A Bounds Approach to Inference Using the Long Run Multiplier

Clayton Webb[1], **Suzanna Linn**[2] **and Matthew Lebo**[3]

[1] *Department of Political Science, University of Kansas, Lawrence, KS 66045, USA. Email: webb767@ku.edu*
[2] *Department of Political Science, Pennsylvania State University, State College, PA 16802, USA. Email: slinn@la.psu.edu*
[3] *Department of Political Science, Stony Brook University, Stony Brook, NY 11794, USA. Email: matthew.lebo@stonybrook.edu*

Abstract

Pesaran, Shin, and Smith (2001) (PSS) proposed a bounds procedure for testing for the existence of long run cointegrating relationships between a unit root dependent variable (y_t) and a set of weakly exogenous regressors $\boldsymbol{x}_t$ when the analyst does not know whether the independent variables are stationary, unit root, or mutually cointegrated processes. This procedure recognizes the analyst's uncertainty over the nature of the regressors but not the dependent variable. When the analyst is uncertain whether y_t is a stationary or unit root process, the test statistics proposed by PSS are uninformative for inference on the existence of a long run relationship (LRR) between y_t and $\boldsymbol{x}_t$. We propose the long run multiplier (LRM) test statistic as a means of testing for LRRs without knowing whether the series are stationary or unit roots. Using stochastic simulations, we demonstrate the behavior of the test statistic given uncertainty about the univariate dynamics of both y_t and $\boldsymbol{x}_t$, illustrate the bounds of the test statistic, and generate small sample and approximate asymptotic critical values for the upper and lower bounds for a range of sample sizes and model specifications. We demonstrate the utility of the bounds framework for testing for LRRs in models of public policy mood and presidential success.

Keywords: time series, unit root test, bounds test, cointegration

The analysis of time series data is often motivated by the desire to test for, and estimate, long run relationships (LRRs) between some scalar process, y_t, and a set of weakly exogenous regressors, $\boldsymbol{x}_t$. A recent exchange in *Political Analysis* highlighted some of the challenges analysts face in pursuit of this goal. Chief among these is that popular approaches assume analysts know the univariate properties of their data. Yet, if the analyst is uncertain whether their data should be classified as stationary, unit root, or fractionally integrated, the appropriate models, tests, and critical values are unclear. While this problem was raised in the exchange, the authors did not offer clear prescriptions for analysts. Our aim is to provide a means of testing for LRRs given uncertainty about univariate dynamics.

Time series modeling emphasizes the importance of pretesting. This is because the appropriate hypothesis tests and critical values depend on whether the data are $I(0)$ stationary processes or $I(1)$ unit root processes. In the case where the time series are all $I(1)$, the analyst proceeds to test for a long run cointegrating relationship between y_t and $\boldsymbol{x}_t$. The Engle–Granger two-step methodology (Engle and Granger 1987) and the single equation error correction model (ECM) (Banerjee, Dolado, and Mestre 1998; Ericsson and MacKinnon 2002) are the most common approaches. If there is evidence of cointegration, the LRR can be estimated in a levels regression and the short run dynamics, including the rate of return to equilibrium, from an ECM (Pesaran and Shin 1998). Absent evidence of cointegration, the analyst concludes no LRR exists between y_t and $\boldsymbol{x}_t$, and inference on short run dynamics proceeds from a regression in first differences.

Authors' note: The authors are grateful to John Freeman and the anonymous reviewers for their thoughtful comments. We would also like to thank Paul Johnson and Dave Armstrong for their help with computing and feedback on the simulation designs. We thank the Center for Research Methods and Data Analysis and the College of Liberal Sciences at the University of Kansas for access to their high performance compute cluster on which many of the calculations reported here were conducted. Replication materials can be found at Webb, Linn, and Lebo (2018).

***Political Analysis* (2019)**
vol. 27:281–301
DOI: 10.1017/pan.2019.3

Published
22 March 2019

Corresponding author
Clayton Webb

Edited by
Jeff Gill

In the case where the time series are all $I(0)$, inferences about, and estimation of, the LRRs proceed in the standard linear regression framework. The analyst may choose to estimate an autoregressive distributed lag model, a generalized ECM, or restricted versions of either (Hendry 1995).

Generally, however, there is uncertainty in the pretesting process. Weak tests, short time series, and ambiguous theory mean diagnosing the unseen data generating process (DGP) with certainty is often impossible and fraught with opportunities for human error. Unit root tests are notorious for having low power, particularly with samples common in applied political science (Evans and Savin 1981, 1984; Campbell and Perron 1991; Stock 1991; DeJong *et al.* 1992; Banerjee *et al.* 1993; Elliott, Rothenberg, and Stock 1996; Perron and Ng 1996; Juhl and Xiao 2003; Box-Steffensmeier *et al.* 2014; Choi 2015; Lebo and Kraft 2017). The choices analysts make about serial correlation, the presence of deterministic components in test regressions, and appropriate levels of significance influence the results of these tests. Time series with upper and lower limits (Cavaliere and Xu 2014), fractional integration (Box-Steffensmeier and Smith 1996; Lebo, Walker, and Clarke 2000), and near-integration (De Boef and Granato 1997) further complicate pretesting. Finally, the arsenal of pretests often provides inconsistent evidence for and against unit roots. Even when all pretesting indicates the data are unit root processes, misclassification is a significant risk (Perron and Ng 1996).

The way forward is less clear in these cases. Some uncertainty can be accommodated by existing methods. If one is certain that y_t is a unit root but unsure about $\boldsymbol{x}_t$, Pesaran, Shin, and Smith (2001) (PSS) offer a framework for testing hypotheses about the existence of a long run cointegrating relationship between y_t and $\boldsymbol{x}_t$. Thus, the model allows for uncertainty about the dynamic properties of the regressors: $\boldsymbol{x}_t$ may be stationary, unit roots, or mutually cointegrated (see also Pesaran and Shin 1998; Pesaran and Smith 1998).[1] The authors derived the limiting distributions for the ECM t- and Wald (F-) statistics used to test the significance of lagged levels in an ECM for the two polar cases in which (a) all regressors are stationary and (b) all regressors are unit roots. The results are presented as critical value bounds for the null hypothesis of no long run cointegrating relationship. If the computed test statistic lies above or below the bounds, inference on the null is conclusive, regardless of the underlying dynamics, but if it lies between the bounds, the test is inconclusive because reliable inference depends on knowing the true dynamics of $\boldsymbol{x}_t$.

Grant and Lebo (2016) and Philips (2018) advocate for the PSS approach and the approach has been used by political scientists (Dickinson and Lebo 2007; Enns and Wlezien 2017). Yet, the approach's reliance on the assumption that the dependent variable is a unit root makes it inflexible. Indeed, Philips (2018) provides a flowchart for analysts (p. 233) where the first question is: "Is the dependent variable stationary?" and the two branches "yes" and "no" lead researchers toward solutions. But the third branch, "I am not sure", does not exist in Philip's diagram or in the time series analyst's toolkit.

As we show below, uncertainty about the univariate dynamics of y_t renders cointegration tests based either on the significance of lagged y_t or the joint significance of lagged y_t and lagged $\boldsymbol{x}_t$ uninformative. As PSS note, the alternative hypothesis for both tests is consistent with multiple types of long run behavior, including degenerate equilibria in which y_t is stationary and independent of $\boldsymbol{x}_t$. The problem occurs because the coefficient on lagged y_t diverges from zero as y_t departs from a unit root such that the null of no cointegration will be rejected with increasing frequency, even if y_t is unrelated to $\boldsymbol{x}_t$ in the long run.

What, then, should analysts do when they are uncertain about the univariate properties of *all* of their data? We propose conducting inference based on the significance of the long run multiplier (LRM) relating each element of $\boldsymbol{x}_t$ to y_t by applying a bounds hypothesis-testing framework to assess the existence of a LRR between y_t and $\boldsymbol{x}_t$. The use of critical bounds applied to the LRM

1 In principle, fractionally integrated $\boldsymbol{x}_t$ processes are also admitted in this framework.

t-test allows a more flexible testing framework that accommodates analysts' uncertainty in the pretesting phase and, as such, applies whether y_t is $I(0)$, $I(1)$, or $I(d)$ and whether the elements of $\boldsymbol{x}_t$ are individually $I(0)$, $I(1)$, $I(d)$, or cointegrated.

We begin by identifying the model and assumptions underlying our analysis. Next, we describe the null and alternative hypotheses and test statistics underlying the PSS analysis and explicate their limitations. We show that neither the t-test nor the Wald test presented by PSS discriminate among a number of alternative long run behaviors.[2] We then propose an alternative approach that uses the LRM t-test. We generate critical value bounds for the test under uncertainty about univariate dynamics. Finally, we demonstrate the utility of this approach as a test for LRRs in models of public policy mood and presidential success.

1 The Model and Assumptions

The data generating process and assumptions underlying our analysis are the same used by PSS. Briefly, we begin with a vector autoregression (VAR) in which each variable in the system $\boldsymbol{z}_t$ is a function of its own lag(s), current and lagged values of all other variables in the system, a constant, and a trend. We assume the highest order of integration of any of the component variables is one and that the error in the model is well behaved. We then express the VAR as a vector error correction model (VECM), which isolates the LRR of interest. We assume a set of variables, $\boldsymbol{x}_t$, are weakly exogenous for the parameters in a conditional model of y_t—the variable of interest—but these variables may be $I(0)$, $I(1)$, or cointegrated. This permits hypothesis testing based on estimation of the conditional ECM. In the next section we describe the hypothesis tests recommended by PSS and show these tests fail when one is uncertain whether y_t is a unit root or stationary process.

Our DGP is a VAR of order p (VAR(p)) for $\{\boldsymbol{z}_t\}_{t=1}^{\infty}$, a $(k+1)$-vector process. Adopting the notation in PSS, we write the model using lag operator notation as follows:

$$\boldsymbol{\Phi}(L)(\boldsymbol{z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma} t) = \boldsymbol{\epsilon}_t, \tag{1}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are unknown $(k+1)$-vectors of intercept and trend coefficients and $\boldsymbol{\Phi}(L)$ is a $(k+1, k+1)$ matrix lag polynomial equal to $\boldsymbol{I}_{k+1} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i L^i$ with $\{\boldsymbol{\Phi}_i\}_{i=1}^{p}$ $(k+1, k+1)$ matrices of unknown coefficients. All variables are at most $I(1)$ (PSS Assumption 1)[3] and the vector error process $\{\boldsymbol{\epsilon}_t\}_{t=1}^{\infty}$ is $N(\boldsymbol{0}, \boldsymbol{\Omega})$, with $\boldsymbol{\Omega}$ positive definite, allowing for contemporaneous correlations in $\boldsymbol{z}_t$ (PSS Assumption 2).

We reparameterize the VAR as a VECM to isolate the long run levels relationships of interest among the variables. Setting $\boldsymbol{\Phi}(L) \equiv -\boldsymbol{\Pi} L + \boldsymbol{\Gamma}(L)(1-L)$, we can express the VAR as an equivalent VECM given by

$$\Delta \boldsymbol{z}_t = \boldsymbol{a}_0 + \boldsymbol{a}_1 t + \boldsymbol{\Pi} \boldsymbol{z}_{t-1} + \Sigma_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{z}_{t-i} + \boldsymbol{\epsilon}_t, \tag{2}$$

where $\Delta \equiv 1 - L$ is the difference operator and the matrix of LRMs is given by $\boldsymbol{\Pi} \equiv -(\boldsymbol{I}_{k+1} - \Sigma_{i=1}^{p} \boldsymbol{\Phi}_i)$.[4,5]

Equations (1) and (2) specify a system of equations such that each variable responds to all others. Often analysts are only interested in the long run behavior of a single variable,

2 PSS recognize the ambiguity in the alternative hypotheses but note that if y_t is known to be a unit root and we adopt a combination of tests, this ambiguity is of little concern. Given uncertainty over the univariate properties of y_t, we show that the ambiguity is problematic.

3 Formally, the roots of $|\boldsymbol{\Phi}(L) = \boldsymbol{I}_{k+1} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i z^i| = 0$ are either outside or on the unit circle.

4 In order to allow the deterministic components of the model to contribute to the LRR, we must restrict μ and γ to be linear combinations of the elements in the long run (cointegrating) vector. This implies that we must similarly restrict a_0 and a_1 in the VECM such that $\boldsymbol{a}_0 \equiv -\boldsymbol{\Pi}\boldsymbol{\mu} + (\boldsymbol{\Gamma} + \boldsymbol{\Pi})\boldsymbol{\gamma}$, $\boldsymbol{a}_1 \equiv -\boldsymbol{\Pi}\boldsymbol{\gamma}$.

5 The short run matrix lag polynomial is given by $\boldsymbol{\Gamma}(L) \equiv \boldsymbol{I}_{k+1} - \Sigma_{i=1}^{p-1} \boldsymbol{\Gamma}_i L^i$, where $\boldsymbol{\Gamma}_i = -\Sigma_{j=i+1}^{p} \boldsymbol{\Phi}_j$. The sum of the short run coefficient matrices in equation (2) $\boldsymbol{\Gamma} \equiv \boldsymbol{I}_m - \Sigma_{i=1}^{p-1} \boldsymbol{\Gamma}_i = -\boldsymbol{\Pi} + \Sigma_{i=1}^{p} i \boldsymbol{\Phi}_i$.

y_t, in response to a set of exogenous regressors, $\boldsymbol{x}_t$, which themselves may or may not be endogenously related. In order to estimate a single equation for y_t, we assume the elements of $\boldsymbol{x}_t$ are weakly exogenous for the parameters of a conditional model for y_t that also accounts for any contemporaneous correlations among y_t and $\boldsymbol{x}_t$. We partition $\boldsymbol{z}_t = (y_t, \boldsymbol{x}_t')'$, the error, deterministic components, and coefficient matrices conformably and restrict the k-vector of coefficients on lagged levels of y_t in the equations for each $\boldsymbol{x}_t$ to be 0, $\boldsymbol{\pi}_{xy} = 0$ (PSS Assumption 3). This eliminates the possibility of feedback from y_t to $\boldsymbol{x}_t$ and guarantees that any long run equilibrium involving y_t is unique. The marginal model for $\boldsymbol{x}_t$ is thus given by

$$\Delta \boldsymbol{x}_t = \boldsymbol{a}_{x0} + \boldsymbol{a}_{x1} t + \boldsymbol{\Pi}_{xx} \boldsymbol{x}_{t-1} + \Sigma_{i=1}^{p-1} \boldsymbol{\Gamma}_{xi} \Delta \boldsymbol{z}_{t-i} + \boldsymbol{\epsilon}_{xt}. \tag{3}$$

After conditioning on any contemporaneous correlations in the errors of y_t and $\boldsymbol{x}_t$, we can specify—and test hypotheses using—an ECM for y_t conditional on the $\boldsymbol{x}_t$.[6]

$$\Delta y_t = c_0 + c_1 t + \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \boldsymbol{x}_{t-1} + \Sigma_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta \boldsymbol{z}_{t-i} + \boldsymbol{\omega}' \Delta \boldsymbol{x}_t + u_t, \tag{5}$$

where π_{yy} is the familiar error correction rate and $\boldsymbol{\pi}_{yx.x}$ is the vector of coefficients that describe the net effect of $\boldsymbol{x}_t$ on y_t after controlling for any LRR among $\boldsymbol{x}_t$ and any contemporaneous correlations in the errors.[7] More specifically, $\boldsymbol{\pi}_{yx.x} \equiv \boldsymbol{\pi}_{yx} - \boldsymbol{\omega}' \boldsymbol{\Pi}_{xx}$ where $\boldsymbol{\pi}_{yx}$ is a vector that describes the direct effect of $\boldsymbol{x}_{t-1}$ on y_t and $\boldsymbol{\omega}$ describes the contemporaneous correlations among the variables in the system: $\boldsymbol{\omega} = \text{cov}(e_{yt}, e_{xt})/\text{var}(e_{xt})$. If the errors are uncorrelated, $\boldsymbol{\omega} = 0$ and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx}$. $\boldsymbol{\Pi}_{xx}$ specifies the LRRs among the $\boldsymbol{x}_t$.

Consistent with our uncertainty over the dynamics in $\boldsymbol{x}_t$, we wish to allow the $\boldsymbol{x}_t$ to be $I(0)$, $I(1)$ and not cointegrated, or $I(1)$ and cointegrated. $\boldsymbol{\Pi}_{xx}$, the long run coefficient matrix for $\boldsymbol{x}_t$, may thus have rank $0 \leqslant r_x \leqslant k$ (PSS Assumption 4). If $r_x = 0$, there are no cointegrating relationships and the $\boldsymbol{x}_t$ are purely $I(1)$ such that $\boldsymbol{\Pi}_{xx} = 0$ and $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx}$. If $r_x = k$ (the number of independent variables in the system), the $\boldsymbol{x}_t$ are all $I(0)$. If $0 < r_x < k$, then there are r_x cointegrating relationships in $\boldsymbol{x}_t$.

Given our assumptions, if there is a LRR between y_t and $\boldsymbol{x}_t$ it is given by:

$$LRR = \pi_{yy} y_{t-1} + \boldsymbol{\pi}_{yx.x} \boldsymbol{x}_{t-1} \tag{6}$$

$$= \pi_{yy} y_{t-1} + (\boldsymbol{\pi}_{yx} - \boldsymbol{\omega}' \boldsymbol{\Pi}_{xx}) \boldsymbol{x}_{t-1}. \tag{7}$$

Conversely, there is no LRR between y_t and $\boldsymbol{x}_t$ only if *both* $\pi_{yy} = 0$ *and* $\boldsymbol{\pi}_{yx.x} = \boldsymbol{\pi}_{yx} - \boldsymbol{\phi}' \boldsymbol{\Pi}_{xx} = \boldsymbol{0}'$ for some k-vector $\boldsymbol{\phi}$, in which case the ECM reduces to a model in first differences.[8]

6 For estimation of the conditional model to produce the same inferences as the full system, we must condition on any contemporaneous correlation in the error term $\boldsymbol{\epsilon}_t$. Following PSS, let the variance covariance matrix of the errors be given by $\boldsymbol{\Omega}$ as

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_{yy} & \boldsymbol{\omega}_{yx} \\ \boldsymbol{\omega}_{xy} & \boldsymbol{\Omega}_{xx} \end{pmatrix}.$$

We can then express ϵ_{yt} in terms of the errors for the marginal model ($\boldsymbol{\epsilon}_{xt}$) as:

$$\epsilon_{yt} = \boldsymbol{\omega}_{yx} \boldsymbol{\Omega}_{xx}^{-1} \boldsymbol{\epsilon}_{xt} + u_t, \tag{4}$$

where $u_t \sim IN(0, \omega_{uu})$, $\omega_{uu} \equiv \omega_{yy} - \boldsymbol{\omega}_{yx} \boldsymbol{\Omega}_{xx}^{-1} \boldsymbol{\omega}_{xy}$ and u is independent of $\boldsymbol{\epsilon}_{xt}$. Substitution of equation (4) into equation (2) produces the conditional model in equation (5) where $\boldsymbol{\omega} \equiv \boldsymbol{\Omega}_{xx}^{-1} \boldsymbol{\omega}_{xy}$ and $\boldsymbol{\psi}_i' \equiv \boldsymbol{\gamma}_{yi} - \boldsymbol{\omega}' \boldsymbol{\Gamma}_{xi}$, $i = 1, \ldots, p-1$.

7 The constant and trend terms are now modified to:

$$c_0 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x}) \boldsymbol{\mu} + [\boldsymbol{\gamma}_{yx} + (\pi_{yy}, \boldsymbol{\pi}_{yx.x})] \boldsymbol{\gamma}, c_1 = -(\pi_{yy}, \boldsymbol{\pi}_{yx.x}) \boldsymbol{\gamma}.$$

8 Given rank r_x, it follows that the rank of the long run coefficient matrix in the full system rank $\boldsymbol{\Pi}$ must be at least r_x and no more than $r_x + 1$. PSS further specify the conditions that must hold to ensure that the maximum order of integration in the system is one in each case. See Pesaran, Shin, and Smith (2001) for further details.

Table 1. The PSS F-test.

Specification	Long run relationship	Conclusion
$H_{0,F}: \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} = 0$	—	No equilibrating relationship, $y_t \sim I(1)$
$H_{A_1,F}: \pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$	$(\boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\boldsymbol{\Pi}_{xx})\boldsymbol{x}_{t-1}$	Nonsense equilibrating relationship, $y_t \sim I(1)$
$H_{A_2,F}: \pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} = 0$	$\pi_{yy}y_{t-1}$	Degenerate equilibrating relationship, $y_t \sim I(0)$ or trend stationary
$H_{A_{3a},F}: \pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$	$\pi_{yy}y_{t-1} + (\boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\boldsymbol{\Pi}_{xx})\boldsymbol{x}_{t-1}$	Cointegrating equilibrating relationship, $y_t \sim I(1)$
$H_{A_{3b},F}: \pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$	$\pi_{yy}y_{t-1} + (\boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\boldsymbol{\Pi}_{xx})\boldsymbol{x}_{t-1}$	Equilibrating relationship, $y_t \sim I(0)$ or trend stationary

Note: $H_{0,F}$ is based on the conditional ECM (equation (5)). Table adapted from Eviews (2017).

2 PSS Hypothesis Tests and their Limits

PSS proposed analysts test the null hypothesis of no (cointegrating) LRR between y_t and $\boldsymbol{x}_t$ assuming y_t is a unit root using a Wald (F-) test, where $H_{0F}: \pi_{yy} = \boldsymbol{\pi}_{yx.x} = 0$. The alternative hypothesis is that *either or both* are nonzero: $H_{AF}: \pi_{yy} \neq 0$ or $\boldsymbol{\pi}_{yx.x} \neq 0$ or both. Critical values for the test are unavailable for an arbitrary mix of $I(0)$ and $I(1)$ regressors. However, two polar cases establish bounds for the F-test. The lower bound is associated with $r_x = k$, in which case the elements of $\boldsymbol{x}_t$ are $I(0)$. The upper bound is associated with $r_x = 0$, in which case they are $I(1)$ and not cointegrated. Of course, the truth may lie between, in which case there is at least one cointegrating relationship among the elements of $\boldsymbol{x}_t$.

To test for a long run cointegrating relationship, the analyst estimates the conditional ECM, computes the F-statistic for the lagged level variables, and compares the result to the bounds.[9] If F is below the lower bound, we cannot reject the null regardless of whether $\boldsymbol{x}_t \sim I(0)$, $I(1)$, or is cointegrated. If F is greater than the upper bound, we can infer the existence of a LRR regardless of the dynamic properties of $\boldsymbol{x}_t$. If F is between the bounds, without knowing the dynamic properties of $\boldsymbol{x}_t$, we cannot determine whether to reject or fail to reject. If we knew the elements of $\boldsymbol{x}_t$ were $I(0)$, then we would reject the null. If we knew the elements of $\boldsymbol{x}_t$ to be $I(1)$, we would fail to reject the null.

Rejection of the null hypothesis does not, however, guarantee a *valid* long run equilibrium.[10] The alternative hypothesis is consistent with four types of LRR. PSS describe two of these as *degenerate*: the LRR is either nonsensical or of a *simpler class* in which y_t is independent of $\boldsymbol{x}_t$. Degenerate equilibria occur when we reject the null hypothesis because either $\pi_{yy} \neq 0$ or $\boldsymbol{\pi}_{yx.x} \neq 0$ but not both; a nondegenerate relationship requires both $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$. We describe each type of LRR permitted under the alternative hypothesis and present the possible relationships between y_t and $\boldsymbol{x}_t$ in Table 1.

Alternatives A_1 and A_2 describe degenerate LRRs. Under A_1, $\pi_{yy} = 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$ and the LRR given in equation (6) reduces to $(\boldsymbol{\pi}_{yx} - \boldsymbol{\omega}'\boldsymbol{\Pi}_{xx})\boldsymbol{x}_{t-1}$. In this case, y_t is a unit root process but not cointegrated with $\boldsymbol{x}_t$. The $\boldsymbol{x}_t$ are either jointly cointegrated or all individually stationary and

9 PSS show different sets of critical values apply given the deterministic relationship specified and allow for constants and trends to be omitted, unrestricted, or to apply to the LRR. Narayan (2005) derives bounds for smaller samples.
10 PSS acknowledge this possibility on p. 295.

influence Δy_t only in the short run.[11] If A_2 holds, $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} = 0$ and the LRR reduces to $\pi_{yy} y_{t-1}$. In this case, y_t is stationary and independent of $\boldsymbol{x}_t$ in the long run, regardless of the dynamic properties of $\boldsymbol{x}_t$. Changes in $\boldsymbol{x}_t$ may affect changes in y_t in the short run but y_t returns to its unconditional mean in the long run.

The remaining specifications characterize nondegenerate long run equilibria between y_t and $\boldsymbol{x}_t$ that typically motivate our hypothesis tests. Both $\pi_{yy} \neq 0$ and $\boldsymbol{\pi}_{yx.x} \neq 0$ and the LRR is given by equation (6). Under alternative A_{3a}, y_t is a unit root process and cointegrated with $\boldsymbol{x}_t$. Under alternative A_{3b}, a second type of nondegenerate equilibrium holds in which y_t is stationary and dependent on $\boldsymbol{x}_t$. The $\boldsymbol{x}_t$ may be stationary or cointegrated but in either case their influence on y_t is via a linear combination of the $\boldsymbol{x}_t$ that is stationary.[12]

PSS propose using the familiar ECM test for cointegration, $H_{0t} : \pi_{yy} = 0$, to arbitrate among a subset of the alternatives.[13] Like the F-test, critical values for this t-test depend on the nature of $\boldsymbol{x}_t$ and PSS derive bounds for this test as well. If we fail to reject the null, rejection of the F-test implies $H_{A_{1},F}$ holds and the long run equilibrium is undefined. If we reject both null hypotheses, either $H_{A_2,F}$, $H_{A_{3a},F}$ or $H_{A_{3b},F}$ holds. We can only rule out A_2 and A_{3b} when we are certain y_t is $I(1)$ (as PSS assume). This presents a dilemma for the analyst who is uncertain of the dynamics of y_t.

Researchers frequently use the ECM to test for the existence of LRRs, primarily relying on the test of the null $H_0 : \pi_{yy} = 0$ and using either MacKinnon critical values (Banerjee *et al.* 1993; Ericsson and MacKinnon 2002; Lebo and Grant 2016) or the PSS critical values for inference (Dickinson and Lebo 2007; Philips 2018). However, a researcher may make an incorrect judgment and y_t may be truly stationary. If so, this presents problems for valid inference and one must appreciate the different types of long run behavior that may lead to rejection of the null. The above discussion makes clear that if both theory and univariate tests are inconclusive as to whether y_t is $I(0)$ or $I(1)$, it is a dangerous strategy to conclude a LRR exists between y_t and $\boldsymbol{x}_t$ based on either the ECM test or the F-test proposed by PSS.

3 The LRM Test

What should analysts do if they are uncertain whether y_t is a stationary or unit root process and wish to draw inference about the existence of a valid LRR? We propose a test for the existence of a valid LRR between y_t and $\boldsymbol{x}_t$ based on the LRM. To understand the appeal of the LRM, it is helpful to express the conditional ECM in equation (5) to isolate the LRR:

$$\Delta y_t = c_0 + c_1 t + \pi_{yy}\left(y_{t-1} + \frac{\boldsymbol{\pi}_{yx.x}}{\pi_{yy}}\boldsymbol{x}_{t-1}\right) + \Sigma_{i=1}^{p-1}\boldsymbol{\psi}_i'\Delta\boldsymbol{z}_{t-i} + \boldsymbol{\omega}'\Delta\boldsymbol{x}_t + \mu_t, \tag{8}$$

where $(y_{t-1} + (\boldsymbol{\pi}_{yx.x}/\pi_{yy})\boldsymbol{x}_{t-1})$ gives the long run, and possibly cointegrating, relationship, π_{yy} gives the rate of return to equilibrium, and $\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ is the LRM.

As we discussed above, a valid LRR requires *both* $\pi_{yy} \neq$ *and* $\boldsymbol{\pi}_{yx.x} \neq 0$. This implies $\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ must also be nonzero. In this case $\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ describes the links between $\boldsymbol{x}_t$ and y_t and π_{yy} tells us how this linkage drives change in y_t. In contrast, if $\pi_{yy} = 0$ the equilibrium term drops out of the equation and $\boldsymbol{\pi}_{yx.x}/\pi_{yy}$ is undefined. If $\boldsymbol{\pi}_{yx.x} = 0$, the LRMs are zero and y_t is not a function of $\boldsymbol{x}_t$. *Thus, a nondegenerate, or valid, equilibrium relationship between y_t and $\boldsymbol{x}_t$ requires the LRM*

11 If either $\boldsymbol{\omega} = 0$ (y_t and $\boldsymbol{x}_t$ are not contemporaneously correlated) or $\boldsymbol{x}_t$ are unit roots but not cointegrated ($r_x = 0$ such that $\boldsymbol{\Pi}_{xx} = 0$), the LRR is given by the coefficients that describe the relationships among the $\boldsymbol{x}_{t-1}$ and y_t, $\boldsymbol{\pi}_{yx}$.

12 The distinction between cases 3a and 3b depends on the rank of the rank factorized matrix of LRRs. Specifically, we can write $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ give the speed of adjustment to disequilibrium—the error correction coefficient—and $\boldsymbol{\beta}$ comprise the vector of LRRs. If rank$(\boldsymbol{\beta}_{yx}, \boldsymbol{\beta}_{xx}) = r_x$, the vector involving y_t duplicates the LRRs in $\boldsymbol{x}_t$ and all series are stationary. If rank$(\boldsymbol{\beta}_{yx}, \boldsymbol{\beta}_{xx}) = r_x + 1$ then the vector contains independent information and all series are unit roots and y_t is cointegrated with $\boldsymbol{x}_t$. See Eviews (2017) for further details.

13 Banerjee, Dolado, and Mestre (1998) first developed the test but only considered the case where $\boldsymbol{x}_t$ were all unit roots.

Table 2. The LRM test.

	Specification	Conclusion	Integration order (y_t)
$H_{0,LRM}$	$\frac{\boldsymbol{\pi}_{yx.x}}{\pi_{yy}} = 0$	No equilibrating relationship	$y_t \sim I(1)$ or $I(0)$
$H_{A_1,LRM}$	$\frac{\boldsymbol{\pi}_{yx.x}}{\pi_{yy}} \neq 0$	Nondegenerate cointegrating equilibrating relationship	$y_t \sim I(1)$
$H_{A_2,LRM}$	$\frac{\boldsymbol{\pi}_{yx.x}}{\pi_{yy}} \neq 0$	Nondegenerate equilibrating relationship	$y_t \sim I(0)$

Note: $H_{0,LRM}$ is based on the Bewley instrumental variables regression given in equation (9).

to be nonzero. It follows immediately that inference on the existence of a valid LRR between y_t and $\boldsymbol{x}_t$ can be made based on the hypothesis test $H_{0,LRM} : \boldsymbol{\pi}_{yx.x}/\pi_{yy} = 0$.[14]

This is true whether y_t is $I(0)$ or $I(1)$. If y_t is a unit root process, the only way π_{yy} can be nonzero is if y_t is linked to $\boldsymbol{x}_t$ in the long run. In other words, it must be the case that $\boldsymbol{\pi}_{yx.x} \neq 0$ and $\boldsymbol{\pi}_{yx.x}/\pi_{yy} \neq 0$ such that y_t has a long run, cointegrating relationship with $\boldsymbol{x}_t$. This is the logic underlying the ECM test for cointegration (Banerjee *et al.* 1993; Banerjee, Dolado, and Mestre 1998; Ericsson and MacKinnon 2002). In the stationary case, π_{yy} is, by definition, nonzero: y_t will always return to its mean in the long run, whether that mean is conditional on $\boldsymbol{x}_t$ or not. Only if $\boldsymbol{\pi}_{yx.x}$, and thus $\boldsymbol{\pi}_{yx.x}/\pi_{yy} \neq 0$, will the long run value of y_t be conditional on $\boldsymbol{x}_t$.[15]

We present the null and alternative hypotheses for the LRM test in Table 2. If we cannot reject the null, then y_t does not have a valid equilibrium with $\boldsymbol{x}_t$, regardless of whether the data are $I(1)$ or $I(0)$ and whether the regressors are cointegrated. If we can reject the null, we infer a LRR between y_t and $\boldsymbol{x}_t$. We cannot, however, distinguish a long run cointegrating relationship, $H_{A_1,LRM}$, from a LRR between a set of stationary variables, $H_{A_2,LRM}$, using this test. Our contention is that this uncertainty is an inevitable—and appropriate—consequence of uncertainty in the pretesting phase. Perhaps analysts can appeal to theory to overcome this bind, but there are few persuasive theoretical arguments that make this effort in political science (for an example see Erikson, MacKuen, and Stimson 2002).

The LRMs are not estimated directly in the ECM or the equivalent ADL. While we can calculate the LRMs from these models, the standard error is more problematic as there is no simple formula calculating the standard error of a ratio of coefficients. Various methods exist to approximate the variance of a quotient of items with known variances. One option is to apply the delta method. Alternatively, one can use instrumental variables to estimate the Bewley transformation of the model, which estimates the LRM and its standard error directly (Bewley 1979).[16] The Bewley transformation for the general case with a constant and trend is written as:

$$y_t = \phi_0 + \tau t - \phi_1 \Delta y_t + \psi_0 \boldsymbol{x}_t - \psi_1 \Delta \boldsymbol{x}_t + \mu_t, \tag{9}$$

where $\psi_0 = -(\pi_{yx.x}/\pi_{yy}) = LRM$, $\phi_0 = -(c_0/\pi_{yy})$, $\tau = c_1/\pi_{yy}$, $\phi_1 = -(\pi_{yy} + 1/\pi_{yy})$, $\psi_1 = \pi_{yx}$, and $\mu = -(e/\pi_{yy})$ in the conditional ECM. A constant, trend, $\boldsymbol{x}_t$, $\boldsymbol{x}_{t-1}$, and y_{t-1} are used as instruments to estimate the model (Banerjee *et al.* 1993; De Boef and Keele 2008).

14 It is also possible that the LRM is undefined under the null.
15 This is, of course, a common occurrence. But as we will see below, the appropriate critical values on the hypothesis test on the LRM will often be nonstandard, a fact that has eluded most applied research.
16 Banerjee *et al.* (1993) prove that estimates of the standard error based on approximations such as the delta method applied to either the ADL or ECM are equivalent to that obtained directly from the Bewley transformation using instrumental variables.

Because the LRM is a ratio of coefficients and the coefficients are a function of time series with potentially varying dynamic behavior, the form of the distribution of the LRM t-test is not obvious. If y_t is a unit root, it is likely to have a nonstandard distribution, like the t and F-tests evaluated by PSS. It is also unclear how the sample size, presence or absence of deterministic components, or number of regressors will affect the distribution.[17] In the next section we calculate the appropriate critical values for the LRM test, allowing each of these features of the data and model to vary, and we offer a bounds testing framework for inference.

Before doing so, we summarize three advantages of focusing on the significance of the LRM as a test for a valid long run equilibrium. First, inferences do not depend on whether any given time series is $I(0)$ or $I(1)$. Second, the LRM test has a specific advantage over the ECM test for cointegration when we know y_t is $I(1)$ and we have multiple independent variables in the model. While rejecting the null $H_0 : \pi_{yy} = 0$ implies y_t is cointegrated with a vector of $\boldsymbol{x}_t$, it does not tell us which element(s) of $\boldsymbol{x}_t$ contribute to the cointegrating relationship. The LRM test allows us to draw inferences about whether there is a LRR between y_t and any element of $\boldsymbol{x}_t$. Third, rejection of the null implies a nondegenerate LRR between y_t and $\boldsymbol{x}_t$.

4 The Distribution of the LRM t-Test

We compute the distribution of the LRM test by estimating the sampling distribution of the LRM test statistic in the Bewley IV regression given in equation (9) under the true null hypothesis that there is no LRR between y_t and $\boldsymbol{x}_t$ ($\boldsymbol{\pi}_{yx.x}/\pi_{yy} = 0$) under a range of conditions.

4.1 The Importance of Autocorrelation and Existence of Bounds

Our first set of stochastic simulations demonstrates the sensitivity of the test's behavior to the strength of autocorrelation in the data and to sample size. We generate critical values for the LRM t-test for varying degrees of autocorrelation in y_t and a single x_t for sample sizes of 75 and 1000. The smaller sample size is common in applied work while the larger sample size produces critical values that approximate the asymptotic distribution. We generate two independent autoregressive processes, $y_t = \rho_y y_{t-1} + e_{yt}$ and $x_t = \rho_x x_{t-1} + e_{xt}$ with the errors drawn from independent standard normal distributions. We vary the values of ρ_y and ρ_x from 0 to 0.90 in increments of 0.10 and from 0.90 to 1.0 in increments of 0.01. For each combination of ρ_y and ρ_x, we simulate the sampling distribution of the LRM t-statistic using 50,000 replications. The LRM t-value is estimated as the t-value on x_t in the Bewley ECM, equation (9), in a model with an unrestricted constant and no trend. Figure 1 presents the simulated critical values associated with the 97.5 percentile of the distribution for values of ρ_x as ρ_y varies for both sample sizes.[18]

We draw three conclusions from our results. First, critical values are approximately standard normal for both sample sizes when y_t is white noise, regardless of the dynamics in x_t. However, as y_t becomes more autoregressive, the appropriate critical values fan out based on the degree of autocorrelation in x_t: for smaller ρ_x they are closer to zero and for larger ρ_x they are farther from zero than standard normal critical values.[19] Second, patterns are the same in both panels of the figure, although the range of autocorrelation for which critical values depart from standard normal is smaller for $T = 1000$. In fact, the range of autocorrelation in y_t for which standard critical values are appropriate can be quite small when sample sizes take on values typical in applied work, suggesting the possibility that our confidence in some published findings regarding the

17 Pesaran and Shin (1998) show that the distribution of the estimate of the LRM itself is mixture normal asymptotically and that it is super consistent.

18 The data used to produce these figures can be found in the Supplementary Appendix, Section 1. Replication materials are available at Webb, Linn, and Lebo (2018).

19 The largest value of the t-statistics occurs when $\rho_y = \rho_x$ and both are large. This is when the series are most similar and, therefore, most vulnerable to the spurious regression problem. But, at their peak, these values are smaller than the cases when both series are unit roots. The similarities of the ρ_y and ρ_x create the peak patterns observed in the plots.

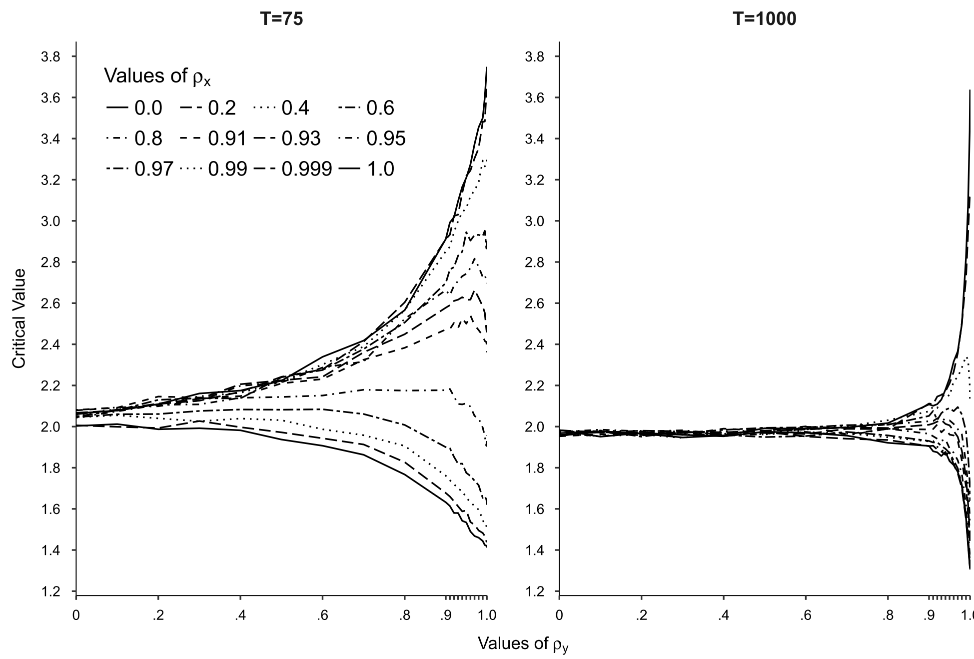

Figure 1. Simulated critical values for the LRM t-test (95th percentile). *Note:* Critical values are computed via stochastic simulations using 50,000 replications for the LRM t-statistic in the Bewley instrumental variables regression in equation (9). The time series y_t and x_t are generated from: $y_t = \rho_y y_{t-1} + e_{yt}$ and $x_t = \rho_x x_{t-1} + e_{xt}$ where the errors are drawn from independent standard normal distributions.

significance of the LRM may be overstated. Third, the results establish lower and upper bounds that are similar regardless of sample size, about 1.30 and 3.65, respectively. If the LRM t-statistic is greater than about 3.65, we can infer a LRR between y_t and x_t. If it is less than about 1.30 we can infer the absence of a LRR. For t-statistics within the bounds, we need to know sample size and degree of autocorrelation in each series to draw an inference. Next, we assess how these bounds behave as more independent variables are added to the model, as the dynamic behaviors of y_t and $\boldsymbol{x}_t$ vary, and as the specification of deterministic components varies.

4.2 The Conditions That Set The Bounds

Critical values could be calculated for the LRM t-statistic if we knew the dynamic properties of the data, but this information is not available in applied settings. We can, however, establish the lowest and highest critical values associated with the LRM t-statistic under a number of conditions. Table 3 shows quantiles for the empirical distributions of the ECM t-tests estimated from a model with three independent variables ($k = 3$) and a constant. The rows of the table give the possible permutations of $I(0)$ and $I(1)$ variables ranging from the case where all the variables are independent white noise processes, to the case where all the variables are independent unit roots. We show the 2.5th and 97.5th percentiles for each LRM t-test for each DGP. We simulated the sampling distributions using 100,000 replications of $T = 1000$.

The first column shows the percentiles for the t-statistic on π_{yy}, the error correction rate. The percentiles for the t-statistics in the top half provide empirical estimates of the t-test on π_{yy} when $\rho_y = 0$. The null hypothesis is false in these cases: $\pi_{yy} \neq 0$. As such, the magnitudes of the observed t-statistics are large. The simulated critical values reported in the bottom half of the table correspond to those reported by PSS for the ECM t-test under the true null $\pi_{yy} = 0$ when $\rho_y = 1$. Our bounds (−3.12 and −4.01) are approximately equal to the bounds reported by PSS (−3.13 and −4.05).[20] The lower bound for the t-statistic on π_{yy} is set by the case where all the independent

20 See Pesaran, Shin, and Smith (2001) Table CII(iii) Case III page 303.

Table 3. The empirical distribution of the ECM t-test and simulated critical values for the LRM t-test: identifying the bounds conditions.

	y_{t-1}		$x_{1,t-1}$		$x_{2,t-1}$		$x_{3,t-1}$	
	2.5%	**97.5%**	**2.5%**	**97.5%**	**2.5%**	**97.5%**	**2.5%**	**97.5%**
$\rho_y = 0$								
$\rho_{x1} = 0 \mid \rho_{x2} = 0 \mid \rho_{x3} = 0$	−33.54	−29.61	−1.96	1.97	−1.96	1.97	−1.97	1.97
$\rho_{x1} = 1 \mid \rho_{x2} = 0 \mid \rho_{x3} = 0$	−33.57	−29.65	−1.98	1.98	−1.96	1.97	−1.96	1.96
$\rho_{x1} = 0 \mid \rho_{x2} = 1 \mid \rho_{x3} = 0$	−33.59	−29.63	−1.97	1.96	−1.97	1.97	−1.96	1.97
$\rho_{x1} = 0 \mid \rho_{x2} = 0 \mid \rho_{x3} = 1$	−33.57	−29.63	−1.97	1.96	−1.96	1.98	−1.96	1.98
$\rho_{x1} = 1 \mid \rho_{x2} = 1 \mid \rho_{x3} = 0$	−33.59	−29.68	−1.99	1.98	−1.96	1.97	−1.97	1.97
$\rho_{x1} = 1 \mid \rho_{x2} = 0 \mid \rho_{x3} = 1$	−33.60	−29.67	−1.97	1.98	−1.97	1.96	−1.98	1.97
$\rho_{x1} = 0 \mid \rho_{x2} = 1 \mid \rho_{x3} = 1$	−33.61	−29.67	−1.97	1.97	−1.97	1.98	−1.97	1.98
$\rho_{x1} = 1 \mid \rho_{x2} = 1 \mid \rho_{x3} = 1$	−33.63	−29.69	−1.96	1.97	−1.97	1.97	−1.97	1.96
$\rho_y = 1$								
$\rho_{x1} = 0 \mid \rho_{x2} = 0 \mid \rho_{x3} = 0$	−3.13	0.24	−1.30	1.30	−1.29	1.29	−1.30	1.29
$\rho_{x1} = 1 \mid \rho_{x2} = 0 \mid \rho_{x3} = 0$	−3.50	0.08	−3.65	3.62	−1.37	1.37	−1.38	1.38
$\rho_{x1} = 0 \mid \rho_{x2} = 1 \mid \rho_{x3} = 0$	−3.49	0.07	−1.38	1.37	−3.63	3.65	−1.38	1.37
$\rho_{x1} = 0 \mid \rho_{x2} = 0 \mid \rho_{x3} = 1$	−3.49	0.07	−1.38	1.38	−1.37	1.37	−3.64	3.67
$\rho_{x1} = 1 \mid \rho_{x2} = 1 \mid \rho_{x3} = 0$	−3.78	−0.13	−3.42	3.46	−3.44	3.42	−1.44	1.43
$\rho_{x1} = 1 \mid \rho_{x2} = 0 \mid \rho_{x3} = 1$	−3.77	−0.13	−3.44	3.43	−1.43	1.43	−3.44	3.38
$\rho_{x1} = 0 \mid \rho_{x2} = 1 \mid \rho_{x3} = 1$	−3.77	−0.13	−1.44	1.43	−3.41	3.43	−3.40	3.46
$\rho_{x1} = 1 \mid \rho_{x2} = 1 \mid \rho_{x3} = 1$	−4.04	−0.33	−3.27	3.32	−3.27	3.29	−3.28	3.27

Note: Critical values are computed via stochastic simulations using 100,000 replications of $T = 1000$ for the LRM t-statistic in the Bewley instrumental variables regression in equation (9). A constant $\boldsymbol{x}_t$, $\boldsymbol{x}_{t-1}$, and y_{t-1} are used as instruments. The time series y_t and x_t are generated from: $y_t = \rho_y y_{t-1} + e_{yt}$ and $x_{i,t} = \rho_{x_i} x_{i,t-1} + e_{x_i,t}$ for $i = 1, 2, 3$, where the errors are drawn from independent standard normal distributions.

variables are white noise while the upper bound is set by the case where all the regressors are unit roots.

The next three columns show the behavior of the LRM t-statistic. We make four observations. First, the empirical distributions of the LRM t-statistics are (roughly) symmetric. Second, as in Figure 1, when $\rho_y = 0$ the critical values correspond to the standard t-distribution, regardless of the dynamics in $\boldsymbol{x}_t$. Third, the critical values are nonstandard for cases where $\rho_y = 1$. Finally, the shapes of these nonstandard distributions change with the number of $I(1)$ variables in the model. We can use these results to determine the values of the LRM t-statistic that establish a lower bound below which we fail to reject the null and an upper bound beyond which we can reject the null. These bounds can be applied in the absence of knowledge of the autoregressive properties of both $\boldsymbol{x}_t$ and y_t.

The simulated bounds for the LRM t-statistics are set under different conditions than those for the t- and F-tests reported by PSS. The lower bound for the LRM test is similarly set by the case where all the independent variables are white noise and y_t is $I(1)$, but the upper bound is set by the case where y_t is $I(1)$ and *exactly one* independent variable is $I(1)$. For example, in the case where $\rho_{x1} = \rho_{x2} = 0$ and $\rho_{x3} = \rho_y = 1$, the t-statistics for x_1 and x_2 are (roughly) equal at 1.38. The t-statistic for x_3 is much higher (3.70). This same pattern exists regardless of which element of $\boldsymbol{x}_t$ is the unit root process. Critical values for all other t-statistics, including the standard t-statistics in the top half of the table, fall between these bounds.

Table 4. Upper and lower bounds for the LRM t-test by k and T.

	$T = 75$		$T = 150$		$T = 1000$	
k	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
1	1.40	3.69	1.35	3.69	1.29	3.65
2	1.40	3.68	1.34	3.63	1.30	3.60
3	1.40	3.62	1.35	3.63	1.30	3.65
4	1.40	3.61	1.34	3.59	1.30	3.61

Note: Critical values are computed via stochastic simulations using 100,000 replications for the LRM t-statistic in the Bewley instrumental variables regression in equation (9). A constant x_t, x_{t-1}, and y_{t-1} are used as instruments. The time series y_t and x_t are generated from: $y_t = \rho_y y_{t-1} + e_{yt}$ and $x_{i,t} = \rho_{x_i} x_{i,t-1} + e_{x_i,t}$ for $i = 1, 2, 3, 4$, where the errors are drawn from independent standard normal distributions.

Why is the upper bound for the LRM test different from the test statistics considered by PSS? Recall that the LRM is a ratio. The variance of a ratio may be approximated by

$$\frac{a}{b} = \frac{1}{b^2}\mathrm{Var}(a) + \frac{a^2}{b^2}\mathrm{Var}(b) - 2\frac{a}{b^3}\mathrm{Cov}(a, b). \tag{10}$$

The first quantity in equation (10), $(1/b^2)\mathrm{Var}(a)$, shows that the variance of a ratio increases with the variance of the numerator, here the coefficient on x_t. Spurious correlations among multiple $I(1)$ independent variables increase the variance of the associated LRM in the same way correlation among the independent variables in a regression model increases the variances of the estimated parameters for those variables. This spurious correlation in $\boldsymbol{x}_t$ does not affect the ECM t-statistic, which is not a function of the variance of a or the covariance of a and b in equation (10). Thus, the largest LRM t-statistic will occur when there is exactly one unit root in $\boldsymbol{x}_t$.[21]

Table 4 shows how the bounds for the LRM t-statistic change with T and k. The number of independent variables increases from 1 to 4 along the vertical dimension of the table. The sample size increases from 75 to 150 to 1000 along the horizontal dimension. We simulated the sampling distributions of the t-statistics using 100,000 replications of each sample size.

The bounds are similar across conditions in Table 4, but with some notable patterns. The critical value of the LRM t-statistic associated with the lower bound is essentially the same as k increases from 1 to 4, but it slowly declines as the sample size increases from 75 to 1000. The small values of the lower bound reflect the low probability of finding a relationship between a white noise variable and a random walk, a probability that declines slightly with T. The upper bounds behave differently. No clear patterns emerge as the sample size changes, but the upper bounds decline gradually as the number of independent variables increases from 1 to 4.[22] The upper bounds for $T = 75$, $T = 150$, and $T = 1000$ are 3.69, 3.69, and 3.65 when $k = 1$. These values fall to 3.61, 3.59, and 3.61 when $k = 4$.

These results suggest an inferential strategy in which the analyst estimates any completely specified dynamic regression, calculates the LRM, and obtains the t-statistics for each using either the delta method or the Bewley IV regression. Next, the results are compared to the bounds. If the LRM t-statistic falls below the upper bound, the null hypothesis cannot be rejected regardless of whether y_t is $I(0)$ or $I(1)$ and regardless of the dynamic behavior of the regressors. If the LRM t-statistic is above the upper bound, the analyst can reject the null hypothesis and infer a LRR,

21 If one or more independent variables are cointegrated, the linear combination of the cointegrated variables is stationary. This is equivalent to the case where y is $I(1)$ and $\boldsymbol{x}_t$ is stationary.

22 While the upper bounds are not identical across the three sample sizes, given the similarity in the simulated bounds, the absence of any pattern in their behavior, and the random variation inherent in stochastic simulations, we consider the upper bounds to be "the same". Smaller sample sizes may induce different behavior.

Table 5. Upper and lower bounds for the LRM t-test by k and T and deterministic components.

	$T=75$		$T=150$		$T=1000$		$T=75$		$T=150$		$T=1000$	
k	**LB**	**UB**	**LB**	**UB**	**LB**	**UB**	**LB**	**UB**	**LB**	**UB**	**LB**	**UB**
	Model: $c_0 \neq 0.0$ $c_1 = 0.0$						Model: $c_0 \neq 0.0$ $c_1 \neq 0.0$					
	DGP: $c_0 = 0.0$ $c_1 = 0.0$						DGP: $c_0 = 0.0$ $c_1 \neq 0.0$					
1	1.40	3.69	1.35	3.69	1.29	3.65	1.05	1.94	1.01	1.89	0.99	1.85
2	1.40	3.68	1.34	3.63	1.30	3.60	1.05	1.99	1.01	1.90	0.99	1.88
3	1.40	3.62	1.35	3.63	1.30	3.65	1.05	1.93	1.01	1.88	0.99	1.86
4	1.40	3.61	1.34	3.59	1.30	3.61	1.05	1.91	1.01	1.89	0.99	1.86
	Model: $c_0 \neq 0.0$ $c_1 = 0.0$						Model: $c_0 \neq 0.0$ $c_1 \neq 0.0$					
	DGP: $c_0 \neq 0.0$ $c_1 = 0.0$						DGP: $c_0 \neq 0.0$ $c_1 \neq 0.0$					
1	1.06	3.25	1.01	3.04	0.99	2.89	1.06	1.95	1.02	1.90	0.98	1.86
2	1.06	3.26	1.02	3.04	0.98	2.88	1.05	1.92	1.01	1.91	0.99	1.86
3	1.06	3.24	1.01	3.07	0.98	2.87	1.05	1.97	1.01	1.90	0.98	1.87
4	1.07	3.29	1.01	3.09	0.99	2.90	1.06	1.92	1.02	1.89	0.98	1.87

Note: Critical values are computed via stochastic simulations using 100,000 replications for the LRM t-statistic in the Bewley instrumental variables regression in equation (9). A constant x_t, x_{t-1}, and y_{t-1} are used as instruments. The time series y_t and x_t are generated from: $y_t = c_0 + c_1 t + \rho_y y_{t-1} + e_{yt}$ and $x_{i,t} = \rho_{x_i} x_{i,t-1} + e_{x_i,t}$ for $i = 1, 2, 3, 4$, where the errors are drawn from independent standard normal distributions. c_0 denotes the constant and c_1 the trend. The constant in the DGP (c_0) took values of 0 and 1.

again regardless of the dynamics of y_t or $\boldsymbol{x}_t$. However, if the estimated statistic falls between the bounds, no conclusion can be drawn absent full knowledge of the dynamic properties of all the variables in the model.

4.3 The Effects of Deterministic Features of y on the Bounds

Changes in the deterministic components in the DGP are more consequential. The bounds presented in Table 4 were derived assuming a DGP for y_t that contained neither a constant nor a trend ($c_0 = c_1 = 0$ in equation (5)).[23] The results presented in this section show that the critical values associated with a 95% confidence interval for the LRM t-test shift toward zero if the DGP for y_t includes a nonzero constant, trend, or both, *as long as the estimated model encompasses the DGP*. Omitting a constant or trend from the model when it is a feature of the DGP produces misspecification bias and renders any tests of the null hypothesis ($H_0 : LRM = 0$) meaningless. Further, excluding a constant, even when it is zero, produces biased estimates of the effects of the $\boldsymbol{x}_t$ on y_t (Greene 2017).

Table 5 presents bounds for the four cases in which the DGP includes (a) neither a constant nor a trend ($c_0 = c_1 = 0$, top left); (b) a constant but no trend ($c_0 = 1$, $c_1 = 0$, bottom left); (c) a trend but no constant ($c_0 = 0$, $c_1 = 1$, top right); and (d) both a constant and trend ($c_0 = c_1 = 1$, bottom right). The regression model used to estimate the values in the left column includes a constant. Those in the right column add a trend. For each DGP and regression model, we present the bounds for $k = \{1, 2, 3, 4\}$ independent variables (along the vertical dimension) for sample sizes $T = \{75, 150, 1000\}$ (along the horizontal dimension). We present values for the 97.5th percentiles of the sampling distributions of the t-statistics for the LRMs. The distributions are symmetric. As above, the lower bound for each model is set by the case where y_t is $I(1)$ and each of the independent variables is white noise while the upper bound is set by the case where y_t and exactly one independent variable are $I(1)$.

23 The bounds presented in Table 4 also assume that the DGP for $\mathbf{x}_t$ does not include a constant or a trend. Along with the assumptions about the DGP for y_t, these assumptions imply $\boldsymbol{a}_0 = \boldsymbol{a}_1 = 0$ in equation (2).

The bounds presented in the upper left of Table 5 correspond to those presented in Table 4 and have been discussed in detail.[24] In all other cases, the bounds shift closer to zero, but maintain a similar width. The differences can be understood in terms of the spurious regression problem (Granger and Newbold 1974) where two uncorrelated unit roots appear related more frequently than chance would predict. The addition of a constant or a trend in the DGP for y_t, e.g., $c_0 \neq 0$ and/or $c_1 \neq 0$, causes the trajectory of y_t to diverge from the trajectory of x_t, which contains neither a constant nor a trend. As a consequence, the dynamics of y_t and x_t will be more distinct than in the classic case, pulling the distribution of the t-statistic, and the value of the bounds, toward zero. As a result, the upper bounds presented in Table 4 and the top left panel of Table 5, when neither y_t nor $\boldsymbol{x}_t$ contain a constant or trend, give the limiting case. The results also suggest that the lower bound hits a floor. Once any deterministic component enters the DGP, the behaviors of y_t and $\boldsymbol{x}_t$ are so different that y_t will very seldom appear to be related to mean reverting x_t variables.[25]

The results presented in Tables 4 and 5 illustrate how the critical values associated with the bounds for the LRM t-statistic change based on the dynamics of y_t. One might conclude that the changing critical values complicate our proposed hypothesis-testing procedure, that we are giving up a complicated set of preestimation procedures in exchange for a set of complicated postestimation procedures. In the next section we show the results above simplify hypothesis testing, allowing analysts to avoid making tenuous assumptions about the DGP for y_t.

5 Inference Using the LRM: A Bounds Approach

The critical values for the LRM t-statistic change based on the dynamics of y_t and $\boldsymbol{x}_t$. If y_t is white noise, standard critical values apply. If y_t is a unit root, critical values for the LRM t-statistic change based on the number of $I(1)$ independent variables. The critical values are closer to zero if all the independent variables are white noise, farther from zero when multiple independent variables are $I(1)$, and even farther from zero when exactly one the independent variables is $I(1)$. The critical values change further based on the deterministic features of the DGP. The analyst must know all of this information to select the correct set of critical values. This would seem to leave the analyst at an impasse. None of this information can be known. The solution to this problem is to accept uncertainty associated with the features of $\boldsymbol{x}_t$ and y_t and use a hypothesis-testing procedure that acknowledges this uncertainty.

We propose a general bounds testing procedure to accommodate dynamic uncertainty. The critical values furthest from 0 in Table 5 occur in the cases where neither $\boldsymbol{x}_t$ nor y_t contain a constant or a trend ($a_0 = a_1 = 0$). These are the cases where the series are most similar and that set the upper bounds for the procedure. The critical values closest to zero in Table 5 occur in the cases where y_t contains a trend ($c_1 \neq 0$), constant ($c_0 \neq 0$), or trend and constant ($c_0 \neq 0$ and $c_1 \neq 0$). These cases set the lower bounds for the procedure. In practical terms, the change in the lower bound is unimportant for the analyst. Whether the test statistic falls within the bounds or below the lower bound, one fails to reject the null hypothesis. The only case where one can reject the null hypothesis while accounting for the uncertainty inherent in the classification of time series, is the case where the calculated test statistic falls beyond the upper bound. Thus, the upper limit of the bounds is set by the DGP with no trend and no constant and the lower bound is set by any of the DGPs that include deterministics. These combined bounds are presented in Table 6.

24 The DGP excludes a constant, but a constant is included in the estimated regression. This is the same DGP used by PSS to derive the bounds for the ECM t-test and the F-test.

25 The small critical values on the right side of Table 5 are atypical but simply reflect the fact that it is very rare to find evidence for a LRR when y_t is drifting or trending and the elements of $\boldsymbol{x}_t$ are not. Given the difficulty of distinguishing drifting and trending time series, we contend that it would be difficult to justify relying on the bounds that assume a trend in the DGP for y_t.

Table 6. Bounds given uncertainty about deterministic features of the DGP.

	T = 75		T = 150		T = 1000	
k	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
1	1.05	3.69	1.01	3.69	0.98	3.65
2	1.05	3.68	1.01	3.63	0.98	3.60
3	1.05	3.62	1.01	3.63	0.98	3.65
4	1.05	3.61	1.01	3.59	0.98	3.61

Note: Critical values are computed via stochastic simulations using 100,000 replications for the LRM t-statistic in the Bewley instrumental variables regression in equation (9). A constant x_t, x_{t-1}, and y_{t-1} are used as instruments. The time series y_t and x_t are generated from: $y_t = \rho_y y_{t-1} + e_{yt}$ and $x_{i,t} = \rho_{x_i} x_{i,t-1} + e_{x_i,t}$ for $i = 1, 2, 3, 4$, where the errors are drawn from independent standard normal distributions.

The bounds presented in Table 6 facilitate every type of analytical uncertainty that typically vex time series analysts. One does not need to know whether a series is a stationary or unit root process. One does not need to know whether a series is characterized as a random walk, a random walk with drift, or a random walk with trend and drift. These bounds allow analysts to focus on the theoretical questions at the heart of political analysis, the existence of LRRs.

Applying the bounds is simple. The analyst must make a decision about whether to include a trend in the regression model. This can occur as part of the typical general-to-specific modeling procedure that should govern dynamic specification (Hendry 1995). The analyst includes a trend in the first model. If the trend is not significant, the trend can be removed. If the analyst is uncertain, the trend can be left in the model. Including the trend in the regression model when the trend is not part of the DGP does not affect the bounds. The second step is the estimation of the Bewley regression or application of the delta method to the estimated model. The t-statistics for the estimated LRMs can be compared to the bounds presented in Table 6. This allows for inferences about the LRRs between the outcome and the independent variables that recognize the analyst's uncertainty about the dynamics of the data.

The bounds procedure comes at a cost: an area of indeterminacy. There is a real possibility analysts will find themselves in a situation where they cannot reach a definitive conclusion. This may seem like a major shortcoming of the procedure. But the uncertainty associated with the dynamic properties of the variables has always affected applied time series analysis. The benefit of the bounds procedure is that this uncertainty is reflected in the hypothesis test. In the next section we demonstrate the utility of our proposed bounds procedure using data from previously published work.

6 Applications

We demonstrate our approach using two examples. In the first example we analyze the dynamics of public policy mood in the United States and in the second we look at explanations of presidential success in Congress.

Public policy mood is conceptualized as the overall predisposition among the public for an activist government role in solving society's problems (Stimson 1991, 1998). It is measured from thousands of survey questions asking about preferences for more or less government in the domain of social policy. Measures of policy mood have been developed for a multitude of countries, and efforts to explain the dynamics of policy mood have proliferated to countries as diverse as Britain (Bartle, Dellepiane-Avellaneda, and Stimson 2011; Green and Jennings 2012), France (Stimson, Tiberj, and Thiébaut 2010; Stimson, Thiebaut, and Tiberj 2012; Brouard and Guinaudeau 2015), Mexico (Baker 2015), Spain (Bartle, Bosch, and Orriols 2014), Portugal, Germany, and Japan.

https://doi.org/10.1017/pan.2019.3 Published online by Cambridge University Press

Table 7. Unit root and stationary tests: public policy mood, public policy outcomes, unemployment rate, and inflation: second quarter 1968 through the fourth quarter 2010 ($T = 168$).

Test	Mood	Policy outcomes	Unemployment rate	Inflation rate
Dickey–Fuller				
τ_τ	−2.95	−1.86	-3.14^{+}	−2.55
ϕ_3	4.70	1.82	5.05	3.34
ϕ_2	3.30	1.23	3.41	2.26
τ_μ	-2.95^{*}	−1.91	-3.16^{*}	−1.31
ϕ_1	4.59^{+}	1.86	5.07^{*}	0.91
τ	−0.82	0.11	0.04	−0.88
KPSS				
τ, long	0.16^{*}	0.13^{+}	0.14^{+}	0.09
τ, short	0.36^{**}	0.32^{**}	0.26^{**}	0.15^{*}
μ, long	0.17	0.14	0.15	0.76^{**}
μ, short	0.38^{+}	0.35^{+}	0.28	1.63^{**}

Note: Shown are (augmented) Dickey–Fuller (Dickey and Fuller 1979) test results for the null hypothesis that the series is a unit root (τ) possibly with drift (τ_μ) and trend (τ_τ). Also reported are tests of the null hypothesis that the constant, trend, and lagged dependent variable are jointly zero (ϕ_2), that the trend and lagged dependent variable are jointly zero (ϕ_3), and that the constant and lagged dependent variable are zero (ϕ_1). The lag length for the test was selected using the AIC (maximum of 12 lags). The KPSS (Kwiatkowski *et al.* 1992) test is of the null hypothesis that the series is stationary around a trend (τ) or a mean (μ). We present test results for both a long and short lag truncation. $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$.

Durr (1992) was the first to elucidate a theory to explain the dynamics of public policy mood and model its behavior in the United States. He argued that economic expectations and policy output exhibit long run equilibrium relationships with policy mood. According to Durr, perceptions of economic security pave the way for the implementation of a more expensive liberal domestic policy agenda "by fostering a willingness among the public to pay for such policies" (Durr 1992, 159). His analysis (and others) also recognizes the "thermostatic" nature of policy preferences: the more (less) spending on domestic policy, the less (more) Americans demand it (Wlezien 1995). Since Durr's seminal analysis, many others have extended his work (Erikson, MacKuen, and Stimson 2002; Enns and Kellstedt 2008; Ellis and Faricy 2011; Ferguson, Kellstedt, and Linn 2013; Owen and Quinn 2016). Here we reanalyze Ferguson, Kellstedt, and Linn (2013)'s replication and extension of Durr (1992) and Erikson, MacKuen, and Stimson (2002). We focus on their model of mood as a function of inflation, unemployment, and policy outcomes from 1968 (second quarter) through 2010.[26]

Policy mood has been treated as a unit root (Durr 1993) and as a stationary time series (Erikson, MacKuen, and Stimson 2002; Ferguson, Kellstedt, and Linn 2013), with some noting that "mood potentially has a unit root" (Owen and Quinn 2016, p. 107) and others omitting any discussion of the question (Ellis and Faricy 2011). Grant and Lebo (2016) note that, if policy mood is a unit root, it is a bounded unit root because it has upper and lower limits (Cavaliere and Xu 2014). The nature of the dynamics of the public policy mood time series is not obvious from the usual battery of statistical tests. In Table 7 we present the evidence on this score.

26 We use the measure of policy outcomes adopted by Ferguson, Kellstedt, and Linn (2013). The measure is an index that averages the following four time series: the percentage of total federal outlays dedicated to human resources, 100% minus the percentage of total federal outlays dedicated to defense, the percentage of total state and local revenues consisting of federal grants-in-aid, and the top marginal tax rate for married persons filing jointly.

We first consider tests of the null hypothesis that mood contains a unit root. We use the Dickey–Fuller test and adopt an iterative testing procedure, assuming we are agnostic about whether the series contains a constant or trend under the null (Dickey and Fuller 1979). We begin with the most general form of the test, including a constant and trend in the test regression.[27] Based on the ϕ_2 and ϕ_3 joint hypothesis tests, we conclude mood is not trending and so estimate a test regression omitting the trend. We then use ϕ_1 to test for the inclusion of a constant in the test regression. Here the results are ambiguous. If we adopt a 0.05 significance level, we cannot reject the null, in which case we draw inferences from τ and conclude mood is a unit root. But if we adopt a 0.10 significance level, inference relies on τ_μ and we thus conclude the series is not a unit root in favor of the alternative that it is stationary around a long run mean. The KPSS test does not help clarify our inference (Kwiatkowski *et al.* 1992). Assuming no trend, the test provides different inferences for each lag truncation parameter. Unsurprisingly, given the disparate treatment of the dynamic properties of mood in the literature, we find that inferences about the dynamic properties of mood vary based on the test used and the level of significance adopted by the analyst.

The dynamic properties of the independent variables also affect inference about LRRs. We report test results for inflation, unemployment, and policy outcomes in Table 7. Like mood, there is inconsistent evidence regarding the dynamic properties of policy outcomes. The Dickey–Fuller test suggests policy outcomes follow a simple random walk while the KPSS test (omitting a trend) provides evidence the series may be stationary. In contrast, test results for unemployment and inflation are unambiguous: unemployment is stationary around a long run mean, while inflation is a random walk with neither drift nor trend. Based on the evidence as a whole, we proceed by assuming none of the series contain a deterministic trend. We are otherwise uncertain about the dynamics of mood and policy outcomes, but are willing to conclude that unemployment is stationary and that inflation contains a unit root.

We begin our analysis of public policy mood under the assumption that it is weakly exogenous[28] such that we can estimate the conditional ECM where m_t is policy mood, v_t is an intervention for the Vietnam war,[29] $\boldsymbol{x}_t = (inflation, unemployment, policy)$ and $\boldsymbol{z}_t = (mood, Vietnam, inflation, unemployment, policy)$:

$$\Delta m_t = c_0 + d_1 v_t + \pi_{mm} m_{t-1} + \boldsymbol{\pi}_{\boldsymbol{mx\cdot x}} \boldsymbol{x}_{t-1} + \Sigma_{i=1}^{p-1} \boldsymbol{\psi}_i \Delta \boldsymbol{z}_{t-1} + \boldsymbol{\delta}' \Delta \boldsymbol{x}_t + \mu_t. \quad (11)$$

Estimates of the full model and each LRM are reported in Table 8.[30] Columns one and two contain estimates from the Generalized Error Correction Model (GECM), column three presents estimates of the LRM and their standard errors (equivalently estimated using the delta method and the Bewley IV regression), and column four presents the resulting test statistic. We focus our attention on the LRM test of the null hypothesis of no valid LRR. Using standard critical values, Ferguson et al. reject the null at the 0.05 level for policy outcomes, at the 0.10 level for inflation, and fail to reject the null for unemployment. However, given our uncertainty about the dynamic properties of mood and policy outcomes, the LRM statistics should be compared to the critical bounds given in the top left panel of Table 5 ($T = 150$: 1.35, 3.70). The resulting inference on unemployment ($t = -0.77$) is unchanged, we cannot reject the null hypothesis, regardless of the dynamic properties of either mood or unemployment. The test statistic on the LRM for inflation ($t = 1.96$) and policy outcomes ($t = 2.32$) are well inside the bounds; unless we can be confident

27 In fact, there is little reason to believe mood contains a deterministic trend, a fact confirmed by the test.

28 Granger causality tests are ambiguous as to the legitimacy of this assumption but for pedagogical purposes we proceed as such.

29 The Vietnam intervention is coded as zero in the second quarter of 1968, grows exponentially from the third quarter of 1968 until the second quarter of 1970, takes a value of one through 1975, and returns to zero for the remainder of the data.

30 This specification assumes the disturbances are uncorrelated. Breusch-Godfrey tests indicate that we cannot reject the null hypothesis that the residuals are uncorrelated.

Table 8. A model of domestic policy mood: second quarter 1968 through fourth quarter 2010.

	z_{it-1}	Δx_{it}	LRM x_{it}	LRM t-statistic
Policy mood	−0.229			
	(0.050)			
Inflation	−0.120	−0.118	−0.522	−1.96
	(0.066)	(0.208)	(0.266)	(Between)
Unemployment	−0.082	0.925	−0.360	-0.77
	(0.104)	(0.488)	(0.469)	(Below)
Policy outcomes	−0.103	−0.170	−0.449	−2.23
	(0.046)	(0.204)	(0.201)	(Between)
Vietnam war$_{t-1}$	1.802			
	(0.652)			
Constant	19.35			
	(4.17)			
Breush–Godfrey (8 lags)	0.13			
RMSE	1.96			
$\bar{R}^2$	0.11			
T	168			

Note: The LRM, LRM$_{SE}$, and t-LRM, are equivalently estimated using the delta method and the Bewley instrumental variables regression. Standard errors in parentheses. The t-statistics are reported as "Below" when $|t| < 1.35$, "Between" when $1.35 < |t| < 3.70$, and "Beyond" when $|t| > 3.70$.

mood is a stationary series, we must admit our uncertainty about whether there is a LRR between either of these variables and mood. The ambiguity of these latter inferences is consistent with Ferguson et al.'s findings that Durr's original results are not robust over time or across alternate measures of economic performance.

Our second application examines presidential success in Congress, the percentage of times a president wins on votes on which he took a position in the House of Representatives each year, from 1953 to 2006 (Ornstein, Mann, and Malbin 2008). A long and old debate examines the importance of presidential approval to presidential success. For Neustadt (1960), the power of the president is the power to persuade, which might lead us to expect that political capital—in the form of public support—will lead to legislative accomplishments. Empirical studies have sometimes found support for the approval-success link (Bond and Fleisher 1984; Ostrom, Jr and Simon 1985). In *The Strategic President*, however, George Edwards argues that high approval levels might sometimes matter but that "presidential legislative leadership operates in an environment largely beyond the President's control" (Edwards 2009, p. 150). Lebo and O'Geen (2011) investigate several of these relationships including the effects of institutional factors, specifically, the partisan and ideological divisions within Congress. Tests of LRRs will help us identify the factors that shape the legacies of successes and failures of modern presidents.

It is unclear whether presidential success is a stationary or unit root process, however, and pretesting does not clear up our questions about univariate dynamics. Using the Dickey–Fuller tests we cannot reject the null hypothesis that presidential success is a unit root (with neither drift nor trend) but this result is contradicted by the KPSS μ test. Unless we rely on the short lag truncation parameter and an $\alpha = 0.10$, we cannot reject the null that the series is stationary around a long run mean. The question of whether presidential success is a unit root or a stationary series over this period has no definitive answer.

Pretests for the explanatory variables indicate that we can reject the unit root null hypothesis for President's party's seat share in the House of Representatives and presidential approval. The

https://doi.org/10.1017/pan.2019.3 Published online by Cambridge University Press

Table 9. A model of presidential success, 1953–2006.

	z_{it-1}	Δx_{it}	LRM x_{it}	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)			
Breush–Godfrey (8 lags)	0.43			
RMSE	10.84			
$\bar{R}^2$	0.61			
T	54			

Note: The LRM, LRM_{SE}, and t-LRM, are equivalently estimated using the delta method and the Bewley instrumental variables regression. Standard errors in parentheses. The t-statistics are reported as "Below" when $|t| < 1.40$, "Between" when $1.40 < |t| < 3.62$ and "Beyond" when $|t| > 3.62$.

KPSS test supports this inference for the president's House share but for presidential approval it suggests we can reject the null of stationarity around a trend but not around a long run mean. Both the Dickey–Fuller and KPSS test results for the conditional party government index (CPG) (Aldrich, Berger, and Rohde 2002) suggest that the series is stationary. (See the Supplementary Appendix, Section 2 for full results.)

We incorporate the uncertainty in the pretesting stage in our tests of the null hypothesis of no valid LRR between presidential success and each of our three independent variables by comparing the LRM t-statistics from a GECM of presidential success to the critical bounds given in the top left panel of Table 5 ($T = 75$: 1.40, 3.62). Columns one and two of Table 9 contain estimates from the GECM, column three presents estimates of the LRM and their standard errors (equivalently estimated using the delta method and the Bewley IV regression), and column four presents the resulting test statistic. Our results indicate that there is ambiguity as to the existence of a valid LRR between CPG and presidential success: the LRM t-statistic for CPG (3.17) lies between the bounds. Without knowledge of the univariate properties of both series, we cannot draw a definitive conclusion. Inference on both presidential approval and the president's House share are, however, conclusive. We cannot reject the null that a president's approval is unrelated to his success in the House: the LRM t-statistic (0.53) lies below the bounds. The test statistic for the president's House share (5.38) is above both bounds and supports the existence of a valid LRR in which each point increase in the president's party's share of the House of Representatives increases his success rate by just over 2.3 points in the long run. Both conclusions hold *regardless of what the dynamic properties of the individual series might be*.

These applications illustrate how applying a bounds hypothesis-testing framework to the LRM t-test allows us to be transparent about how the uncertainty in pretesting translates into uncertainty about LRRs between time series. Is an area of indeterminacy in hypothesis testing unsatisfying? Yes, but it also reduces the risk of type I errors. A bounds framework also provides a firmer foundation for conclusions under the alternative hypothesis. Put differently, when the LRM t-statistic is above the bounds, we can be more confident in the reliability of our findings, independent of any conclusions we might draw when pretesting.

7 Conclusion

Time series analysis typically begins with the analyst conducting pretests designed to determine the dynamic properties of one's data. Such tests, one is led to believe, produce clear diagnoses that neatly dictate appropriate modeling and hypothesis-testing strategies. But classification is complicated. Most theories are ambiguous about the univariate properties of data, many political time series are short, and tests often produce conflicting results. When analysts are uncertain whether their time series are $I(0)$, $I(1)$—or some combination of both—the textbook strategies for inference regarding LRRs are untenable. If one can convincingly establish that the dependent variable contains a unit root, PSS's hypothesis-testing framework is a workable alternative for inference that has become extremely popular in economics with over 8,000 citations on Google Scholar. But if, as in our examples above, the dynamic properties of the dependent variable are uncertain, this strategy, too, is untenable and we caution political scientists against the PSS tests.

Instead we offer the following suggestions for empirical researchers uncertain about the dynamics of their data. First, analysts should admit the uncertainty that hypothesis tests often conducted behind closed doors suggest. Second, analysts should adopt a modeling strategy and hypothesis-testing framework that account for that uncertainty. The LRM test combined with the bounds testing framework we recommend here generalizes and improves upon the strategy advocated by PSS by not insisting on certainty over the univariate characteristics of *any* of the variables in the model. Third, analysts should accept the possibility that inference will not permit a definitive conclusion on a hypothesis test. If analysts follow these guidelines we will be more transparent about uncertainty, reduce the false discovery rate, and increase the likelihood that significant findings are reproducible (Esarey 2017).[31]

The procedure we advocate sacrifices the power of the test, increasing the risk of a type II error. The likelihood analysts find themselves in this position will depend on the strength of any LRR, the length of the time series, and the similarity in the dynamics of y_t and $\boldsymbol{x}_t$. This fact suggests that analysts may need larger samples to identify LRRs but it should also push analysts to develop stronger theories about both the univariate dynamics of their data and the nature of LRRs between them. This, in turn, might be used to justify the use of a particular test statistic and critical values. Out of sample forecasting presents another tool for assessing the performance of models when the LRM test statistic falls between the bounds.

In all, classifying political time series typically involves too much guesswork and time series analysts spend much too much time sparring over the nature of their data. Our work provides a way forward that recognizes that uncertainty. Further, when the LRM test statistic is above both bounds, one can reject the null hypothesis, learn about important dynamic relationships, and be believed. Taking uncertainty seriously and following a method that does not rely on tenuous conclusions from pretesting is the best way forward.

Supplementary material

For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2019.3.

References

Aldrich, J. H., M. M. Berger, and D. W. Rohde. 2002. "The Historical Variability in Conditional Party Government, 1874–1944." In *Party, Process, and Political Change in Congress*, edited by D. W. Brady and M. D. McCubbins, 17–35. Stanford, CA: Stanford University Press.

Baker, A. 2015. "Public Mood and Presidential Outcomes in Mexico." In *Mexico's Evolving Democracy: A Comparative Study of the 2012 Elections*, edited by J. I. Dominguez, K. F. Greene, C. H. Lawson, and A. Moreno, 107–127. Baltimore, MD: Johns Hopkins University Press.

31 Analysts may also choose a more stringent test size, whether 0.01 or 0.05 (Benjamin *et al.* 2018).

Banerjee, A., J. Dolado, and R. Mestre. 1998. "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework." *Journal of Time Series Analysis* 19(3):267–283.
Banerjee, A. A., J. Dolado, J. W. Galbraith, and D. F. Hendry. 1993. *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.
Bartle, J., A. Bosch, and L. Orriols. 2014. "The Spanish Policy Mood, 1978–2012." In *8th ECPR General Conference*, 3–6. Glasgow: University of Glasgow.
Bartle, J., S. Dellepiane-Avellaneda, and J. Stimson. 2011. "The Moving Centre: Preferences for Government Activity in Britain, 1950–2005." *British Journal of Political Science* 41(2):259–285.
Benjamin, D. J., J. O. Berger, M. Johannesson, B. A. Nosek, E.-J. Wagenmakers, R. Berk, K. A. Bollen, B. Brembs, L. Brown, and C. Camerer et al. 2018. "Redefine Statistical Significance." *Nature Human Behaviour* 2(1):6–10.
Bewley, R. A. 1979. "The Direct Estimation of the Equilibrium Response in a Linear Model." *Economic Letters* 3:357–361.
Bond, J. R., and R. Fleisher. 1984. "Presidential Popularity and Congressional Voting: A Reexamination of Public Opinion as a Source of Influence in Congress." *Western Political Quarterly* 37(2):291–306.
Box-Steffensmeier, J. M., J. R. Freeman, M. P. Hitt, and J. C. Pevehouse. 2014. *Time Series Analysis for the Social Sciences*. New York: Cambridge University Press.
Box-Steffensmeier, J. M., and R. M. Smith. 1996. "The Dynamics of Aggregate Partisanship." *American Political Science Review* 90(3):567–580.
Brouard, S., and I. Guinaudeau. 2015. "Policy Beyond Politics? Public Opinion, Party Politics and the French Pro-Nuclear Energy Policy." *Journal of Public Policy* 35(1):137–170.
Campbell, J. Y., and P. Perron. 1991. "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots." *NBER Macroeconomics Annual* 6:141–201.
Cavaliere, G., and F. Xu. 2014. "Testing for Unit Roots in Bounded Time Series." *Journal of Econometrics* 178(2):259–272.
Choi, I. 2015. *Almost All About Unit Roots*. New York: Cambridge University Press.
De Boef, S., and J. Granato. 1997. "Near-Integrated Data and the Analysis of Political Relationship." *American Journal of Political Science* 41(2):619–640.
De Boef, S., and L. Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52(1):184–200.
Dejong, D. N., J. C. Nankervis, N. E. Savin, and C. H. Whiteman. 1992. "The Power Problems of Unit Root Test in Time Series With Autoregressive Errors." *Journal of Econometrics* 53(1–3):323–343.
Dickey, D. A., and W. A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." *Journal of the American Statistical Association* 74:427–431.
Dickinson, M. J., and M. J. Lebo. 2007. "Reexamining the Growth of the Institutional Presidency, 1940–2000." *Journal of Politics* 69(1):206–219.
Durr, R. H. 1992. "Of Forests and Trees." *Political Analysis* 4:255–258.
Durr, R. H. 1993. "What Moves Policy Sentiment? *American Political Science Review* 87(1):158–170.
Edwards, G. C. 2009. *The Strategic President*. Princeton, NJ: Princeton University Press.
Elliott, G., T. Rothenberg, and J. H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root." *Econometrics* 64:813–836.
Ellis, C. R., and C. Faricy. 2011. "Social Policy and Public Opinion: How the Ideological Direction of Spending Influences Public Mood." *The Journal of Politics* 73(4):1095–1110.
Engle, R. F., and C. W. J. Granger. 1987. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55:251–276.
Enns, P. K., and P. M. Kellstedt. 2008. "Policy Mood and Political Sophistication: Why Everybody Moves Mood." *British Journal of Political Science* 38(3):433–454.
Enns, P. K., and C. Wlezien. 2017. "Understanding Equation Balance in Time Series Regression." *The Political Methodologist* 24(2):2–12.
Ericsson, N. R., and J. G. MacKinnon. 2002. "Distributions of Error Correction Tests for Cointegration." *The Econometrics Journal* 5(2):285–318.
Erikson, R. S., M. B. MacKuen, and J. A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.
Esarey, J. 2017. "Lowering the Threshold of Statistical Significance to $p < 0.005$ to Encourage Enriched Theories of Politics." *The Political Methodologist*, August 7. https://thepoliticalmethodologist.com/2017/08/07/in-support-of-enriched-theories-of-politics-a-case-for-lowering-the-threshold-of-statistical-significance-to-p-0-005/.
Evans, G., and N. E. Savin. 1981. "Testing for Unit Roots: 1." *Econometrica: Journal of the Econometric Society* 49(3):753–779.
Evans, G., and N. Savin. 1984. "Testing for Unit Roots: 2." *Econometrica: Journal of the Econometric Society* 52(5):1241–1269.
EViews. 2017. "AutoRegressive Distributed Lag (ARDL) Estimation. Part 2—Inference." *EViews*, May 8. http://blog.eviews.com/2017/05/autoregressive-distributed-lag-ardl_8.html#mjx-eqn-eq.ardl.20.

Ferguson, G., P. M. Kellstedt, and S. Linn. 2013. "How Does the Economy Shape Policy Preferences? *Electoral Studies* 32(3):544–550.
Granger, C. W., and P. Newbold. 1974. "Spurious Regression in Econometrics." *Journal of Econometrics* 2(2):111–120.
Grant, T., and M. J. Lebo. 2016. "Error Correction Methods with Political Time Series." *Political Analysis* 24(1):3–30.
Green, J., and W. Jennings. 2012. "Valence as Macro-Competence: An Analysis of Mood in Party Competence Evaluations in Great Britain." *British Journal of Political Science* 42(2):311–343.
Greene, W. H. 2017. *Econometric Analysis*. 8th ed. New York: Pearson.
Hendry, D. F. 1995. *Dynamic Econometrics*. Oxford: Oxford University Press.
Juhl, T., and Z. Xiao. 2003. "Power Functions and Envelopes for Unit Root Tests." *Econometric Theory* 19(2):240–253.
Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics* 54:159–178.
Lebo, M. J., and T. Grant. 2016. "Equation Balance and Dynamic Political Modeling." *Political Analysis* 24(1):69–82.
Lebo, M. J., and P. W. Kraft. 2017. "The General Error Correction Model in Practice." *Research & Politics* 4(2): 2053168017713059.
Lebo, M. J., and A. J. O'Geen. 2011. "The President's Role in the Partisan Congressional Arena." *The Journal of Politics* 73(3):718–734.
Lebo, M. J., R. W. Walker, and H. D. Clarke. 2000. "You Must Remember This: Dealing with Long Memory in Political Analyses." *Electoral Studies* 19(1):31–48.
Narayan, P. K. 2005. "The Saving and Investment Nexus for China: Evidence from Cointegration Tests." *Applied Economics* 37(17):1979–1990.
Neustadt, R. E. 1960. *Presidential Power*. New York, NY: Wiley.
Ornstein, N. J., T. E. Mann, and M. J. Malbin. 2008. *Vital Statistics on Congress 2008*. Washington, DC: Brookings Institution Press.
Ostrom, C. W. Jr., and D. M. Simon. 1985. "Promise and Performance: A Dynamic Model of Presidential Popularity." *American Political Science Review* 79(2):334–358.
Owen, E., and D. P. Quinn. 2016. "Does Economic Globaliszation Influence the US Policy Mood? A Study of US Public Sentiment 1956–2011." *British Journal of Political Science* 46(1):95–125.
Perron, P., and S. Ng. 1996. "Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties." *The Review of Economic Studies* 63(3):435–463.
Pesaran, M. H., and Y. Shin. 1998. "An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis." *Econometric Society Monographs* 31:371–413.
Pesaran, M. H., Y. Shin, and R. J. Smith. 2001. "Bounds Testing Approaches to the Analysis of Level Relationships." *Journal of Applied Econometrics* 16(3):289–326.
Pesaran, M. H., and R. P. Smith. 1998. "Structural Analysis of Cointegrating Vars." *Journal of Economic Surveys* 12(5):471–505.
Philips, A. Q. 2018. "Have Your Cake and Eat it Too? Cointegration and Dynamic Inference From Autoregressive Distributed Lag Models." *American Journal of Political Science* 62(1):230–244.
Stimson, J. A. 1991. *Public Opinion in America: Moods, Cycles, and Swings*. Boulder, CO: Westview Press.
Stimson, J. A. 1998. *Public Opinion in America: Moods, Cycles, and Swings*. 2nd ed. Boulder, CO: Westview Press.
Stimson, J. A., C. Thiebaut, and V. Tiberj. 2012. "The Evolution of Policy Attitudes in France." *European Union Politics* 13(2):293–316.
Stimson, J. A., V. Tiberj, and C. Thiébaut. 2010. "Au Service de l'Analyse Dynamique des Opinions." *Revue française de science politique* 60(5):901–926.
Stock, J. H. 1991. "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series." *Journal of Monetary Economics* 28(3):435–459.
Webb, C., S. Linn, and M. Lebo. 2018. "Replication Data for: A Bounds Approach to Inference Using the Long Run Multiplier." https://doi.org/10.7910/DVN/4RCPSE, Harvard Dataverse, V1.
Wlezien, C. 1995. "The Public as Thermostat: Dynamics of Preferences for Spending." *American Journal of Political Science* 39(4):981–1000.