonstrate convincingly that a clear test of the presence on long run relationship between X_t and Y_t is captured by the significance of the LRM. As they note *"Thus, a nondegenerate, or valid, equilibrium relationship between y_t and x_t requires the LRM to be nonzero"* (p. 7, emphasis in original). This is a vitally important result. They demonstrate that this is true regardless of the stationarity of y_t, helping resolve much of the uncertainty in pre-analysis specification tests of our time series. Additionally, because the LRM is calculated separately for each of the independent variables, this approach allows the researcher to know which of the variables have a significant long run relationship with y_t, while the ECM based tests only indicate that at least one independent variable has a long run relationship.

Given the importance of the LRM for specification testing, WLL 2019 also empirically explore the appropriate distribution of the test statistics for the LRM based on the Bewley transformation. While the exact amount of information in the uncertainty estimates and the critical values for the LRM depend on the sample size and the degree of autoregression y_t and x_t, in general they find that the critical values do no follow a standard distribution. Instead, they estimate these critical values via a stochastic simulation to determine the bounds of the test. Their conclusion is that most empirical tests of the LRM are likely over confident in the hypothesis tests. More importantly, they develop the bounds for the hypothesis tests of the long run relationship. These bounds will guide a researcher to conclude that whether or not there is a long run relationship.

This is a tremendous step forward for applied time series. What may be unsatisfying for many researchers, however, is that the bounds have a relatively large range of middle

[4] The Bewley transformation is a computational convenience and is not a theoretical model.

values that are inconclusive. For many empirical research questions, the conclusion may end up with the unsatisfying result of a test statistic between the bounds and an uncertain conclusion. This is the intellectually honest answer, but one we think will be frustrating for many researchers. If an alternative estimation of the uncertainty in the LRM can provide a more precise test of the significance of the LRM, this might be a more fruitful approach for the applied researcher.

3. BAYESIAN APPROACH

Our approach to this problem is to use a Bayesian framework. This allows us to fully integrate the stationarity of the dependent variable by adding a prior on the coefficient that will constrain it to be strictly between 0 and 1. In the example presented here, we use the prior such that $\alpha_1 \sim \mathcal{B}(1,1)$ to accomplish this. This is a diffuse prior that places equal probability on all values between 0 and 1. A more informed prior, which places more weight to values closer to 1 may be practical, but for this example we illustrate the influence of this design with a flat prior.[5]

One way to think about this prior is that it is simply the formalization of stationarity. When a researcher treats a series a stationary, she is assume that the root of the characteristic equation of the time series is less than one. The use of this prior constrains the estimate of α_1 to be less than 1, precisely the implication of treating the dependent variable as stationary.

The actual estimation of the model is carried out via a Markov chain Monte Carlo (MCMC). This allows us to calculate the distribution of the posterior for all of the coefficients directly. Rather than using an asymptotic equivalent to the confidence interval of the LRM or relying on a formula that is not easily available in most statistical output, we can calculate the LRM for each of the draws from the posterior in the MCMC and using this distribution to summarize the credible region of the LRM.[6] This is one of the virtues of inference from the posterior of an MCMC. A researcher can estimate the distribution of functions (such as ratios) of an unknown parameter or parameters directly from the posterior distribution of the MCMC (Gelfand et al, 1990). As such, we do not need to rely on the asymptotic properties of the variance estimator of the LRM and should provide more accurate estimates of the credible region for the LRM than either the Bewley method or the Delta method.

At this point, this claim is little more than conjecture. In the next section we provide some initial Monte Carlo evidence of this, but this is most likely not satisfying to the reader. Our intended next step is to replicate the dynamic simulations of WLL 2019, but estimating the distribution of the LRM via an MCMC approach instead of OLS.

4. MONTE CARLO ANALYSIS

To compare the properties of this approach versus the standard techniques, we conduct two Monte Carlos. We generate data following an ADL(1,1;1). To explore the small sample

[5]If one has theoretical reasons to expect the coefficient on the lagged dependent variable to approach 1, a Beta prior assuming a pmf massed near 1, such as $\mathcal{B}(5,2)$, could be used. For the sake of demonstrating the approach, however, we use $\mathcal{B}(1,1)$, which will calculate a more conservative LRM.

[6]Implementing this approach can be easily accomplished in common statistical programs, such as R or Stata.

properties, we set $T = 50$. We generate the exogenous variable so that $X_t = \gamma X_{t-1} + \eta_t$ where $\gamma = 0.5$.[7] We generate the endogenous variable so that $Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t$ where $\alpha_0 = 0$, $\beta_0 = 0.5$, and $\beta_1 = 0.25$. η_t and ϵ_t are drawn from a standard normal where $\text{cov}(\eta_t, \epsilon_t = 0)$. We vary $\alpha_1 = \{0.5, 0.9\}$. In the case where $\alpha_1 = 0.5$, the LRM is $\frac{0.5+0.25}{1-0.5} = 1.5$; when $\alpha_1 = 0.5$, the LRM is $\frac{0.5+0.25}{1-0.9} = 7.5$.

We estimate an ADL, ECM, and Bayesian ADL (B-ADL) model.[8] Priors for the Bayesian ADL are: $\beta_j \sim \mathcal{N}(0, 20)$, $\alpha_0 \sim \mathcal{N}(0, 20)$, $\alpha_1 \sim \mathcal{B}(1, 1)$, and $\epsilon \sim \mathcal{G}(1, 10)$. The Beta prior has a uniform probability mass function (PMF) strictly between 0 and 1. Each Bayesian ADL uses 10,000 MCMCs after a 2,500 burnin. Estimates are based on 250 simulations for each value of α_1.

5. RESULTS

Table 1 reports the results from the Monte Carlos. Each of the models recovers the expected coefficients when $\alpha_1 = 0.5$. Moreover, the LRM is also correctly estimated by each of the models. When $\alpha_1 = 0.9$, we again find that each of the models is able to recover the correct coefficients. The LRM is the same for the ADL and ECM (both using the calculation of the variance of the ratio of coefficients, or *ECM-ratio*, and the Bewley transformation, or *ECM-Bewley*), while the B-ADL somewhat underestimates the LRM. The latter result stems from the inclusion of the flat prior in the lower power context. Overall, each of the models performs well in terms of bias.

Turning to the estimates of uncertainty, we find that the standard errors on the coefficients are consistent across models for both $\alpha_1 = 0.5$ and $\alpha_1 = 0.9$. With small sample sizes, none of the models indicate that β_1 is statistically significant, even though it is a predictor of Y_t in the true data generating process. This is potentially problematic, as applied researchers often disregard insignificant lags of independent variables, which would result in biased estimates of the LRM.

Uncertainty estimates for the LRM differ substantially across models. Figure 1 report the estimated standard errors for the B-ADL, ECM-ratio, and ECM-Bewley for $\alpha_1 = 0.5$ and $\alpha_1 = 0.9$. When $\alpha_1 = 0.5$, the B-ADL model has a SE that over twice as large as the ECM-ratio, and the ECM-ratio is almost twice as large as ECM-Bewley. The former is accounted for by the influence of the prior given the small sample size. The latter result reflects the fact that ECM-Bewley is an instrumental approach and thus tends to underestimate the level of uncertainty in a small sample setting. The difference between the ECM-ratio and ECM-Bewley techniques demonstrates that, while they recover the same asymptotic results, they differ in small samples.

When $\alpha_1 = 0.9$, differences in the estimates of the SE are even more stark. SEs for the LRM are approximately half the size of the coefficient for B-ADL, while the SEs for ADL are over 5.5, reflecting the fact that ADL contains some simulations where the confidence interval are "mildly explosive," or greater than 1. Figure 2 makes this point more clear. The prior on B-ADL forces estimates of the standard errors to remain within the range of

[7]The initial X is drawn from a standard normal distribution.

[8]The LRM for the B-ADL is calculated by recovering the median posterior. We also calculate the LRM for the ECM using the Bewley transformation.

Table 1: B-ADL, ADL, ECM Estimates.

	$\alpha_1 = 0.5$				$\alpha_1 = 0.9$		
Variable	B-ADL	ADL	ECM	Variable	B-ADL	ADL	ECM
Y_{t-1}	0.46	0.46	-0.54	Y_{t-1}	0.87	0.87	-0.13
	(0.11)	(0.11)	(0.11)		(0.05)	(0.05)	(0.05)
X_t	0.49	0.49		X_t	0.48	0.48	
	(0.17)	(0.17)			(0.17)	(0.17)	
X_{t-1}	0.28	0.28	0.77	X_{t-1}	0.29	0.28	0.77
	(0.20)	(0.19)	(0.16)		(0.17)	(0.17)	(0.13)
ΔX_t			0.49	ΔX_t			0.48
			(0.17)				(0.17)
Const.	0.00	0.00	0.00	Const.	0.03	0.03	0.03
	(0.15)	(0.15)	(0.15)		(0.16)	(0.16)	(0.16)
LRM	1.47	1.47	1.47	LRM	7.08	7.29	7.29
	(0.81)		(0.29)		(3.41)		(5.53)
T	49	49	49	T	49	49	49

Note: Bewley LRM($\beta_{\alpha_1=0.5}$)=1.47, LRM($SE_{\alpha_1=0.5}$)=0.15 and Bewley LRM($\beta_{\alpha_1=0.9}$)=7.29, LRM($SE_{\alpha_1=0.9}$)=0.39. B-ADL LRM: HDP 95%$CI_{\alpha_1=0.5}$=[0.89, 2.20], HDP 95%$CI_{\alpha_1=0.9}$=[2.27,32.02].

theoretically possible outcomes, whereas a non-insignificant proportion of the upper 95% confidence intervals for ADL are greater than 1. In contrast, the SEs for ECM-Bewley are much smaller than those of either the B-ADL or ADL models.

6. APPLICATIONS

We demonstrate the value of our approach with 3 applications. The first shows the potential difficulty of calculating the variance estimate of the LRM from an ADL with a small sample

Figure 1: Standard Error on Long Run Multiplier

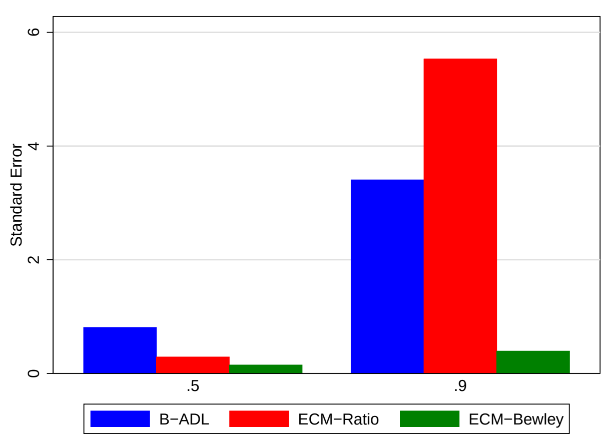
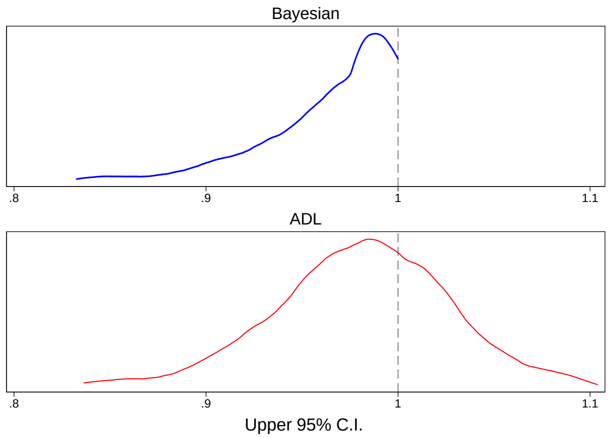


Figure 2: Upper C.I. on Lagged DV when $\alpha_1 = 0.9$



Note: 1 is the theoretical upper bound.

of mildly explosive error variance. The others apply our approach to the same applications as WLL 2019.

6.1. *Guns, Butter, and Deficits*

Heo and Bohte 2012 examine the trade off between military expenditures, domestic spending, and the role of deficits. While most of their emphasis is on the guns versus butter tradeoff, they also demonstrate that increased defense spending is related to increased deficit spending. It is this latter finding that we are exploring. We make this choice because it fits the profile of the type of ADL model we highlight above. The specifications tests that Heo and Bohte conduct indicate that the deficit spending measure is stationary and they analyze the data as a straightforward model with a lagged dependent variable.[9] The point estimate of the lagged dependent variable in their results is 0.82. While the lagged dependent variable in a model with an integrated series will have a downward bias, this is still far enough away from 1.0 to probably not generate much skepticism. The difficulty is that with only 59 observations and 11 independent variables, the model has potentially inflated standard errors. The standard error on the effect of the lagged dependent variable is 0.12, which means that the confidence interval of the effect of the lagged dependent variable contains 1.0.

Table 2 presents the results from three models. The first is the results from the original Heo and Bohte paper, which we can successfully replicate from their data available on Dataverse. The second column presents the same model, but instead of estimating the full seemingly unrelated regression model of all three equations, we estimate a simple ADL model. In this design, the coefficient on the lagged dependent variable is substantively smaller, but the standard errors are still large enough so that the confidence interval includes 1.0. The

[9] To be clear, they estimate a seemingly unrelated regression for three different models. The other two are models with differenced dependent variables capturing the guns versus butter tradeoff. We have estimated our application as a seemingly unrelated regression and reach similar conclusions, but this application is more straightforward.

Table 2: B-ADL, ADL, ECM Estimates.

Variable	ADL (SUR)	ADL	B-ADL
Y_{t-1}	0.82	0.75	0.82
	(0.12)	(0.14)	(0.04)
Δ(Milex/GDP)	-0.42	-0.38	0.41
	(0.12)	(0.13)	(0.04)
War	-0.00	- 0.00	-0.00
	(0.00)	(0.00)	(0.00)
Δ GDP$_{t-1}$	0.36	0.34	0.36
	(0.05)	(0.06)	(0.03)
Δ Unemployment$_{t-1}$	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
Inflation	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
DPresRCong$_{t-1}$	0.00	0.01	0.00
	(0.00)	(0.01)	(0.00)
RPresDCong$_{t-1}$	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
RPresSplitCong$_{t-1}$	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)
ΔCongress$_{t-1}$	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Δ (Debt/GDP)$_{t-1}$	0.07	0.04	0.07
	(0.05)	(0.06)	(0.03)
Const.	-0.01	0.00	-0.01
	(0.003)	(0.03)	(0.00)
LRM δ(Milex/GDP)	-2.34	-1.70	-2.30
	(1.68)	(1.13)	(0.70)
T	59	59	59

third column estimates the Bayesian ADL. At the bottom of the table, we report the LRM and its standard error (calculated via the delta method) for the change in the ratio of military expenditures to the GDP, which is the key independent variable in Heo and Bohte's discussion.

The results in the first two columns are substantively consistent. In both cases, the coefficient on the change in the ratio of military expenditures to the GDP is significant, with a t-statistic of greater than 3.5. The LRM for this coefficient, however, is not statistically significant in either the seemingly unrelated regression model or the standard ADL. If the LRM is the key calculation of how changes in the independent variable accumulate in the dependent variable, the implication is that there is not much of an effect.[10]

The B-ADL model, however, has a very different conclusion. First, the standard errors

[10]Tests of the significance of the LRM for the other variables in the model all also fail to reject the null hypothesis.

are smaller in the MCMC based model than in either of the frequentist models. This is not surprising since MCMC models are often more efficient with small samples (McNeish:2016). More importantly, the confidence region of the LRM clearly excludes zero, leading to the conclusion that there is a reliable long run relationship between the change in the ratio of military expenditures to the GDP and the size of the deficit.

6.2. *Replications from Webb, Linn, and Lebo (2019)*

In this section, we replicate the models that WLL 2019 use in their article. The first of these focuses on public policy mood (Stimson, 1998) and the second addresses presidential success in Congress. In each case, the authors take standard models in the literature and use their bounds approach to reach conclusions about the long run relationships between several standard independent variables and the series of interest. They find mixed evidence with many of the tests being inconclusive about the presence of a long run relationship. Our approach has the potential to provide a clearer test without this indeterminate range of results.

Public policy mood has become the dominant measure of the American public's preferences for a more or less expansive role from the federal government. The measure itself relies on a large number of individual survey items and uses the dyadic ratio algorithm to find the common over time variance in the individual measures. Mood helps explain election outcomes and movement in government policy (Erikson, Mackuen, and Stimson, 2002). Given its prominent role in understanding American politics, it is no surprise that there has been a lot of attention to understanding its predictors. The specific replication that WLL examine is the work of Ferguson, Kellstedt, and Linn (2013). They model public policy mood as a quarterly series with an ECM predicted by inflation, unemployment, and policy outcomes.

Table 3 presents the results reported by WLL and Table 4 presents our reanalysis. The results in Table 3 are exactly the same as the ones reported by WLL. The coefficient estimates for the MCMC model are approximately the same, with minor deviations due to the sampling variably in the MCMC approach. The point estimates of the LRM are also essentially the same in the two models. The main difference, however, is in the conclusions reached based on the uncertainty in the estimate of the LRM. For WLL, there is only a conclusive result for one of the three independent variables. Their results are inconclusive about the LRR that inflation and policy outcomes have with policy mood. Our results, in contrast, provide evidence that while there are no LRRs with either of the economic predictors, there is a LRR between policy output and policy mood. In this case, the more conclusive results from the MCMC approach lead to conclusions where the bounds approach is inconclusive.

The second application in WLL is a model predicting presidential success in Congress. The dependent variable in this model is the percentage of times in a year the president was successful in the House of Representatives when he took a public position. The model includes an index of conditional party government, the president's party seat share and presidential approval as independent variables. As with the model explaining public policy mood, we are able to accurately replicate the WLL results. Based on their bound approach, the results indicate that there is a LRR between the president's party share of the House and the president's success rate. The results for the effect of conditional party government are inconclusive and the results for the effect of presidential approval suggest that there is

Table 3: Replication of WLL model of domestic policy mood

Variable	Lagged value	Differenced	LRM	LRM t-statistic
Policy Mood	-0.23			
	(0.05)			
Inflation	-0.12	-0.12	-0.52	-1.96
	(0.07)	(0.21)	(0.27)	(Between)
Unemployment	-0.08	0.91	-0.36	-0.78
	(0.10)	(0.48)	(0.47)	(Below)
Policy Outcome	-0.10	-0.17	-0.45	-2.23
	(0.05)	(0.21)	(0.20)	(Between)
Vietnam war$_{t-1}$	1.81			
	(0.65)			
Constant	19.37			
	(4.20)			
R^2	0.11			
T	168			

Note: All of the result of the LRMs are equivalently estimated via the Bewley method and the delta method. Standard errors in parenthesis. The determination of the LRMs t-statistics is based on the dynamic simulation results reported by WLL (2019).

no LRR between approval and the success rate.

The Bayesian model leads to slightly different conclusions, however. Consistent with the bounds approach we find that here is a LRR with House share and success and no LRR with approval and success. The difference is with the effect of conditional party government. Our results show that the credible region for the LRM of conditional party government excludes zero. In fact, less than 0.7% of the draws from the posterior of that LRM are negative. Our results suggest that the inference from the posterior distribution of the LRM of conditional party government indicates that there is an LRR.

7. CONCLUSION

We are not entirely confident in this approach. We think the theoretical motivation of WLL is an impressive one. If they are correct, then the hypothesis test of an LRM will directly test the LRR between the independent variable and the dependent variable, regardless of the specifics of the stationarity of the variables. As we noted above, this is an important insight that we think moves the applied testing of time series forward tremendously. The one shortcoming, we think, is the use of the Bewley transformation for this hypothesis test. We believe that the MCMC approach gives a more accurate estimate of the uncertainty of the LRM than the Bewley transformation, which relies on asymptotic theory. Our Monte Carlo evidence is suggestive of this, but nowhere near definitive. The obvious next step that we have no implemented yet is to replicate the dynamic simulations of WLL but using our

Table 4: MCMC model of domestic policy mood

Variable	Lagged value	Differenced	LRM	LRM Credible Region
Policy Mood	-0.23			
	(0.05)			
Inflation	-0.12	-0.12	-0.52	(-1.11, 0.06)
	(0.07)	(0.21)		
Unemployment	-0.08	0.93	-0.36	(-1.51, 0.54)
	(0.10)	(0.49)		
Policy Outcome	-0.10	-0.17	-0.45	(-0.93, -0.06)
	(0.05)	(0.20)		
Vietnam war$_{t-1}$	1.80			
	(0.65)			
Constant	19.35			
	(4.17)			
T	168			

Note: Results from an MCMC model with a burn in of 20,000. The coefficients report the mean of the posterior distribution with 10,000 samples. The standard deviation of the parameter estimates in parenthesis. The LRM is calculated from the distribution of the ratio of the coefficient for the differenced term divided by -1 multiplied by the coefficient on the lagged dependent variable. The credible region for the LRM reports the 95% credible region.

Table 5: Replication of WLL model of presidential success

Variable	Lagged value	Differenced	LRM	LRM t-statistic
Presidential success	-0.58			
	(0.12)			
Conditional party government	7.51	11.14	12.96	3.17
	(2.90)	(2.76)	(4.08)	(Between)
President's party house share	1.35	1.96	2.33	5.38
	(0.38)	(0.27)	(0.43)	(Beyond)
Presidential approval	0.09	0.30	0.15	0.53
	(0.17)	(0.18)	(0.29)	(Below)
Constant	-34.77			
	(18.56)			
R^2	0.61			
T	54			

Note: All of the result of the LRMs are equivalently estimated via the Bewley method and the delta method. Standard errors in parenthesis. The determination of the LRMs t-statistics is based on the dynamic simulation results reported by WLL (2019).

Table 6: MCMC model of presidential success

Variable	Lagged value	Differenced	LRM	LRM Credible Region
Presidential success	-0.58			
	(0.13)			
Conditional party government	7.51	11.17	12.96	(3.73, 22.01)
	(2.94)	(2.81)		
President's party house share	1.35	1.96	2.33	(1.38, 3.30)
	(0.39)	(0.28)		
Presidential approval	0.09	0.30	0.15	(-0.53, 0.74)
	(0.17)	(0.19)		
Constant	-34.87			
	(18.85)			
T	54			

Note: Results from an MCMC model with a burn in of 20,000. The coefficients report the mean of the posterior distribution with 10,000 samples. The standard deviation of the parameter estimates in parenthesis. The LRM is calculated from the distribution of the ratio of the coefficient for the differenced term divided by -1 multiplied by the coefficient on the lagged dependent variable. The credible region for the LRM reports the 95% credible region.

model instead. The properties of the Gibbs Sampler and MCMC models implies that we should get accurate estimates of the uncertainty in the estimate, but this is empirical and testable.

Our results demonstrate that when the coefficient on the lagged DV is moderate to small, each approach is able to recover correct parameter estimates, including for the LRM. When the coefficient on the lagged DV is large, however, B-ADL is preferable, as it is the only approach that provides theoretically appropriate estimates of the uncertainty on the LRM; ECM-ratio suffers from "explosive" SEs and ECM-Bewley underestimates the degree of uncertainty when sample sizes are small.

In addition, the results reveal that, in small samples, analysts may underestimate the LRM because the coefficient for X_{t-1} is not statistically significant in any of the models at either value of α. This suggests that applied researchers should be especially careful when selecting the lag structure of the independent variables, as the coefficients on such variables may be incorrectly identified as lacking statistical significance owing to low power. Discarding such variables, of course, would lead to inaccurate estimates of the LRM.

REFERENCES

Bewley, Ronald A. 1979. The direct estimation of the equilibrium response in a linear dynamic model. *Economics Letters* 3(4): 357-361.

De Boef, Suzanna and Luke Keele. 2008. Taking Time Seriously. *American Journal of Political Science* 52(1): 184-200.

Enes, Peter K., Nathan J. Kelly, Takaaki Masaki, and Patrick C. Wohlfarth. 2016. Don't Jettison the General Error Correction Model Just Yet: A Practical Guide to Avoiding Spurious Regression with the GECM. *Research and Politics* 3(2): 10.1177/2053168016643345.

Erikson, Robert S., Michael B. Mackuen, and James A. Stimson. 2002. *The Macropolity.* New York: Cambridge University Press.

Ferguson, Grant P., Paul M. Kellstedt, and Suzanna Linn. 2013. How Does the Economy Shape Policy Preferences? *Electoral Studies* 32(3): 544-550.

Freeman, John R. 2016. Progress in the Study of Nonstationary Political Time Series: A Comment. *Political Analysis* 24(1): 42-49.

Gelfand, Alan, Susan E. Hills, Amy Racine-Poon, and Adrian F. M. Smith. 1990. Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of the American Statistical Association* 85(412): 972-985.

Grant, Taylor and Matthew J. Lebo. 2016. Error Correction Methods with Political Time Series. *Political Analysis* 24(1): 3-30.

Heo, Uk and John Bohte. 2012. Who Pays for National Defense? Financing Defense Programs in the United States, 19472007 *Journal of Conflict Resolution* 56(3): 413-438.

Hill, Jonathan and Liang Peng. 2014. Unified Interval Estimation for Random Coefficient Autoregressive Models. *Journal of Time Series Analysis* 35(3): 282-297.

Hill, Jonathan, Deyuan Li, and Liang Peng. 2016. Uniform Interval Estimation for an AR(1) Process with AR Errors. *Statistica Sinica* 26(1): 119-136.

Keele, Luke, Suzanna Linn, and Clayton McLaughlin Webb. 2016. Treating Time with All Due Seriousness. *Political Analysis* 24(1): 31-41.

McNeish, Daniel. 2016. *On Using Bayesian Methods to Address Small Sample Problems. Structural Equation MOdeling: A Multidisciplinary Journal* 23(5): 750-773.

Stimson, James A. 1998. *Public Opinion in America: Moods, Cycles, and Swings.* 2nd. edition. Boulder, CO: Westview Press.

Webb, Clayton, Suzanna Linn, and Matthew Lebo. 2019. A Bounds Approach to Inference Using the Long Run Multiplier. *Political Analysis*