# Understanding Equation Balance in Time Series Regression: An Extension*

Peter K. Enns
peterenns@cornell.edu
Associate Professor
Department of Government
Cornell University

Christopher Wlezien
wlezien@austin.utexas.edu
Hogg Professor of Government
University of Texas at Austin

**Abstract:** Most contributors to a recent *Political Analysis* symposium on time series analysis suggest that in order to maintain equation balance, one cannot combine stationary, integrated and/or fractionally integrated variables with general error correction models (GECMs) and the equivalent autoregressive distributed lag (ADL) models. This definition of equation balance implicates most previous uses of these models in political science and circumscribes their use moving forward. The claim thus is of real consequence and worthy of empirical substantiation, which the contributors did not provide. Here we address the issue. First, we highlight the difference between estimating unbalanced equations and mixing orders of integration, the former of which clearly is a problem and the latter of which is not, at least not necessarily. Second, we assess some of the consequences of mixing orders of integration by conducting simulations using stationary, integrated, and combined (stationary plus integrated) time series. Our simulations show that with an appropriately specified model, regressing a stationary variable on an integrated one or the reverse does not increase the risk of spurious results and that such regressions can detect true relationships when they exist. We then illustrate the potential importance of these conclusions with an applied example—income inequality in the United States.

*This is a revised and extended version of an article published in *The Political Methodologist*. A previous version of this paper was presented at the Texas Methods Conference, 2017. We would like to thank Neal Beck, Patrick Brandt, Harold Clarke, Justin Esarey, John Freeman, Nate Kelly, Jamie Monogan, Andy Philips, Mark Pickup, Pablo Pinto, Randy Stevenson, Thomas Volscho, the *PSRM* editor, Vera Troeger, and two anonymous reviewers for helpful comments and suggestions.

*Political Analysis* (PA) recently hosted a symposium on time series analysis that built upon De Boef and Keele's (2008) influential time series article in the *American Journal of Political Science*. Equation balance was an important point of emphasis throughout the symposium. In their classic work on the subject, Banerjee, Dolado, Galbraith and Hendry (1993, 164) explain that an unbalanced equation is a regression, "in which the regressand is not the same order of integration as the regressors, or any linear combination of the regressors." The contributors to this symposium were right to emphasize the importance of equation balance, as unbalanced equations can produce serially correlated residuals (e.g., Pagan and Wickens 1989) and spurious relationships (e.g., Banerjee et al. 1993, 79).

Throughout the *PA* symposium, however, equation balance is defined and applied in different ways. Grant and Lebo (2016, 7) follow Banerjee, et al's definition when they explain that a general error correction model (GECM)—or autoregressive distributed lag (ADL)—is balanced if co-integration is present.[1] Keele, Linn and Webb (2016*a*, 83) implicitly make this same point in their second contribution to the symposium when they cite Bannerjee et al. (1993) in their discussion of equation balance. Yet, other parts of the symposium seem to apply a stricter standard of equation balance, stating that when estimating a GECM/ADL all time series must be the same order of integration. As Grant and Lebo write in the abstract of their first article, "Time series of various orders of integration—stationary, non-stationary, explosive, near- and fractionally integrated—should not be analyzed together... That is, without equation balance the model is misspecified and hypothesis tests and long-run-multipliers are unreliable." Keele, Linn and Webb (2016*b*, 34) similarly write, "no regression model is appropriate when the orders of integration are mixed because no long-run relationship can exist when the equation is unbalanced." Box-Steffensmeier and Helgason (2016, 2) make the point by stating, "when studying the relationship between two (or more) series, the analyst must ensure that they are of the same level of integration; that is, they have

[1]The GECM and ADL are the same model (e.g., Banerjee et al. 1993, De Boef and Keele 2008, Esarey 2016). However, since the two models estimate different quantities of interest (Enns, Kelly, Masaki and Wohlfarth 2016), they are often discussed as two separate models.

to be balanced." Although Freeman (2016) offers a more nuanced perspective on equation balance, many of the symposium contributors could be interpreted as recommending that scholars never mix orders of integration.[2] Indeed, in their concluding article, Lebo and Grant write, "One point of agreement among the papers here is that equation balance is an important and neglected topic. One cannot mix together stationary, unit-root, and fractionally integrated variables in either the GECM or the ADL" (p.79).

It is possible that these authors did not mean for these quotes to be taken literally. However, we both have recently been asked to review articles that have used these quotes to justify analytic decisions with time series data.[3] Thus, we think the claims should be reviewed carefully. This is especially the case because Grant and Lebo could be interpreted as applying these strict standards in some of their empirical applications. For example, in their discussion of Sánchez Urribarrí, Schorpp, Randazzo and Songer (2011), Grant and Lebo write, "both the UK and US models are unbalanced—each DV is stationary, and the inclusion of unit-root IVs has compromised the results" (Supplementary Materials, p.36). Researchers might take this statement to imply that including stationary and unit root variables automatically produces an unbalanced equation.

In addition to holding implications for practitioners, the strict interpretation of equation balance holds implications for the vast number of existing time series articles that employ GECM/ADL models without pre–whitening the data to ensure equal orders of integration across all series. Lebo and Grant (2016, 79) point out, for example, "FI [fractional integration] methods allow us to create a balanced equation from dissimilar data. By filtering each series by its own $(p, d, q)$ noise model, the residuals of each can be rendered (0, 0, 0) so that you can investigate how $X$'s deviations from its own time-dependent patterns affect $Y$'s

[2] Specifically, Freeman (2016, 50) explains, "KLWs [Keele, Linn, and Webb] claim that unbalanced equations are 'nonsensical' (16, fn. 4) and GLs [Grant and Lebo] recommendation to 'set aside' unbalanced equations (7) are a bit overdrawn. Banerjee et al. (1993) and others discuss the estimation of unbalanced equations. They simply stress the need to use particular nonstandard distributions in these cases."

[3] Given the prominence of the authors as well as the *Political Analysis* journal, it is perhaps not surprising that practitioners have begun to adopt these recommendations.

deviations from its own time-dependent patterns." Fortunately, existing time series analysis that does not pre-whiten the data need not be automatically dismissed. The strict interpretation of equation balance—i.e., that mixing orders of integration is always problematic with the GECM/ADL—is not accurate. As noted above, the contributors to the symposium may indeed understand this point. But based on the quotes above, we feel that it is important to clarify for practitioners that an unbalanced equation is not synonymous with mixing orders of integration. While related, they are not the same, and while the former is always a problem the latter is not.

We begin by showing that equation balance does not necessarily require that all series have the same order of integration with the GECM/ADL. This is important because the classic examples in the literature of unbalanced equations include series of different orders of integration (see, for example, Banerjee et al. (1993, 79,164), Maddala and Kim (1998, 252), Mankiw and Shapiro (1986)). But our results are not at odds with these scholars, as their examples all assume a relationship with no dynamics. When using a GECM/ADL to model dynamic processes, even mixed orders of integration can produce balanced equations. This conclusion is consistent with Banerjee et al. (1993), who write, "The moral of the econometricians' story is the need to keep track of the orders of integration on both sides of the regression equation, *which usually means incorporating dynamics*; models that have restrictive dynamic structures are relatively likely to give misleading inferences simply for reasons of inconsistency of orders of integration" (p.192, italics ours).

We believe the *PA* symposium was not sufficiently clear that adding dynamics can solve the equation balance problem with mixed orders of integration. Thus, a key contribution of our article is to show how appropriate model specification can be used to ensure equation balance and avoid inflating the rate of spurious regression—even when the model includes series with different orders of integration. Our particular focus is analysis that mixes stationary $I(0)$, integrated $I(1)$, and combined (stationary plus integrated) time series. In practice, researchers might encounter other types of time series, such as fractionally integrated, nearintegrated, or explosive series. Evaluating every type of time series and the vast number of ways different orders of integration could appear in a regression model is beyond the scope of this paper. Our goal is more basic, but still important. We aim to demonstrate that there are exceptions to the claim that, "The order of integration needs to be consistent across all series in a model" (Grant and Lebo 2016, 4) and that these exceptions can hold important implications for social science research.

More specifically, we show that when data are either stationary or (first order) integrated, scenarios exist when a GECM/ADL that includes both types of series can be estimated without problem. Our simulations show that regressing an integrated variable on a stationary one (or the reverse) does not increase the risk of spurious results when modeled correctly. While this may be a simple point, we think it is a crucial one. As mentioned above, if readers interpreted the previous quotes from the *PA* symposium as defining equation balance to mean that different orders of integration cannot be mixed, most existing research that employs the GECM/ADL model would be called into question. Given the fact that *Political Analysis* is one of the most cited journals in political science and the symposium included some of the top time series practitioners in the discipline, we believe it is valuable to clarify that mixing orders of integration is not always a problem and that existing time series research is not inherently flawed. Furthermore, the one article that has responded to particular claims made in the symposium contribution did not address the symposium's definition of equation balance (Enns et al. 2016).[4] We hope our article helps clarify the concept of equation balance for those who use time series analysis.

In addition to considering Type I errors, our simulations show that even with mixed orders of integration, as long as the equation is balanced, it is possible to identify true relationships between series with a GECM/ADL. This is an important and often overlooked step among those offering recommendations to time series practitioners. Although time

[4]Enns et al. (2016) focused on how to correctly implement and interpret the GECM, a topic we do not consider here.

series researchers are typically (and appropriately) most concerned with avoiding Type I errors, those recommending specific methods must also show that the proposed methods can identify true relationships in the data.

To illustrate the importance of our findings, we consider an applied example—income inequality in the United States. The example illustrates how the use of pre-whitening to force variables to be of equal orders of integration (when the equation is already balanced) can be quite costly, leading researchers to fail to detect relationships.[5]

## Clarifying Equation Balance

The contributors to the *PA* symposium were all correct to emphasize equation balance. Time series analysis requires a balanced equation. An unbalanced equation is mis-specified by definition, typically resulting in serially correlated residuals and an increased probability of Type I errors.[6] As noted above, Banerjee et al. (1993, 164, italics ours) explain that an unbalanced equation is a regression, "in which the regressand is not the same order of integration as the regressors, *or any linear combination of the regressors*." Our primary concern is that much of the discussion in the *PA* symposium seems to focus on the order of integration of *each* variable in the equation without acknowledging that a "linear combination of the regressors" can also produce equation balance. We worry that researchers might interpret this focus to mean that equation balance requires each series in the model to be the same order of integration. Such a conclusion would be wrong. As the previous quote from Banerjee et al. (1993) indicates (also see, Maddala and Kim (1998, 251), if the regressand and the regressors are *not* the same order of integration, the equation will still be balanced if a *linear combination* of the variables is the same order of integration.

[5]Of course, equation balance is not the only relevant consideration. Researchers must check that their model satisfies other assumptions, such as no autocorrelation in the residuals and no omitted variables.

[6]See, e.g., Banerjee et al. (1993, 164-168), Maddala and Kim (1998, 251-252), Mankiw and Shapiro (1986), and Pagan and Wickens (1989, 1002).

As Grant and Lebo (2016, 7) and Keele, Linn and Webb (2016*a*, 83) acknowledge, cointegration offers a useful illustration of how an equation can be balanced even when the regressand and regressors are *not* the same order of integration.[7] Consider two integrated $I(1)$ variables, $Y$ and $X$, in a standard GECM model:

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

Clearly, the equation mixes orders of integration. We have a stationary regressand ($\Delta Y_t$) and a combination of integrated ($Y_{t-1}$, $X_{t-1}$) and stationary ($\Delta X_t$) regressors. However, if $X$ and $Y$ are cointegrated, the equation is still balanced. To see why, we can rewrite Equation 1 as:

$$\Delta Y_t = \alpha_0 + \alpha_1 (Y_{t-1} + \frac{\beta_1}{\alpha_1} X_{t-1}) + \gamma_1 \Delta X_t + \epsilon_t. \quad (2)$$

$X$ and $Y$ are cointegrated when $X$ and $Y$ are both integrated (of the same order) and $\alpha_1$ and $\beta_1$ are non-zero (and $\alpha_1 < 0$). Because cointegration ensures that $Y$ and $X$ maintain an equilibrium relationship, a linear combination of these variables exists that is stationary (that is, if we regress $Y$ on $X$, in levels, the residuals would be stationary).[8] As noted above, this (stationary) linear combination is captured by $(Y_{t-1} + \frac{\beta_1}{\alpha_1} X_{t-1})$. Additionally, since $Y$ and $X$ are both integrated of order one, $\Delta Y$ and $\Delta X$ will be stationary. Thus, cointegration ensures that the equation is balanced: the regressand ($\Delta Y_t$) and either the regressors ($\Delta X_t$) or a linear combination of the regressors $(Y_{t-1} + \frac{\beta_1}{\alpha_1} X_{t-1})$ are *all* stationary. Importantly, if we added a stationary regressor to the model, e.g., if we thought innovations in $Y$ were also influenced by a stationary variable, the equation would still be balanced.

The fact that the GECM—which mixes stationary and integrated regressand and regressors—is appropriate when cointegration is present demonstrates that equation balance does not require the series to be the same order of integration. As we have mentioned, Grant and

[7] See Murray (1994) for a discussion of cointegration.

[8] This, in fact, is the first step of the Engle-Granger two-step method of testing for cointegration.

Lebo (2016, 7) acknowledge that a GECM is balanced if co-integration is present and Keele, Linn and Webb (2016*a*, 83) make this point in their second contribution to the symposium citing Bannerjee et al. (1993) in their discussion of equation balance. However, as noted above, we have begun to encounter research that interprets other statements in the symposium to mean that analysts can never mix orders of integration. For example, in their discussion of Volscho and Kelly (2012), Grant and Lebo write that the "data is a mix of data types (stationary and integrated), so any hypothesis tests will be based on unbalanced equations" (supplementary appendix, p. 48). But is this this really the case? The above example shows that when cointegration is present, equation balance can exist even when the orders of integration are mixed.

Below, we use simulations to illustrate three seemingly less well-known scenarios when equation balance exists despite different orders of integration. Again, our goal is not to identify all cases where different orders of integration can result in equation balance. Rather, we want to show that researchers should not automatically equate different orders of integration with an unbalanced equation. Situations exist where it is completely appropriate to estimate models with different orders of integration.

# Equation Balance with Mixed Orders of Integration: Simulation Results

We begin with an integrated $Y$ and a stationary $X$. At first glance, estimating a relationship between these variables, which requires mixing an $I(1)$ and $I(0)$ series, might seem problematic. Grant and Lebo (2016, 4) explain, "Mixing together series of various orders of integration will mean a model is misspecified" and in econometric texts, mixing $I(1)$ and $I(0)$ series offers a classic example of an unbalanced equation (Banerjee et al. 1993, 79, Maddala and Kim 1998, 252).[9]

[9] Interestingly, existing simulations show that despite being unbalanced regressions, we will *not* find evidence that unrelated I(0) and I(1) series are (spuriously) related in a simple bivariate regression if the I(0)

It is still possible to estimate the relationship between an integrated $Y$ and a stationary $X$ in a correctly specified and balanced equation. First, we must recognize that when Banerjee et al. (1993) (see also Mankiw and Shapiro (1986) and Maddala and Kim (1998)) state that an $I(1)$ and $I(0)$ series represent an unbalanced equation, they are modeling the equation:

$$y_t = \alpha + \beta x_{t-1} + u_t. \tag{3}$$

Equation 3 is indeed unbalanced (and thus misspecified) as the regressand is integrated and the regressor is stationary. This result does not, however, mean that we cannot consider these two series. A stationary series, $X$, might be related to *innovations* in an integrated series, $Y$. If so, we could model this process with an autoregressive distributed lag model:

$$Y_t = \alpha + \alpha_1 Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \epsilon. \tag{4}$$

Much as before, this might appear to still be an unbalanced equation. We continue to mix $I(1)$ and $I(0)$ series, which seemingly violates Lebo and Grant's (2016, 71) conclusion that, "One cannot mix together stationary, unit–root, and fractionally integrated variables in either the GECM or the ADL."[10] However, since $Y$ is $I(1)$, $\alpha_1 = 1$, which means $Y_t - \alpha_1 Y_{t-1} = \Delta Y$. Thus, we can rewrite the equation as,

$$\Delta Y_t = \alpha + \beta_1 X_t + \beta_2 X_{t-1} + \epsilon. \tag{5}$$

Because $Y$ is an integrated, $I(1)$, series, $\Delta Y$ must be stationary. Thus, the regressand and

variable is AR(0) (see, e.g., Banerjee et al. (1993, 79), Granger, Hyung and Jeon (2001, 901), and Maddala and Kim (1998, 252)). Banerjee et al. explain that the only way in which OLS can make the regression consistent and minimize the sum of squares is to drive the coefficient to zero (p.80). However, with AR(1) and I(1) series, our own simulations (not shown) confirm that when estimating unbalanced regressions both serial correlation and inflated Type I error rates emerge.

[10]Although fractionally integrated variables may also be of interest to researchers, this example focuses on stationary and integrated processes, which offer a clear illustration of the consequences of mixing orders of integration.

regressors in Equation 5 are all $I(0)$ series. As Banerjee et al. (1993, 169) explain, "regressions that are linear transformations of each other have identical statistical properties. What is important, therefore, is the *possibility* of transforming in such a way that the regressors are integrated of the same order as the regressand."[11] Thus, Equation 5 shows that the ADL in Equation 4 is indeed balanced. (Because the GECM is algebraically equivalent to the ADL, the GECM would—by definition—also be balanced in this example.[12])

The above discussion suggests that we can use an ADL to estimate the relationship between an integrated $Y$ and stationary $X$. To test these expectations, we conduct a series of Monte Carlo experiments. We generate an integrated $Y$ with the following DGP:

$$Y_t = Y_{t-1} + \epsilon_{yt},\ \epsilon_{yt}\, N(0,1). \tag{6}$$

We generate the stationary time series $X$, with the following DGP, where $\theta$ equals 0.0 or 0.5:

$$X_t = \theta_x X_{t-1} + \epsilon_{xt},\ \epsilon_{xt}\, N(0,1). \tag{7}$$

Notice that $X$ and $Y$ are independent series. Particularly with dependent series that contain a unit root (as is the case here), the dominant concern in time series literature is the potential for estimating spurious relationships (e.g., Granger and Newbold 1974, Grant and Lebo 2016, Yule 1926). Thus, our first simulations seek to identify the percentage of analyses that would *incorrectly* reject the null hypothesis of no relationship between a stationary $X$ and integrated $Y$ with an ADL. As noted above, in light of the recommendations in the *PA* symposium to never mix orders of integration, this approach seems highly problematic. However, if the equation is balanced as we suggest, the false rejection rate in our simulations should only be

[11] Banerjee et al. (1993) wrote this in the context of a discussion of equation balance among cointegrated variables, but the point applies equally well in this context.

[12] Despite the mathematical equivalence of the ADL and GECM, as noted above the parameters in these models estimate different quantities of interest (see, e.g., Enns, Masaki and Kelly 2014, Enns et al. 2016, De Boef and Keele 2008, Lebo and Grant 2016).

about 5 percent.

In the following simulations, $T$ is set to 50 and then 1,000. These values allow us to evaluate both a short time series that political scientists often encounter and a long time series that will approximate the asymptotic properties of the series. We use the DGP from Equations 6 and 7, above, to generate 1,000 simulated data sets. Recall that in our stationary series, $\theta_x$ equals 0.5 or 0.0 and $Y$ and $X$ are never related. To evaluate the relationship between $X$ and $Y$, we estimate an ADL model in Equation 4.[13]

Table 1: The Percent of Spurious Regressions with an Integrated $Y$ and Stationary $X$

| | $\theta_x = 0$ | | | | $\theta_x = 0.5$ | | | |
|---|---|---|---|---|---|---|---|---|
| $T =$ | $\hat{\beta}_1$ | % | $\hat{\beta}_2$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_2$ | % |
| 50 | .0032 | 4.2% | -.0021 | 5.4% | 0.0025 | 4.2% | -0.0035 | 4.3% |
| 1,000 | -.0011 | 4.8% | 0.0003 | 5.0% | -0.0011 | 4.7% | 0.0010 | 4.3% |

*Notes:* $\hat{\beta}$ represents the mean coefficient estimate across 1,000 simulations. % represents the percent of the simulations for which we *incorrectly* reject the null hypothesis of no relationship between X and Y. $\theta_x$ represents the autoregressive parameter in Equation 7.

Table 1 reports the average estimated relationship across all simulations between $X$ and $Y$ ($\hat{\beta}_1$ and $\hat{\beta}_2$ in Equation 4) and the percent of simulations in which these relationships were statistically significant. The mean estimated relationship is close to zero and the Type I error rate is close to 5 percent. With this ADL specification, when $Y$ is integrated and $X$ is stationary, mixing integrated and stationary time series does not increase the risk of spurious regression.[14]

Table 2 shows that the same pattern of results emerges when $X$ is integrated and $Y$ is stationary as the mean of $\hat{\beta}_1$ and $\hat{\beta}_2$ again are close to zero and the Type I error rate is

[13]The ADL is mathematically equivalent to the general error correction model (GECM), so the GECM would produce the same results, as long as the parameters are interpreted correctly (see Enns et al. 2016).

[14]The simulations reported in Table 1 also indicate that the ADL specification addresses the issue of serially correlated residuals, which would not be the case with an unbalanced regression. When $\theta_x$=0 and $T = 50$, a Breusch-Godfrey test rejects the null of *no* serial correlation just 6.5% of the time. When $T$=1,000, we find evidence of serially correlated residuals in just 4.6% of the simulations. When $\theta_x$=0.5, the corresponding rates are 6.4% ($T$=50) and 4.6% ($T$=1,000).

close to 5 percent.[15] Most time series analysis in the political and social sciences could be accused of mixing orders of integration. Thus, the recommendations of the *PA* symposium could be interpreted as calling this research into question. We have shown, however, that mixing orders of integration does not automatically imply an unbalanced equation. It also does not automatically lead to spurious results.[16]

Table 2: The Percent of Spurious Regressions with a Stationary $Y$ and an Integrated $X$

| | $\theta_y = 0$ | | | | $\theta_y = 0.5$ | | | |
|---|---|---|---|---|---|---|---|---|
| $T =$ | $\hat{\beta}_1$ | % | $\hat{\beta}_2$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_2$ | % |
| 50 | -.0023 | 6.0% | .0013 | 5.5% | -0.0034 | 6.4% | 0.0015 | 6.0% |
| 1,000 | .0011 | 5.6% | -0.0010 | 5.3% | 0.0011 | 5.5% | -0.0009 | 5.2% |

*Notes:* $\hat{\beta}$ represents the mean coefficient estimate across 1,000 simulations. % represents the percent of the simulations for which we *incorrectly* reject the null hypothesis of no relationship between X and Y. $\theta_y$ represents the autoregressive parameter in the DGP of $Y$.

## Identifying True Relationships with Different Orders of Integration

The previous simulations demonstrated that unrelated stationary and (first order) integrated time series can (in some cases) be analyzed together with an ADL model without concerns for spurious regression. In this section, we evaluate whether an ADL model can identify a true relationship between series that are different orders of integration. Specifically, we consider a regression model with a stationary variable on the right hand side and a dependent variable that includes both stationary and unit root components. Wlezien (2000) refers to such a variable as a "combined" time series process. Here, the shock to a combined time series, $e_t$, can be separated into an integrated component $e_t^I$ that carries over indefi-

[15] The fact that we do not observe evidence of an increased rate of spurious regression in Table 2, particularly when $Y$ is AR(1), implies that we do not have an equation balance problem. We also find that the simulations in Table 2 tend not to produce serially correlated residuals (we only reject the null of no serial correlation in 6.3% and 5.2% of simulations when $T$=50 and 4.9% and 3.3% of simulations when $T$=1,000).

[16] Also note that that the mean estimates for $\alpha_1$ in the simulations reflect the DGP for $Y$, particularly as $T$ increases. With integrated $Y$, per Table 1, $\hat{\alpha_1}$=0.90 and 0.89 when $T$=50 and 0.99 and 0.99 when $T$=1,000. With stationary $Y$, per Table 2, $\hat{\alpha_1}$=-0.04 for $\theta_y = 0.0$ and 0.42 for $\theta_y = 0.5$ when $T$=50 and $\hat{\alpha_1}$=-0.002 and 0.496, respectively, when $T$=1,000.

nitely and a stationary component $e_t^S$ that decays (Wlezien 2000, 79). In theory, such series are integrated (Granger 1980).

There are many scenarios in the political and economic world that can produce combined time series. To begin with, consider that any process that includes long-term change and measurement error is such a series. Putting aside measurement error, there are reasons to suppose that numerous processes combine both long-term and short-term change. Theories of people's attitudes over time reflect distinctions between long-term stability versus short-term change (Converse 1964, Achen 1975, Erikson 1979). Characterizations of party identification also reflect these distinctions, and some scholars (Erikson, MacKuen and Stimson 1998) explicitly conceive of macro-partisanship as a combined process. The same is true for electoral preferences, which clearly change over time, some of which lasts to impact the outcome and some does not (Erikson and Wlezien 2012). There are numerous other such cases. Combined time series are common—and important—for political research. They also are ideally suited for GECMs, which estimate both short run (stationary) and long run (integrated) components of the combined time series.

Just as important for our purposes, however, the DGP of combined time series allow us to conduct simulations where the left and right hand side variables are related and of different orders of integration. For our simulations,

$$Y_t = e_t^I + e_t^S \tag{8}$$

$$e_t^I = e_{t-1}^I + u1_t,\; u1_t\, N(0,1) \tag{9}$$

$$e_t^S = \rho e_{t-1}^S + u2_t,\; u2_t\, N(0,1) \tag{10}$$

where $\rho$ equals 0.2, 0.5, or 0.8.

Because we want to evaluate whether the ADL can recover true relationships when the orders of integration on the right and left hand side of the equation are mixed, we estimate the equation,

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_1 e_t^S + \beta_2 e_{t-1}^S + \delta. \tag{11}$$

We do not include $e_t^I$ in the regression equation, which means we are mixing a combined time series, $Y$, which in theory is integrated (Granger 1980), with a stationary time series, $e^S$. Notice that by omitting $e_t^I$, we are creating an omitted variable bias problem. Clearly, when analyzing combined time series (as with all data types), researchers should aim to model all explanatory factors. Our goal, however, is to model related series that are of different orders of integration and omitting $e_t^I$ ensures this scenario. Also keep in mind that the equation is one we might estimate in the absence of knowledge about the true data generating process. Thus, although Equation 11 would not be recommended if we knew the true DGP, this specification allows us to evaluate the performance of the GECM/ADL with mixed orders of integration and a common but imperfect specification (which reflects the realities researchers often face).

Since the true relationship between $e_t^S$ and $Y_t$ is 1.0, in our simulations we expect $\hat{\beta}_1$ to equal 1.0. (To be clear, the equation reveals the contribution the independent variable makes to our outcome variable, not the autoregressive parameter of the component.) Relatedly, we expect $\hat{\beta}_2$ to equal $-1.0$. That this is true can be seen by substituting for $Y_{t-1}$, which equals $e_{t-1}^I + e_{t-1}^S$, as follows,

$$Y_t = \alpha_0 + \alpha_1[e_{t-1}^I + e_{t-1}^S] + \beta_1 e_t^S + \beta_2 e_{t-1}^S. \tag{12}$$

Assuming $\alpha_1 = 1$, the equation reduces to

$$Y_t = \alpha_0 + e_{t-1}^I + e_{t-1}^S + \beta_1 e_t^S + \beta_2 e_{t-1}^S. \tag{13}$$

Given that $Y_t = e_t^I + e_t^S$, by construction, we expect $\beta_2$ to equal $-\beta_1$, which cancels the portion of $e_{t-1}^S$ in $Y_{t-1}$ (since $e_{t-1}^S$ does not enter the DGP in Equation 8). Notice that this generalizes across combinations of $e_t^I$ and $e_t^S$. For example, were the true relationship between $e_t^S$ and $Y_t$ equal to 2.0 (and $e_t^I$ and $Y_t$ equal to 1.0), we would expect $\beta_1$ to equal 2 and $\beta_2$ to equal -2, which would again cancel out the portion of $e_{t-1}^S$ in $Y_{t-1}$. Finally, we expect $\alpha_0$ to equal 0.0.

Considering the statements regarding not mixing orders of integration in the *PA* symposium, scholars might conclude that the Equation 11 should not be estimated. However, despite different orders of integration on both sides of the equation, it is easy to see that the equation is indeed balanced. Substituting Equation 8 for $Y$, Equation 11 can be rewritten as follows,

$$e_t^S + e_t^I = e_{t-1}^S + e_{t-1}^I + \beta_1 e_t^S + \beta_2 e_{t-1}^S + \delta. \tag{14}$$

By rearranging Equation 14, in Equation 15, we see that we now have stationary series on both sides of the equation.

$$\Delta e_t^I + e_t^S = \beta_1 e_t^S + (1 + \beta_2) * e_{t-1}^S + \delta. \tag{15}$$

Thus, we would expect to be able to identify the true relationships between $e^S$ and $Y$ described above. This does not mean that estimating equation 11 will correctly represent the data generating process (DGP), as it clearly does not, since we omit $e_t^I$ to ensure that we mix orders of integration. We are intending only to illustrate that, just as an ADL (or GECM) does not necessarily induce spurious results when orders of integration are mixed, it can reveal true relationships with mixed orders of integration, at least when the equation is balanced.

Table 3 presents the results of our simulations.[17] The results are as expected.[18] This is most clear in the far-right columns of the table, where T=5,000. There we can see that we always accept the hypothesis of a relationship and the mean coefficients are precisely what we expect. First, $\hat{\alpha}_1$=1.0, which is exactly as we expect with a combined time series that includes a unit root. Second, consistent with Equation 8, $\hat{\beta}_1 = 1$. Finally, per our discussion of Equation 13 above, $\hat{\beta}_2 = -1$, which ensures the effects of $e^S_{t-1}$ cancel out on the right-hand side. These results holds regardless of the $\rho$ of the stationary component. Things are only a little less clear when T=100 and a little less clear still when dropping to T=50. We nevertheless tend to find exactly what we expect on average given the construction of the series, once again regardless of the $\rho$ of the stationary component. Interestingly, when we see departures from the mean expected value, the departures for the $Y_{t-1}$ and $X_{t-1}$ are equal and opposite. Ultimately, despite different orders of integration on the two sides of the equation, we are able to correctly identify the relationship between the stationary $e^S_t$ and the combined time series $Y$ (which in theory is integrated), even with fairly finite samples.[19] As discussed above, Equation 11 is misspecified by construction, because it omits $e^I_t$.[20] Clearly, theory must guide model specification and it would be wrong to conclude that these results imply that the GECM/ADL is always appropriate. They key point is that even with mixed orders of integration, we find that the ADL in Equation 11 is balanced and avoids Type I errors and also produces serially independent residuals.

[17] The appendix reports analogous results where a disturbance term, $q$, is added to the DGP of $Y_t$.

[18] Across all simulations, Breusch-Godfrey tests indicate that we reject the null hypothesis of no serial correlation between 4.2 and 6.3 percent of the time.

[19] Separate analyses indicate that this result generalizes across combinations of the integrated and stationary series.

[20] Thus, any correlation between $u1_t$ and $u2_t$ in Equations 9 and 10 would bias the estimate of $\beta_1$ in Equation 11.

Table 3: Identifying a True Relationship ($\beta_1$=1.0) between $X$ and $Y$ when $X$ is Stationary and $Y$ Contains Stationary *and* Unit Root Properties

| | $T = 50$ | | | | | | $T = 100$ | | | | | | $T = 5,000$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | |
| | coef | % | coef | % | coef | % | coef | % | coef | % | coef | % | coef | % | coef | % | coef | % |
| $y_{t-1}$ ($\hat{\alpha}_1$) | 0.89 | 100 | 0.89 | 100 | 0.88 | 100 | 0.95 | 100 | 0.95 | 100 | 0.94 | 100 | 1.00 | 100 | 1.00 | 100 | 1.00 | 100 |
| $x1_t$ ($\hat{\beta}_1$) | 1.00 | 100 | 1.00 | 100 | 1.00 | 100 | 1.00 | 100 | 1.01 | 100 | 1.00 | 100 | 1.00 | 100 | 1.00 | 100 | 1.00 | 100 |
| $x1_{t-1}$ ($\hat{\beta}_2$) | -0.89 | 99.8 | -0.90 | 99.5 | -0.89 | 99.2 | -0.94 | 100 | -0.95 | 100 | -0.94 | 100 | -1.00 | 100 | -1.00 | 100 | -1.00 | 100 |

*Notes:* coef represents the mean estimate of $\alpha_1$, $\beta_1$, or $\beta_2$ across 1,000 simulations. % represents the percent of the simulations for which we *correctly* reject the null hypothesis of no relationship between X and Y.

## The Rise of the Super Rich: Reconsidering Volscho and Kelly (2012)

We think the foregoing discussion and analyses offer compelling evidence that, despite the range of statements about equation balance in the *PA* symposium, mixing orders of integration when using a GECM/ADL does not automatically pose a problem to researchers. Of course, to a large degree the previous sections reiterate and unpack what econometricians have shown mathematically (e.g., Sims, Stock and Watson 1990), and so may come as little surprise to some readers (especially those who have not read the *PA* Symposium). Here, we use an applied example to illustrate the importance of correctly understanding equation balance. We turn to a recent article by Volscho and Kelly (2012) that analyzes the rapid income growth among the super-rich in the United States (US). They estimate a GECM of pre-tax income growth among the top 1% and find evidence that political, policy, and economic variables influence the proportion of income going those at the top. Critically for our purposes, they include stationary and integrated variables on the right-hand side, which Grant and Lebo (2016, 26) actually single out as a case where the "GECM model [is] inappropriate with mixed orders of integration." Grant and Lebo go on to assert that Volscho and Kelly's "data is a mix of data types (stationary and integrated), so any hypothesis tests will be based on unbalanced equations" (supplementary appendix, p. 48). Based on the conclusion that mixing orders of integration produces an unbalanced equation, Grant and Lebo employ fractional error correction technology and find that none of the political or policy variables (and only some economic variables) matter for incomes among the top 1%.

These are very different findings, ones with potential policy consequences, and so it is important to reconsider what Volscho and Kelly did—and whether the mixed orders of integration pose a problem for their analysis. To begin our analysis, we present the dependent variable from Voschlo and Kelly, the total pre-tax income share of the top 1% for the period between 1913 and 2008.[21] In Figure 1 we can see that income shares start off quite high and then drop and then return to inter-war levels toward the end of the series. The variable thus exhibits none of the trademarks of a stationary series, i.e., it is not mean-reverting, and looks to contain a unit root instead. Notice that the same is true for the shorter period encompassed by Volscho and Kelly's analysis, 1949-2008. Augmented Dickey-Fuller (ADF) and Phillips–Perron unit root tests confirm these suspicions, and are summarized in the first row of Table 4, below.[22]

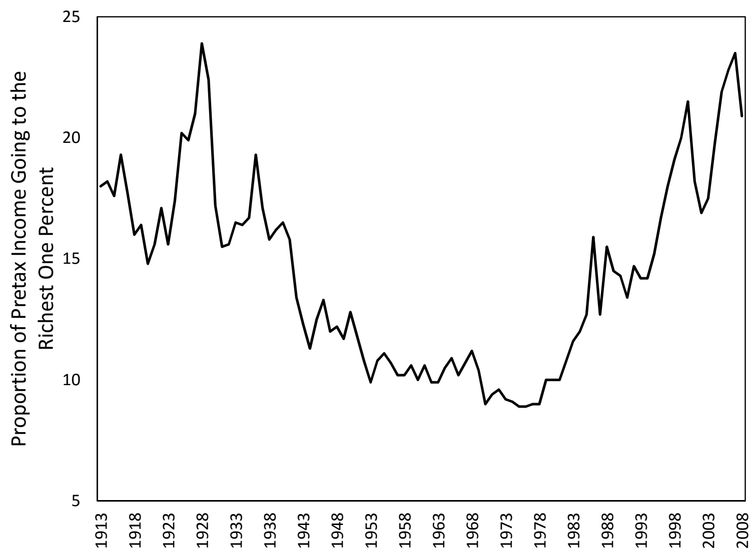


Figure 1: The Top 1 Percent's Share of Pre-tax Income in the United States, 1913 to 2008

What about the independent variables? Here, we find a mix (see Table 4). Some variables clearly are nonstationary and also appear to contain unit roots: the capital gains tax rate, union membership, the Treasury Bill rate, Gross Domestic Product (logged), and the Standard and Poor 500 composite index. The top marginal tax rate also is clearly nonstationary and we cannot reject a unit root even when taking into account the secular (trending) decline over time.[23] The results for the Shiller Home Price Index are mixed and trade openness is on

[21]These data, which come from Voschlo and Kelly, were originally compiled by Piketty and Saez (2003).

[22]These results are consistent with the unit root tests Volscho and Kelly report in the supplementary materials to their article. Grant and Lebo's analysis also supports this conclusion. In their supplementary appendix, Grant and Lebo estimate the order of integration $d$=0.93 with a standard error of (0.10), indicating they cannot reject the null hypothesis that $d$=1.0.

[23]Although the dependent variable is pre-tax income, Volscho and Kelly identify several mechanisms that could lead tax rates to influence pre-tax income share (also see, Mertens 2015, Piketty, Saez and

the statistical cusp, and there is reason—based on the size of the autoregressive parameter (-0.29) and the fact that we reject the unit root over a longer stretch of time—to assume that the variable is stationary. For the other variables included in the analysis, we reject the null hypothesis of a unit root: Democratic president, and the Percentage of Democrats in Congress. The pattern of results comports with what Volscho and Kelly found (see their supplementary materials), and is not at all unreasonable, as we expect many of the economic variables to be integrated and the political variables to be stationary.

Table 4: Time Series Properties of Variables Analyzed by Volscho and Kelly

| Variable | ADF test | Phillips-Perron test | Conclusion |
|---|---|---|---|
| Top 1% Share | 0.5121 | 0.5794 | Integrated |
| Democratic President | 0.0462 | 0.0254 | Stationary |
| Divided Government | 0.0082 | 0.0034 | Stationary |
| Top Marginal Tax Rate | 0.1563 | 0.3351 | Integrated |
| Capital Gains Tax Rate | 0.4220 | 0.5812 | Integrated |
| 3-Month Treasury Bill | 0.3654 | 0.2712 | Integrated |
| Trade Openness | 0.0608 | 0.0597 | Borderline |
| Log Real GDP | 0.2690 | 0.2323 | Integrated |
| Real S&P 500 Index | 0.2149 | 0.6427 | Integrated |
| Shiller Home Price Index | 0.0000 | 0.5616 | Unclear |
| Union Membership | 0.6710 | 0.8185 | Integrated |
| % Congressional Dem. | 0.0209 | 0.0147 | Stationary |

*Notes:* The Null hypothesis for ADF and Phillips-Perron tests is a unit root. When statistically significant, a trend and/or additional lags were included in the tests.

Volscho and Kelly proceed to estimate a GECM of the top 1% income share including current first differences and lagged levels of the stationary and integrated variables. So far, the diagnostics (integrated DV and some IVs are integrated) support their decision. However, given the inclusion of integrated variables on both sides of the equation, the GECM is only appropriate if cointegration is present. Some might wonder about the expectation of a cointegrating relationship between a range of economic variables. We are not absolutely sure

Stantcheva 2014). Based on existing research, it also would not be surprising if we observed evidence of a relationship between the top 1 percent's income share and union strength (for recent examples, see Jacobs and Myers 2014, Pontusson 2013, Western and Rosenfeld 2011) and the partisan composition of government (Bartels 2008, Hibbs 1977, Kelly 2009).

based on Volscho and Kelly, but, as discussed above, there is reason to expect those variables to be integrated and also related to income shares. Furthermore, Volscho and Kelly's (2012) preferred specification (Model 5), which we replicate in Column 1 of Table 5, shows evidence of cointegration.[24]

If the integrated variables Volscho and Kelly analyzed form a cointegrating relationship, the the inclusion of stationary variables on the right hand side would not affect equation balance. This expectation follows our discussion of the GECM in Section 2 (first-differencing an integrated variable produces a stationary regressand and the cointegrating relationship on the right hand side produces a linear combination of variables that is stationary). Of course, we cannot know the true DGPs the variables Volschlo and Kelly analyzed. We can, however, use an alternate modeling approach that is designed for mixed orders of integration. Specifically, we re-consider their data with Pesaran and Shin's ARDL (Autoregressive Distributed Lag) critical bounds testing approach (Pesaran, Shin and Smith 2001).

For their bounds test of cointegration, Pesaran and Shin estimate the model as a GECM.[25] The ARDL approach is one of the approaches recommended by Grant and Lebo and is especially advantageous in the current context because two critical values are provided, one which assumes all stationary regressors and one which assumes all integrated regressors. Values in between these "bounds" correspond to a mix of integrated and stationary regressors, meaning the bounds approach is especially appropriate when the analysis includes both types of

[24]When the dependent and independent variables contain a unit root and the correct Ericsson and MacKinnon (2002) critical values are used, the t-statistic associated with the error correction parameter ($Y_{t-1}$) in the GECM can be used to test for cointegration (Note that when $Y$ does not contain a unit root, the coefficient on $Y_{t-1}$ *cannot* be used as a test of cointegration (see, e.g., Enns, Masaki and Kelly 2014, Enns et al. 2016, Enns, Kelly, Masaki and Wohlfarth 2017, Lebo and Grant 2016)). With this approach, the absolute value of the t-statistic associated with the error correction parameter (6.5) exceeds the absolute value of the critical value from Ericsson and MacKinnon (2002, 304), when T=60 and with 12 independent variables (5.11). Although Ericsson and MacKinnon's (2002) critical values vary according to the number of independent variables, the values they calculate assume integrated predictors. Although some research suggests that stationary predictors do not affect the rate of spurious regressions (Enns, Masaki and Kelly 2014), this scenario offers another reason to consider the ARDL approach discussed below.

[25]Although political scientists typically refer to the autoregressive distributed lag model as an ADL, Pesaran, Shin and Smith (2001) prefer ARDL. Thus, the ADL, ARDL, and GECM all refer to equivalent models

regressors. Grant and Lebo (2016, 19) correctly acknowledge that "With the bounds testing approach, the regressors can be of mixed orders of integration—stationary, non-stationary, or fractionally integrated—and the use of bounds allow the researcher to make inferences even when the integration of the regressors is unknown or uncertain."[26] Since Table 4 indicates we have a mix of stationary and integrated regressors, if our critical value exceeds the highest bound, we will have evidence of cointegration.

The ARDL approach proceeds in several steps.[27] First, if the dependent variable is integrated, the ARDL model (which is equivalent to the GECM) is estimated. Next, if the residuals from this model are stationary, an F-test is conducted to evaluate the null hypothesis that the combined effect of all lagged variables in the model equals zero. This F statistic is compared to the appropriate critical values (Pesaran, Shin and Smith 2001). We rely on the small-sample critical values from Narayan (2005). If the F statistic shows evidence of cointegration, Pesaran, Shin and Smith (2001) recommend a bounds t-test to further evaluate this conclusion. If there is evidence of cointegration, both long and short-run relationships from the initial ARDL (i.e., GECM/ADL) model can be evaluated.

Our analysis focuses on Column 5 from Volscho and Kelly's Table 1, which is their preferred model. The first column of our Table 5, below, shows that we successfully replicate their results. The ARDL analysis appears in Column 2.[28] The key difference between this specication and that of Volscho and Kelly's is that they (based on a Breusch-Godfrey test) employed the Prais-Winsten estimator to correct for serially correlated errors and we do not. Our decision reflects the fact that other tests do not reject the null of white noise, e.g., the Portmanteau (Q) test produces a p-value of 0.12, and it allows us to compare the results with and without the correction. Also note that an expanded model including lagged

[26] It is not clear why Grant and Lebo seemingly contradict their statement that "Mixing together series of various orders of integration will mean a model is misspecified" (p.4) in this context, especially since the ARDL is equivalent to the GECM, but they are correct to do so.

[27] For a concise overview of the ARDL approach, see `http://davegiles.blogspot.ca/2013/06/ardl-models-part-ii-bounds-tests.html`.

[28] We exactly follow their lag structure and the assumption of a single endogenous variable, which seemingly is incorrect but possibly intractable.

differenced dependent and independent variables (see Appendix Table A-1) produces very similar estimates to those shown in column 2 of Table 4, and a Breusch-Godfrey test indicates that the resulting residuals are uncorrelated.

To begin with, we need to test for cointegration. For this, we compare the F-statistic from the lagged variables (6.54) with the Narayan (2005) upper ($I(1)$) critical value (3.82), which provides evidence of cointegration.[29] The bounds t-test also supports this inference, as the t-statistic (-7.75) for the $Y_{t-1}$ parameter is greater (in absolute terms) than the $I(1)$ bound tabulated by Pesaran, Shin and Smith (2001, 303). Returning to the results in Column 2, we see that the ARDL approach produces similar conclusions to Column 1. (Philips (2016) uses the ARDL approach to re-consider the first model in Volscho and Kelly's (2012) Table 1 and also obtains similar results.) The coefficients for all but two of the independent variables have similar effects, i.e., the same sign and statistical significance.[30] The exceptions are Divided Government$_{t-1}$ and $\Delta$Trade Openness, for which the coefficients using the two approaches are similar but the standard errors differ substantially. Consistent with the existing research on the subject, we find evidence that economics, politics, and policy matter for the share of income going to the top 1 percent.

Although Grant and Lebo (2016, 18) recommend both the ARDL approach and a three-step fractional error correction model (FECM) approach, they only report the results for the latter in their re-analysis of Volscho and Kelly.[31] It turns out that the two approaches

[29] The 5 percent critical value when T=60 with an unrestricted intercept and no trend is 3.823. Narayan (2005) only reports critical values for up to 7 regressors. However, the size of the critical value *decreases* as the number of regressors increases (Narayan 2005, Pesaran, Shin and Smith 2001), so our reliance on the the critical value based on 7 regressors is actually a conservative test of cointegration. We also tested for integration allowing for short-run effects of all integrated variables and we again find evidence of cointegration (F= 4.32).

[30] This reveals that explicitly taking into account serial correlation, which Volscho and Kelly did, has modest consequences.

[31] As Grant and Lebo (2016, 18) explain, the three–Step FECM proceeds as follows. First, $Y$ is regressed on $X$ and the residuals are obtained. The fractional difference parameter, $d$, is then estimated for each of the three series ($Y$, $X$, and the residuals). Grant and Lebo explain that if $d$ for the residuals is less than $d$ for both $X$ and $Y$, then error correction is occurring. If this is the case, the researcher then fractionally differences $Y$, $X$, and the residual by each ones own $d$ value. Finally, the researcher regresses the fractionally differenced $Y$ and the fractionally differenced $X$, and the lag of the fractionally differenced residual (Grant and Lebo 2016, 18). This regression produces the results reported in Column 3 of Table 5.

Table 5: Replication of Volscho & Kelly (2012) Table 1, Column 5 with the ARDL Bounds Test and the Three-Step FECM employed by Grant and Lebo

| | (1)<br>V&K<br>Replication | (2)<br>ARDL | (3)<br>G&L 3-Step<br>FECM | |
|---|---|---|---|---|
| Top 1% Share$_{t-1}$ | -0.65* | -0.93* | | |
| | (0.10) | (0.12) | | |
| % Congressional Dem.$_{t-1}$ | -0.05* | -0.06* | $\Delta^d$ % Congressional Dem. | -0.01 |
| | (0.02) | (0.02) | | (0.04) |
| Divided Government$_{t-1}$ | -0.37* | -0.42 | $\Delta^d$ Divided Government | 0.15 |
| | (0.17) | (.24) | | (0.36) |
| Union Membership$_{t-1}$ | -0.28* | -0.41* | $\Delta^d$ Union Membership | -0.06 |
| | (0.07) | (0.09) | | (0.26) |
| Top Marginal Tax Rate$_{t-1}$ | -0.03* | -0.05* | $\Delta^d$ Top Marginal Tax Rate | -0.02 |
| | (0.01) | (0.02) | | (0.03) |
| $\Delta$ Capital Gains Tax Rate | -0.03 | -0.05 | $\Delta^d$ Capital Gains Tax Rate | -0.07 |
| | (0.03) | (0.04) | | (0.05) |
| Capital Gains Tax Rate$_{t-1}$ | -0.06* | -0.08* | | |
| | (0.02) | (0.02) | | |
| 3-Month Treasury Bill$_{t-1}$ | 0.01 | 0.02 | $\Delta^d$ 3-Month Treasury Bill | -0.13 |
| | (0.04) | (0.06) | | (0.14) |
| $\Delta$ Trade Openness$_t$ | 0.20* | 0.22 | $\Delta^d$ Trade Openness | 0.38* |
| | (0.01) | (0.12) | | (0.19) |
| Log Real GDP$_{t-1}$ | -5.04* | -8.12* | $\Delta^d$ Log Real GDP | 3.35 |
| | (1.45) | (1.87) | | (6.89) |
| $\Delta$ Real S&P 500 Index$_t$ | 0.06* | 0.06* | $\Delta^d$ Real S&P 500 Index | 0.07* |
| | (0.01) | (0.01) | | (0.01) |
| Real S&P 500 Index$_{t-1}$ | 0.03* | 0.05* | | |
| | (0.01) | (0.01) | | |
| Shiller Home Price Index$_{t-1}$ | 0.28* | 0.43* | $\Delta^d$ Shiller Home Price Index | 0.32 |
| | (0.07) | (0.10) | | (0.27) |
| | | | FECM | -0.35* |
| | | | | (0.14) |
| Constant | 58.99* | 91.13* | | 0.01 |
| | (14.27) | (18.10) | | (0.18) |
| Adj. $R^2$ | 0.76 | 0.67 | | 0.36 |

*Notes:* Regression coefficients with standard errors in parentheses, * $p < .05$. The V&K replication (1) is based on Column 5 of their Table 1 and relies on the Prais-Winsten (GLS) estimator. The G&L results (3) come from the Supplementary Materials (p.50) to Grant and Lebo (2016). FECM reflects the long-term equilibrium relationship of both Trade Openness and Real S&P Composite Index and the other coefficients reflect estimated short run effects based on G&L's Three-Step FECM.

produce very different results. This can be seen in column 3 of Table 5, which reports Grant and Lebo's FECM reanalysis of Volscho and Kelly Model 5 (from Grant and Lebo's supplementary appendix, p. 50). With their approach, only the change in stock prices (Real S&P 500 Index) and Trade Openness are statistically significant ($p<.05$) predictors of income shares, though levels of stock prices and trade openness also matter via the FECM component, which captures disequilibria between those variables and lagged income shares. Despite theoretical and empirical evidence suggesting that the marginal tax rate (Mertens 2015, Piketty, Saez and Stantcheva 2014), union strength (Jacobs and Myers 2014, Pontusson 2013, Western and Rosenfeld 2011), and the partisan composition of government (Bartels 2008, Hibbs 1977, Kelly 2009) can influence the pre-tax income of the upper 1 percent, we would conclude that only trade openness and stock prices influence the pre-tax income share of richest Americans. Of course, analysts might reasonably prefer alternative models to the ones Volscho and Kelly estimate, perhaps opting for a more parsimonious specification, allowing endogenous relationships, and/or including alternate lag specifications and we hope future research on the income shares of the super rich takes these considerations into account. Future research should also consider the key points from our replication—i.e., mixing stationary and integrated variables does not necessarily produce an unbalanced equation and the ARDL and three–step FECM produce very different estimates.

## Conclusions

In his contribution to the *PA* symposium, John Freeman wrote, "It now is clear that equation balance is not understood by political scientists" (Freeman 2016, 50). Our goal has been to help clarify misconceptions about equation balance. In particular, we have shown that mixing orders of integration in a GECM/ADL model does *not* automatically lead to an unbalanced equation. The many quotes we have highlighted and the title of Lebo and Grant's second contribution to the symposium ("Equation Balance and Dynamic Political Modeling") illustrate that equation balance was an important theme of the symposium. Although others have responded to particular criticisms within the *PA* symposium (e.g., Enns et al. 2016), this article is the first to address the symposium's discussion and recommendations related to equation balance.

Because they are related, it is easy to (erroneously) conclude that mixing orders of integration is synonymous with an unbalanced equation. It would be wrong, however, to reach this conclusion. We have focused on three types of time series: stationary, unit–root, and combined series (ones that contain both stationary and unit–root properties) and we have found that situations exist when it is unproblematic—and inconsequential—to mix these types of series (because the equation is balanced).[32]

These results help clarify existing time series research (e.g., Banerjee et al. 1993, Sims, Stock and Watson 1990) by showing that when we use a GECM/ADL to model dynamic processes, even mixed orders of integration can produce balanced equations. The findings also lead to three recommendations for researchers. First, scholars should not automatically dismiss existing time series research that mixes orders of integration. Even when series are of different orders of integration or when the equation transforms variables in a way that leads to different orders of integration, the equation may still be balanced and the model correctly specified. In fact, we identified, and our simulations confirmed, specific scenarios when integrated, stationary, and combined time series can be analyzed together. Second, as we showed with our simulations and with our applied example, researchers must evaluate whether they have equation balance based on *both* the univariate properties of their variables *and* the model they specify. Third and finally, our results show that researchers do *not* always need to pre-whiten their data to ensure equation balance. Although pre-whitening time series will sometimes be appropriate, we have shown that this step is not a necessary condition for equation balance. This is important because such data transformations are potentially quite costly, specifically, in the presence of equilibrium relationships. As we saw above, Grant and Lebo's decision to pre-whiten Volscho and Kelly's data with their three-step FECM may be

[32]Of course, other statistical assumptions must also be satisfied.

one such example.

## References


Achen, Christopher. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218–1231.

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

Bartels, Larry M. 2008. *Unequal Democracy.* Princeton: Princeton University Press.

Box-Steffensmeier, Janet and Agnar Freyr Helgason. 2016. "Introduction to Symposium on Time Series Error Correction Methods in Political Science." *Political Analysis* 24(1):1–2.

Converse, Philip E. 1964. The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*, ed. David E. Apter. Ann Arbor: University of Michigan Press.

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

Enns, Peter K., Nathan J. Kelly, Takaaki Masaki and Patrick C. Wohlfarth. 2016. "Don't Jettison the General Error Correction Model Just Yet: A Practical Guide to Avoiding Spurious Regression with the GECM." *Research and Politics* 3(2):1–13.

Enns, Peter K., Nathan J. Kelly, Takaaki Masaki and Patrick C. Wohlfarth. 2017. "Moving Forward with Time Series Analysis." *Research and Politics* .

Enns, Peter K., Takaaki Masaki and Nathan Kelly. 2014. "Time Series Analysis and Spurious Regression: An Error Correction." Paper presented at the Annual Meeting of the Southern Political Science Association, New Orleans.

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

Erikson, Robert S. 1979. "The SRC Panel Data and Mass Political Attitudes." *British Journal of Political Science* 9:89–114.

Erikson, Robert S. and Christopher Wlezien. 2012. *The Timeline of Presidential Elections.* Chicago: University of Chicago Press.

Erikson, Robert S., Michael B. MacKuen and James A. Stimson. 1998. "What Moves Macropartisanship? A Reply to Green, Palmquist, and Schickler." *American Political Science Review* 92:901–912.

Esarey, Justin. 2016. "Fractionally Integrated Data and the Autodistributed Lag Model: Results from a Simulation Study." *Political Analysis* 24(1):42–49.

Freeman, John R. 2016. "Progress in the Study of Nonstationary Political Time Series: A Comment." *Political Analysis* 24(1):50–58.

Granger, Clive W.J., Namwon Hyung and Yongil Jeon. 2001. "Spurious Regressions with Stationary Series." *Applied Economics* 33:899–904.

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.

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

Hibbs, Jr., Douglas A. 1977. "Political Parties and Macroeconomic Policy." *American Political Science Review* 71(4):1467–1487.

Jacobs, David and Lindsey Myers. 2014. "Union Strength, Neoliberalism, and Inequality." *American Sociological Review* 79(4):752–774.

Keele, Luke, Suzanna Linn and Clayton McLaughlin Webb. 2016*a*. "Concluding Comments." *Political Analysis* 24(1):83–86.

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

Kelly, Nathan J. 2009. *The Politics of Income Inequality in the United States.* New York: Cambridge University Press.

Lebo, Matthew J. and Taylor Grant. 2016. "Equation Balance and Dynamic Political Modeling." *Political Analysis* 24(1):69–82.

Maddala, G.S. and In-Moo Kim. 1998. *Unit Roots, Cointegration, and Structural Change*. 1 ed. New York: Cambridge University Press.

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.

Mertens, Karel. 2015. "Marginal Tax Rates and Income: New Time Series Evidence." `https://mertens.economics.cornell.edu/papers/MTRI_september2015.pdf`.

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

Narayan, Paresh Kumar. 2005. "The Saving and Investment Nexus for China: Evidence from Cointegration Tests." *Applied Economics* 37(17):1979–1990.

Pagan, A.R. and M.R. Wickens. 1989. "A Survey of Some Recent Econometric Methods." *The Economic Journal* 99(398):962–1025.

Pesaran, Hashem M., Yongcheol Shin and Richard J. Smith. 2001. "Bounds Testing Approaches to the Analysis of Level Relationships." *Journal of Applied Econometrics* 16(3):289–326.

Philips, Andrew Q. 2016. "Have Your Cake and Eat it Too? Cointegration and Dynamic Inference from Autoregressive Distributed Lag Models." *Working Paper* .

Piketty, Thomas and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics* 118(1):1–39.

Piketty, Thomas, Emmanuel Saez and Stefanie Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6(1):230–271.

Pontusson, Jonas. 2013. "Unionization, Inequality and Redistribution." *British Journal of Industrial Relations* 51(4):797–825.

Sánchez Urribarrí, Raúl A., Susanne Schorpp, Kirk A. Randazzo and Donald R. Songer. 2011. "Explaining Changes to Rights Litigation: Testing a Multivariate Model in a Comparative Framework." *Journal of Politics* 73(2):391–405.

Sims, Christopher A., James H. Stock and Mark W. Watson. 1990. "Inference in Linear Time Series Models with Some Unit Roots." *Econometrica* 58(1):113–144.

Volscho, Thomas W. and Nathan J. Kelly. 2012. "The Rise of the Super-Rich: Power Resources, Taxes, Financial Markets, and the Dynamics of the Top 1 Percent, 1949 to 2008." *American Sociological Review* 77(5):679–699.

Western, Bruce and Jake Rosenfeld. 2011. "Unions, Norms, and the Rise of U.S. Wage Inequality." *American Sociological Review* 76(4):513537.

Wlezien, Christopher. 2000. "An Essay on 'Combined' Time Series Processes." *Electoral Studies* 19(1):77–93.

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:

## Understanding Equation Balance in Time Series Regression: An Extended Examination

Peter K. Enns
peterenns@cornell.edu
Associate Professor
Department of Government
Cornell University

Christopher Wlezien
wlezien@austin.utexas.edu
Hogg Professor of Government
University of Texas at Austin

## Contents

# Appendix 1 Alternate ARDL Model Specification

Table A-1: Alternate Specification of the ARDL Model in Table 4

| | |
|---|---|
| Top 1% Share$_{t-1}$ | -0.88* |
| | (0.13) |
| $\Delta$ Top 1% Share$_{t-1}$ | -0.24* |
| | (0.10) |
| % Congressional Dem.$_{t-1}$ | -0.06* |
| | (0.02) |
| Divided Government$_{t-1}$ | -0.42 |
| | (0.22) |
| Union Membership$_{t-1}$ | -0.36* |
| | (0.09) |
| Top Marginal Tax Rate$_{t-1}$ | -0.04* |
| | (0.02) |
| $\Delta$ Capital Gains Tax Rate$_{t}$ | -0.08* |
| | (0.04) |
| $\Delta$ Capital Gains Tax Rate$_{t-1}$ | 0.06 |
| | (0.04) |
| Capital Gains Tax Rate$_{t-1}$ | -0.11* |
| | (0.02) |
| 3-Month Treasury Bill$_{t-1}$ | -0.01 |
| | (0.06) |
| $\Delta$ Trade Openness$_{t}$ | 0.17 |
| | (0.11) |
| $\Delta$ Trade Openness$_{t-1}$ | 0.14 |
| | (0.11) |
| Log Real GDP$_{t-1}$ | -6.54* |
| | (1.85) |
| $\Delta$ Real S&P 500 Index$_{t}$ | 0.06* |
| | (0.01) |
| $\Delta$ Real S&P 500 Index$_{t-1}$ | 0.03* |
| | (0.01) |
| Real S&P 500 Index$_{t-1}$ | 0.04* |
| | (0.01) |
| Shiller Home Price Index$_{t-1}$ | 0.38* |
| | (0.10) |
| Constant | 77.52* |
| | (18.17) |
| Breusch-Godfrey | 0.13 |

*Notes:* The null hypothesis for the Breusch-Godfrey LM test is *no* autocorrelation.

# Appendix 2 Replication Code for Table 1: Integrated $Y$ and Stationary $X$

All simulations are conducted in Stata 13.

## No Relationship, $\theta_y = 1, \theta_x = 0$, T=50

```
set seed 4545
program define unitroot, rclass
drop _all
set obs 50
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
*critical value set to 2.01 because T=50
sum tstat_x1 if tstat_x1>=2.01
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01
```

## No Relationship, $\theta_y = 1, \theta_x = 0$, T=1,000

```
set seed 4545
program drop unitroot
program define unitroot, rclass
drop _all
```

```
set obs 1000
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```

## No Relationship, $\theta_y = 1, \theta_x = .5$, T=50

```
set seed 4545
program drop unitroot
program define unitroot, rclass
drop _all
set obs 50
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=.5*x[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
```

```
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=2.01
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01
```

## No Relationship, $\theta_y = 1, \theta_x = .5$, T=1000

```
set seed 4545
program drop unitroot
program define unitroot, rclass
drop _all
set obs 1000
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=.5*x[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
```

```
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```

## No Relationship, $\theta_y = 0, \theta_x = 1$, T=50

```
set seed 5656
program drop unitroot
program define unitroot, rclass
drop _all
set obs 50
gen t = _n
gen u=invnorm(uniform())
gen y=u
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=x1[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=2.01
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01
```

## No Relationship, $\theta_y = 0, \theta_x = 1$, T=1,000

```
set seed 5656
program drop unitroot
program define unitroot, rclass
```

```
drop _all
set obs 1000
gen t = _n
gen u=invnorm(uniform())
gen y=u
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=x1[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```

## No Relationship, $\theta_y = 0.5, \theta_x = 1$, T=50

```
set seed 5656
program drop unitroot
program define unitroot, rclass
drop _all
set obs 50
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=.5*y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=x[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
```

```
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=2.01
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01
```

## No Relationship, $\theta_y = 0.5, \theta_x = 1$, T=1,000

```
set seed 5656
program drop unitroot
program define unitroot, rclass
drop _all
set obs 1000
gen t = _n
gen u=invnorm(uniform())
gen y=u if t==1
replace y=.5*y[_n-1] + u if t>1
gen e1=invnorm(uniform())
gen x1=e1 if t==1
    replace x1=x[_n-1] + e1 if t>1
tsset t
reg y l.y x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "unitroot" N times and save the betas and standard errors.
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: unitroot

*Generate t-statistic for each simulated regression and evaluate how many
*regressions we would incorrectly reject the null hypothesis
sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _b_x1
```

```
sum _sim_3
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```

# Appendix 3 Combined time series with additional innovation ($q$)

For the combined time series analysis, the results in the text use the following DGP,

$$Y_t = (e_t^I + e_t^S) \tag{16}$$

We did not add a disturbance to this DGP because the DGP of both $e_t^I$ and $e_t^S$ contain disturbance terms. Nevertheless, we wanted to be sure that our results were not sensitive to this decision. Thus, we conducted the same simulations where the DGP for $Y$ was,

$$Y_t = (e_t^I + e_t^S) + q, q\, N(0,1) \tag{17}$$

The results appear in Table A-1. Not surprisingly, adding the additional disturbance term makes us less likely to observe the true relationship in small samples. However, the overall pattern of results parallels the findings in Table 3.

Table A-1: Identifying a True Relationship ($\beta_1$=0.5) between $X$ and $Y$ when $X$ is Stationary and $Y$ Contains Stationary *and* Unit Root Properties (When $Y_t$ contains the additional innovation $q$.)

| | $T = 50$ | | | | | | $T = 100$ | | | | | | $T = 5,000$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | | $\rho = 0.2$ | | $\rho = 0.5$ | | $\rho = 0.8$ | |
| | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % | $\hat{\beta}_1$ | % |
| $y_{t-1}$ | 0.53 | 85.4 | 0.52 | 84.4 | 0.50 | 82.9 | 0.70 | 99.7 | 0.69 | 99.4 | 0.68 | 99.8 | 0.99 | 100 | 0.99 | 100 | 0.99 | 100 |
| $x1_t$ | 0.51 | 73.3 | 0.51 | 72.5 | 0.50 | 71.0 | 0.50 | 93.3 | 0.50 | 92.4 | 0.50 | 93.7 | 0.50 | 100 | 0.50 | 100 | 0.50 | 100 |
| $x1_{t-1}$ | -0.27 | 28.7 | -0.26 | 27.4 | -0.26 | 27.9 | -0.34 | 64.4 | -0.35 | 63.3 | -0.34 | 60.8 | -0.50 | 100 | -0.50 | 100 | -0.49 | 100 |

*Notes:* $\hat{\beta}_1$ represents the mean estimate of $\beta_1$ across 1,000 simulations. % represents the percent of the simulations for which we (correctly) reject the null hypothesis of no relationship between X and Y.

# Appendix 4 Replication Code for Table 3: Combined Time Series, True Relationship

```
************************
**T=50; x1, rho=.2, .5, .8
************************
set seed 4545

******
*rho=.2
*****
*program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 50
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.2*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "combined" N times and save the betas and standard errors.
*Test whether an equation with mixed orders of integration (combined z, I(0) x1, I(1) x2
*can correctly identify TRUE relationships
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

*Generate t-statistic for each simulated regression
gen tstat_x1 = abs(_b_x1/_se_x1)
```

```
sum tstat_x1 if tstat_x1>=2.01
gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=2.01
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01

******
*rho=.5
*****
program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 50
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.5*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=2.01

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=2.01
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01
```

```
******
*rho=.8
*****
program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 50
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.8*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=2.01

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=2.01
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=2.01


************************
**T=100; x1, rho=.2, .5, .8
************************
```

```
******
*rho=.2
*****
program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 100
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.2*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "combined" N times and save the betas and standard errors.
*Test whether an equation with mixed orders of integration (combined z, I(0) x1, I(1) x2
*can correctly identify TRUE relationships
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

*Generate t-statistic for each simulated regression
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```

```
******
*rho=.5
*****
program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 100
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.5*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96

******
*rho=.8
*****
program drop combined_noq
program define combined_noq, rclass
```

```
drop _all
set obs 100
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.8*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96


************************
**T=5,000; x1, rho=.2, .5, .8
************************

******
*rho=.2
*****
program drop combined_noq
program define combined_noq, rclass
```

```
drop _all
set obs 5000
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.2*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

*Simulate the program "combined" N times and save the betas and standard errors.
*Test whether an equation with mixed orders of integration (combined z, I(0) x1, I(1) x2
*can correctly identify TRUE relationships
simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

*Generate t-statistic for each simulated regression
gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96


******
*rho=.5
*****
program drop combined_noq
program define combined_noq, rclass
```

```
drop _all
set obs 5000
gen t = _n
*gen stationary time series (x1)
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.5*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96


******
*rho=.8
*****
program drop combined_noq
program define combined_noq, rclass
drop _all
set obs 5000
gen t = _n
*gen stationary time series (x1)
```

```
gen e1=invnorm(uniform())
gen x1=e1 if t==1
replace x1=.8*x1[_n-1] + e1 if t>1
*gen integrated time series (x2)
gen u=invnorm(uniform())
gen x2=u if t==1
replace x2=x2[_n-1] + u if t>1
*gen combined times series (z) that his a function of x1 and x2
gen q=invnorm(uniform())
gen z = x1 + x2
tsset t
reg z l.z x1 l.x1
estat bgodfrey
mat P = r(p)
return scalar pvalue_bg = P[1,1]
end

simulate pvalue_bg=r(pvalue_bg) _b _se, reps(1000) nodots: combined_noq

sum _eq2_pvalue_bg if _eq2_pvalue_bg<0.05
sum _sim_1 _b_x1 _sim_3

gen tstat_ly = abs(_sim_1/_sim_5)
sum tstat_ly if tstat_ly>=1.96

gen tstat_x1 = abs(_b_x1/_se_x1)
sum tstat_x1 if tstat_x1>=1.96
*
gen tstat_lx1 = abs(_sim_3/_sim_7)
sum tstat_lx1 if tstat_lx1>=1.96
```