Equation Balance in Time Series Analysis: What It Is and How to Apply It*

Mark Pickup[†]
Paul M. Kellstedt[‡]

January 27, 2020

Abstract

The recent symposium in Political Analysis (volume 24, number 1) on potential problems with the use of the generalized error correction model (GECM) raised several important issues, not all of which were fully resolved. One of those is the issue of equation balance. While it is true that most of the participants in the symposium agreed that ensuring equation balance is important, there has not been a consistent and agreed-upon definition of what equation balance is, or an understanding of exactly why it is important or how it can be applied. In this paper, we treat equation balance as a theoretical matter, not merely an empirical one—including two ways to use the concept of equation balance to test theoretical propositions before data have been gathered and before any model has been estimated. We illustrate our ideas with two influential time-series studies.

Keywords: time series; equation balance; general error correction model.

*We thank John Freeman, Matthew Lebo, and Dominik Hangartner for useful insights on earlier drafts of this paper. All errors remain our own.

[†]Department of Political Science, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada, e-mail: mark.pickup@sfu.ca

[‡]Department of Political Science, Texas A&M University, College Station, TX 77843, USA, e-mail: kellstedt@tamu.edu

1 Introduction

Since DeBoef and Keele's (2008) influential article "Taking Time Seriously," the use of the generalized error correction model (GECM) in applied research in political science has proliferated.[1] But in recent years, these applications have generated considerable controversy. Nowhere was this more evident than in the "Symposium on Time Series Error Correction Methods in Political Science" in *Political Analysis* (volume 24, number 1) where seven articles debated the situations under which the use of the GECM are appropriate.

Despite the disagreement evident in the symposium, there was one subject upon which all of the participating authors agreed: The necessity of estimating models with *balanced* equations. Not one disputant in the symposium, in fact, disagrees with this foundational claim. In fact, in the abstract of the very first article in the symposium, Grant and Lebo (2016, p. 3) note: "... without equation balance the model is misspecified and hypothesis tests and long-run-multipliers are unreliable." In their initial response, Keele, Linn, and Webb (2016a, p. 34) emphasize that "Stable long-run relationships in turn imply balanced equations." In their rejoinders, Lebo and Grant (2016, p. 71) note optimistically that "One point of agreement among the papers here is that equation balance is an important and neglected topic." Keele, Linn, and Webb (2016b, p. 83) pick up the olive branch, agreeing that "We believe the discussion of equation balance was an important part of the initial exchange." And yet, one of the Symposium's commenters on the controversy (Freeman 2016, p. 50) lays it bare: "It now is clear that equation balance is not understood by political scientists."

Despite the agreement about its importance, the term "equation balance" went undefined in the initial pieces by both Grant and Lebo (2016) and Keele, Linn, and Webb (2016a). It remained undefined, in fact, until Freeman's (2016) contribution—and even then, it was defined in a footnote.[2] Keele, Linn, and Webb's "Concluding Comments"

[1] Grant and Lebo (2016) chronicle both the growth and shifts in its use in their paper. We will not revisit that topic here.

[2] See Freeman's note 1, p. 50.

(2016b, p. 83) does contain a more complete definition of the term, though this is obviously after-the-fact. And it is worth emphasizing that the original DeBoef and Keele (2008) article did not mention equation balance at all—which may be part of the reason that the controversy arose in the first place.

In this paper, we seek to bring the issue of "equation balance" into sharp focus, with the hope of providing concrete guidance to applied researchers who use time-series models that may have been lacking at the end of the Symposium in *Political Analysis*. We begin by reiterating the clearest definition of equation balance that is available. Next, we show why balance matters for applied researchers, discussing the notion of equation balance both as a theoretical matter and an empirical one. To preview the argument: Before a researcher even considers the best estimator for her empirical model or empirically determining the order of integration of her variables, she can determine if her theory implies a balanced theoretical model. Also, once she has empirically tested the order of integration of her variables (but before she estimates her model), she can determine if her empirical model is balanced. If the answer is no in either case, she need not continue on to estimating the model. She already knows at this point that her theory (as specified) is incorrect. These two procedures are a powerful way of testing the validity of a model without even having to worry about model estimation. Unfortunately, because balance is not a well understood concept, researchers rarely take advantage of these procedures. In our last section, we explain how these two procedures can be applied to two influential time-series studies.

2 What is balance?

A model is defined as balanced "if and only if the regressand and the regressors (either individually or collectively, as a co-integrated set) are of the same order of integration" (Banerjee *et al.*, 1993, page 166). This definition needs some detailed explanation and requires us to understand the concept *order of integration* for both an individual variable and a collection of variables.

Starting with an individual variable, the order of integration d of a variable (Y) equals the number of times the data needs to be differenced in order to transform it into a stationary process, and is denoted as $Y \sim I(d)$ (Greene, 2012, page 943).[3] If the variable is already stationary, it is defined as zero-ordered integrated, $Y \sim I(0)$. A variable that needs to be differenced once is first-order integrated, $Y \sim I(1)$, differenced twice second-order integrated, $Y \sim I(2)$, and so on. For example, a time series of aggregated poll measures of consumer sentiment is $I(0)$ if it is stationary without any differencing. As another example, a time series of coded news reports of conflictual foreign policy acts is I(1) if it is *not* stationary without differencing, but the first difference of the series ($Y_t - Y_{t-1}$) is stationary. If the order of integration, d, required to difference a variable, (Y), to produce a stationary variable is a non-integer number between 0 and 1 then Y is called a fractionally integrated process (Pesaran, 2015). This means the variable must be fractionally differenced to produce a stationary variable (Box-Steffensmeier and Smith, 1998).[4]

Moving from consideration of a single variable to a collection of variables, the linear combination of two or more variables with the same order of integration may have a different (lower) order of integration then each of the variables individually (Engle and Granger, 1987). The most simple example of such a relationship is of two $I(1)$ variables adding up to produce a $I(0)$ process. Say $Y \sim I(1)$ and $X \sim I(1)$ but $\beta_1 X + \beta_2 Y = Z \sim I(0)$. This is called cointegration, and the *collective* order of integration of Y and X is zero even though they are *individually* $I(1)$. For example, if time series of aggregated poll measures of economic expectations and of the public's appetite for liberal policy are both $I(1)$, but there is a linear combination of the two series that produce an $I(0)$ series, then the two original series are cointegrated. This cointegration may take place

[3]By stationary we mean a covariance stationary process. For a process to be covariance stationary, its expected value, variance, and autocovariances must be constant over time.

[4]Such processes are called long-memoried, but note that there are also long-memory processes that are covariance stationary without fractional differencing.

because, even though both liberal policy appetite and economic expectations may wander up and down without returning to any constant equilibrium (expected value), they may covary such that when one goes up(down), so does the other. For example, as the public's expectations of the economy increase, their appetite for liberal policy also increases. If this occurs in such a way that the difference between the two variables is equal to a constant plus a random perturbation up or down, then this difference is a stationary process with a constant equilibrium. The combination of the variables is $I(0)$.

Cointegration can occur between more than two variables. For example, if $X_1 \sim I(1)$ and $X_2 \sim I(1)$ and $Y \sim I(1)$ and $\beta_1 X_1 + \beta_2 X_2 + \beta_3 Y = Z \sim I(0)$, then X_1 and X_2 and Y are said to cointegrate. More complicated still is the possibility of "multiple cointegration." In multiple cointegration, two or more variables with the same order of integration add up to a lower order of integration and then the resulting process combines with one or more other variables with the same order of integration to add up to an even lower order of integration. For example: $X_1 \sim I(2)$ and $X_2 \sim I(2)$ combine to produce $Y_1 \sim I(1)$. And then, $Y_1 \sim I(1)$ and $Y_2 \sim I(1)$ combine to produce $Z \sim I(0)$. If this occurs, the combination of X_1 and X_2 and Y_2 is of order $I(0)$.[5] Multiple cointegration is a possibility but, in political science data, is very rare.

A model is balanced when the collection of variables on the right-hand side (the regressors) are *collectively* of the same order of integration as the variable on the left-hand side (the regressand).[6] From a theoretical perspective, this is the only requirement for a model to be balanced. There are some additional considerations when estimating the model. We discuss these in the next section.

[5]So far, we have discussed single equation models with weakly exogenous regressors. We will continue to do so for the remainder of the paper but the concept of balance applies equally to multivariate models. The condition of balance must be met for the equation for each of the endogenous variables in the system of equations.

[6]As Enns and Wlezien (2017) note, this does not mean the order of integration of all variables on the right-hand side has to be the same as the variable on the left-hand side. It is the *collective* order of integration on the right-hand side that has to match the left-hand side.

3 Why does balance matter?

The concept of balance can be applied to both the theoretical model and the empirical model. Balance matters, we will argue, because a theoretical or empirical model that is not balanced is wrong or—at the very least—incomplete in an important way. Before discussing this in detail, an analogy may be helpful. Balance also applies to chemical equations. A chemical equation describes how a combination of entities react to produce new entities. The entities on the left-hand side of the equation represent the chemicals being combined, and the entities on the right-hand side represent the chemicals that are produced. The law of conservation of mass requires the same amount of mass before and after the reaction. This means that the number of atoms on the left-hand side of the equation must add up to the number of atoms on the right-hand side. This is called equation balance and is a necessary condition for a theorized chemical equation to be correct. If it is not balanced, it violates the law of conservation of mass and has no possibility of being correct. If the chemist has a theory that implies an unbalanced chemical equation, she does not even need to enter the lab to know her theory is incorrect.

For statistical models, it is not the law of conservation of mass that applies, but there is an equivalent (if less famous) law of conservation of order of integration. The intuition is simply that a variable on the left-hand side that does not trend cannot be the product of right hand side variables that all trend, unless those trending variables combine with one another to produce a non-trending process. In the same way an $I(0)$ left-hand side variable with a constant equilibrium cannot be the product of $I(1)$ right-hand side variables without equilibria, unless those right-hand side variables add up (cointegrate) to produce an $I(0)$ process with a constant equilibrium.

If the order of integration on the left-hand side differs from that on the right-hand side, this tells you that your model is either wrong or incomplete in a way that will prevent a meaningful interpretation of the model. For example, if the regressand is an $I(0)$ media tone variable, it cannot be the product of (equal to) an $I(1)$ unemployment level regressor and an $I(0)$ error term:

$$tone_t = \beta_0 + \beta_1 unemp_t + \epsilon_t \tag{1}$$

This model, empirical or theoretical, cannot be correct. As an $I(1)$ process, shocks to unemployment would not dissipate (by definition). If these changes have an effect on media tone, then these effects would also not dissipate and media tone would then also have to be an $I(1)$ process.

As another example, if media tone were in fact $I(1)$, then it could not be the product of an $I(0)$ consumer sentiment regressor and an $I(0)$ error term:

$$tone_t = \beta_0 + \beta_1 CS_t + \epsilon_t \tag{2}$$

All changes in consumer sentiment will dissipate over time and so cannot explain the non-dissipating changes in media tone. If one were willing to allow the error term to be $I(1)$, the model would not be balanced by the strict definition (that requires the regressand and regressors collectively to have the same order of integration) but it would achieve balance in a looser sense though through the error term. However, this model is very much incomplete. It implies that the non-dissipating changes in media tone are being driven by some $I(1)$ variable that has been excluded from the model (resulting in an $I(1)$ error term). The $I(0)$ consumer sentiment regressor may have an effect on the dependent variable (media tone) but only in that it explains short-term deviations from the underlying long-term changes. For example:

$$\Delta tone_t = \beta_0 + \beta_1 CS_t + \epsilon_t \tag{3}$$

or:

$$\Delta tone_t = \beta_0 + \beta_1 \Delta CS_t + \epsilon_t \tag{4}$$

But these are both very different models from (2), suggesting very different relationships.

Any theory that includes a long-run relationship between consumer sentiment and media tone would be incorrect and, from the perspective of the theoretical model, an important component is missing. From the perspective of the empirical model, the $I(1)$ error will produce problems for estimation and inference, as standard estimation procedures assume stationary errors. As Maddala and Kim note, one should avoid estimating an unbalanced regression (again, in the strict sense) if it can be avoided (Maddala and Kim, 1998, page 252). On the basis that the concept of balance can tell us a great deal about both the theoretical and empirical model even before we estimate the empirical model, we propose a three-step process.

4 Three ways to apply balance before estimating your empirical model

The principles of equation balance can be applied at multiple stages of the research process, which we describe below.

4.1 Using balance to test the theoretical model before collecting data

When a researcher is developing a theory, she should not only develop expectations regarding what concepts causally relate to others, but also about the dynamic nature of those concepts. For example, is public opinion towards the Affordable Care Act $I(0)$ stationary, trend stationary, $I(1)$, or some other process? By doing this, the researcher can place a check on her theory. If she has good theoretical reason to believe that public opinion towards the ACA is $I(0)$, a theory that stipulates it is caused by levels of media coverage of the ACA is incorrect or incomplete, if she strongly believes that media coverage is an $I(1)$ process, unless her theory also includes other causal factors (of $I(1)$ or higher order) that combined with media coverage to produce an $I(0)$ process.[7]

Returning to the analogy of balancing a chemical equation, when a chemist has a theory that chemical X combines with chemical Y to produce chemical Z: $(X + Y = Z)$, a first check is that the chemical equation is balanced. This means making sure that there are just as many atoms of each type on the left-hand side as there are on the right-hand side. If the equation is not balanced, she does not need to test the reaction between X and Y to know that her theory is wrong.

Similar to the chemist, the political scientist should ask: 1. What type of data-generating process (DGP) do I believe produced my variables? and 2. Given 1, is my theoretical model balanced?

4.2 Using balance to test the empirical model before estimation

Once a researcher has developed a theory that passes the balance test, she should test the order of integration of each of her variables and, if it is part of her theory, whether or not any right-hand side variables that are $I(1)$ (or higher) cointegrate. There are issues of power with many of the tests of integration and cointegration, and so a grain of salt (and sometimes many grains) should be applied when interpreting those results. Fortunately, there have been recent advances in this area. If the empirical evidence is relatively unambiguous on the fact that the regressand is $I(1)$, but the order of integration of the regressors is unknown, the researcher can use the procedure described by Pesaran *et al.* (2001) and Philips (In Press) to determine if the regressors cointegrate with the lag of the regressand to produce a balanced model. Further, Webb *et al.* (2017) are investigating a method to test for cointegration between the regressand and regressors when the order of integration of one or both are unknown. Unfortunately, this method can result in an indeterminate result, but the concept of balance may be of assistance here as well. Care-

[7] In our example, it is also possible that public opinion might be caused by *shifts in* (i.e., the first difference of) media coverage, which would be an $I(0)$ variable.

ful consideration of the ways in which variables in the model may theoretically combine to produce a balanced equation may narrow down the ways those variables may cointegrate. Reducing these possibilities may lessen the possibility of indeterminate results. The Johansen test (1991) also provides a means of testing for cointegration when the order of integration of the variables is unknown. This last test also has the advantage of being applicable to multiple time series models.

A Bayesian approach (Brandt and Freeman, 2009, 2006; Sims and Zha, 1999) allows the researcher to avoid making a definitive decision regarding the order of integration and cointegration of the variables. Instead, she can use the theoretical expectations and empirical evidence to place priors on one's model that reflect her uncertain beliefs. Here also the need to achieve balance can guide the researcher. For example, the priors placed on the model should suggest a balanced model. We will also see below that, when the theoretical expectations and empirical evidence regarding a variable's order of integration conflict, the concept of balance can be used as a touchstone by which to interpret the empirical tests.

Once the researcher has updated her theoretical expectations of the order of integration of the left-hand and right-hand side variables on the basis of the empirical evidence, she is in a position to check whether or not the empirical model is balanced.[8] Extending the chemical equation analogy, if the chemist's theory implies a balanced chemical equation ($X + Y = Z$) based on what they believe about chemicals X, Y and Z, but then, after examining the chemicals, discovers X has more atoms than originally believed, the chemist knows that their theory is wrong without mixing the chemicals. The political scientist should use tests of integration, fractional integration, and cointegration to check the evidence regarding the DGP for the variables. On the basis of these tests, she should ask: is my empirical model balanced?

[8] It should be noted that if the variables are bounded and regularly "bumping up against" their bounds, then tests of integration and cointegration will be problematic. Such variables do not have straightforward $I(0)$ or $I(1)$, or even $I(d)$, properties. The topic of such bounded variables that regularly collide with their boundaries is beyond the scope of this paper.

If the theoretical or empirical model do not pass the balance test, this indicates to the researcher that further theorizing is required. Assuming that the theoretical and empirical models do pass the balance test, the researcher needs to make one further check.

4.3 Using balance to test the re-parameterized empirical model before estimation

For the purposes of estimation, it is not only necessary that the equation is balanced, it is also necessary that there is a re-parameterization of the empirical model in which the regressand is $I(0)$ and the equation is balanced (Banerjee *et al.*, 1993, 167-168). We call this "$I(0)$ balance." If this is not the case, the distributions for some or all of the usual tests of statistical inference—most commonly t and F statistics—will not have standard distributions.[9]. Typical examples of re-parameterized models are standard and general error correction models. The researcher should seek out a re-parameterization in which the regressand is $I(0)$ and determine if the equation is still balanced and, if not, what restrictions need to be placed to make it so. These would be restrictions on one or more parameters in the model. For example, if $y_t \sim I(1)$ and $x_t \sim I(0)$, a re-parameterization of the lagged dependent variable (LDV) model:

$$y_t = \alpha_1 y_{t-1} + \beta_0 + \beta_1 x_t + \epsilon_t \tag{5}$$

in which the regressand is $I(0)$ and the equation is balanced can be achieved by placing the restriction $\alpha_1 = 1$.

$$y_t = y_{t-1} + \beta_0 + \beta_1 x_t + \epsilon_t \tag{6}$$

[9]If this condition is not met, tests of inference for some of the parameters may still have standard distributions. See (Enders, 2004, page 285–287)

$$y_t - y_{t-1} = \beta_0 + \beta_1 x_t + \epsilon_t \tag{7}$$

$$\Delta y_t = \beta_0 + \beta_1 x_t + \epsilon_t \tag{8}$$

Recall that if $y_t \sim I(1)$, then $\Delta y_t \sim I(0)$. Then in this re-parameterization the regressand Δy_t is $I(0)$ and the regressor x_t is $I(0)$.

Continuing from a previous example, we noted that the following model was not balanced:

$$tone_t = \beta_0 + \beta_1 CS_t + \epsilon_t \tag{9}$$

when $tone_t \sim I(1)$ and $CS_t \sim I(0)$ and that there was at least one $I(1)$ regressor missing from the right-hand side. The missing regressor may simply be a lag of the regresand:

$$tone_t = \beta_0 + \alpha_1 tone_{t-1} + \beta_1 CS_t + \epsilon_t \tag{10}$$

This is balanced, but it is only $I(0)$ balanced if we can place the restriction $\alpha_1 = 1$.[10]

In general, if restrictions are required, the researcher must decide empirically and/or theoretically if they are valid. It is only at this point that the researcher should proceed with estimating a model. Note that if such a re-parameterization exists, it is not necessary for the researcher to use the re-parameterized form for estimation. It is sufficient that it exists (Banerjee *et al.*, 1993, 167-168).

[10]If this restriction is not placed on the model before estimation, OLS and MLE will tend to underestimate the value of α_1, so that it appears to be less than 1. Bayesian estimation approaches may be able to overcome this problem but they can be sensitive to the prior chosen for α_1 (Maddala and Kim, 1998, pages 263-295).

5 Determining if a model is balanced

The following procedure determines if a model is balanced. It applies equally if it is a theoretical model—a data-generating process—or an empirical model for which you are checking balance. However, a prerequisite for checking balance is determining orders of integration and cointegration. For the data generating process, this is based on theoretical expectations but in the case of empirical models, this is based on tests (such as Dickey-Fuller tests) that can lead to uncertainty. For now, we assume the order of integration and cointegration is knowable and return to the issue of uncertainty at the end. Accordingly, the researcher should proceed by:

First, determining the order of integration of the variable on the left-hand side, theoretically or empirically.

Second, determining the order of integration of variables on the right-hand side. Without cointegration, the order of integration of the right-hand side is equal to the *highest* order of integration of all variables on the right-hand side. With cointegration, the order of integration may be lower. For example, if all $I(1)$ variables on the right-hand side combine to produce an $I(0)$ process and the only remaining variables are $I(0)$, the order of integration for the right-hand side is $I(0)$. However, if $X_1 \sim I(1)$ and $X_2 \sim I(1)$ cointegrate to produce an $I(0)$, but $X_3 \sim I(1)$ is also on the right-hand side, the order of integration for the right-hand side variables is $I(1)$.

Third, following some simple rules for checking model balance.[11]

i) If the regressand and all regressors individually are $I(0)$, you have balance. For example, if $y_t \sim I(0)$ and $x_t \sim I(0)$, the autoregressive distributed lag (ADL) model with one lag of the independent variable and one lag of the dependent variable :

[11]Note that we assume throughout that the error term ϵ_t is stationary $I(0)$. This is a requirement for most standard estimators and so it is a practical assumption.

$$y_t = \alpha_1 y_{t-1} + \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \epsilon_t \tag{11}$$

is balanced, and the standard error correction model (ECM)

$$\Delta y_t = \alpha_0 + \gamma(y_{t-1} + \kappa_1 x_{t-1}) + \kappa_0 \Delta x_t + \epsilon_t \tag{12}$$

is balanced, and the general error correction model (GECM)

$$\Delta y_t = \gamma y_{t-1} + \alpha_0 + \delta_1 x_t + \delta_2 \Delta x_{t-1} + \epsilon_t \tag{13}$$

is balanced. Note equations 11, 13, and 12 are all different parameterizations of the same process.

It is also the case that if the regressand and all regressors are I(0), the first difference (FD) model is balanced:

$$\Delta y_t = \alpha_0 + \kappa_0 \Delta x_{1t} + \epsilon_t \tag{14}$$

Recall that if $y_t \sim I(0)$, then $\Delta y_t \sim I(0)$. As is the case for x_t. However, it is important to note that equation 14 represents a different relationship between X and Y than equations 11, 12 or 13.

It is also the case that these are all equations that have an $I(0)$ regressand and are balanced.

ii) If the regressand is $I(0)$ but some regressors are not, ask yourself: is there a linear combination of these non-$I(0)$ regressors that is $I(0)$? If yes, you have balance. If no such linear combination exists, you do not have balance.

If $y_t \sim I(0)$, $x_{1t} \sim I(1)$ and $x_{2t} \sim I(1)$:

$$y_t = \alpha_1 y_{t-1} + \beta_0 + \beta_1 x_{1t} + \beta_2 x_{1t-1} + \beta_3 x_{2t} + \beta_4 x_{2t-1} + \epsilon_t \tag{15}$$

is balanced if and only if: $\beta_1 x_{1t} + \beta_2 x_{1t-1} + \beta_3 x_{2t} + \beta_4 x_{2t-1} \sim I(0)$. If it is balanced, it is balanced with an $I(0)$ regressand.

iii) If the regressand is $I(1)$ and all regressors are $I(0)$, you do not have balance. For example, if $y_t \sim I(1)$ and $x_t \sim I(0)$, the finite distributed lag (FDL) model with one lag of the independent variable:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \epsilon_t \tag{16}$$

is not balanced.

iv) If the regressand is $I(1)$ and regressors are collectively $I(1)$, the equation is balanced. For example, if $y_t \sim I(1)$ and $x_t \sim I(1)$, the ADL(1,1):

$$y_t = \alpha_1 y_{t-1} + \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \epsilon_t \tag{17}$$

is balanced.

However, when we seek a re-parameterization with an $I(0)$ regressand, we find that additional restrictions are required. One possible re-parameterization is the following ECM:

$$\Delta y_t = \alpha_0 + \gamma(y_{t-1} + \kappa_1 x_{t-1}) + \kappa_0 \Delta x_t + \epsilon_t \tag{18}$$

The regressand is $I(0)$ and Δx_t is $I(0)$. In order for the regressors to be collectively $I(0)$, it must either be the case that $(y_{t-1} + \kappa_1 x_{1t-1}) \sim I(0)$ or $\gamma = 0$. The first restriction is to say that Y and X cointegrate. If this restriction cannot be made then we must assume that $\gamma = 0$ and the appropriate equation is the FD model:

$$\Delta y_t = \alpha_0 + \kappa_0 \Delta x_t + \epsilon_t \tag{19}$$

Similarly, if $y_t \sim I(1)$, $x_{1t} \sim I(1)$ and $x_{2t} \sim I(0)$, the ADL(1,1):

$$y_t = \rho y_{t-1} + \beta_0 + \beta_1 x_{1t} + \beta_2 x_{1t-1} + \beta_3 x_{2t} + \beta_4 x_{2t-1} + \epsilon_t \tag{20}$$

is balanced but the following ECM re-parameterization with a $I(0)$ regressand:

$$\Delta y_t = \alpha_0 + \gamma(y_{t-1} + \kappa_1 x_{1t-1}) + \kappa_0 \Delta x_{1t} + \beta_3 x_{2t} + \beta_4 x_{2t-1} + \epsilon_t \tag{21}$$

requires either $(y_{t-1} + \kappa_1 x_{1t-1}) \sim I(0)$ or $\gamma = 0$

As we discussed at the outset, empirically determining the order of integration and cointegration can be difficult. There are a number of tests for both, and the appropriateness of each depends on assumptions about the deterministic elements in the DGP (e.g., structural breaks and trending). Enders (2004) (201-214) provides a procedure to follow when testing the order of integration; Philips (2017) outlines the PSS procedure (Pesaran *et al.*, 2001) for testing cointegration when the order of integration of the regressors is uncertain; Webb *et al.* (2017) outline a bounds procedure when the orders of integration of both the regresand and regressors are uncertain; and Johansen (1991) provides a rank test that does the same. These procedures are useful, but testing is complicated by the possibility of fractional integration, and the low power these tests all suffer from when T is small. It is not our intention to provide a complete pre-test procedure, but we do note how the concept of balance can assist the researcher when interpreting these empirical tests.

When a researcher has established theoretical expectations regarding the order of integration of each variable in the theoretical model, these can be used as priors when interpreting the tests of integration and cointegration. The requirement that the theoretical and empirical models be balanced provides an additional prior that can be applied. For example, consider the case in which tests of integration confirm the theoretical expectation that the regressand is $I(0)$ and that the same holds for all regressors but one. For that final regressor, X, tests of integration are unclear. In this situation, balance requires that either X is $I(0)$ or there is an $I(1)$ covariate that cointegrates with X that is missing from the model. If the researcher believes that the second possibility is theoretically unlikely, this then suggests that X is likely $I(0)$. Alternatively, it tells the researcher that they need to rethink their theoretical model, such that it is balanced with $X \sim I(1)$.

6 Examples

We now describe how two influential papers published in political science using time series data that may have produced different findings if the authors followed our advice on considering balance in their theoretical and/or empirical models.

6.1 Inequality and policy attitudes in the U.S.

In a well known and provocative study, Kelly and Enns (2010) use time-series data from the U.S. and single-equation ECM estimation methods to examine the relationship between the degree of economic inequality and the public's preferences for activist government. With annual data, they measure inequality with the widely used Gini coefficient, and measure public opinion with Stimson's (1991) Policy Mood index. The theoretical model postulated is:

$$Mood_t \sim f(Inequality_t, Policy_t, Unemployment_t, Inflation_t, \epsilon_t) \quad (22)$$

where *Policy* is included to capture thermostatic feedback effects (Wlezien, 1995), and the other variables are self-explanatory.

If the authors had followed the procedures outlined above, they would theorize the order of integration of their variables and ask if the theoretical model was balanced. They are, in fact, either silent or (in one case) agnostic as to whether the key variables in the system contain unit roots, or whether any variables might cointegrate. With respect to *Mood*, their dependent variable, the authors write (note 12, p 863), that:

> Our dependent variable is, theoretically, a stationary process (Wlezien, 1995). Unit root tests of public mood provide mixed results, meaning that there is some possibility that mood has a long memory.

The authors are silent about the univariate properties of the variables on the right-hand-side. These variables—*Inequality*, *Policy*, *Unemployment*, and *Inflation*—are included in the GECM both in first differences and in lagged levels, which is standard. Those variables in levels almost certainly contain some, perhaps several, series that are $I(1)$. Their key theoretical variable, *Inequality* (the Gini coefficient), is theoretically $I(1)$, and most likely empirically as well. The same would be true for *Inflation*, which Economists routinely treat as $I(1)$. Wlezien's *Policy* variable is a spending measure, and may have the same properties but has been treated as stationary in past work. *Unemployment* has debatable stationarity properties as well. This raises important questions about how these variables theoretically may combine to produce a balanced model. These questions go unanswered, missing an opportunity to flesh out the theoretical model.

The authors opt to estimate a GECM to model the dynamic process, though not because of theoretical expectations about unit roots or cointegration in the system. As the authors continue (note 12, p 863):

> Using an ECM, then, also hedges against results driven by the time-series properties of mood since the first difference of mood is the dependent variable in an ECM.

Mood, that is, may be either $I(0)$ or $I(1)$. But in their GECM setup, because the left-hand-side variable is first-differenced, Δ *Mood* will be $I(0)$ (because the first difference of either an $I(0)$ or $I(1)$ process is $I(0)$.

$$\begin{aligned}\Delta Mood_t = \alpha_1 Mood_{t-1} + \alpha_0 + \beta_1 Inequality_t + \beta_2 \Delta Inequality_{t-1} \\ + \beta_3 Policy_t + \beta_4 \Delta Policy_{t-1} + \beta_5 Unemployment_t + \beta_6 \Delta Unemployment_{t-1} \\ + \beta_7 Inflation_t + \beta_8 \Delta Inflation_{t-1} + \epsilon_t \quad (23)\end{aligned}$$

The authors implicitly assume ϵ_t is stationary and present three GECMs in their Table 1 (p 864) with Δ *Mood* as the dependent variable. Two of these (columns 2 and 3) contain the key *Inequality* variable to test their theory. (Column 1 is a baseline model.) Are these equations balanced? The left-hand-side variable is clearly $I(0)$, but the right-hand side of the models in columns 2 and 3 of their table contains at least one variable that is theoretically $I(1)$. This means either that a linear combination of the $I(1)$ variables on the right hand side must combine to produce an $I(0)$ series—that is, that they are cointegrated—or that the order of integration of the right-hand-side is 1, and that the model is unbalanced. Again, if the authors were to follow our advice, they would empirically consider the orders of integration of these variables and ask if the empirical model could plausibly be balanced.

Using their publicly available data, we test whether the $I(1)$ variables on the right-hand-side might be cointegrated. Of the established procedures, there are two that we would recommend: the PSS bounds test by Pesaran *et al.* (2001) and the Johansen rank test (Johansen 1991). Philips (In Press) shows the sensitivity of these methods to both Type I and Type II errors. We discuss the bounds test procedure proposed by Webb *et al.* (2017) at the end. Of the Johansen and PSS tests, the latter assumes that the left-hand-side variable, before first differencing, is $I(1)$. If *Mood* is in fact $I(0)$, the PSS test is inappropriate.

Beginning with the first model presented by Kelly and Enns (2010) containing *Mood* and *Policy* only, the Johansen test indicates that all variables are stationary. On that basis, the equation is balanced. When *Inequality* (Gini coefficient) is added to the model, the Johansen test indicates that cointegration is not achieved.[12] This makes sense if we believe *Mood* and *Policy* are stationary and *Inequality* is $I(1)$. There is nothing with which *Inequality* can cointegrate. The result is that this equation is not balanced. It is $I(0)$ on the left-hand side and $I(1)$ on the right-hand side. When *Inflation* and *Unemployment* are added into the model, the Johansen test indicates that cointegration is achieved (with a single cointegrating equation). This suggests that *Inequality* integrates with *Inflation*, *Unemployment* or both. The consequence is a balanced equation.

The PSS bounds test fails to find cointegration for all models but if we believe that *Mood* is stationary, as the Johansen test indicates in the first model, then the PSS bounds test is inappropriate. If, after viewing the evidence from the Johansen test, the researcher is not convinced that *Mood* is stationary but does not want to commit to it being $I(1)$, the procedure being developed by Webb *et al.* (2017) allows the researcher to test if equilibrium is achieved between *Mood* and the right-hand-side variables, either because they are all stationary or because the right-hand side variables cointegrate, while remaining agnostic as to which of the two is the case. This procedure may develop into a useful tool for testing balance, but the procedure has a relatively high probability of returning an inconclusive result. And even if it does not, it leaves the researcher in the dark about the nature of the equilibrium relationship between the variables. Hopefully, though, giving careful consideration to balance may limit the range of possible data generating processes, and in doing so, reduce the probability of an inconclusive result, and provide some direction to the researcher about the nature of the equilibration (if it is found).

Overall, we believe the results indicate that the second model should never have been estimated. It is not balanced. The results indicate that the first model, which is a replication of the thermostatic model, is balanced. They also indicate that the third (final) model is balanced. That balance is achieved through more than one $I(1)$ independent

[12] A 0.01 significance level was used to interpret the rank tests.

variable cointegrating, rather than one or more independent variable cointegrating with the lagged dependent variable, has consequences for how the equilibrium between the variables is interpreted. The latter (incorrectly) suggests that *Mood* itself has no equilibrium but has an equilibrium with respect to one or more of the independent variables. This is the interpretation given by Kelly and Enns (2010). The former (correctly suggested by the Johansen tests) reveals that *Mood* has an equilibrium that is conditional on variables that are (in combination) stationary. Not only does this have important consequences for how we understand the dynamics in *Mood*, we believe that the type of consideration to balance that we have given may have avoided the controversy with which the publication of Kelly and Enns (2010) was met. Much of that controversy centred on how the results of the GECM should be interpreted and, as the *Political Analysis* symposium revealed, questions of balance.

6.2 Media coverage and economic performance

In a recent example, Soroka *et al.* (2015) examine the relationship between both the volume and tone of media coverage of the economy and various aspects of economic performance. They ask, in effect: How does economic performance shape economic news? In their Table 3, they present six models—three for volume of media coverage, three for tone of media coverage—examining lagging, coincident, and leading indicators of the economy. The dependent variables representing media coverage are all first-differences. The independent variables representing the past, present, and future of the economy are in both changes and in levels. The implied data generating process is:

$$media_t \sim f(media_{t-1}, economy(past)_t, economy(current)_t, economy(future)_t, \varepsilon_t).$$

As in the previous example, Soroka *et al.* (2015) do not discuss the possibility of their theoretical model being balanced, but they do empirically test the univariate properties of their dependent and independent variables. They determine that the dependent variable ($media_t$) is stationary ($I(0)$) while the economic variables on the right-hand-side are nonstationary ($I(1)$) in levels (see their Appendix B). They implicitly assume that ε_t is stationary. Taken together, and bearing in mind that an $I(0)$ variable ($media_{t-1}$) and an $I(1)$ economy variable cannot cointegrate to produce an $I(0)$ process, this means that the implied empirical model is not balanced *unless* the economic variables cointegrate—and no case is made for this. The left-hand side is $I(0)$ and the right-hand side is $I(1)$. This means that, if the economic variables do belong in the model, they must either cointegrate with each other or the model is missing one or more nonstationary variables that cointegrate with the economic variables. Determining which is the case is necessary to understand how the economic variables affect media coverage. If $I(1)$ variables have been mistakenly excluded from the model, ε_t is not stationary and the t-statistics calculated and used for inferential purposes are not distributed as expected.

Based in part on the conclusions from their Table 3, the authors go on to estimate two further models (in their Table 6) that control for consumer confidence (both prospective and retrospective), to see if media coverage is influenced by (and not just influences, as prior studies have found) consumer confidence. As before, the dependent variables in the two models are the volume and tone of media coverage of the economy. The independent variables are the leading index of economic indicators (LEI), and both prospective and retrospective consumer confidence (all in both levels and first differences). The implied data generating process is: $media_t \sim f(media_{t-1}, LEI_t, pros(confidence)_t, retro(confidence)_t, \varepsilon_t)$. Soroka *et al.* (2015) test the dependent and independent variables and determined that the dependent variable is stationary while the two consumer confidence variables are stationary and the single economic variable is nonstationary (see their Appendix B). Three stationary and one nonstationary variable cannot cointegrate to produce an $I(0)$ process, so the three independent variables and the lag of the dependent variable produce a $I(1)$ process. The empirical model is not balanced. The left-hand side is $I(0)$ and the right-hand side is $I(1)$. This produces the same problems as those described above for Table 3, except that there isn't even the possibility of multiple independent variables cointegrating.

Unlike the Kelly and Enns work, these authors pre-tested their data for unit roots, and drew conclusions about the $I(0)$ and $I(1)$ properties of their univariate series. This represents a positive step for the community of scholars to be able to evaluate the models considered. And yet, as we have shown above, the key tables in the article do not meet the requirements to have balanced equations. It should be noted that these authors have since written an erratum that addresses and corrects exactly this issue (Soroka *et al.*, 2016). Specifically, the authors applied additional tests of integration to the variables and, keeping the principle of balance in mind, came to the conclusion that the economic variables were, in fact, stationary. This then balanced their empirical models—both the left-hand-side and right-hand-side are $I(0)$. Importantly, it also leads to a different conclusion regarding the relationship between the economy and media tone. While the original results suggested that media tone with a constant long-run equilibrium was being driven by variables without any such long-run equilibria (an impossibility), or that the economic variables cointegrated with each other or some unknown variable to drive media tone, the new results suggest that media tone is driven by economic variables which also have constant long-run equilibria.

7 Conclusion

The recent symposium in *Political Analysis* highlighted the problematic use of the GECM in a large body of the applied literature that uses time series models, and identified the issue of equation balance as among the largest of those problems. And yet many practitioners may have been left without clear guidance about the definition of equation balance, and how to assess whether a theoretical and empirical model is balanced. We hope this project begins to fill this lingering void.

We hope, too, that this work will help to steer applied researchers away from the knife-edged question of whether or not they should use the GECM in their time-series models.

Rather, our view is that the GECM is a perfectly defensible modeling choice when the equations are balanced. (In that context, that either requires the presence of cointegration among the right-hand-side variables to produce a right-hand-side that is integrated at order 0, or that all right-hand-side variables be $I(0)$.[13]) The GECM can be a useful model in these circumstances.

Moving beyond the question of the utility of the GECM that so dominated the symposium, it should be evident that the principles of balance are a useful tool regardless of which time series model the researcher is considering using. And yet we hasten to add that equation balance is merely a necessary, and not a sufficient, condition for having a good model. Balance is the beginning of the process to determine if an estimated model is a good representation of reality; it is not the end of that process.

References

BANERJEE, A., DOLADO, J., GALBRAITH, J. W. and HENDRY, D. (1993). *Co-integration, Error-correction, and the Econometric Analysis of Non-stationary Data*. Oxford, UK: Oxford University Press.

BOX-STEFFENSMEIER, J. M. and SMITH, R. M. (1998). Investigating political dynamics using fractional integration methods. *American Journal of Political Science*, **42** (2), 661–689.

BRANDT, P. T. and FREEMAN, J. R. (2006). Advances in bayesian time series modelling and the study of politics: Theory testing, forecasting, and policy analysis. *Political Analysis*, **14**, 1–36.

— and — (2009). Modelling macro-political dynamics. *Political Analysis*, **17**, 113–142.

DE BOEF, S. and KEELE, L. (2008). Taking time seriously. *American Journal of Political Science*, **52** (1), 184–200.

ENDERS, W. (2004). *Applied Econometric Time Series*. Hoboken, NJ: John Wiley & Sons, Inc., 2nd edn.

ENGLE, R. F. and GRANGER, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, **55** (2), 251–276.

[13] And in the latter case, we agree with the critics of the use of the GECM that the interpretation of the error-correction coefficient changes substantially.

ENNS, P. K. and WLEZIEN, C. (2017). Understanding equation balance in time series regression. https://thepoliticalmethodologist.com/2017/06/23/understanding-equation-balance-in-time-series-regression/.

FREEMAN, J. R. (2016). Progress in the study of nonstationary political time series: A comment. *Political Analysis*, **24** (2), 50–58.

GRANT, T. and LEBO, M. J. (2016). Error correction methods with political time series. *Political Analysis*, **24** (1), 3–30.

GREENE, W. H. (2012). *Econometric Analysis*. Upper Saddle River, NJ, US: Prentice Hall, seventh edn.

JOHANSEN, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Political Analysis*, **59** (16), 1551–1580.

KEELE, L., LINN, S. and WEBB, C. M. (2016a). Concluding comments. *Political Analysis*, **24** (1), 83–86.

—, — and — (2016b). Treating time with all due seriousness. *Political Analysis*, **24** (1), 31–41.

KELLY, N. J. and ENNS, P. K. (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.

LEBO, M. J. and GRANT, T. (2016). Equation balance and dynamic political modeling. *Political Analysis*, **24** (1), 69–82.

MADDALA, G. and KIM, I.-M. (1998). *Unit Roots, Cointegration, and Structural Change*. Princeton, N.J.: Cambridge University Press.

PESARAN, M., SHIN, Y. and SMITH, R. (2001). Balance testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, **16** (3), 289–326.

PESARAN, M. H. (2015). *Time Series and Panel Data Econometrics*. Oxford: Oxford University Press.

PHILIPS, A. Q. (In Press). Have your cake and eat it too? cointegration and dynamic inference from autoregressive distributed lag models. *American Journal of Political Science*, pp. 1–15.

SIMS, C. A. and ZHA, T. A. (1999). Error bands for impulse responses. *Econometrica*, **67** (5), 1113–1156.

SOROKA, S. N., STECULA, D. A. and WLEZIEN, C. (2015). It's (change in) the (future) economy, stupid: Economic indicators, the media and public opinion. *American Journal of Political Science*, **59** (2), 457–474.

—, — and — (2016). Erratum to "it's (change in) the (future) economy, stupid: Economic indicators, the media and public opinion" american journal of political science, vol. 59, no. 2, april 2015, pp. 457–474. *American Journal of Political Science*, **60** (3), 803–804.

STIMSON, J. A. (1991). *Public opinion in America: Moods, cycles, and swings*, vol. 14. Westview Press.

WEBB, C. M., LINN, S. and LEBO, M. (2017). Out of bounds? testing for long run relationships under uncertainty over univariate dynamics. Madison, WI: Political Methodology Society.

WLEZIEN, C. (1995). The public as thermostat: Dynamics of preferences for spending. *American journal of political science*, pp. 981–1000.