Time Series Analysis and Spurious Regression: An Error Correction[*]

Peter K. Enns
pe52@cornell.edu
Cornell University

Takaaki Masaki
tm427@cornell.edu
Cornell University

Nathan J. Kelly
nathan.j.kelly@gmail.com
University of Tennessee

September 25, 2014

Abstract: For scholars interested in longitudinal analysis, the single-equation (i.e., one way) error correction model (ECM) has become an important analytic tool. The popularity of this method has stemmed, in part, from evidence that the ECM can be applied to both stationary and nonstationary data (e.g., De Boef and Keele, 2008). We highlight an under-appreciated aspect of the ECM—when data are integrated or near-integrated, the ECM *only* avoids the spurious regression problem when (near-)cointegration is present. In fact, spurious regression results emerge close to 20 percent of the time when integrated or near integrated data that are not cointegrated are analyzed with an ECM. Given the near absence of cointegration tests in recent political science publications, we believe this is a fundamental point to emphasize. We also show, however, that when evidence of cointegration is present, the ECM performs well across a variety of data scenarios that political scientists are likely to encounter.

[*]All files necessary to replicate our simulations and produce the figures and tables in this manuscript are available at: `http://thedata.harvard.edu/dvn/dv/ECM`.

As political scientists, we are often interested in change. Why has inequality increased? Why do some authoritarian regimes persist and others transition to democracy? Why do attitudes toward the death penalty, immigration, or welfare vary over time? For researchers with an empirical approach to these types of political dynamics, the single equation (i.e., one way) error correction model (ECM) has become a popular analytic tool. The popularity of the ECM stems, in large part, from the long standing concern of observing "nonsense-correlations" (Yule, 1926) or "spurious regression" (Granger and Newbold, 1974) in non-stationary time series. In fact, one of the most common reasons scholars cite for selecting the ECM is its applicability to nonstationary data (See, for example, Enns and McAvoy (2012, 634), Hellwig (2007, 150), Kayser (2009, 957), Kelly (2005, 873), Kono (2007, 751), McGuire (2004, 133), Ramirez (2009, 686), Voeten and Brewer (2006, 820)). In this article we demonstrate that despite attempts to avoid spurious findings, many prominent applications do not correctly apply the single equation ECM and thus the analyses do not fully address the spurious regression problem.[1]

Nathaniel Beck's (1991) "Comparing Dynamic Specifications" and a subsequent *Political Analysis* symposium on error correction models (Beck, 1992; Durr, 1992*a*; Durr, 1992*b*; Ostrom and Smith, 1992; Smith, 1992; Williams, 1992) helped introduce ECMs to the political science community. Following the work of Clive Granger and Robert Engle (e.g., Granger, 1986; Engle and Granger, 1987), these articles emphasized the connection between integration, cointegration, and error correction. That is, if two (or more) series are *integrated* and they maintain a long run equilibrium relationship (i.e., they are *cointegrated*), this process can be specified with an *error correction* model.[2] Smith (1992, 253-254), for example, explained, "Although Davidson et al. (1978) popularized the ECM before the concept of

[1]All of our analyses focus on standard time series models, but the arguments are also relevant to time series cross sectional analysis.

[2]A classic example of two integrated series that are also cointegrated is a drunk walking a dog (Murray, 1994). The path of the drunk and the dog each reflect an integrated time series, where the position at each time point is a function of the position at the previous time point plus the current stochastic step (integrated time series are often referred to as "random walks" or as having a "unit root"). Now, suppose every time the drunk calls for the dog, the dog moves a little closer to the drunk and every time the dog barks, the drunk aims in the direction of the dog. These movements represent the "error correction" and keep the two

cointegration was developed, they impose restrictions on the parameters of their model that imply integration and cointegration." Durr (1992*b*, 255) similarly noted, "the Granger representation theorem suggests that the appropriateness of the cointegration/error correction methodology rests squarely on our capacity to ascertain whether an individual time-series is integrated, and whether a group of such series are cointegrated."

In addition to offering a statistical procedure that is appropriate when analyzing cointegrated time series, the logic of cointegration offers a relevant theoretical account of many of our series of political interest. Error correction, for example, is central to the theory of thermostatic opinion responsiveness (Jennings 2009; Jennings 2013; Soroka & Wlezien 2010, 177-179; Wlezien 2004). Similarly, theories of democratic representation often (implicitly) assume an equilibrium relationship between the public's preferences and government outputs. If policy moves too far away from the public's preferences, voters will notice and respond by electing new candidates who will bring policy back toward the public's preferences.

Interestingly, despite the empirical and theoretical roots in cointegration, as political scientists have increasingly utilized the ECM, an important development has been a move *away* from the original application to cointegration methodology. De Boef and Granato (2000, 114), for example, conclude, "the [single-equation] ECM can be estimated with stationary time series, and where the characterization of the univariate series is in doubt... it is not necessary to use the language of cointegration to use the language of error correction." By emphasizing the algebraic equivalence of the ECM and the autoregressive distributed lag (ADL) model, De Boef and Keele (2008) also stress that the ECM does not necessitate integration or cointegration. These scholars are correct that when the dependent variable is stationary, an ECM can be estimated without concern for cointegration. However, many scholars have interpreted this research to mean that when the dependent series is integrated or near-integrated (i.e., a root close to but not quite unity), cointegration does not need to

integrated series in a long term equilibrium (i.e., the distance between the two paths is stationary). As a result, we have two integrated series that are *cointegrated*.

be established prior to estimating the ECM (see, e.g., Kelly and Enns (2010, 863), Ramirez (2009, 686), and Ura and Wohlfarth (2010, 950)). Indeed, when analyzing long–memoried series, failure to test for cointegration has become common practice in the discipline. Between 1993 and 2012, 63 articles in the discipline's top journals included a single equation ECM. Of these, only 8 reported a test for cointegration.[3] We show that this practice of not testing for cointegration when the dependent variable is integrated or near-integrated poses a serious threat to ECM analyses.

Given the move in political science to disconnect the ECM from cointegration, our claim that cointegration must be established prior to estimating an ECM with (near-)integrated data may come as a surprise. After all, it is now well known that the ECM and the ADL model are algebraically equivalent (Banerjee, Dolado, Galbraith and Hendry, 1993; De Boef and Granato, 2000; De Boef and Keele, 2008). Furthermore, since the ADL only produces accurate inferences when the dependent series is stationary (e.g., Keele and Kelly, 2006), the ECM (which is simply an alternate parameterization of the ADL) must also be applicable to stationary time series which, by definition, are not cointegrated. As we show in Appendix 2, this logic is sound. The inverse, however, is also true. If the ADL can produce spurious results with integrated data, and if the ADL and ECM are equivalent, the ECM will also produce spurious results with integrated data—unless the integrated series are cointegrated.

The following section further develops this intuition. First, we explain why cointegration *must* be established prior to estimating an ECM when the dependent variable has a unit root. Although some time series practitioners will not be surprised by this claim, as we indicated above, this statement offers a major revision to most current political science applications of the ECM. Second, we demonstrate the importance of first testing the time series properties

[3] 34 of these articles appeared in the *American Journal of Political Science*, *American Political Science Review*, and *Journal of Politics*. The rest appeared in the *British Journal of Political Science*, *Comparative Politics*, *International Organization*, *International Studies Quarterly*, *Journal of Conflict Resolution*, and *World Politics*. To identify relevant articles, we searched JSTOR for the word, "error correction." We then read the methods and results sections to verify that a single equation ECM was estimated and to determine if a cointegration test was reported. We also searched the PDF for "coint" to identify cointegration tests.

of the dependent variable before testing for cointegration. This step is necessary because standard cointegration tests will produce *incorrect* inferences if the dependent variable is stationary. Third, we develop the implications of estimating an error correction model with near-integrated data. We show that when near-integrated data maintain an equilibrium relationship (i.e., the series are near-cointegrated), the ECM will provide accurate statistical inferences regarding the relationships between predictor and outcome variables. We also address an often ignored question in ECM analysis—in the multivariate context, what are the implications of a cointegrating relationship between two variables but not other variables in the model? We show that when two (near-)integrated series are (near-)cointegrated and a third (near-)integrated series is not related to the dependent variable of interest, the ECM greatly diminishes the spurious regression problem.

Given the numerous political and economic series that behave as if they are near-integrated—e.g., oil consumption (Duffield and Hankla, 2011), tariff rates (Ehrlich, 2007), social spending (Faricy, 2011; Nooruddin and Simmons, 2009), the level of employment (Iversen and Wren, 1998), trust in the government (Keele, 2007), presidential approval (Lebo, 2008, 8), the index of consumer sentiment (Lebo, Walker and Clarke, 2000, 40), and foreign debt (Oatley, 2012)—we believe our focus on the behavior of the ECM with near-integrated data will be of importance to many political and economic applications. The analysis uses a series of Monte Carlo experiments to test the previous claims. We then turn to two important empirical applications to illustrate our proposed use of the ECM. The conclusion discusses the implications of our argument as well as limitations and caveats.

Spurious Regression in the Context of ECMs

The single-equation (i.e., one way) ECM for the bivariate case can be expressed as follows:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 \Delta X_t + \beta_1 X_{t-1} + \epsilon_t \tag{1}$$

where, γ_1 and β_1 represent the short-run and long-run effect of X, respectively. The long-run estimate (β_1) can be thought of the combined effect of X_t and X_{t-1}. Since this combined effect is distributed over future time periods (through Y_{t-1}), we can estimate the total effect of X with the long-run multiplier (LRM), which equals $\frac{\beta_1}{\alpha_1}$.

We begin our discussion assuming Y_t and X_t contain a unit root. We then consider the implications for Equation 1 when X_t and Y_t are near-integrated. Granger and Newbold's (1974) classic article on spurious regression focused on the problems associated with regressing an integrated Y_t in levels on an integrated X_t in levels. Most notably, regressing one integrated series on another integrated series produces integrated—I(1)—residuals, which produce the spurious regression problem that Granger and Newbold highlighted. As Enders (2004, 171) explains, "the variance of the error becomes infinitely large as t increases... Hence, the assumptions embedded in the usual hypothesis tests are violated."

The ECM, however, differences Y_t, producing a stationary series on the left hand side of the equation. We are no longer regressing levels on levels. Since ΔY_t is stationary, we should expect ϵ_t to also be stationary, eliminating the source of spurious regression highlighted by Granger and Newbold (1974).[4] Since the ECM addresses Granger and Newbold's (1974) concern of serially correlated residuals, it is not surprising that so many scholars have concluded that the ECM will avoid spurious regressions with long–memoried time series (see, e.g., De Boef and Keele, 2008, 195). However, absent cointegration, the ECM still produces spurious results when Y_t is integrated—*even when the residuals are stationary.*

To illustrate this point empirically, we generated two integrated series Y_t and X_t through the following data generating processes (DGPs):

[4] The simulations below confirm that the ECM largely eliminates serial correlation in the residuals when Y_t has a unit root. This result emerges because if Y_t is $I(1)$, ΔY_t is $I(0)$ and regressing a stationary, $I(0)$, process on an integrated variable will typically result in stationary residuals.

$$Y_t = Y_{t-1} + u_{1,t} \tag{2}$$

$$X_t = X_{t-1} + u_{2,t} \tag{3}$$

where

$$\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \sim IN \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \tag{4}$$

By design, there is no short-run or long-run effect of X_t on Y_t because these two series are generated with independent DGPs. We conduct 2,000 simulations for the sample size T of 30, 60, and 100.[5] These values of T represent common lengths of time series in political science data. For each experiment, we estimate the single equation ECM in Equation 1. We would expect to observe a significant relationship between X_{t-1} and ΔY_t about 5 percent of the time. Instead, β_1 is significant in 14.9%, 16.7%, and 17% of the simulations when T=30, 60, and 100, respectively. If we look at the total expected effect of X_{t-1} on Y_t (which is captured by the LRM, $\frac{\beta_1}{\alpha_1}$), the proportion of spurious results increases to 18.1%, 18.5%, and 19.7%, respectively (See Appendix 1 for the complete results). As suggested above, autocorrelation in the residuals does *not* account for these spurious results. A Portmanteau (Q) test for white noise indicates that the residuals are white noise in almost 95 percent of the observed spurious regressions. These spurious results emerge because even though the ECM largely solves the serial correlation problem by differencing Y_t, unless cointegration is present, the ECM produces what Banerjee et al. (1993, 164) describe as an "unbalanced regression," which they define as a regression equation, "in which the regressand is not of the same order of integration as the regressors, or any linear combination of the regressors." As the above results demonstrate, unbalanced regressions (which are sometimes also referred to as inconsistent regressions) produce a substantial spurious regression problem (Banerjee, 1995;

[5] We first generated Y_t and X_t so T=130, 160, or 200, and then discarded the first 100 observations to avoid the potential influence of initial values (Jönsson, 2005; Larsson, Lyhagen and Löthgren, 2001).

Mankiw and Shapiro, 1985; Mankiw and Shapiro, 1986; Marmol, 1996; Ventosa-Santaulària, 2012). As Maddala and Kim (1998, 251) explain, "a requirement in order to obtain a meaningful estimation with integrated variables is balance in the orders of integration of the variables on the left-hand and right-hand side of the regression equation."

Unless a cointegrating relationship exists between Y_t and X_t, Equation 1 is unbalanced and standard t-statistics do not apply. By contrast, when cointegration is present, the ECM is balanced, and Equation 1 will produce accurate standard errors that yield valid inferences. To see why cointegration balances Equation 1, it is helpful to rewrite the equation in the following form:

$$\Delta Y_t = \alpha_0 + \alpha_1(Y_{t-1} + \frac{\beta_1}{\alpha_1}X_{t-1}) + \gamma_1 \Delta X_t + \epsilon_t \tag{5}$$

X_t and Y_t are cointegrated when X_t and Y_t are both integrated (of the same order) and α_1 and β_1 are non-zero—which implies that the LRM effect of X_t on Y_t is also non-zero. Any deviation from the long-run equilibrium, expressed by $(Y_{t-1}+\frac{\beta_1}{\alpha_1}X_{t-1})$, is corrected at the rate of α_1, which is called the error correction term, which can range between 0 and -1.0. Due to the presence of this error correction term, which serves as a leash that binds them together, these two series never drift too far from each other.[6] Furthermore, because cointegration ensures that Y_t and X_t maintain an equilibrium relationship, a linear combination of these variables exists that is stationary (that is, if we regress Y_t on X_t, in levels, the residuals would be stationary).[7] As noted above, this (stationary) linear combination is captured by $(Y_{t-1} + \frac{\beta_1}{\alpha_1}X_{t-1})$. Additionally, since Y_t and X_t are both integrated of order one, ΔY_t and ΔX_t will be stationary. Thus, cointegration ensures that the equation is balanced: the regressand (ΔY_t) and either the regressors (ΔX_t) or a linear combination of the regressors

[6]When applied to stationary data, the error correction term is still of interest because it tells how fast the effect of X_t on Y_t diminishes over future time periods, but these two series do not contain a long-run equilibrium that they always revert back to. Thus, while cointegration entails error correction, error correction does not necessarily entail cointegration. For this reason, Hendry (2008) differentiates between error correction and equilibrium correction, where equilibrium correction specifies cointegration.

[7]This, in fact, is the first step of the Engle-Granger two-step method of testing for cointegration. This procedure highlights that a stationary linear combination between Y_t and X_t is *not* the same as stationary residuals, ϵ_t, in Equation 5.

$((Y_{t-1} + \frac{\beta_1}{\alpha_1} X_{t-1}))$ are *all* stationary.[8] By contrast, when Y_t and X_t are integrated and there is no evidence of cointegration, Equation 5 reduces to the differences model, which only includes a constant and ΔX_t as regressors. This differences model should be estimated if the data are integrated but not cointegrated.

Identifying Cointegration

We have seen that in order to ensure a balanced equation, it is crucial to establish cointegration prior to estimating an ECM with integrated data. To test for cointegration, we examine whether the error correction term (i.e., α_1 in Equation 1) is statistically significant (see, e.g., Ericsson and MacKinnon (2002)). We refer to this as the error correction (EC) statistic. For the EC statistic, the distribution under the null is Dickey-Fuller, so it is crucial to apply the correct critical values (Banerjee, Dolado and Mestre, 1998; Ericsson and MacKinnon, 2002; Kremers, Ericsson and Dolado, 1992). We rely on the critical values reported in Ericsson and MacKinnon (2002).[9]

Although previous scholars have drawn attention to the importance of using the correct critical values with the EC statistic, another important condition of the EC statistic has not been sufficiently emphasized. Researchers must first identify the time series properties of the dependent variable before interpreting the EC statistic, because when the dependent variable is stationary, α_1 will have a large t-statistic and be close to negative one (even though cointegration is *not* present). To see why, suppose that the DGP for Y_t is:

$$Y_t = \rho Y_{t-1} + u_t \sim IN(0, 1) \tag{6}$$

If we subtract Y_{t-1} from both sides of Equation 6, we see that $\Delta Y_t = (\rho - 1)Y_{t-1} + u_t$. Thus,

[8] The single equation ECM also assumes weak exogeneity (e.g., Boswijk and Urbain, 1997; Engle, Hendry and Richard, 1983). Interestingly, De Boef (2001, 91) finds that as long as X_t is (near-)integrated (i.e., the autoregressive parameter is >0.90) and the covariance of the residuals in the DGP of Y_t and X_t are <0.30, "long-run relationships can be estimated reasonably well."

[9] The Engle-Granger (EG) cointegration test (Engle and Granger, 1987) can also be used to test for cointegration.

when we estimate the ECM in Equation 1, our estimate of $\widehat{\alpha}_1$ includes the error correction rate *plus* $(\rho - 1)$. When Y_t has a unit root (i.e., $\rho = 1$), $\widehat{\alpha}_1$ correctly estimates the error correction rate because $\widehat{\alpha}_1 = (1 - 1) + \alpha_1$. However, when $\rho < 1$, $\widehat{\alpha}_1$ estimates $(\rho - 1) + \alpha_1$. For example, if $\rho = 0.1$, $\widehat{\alpha}_1$ will equal -0.9 plus the actual error correction rate. Even if there was *no* error correction in the DGP, our estimate of α_1 would approximate $-.9$ because Y_t is stationary, not because a cointegrating relationship exists between X_t and Y_t. We should emphasize that because the ECM is equivalent to the ADL, like the ADL, the ECM can be estimated with stationary data. Thus, the ECM produces correct inferences when Y_t is stationary *or* when Y_t is integrated and cointegration is present. However, if applying the ECM when the outcome variable is stationary, researchers must remember that in these cases the coefficient on Y_{t-1} includes $(\rho - 1)$ and thus *cannot* be used as a test of cointegration.

Expectations for Near-Integrated Data

So far, we have considered the ECM when Y_t and X_t have a unit root and the implications for α_1 when Y_t is stationary. Yet, many political series cannot be strictly classified as integrated of order zero, I(0), or of order one, I(1). Fractional integration allows the order of integration (d) to range between 0 and 1. Factional integration characterizes many important political series and when series demonstrate fractional integration processes ("i.e., the value of d is significantly greater than 0 and significantly less than 1.0" (Lebo and Clarke, 2000, 4)), appropriate fractional integration methods should be employed (e.g., Box-Steffensmeier and Smith, 1998; Box-Steffensmeier and Tomlinsonl, 2000; Lebo, Walker and Clarke, 2000). Our interest lies, however, in a more specific—albeit important—dynamic process: near-integrated time series. Near-integrated series are distinct from fractionally integrated processes because they have a root close to but not quite unity (Phillips, 1988).[10] Although near-integrated series can include both strongly autoregressive and mildly explo-

[10] Although both fractionally integrated series and near-integrated series are characterized as long memory processes, fractionally integrated processes are associated with hyperbolically decaying autocorrelations while near-integrated processes are associated with geometric or exponential rates of decay (Baille, 1996, 5-6).

sive processes, we limit our focus to conditions when d is close to but less than one. As we have noted above, many important political series behave as if they are near-integrated. In addition to being relevant to political science research, we are interested in near-integrated series because we propose that despite not having a unit root, the single equation (one-way) ECM will produce valid inferences with near-integrated series that form a cointegrating relationship. De Boef and Granato (2000, 103) refer to this as *near-cointegration*, meaning that a linear combination of the near-integrated series is stationary.

The intuition behind the claim that the ECM can be used with near-integrated series that are near-cointegrated is straightforward. In finite samples, near-integrated series and integrated series are almost indistinguishable (e.g., Banerjee et. al. 1993, 96; DeBoef & Granato 2000, 102; Phillips 1987, 544). De Boef and Granato (1997, 620), in fact, find that "statistical distinctions between integrated and near-integrated data occur only asymptotically when the sample size approaches infinity" and Hamilton (1994, 446) refers to the "identical observable implications" of these series. For statistical analysis, it is the sample properties—not the asymptotic properties—of variables that matter. Thus, although near-integrated series are stationary asymptotically, because they behave like integrated series when T is small (which is the case with most time series in political science), we should analyze them as if they were integrated.

To illustrate this point, consider a long-memoried series like presidential approval. Theoretically, presidential approval cannot have a unit root. Since the percent approving of the president is bound between 0 and 100, by definition, the series has a finite variance. Additionally, because presidential approval is an aggregation of individual attitudes, which likely follow different autoregressive processes, we should not expect aggregate approval to have a unit root (Granger, 1980). Yet, for monthly presidential approval between 1995 and 2005, Lebo (2008) reports a fractional differencing parameter, d, of 0.94 and explains that "The presidential approval series is nearly a unit-root and clearly has long, if not quite perfect, memory" (8). Despite its asymptotic properties, in the observed data presidential approval is near-integrated and behaves as if it is integrated.

We propose that with series like presidential approval, which have a long memory, an augmented Dickey Fuller (DF) test should be conducted to evaluate whether the null of a unit root can be rejected. At first, our recommendation to employ a DF test may seem surprising. It is well known that DF tests are underpowered against the alternative hypothesis of stationarity (Cochrane, 1992; Blough, 1992). Thus, the obvious concern is that we may incorrectly conclude that a stationary or fractionally integrated series, which does not mimic the statistical properties of a unit root, behaves like an integrated series. However, for our specific application, the low power of the DF does *not* pose a problem.

To see why the DF test is appropriate, we must consider the implications of the DF test for the EC test statistic. Recall the previous section which demonstrated that when the order of integration of Y_t is less than one (i.e., $\rho < 1$), the t-statistic associated with $\widehat{\alpha}_1$ in Equation 5 will increase because $\widehat{\alpha}_1$ reflects the error correction rate (α_1) *plus* the autoregressive component of Y_t minus 1 (i.e., $\rho - 1$). Even though $\widehat{\alpha}_1$ includes $\rho - 1$, if a DF test accepts the null of a unit root, the t-statistic associated with $\widehat{\alpha}_1$ will still provide a correct test of cointegration. This result emerges because the DF test is based on $\rho - 1$. Specifically, the DF test estimates,

$$\Delta Y_t = \phi + \delta Y_{t-1} + \epsilon_t, \tag{7}$$

where δ is equivalent to $\rho - 1$ and the DF test rejects the null hypothesis of a unit root if the t-statistic associated with δ is significant.[11]

With this in mind, suppose the true error correction rate, α_1, is zero. When Y_t is stationary, adding $(\rho - 1)$ to 0 can produce a significant $\widehat{\alpha}_1$. However, if a DF accepts the null of a unit root, $\widehat{\delta}$ (which equals $(\rho - 1)$) is—by definition—not statistically different

[11] The DF test does not follow a normal distribution so the correct critical values must be used. If serial correlation is present in ϵ_t, lagged values of ΔY_t should be included in the test to produce stationary residuals, producing an augmented Dickey Fuller test.

from zero. If the error correction rate is zero, adding an insignificant $\widehat{\delta}$ will not produce a significant estimate of $\widehat{\alpha}_1$; i.e., we will not find incorrect evidence of cointegration.[12] As a consequence, despite the DF test's lack of power against the alternative hypothesis, for this particular application, the DF test is well suited. Even though a DF test cannot tell us the exact properties of the time series, when a DF test does not reject the null hypothesis of a unit root in Y_t, the t-statistic associated with $\widehat{\alpha}_1$ will be accurate (provided the correct critical values are used). Interestingly, we do *not* need to test the time series property of X_t. Recall that $\widehat{\alpha}_1$ will only be significant if a linear combination of Y_t and X_t is I(0) (provided that the DF test finds evidence of integration in Y_t), which is only possible if Y_t and X_t are integrated of the same order. Thus, a significant $\widehat{\alpha}_1$ implies that Y_t and X_t behave as if they are integrated of the same order.

Although $\widehat{\alpha}_1$ will produce correct inferences when used as a test of cointegration, with near-integrated data, $\widehat{\alpha}_1$ will over-estimate the rate of equilibrium error correction. The bias will be directly in proportion to Y_t's deviation from I(1) (or (ρ-1)). Our simulation results show, however, that with near–integrated data, this bias is relatively minor. Nevertheless, in Appendix 3 and in our first empirical example we show the conditions when this bias is most likely to be a problem and how it can be addressed by estimating $\rho - 1$ and subtracting this value from $\widehat{\alpha}_1$ to recover the unbiased estimate of α_1.[13]

Summary of Predictions

1. When Y_t is (near-)integrated, (near-)cointegration must be established prior to estimating an ECM.

[12]We have tested whether this theoretical argument holds true through Monte Carlo simulations. Using the same DGPs for Y_t and X_t as defined in Equations (2), (3), and (4), we compute the probabilities of finding evidence of cointegration (i.e., a statistically significant coefficient for the error correction estimate based on the MacKinnon critical values) after testing for a unit root in Y_t using the DF test. We let the memory of Y_t (ρ) and the memory of X_t (q) each take the values of 0.1, 0.4, 0.7, 0.9, 0.95, 0.99, and 1.0. We find that the rate of finding false evidence of cointegration never exceeds 5.5%, confirming that even with near-integrated series, using the DF test as the first step guards against falsely inferring cointegration.

[13]In addition to Monte Carlo simulations to evaluate our adjustment procedure, we also examined if fractional error correction models (FECMs) produce more accurate estimates of the true error correction term. We find that both the original and adjusted estimates of the error correction term in the ECM consistently outperform the estimates from the FECMs. See Appendix 3 and Appendix 4 for further details.

2. If an augmented Dickey Fuller test does not reject the null of a unit root in Y_t, we can treat the series as if it is near-integrated.

3. When Y_t is near-integrated, the EC test for cointegration (based on $\widehat{\alpha}_1$) is accurate.

Monte Carlo Experiments

In this section, we test the predictions developed above with a series of Monte Carlo Experiments. We consider a three variable system (Y_t, $X_{1,t}$, and $X_{2,t}$), though—as we demonstrate in Appendix 6—the intuition and findings generalize beyond three variables. The following three variant situations are examined in our Monte Carlo experiments:[14]

Case 1: Y_t, $X_{1,t}$, and $X_{2,t}$ are all (near-)integrated and unrelated to each other.

Case 2: Y_t, $X_{1,t}$, and $X_{2,t}$ altogether form one cointegrating relationship.

Case 3: Y_t is (near-)integrated and (near-)cointegrated with just one of the X's.

In each case, we conduct 2,000 simulations for the sample size T of 30, 60, and 100. The following single-equation ECM is estimated for the trivariate system:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma_1 \Delta X_{1,t} + \beta_1 X_{1,t-1} + \gamma_2 \Delta X_{2,t} + \beta_2 X_{2,t-1} + \epsilon_{1,t} \tag{8}$$

Case 1: Y_t, $X_{1,t}$, and $X_{2,t}$ are all (near-)integrated and unrelated to each other.

We first examine the situation where Y_t, $X_{1,t}$, and $X_{2,t}$ are (near-)integrated and have no relationships amongst themselves. That system is:

$$Y_t = \rho Y_{t-1} + u_{1,t}, \text{where } u_{1,t} \sim iid\, \mathrm{N}(0,1) \tag{9}$$

$$X_{1,t} = q X_{1,t-1} + u_{2,t}, \text{where } u_{2,t} \sim iid\, \mathrm{N}(0,1) \tag{10}$$

$$X_{2,t} = \xi X_{2,t-1} + u_{3,t}, \text{where } u_{3,t} \sim iid\, \mathrm{N}(0,1) \tag{11}$$

[14] A fourth possibility is that more than one cointegrating vector exists. Because of our focus on the single equation ECM, we do not elaborate on multiple cointegrating vectors. However, a Johansen test can be used to identify the number of cointegrating vectors (e.g., Johansen, 1995). If there is more than one cointegrating vector, a system of equations needs to be estimated simultaneously (e.g., a VECM model).

Because we are interested in (near-)integrated processes, we let ρ, q, and ξ each take values of 0.9 to 1.0 by increments of 0.05. (Importantly, Appendix 2 shows that regardless of the autoregressive parameter in the DGP of Y_t and X_t, as long as a DF test indicates Y_t has a unit root and the EC statistic indicates that cointegration is present, the ECM does not exceed the expected 5 percent spurious correlations.)

If we find statistically significant effects of $X_{1,t}$ or $X_{2,t}$ on Y_t in this system, the findings are spurious. Due to the absence of cointegration, we expect that unbalanced regressions will lead to spurious regressions (as we saw with the bivariate simulations reported above). However, by combining a DF test of Y_t with a test for cointegration, we expect to eliminate the spurious regression problem. Since each of the X_t variables is a mirror image of the other, we only report spurious findings on $X_{1,t}$. Also, to save space, we only report the results from the simulations where ρ, q, and ξ take the values of either 0.9 or 1.0. In Table A-4 in the Appendix, we report the results from the full set of simulations.

The results in Table 1 are consistent with expectations. When $X_{t,1}$ and Y_t are (near-)integrated but not (near-)cointegrated, failure to test for a unit root in Y_t and for cointegration prior to estimating an ECM (Column 1) leads to spurious findings. The rate of spurious results is roughly the same as in the bivariate case. Column 2 shows, however, that when we only analyze dependent series that show evidence of a unit root and we use the correct critical values to test for cointegration, we eliminate the spurious regression problem.[15] In fact, all results are below 5 percent. The reason that our rate of spurious correlation appears conservative is as follows. Given the small sample sizes and the strongly autoregressive DGP in Y_t, we expect the DF test to (correctly) identify a unit root in Y_t 95% of the time. We also expect the EC test to (falsely) identify cointegration 5% of the time and the t-test on β_1 to (falsely) identify the relationship between ΔY_t and X_{t-1} 5% of the time. So the expected rate of false rejection should be $0.95 \times 0.05 \times 0.05 = 0.0023$ (or 0.23%). In practice, we

[15]Table A-4 in the Appendix shows that examining the EC statistic with correct critical values is crucial. If we use standard t critical values, we would reject the true null between 7 and 14 percent of the time.

Table 1: Case 1 (β_1=0). Rejection Frequencies (%) for the True Null Hypothesis:

Y	$X_{1,t}$	$X_{2,t}$	No Test	EC$_M$
ρ	q	ξ	(1)	(2)
			T=30	
0.9	0.9	0.9	11.95	1.70
0.9	0.9	1.0	12.00	1.50
0.9	1.0	0.9	13.95	2.15
0.9	1.0	1.0	12.90	2.00
1.0	0.9	0.9	13.65	1.65
1.0	0.9	1.0	12.05	1.85
1.0	1.0	0.9	15.20	2.00
1.0	1.0	1.0	13.95	1.70
			T=60	
0.9	0.9	0.9	11.45	1.90
0.9	0.9	1.0	11.90	2.70
0.9	1.0	0.9	12.90	2.75
0.9	1.0	1.0	13.25	2.55
1.0	0.9	0.9	10.85	1.20
1.0	0.9	1.0	12.55	1.45
1.0	1.0	0.9	17.15	2.45
1.0	1.0	1.0	15.90	1.70
			T=100	
0.9	0.9	0.9	9.25	1.30
0.9	0.9	1.0	9.70	1.60
0.9	1.0	0.9	11.70	2.85
0.9	1.0	1.0	12.15	3.20
1.0	0.9	0.9	9.55	0.70
1.0	0.9	1.0	10.80	1.30
1.0	1.0	0.9	18.10	1.55
1.0	1.0	1.0	15.45	1.70

Notes: EC$_M$ uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). ρ, q and ξ correspond to the memories of Y_t, $X_{1,t}$ and $X_{2,t}$, respectively. In No Test, we do not conduct pretests for cointegration or integration in Y_t.

observe values greater than this (but less than 5%) because the probabilities of getting a null finding across these three tests are not independent from each other.

Case 2: Y_t, $X_{1,t}$, and $X_{2,t}$ altogether form one cointegrating relationship

In Case 2, (near-)cointegration (and thus a long run relationship) exists between $X_{1,t}$, $X_{2,t}$, and Y_t. We use the term (near-)cointegration to acknowledge that the memory of X_t

will be close to, but is not necessarily one. Our interest in this section is the power of the ECM (when coupled with tests for cointegration and the time series properties of Y_t) to identify the existence of a relationship between X_t and Y_t. The DGP for Y_t now becomes:

$$Y_t = Y_{t-1} + \alpha_1(Y_{t-1} + \frac{\beta_1}{\alpha_1} X_{1,t-1} + \frac{\beta_2}{\alpha_1} X_{2,t-1}) + u_{5,t}, \text{where } \; u_{5,t} \sim iid\,\mathrm{N}(0,1) \qquad (12)$$

As in Case 1 (Equations 10 and 11), we allow the memory of the X_t variables to equal 0.9, 0.95, or 1.0. However, we now set $\beta_1 = \beta_2 = 0.3$. The decision to set the long run relationship (i.e., β_1 and β_2) to 0.3 provides a conservative test of the ability of the ECM to identify a relationship. For example, De Boef (2001, 86) uses 0.5 in her simulations.[16] The error correction (or cointegration) term α_1 takes values from -0.1 to -0.4 by increments of 0.1. We select these values because they represent relatively modest error correction rates in political science research and thus offer a conservative test.[17] If we are able to uncover evidence of cointegration and long run relationships in our simulations, we can be confident that the methods are relevant for political science applications. For each value of α_1, we estimate the ECM in Equation 8 and examine if $\hat{\beta}_1$ is statistically significant at the .05 level (we do not report β_2 because β_1 and β_2 are mirror images of each other).

Table 2 presents results for this second case (full results appear in Table A-5). The table reports the frequencies of correctly identifying the relationship between $X_{1,t}$ and Y_t. Column 2 indicates that the additional steps of testing for evidence of a unit root in Y_t and testing for cointegration come at some cost, reducing the probability of correctly finding a significant long-run relationship between Y_t and $X_{1,t}$. This is especially the case when the error correction term α_1 and T are both small. When α_1=-0.1 and T=30, the rejection

[16]Formally, De Boef (2001, 86) sets β to 1.0. However, in the DGP used by De Boef (2001), β is multiplied by the error correction term, which is set to 0.5. Thus, the long run relationship in her simulations is 0.5.

[17]Ahlquist (2006), for example, finds error correction rates of -0.62 and -0.65 for portfolio capital flows and foreign direct investment, respectively. Ura and Wohlfarth (2010, 951) report an error correction rate of -0.68 in their analysis of U.S. Supreme Court institutionalization and Keele (2007, 249) reports an error correction rate of -0.39 in his analysis of trust in government. An error correction rate of -0.4 implies that just over 90 percent of the total effect of X_t on Y_t will be realized after five periods.

frequencies are between 6% and 9% (compared to 51%—59% for situations where we do not test for either integration or cointegration). Importantly, the potential cost of low power of the ECM (when paired with necessary pretests) becomes negligible as the error correction rate increases (i.e., α_1 decreases) and T becomes larger. For instance, for $\alpha_1 = -0.4$ and T =60, we correctly reject the false null hypothesis more than 90% of the time. When T =100, the ECM's ability to correctly identify true relationships approaches 100%. For sufficiently negative α_1, large T, *or* large β_1, employing the ECM after finding evidence of a unit root in Y_t and evidence of cointegration does not seem to significantly impede our ability to correctly identify long-run relationships. This is an important result as many time series studies focus exclusively on the spurious regression problem without evaluating whether the proposed method can identify true relationships in the data.[18]

In addition to identifying true relationships between X_t and Y_t, we also want to be able to correctly estimate the error correction parameter. As discussed above, when the autoregressive parameter of Y_t is less than one, the estimated error correction rate ($\widehat{\alpha}_1$) will be biased downward (implying stronger error correction) because the estimate includes the true error correction rate plus $(\rho - 1)$. The simulation results indicate that with our near–integrated data, when T is reasonably large this bias is not very severe. For example, when T=60, the bias on $\widehat{\alpha}_1$ ranges between –0.03 and –0.05. When T=100, the bias ranges from −0.01 to −0.03. When T=30, however, the bias is quite substantial. When T is small, scholars must remember that their estimate of the error correction rate will be biased downward, leading to incorrect inferences about the speed of error correction. It is possible, however, to estimate $\rho - 1$ and then subtract this from $\widehat{\alpha}_1$ to correct the bias. In Appendix 3 we discuss our specific recommendations for addressing the potential bias in $\widehat{\alpha}_1$.

[18]Given our interest in long memoried series that are not true unit roots, we were also interested in the ability of fractional co-integration techniques (e.g., Clarke and Lebo, 2003) to correctly identify the DGP. In all cases, fractional error correction models (FECMs) were much less likely than the single equation ECM to correctly identify true relationships. In each case, FECMs also produce less accurate estimates of α_1. We report these results in the Appendix 3 and Appendix 4.

Table 2: Case 2 (β_1=0.3). Rejection Frequencies (%) for the False Null Hypothesis and Estimates for α_1

Y α_1	$X_{1,t}$ q	$X_{2,t}$ ξ	No Test (1)	EC$_M$ (2)	Means of $\widehat{\alpha}_1$ (3)
				T=30	
-0.1	0.9	0.9	50.94	5.59	-0.267
-0.1	0.9	1.0	49.22	6.17	-0.230
-0.1	1.0	0.9	59.29	8.89	-0.234
-0.1	1.0	1.0	53.16	8.57	-0.223
-0.4	0.9	0.9	55.04	30.41	-0.595
-0.4	0.9	1.0	53.72	31.89	-0.599
-0.4	1.0	0.9	64.01	36.83	-0.599
-0.4	1.0	1.0	63.17	38.71	-0.598
				T=60	
-0.1	0.9	0.9	89.23	43.15	-0.142
-0.1	0.9	1.0	86.96	53.39	-0.130
-0.1	1.0	0.9	94.77	57.37	-0.131
-0.1	1.0	1.0	94.02	62.23	-0.125
-0.4	0.9	0.9	93.32	91.62	-0.443
-0.4	0.9	1.0	91.69	90.37	-0.451
-0.4	1.0	0.9	98.05	96.15	-0.454
-0.4	1.0	1.0	95.94	94.29	-0.452
				T=100	
-0.1	0.9	0.9	99.39	87.06	-0.116
-0.1	0.9	1.0	99.21	95.37	-0.111
-0.1	1.0	0.9	99.90	96.19	-0.111
-0.1	1.0	1.0	99.79	97.27	-0.109
-0.4	0.9	0.9	99.91	99.91	-0.415
-0.4	0.9	1.0	99.29	99.29	-0.430
-0.4	1.0	0.9	100.00	100.00	-0.427
-0.4	1.0	1.0	99.89	99.89	-0.430

Notes: EC$_M$ uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). α_1 is the error correction as denoted in Equation 12. In No Test, we do not conduct pretests for cointegration or integration in Y_t. Column 3 reports the means of estimated $\widehat{\alpha}_1$ coefficients for series where we find evidence of both integration and cointegration.

Case 3: Y_t is (near-)integrated and (near-)cointegrated with just one of the X's

In Case 3, we introduce cointegration between Y_t and $X_{2,t}$. $X_{1,t}$, by contrast, is not related to Y_t. The DGP for Y_t is now re-defined as follows:

$$Y_t = Y_{t-1} + \alpha_1(Y_{t-1} + \frac{\beta_2}{\alpha_1} X_{2,t-1}) + u_{4,t}, \text{where } u_{4,t} \sim iid\,\mathrm{N}(0,1) \tag{13}$$

where the DGPs for $X_{1,t}$ and $X_{2,t}$ remain the same as the previous simulations. Our interest lies in finding how the presence of this (near-)cointegrating relationship between Y_t and $X_{2,t}$ may affect the frequency of finding a spurious relationship between Y_t and $X_{1,t}$. This is an especially important case, because researchers may find that not all right hand side variables are cointegrated with the dependent series. We propose the ECM is appropriate in this scenario because as long as at least one right hand side variable is cointegrated with Y_t the equation is balanced. To see this, we can re-write the right hand side of Equation 13 as the following cointegrating vector, $\alpha_1(Y_{t-1} + \frac{\beta_1}{\alpha_1}X_{1,t-1} + \frac{\beta_2}{\alpha_1}X_{2,t-1})$. Because β_1=0 in the DGP, even if $X_{1,t}$ is (near-)integrated, the linear combination of Y_t $X_{t,1}$, and $X_{t,2}$ would still be expected to be stationary.

We again set α_1 to range from -0.1 and -0.4, at increments of 0.1, and β_2=0.3. The results in Table 3 answer the question of whether ECMs produce spurious regression results in a multivariate system in which one explanatory variable is cointegrated with the dependent variable but the other is not. (Table A-6 in the Appendix reports the results from the full set of simulations.) Even with no pretest on Y_t or for cointegration, false rejection rates are typically less than 10 percent. We observe more than the expected 5 percent false rejection rate, but these percentages represent a much less severe bias toward incorrectly inferring a relationship between two unrelated variables than in previous simulations. These relatively low rates of spurious correlations between $X_{1,t}$ and Y_t result because $X_{2,t}$ and Y_t *are* cointegrated. In other words, although we do not conduct pretests in Column 1, we know from the DGP that $X_{2,t}$ and Y *are* cointegrated. Furthermore, as described above, when a cointegrating relationship exists between one predictor and the dependent variable, the equation is no longer unbalanced, even if other predictors not related to Y_t are (near-)integrated. Of course, given the stochastic component, $u_{4,t}$ in Equation 13, the DGP does not always produce a cointegrating relationship—especially when T is small. Thus, testing for integration in Y_t and then conducting the EC test with correct critical values (Column 2) offers a notable improvement (especially when T=30) over Column 1. The rate of spurious regression is roughly 5 percent if the single-equation ECM is combined with appropriate tests for cointegration and integration in Y_t.

Table 3: Case 3 (β_1=0, β_2=0.3). Rejection Frequencies (%) for the True Null Hypothesis (No Long Term Effect of $X_{1,t}$ on Y_t), for the False Null Hypothesis (Long Term Effect of $X_{2,t}$ on Y_t), and Estimates for α_1

			β_1=0 (Correct Value=5%)		β_2=0.3 (Correct Value=95%)		
Y	$X_{1,t}$	$X_{2,t}$	No Test	EC_M	No Test	EC_M	Means of $\widehat{\alpha}_1$
α_1	q	ξ	(1)	(2)	(3)	(4)	(5)
				T=30			
-0.1	0.9	0.9	8.35	1.60	55.84	5.21	-0.350
-0.1	0.9	1.0	7.25	1.65	59.60	6.76	-0.299
-0.1	1.0	0.9	10.30	2.30	52.48	4.79	-0.397
-0.1	1.0	1.0	7.75	2.00	58.06	6.60	-0.300
-0.4	0.9	0.9	7.85	4.30	60.69	27.00	-0.661
-0.4	0.9	1.0	7.45	4.60	70.54	32.83	-0.659
-0.4	1.0	0.9	8.05	4.30	61.89	28.45	-0.654
-0.4	1.0	1.0	8.80	5.85	66.92	30.63	-0.669
				T=60			
-0.1	0.9	0.9	6.50	3.05	92.36	24.95	-0.183
-0.1	0.9	1.0	5.80	3.90	96.35	46.27	-0.143
-0.1	1.0	0.9	7.90	3.70	91.39	23.17	-0.189
-0.1	1.0	1.0	8.35	5.45	96.76	43.03	-0.152
-0.4	0.9	0.9	7.80	4.90	96.47	89.95	-0.450
-0.4	0.9	1.0	5.90	4.55	98.70	93.70	-0.469
-0.4	1.0	0.9	7.30	3.70	94.59	88.72	-0.454
-0.4	1.0	1.0	7.50	6.10	98.04	92.48	-0.473
				T=100			
-0.1	0.9	0.9	6.05	4.50	99.53	64.33	-0.132
-0.1	0.9	1.0	6.10	5.50	99.95	91.44	-0.115
-0.1	1.0	0.9	7.45	6.20	99.38	63.15	-0.137
-0.1	1.0	1.0	5.50	4.85	99.84	88.65	-0.117
-0.4	0.9	0.9	5.90	1.70	100.00	100.00	-0.413
-0.4	0.9	1.0	5.95	4.65	99.87	99.87	-0.434
-0.4	1.0	0.9	6.30	1.35	100.00	100.00	-0.408
-0.4	1.0	1.0	6.10	5.05	99.94	99.94	-0.438

Notes: EC_M uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). q and ξ correspond to the memories of $X_{1,t}$ and $X_{2,t}$, respectively. α_1 is the error correction as denoted in Equation 13. In No Test, we do not conduct pretests for cointegration or integration in Y_t. Column 5 reports the means of estimated $\widehat{\alpha}_1$ coefficients for series where we find evidence of both integration and cointegration.

Columns 3 and 4 report our ability to correctly identify the true relationship between $X_{2,t}$ and Y_t (β_2=0.3). (See Table A-7 for the results from the full set of simulations.) Again, we see that conducting the additional steps of testing for cointegration and integration in

Y_t reduce the rate of recovering a relationship between $X_{2,t}$ and Y_t. If T is close to 30, researchers must be aware that they will only be able to detect long term relationships between X_t and Y_t when this relationship is strong or when the rate of error correction is high. This problem is much less of a concern when $\alpha_1 = -0.4$ and T is greater than 60. As long as we have a moderate level of error correction and a sufficiently large T (in this case, $T \geq 60$), the ECM can identify the relationship between $X_{2,t}$ and Y_t roughly 90% of the time or more. When T=100, the ECM identifies approximately 95% of true relationships when $\alpha_1 = -0.2$ (see Table A-7).

Column 5 reports the mean estimate of $\widehat{\alpha}_1$. When $T = 100$ the bias is quite small. The bias is more severe when $T \leq 60$. Importantly, correct rates of spurious regression in Column 2 confirm that the bias in $\widehat{\alpha}_1$ does not reduce the accuracy of the EC test for cointegration. Nevertheless, scholars need to be aware that their estimate of the rate of error correction will be biased, especially with small samples. As noted above, in Appendix 3 we propose a correction which yields more conservative estimates of the true error correction rate.

In sum, we have found that when the correct pretests are conducted, the ECM avoids the spurious regression problem. When T is small, however, scholars must be aware of the low power to identify true relationships and bias in $\widehat{\alpha}_1$. Below, we reconsider two prominent political science articles in order to show how to apply the above lessons.

Public Opinion and Supreme Court Decisions

Political scientists and legal scholars have long been interested in the relationship between the public's policy preferences and Supreme Court decisions in the United States. (e.g., Epstein and Martin, 2011; Enns and Wohlfarth, 2013; Giles, Blackstone and Vining, 2008; McGuire and Stimson, 2004; Mishler and Sheehan, 1993; Norpoth and Segal, 1994). Perhaps the strongest statement on this subject comes from Casillas, Enns and Wohlfarth (2011). Based on a series of single equation ECMs, they conclude that, "the public mood directly constrains the justices' behavior and the Court's policy outcomes, even after controlling for the social forces that influence the public *and* the Supreme Court" (86). Casillas, Enns and Wohlfarth (CEW) argue that because cases that are salient to the public tend to be the most politically and legally important, these are the cases where justices may prioritize their ideology over the public's preferences. In nonsalient cases, by contrast, they propose that the justices face an incentive to avoid decisions that draw public scrutiny in order to avoid squandering public support on less influential or less politically important cases. Below, we review CEW's analysis and find strong support for their key claims. We also show, however, that their estimation and interpretation of the error correction rates should be revised to somewhat temper their conclusions about how fast Court decisions reflect shifts in the public's preferences.

The dependent variable in CEW's first analysis is the percent of liberal decisions that reversed the lower court's ruling in each term. CEW do not report the time series properties of this variable, so we begin with an augmented Dickey Fuller test, which indicates that we accept the null hypothesis that the dependent variable has a unit root (p=0.63). Clearly the percent of liberal reversals is bound between 0 and 100, indicating that the series cannot have an infinite variance and thus is not integrated of order one. However, the observed data behave as if they are (near-)integrated, so our next step is to look for evidence of cointegration. We compare the t-statistic on the estimated error correction parameter ($-\frac{0.83}{0.15} = -5.53$) to the corresponding critical values in MacKinnon (2010) (-4.35) and find evidence of cointegration.

The evidence that the dependent variable is (near-)integrated and that (near-)cointegration is present suggests that the ECM was the appropriate model and we have a high degree of confidence in the estimated relationships between the the public's preferences and Supreme Court outcomes. However, given that their analysis only includes 45 observations, we have reason to believe that the error correction rate is biased. As explained previously, this bias does not pose a problem for the cointegration test, but it does affect our interpretation of the magnitude of the error correction rate. To correct the estimated error correction rate, we regress the percent liberal reversals on its lagged value. This yields an estimate of the autoregressive parameter of ρ=0.823. We subtract $(\rho - 1)$, from the estimated error correction parameter, which yields a revised estimate of −0.65. This value still implies a strong error correction rate, but it is more than 20 percent lower than the previous estimate. We would recommend reporting the *unadjusted* error correction rate with its standard error when referring to cointegration tests, but when discussing how fast the estimated long term effect is expected to take place, we would rely on the *adjusted* value.

CEW's final analysis did *not* find a statistically significant relationship between the public's preferences and salient Supreme Court decisions. The error correction rate of −1.27 they obtained (Column 2 of their Table 2) suggests that the dependent variable in this analysis may approach stationarity. Indeed, a DF test rejects the null of a unit root (p=0.002). Thus, the magnitude of this estimate stems from the fact that the dependent variable behaves as if stationary, not because of cointegration. CEW should not have interpreted their results as if cointegration and long run equilibrium relationships were present. Interestingly, their statistical estimates are not wrong—only their use of the language of long run relationships needs to be corrected.

To understand why their results are correct, recall that we have noted throughout that the ECM *is* appropriate with stationary data because the ADL and ECM are equivalent. Table 4 illustrates this point. Column 1 perfectly replicates CEW's salient case analysis. The dependent variable is the change in the percentage of salient liberal decisions issued by the Supreme Court during each term that reversed the lower courts' ruling from 1956 to 2000. Public mood comes from Stimson (1999) and is a measure of public support for liberal policy outcomes. Court Ideology reflects the Segal-Cover (1989) ideological scores for the median justice on the Court. These data are based on a content analysis of editorials about the justices during their nomination process. Social Forces (IV) controls for the social forces that may simultaneously influence both the Supreme Court justices and the mass public.

To generate this variable, CEW use an Instrumental Variable (IV) approach that regresses a measure of the Court's current ideological disposition (based on the Martin-Quinn scores (Martin and Quinn, 2002)) on all factors thought to influence the pubic's policy mood. This regression is then used to generate predicted values of the Martin-Quinn scores, accounting for the variation in the justices' ideology that is associated with the social forces that influence the public's policy mood.

Column 2 estimates a General ADL model that includes both contemporaneous and lagged effects. Because the ADL does not difference the dependent variable, estimating the ADL is equivalent to adding the lagged value of the dependent variable (Percent Liberal_{t-1}) to both sides of the ECM. Thus, the coefficient on Percent Liberal_{t-1} in the ADL equals -0.27 instead of -1.27 in column 1. The estimates for Public Mood, Court Ideology, and Social Forces (IV) at time t are identical to the estimates for these variables when differenced in the ECM. The differences between Columns 1 and 2 at $t-1$ may come as a surprise, but these differences exist because the coefficients represent different quantities. In Column 1, the coefficients on the lagged predictors represent the combined effect of X_t and X_{t-1} in the ADL. For example, if we add the estimates in Column 2 for Public Mood_t and Public Mood_{t-1} we get an estimate of $1.24 + -0.53 = 0.71$ with an estimated standard error of $\sqrt{1.81^2 + 1.46^2 + 2 \times -2.24} = 0.96$ (where -2.24 is the estimated covariance between Public Mood_t and Public Mood_{t-1}). This is the exact result produced by the ECM. CEW were wrong to imply that equilibrium error correction was present, but their estimates were identical to those they would have obtained if they estimated an ADL, which produces accurate inferences with stationary data—as is the case here (Keele and Kelly, 2006).

Although our interest is in the ECM, instead of assuming the dependent variable is $I(0)$ and estimating the ECM or ADL, we also replicated CEW's analysis using fractional differencing methods. We first estimate an autoregressive fractionally integrated moving average (arfima) model to estimate the fractional differencing parameter d for each variable. We then

Table 4: Replication of Casillas, Enns, and Wohlfarth Table 2, Column 2: The Determinants of Salient Supreme Court Decisions, 1956-2000

	ECM (Replication)	ADL	Fractional Differencing
Percent Liberal_{t-1}	-1.27*	-0.27	–
	(0.15)	(0.15)	
Δ Public Mood	1.24	–	–
	(1.81)		
Δ Court Ideology	10.47	–	–
	(10.03)		
Δ Social Forces	7.49	–	–
(Instrumented MQ Scores)	(8.03)		
Public Mood_t	–	1.24	0.97
		(1.81)	(2.35)
Court Ideology_t	–	10.47	12.80
		(10.03)	(12.11)
Social Forces $(\text{IV})_t$	–	7.49	8.09
(Instrumented MQ Scores)		(8.03)	(11.55)
Public Mood_{t-1}	0.71	-0.53	-0.39
	(0.96)	(1.46)	(1.82)
Court Ideology_{t-1}	16.90*	6.42	-6.50
	(5.73)	(9.84)	(11.62)
Social Forces $(\text{IV})_{t-1}$	9.20*	1.72	-3.80
(Instrumented MQ Scores)	(5.09)	(8.51)	(11.75)
Constant	31.02	31.02	25.72
	(58.32)	(58.32)	(80.60)
R^2	0.64	0.63	0.16
N	45	45	45

The dependent variable represents the change in the percentage of salient liberal decisions issued by the Supreme Court during each term, among all reversals. $* = p < .05$ (one–tailed tests), standard errors in parentheses. All variables in Column 3 have been fractionally differenced.

use this estimate to fractionally difference each series.[19] Column 3 presents these results. The standard error of each estimate is larger than column 2. Furthermore, none of the coefficients nor the total estimated effects (i.e., the coefficient at t plus $t-1$) are statistically different from zero. Given the prominence of the attitudinal model (Segal and Spaeth, 1993), which

[19] We began with a model that assumed no autoregressive terms (ρ=0) and no moving average terms (q=0). We added an autoregressive or moving average term if doing so significantly improved model fit.

demonstrates that justices' votes and case outcomes are heavily influenced by the justices' ideological predispositions, the imprecise estimates for the relationship between the ideology of the Court and its salient decisions is surprising. Fractional differencing may be underpowered to observe such relationships when T is small. Consistent with this conclusion, Lebo and Weber (2014, 4) explain that, "estimates of d are less reliable as T drops."[20]

The 1934 RTAA and Tariff Reductions

Our second example stems from the debate on the relationship between the 1934 Reciprocal Trade Agreement Act (RTAA) and subsequent reductions in tariffs in the United States. The RTAA shifted authority over trade policy from Congress to the president, and as Hiscox (1999, 669) explains, this shift seemingly "altered the nature of the policymaking process and drastically changed the future course of U.S. trade relations." Although scholars universally agree that trade policy changed dramatically following the RTAA, the precise relationship between the RTAA and these subsequent changes has been highly debated (e.g., Bailey, Goldstein and Weingast, 1997; Ehrlich, 2008; Ehrlich, 2011; Hiscox, 1999; Hiscox, 2002; Lohmann and O'Halloran, 1994).

Our focus is on important recent work by Ehrlich (2008; 2011). Ehrlich proposes that, relative to supporters of free trade, those in favor of protectionist trade policies are better able to overcome the collective action problem because the benefits of protection are concentrated on the industries receiving trade protection and the costs are dispersed across consumers. Because of their ability to better overcome the collective action problem, Ehrlich argues that protectionist interests will lobby more than free trade interests and that this lobbying advantage grows when there are more access points to the political system. Because the RTAA reduced the number of access points by concentrating trade policy under the president, Ehrlich hypothesizes that the RTAA reduced the lobbying advantage of protectionist

[20] Lebo and Weber (2014) suggest $T > 50$ as a rule of thumb for reliable estimates of d.

interests, leading to trade policy more favorable to free trade interests.

Ehrlich's work offers several tests of this argument, both cross-nationally and within the United States (Ehrlich, 2007; Ehrlich, 2008; Ehrlich, 2011). We focus on one important test that estimates a single equation ECM (Ehrlich, 2008, Model 1 of Table 1). The dependent variable is the annual rate of tariff revenue from 1902 to 1988. This variable is calculated by dividing tariff revenues collected by the government each year by the total value of imports into the United States that year. An augmented DF test indicates that we cannot reject that null hypothesis of a unit root (p=0.22) in this series. Ehrlich estimates an ECM in order to evaluate the long-term relationship between the RTAA and tariff rates, where RTAA is a dichotomous variable coded as zero until 1933 and 1 after. As Ehrlich explains, "the effect of the RTAA (and the delegation it ushered in) should be in the long term as treaties are slowly negotiated and renegotiated with individual countries and tariff levels gradually reduce over time. In other words, the RTAA led to a new, lower equilibrium tariff level" (438-439). However, the EC statistic (i.e., the t-statistic associated with the lagged dependent variable in the ECM) offers no evidence of cointegration ($t = -1.16$), which means the ECM should *not* have been estimated. It is important to emphasize that by not testing for cointegration, Ehrlich was simply following accepted practice in the discipline. As we noted above, in recent years fewer than 15 percent of political science articles have reported tests for cointegration when estimating an ECM. In fact, Ehrlich justified his decision, noting, "as Beck (1991) argues and DeBoef and Keele (2005) prove, ECMs can be used even when unit roots or cointegration are not present" (438).[21]

One goal of this article is to clarify this line of reasoning. While the above statement can be true, when the dependent series behaves as if it has a unit root (as is the case here), cointegration must be established. Given the absence of cointegration, long memory series on both sides of the equation should be differenced. One option would be to treat all long-memoried series as if they are $I(1)$ and first difference these series. Alternatively, long

[21] De Boef and Keele (2005) refers to an earlier version of De Boef and Keele (2008).

memoried series could be fractionally differenced (as we did in Column 3 in the previous example). However, by first-differencing (or fractionally differencing) all long-memoried variables in the model, the analysis would no longer estimate a potential long-run effect of the RTAA on tariff rates. Thus, to estimate the hypothesized long run effect, we add an additional variable to the model. This variable is coded as 0 until 1933 and then the variable values increase following a logistic function, initially increasing slowly, and then rising until the values approach 1. This functional form is designed to capture Ehrlich's prediction that "treaties are slowly negotiated and renegotiated with individual countries and tariff levels gradually reduce over time" (438-439). Figure 1 presents the tariff rate, a dichotomous indicator which would capture an average shift in tariff reduction following the RTAA, and our proposed variable designed to capture the hypothesized relationship between the RTAA and tariffs. While we believe our long term RTAA variable captures the theoretical predictions advanced by Erhlich, our primary aim is not hypothesis testing. Rather, our intent is to illustrate how the hypothesized long term effects could be estimated with a differences model.

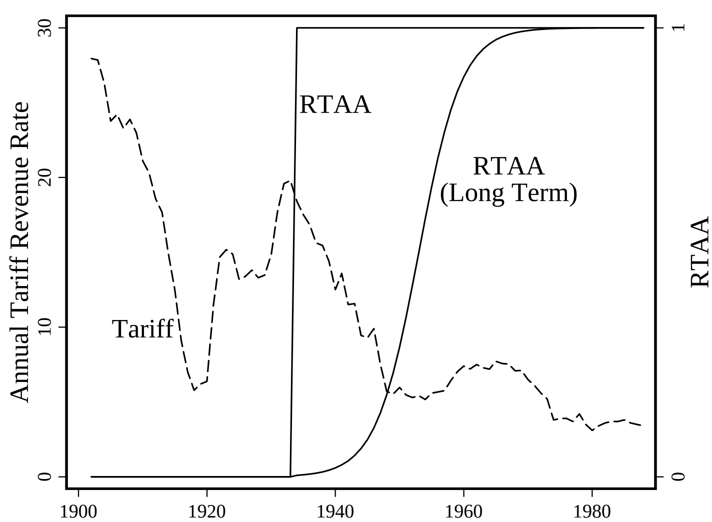

Figure 1: The Rate of Tariff Revenue, the onset of the RTAA, and an Approximation of the Hypothesized Long Term Effect of the RTAA, 1902 to 1988

Figure 2 presents the estimated relationship between the RTAA and tariff rates from three statistical analyses (the full results appear in Appendix 5). The top estimate is based on our replication of Ehrlich (2008, Table 1, Model 1). Ehrlich divides this estimate by the coefficient on the lagged dependent variable and concludes, "The long-term effect of the RTAA is a predicted 17 percent decrease in tariff rates" (439). We have two concerns with this conclusion. First, because there is no evidence of cointegration, this long run multiplier should not be estimated. Second, the absence of cointegration combined with the (near-) integrated dependent variable indicates that the equation is unbalanced and we cannot rely on standard t-statistics. The next two estimates come from the first differences model. The negative and significant coefficient for RTAA (1934) indicates that the onset of the RTAA did correspond with an average decrease in the change in tariff rates. Contrary to expectations, however, the estimate for the long run effect of the RTAA is positive and significant. The final two estimates come from a model which first fractionally differences the variables.[22] We again see a negative and significant estimate for the onset of the RTAA. The estimate for the long term effect is now also negative, but the coefficient is imprecisely estimated.

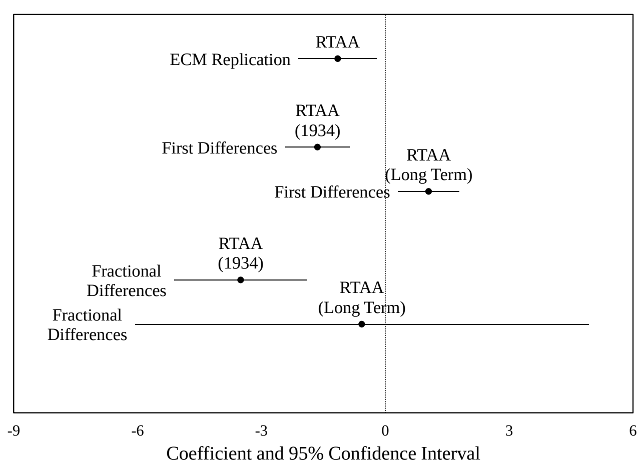

Figure 2: The Expected Average and Long Term Relationship between the RTAA and Tariff Rates (1902 to 1988), Based on Three Separate Analyses

Our inability to find evidence of a long term relationship between the RTAA and tariff rates should be taken lightly. First, as noted above, Ehrlich has provided substantial ad-

[22]We did not fractionally difference dichotomous variables or the variable capturing the hypothesized long term RTAA effect.

ditional support for his access point theory (Ehrlich, 2007; Ehrlich, 2008; Ehrlich, 2011). Additionally, our primary goal was to show how long term relationships *might* be estimated with a differences model. It is entirely possible that an alternate functional form could be specified to better approximate Ehrlich's predictions. Another estimation strategy would be to directly model how the RTAA is hypothesized to influence tariff rates. For example, Ehrlich expects the RTAA to influence tariff rates by decreasing the amount of lobbying by protectionist interests. Although annual lobbying data may not be available (Hiscox, 2002, 143), such a measure might show evidence of cointegration. Given the importance of the debate regarding the RTAA and trade policy, we hope those engaged in this literature will continue this line of research.

Conclusions and Implications

Error correction models continue to be a theoretically important and widely applied methodology in political science (e.g., Beck and Katz, 2011; Blaydes and Kayser, 2011; Jennings, 2013; Layman, Carsey, Green, Herrera and Cooperman, 2010). While scholars have correctly concluded that ECMs can be applicable to both stationary and non-stationary data, there has been a great degree of confusion over how to correctly apply ECMs to near-integrated and integrated data. De Boef and Keele (2008, 195) explain, "the only situation where one would strongly prefer the ECM [relative to the ADL] is if the data are strongly autoregressive." This statement is not wrong, but we have added important caveats. A more precise statement would be, one would strongly prefer the ECM if the data are strongly autoregressive *and near-cointegrated or integrated and cointegrated.*[23] This clarification offers an important change to many recent applications of the ECM. In fact, the majority of recent applications of the ECM in the top political science journals fail to test for cointegration and many of those that do test for cointegration use the incorrect critical values.

[23]De Boef and Keele (2008, 195) also note that "If the data are integrated, an alternate form of the ECM must be estimated so that no regressors are integrated." Again, we would simply clarify that the ECM produces correct inferences with integrated data *if* the series are cointegrated.

The cost of failing to test the presence of cointegration can be high. Based on our simulation results, however, we find that conducting a pretest to determine if Y_t behaves as if it contains a unit root and then testing for cointegration (with the correct critical values) avoids the spurious regression problem. The proposed pretests do come with some cost, leading to excessive Type II errors. Yet, provided the error correction rate is moderately strong (i.e., $\leq$-0.4), the sample size is ≥ 60, *or* the long run relationship is moderately strong (i.e., $\beta > 0.3$), the ECM correctly identifies true relationships in the DGP. These results hold in the bivariate and multivariate context (See Appendix 6 for additional multivariate simulations). Thus, we recommend that political scientists return to earlier practices (see, e.g., Beck, 1992, 238) and ensure that Y_t behaves like a unit root and cointegration is present prior to estimating an ECM with long–memoried dependent series.

The phrase, "behaves like a unit root" is, of course, quite intentional. Unlike many economic time series, in political science we often do *not* have theoretical reasons to believe our series contain a true unit root. Theoretically and empirically, many series are, however, long–memoried. We also know that when data are near-integrated they mimic the behavior of a unit root when T is small (e.g., Banerjee et al., 1993, 96). We have argued that for statistical analysis, it is the sample properties—not the asymptotic properties—of variables that matter. Indeed, the ECM performed well in all of our simulations with near-integrated data, as long as a DF test did not reject the null hypothesis of a unit root.

One potential concern of estimating an ECM with near-integrated data is that the error correction rate, while still providing an accurate test of cointegration, will be biased in a negative direction. Our simulation results demonstrate that as T increases, this bias diminishes. When T is small or of moderate size (e.g., T$\leq$60), we propose a correction, by estimating the autoregressive component of Y_t (i.e., ρ) and then subtracting (ρ-1) from the error correction parameter. We also analyzed what we believe to be an under-studied—but empirically common—scenario: when one predictor is cointegrated with the dependent variable but another predictor is not. We developed the intuition for why this scenario should *not* pose a problem for spurious regression. Consistent with expectations, even if just one predictor is cointegrated with the dependent variable, the ECM can be applied to data where Y_t behaves as if a unit root. This result helps illustrate the importance of the ECM for a variety of time series applications in political science.

We also wish to emphasize the limits of these findings. First, we have focused on near-integrated and integrated series, where the autoregressive parameter in the DGP ranges from 0.90 to 1.0. While the ECM performs quite well with these series when (near-)cointegration is present, many political time series might be better characterized as fractionally integrated. For such series, fractional integration techniques might be more appropriate (e.g., Clarke and Lebo, 2003; Lebo and Clarke, 2000; Lebo, Walker and Clarke, 2000). Although our replication of Casillas, Enns and Wohlfarth (2011) raised some questions about the small sample properties of fractional differencing methods (see also, Lebo and Weber, 2014), it is important to emphasize that the ECM is not appropriate with all types of data. Second, we have focused on the case of one cointegrating vector. It is possible, however, that more than one cointegrating vector exists between variables. When Johansen tests provide evidence of more than one cointegrating vector, a system of equations, such as a VECM model, should be estimated. Although we have demonstrated the appropriateness of the single equation ECM when relevant conditions hold—and these conditions appear to hold both theoretically and empirically in many political science applications—VECM and fractional integration techniques should be utilized when necessary.

References

Ahlquist, John S. 2006. "Economic Policy, Institutions, and Capital Flows: Portfolio and Direct Investment Flows in Developing Countries." *International Studies Quarterly* 50(3):681–704.

Bailey, Michael A., Judith Goldstein and Barry R. Weingast. 1997. "The Institutional Roots of American Trade Policy." *World Politics* 49(3):309–338.

Baille, Richard T. 1996. "Long Memory Processes and Fractional Integration in Econometrics." *Journal of Econometrics* 73(1):5–59.

Banerjee, Anindya. 1995. Dynamic Specification and Testing for Unit Roots and Cointegration. In *Macroeconometrics*, ed. Kevin D. Hoover. Netherlands: Springer chapter 12, pp. 473–516.

Banerjee, Anindya, Juan Dolado, John W. Galbraith and David F. Hendry. 1993. *Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.

Banerjee, Anindya, Juan J. Dolado and Ricardo Mestre. 1998. "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework." *Journal of Time Series Analysis* 19(3):267–283.

Beck, Nathaniel. 1991. "Comparing Dynamic Specifications: The Case of Presidential Approval." *Political Analysis* 3:51–87.

Beck, Nathaniel. 1992. "The Methodology of Cointegration." *Political Analysis* 4(1):237–248.

Beck, Nathaniel and Jonathan N. Katz. 2011. "Modeling Dynamics in Time-Series–Cross-Section Political Economy Data." *Annual Review of Political Science* 14:331–352.

Blaydes, Lisa and Mark Andreas Kayser. 2011. "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55(4):897–908.

Blough, Stephen R. 1992. "The Relationship Between Power and Level for Generic Unit Root Tests in Finite Samples." *Journal of Applied Econometrics* 7:295–308.

Boswijk, H. Peter and Jean-Pierre Urbain. 1997. "Lagrange-Multiplier Tests For Weak Exogeneity: A Synthesis." *Econometric Reviews* 16(1):21–38.

Box-Steffensmeier, Janet M. and Andrew R. Tomlinsonl. 2000. "Fractional Integration Methods in Political Science." *Electoral Studies* 19(1):63–76.

Box-Steffensmeier, Janet M. and Renee M. Smith. 1998. "Investigating Political Dynamics Using Fractional Integration Methods." *American Journal of Political Science* 42(2):661–689.

Casillas, Christopher J., Peter K. Enns and Patrick C. Wohlfarth. 2011. "How Public Opinion Constrains the U.S. Supreme Court." *American Journal of Political Science* 55(1):74–88.

Clarke, Harold D. and Matthew Lebo. 2003. "Fractional (Co)integration and Governing Party Support in Britain." *British Journal of Political Science* 33(2):283–301.

Cochrane, John H. 1992. "A Critique of the Application of Unit Root Tests." *Journal of Economic Dynamics and Control* 15:275–284.

Davidson, James E. H., David F. Hendry, Frank Srba and Steven Yeo. 1978. "Econometric Modelling of the Aggregate Time-series Relationship BetweenConsumer's Expenditure and Income in the United Kingdom." *Economic Journal* 88:661–692.

De Boef, Suzanna. 2001. "Modeling Equilibrium Relationships: Error Correction Models with Strongly Autoregressive Data." *Political Analysis* 9(1).

De Boef, Suzanna and Jim Granato. 1997. "Near-integrated Data and the Analysis of Political Relationships." *American Journal of Political Science* 41(2):619–640.

De Boef, Suzanna and Jim Granato. 2000. "Testing for Cointegrating Relationships with Near-Integrated Data." *Political Analysis* 8(1):99–117.

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

Duffield, John S. and Charles R. Hankla. 2011. "The Efficiency of Institutions: Political Determinants of Oil Consumption in Democracies." *Comparative Politics* 43(2):187–

205.

Durr, Robert H. 1992*a*. "An Essay on Cointegration and Error Correction Models." *Political Analysis* 4:185–228.

Durr, Robert H. 1992*b*. "Of Forests and Trees." *Political Analysis* 4:255–258.

Ehrlich, Sean D. 2007. "Access to Protection: Domestic Institutions and Trade Policy in Democracies." *International Organization* 61(3):571–605.

Ehrlich, Sean D. 2008. "The Tariff and the Lobbyist: Political Institutions, Interest Group Politics, and U.S. Trade Policy." *International Studies Quarterly* 52(2):427–445.

Ehrlich, Sean D. 2011. *Access Points*. New York: Oxford University Press.

Enders, Walter. 2004. *Applied Econometric Time Series*. 2 ed. John Wiley & Sons.

Engle, Robert F. and C. W. J. Granger. 1987. "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55:251–276.

Engle, Robert F., David F. Hendry and Jean-Francois Richard. 1983. "Exogeneity." *Econometrica* 51:277–304.

Enns, Peter K. and Gregory McAvoy. 2012. "The Role of Partisanship in Aggregate Opinion." *Political Behavior* 34(4):627–651.

Enns, Peter K. and Patrick C. Wohlfarth. 2013. "The Swing Justice." *Journal of Politics* 75(4).

Epstein, Lee and Andrew D. Martin. 2011. "Does Public Opinion Influence the Supreme Court? Possibly Yes (But We're Not Sure Why)." *University of Pennsylvania Journal of Constitutional Law* 13:263–281.

Ericsson, Neil R. and James G. MacKinnon. 2002. "Distributions of Error Correction Tests for Cointegration." *Econometrics Journal* 5(2):285–318.

Faricy, Christopher. 2011. "The Politics of Social Policy in America: The Causes and Effects of Indirect versus Direct Social Spending." *Journal of Politics* 73(1):74–83.

Giles, Micheal W., Bethany Blackstone and Richard L. Vining. 2008. "The Supreme Court in American Democracy: Unraveling the Linkages between Public Opinion and Judicial

Decision Making." *Journal of Politics* 70(2):293–306.

Granger, Clive W.J. and Paul Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics* 26:1045–1066.

Granger, C.W.J. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14(2):227–238.

Granger, C.W.J. 1986. "Developments in the Study of Cointegrated Economic Variables." *Oxford Bulletin of Economics and Statistics* 48:213–228.

Hamilton, James D. 1994. *Time Series Analysis.* Princeton, NJ: Princeton University Press.

Hellwig, Timothy. 2007. "Globalization and Perceptions of Policy Maker Competence." *Political Research Quarterly* 60:146–158.

Hendry, David F. 2008. Equilibrium-Correction Models. In *The New Palgrave Dictionary of Economics*, ed. Steven N. Durlauf and Lawrence E. Blume. 2 ed. Basingstoke, England: Palgrave Macmillan pp. 1–15.

Hiscox, Michael J. 1999. "The Magic Bullet? The RTAA, Institutional Reform, and Trade Liberalization." *International Organization* 53(4):669–698.

Hiscox, Michael J. 2002. *International Trade and Political Conflict: Commerce, Coalitions, and Mobility.* Princeton: Princeton University Press.

Iversen, Torben and Anne Wren. 1998. "Equality, Employment, and Budgetary Restraint: The Trilemma of the Service Economy." *World Politics* 50(4):507–546.

Jennings, Will. 2009. "The Public Thermostat, Political Responsiveness and Error-Correction: Border Control and Asylum in Britan, 1994-2007." *British Journal of Political Science* 39(4):847–870.

Jennings, Will. 2013. Error-Correction as a Concept and as a Method: Time Series Analysis of Policy-Opinion Responsiveness. In *Political Science Research Methods in Action*, ed. Michael Bruter and Martin Lodge. New York: Palgrave Macmillan chapter 10, pp. 203–228.

Johansen, Søren. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Jönsson, Kristian. 2005. "Using Panel Data to Increase the Power of Modified Unit Root Tests in the Presence of Structural Breaks." *Applied Mathematics and Computation* 171(2):832–842.

Kayser, Mark Andreas. 2009. "Partisan Waves: International Business Cycles and Electoral Choice." *American Journal of Political Science* 53:950–970.

Keele, Luke. 2007. "Social Capital and the Dynamics of Trust in Government." *American Journal of Political Science* 51(2):241–254.

Keele, Luke and Nathan J. Kelly. 2006. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14(2):186–205.

Kelly, Nathan J. 2005. "Political Choice, Public Policy, and Distributional Outcomes." *American Journal of Political Science* 49(4):865–880.

Kelly, Nathan J. and Peter K. Enns. 2010. "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences." *American Journal of Political Science* 54(4):855–870.

Kono, Daniel Y. 2007. "Making Anarchy Work: International Legal Institutions and Trade Cooperation." *Journal of Politics* 69:746–759.

Kremers, Jeroen J. M., Neil R. Ericsson and Juan J. Dolado. 1992. "The Power of Cointegration Tests." *Oxford Bulletin of Economics and Statistics* 54(3):325–348.

Larsson, Rolf, Johan Lyhagen and Mickael Löthgren. 2001. "Likelihood-based Cointegration Tests in Heterogenous Panels." *Econometrics Journal* 4(1):109–142.

Layman, Geoffrey C., Thomas M. Carsey, John C. Green, Richard Herrera and Rosalyn Cooperman. 2010. "Activists and Conflict Extension in American Politics." *American Political Science Review* 102(2):324–346.

Lebo, Matthew J. 2008. "Divided Government, Unified Approval: The Dynamics of Congressional and Presidential Approval." *Congress & the Presidency* 35(2):1–16.

Lebo, Matthew J. and Christopher Weber. 2014. "An Effective Approach to the Repeated Cross-Sectional Design." *American Journal of Political Science* pp. 1–17.

Lebo, Matthew J. and Harold D. Clarke. 2000. "Modelling memory and volatility: recent advances in the analysis of political time series." *Electoral Studies* 19(1):1–7.

Lebo, Matthew J., Robert W. Walker and Harold D. Clarke. 2000. "You Must Remember This: Dealing with Long Memory Political Analysis." *Electoral Studies* 19(1,).

Lohmann, Susanne and Sharyn O'Halloran. 1994. "Divided Government and U.S. Trade Policy: Theory and Evidence." *International Organization* 48(4):595–632.

MacKinnon, James G. 2010. "Critical Values for Cointegration Tests." University of California San Diego Discussion Paper 90-4.

Maddala, G.S. and In-Moo Kim. 1998. *Unit Roots, Cointegration, and Structural Change.* 6 ed. New York: Cambridge University Press.

Mankiw, N. Gregory and Matthew D. Shapiro. 1985. "Trends, Random Walks, and Tests of the Permanent Income Hypothesis." *Journal of Monetary Economics* 16(2):165–174.

Mankiw, N. Gregory and Matthew D. Shapiro. 1986. "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models." *Economics Letters* 20(2):139–145.

Marmol, Francesc. 1996. "Nonsense Regressions Between Integrated Processes of Different Orders." *Oxford Bulletin of Economics and Statistics* 58(3):525–536.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.

McGuire, Kevin T. 2004. "The Institutionalization of the U.S. Supreme Court." *Political Analysis* 12:128–142.

McGuire, Kevin T. and James A. Stimson. 2004. "The Least Dangerous Branch Revisited: New Evidence on Supreme Court Responsiveness to Public Preferences." *Journal of Politics* 66(4):1018–1035.

Mishler, William and Reginald S. Sheehan. 1993. "The Supreme Court as a Countermajoritarian Institution? The Impact of Public Opinion on Supreme Court Decisions." *American Political Science Review* 87(1):87–101.

Murray, Michael P. 1994. "A Drunk and Her Dog: An Illustration of Cointegration and Error Correction." *The American Statistician* 48(1):37–39.

Nooruddin, Irfan and Joel W. Simmons. 2009. "Openness, Uncertainty, and Social Spending: Implications for the Globalization-Welfare State Debate." *International Studies Quarterly* 53(3):841–866.

Norpoth, Helmut and Jeffrey A. Segal. 1994. "Comment: Popular Influence on Supreme Court Decisions." *American Political Science Review* 88(3):711–716.

Oatley, Thomas. 2012. "Political Institutions and Foreign Debt in the Developing World." *International Studies Quarterly* 54(1):175195.

Ostrom, Jr. Charles W. and Renée M. Smith. 1992. "Error Correction, Attitude Persistence, and Executive Rewards and Punishments: A Behavioral Theory of Presidential Approval." *Political Analysis* 4:127–184.

Phillips, Peter C.B. 1988. "Regression Theory for Near-Integrated Time Series." *Econometrica* 56(5):1021–1043.

Ramirez, Mark D. 2009. "The Dynamics of Partisan Conflict on Congressional Approval." *Journal of Politics* 53:681–694.

Segal, Jeffrey A. and Albert D. Cover. 1989. "Ideological Values and the Votes of U.S. Supreme Court Justices." *American Political Science Review* 83(2):557–565.

Segal, Jeffrey A. and Harold J. Spaeth. 1993. *The Supreme Court and the Attitudinal Model.* New York: Cambridge University Press.

Smith, Renée M. 1992. "Error Correction, Attractors, and Cointegration: Substantive and Methodological Issues." *Political Analysis* 4:249–254.

Soroka, Stuart N. and Christopher Wlezien. 2010. *Degrees of Democracy.* New York: Cambridge.

Stimson, James A. 1999. *Public Opinion in America: Moods, Cycles, and Swings*. 2nd ed. Boulder, CO: Westview Press.

Ura, Joseph Daniel and Patrick C. Wohlfarth. 2010. "'An Appeal to the People': Public Opinion and Congressional Support for the Supreme Court." *Journal of Politics* 74(4):939–956.

Ventosa-Santaulària. 2012. "Unbalanced Regressions and Spurious Inference." *Open Journal of Statistics* 2(3):297–299.

Voeten, Erik and Paul R. Brewer. 2006. "Public Opinion, the War in Iraq, and Presidential Accountability." *Journal of Conflict Resolution* 50:809–830.

Williams, John T. 1992. "What Goes Around Comes Around: Unit Root Tests and Cointegration." *Political Analysis* 4:229–236.

Wlezien, Christopher. 2004. "Patterns of Representation: Dynamics of Public Preferences and Policy." *Journal of Politics* 66(1):1–24.

Yule, G. Udny. 1926. "Why do we Sometimes get Nonsense-Correlations between Time-Series?–A Study in Sampling and the Nature of Time-Series." *Journal of the Royal Statistical Society* 89:1–63.

Supplementary/Online Appendix for:

Time Series Analysis and Spurious Regression: An Error Correction

Contents

Appendix 1 Bivariate Spurious Correlation Rates

The text reported rates of spurious regressions 3 to 4 times conventional levels of acceptability for the long run and total (LRM) effects of X_t on Y_t. The tables below report the full results from these simulations.

Specifically, Table A-1 reports the percent of rejections for the true null hypothesis ($\beta_1 = 0$) for X_t when T=30, 60, and 100. In addition to reporting the proportion of spurious regressions for no tests (Columns 1, 4, and 7) and for when we test for cointegration and the time series properties of Y_t (Columns 3, 6, and 9), we also report the results when standard critical values are used with the EC test (Columns 2, 5, and 8). The high rate of spurious relationships in these columns illustrates the importance of using the correct critical values with the EC test.

We are also interested in the *total* expected relationship between X_t and Y_t, often referred to as the Long Run Multiplier (LRM) (e.g., De Boef and Keele, 2008). Table A-2, below, shows that the problem of spurious correlations is *more substantial* when we examine the LRM.[24] We also find that we can obviate the spurious regression problem with the LRM if we first establish that Y_t behaves like a unit root and that (near-)cointegration exists between (near-)integrated series.

[24]We estimate a LRM effect of X_t on Y_t, which is denoted by $LRM = \frac{\beta_1}{\alpha_1}$, compute its variance, and perform a t-test to examine if $\widehat{LRM}$ is significant at the .05 level. The variance of $\widehat{LRM}$ is derived from the delta method and computed using the following formula (De Boef and Keele, 2008, 191):

$$(1/\alpha_1^2)Var(\beta_1) + (\beta_1^2/\alpha_1^4)Var(\alpha_1) - 2(\beta_1/\alpha_1^3)Cov(\alpha_1, \beta_1) \tag{14}$$

Table A-1: Rejection Frequencies (%) for the True Null Hypothesis: $\beta_1 = 0$

		T=30			T=60			T=100		
ρ	q	No Test (1)	EC_t (2)	EC_M (3)	No Test (4)	EC_t (5)	EC_M (6)	No Test (7)	EC_t (8)	EC_M (9)
0.90	0.90	13.25	8.90	2.25	9.95	7.70	2.60	9.05	6.05	2.30
0.95	0.90	12.95	8.70	2.20	11.35	7.65	1.85	9.70	7.15	1.65
0.99	0.90	11.75	7.80	1.55	10.55	6.40	1.00	9.45	5.65	0.95
1.00	0.90	11.65	7.20	1.80	10.60	6.55	1.20	10.10	5.60	0.85
0.90	0.95	12.85	9.60	2.45	10.50	7.95	2.45	9.10	6.05	2.45
0.95	0.95	14.85	10.45	2.55	12.25	8.95	2.05	11.40	8.35	2.75
0.99	0.95	14.15	8.75	1.70	13.75	9.60	1.80	12.65	7.75	1.35
1.00	0.95	13.15	8.80	1.55	12.95	8.55	1.90	11.50	7.75	1.15
0.90	0.99	13.75	10.55	2.65	12.45	10.60	3.85	10.95	7.25	3.65
0.95	0.99	13.70	9.90	2.95	14.10	10.75	2.20	12.60	10.15	3.45
0.99	0.99	14.60	10.05	2.35	14.60	10.15	2.25	17.25	12.50	2.40
1.00	0.99	15.00	9.90	1.95	15.55	10.60	2.10	15.05	9.95	2.40
0.90	1.00	13.25	10.60	3.20	12.70	10.65	4.05	9.85	7.00	3.70
0.95	1.00	14.45	10.50	2.45	16.00	12.65	3.85	12.75	10.70	3.75
0.99	1.00	15.45	11.10	2.95	16.70	12.25	2.40	16.30	11.90	2.25
1.00	1.00	14.90	10.10	1.95	16.65	11.55	2.40	17.00	11.95	2.25

Notes: EC_M uses the 5% MacKinnon critical values of -3.340, -3.272, and -3.248 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002), while EC_t employs a standard 5% t critical value. ρ and q refer to the memory of Y_t and X_t in Equation 2 and 3, respectively. In No Test, we do not conduct pretests for cointegration or integration in Y_t.

Table A-2: Rejection Frequencies (%) for the True Null Hypothesis: $LRM = 0$

		T=30			T=60			T=100		
ρ	q	No Test (1)	EC_t (2)	EC_M (3)	No Test (4)	EC_t (5)	EC_M (6)	No Test (7)	EC_t (8)	EC_M (9)
0.90	0.90	13.30	11.25	2.30	9.75	8.65	2.75	9.25	6.15	2.60
0.95	0.90	13.45	11.25	2.30	10.60	9.20	1.95	8.30	7.00	1.75
0.99	0.90	11.80	9.85	1.65	8.65	7.30	1.10	6.65	5.70	0.95
1.00	0.90	11.85	8.70	1.80	8.90	6.80	1.30	6.55	5.35	0.85
0.90	0.95	14.65	12.50	2.65	11.40	9.90	2.70	11.80	8.10	2.80
0.95	0.95	15.90	12.55	2.65	13.60	11.85	2.20	12.30	10.35	2.90
0.99	0.95	15.75	11.65	1.75	14.45	12.05	1.90	10.25	8.60	1.40
1.00	0.95	14.65	11.60	1.70	12.75	10.10	2.00	11.00	9.05	1.20
0.90	0.99	16.75	14.30	2.80	14.75	13.05	4.10	13.85	9.30	4.05
0.95	0.99	16.45	12.60	3.15	16.00	13.50	2.50	14.65	13.05	3.70
0.99	0.99	16.75	12.50	2.60	16.35	13.30	2.40	18.40	15.50	2.50
1.00	0.99	17.80	13.10	2.05	17.75	13.45	2.25	15.85	12.35	2.50
0.90	1.00	15.95	14.10	3.30	15.85	13.60	4.45	12.40	8.80	4.15
0.95	1.00	15.85	13.25	2.70	19.00	16.80	4.15	15.70	14.10	3.90
0.99	1.00	18.10	14.55	3.05	19.00	15.45	2.60	18.45	14.75	2.30
1.00	1.00	18.10	13.15	2.10	18.15	14.25	2.60	19.70	15.25	2.35

Notes: EC_M uses the 5% MacKinnon critical values of -3.340, -3.272, and -3.248 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002), while EC_t employs a standard 5% t critical value. ρ and q refer to the memory of Y_t and X_t in Equation 2 and 3, respectively. In No Test, we do not conduct pretests for cointegration or integration in Y_t.

Appendix 2 Evaluating Spurious Regression when the Memory of Y_t Ranges from 0 to 1

This articles focuses on cases when Y_t is near-integrated or integrated. However, we also conducted additional simulations where we varied the memory of Y_t (ρ) and the memory of X_t (q) from 0 to 1 by increments of 0.1 for T=30, 60, and 100. Each simulation was conducted 2,000 times. Figure A-1 plots the rate of spurious regression from these 363 experiments. Several results stand out.

First, looking at the dashed line, which corresponds to *no pretests*, we see that even when T=30, when Y_t is stationary (i.e., $\rho \leq 0.2$) the spurious regression rate for the ECM is around 5 percent. These results strongly support the conclusions of De Boef and Keele (2008) and others who argue that whey Y_t is *stationary* the ECM produces correct inferences. Second, we see that as the autoregressive component of Y_t increases, the rate of spurious regression increases. Importantly, across all autoregressive values of Y_t and X_t, we see that first testing whether Y_t behaves as if it contains a unit root and testing that Y_t and X_t are cointegrated avoids the spurious correlation problem.

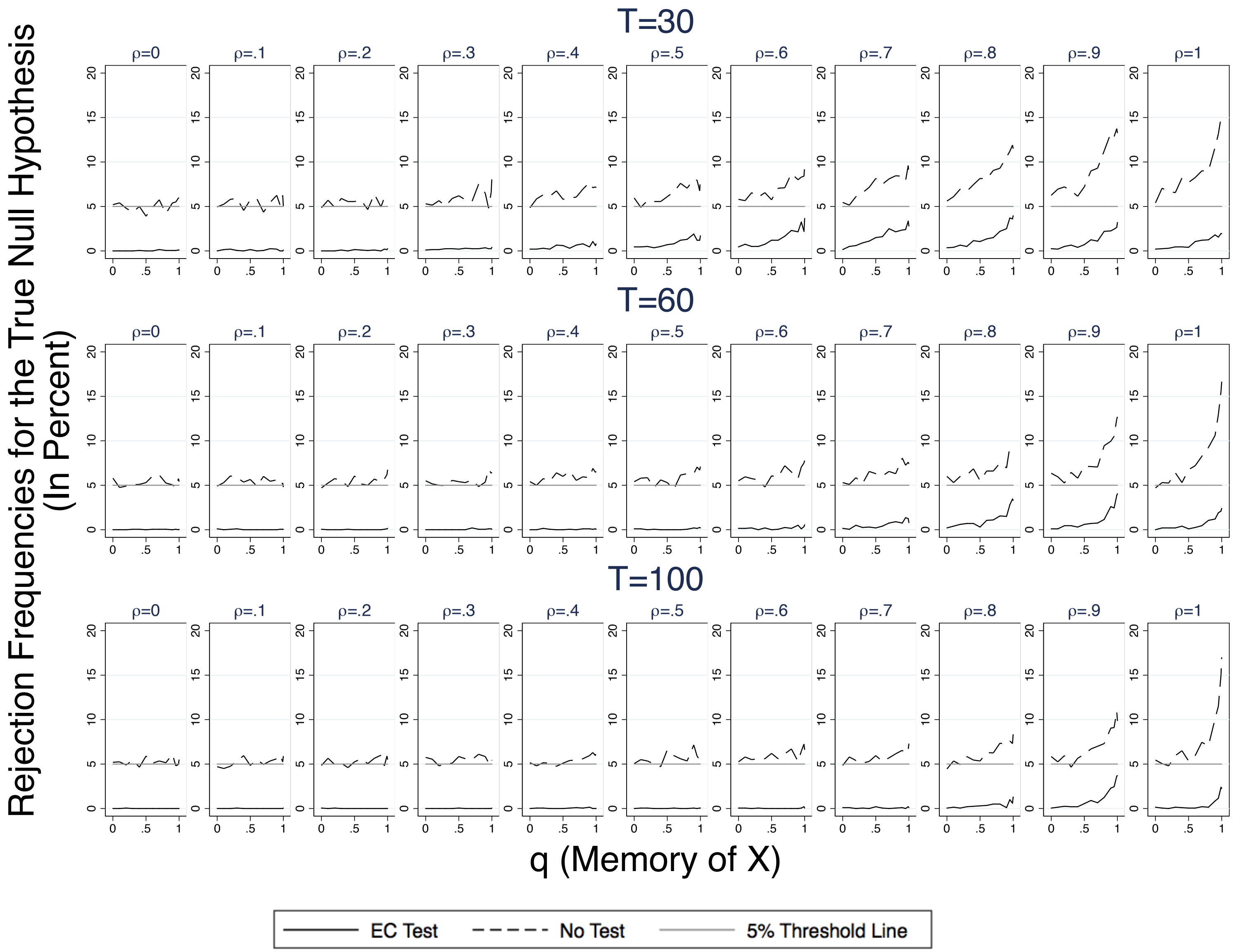

Figure A-1: Rejection Frequencies for the True Null of No Effect of X_t on Y_t ($\beta_1 = 0$)
Note: The dashed line (No Test) represents the rejection frequencies for no pretest for cointegration or integration in Y_t; the solid black line (EC Test) indicates the EC test with the appropriate MacKinnon critical values as a pretest for cointegration when an Augmented Dickey Fuller test indicates that Y_t contains a unit root. q and ρ represent the memory of X_t and Y_t, respectively.

Appendix 3 Estimated Error Correction Rates ($\widehat{\alpha}_1$): Uncorrected, Corrected, and Fractional–ECM

The text demonstrates that when Y_t is near–integrated, although the estimated error correction rate ($\widehat{\alpha}_1$) will produce correct inferences about the presence of cointegration, the estimated rate of error correction will be biased downward when T is small. This results because $\widehat{\alpha}_1$ includes the estimate of the true error correction rate as well as $(\widehat{\rho}-1)$. The text also mentioned a procedure to correct for this bias. In this section, we demonstrate this procedure and report the results from a variety of simulations to illustrate the conditions when the procedure is most likely to improve the estimates of the rate of error correction. We also report simulations where we estimate Fractional Error Correction Models (FECMs). Our aim was to test whether FECM techniques could produce more accurate estimates of the true error correction rate. As we report below, in all simulations the FECM performs worse than the adjusted *and* the unadjusted $\widehat{\alpha}_1$.

To see how well our adjustment procedure performs in alleviating the bias arising from Y_t's deviation from integration, we use Monte Carlo simulations where the DGPs for Y_t and X_t are defined as follows:

$$Y_t = \rho Y_{t-1} + \alpha_1(Y_{t-1} + \frac{\beta_1}{\alpha_1}X_{t-1}) + u_t, \text{where } u_t \sim iid\,\mathrm{N}(0,1) \quad \text{(A-1)}$$

$$X_t = qX_{t-1} + \epsilon_t, \text{where } \epsilon_t \sim iid\,\mathrm{N}(0,1) \quad \text{(A-2)}$$

We vary ρ and q from 0.9 to 1.0 by increments of 0.05 and α_1 from -0.4 to -0.1 by increments of 0.1. The number of iterations for each simulation is set at 2,000. For each simulation, we estimate the single-equation ECM, as defined in Equation 1, and examine if we can correct for the bias in $\widehat{\alpha}_1$ by first regressing ΔY_t on Y_{t-1} to estimate $\hat{\rho}$ and then subtracting $\hat{\rho}-1$ from $\widehat{\alpha}_1$. Since $\widehat{\alpha}_1$ estimates $(\rho-1)+\alpha_1$, subtracting $(\hat{\rho}-1)$ from $\widehat{\alpha}_1$ should correct the bias that results when $\rho \leq 1$.

We also test how Fractional Error Correction Models (FECM) perform in estimating the error correction rate when applied to the analysis of near-integrated or integrated data. Our approach follows directly from Clarke and Lebo (2003). The FECM is defined in the following equation:

$$\Delta^d Y_t = \alpha_0 + \alpha_F \mathrm{FECM}_{t-1} + \beta_F \Delta^d X_t + e_t \quad \text{(A-3)}$$

where $\Delta^d Y_t$ and $\Delta^d X_t$ denote fractionally differenced Y_t and X_t and FECM_{t-1} represents fractionally differenced residuals from the cointegrating regression (Clarke and Lebo, 2003, 295,297). Clarke and Lebo (2003, 294) explain that, "as long as the residuals produced by regressing one fractionally integrated series on another one are of a lower order of integration than are either of the fractionally integrated series, those series can be said to be in dynamic equilibrium and to fractionally cointegrate" (see also footnote 45). Thus, to test whether fractional cointegration is present, we regress Y_t on X_t and evaluate whether the order of integration of the resulting residuals is lower than the order of integration of Y_t or X_t.[25] We adopt ARFIMA modeling procedures to estimate the d parameter for each of the series and fractionally difference them based on their own d values.[26]

Table A-3 presents the means of the estimated error correction rate for each model. Column 1 presents the rate of error correction without any adjustment. Our intuition regarding the potential bias associated with near-integration in Y_t seems to be correct. For $T = 100$, $\rho = 0.9$, and $\alpha_1 = -0.4$, the means of the estimated error correction rates in the ECM range between -0.51 and -0.52, consistent with our prediction that $\widehat{\alpha}_1$ is biased by $0.9 - 1 = -0.1$. In contrast, the adjusted estimates range between -0.3 and -0.43. The adjustment improves the estimate, but also appears to add a conservative bias, meaning we would be more likely to underestimate the error correction rate. Based on this result, we suggest that this adjustment only be made when T is small, which is when the bias in $\widehat{\alpha}_1$ is most severe. Tables A-5 and A-7 (in Appendix 4 below) support this recommendation. These results show that when there are multiple predictors the bias in the *un*adjusted $\widehat{\alpha}_1$ is relatively minor, especially when $T > 60$. Thus, when $T < 60$, scholars may want to report the adjusted $\widehat{\alpha}_1$ as described above, although they should be aware that this is not a perfect fix (given the potential conservative bias of the adjustment).

As described above, as an alternate approach to estimating α_1, we also estimated FECMs. Column 3 presents these results. Both the original and the adjusted estimates of $\widehat{\alpha}_1$ from the ECM consistently outperform the FECM estimates. In fact, all the error correction estimates from the FECM are far from the true error correction rates—often positive when they should be negative. It is not clear why the FECM performs so poorly, but it may relate to the small samples (e.g., Lebo and Weber, 2014). Regardless of the reason, we believe the inaccuracy of the FECM estimates offers an important topic for future research.

[25] In addition to this FECM approach, which Clarke and Lebo (2003) describe on page 294, on page 297 they also discuss estimating the error correction mechanism by regressing fractionally differenced Y_t on fractionally differenced X_t (instead of regressing Y_t on X_t in levels in the first-stage regression). We repeated all FECM simulations following this alternate approach (i.e., fractionally differencing Y_t and X_t prior to estimating the cointegrating relationship). With this approach, the mean value of $\widehat{\alpha}_1$ was farther from the true value than the results reported in A-3. Although we do not report these simulation results here, the code is available at `http://thedata.harvard.edu/dvn/dv/ECM`.

[26] We conduct the likelihood ratio test to see if the inclusion of an autoregressive term for AFRIMA modeling procedures significantly improves the model fit. If it does, we include the autoregressive term for ARFIMA modeling. We use the arfima command in Stata 13.

Table A-3: Means of Estimated Error Correction Rates ($\widehat{\alpha}_1$)

ρ	q	α_1	$\widehat{\alpha}_1$ (1)	Adjusted $\widehat{\alpha}_1$ (2)	FECM $\widehat{\alpha}_1$ (3)
				T=30	
0.90	0.90	-0.1	-0.40	-0.24	0.32
1.00	0.90	-0.1	-0.30	-0.18	0.36
0.90	1.00	-0.1	-0.39	-0.27	0.23
1.00	1.00	-0.1	-0.23	-0.17	0.18
0.90	0.90	-0.4	-0.64	-0.33	0.04
1.00	0.90	-0.4	-0.56	-0.30	0.11
0.90	1.00	-0.4	-0.65	-0.39	0.07
1.00	1.00	-0.4	-0.58	-0.36	0.11
				T=60	
0.90	0.90	-0.1	-0.26	-0.16	0.58
1.00	0.90	-0.1	-0.16	-0.11	0.54
0.90	1.00	-0.1	-0.25	-0.19	0.34
1.00	1.00	-0.1	-0.13	-0.11	0.36
0.90	0.90	-0.4	-0.51	-0.29	0.01
1.00	0.90	-0.4	-0.43	-0.25	0.06
0.90	1.00	-0.4	-0.54	-0.39	-0.06
1.00	1.00	-0.4	-0.45	-0.33	-0.01
				T=100	
0.90	0.90	-0.1	-0.22	-0.14	0.53
1.00	0.90	-0.1	-0.13	-0.08	N/A
0.90	1.00	-0.1	-0.22	-0.19	0.58
1.00	1.00	-0.1	-0.11	-0.10	0.69
0.90	0.90	-0.4	-0.51	-0.30	-0.03
1.00	0.90	-0.4	-0.40	-0.26	0.06
0.90	1.00	-0.4	-0.52	-0.43	-0.04
1.00	1.00	-0.4	-0.43	-0.35	0.00

In this table, we present the means of estimated error correction rates ($\widehat{\alpha}_1$) for all instances where we find evidence of a unit root in Y_t and cointegration. For FECM, we report the means of error correction rates for all instances where we find evidence of fractional cointegration. "N/A" indicates that there was no single simulation where the series were found to be fractionally cointegrated. The number of iteration for all simulations is 2,000.

Appendix 4 Numerical Results for the Trivariate Cases

In order to save space, in the main text we do not report all the results. In the following tables, we report our simulation results from all the permutations of different parameter values that we examined. For Tables A-4—A-7, the columns labeled EC_M correspond with the results when we find evidence of a unit root in Y_t and an EC test with the appropriate MacKinnon critical values provides evidence of cointegration.

In Table A-4, we begin with the results from the trivariate case where there is no relationship among Y_t, $X_{1,t}$, and $X_{2,t}$, as defined in Equations 9-11. We let the values of ρ, q, and ξ vary from 0.9 to 1.0 by increments of 0.05. In Table A-5, we report the results from the case where we introduce one cointegrating relationship that binds Y_t, $X_{1,t}$, and $X_{2,t}$ together (as expressed in Equation 12). Again, q and ξ take the values of 0.9 to 1.0 by increments of 0.05 and α_1 takes the value of -0.4 to -1.0 by increments of 0.1. Lastly, Tables A-6 and A-7 report the trivariate cases where $X_{2,t}$ is cointegrated with Y_t but $X_{1,t}$ is not. As stated in the text, we set $\beta_2 = 0.3$ in Equation 13. Once again, we vary q and ξ from 0.9 to 1.0 by increments of 0.05 while letting α_1 take the value of -0.4 to -1.0 by increments of 0.1. Table A-6 tabulates the rate of spurious regression between Y_t and $X_{1,t}$ (meaning that β_1 is zero in Equation 8), while Table A-7 tabulates the rate of correctly identifying the relationship between Y_t and $X_{2,t}$ with the true value of β_2=0.3.

We also used simulations to evaluate the performance of fractional error correction models (FECMs). In Tables A-4—A-7, the columns labeled FECM report the frequencies of finding evidence of fractional cointegration and a statistically significant effect of fractionally-differenced $X_{1,t}$ variable on fractionally-differenced Y_t. The results show that when no relationship exists between X_t and Y_t, the FECM is even more conservative than the ECM. If the goal is *only* to avoid finding spurious correlations, the FECM will suffice with near–integrated and integrated data. However, we also want to avoid Type II errors. Here, the FECM does *not* perform well. Tables A-5 and A-7 show that even when T=100 and $\alpha_1 = -0.4$, the FECM typically *fails* to find evidence of a true relationship more than 70 percent of the time. When T is < 100 *or* α_1 is greater than -0.4, the Type II error rate is even greater.

Table A-4: Rejection Frequencies (%) for the True Null Hypothesis: $\beta_1 = 0$ (Case 1)

			T=30				T=60				T=100			
ρ	q	ξ	No Test (1)	EC_t (2)	EC_M (3)	FECM (4)	No Test (5)	EC_t (6)	EC_M (7)	FECM (8)	No Test (9)	EC_t (10)	EC_M (11)	FECM (12)
0.90	0.90	0.90	11.95	10.10	1.70	4.50	11.45	10.40	1.90	1.05	9.25	8.80	1.30	0.00
0.90	0.90	0.95	12.45	9.95	2.10	4.45	10.15	9.45	1.90	0.80	10.00	9.80	1.80	0.20
0.90	0.90	1.00	12.00	10.10	1.50	4.20	11.90	11.15	2.70	0.60	9.70	9.60	1.60	0.10
0.90	0.95	0.90	13.50	11.15	2.20	4.75	11.75	11.25	2.20	1.05	10.90	10.55	2.35	0.10
0.90	0.95	0.95	12.60	10.85	2.10	5.55	12.05	10.95	2.35	0.65	10.80	10.55	2.30	0.30
0.90	0.95	1.00	13.35	10.85	2.60	4.80	12.85	11.95	2.45	0.85	11.70	11.60	3.20	0.20
0.90	1.00	0.90	13.95	11.50	2.15	5.05	12.90	12.35	2.75	1.25	11.70	11.65	2.85	0.20
0.90	1.00	0.95	13.60	11.40	3.00	5.35	13.50	12.75	3.00	1.10	13.05	13.05	3.10	0.15
0.90	1.00	1.00	12.90	11.10	2.00	4.40	13.25	12.00	2.55	1.00	12.15	12.10	3.20	0.10
0.95	0.90	0.90	13.25	9.40	1.75	3.25	11.75	9.65	1.20	1.00	11.20	9.85	1.35	0.05
0.95	0.90	0.95	12.40	9.15	1.40	2.90	13.05	10.05	1.80	0.65	9.35	7.95	1.60	0.25
0.95	0.90	1.00	12.40	9.25	1.65	3.75	11.75	9.50	1.85	0.50	11.50	10.05	1.80	0.05
0.95	0.95	0.90	15.10	11.45	1.95	5.00	12.55	9.90	2.10	0.75	11.45	9.80	1.70	0.05
0.95	0.95	0.95	13.90	11.05	1.95	4.40	13.20	10.80	1.55	0.60	12.30	10.90	2.35	0.15
0.95	0.95	1.00	14.70	11.85	2.35	5.80	12.30	10.80	2.60	0.25	12.05	11.00	2.00	0.05
0.95	1.00	0.90	15.05	11.70	2.05	5.15	16.00	13.80	2.20	1.35	12.80	11.65	2.20	0.00
0.95	1.00	0.95	15.60	13.05	1.90	4.50	14.90	12.70	2.50	0.75	15.85	14.65	3.00	0.15
0.95	1.00	1.00	13.35	10.90	2.05	4.95	15.00	13.25	2.75	0.35	15.35	14.10	3.25	0.10
1.00	0.90	0.90	13.65	9.95	1.65	3.50	10.85	7.50	1.20	0.50	9.55	6.60	0.70	0.05
1.00	0.90	0.95	11.15	7.80	1.70	3.10	11.45	7.95	0.95	0.60	10.10	6.60	1.05	0.00
1.00	0.90	1.00	12.05	8.40	1.85	3.25	12.55	8.65	1.45	0.55	10.80	7.40	1.30	0.00
1.00	0.95	0.90	14.30	9.70	1.75	3.35	13.40	9.25	1.40	0.95	12.30	8.55	1.10	0.05
1.00	0.95	0.95	14.30	10.35	2.05	3.85	12.60	9.35	1.50	0.40	10.60	7.70	0.95	0.05
1.00	0.95	1.00	13.45	10.20	2.35	3.55	13.65	10.35	1.55	0.40	10.85	8.10	1.35	0.05
1.00	1.00	0.90	15.20	11.45	2.00	6.30	17.15	13.80	2.45	0.90	18.10	13.15	1.55	0.00
1.00	1.00	0.95	16.20	12.95	2.65	4.85	15.30	11.50	1.60	0.45	17.55	13.55	2.20	0.00
1.00	1.00	1.00	13.95	10.75	1.70	4.65	15.90	12.90	1.70	0.60	15.45	11.90	1.70	0.05

Notes: EC_M uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002) while EC_t employs the standard t critical value. FECM reports the frequencies of finding evidence of fractional cointegration and a statistically significant effect of fractionally-differenced $X_{1,t}$ variable on fractionally-differenced Y_t. ρ, q and ξ correspond to the memories of Y_t, $X_{1,t}$ and $X_{2,t}$, respectively. In No Test, we do not conduct pretests for cointegration or integration.

Table A-5: Rejection Frequencies (%) for the False Null Hypothesis: $\beta_1 = 0.3$ (Case 2)

			T=30				T=60				T=100			
α_1	q	ξ	No Test (1)	EC_M (2)	FECM (3)	$\widehat{\alpha}_1$ (4)	No Test (5)	EC_M (6)	FECM (7)	$\widehat{\alpha}_1$ (8)	No Test (9)	EC_M (10)	FECM (11)	$\widehat{\alpha}_1$ (12)
-0.1	0.9	0.9	50.94	5.59	2.40	-0.267	89.23	43.15	0.51	-0.142	99.39	87.06	0.05	-0.116
-0.1	0.9	1.0	49.22	6.17	3.65	-0.230	86.96	53.39	0.42	-0.130	99.21	95.37	0.05	-0.111
-0.1	1.0	0.9	59.29	8.89	6.62	-0.234	94.77	57.37	0.85	-0.131	99.90	96.19	0.10	-0.111
-0.1	1.0	1.0	53.16	8.57	7.93	-0.223	94.02	62.23	0.69	-0.125	99.79	97.27	0.21	-0.109
-0.2	0.9	0.9	53.89	17.03	4.42	-0.399	90.60	71.44	1.09	-0.251	99.42	98.15	0.58	-0.218
-0.2	0.9	1.0	49.90	15.22	5.18	-0.382	89.15	75.13	0.62	-0.246	99.33	98.87	0.46	-0.218
-0.2	1.0	0.9	62.31	19.62	6.96	-0.383	95.71	79.32	1.40	-0.245	100.00	99.38	0.36	-0.217
-0.2	1.0	1.0	58.79	19.21	7.35	-0.364	95.15	79.45	0.78	-0.244	99.84	99.64	0.26	-0.218
-0.3	0.9	0.9	54.46	24.55	6.00	-0.503	91.35	84.04	2.29	-0.354	99.46	99.46	2.39	-0.322
-0.3	0.9	1.0	50.53	23.99	6.34	-0.497	90.11	84.41	1.44	-0.355	99.52	99.41	2.09	-0.325
-0.3	1.0	0.9	64.62	29.83	9.64	-0.502	96.85	90.81	3.31	-0.353	100.00	100.00	3.47	-0.325
-0.3	1.0	1.0	59.26	29.10	8.47	-0.496	94.84	90.09	1.56	-0.348	99.95	99.95	1.72	-0.325
-0.4	0.9	0.9	55.04	30.41	10.25	-0.595	93.32	91.62	6.28	-0.443	99.91	99.91	8.45	-0.415
-0.4	0.9	1.0	53.72	31.89	10.56	-0.599	91.69	90.37	5.44	-0.451	99.29	99.29	10.53	-0.430
-0.4	1.0	0.9	64.01	36.83	15.58	-0.599	98.05	96.15	9.41	-0.454	100.00	100.00	14.13	-0.427
-0.4	1.0	1.0	63.17	38.71	16.18	-0.598	95.94	94.29	6.32	-0.452	99.89	99.89	8.74	-0.430

Notes: EC_M uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002) while ECM_t employs the standard t critical value for ρ=0.05, two-tail. FECM reports the frequencies of finding evidence of fractional cointegration and a statistically significant effect of fractionally-differenced $X_{1,t}$ variable on fractionally-differenced Y. q and ξ correspond to the memories of $X_{1,t}$ and $X_{2,t}$, respectively. α_1 is the error correction as denoted in Equation 13. In No Test, we do not conduct pretests for cointegration or integration.

Table A-6: Rejection Frequencies (%) for the True Null Hypothesis: $\beta_1 = 0$ (Case 3)

			T=30				T=60				T=100			
α_1	q	ξ	No Test	EC_t	EC_M	FECM	No Test	EC_t	EC_M	FECM	No Test	EC_t	EC_M	FECM
-0.1	0.9	0.9	8.35	6.50	1.60	2.60	6.50	5.70	3.05	0.65	6.05	6.05	4.50	0.10
-0.1	0.9	1.0	7.25	6.15	1.65	3.05	5.80	5.40	3.90	0.40	6.10	6.10	5.50	0.10
-0.1	1.0	0.9	10.30	8.40	2.30	4.15	7.90	7.55	3.70	0.60	7.45	7.45	6.20	0.15
-0.1	1.0	1.0	7.75	6.65	2.00	4.75	8.35	8.00	5.45	0.60	5.50	5.50	4.85	0.10
-0.2	0.9	0.9	9.75	8.90	4.00	4.55	7.05	6.90	5.40	0.75	6.50	6.50	5.70	0.45
-0.2	0.9	1.0	7.75	6.70	3.35	4.75	7.15	7.10	5.40	0.75	6.10	6.10	5.85	0.30
-0.2	1.0	0.9	10.10	9.35	3.45	4.70	7.85	7.75	5.70	1.05	6.70	6.70	6.00	0.40
-0.2	1.0	1.0	8.70	7.85	3.85	4.85	7.75	7.70	6.40	0.70	6.35	6.35	6.25	0.35
-0.3	0.9	0.9	8.90	8.40	4.50	5.55	6.85	6.85	5.00	1.85	6.40	6.40	3.70	1.10
-0.3	0.9	1.0	8.90	8.65	4.85	5.50	6.25	6.25	5.45	1.30	6.45	6.45	5.65	1.20
-0.3	1.0	0.9	9.45	9.15	3.85	6.80	7.50	7.50	5.75	2.15	6.25	6.25	3.15	1.50
-0.3	1.0	1.0	8.75	8.25	4.35	6.25	7.95	7.90	7.10	1.45	6.30	6.30	5.55	1.40
-0.4	0.9	0.9	7.85	7.65	4.30	6.50	7.80	7.80	4.90	3.70	5.90	5.90	1.70	2.60
-0.4	0.9	1.0	7.45	7.35	4.60	7.55	5.90	5.90	4.55	2.75	5.95	5.95	4.65	4.65
-0.4	1.0	0.9	8.05	7.90	4.30	7.15	7.30	7.30	3.70	3.80	6.30	6.30	1.35	2.90
-0.4	1.0	1.0	8.80	8.70	5.85	9.70	7.50	7.50	6.10	3.35	6.10	6.10	5.05	6.50

Notes: EC_M uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002) while EC_t employs the standard t critical value for ρ=0.05, two-tail. FECM reports the frequencies of finding evidence of fractional cointegration and a statistically significant effect of fractionally-differenced $X_{1,t}$ variable on fractionally-differenced Y_t. q and ξ correspond to the memories of $X_{1,t}$ and $X_{2,t}$, respectively.

Table A-7: Rejection Frequencies (%) for the False Null Hypothesis: $\beta_2 = 0.3$ (Case 3)

			T=30				T=60				T=100			
α_1	q	ξ	No Test (1)	EC$_M$ (2)	FECM (3)	$\widehat{\alpha}_1$ (4)	No Test (5)	EC$_M$ (6)	FECM (7)	$\widehat{\alpha}_1$ (8)	No Test (9)	EC$_M$ (10)	FECM (11)	$\widehat{\alpha}_1$ (12)
-0.1	0.9	0.9	55.84	5.21	2.47	-0.350	92.36	24.95	0.67	-0.183	99.53	64.33	0.00	-0.132
-0.1	0.9	1.0	59.60	6.76	4.88	-0.299	96.35	46.27	0.48	-0.143	99.95	91.44	0.05	-0.115
-0.1	1.0	0.9	52.48	4.79	2.95	-0.397	91.39	23.17	0.21	-0.189	99.38	63.15	0.00	-0.137
-0.1	1.0	1.0	58.06	6.60	6.71	-0.300	96.76	43.03	0.53	-0.152	99.84	88.65	0.16	-0.117
-0.2	0.9	0.9	57.78	12.43	4.38	-0.483	93.27	53.99	1.29	-0.285	99.66	93.32	0.79	-0.228
-0.2	0.9	1.0	64.76	14.92	8.55	-0.453	97.43	67.47	1.63	-0.268	100.00	98.24	0.72	-0.225
-0.2	1.0	0.9	55.48	11.78	5.02	-0.510	91.06	52.26	0.75	-0.294	99.78	94.33	0.73	-0.232
-0.2	1.0	1.0	64.56	15.33	7.61	-0.462	96.13	65.65	0.99	-0.276	99.84	97.91	0.42	-0.229
-0.3	0.9	0.9	60.52	19.42	9.16	-0.579	94.20	75.94	4.24	-0.375	99.91	99.74	3.29	-0.322
-0.3	0.9	1.0	65.08	24.32	12.46	-0.575	97.36	82.70	4.94	-0.379	100.00	99.78	4.34	-0.334
-0.3	1.0	0.9	60.48	19.07	9.36	-0.599	94.19	75.56	2.94	-0.379	100.00	99.48	3.69	-0.323
-0.3	1.0	1.0	65.45	22.26	11.35	-0.572	96.71	81.13	2.86	-0.378	100.00	99.84	4.67	-0.335
-0.4	0.9	0.9	60.69	27.00	14.32	-0.661	96.47	89.95	11.87	-0.450	100.00	100.00	17.15	-0.413
-0.4	0.9	1.0	70.54	32.83	21.23	-0.659	98.70	93.70	15.69	-0.469	99.87	99.87	30.67	-0.434
-0.4	1.0	0.9	61.89	28.45	15.04	-0.654	94.59	88.72	10.55	-0.454	100.00	100.00	14.58	-0.408
-0.4	1.0	1.0	66.92	30.63	21.12	-0.669	98.04	92.48	12.35	-0.473	99.94	99.94	30.16	-0.438

Notes: EC$_M$ uses the 5% MacKinnon critical values of -3.642, -3.566, and -3.540 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). FECM reports the frequencies of finding evidence of fractional cointegration and a statistically significant effect of fractionally-differenced $X_{1,t}$ variable on fractionally-differenced Y_t. q and ξ correspond to the memories of $X_{1,t}$ and $X_{2,t}$, respectively. α_1 is the error correction as denoted in Equation 13. In No Test, we do not conduct pretests for cointegration or integration. Columns 4, 8, and 12 report the means of estimated $\widehat{\alpha}_1$ coefficients for series where we find evidence of both integration and cointegration.

Appendix 5 Full Results for the Ehrlich (2008) Replication

Figure 2 in the text reported the estimated relationship between the RTAA and U.S. Tariff Rates from three separate models (Ehrlich's ECM, First Differences, and Fractionally Differenced). Table A-8, below, reports the full results from these models. Column 1 replicates Ehrlich's ECM. Because we do not see evidence of cointegration and because an augmented Dickey Fuller test cannot reject the null hypothesis that the dependent variable (tariff rates) contains a unit root, the model in Column 2 first differences all variables for which we cannot reject the null of a unit root. Variables that behaved as if stationary were not differenced. Column 3 presents the results of an analysis where we fractionally difference all variables that were not dichotomous or trichotomous.[27] As discussed in the text, RTAA (Long Term) is a logistical functional form (see, e.g., Figure 1) to capture the hypothesized long term influence of the RTAA on tariff rates. All other variables come from Ehrlich (2008).

Table A-8: The Relationship between the RTAA and U.S. Tariff Rates, 1902 to 1988

	ECM (Replication)	First Differences	Fractional Differencing
Δ Tariff_{t-1}	0.25	–	–
	(0.15)		
Tariff_{t-1}	-0.07	–	–
	(0.06)		
Δ Unemployment	0.09*	–	0.06
	(0.04)		(0.07)
Δ Inflation	-0.65	–	-0.12
	(1.74)		(2.34)
Δ Capital Labor Ratio	-0.00	-0.00	-0.001*
	(0.00)	(0.00)	(0.000)
Δ Terms of Trade	0.02	–	0.09*
	(0.02)		(0.03)
Δ Exports	-0.41*	–	-0.97*
	(0.12)		(0.16)
Δ Protectionist Party Control	-0.27	-0.08	0.02
	(0.46)	(0.51)	(0.94)
Δ Protectionist Pres./Free Trader Cong.	0.25	0.38	0.19
	(0.29)	(0.42)	(0.78)
Δ Free Trader Pres./Protectionist Cong.	-0.94*	-1.05*	-0.08
	(0.34)	(0.48)	(0.89)
Δ War	-0.43	-0.48	-0.98

Continued on Next Page…

[27] As with previous fractional integration analyses, we conduct the likelihood ratio test to see if the inclusion of an autoregressive term for ARFIMA modeling procedures significantly improves the model fit. If it does, we include the autoregressive term. We use the arfima command in Stata 13.

Table A-8 – Continued

	ECM (Replication)	First Differences	Fractional Differencing
	(1.12)	(0.52)	(0.97)
GATT	0.13	–	–
	(0.12)		
Δ GATT	–	0.001	0.35
		(0.446)	(0.63)
RTAA	-1.15*	-1.64*	-3.50*
	(0.47)	(0.39)	(0.81)
RTAA (Long Term)	–	1.05*	-0.57
		(0.37)	(2.76)
Income Tax	0.82	1.78*	-1.25
	(0.78)	(0.43)	(0.75)
$\text{Unemployment}_{t-1}$	0.01	–	–
	(0.03)		
Inflation_{t-1}	-0.28	–	–
	(2.09)		
$\text{Capital Labor Ratio}_{t-1}$	-0.00	–	–
	(0.00)		
$\text{Terms of Trade}_{t-1}$	0.01	–	–
	(0.01)		
Exports_{t-1}	-0.15	–	–
	(0.11)		
Unemployment_{t}	–	0.01	–
		(0.03)	
Inflation_{t}	–	-1.33	–
		(1.32)	
$\text{Terms of Trade}_{t}$	–	0.003	–
		(0.012)	
Exports_{t}	–	-0.16	–
		(0.09)	
Constant	0.36	-0.64	19.05*
	(1.29)	(1.39)	(3.79)
R^2		0.45	0.62
N	87	87	88

Notes: Column 1 estimates NeweyWest robust standard errors with three lags (Erhlich 2008, 439). Column 2 first–differences all variables for which we cannot reject the null of a unit root. Column 3 fractionally differences all variables that are not dichotomous or trichotomous. Thus, in Columns 1 and 2, Δ indicates first differencing and in Column 3, Δ indicates fractional differencing. RTAA and RTAA (Long Term) are never differenced. $* = p < .05$ (two–tailed tests), standard errors in parentheses.

Appendix 6 Cases with More Predictors

To ensure that the trivariate simulations reported in the text generalize, we also conducted simulations for systems where we doubled the number of predictors to four. The main results are very similar to what we have obtained in the bivariate and trivariate cases. In this section, we present the results from cases where we generate five series: Y_t, $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, and $X_{4,t}$. Similar to our trivariate cases, the following situations are considered:

Case 1: The five series are all (near-)integrated and unrelated to each other.

Case 2: The five series altogether form one cointegrating relationship.

Case 3: Y_t is (near-)integrated and (near-)cointegrated with one of the X's.

In each case, we conduct 2,000 replications for the sample size T of 30, 60, and 100. For Case 1 where Y_t, $X_{1,t}$, $X_{2,t}$, X_3, and X_4 are (near-)integrated and have no relationships amongst themselves, this system is:

$$Y_t = \rho Y_{t-1} + u_{1,t}, \text{where } u_{1,t} \sim iid\,\mathrm{N}(0,1) \tag{A-4}$$
$$X_{1,t} = q_1 X_{1,t-1} + u_{2,t}, \text{where } u_{2,t} \sim iid\,\mathrm{N}(0,1) \tag{A-5}$$
$$X_{2,t} = \xi_1 X_{2,t-1} + u_{3,t}, \text{where } u_{3,t} \sim iid\,\mathrm{N}(0,1) \tag{A-6}$$
$$X_{3,t} = q_2 X_{3,t-1} + u_{4,t}, \text{where } u_{4,t} \sim iid\,\mathrm{N}(0,1) \tag{A-7}$$
$$X_{4,t} = \xi_2 X_{4,t-1} + u_{5,t}, \text{where } u_{5,t} \sim iid\,\mathrm{N}(0,1) \tag{A-8}$$

We let ρ, q_1, q_2, ξ_1, and ξ_2 each take values of 0.9 or 1.0. We replicate the same procedure as in the bivariate and trivariate cases examined in the text. The following single-equation ECM is estimated for this system:

$$\begin{aligned}\Delta Y_t = \alpha_0 + \alpha_1 y_{t-1} + \gamma_1 \Delta x_{1,t} + \beta_1 x_{1,t-1} + \gamma_2 \Delta x_{2,t} + \beta_2 x_{2,t-1} \\ + \gamma_3 \Delta x_{3,t} + \beta_3 x_{3,t-1} + \gamma_4 \Delta x_{4,t} + \beta_4 x_{4,t-1} + \epsilon_{1,t}\end{aligned} \tag{A-9}$$

In Table A-9, we report the results for Case 1. That is, the rate of finding a spurious relationship between Y_t and $X_{1,t}$ (or a statistically significant $\hat{\beta}_1$) in the estimated ECM for Equation A-9. We do not report the results for the rest of the predictors as the results are mirror images of $X_{1,t}$. The results indicate that our ability of avoid spurious conclusions remains.

Table A-9: Rejection Frequencies (%) for the True Null Hypothesis: $\beta_1 = 0$ (Case 1)

Y	$X_{1,t}$	$X_{2,t}$	X_3	X_4	T=30		T=60		T=100	
ρ	q_1	ξ_1	q_2	ξ_2	No Test	EC$_M$	No Test	EC$_M$	No Test	EC$_M$
					(1)	(2)	(3)	(4)	(5)	(6)
0.90	0.90	0.90	0.90	0.90	10.95	1.30	11.90	1.55	10.20	1.15
0.90	0.90	0.90	0.90	1.00	13.90	2.00	13.95	1.85	11.40	0.65
0.90	0.90	0.90	1.00	1.00	12.25	1.35	13.75	1.70	11.25	1.60
0.90	0.90	1.00	1.00	1.00	12.90	1.75	13.45	2.45	11.60	1.70
0.90	1.00	1.00	1.00	1.00	14.30	2.30	14.30	2.00	12.60	2.45
1.00	1.00	1.00	1.00	1.00	13.40	2.05	15.30	1.80	17.30	1.85

Notes: This table shows the rate of finding a spurious relationship between Y_t and $X_{1,t}$ when all the X_t variables are unrelated to Y_t (Case 1). EC$_M$ uses the 5% MacKinnon critical values of -4.119, -4.037, and -4.014 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). ρ, q_1, ξ_1, q_2, and ξ_2 correspond to the memories of Y_t, $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, and $X_{4,t}$, respectively. In No Test, we do not conduct pretests for cointegration or integration.

In Case 2, the DGP for Y_t is modified as:

$$Y_t = Y_{t-1} + \alpha_1(Y_{t-1} + \frac{\beta_1}{\alpha_1}X_{1,t-1} + \frac{\beta_2}{\alpha_1}X_{2,t-1} + \frac{\beta_3}{\alpha_1}X_{3,t-1} + \frac{\beta_4}{\alpha_1}X_{4,t-1}) + u_{7,t}, \text{ where } u_{7,t} \sim iid\,\mathrm{N}(0,1) \quad \text{(A-10)}$$

such that all five series constitute one cointegrating relationship. Just as in Case 1, we allow the memory of the X_t variables to equal 0.9 or 1.0 and set $\alpha_1 = (-0.1, -0.4)$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0.3$. Table A-10 presents the rate of correctly identifying the relationship between Y_t and $X_{1,t}$ in Case 2.

The results in Table A-10 reinforce previous findings. When T=100, the power of the ECM to identify true relationships is close to or exceeds the expected 95%. When $T < 100$, testing for integration in Y_t and cointegration comes at some cost. Of course, the cost depends on the strength of the relationship between the predictors and Y_t and on the error correction rate. As noted in the text, we have chosen conservative values for both of these parameters in our simulations.

Table A-10: Rejection Frequencies (%) for the False Null Hypothesis: $\beta_1 = 0.3$ (Case 3)

Y	$X_{1,t}$	$X_{2,t}$	X_3	X_4	T=30		T=60		T=100	
α_1	q_1	ξ_1	q_2	ξ_2	No Test	EC$_M$	No Test	EC$_M$	No Test	EC$_M$
					(1)	(2)	(3)	(4)	(5)	(6)
-0.1	0.90	0.90	0.90	0.90	32.71	2.55	82.29	35.88	98.45	87.63
-0.1	0.90	0.90	0.90	1.00	33.99	3.35	78.99	40.02	98.39	91.62
-0.1	0.90	0.90	1.00	1.00	30.01	3.31	78.82	40.97	97.95	93.50
-0.1	0.90	1.00	1.00	1.00	32.21	3.79	75.82	41.90	97.04	93.23
-0.1	1.00	1.00	1.00	1.00	35.97	3.42	85.36	47.75	99.53	95.83
-0.4	0.90	0.90	0.90	0.90	38.70	19.73	83.98	81.75	98.89	98.89
-0.4	0.90	0.90	0.90	1.00	40.56	19.96	83.57	81.91	98.55	98.55
-0.4	0.90	0.90	1.00	1.00	36.54	17.90	80.72	78.90	98.03	98.03
-0.4	0.90	1.00	1.00	1.00	38.69	18.92	80.46	78.22	97.75	97.75
-0.4	1.00	1.00	1.00	1.00	41.00	20.45	90.04	88.32	99.48	99.48

Notes: This table shows the rate of correctly identifying the relationship between Y_t and $X_{1,t}$ when Y_t and all the X_t's altogether form one cointegrating relationship (Case 2). EC$_M$ uses the 5% MacKinnon critical values of -4.119, -4.037, and -4.014 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). q_1, ξ_1, q_2, and ξ_2 correspond to the memories of $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, and $X_{4,t}$, respectively. α_1 is the error correction as denoted in Equation A-10. In No Test, we do not conduct pretests for cointegration or integration.

In Case 3, we introduce (near-)cointegration between Y_t and $X_{2,t}$. The DGP for Y_t is now re-defined as follows:

$$Y_t = Y_{t-1} + \alpha_1 (Y_{t-1} + \frac{\beta_2}{\alpha_1} X_{2,t-1}) + u_{6,t}, \text{where } u_{4,t} \sim iid\, \mathrm{N}(0,1) \qquad \text{(A-11)}$$

where the DGPs for the other X's remain the same as above. We then estimate the ECM as specified in Equation A-9. We test how the presence of this (near-)cointegrating relationship between Y_t and $X_{2,t}$ may affect the frequency of finding a spurious relationship between Y_t and $X_{1,t}$. We set α_1 at -0.1 and -0.4 and $\beta_2 = 0.3$. We let the memory of each X_t variable equal 0.9 or 1.0. Table A-11 shows the rate of finding a spurious relationship between Y_t and $X_{1,t}$ (i.e., a statistically significant $\hat{\beta}_1$). Table A-12 tabulates the frequency of correctly identifying the long-run effect of $X_{2,t}$ on Y_t (i.e., a statistically significant $\hat{\beta}_2$).

Table A-11 shows that even when just one predictor is cointegrated with Y_t and three other unrelated (near-)integrated variables are included in the model, as long as we verify that Y_t behaves like a unit root and we find evidence of a cointegrating relationship, the rate of spurious regression in the ECM is around five percent or less. Consistent with previous findings, Table A-12 shows that the recommended statistical tests do come at some cost in terms of identifying the true relationship. Adding three unrelated predictors to the model makes it even more difficult to correctly identify the true relationship, especially when T is small and the error correction rate is weak. This low power appears to be the cost of avoiding spurious regression.

Table A-11: Rejection Frequencies (%) for the True Null Hypothesis: $\beta_1 = 0$ (Case 2)

Y	$X_{1,t}$	$X_{2,t}$	X_3	X_4	T=30		T=60		T=100	
α_1	q_1	ξ_1	q_2	ξ_2	No Test	EC_M	No Test	EC_M	No Test	EC_M
					(1)	(2)	(3)	(4)	(5)	(6)
-0.1	0.90	0.90	0.90	0.90	10.10	1.20	8.50	2.80	6.55	3.75
-0.1	0.90	0.90	0.90	1.00	8.80	0.90	7.85	1.95	7.00	4.05
-0.1	0.90	0.90	1.00	1.00	9.10	1.35	8.75	2.00	7.25	3.80
-0.1	0.90	1.00	1.00	1.00	9.35	1.15	7.50	2.50	6.30	4.40
-0.1	1.00	1.00	1.00	1.00	9.45	1.65	7.55	2.95	6.50	5.25
-0.4	0.90	0.90	0.90	0.90	8.45	2.95	8.75	4.05	6.50	1.75
-0.4	0.90	0.90	0.90	1.00	9.20	3.65	8.00	3.85	7.70	1.65
-0.4	0.90	0.90	1.00	1.00	9.15	3.40	8.90	4.70	8.30	1.45
-0.4	0.90	1.00	1.00	1.00	10.20	4.25	7.80	5.80	6.20	4.95
-0.4	1.00	1.00	1.00	1.00	10.70	4.20	8.40	6.05	7.00	5.55

Notes: This table shows the rate of finding a spurious relationship between Y_t and $X_{1,t}$ when $X_{2,t}$ is (near-)cointegrated with Y_t and the rest of the X_t variables are unrelated to Y_t (Case 3). EC_M uses the 5% MacKinnon critical values of -4.119, -4.037, and -4.014 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). q_1, ξ_1, q_2, and ξ_2 correspond to the memories of $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, and $X_{4,t}$, respectively. α_1 is the error correction as denoted in Equation A-11. In No Test, we do not conduct pretests for cointegration or integration.

Table A-12: Rejection Frequencies (%) for the False Null Hypothesis: $\beta_2 = 0.3$ (Case 2)

Y	$X_{1,t}$	$X_{2,t}$	X_3	X_4	T=30		T=60		T=100	
α_1	q_1	ξ_1	q_2	ξ_2	No Test	EC_M	No Test	EC_M	No Test	EC_M
					(1)	(2)	(3)	(4)	(5)	(6)
-0.1	0.90	0.90	0.90	0.90	39.51	1.95	85.41	11.55	98.61	39.56
-0.1	0.90	0.90	0.90	1.00	41.22	2.19	84.13	9.82	99.02	40.87
-0.1	0.90	0.90	1.00	1.00	39.43	2.61	82.54	10.02	98.81	38.31
-0.1	0.90	1.00	1.00	1.00	44.94	2.18	90.25	19.29	99.63	66.30
-0.1	1.00	1.00	1.00	1.00	42.47	2.79	89.43	16.77	99.63	65.53
-0.4	0.90	0.90	0.90	0.90	47.24	13.90	91.84	70.64	99.79	99.57
-0.4	0.90	0.90	0.90	1.00	46.52	15.35	94.20	74.14	99.77	99.32
-0.4	0.90	0.90	1.00	1.00	48.31	15.28	90.60	70.99	99.02	98.53
-0.4	0.90	1.00	1.00	1.00	51.69	17.98	94.96	77.80	99.87	99.54
-0.4	1.00	1.00	1.00	1.00	51.14	16.85	92.24	76.54	99.87	99.68

Notes: This table shows the rate of correctly identifying the relationship between Y_t and $X_{2,t}$ when only $X_{2,t}$ is (near-)cointegrated with Y_t and the rest of the X_t variables are unrelated to Y_t (Case 3). EC_M uses the 5% MacKinnon critical values of -4.119, -4.037, and -4.014 for T=30, T=60, and T=100, respectively, computed from Ericsson and MacKinnon (2002). q_1, ξ_1, q_2, and ξ_2 correspond to the memories of $X_{1,t}$, $X_{2,t}$, $X_{3,t}$, and $X_{4,t}$, respectively. α_1 is the error correction as denoted in Equation A-11. In No Test, we do not conduct pretests for cointegration or integration.