# Fractional Time Series Applications in Political Science

A Dissertation presented

by

**Taylor Grant**

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

**Doctor of Philosophy**

in

**Political Science**

Stony Brook University

**August 2015**

**Stony Brook University**

The Graduate School

Taylor Grant

We, the dissertation committe for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

**Matthew Lebo - Dissertation Advisor**
**Professor, Political Science**

**Jeffrey Segal - Chairperson of Defense**
**Distinguished Professor, Political Science**

**Michael Peress**
**Associate Professor, Political Science**

**Paul Kellstedt - Outside Advisor**
**Associate Professor, Political Science, Texas A&M**

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

**Fractional Time Series Applications in Political Science**

by

**Taylor Grant**

**Doctor of Philosophy**

in

**Political Science**

Stony Brook University

**2015**

This dissertation is focused on the estimation of time series models with the main area of interest being the use of fractional methods in political science. Since the work of Granger (1980), Granger and Joyeux (1980) and Hosking (1981) initially introduced the concept of fractional integration the area of study has continued to grow in economics. In political science however, research into fractional integration has failed to advance beyond an initial flurry of study that began in the late 1990s and early 2000s. This dissertation contains three stand alone chapters that are focused on the use of fractional methods with political time series. Each attempts to address concerns that have previously dissuaded broader usage of fractional techniques.

Chapter 2 attempts to move beyond the aggregation work of Granger (1980) by offering a second justification for the presence of fractionally integrated political series - the error duration process of Parke (1999). Using the error duration model to investigate the survival of federal policies, survival probabilities of American federal programs are found to be fractionally integrated. Chapter 3 provides the results of a Monte Carlo study of various parametric and semiparametric estimators of long-memory with smaller sample sizes commonly found in political science. This is the first comprehensive simulation of multiple estimators with samples of less than 100. The results indicate that fractional integration can be reliably estimated with data of smaller samples, however the proper frequency domain estimators must be used depending on the data type. Chapter 4 is co-authored with Matthew Lebo. All model estimates, Monte Carlo simulations, figures, and all material found in the appendix for this chapter were solely my own work product. The chapter discusses the single-equation error correction model that is now common in political time series, and diagnoses several problems with the way the method has been employed up to this point. Alternative models are offered in place of the general error correction model.

## Dedication Page

Thanks to my wife, Katie, for putting up with me throughout this whole ordeal.

# Contents

# List of Tables

# List of Figures

## Chapter 1

# Introduction

The process by which political scientists use new statistical methods is predominately one of assimilation rather than pure development. Methodological and theoretical advancements are made - most often in the field of econometrics - and following a lag of several years, some intrepid researcher integrates the new method into the field of political science. And nowhere is the borrowing of methods more apparent than in political time series where almost all methods were first developed by econometricians. For example, Box and Jenkins (1970) developed and popularized ARMA and ARIMA modeling (see also, Box and Pierce 1970; Anderson 1971), and several years later one first encountered these models in the *American Journal of Political Science* (Li 1976). The same holds for other methodological advances such as Kalman filtering,[1] autoregressive conditional heteroscedastic (ARCH) models,[2] the single-equation error correction model,[3] and cointegration[4] to name just a few.

These methods are now widely accepted within political science, with the result being that political time series research has grown increasingly sophisticated. And such sophistication has undoubtedly yielded positive results, not only in terms of research quality, but also by expanding the range and scope of testable hypotheses. It is not uncommon for newer time series research to move beyond the simple estimation of means and delve into the dynamics

[1]Possibly developed by Kalman and Bucy (1961) and popularized by Jones (1980), the method found its way to political science with Beck (1989) and Kellstedt, McAvoy, and Stimson (1993).

[2]Established by Engle (1982), and first recognized by Beck (1983), but not applied to political data until a working paper by Brehm and Gronke (1994)

[3]First developed in econometrics by Davidson, Hendry, Srba, and Yeo (1978) and Hendry and Mizon (1978) it was introduced to political science, like so many other time series methods, by Beck (1992).

[4]Formalized by Granger (1986) and Engle and Granger (1987), and first discussed in a special edition of *Political Analysis* (1992).

and volatilities of a series. Questions previously out of reach are now being addressed thanks in part to the adoption of advanced econometrics. And yet, from this methodological toolbox which political scientists may now draw from, one method is conspicuously absent - the estimation of the long-memory, or fractionally integrated process.[5]

Initially, the adoption and integration of fractional methods into political science followed a trajectory similar to those methods outlined above. The first influential contributions came from Granger (1980), which justified the presence of fractionally integrated series through an aggregate process, as well as the work of Granger and Joyeux (1980) on long-memory series. Hosking (1981) simultaneously made advances in ARFIMA modeling from the field of hydrology. By 1996 the concept of fractional integration was sufficiently well-known in economics to justify a special edition of the *Journal of Econometrics* devoted to the subject. That same year, the benefits of fractional methods were first recognized in the political science literature. As an alternative to the $I(0)/I(1)$ dichotomy, it was argued that fractional integration offered a more accurate method of estimating and accounting for the degree of persistence in time series data such as macropartisanship (Box-Steffensmeier and Smith 1996, (BSS)).

Prior to BSS, theories of macropartisanship were predicated on the strength of political identification at the individual level; if individual identification was strong and enduring then aggregate partisanship would be steady as well, whereas if individuals were more fickle in their affiliations then mass partisanship was potentially fluid. BSS built their theory on an expectation that individual partisan identification was heterogeneous and that, as a consequence, exogenous events would not elicit a uniform response across all voters. Not only would voters differ in their perception of which events were important, voters would also differ in the strength of their partisan identification. Based on Granger's (1980) work on micro-aggregation, BSS argued that this variance in the strength and persistence of individual partisanship would lead to fractionally integrated series in the aggregate. Their

[5] I use the terms long-memory and fractionally integrated interchangeably throughout this dissertation. An in-depth primer on fractional integration and its properties is provided in Appendix A.

estimates of the level of memory of macropartisanship told a story in which the political allegiances of the population were neither permanent nor transitory. Instead, the public's partisan identification was persistent, but given a long enough time horizon, still considered mean reverting.

Independent of BSS, Byers, Davidson, and Peel (1997) pursued the same line of reasoning for their work on political popularity in the United Kingdom.[6] Identifying voters as either "committed" or "floating," Byers et al., hypothesized that voters would not react uniformly to changes in political circumstances and that when the opinions of these various voters were aggregated, the series would fit within Granger's theroem. Byers et al., estimated the popularity of the Conservative and Labour parties and found that each was a fractionally integrated process, and the estimated persistence of their series were strikingly similar to those of BSS. Thus, in very short order political scientists had evidence of several long memory time series, and they also had a valid theoretical process justifying their generation.

Spurred on by this research, others began investigating the fractional dynamics of various commonly used political time series. Lebo, Walker, and Clarke (2000) demonstrated that series such as presidential approval, macropartisanship, Supreme Court ideology, and consumer sentiment were all fractionally integrated; Box-Steffensmeier and Tomlinson (2000) estimated that congressional approval and economic expectations were potentially fractionally integrated; and Byers, Davidson, and Peel (2000) went on to investigate the political popularity series of 26 parties in 8 different countries, finding that many of those series were fractional processes. By the end of the decade, research into fractional integration was becoming commonplace.

This line of research also focused a good deal of attention on best practices and the empirical and inferential consequences of failing to account for long memory. Box-Steffensmeier and Smith (1998) demonstrated the significance of modeling choices; ignoring the presence of fractional integration in favor of ARMA or ARIMA models produced drastically differ-

[6]What Byers, Davidson, and Peel (1997) refer to as "popularity" is measured as the individual response to a question of voting intention, a binary choice of voting for either the Labour or Conservative Party.

ent estimates for the independent variables, and it also opened the researcher up to issues caused by over-differencing. Lebo, Walker, and Clarke (2000) used Monte Carlo simulations to show the increased rate of spurious regressions when series are fractionally integrated. And a consistent theme throughout all of this research was the importance of using multiple methods when testing hypotheses - either in using multiple stationarity tests on series, different estimators when estimating the level of memory, or various information criteria when investigating the presence of higher order approximations.

Despite this promising beginning however, fractional methods failed to gain widespread acceptance with political scientists. Indeed, many at the time were concerned that the small sample properties of fractional methods were unknown, a particularly relevant concern considering the limited number of observations often found in political time series. Others questioned whether fractional methods provided any new insights (Maddala 1998). Lacking a sufficient buy-in from researchers further development of fractional methods within political science largely stagnated. It has since become a niche field of study within political time series; researchers are generally aware of the concept of fractional integration, but few apply it to their own empirical work.[7]

The same is not true in econometrics where research into fractional methods continues apace. Empirical work on fractional integration and cointegration can be found on foreign exchange markets, inflation rates, income and consumption, volatilities, commodities pricing, and internet traffic flows among others (see Gil-Alana and Hualde 2009, for a comprehensive survey of applied research), and as the properties of fractional integration have become better understood the methods have evolved. Best practices are more well defined, pitfalls of estimation - and their possible solutions - have been identified, and researchers, once limited in their options, now have a suite of estimators to choose from.

In spite of these advancements, fractional methods have been relegated to the metaphorical back bench of political science, but it is worth reconsidering their utility. There are strong

[7] Or, in the worst case, researchers misapply fractional methods and draw faulty conclusions from their results (see e.g., Volscho and Kelly 2012).

theoretical justifications for the expectation that (some) political time series are fractionally integrated, and these theories are strongly supported by empirical research. Further, researchers face the threat of inferential errors and spurious results should they fail to account for the presence of fractional integration. Turning a blind eye to the potential for fractional integration needlessly invites problems in modeling. Finally, the persistence of a series is a question that is inherently interesting on its own, and an ability to describe and quantify its properties provides valuable and useful information (e.g., Box-Steffensmeier and Smith 1996).

What follows is a series of papers, in chapter format, focused on the use of fractional methods with political time series. Each attempts to address concerns that have previously dissuaded broader usage of fractional techniques. Chapter 2 attempts to move beyond the aggregation work of Granger (1980) by offering a second justification for the presence of fractionally integrated political series - the error duration process of Parke (1999). The error duration model assumes a process that generates a series of stochastic shocks within a particular activity or area of focus. These shocks are of varying duration and strength and the variable of interest is the sum of all shocks that survive to time $t$. If some shocks survive for long durations, it is possible that the process contains long memory, i.e., the autocovariances are nonsummable. If this stochastic process contains long memory it is expected to have downstream effects within the system, imparting long memory into other series reliant on the original process. Using the error duration model to investigate the survival of federal policies, survival probabilities of American federal programs are found to be fractionally integrated. Based on this finding and the expectation that reliant series inherit the fractional properties of the underlying process, the persistence of quarterly non-defense federal consumption expenditures is tested next and it is also found to be fractionally integrated with a similar level of memory as our EDM estimates.

Chapter 3 provides the results of a Monte Carlo study of various parametric and semi-parametric estimators of long-memory with smaller sample sizes commonly found in political science. This is the first comprehensive simulation of multiple estimators with samples of less than 100. The results indicate that pure fractional series can be reliably estimated by multiple estimators. Such a finding is substantively important given the findings of Lebo, Walker, and Clarke (2000) and Byers, Davidson, and Peel (2000) that the vast majority of political time series are accurately approximated as (0,$d$,0) fractional processes. The simulation results urge caution for estimation in the the presence of higher frequencies as the semiparametric and exact maximum likelihood of Sowell (1992) are negatively biased and unreliable. However, in the presence of higher frequencies the frequency-domain Whittle maximum likelihood estimator is generally unbiased and outperforms all other estimators.

Chapter 4 discusses the single-equation error correction model that is now common in political time series. As we demonstrate, this model is severely limited in its applicability, however political scientists have assumed it can be fit to any type of time series data. The chapter discusses the misuse of the GECM, the common misunderstanding of its underlying assumptions in political science, and we test the model in a variety of circumstances using Monte Carlo simulations. We also provide alternative model specifications to test for the presence of error correction. The first is the fractional ECM (FECM), which relaxes the assumption of $I(1)$ to $I(0)$ cointegration and allows for fractional cointegration. Second is the autoregressive distributed lag model (ARDL) bounds test of Pesaran, Shin, and Smith (2001) which offers the researcher more flexibility in terms of the stationarity properties of the IVs used on the right hand side of the model.

Chapter 5 offers concluding remarks.

## Bibliography

Anderson, T. W. 1971. *The Statistical Analysis of Time Series.* New York: Wiley.

Beck, Nathaniel. 1983. "Time-Varying Parameter Regression Models." *American Journal of Political Science* 27(3): 557–600.

Beck, Nathaniel. 1989. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1(1): 121–156.

Beck, Nathaniel. 1992. "Comparing Dynamic Specifications: The Caseof Presidential Approval." *Political Analysis* 3: 27–50.

Box, George E. P., and Gwilym M. Jenkins. 1970. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Box, George EP, and David A Pierce. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." *Journal of the American Statistical Association* 65(332): 1509–1526.

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

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

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

Brehm, John, and Paul Gronke. 1994. "History, Heterogeneity, and Presidential Approval: a Modified ARCH Approach.".

Byers, David, James Davidson, and David Peel. 1997. "Modelling Political Popularity: An Analysis of Long-range Dependence in Opinion Poll Series." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160(3): 471–490.

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

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

Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica: Journal of the Econometric Society* pp. 987–1007.

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

Gil-Alana, Luis A, and Javier Hualde. 2009. "Fractional Integration and Cointegration. An Overview and an Empirical Application." In *Palgrave Handbook of Econometrics*, ed. Terence C. Mills, and Kerry Patterson. Vol. 2 MacMillan Publishers London pp. 434–469.

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

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

Hendry, David F, and Grayham E Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England." *The Economic Journal* 88(351): 549–563.

Hosking, Jonathan RM. 1981. "Fractional Differencing." *Biometrika* 68(1): 165–176.

Jones, Richard H. 1980. "Maximum Likelihood Fitting of ARMA Models to Time Series With Missing Observations." *Technometrics* 22(3): 389–395.

Kalman, Rudolph E, and Richard S Bucy. 1961. "New Results in Linear Filtering and Prediction Theory." *Journal of Fluids Engineering* 83(1): 95–108.

Kellstedt, Paul, Gregory E. McAvoy, and James A. Stimson. 1993. "Dynamic Analysis with Latent Constructs: The Kalman DYMIMIC Specification." *Political Analysis* 5: 113–150.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

Li, Richard P. Y. 1976. "A Dynamic Comparative Analysis of Presidential and House Elections." *American Journal of Political Science* 20(4): 671–691.

Maddala, G.S. 1998. "Rcent Developments in Dynamic Econometric Modelling: A Personal Viewpoint." *Political Analysis* pp. 59–87.

Parke, William R. 1999. "What is Fractional Integration?" *Review of Economics and Statistics* 81(4): 632–638.

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

Sowell, Fallaw. 1992. "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models." *Journal of econometrics* 53(1): 165–188.

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.

Chapter 2

# Error Duration Models: A New Justification for Fractional Integration in Political Time Series

**Abstract**

The primary justification for the expectation that political time series may be fractionally integrated is based upon the aggregation theorem of Granger (1980), which has been demonstrated empirically with public opinion data by Box-Steffensmeier and Smith (1996) as well as Byers, Davidson, and Peel (1997). This chapter investigates a secondary justification for the presence of fractional integration - the error duration model (EDM) of Parke (1999). The variable of interest in such a model is the sum of shocks which survive to a specific time point, $T$. Given a process that generates a sequence of stochastic shocks, which are themselves stochastic in terms of duration and magnitude, a process can exhibit long-memory if a small percentage of shocks survive for long durations. Such a representation potentially explains the presence of fractionally integrated political time series that are not the product of aggregation. If events generate shocks of varying duration, it is possible for these shocks to then transmit long-memory to other political variables. The EDM representation is applied to a data set capturing the post-enactment history of federal domestic spending programs from 1971 to 2004. The survival rates of federal policies are demonstrated to be fractionally integrated.

## 2.1 Introduction

What sort of process generates fractionally integrated (FI) data? The most common mechanism is based upon the aggregation theorem of Granger (1980), in which heterogeneous AR(1) series are aggregated and the aggregate that is formed is then long-memoried.[1] Based upon this theorem a number of aggregated political opinion series have been tested, and many have been found to be fractionally integrated (Box-Steffensmeier and Smith 1996; Byers, Davidson, and Peel 1997, 2000; Lebo, Walker, and Clarke 2000; Lebo and Cassino 2007). The aggregation theorem is useful in explaining series such as macropartisanship and party popularity, but other series within political science, which are not the product of aggregation, have also been found to be fractionally integrated. For instance, Dickinson and Lebo (2007) argue that the time series capturing the size of the Executive Office of the Presidency is a fractionally integrated series. But the EOP series is not an aggregate in the traditional sense of unemployment or public opinion, which indicates that there may be some other underlying process generating its long memory.[2]

One particularly promising answer comes from Parke (1999), who introduced the error-duration representation for fractionally integrated processes. In this model, sequences of stochastic shocks of varying magnitude and duration can generate a fractionally integrated series if a certain percentage of of the shocks survive over long durations. If the survival probabilities decay slower than an AR(1) series (the rate is given more specifically below), the series is a candidate to be fractionally integrated. The difference in survival probabilities between the FI and AR series is due to the conditional survival probabilities for those $I(d)$ shocks that survive the first few periods. If the series is FI, the conditional survival probabilities will increase towards unity as the length of time increases. Parke (1999) applies the error-duration representation to the survival rate of United States businesses as a demonstration of how aggregate employment can be fractionally integrated.

[1] For an explanation of Granger's aggregation theorem see Appendix A.2.2.

[2] Dickinson and Lebo (2007) argue that the EOP series is an aggregate series, however the error duration representation is a much more plausible process for the generation of fractional integration.

This chapter uses the error-duration model to estimate the persistence of federal programs created or already in existence between the years of 1971 and 2004 (Berry, Burden, and Howell 2010, (BBH)). BBH note previous research within the field of American Policy Development that policy change is "sticky," and that the inherent inertia and path dependence of government programs make policy resilient to change (Jones, Sulkin, and Larsen 2003; Pierson 2004). The authors employ a survival analysis and find the conventional wisdom that policies, once enacted, are "immortal" is overblown, and that a substantial number of programs do in fact face substantive changes. However, BBH also find that following an introductory period of approximately five congresses, programs are increasingly likely to endure without alteration. BBH are effectively describing an increasing conditional survival probability, which strongly indicates that the survival rates of federal policy are fractionally integrated.

The rest of this chapter proceeds with a brief discussion of policy development and survival followed by a description of the error-duration model of Parke (1999). I then evaluate the BBH data and provide evidence that the "stickiness" of policy development is evidence of a long memory process.[3]

## 2.2 Policy and Program Survival

To what extent do federal policies and programs survive once enacted? It is a question that until recently had not received much rigorous attention. In the past, when the topic of survival was taken up the focus often tended towards agency lifespans as opposed to those of programs. The work was largely atheoretical and based on a limited number of case studies, and the common wisdom that prevailed was one of agency immortality. One of the first analytical reviews of the life cycle of executive departments found that over a fifty-year period, of 175 organizations under review only 27 had been terminated (Kaufman 1976). Kaufman argued that the survival rates of governmental agencies were the product of incrementalism, whereby limitations implicit in the governmental decision-making process

[3] An in depth primer on fractional integration is provided in Appendix A.

served to insulate organizations from termination and that any deaths that did occur were more likely due to bad luck than any systemic factor. Kaufman's work was instrumental in creating the common perception of stasis within governmental organizations.

Incrementalism assumes that governmental durability is a product of the inertia of the status quo, in which varying interests align to make change difficult. At any time, legislators face considerable political and informational uncertainty over their decisions, meaning that the consequences of disruption are often in doubt. This uncertainty limits the range of decision options (Simon 1985; Jones 1999). As strategic actors, politicians may therefore be dissuaded from any attempt to deviate too far from the current equilibrium (Wildavsky 1964; Knight 1992).

We should also expect that additional contextual factors will determine programmatic stability. Details such as the direction that the potential disruption comes from, the type of programs in question, and the constituency affected by the disruption are each relevant considerations (Corder 2004; Maltzman and Shipan 2008). Thus, the status quo at the heart of incrementalism is not the sole determinant; additional procedural requirements and roadblocks such as super-majority requirements or multiple veto points throughout the system will also influence the extent of overall stability of government (Binder 1997; Baron and Ferejohn 1989).

The field of American Political Development (APD), which itself borrows from evolutionary biology, provides additional explanations for governmental stability.[4] In a system with a large number of organizations and/or policies fighting for a finite supply of resources, it is those programs that can demonstrate early successes or raise support from broad or powerful constituencies that benefit. In such a system, there is a stability to government, but it does not apply to all institutions or programs uniformly. Nascent programs and agencies face tighter competition, and those that become better established also become more difficult to alter as they age. Furthermore, the political environment in which institutions and policies

[4] APD is focused primarily on institutions, as opposed to policy, but the theories are informative.

survive is complex and multi-layered. A multiplicity of jurisdictions and motivations serve to insulate institutions from change, and over time institutions and policies become more secure (Baumgartner and Jones 1993). As a result, policy is "sticky" and generally resistant to change (Jones, Sulkin, and Larsen 2003).

From the original conception of "immortality," theories of institutional and policy development have eventually shifted to one characterized by an initial hazard followed by increasing durability over time. Such a theory has generally found support in empirical studies. Corder (2004) applies survival analysis to federal credit programs and finds that the hazard of program termination spikes early and then decreases after an initial period of program immaturity. Maltzman and Shipan (2008) study the evolution and amendment of major laws, and while their models do not explicitly model the age of the laws in question, their Figure 1 demonstrates that laws are more likely to be amended in the years immediately following passage with diminishing probability of amendment as age increases. And in the most comprehensive study to date, Berry, Burden, and Howell (2010) investigate the survival of all federal programs in existence from 1971 to 2003 and find that the probability of survival increases over the lifetime of the program. After surviving an initial period of 10 to 12 years, federal programs become increasingly resilient to modification or termination.

**Figure 2.1: Smoothed Hazard Function**


Figure represents the smoothed hazard function of all programmatic mutations and deaths since enactment. Data and code for figure from Berry, Burden, and Howell (2010).

These theories underlying institutional and policy development as well as their empirical findings raise the distinct possibility that policy development, as measured by program survival, is a fractionally integrated process. First, as recognized by the BBH study, at any point in a congressional term members of Congress are working in a policy environment that comes pre-stocked with programs. BBH's findings, evidenced in Figure 2.1, indicate that after an initial period during which programs face the greatest threat to their survival, the hazard rate declines as programs and policies age. Considering that the policy environment varies greatly in programmatic age, this means that at any given time there are certain programs that are virtually guaranteed survival, while others that were more recently enacted are under greater threat. This finding suggests a conditional survival probability that is increasing over time; every additional year that a program survives it becomes more secure and less likely to undergo alteration or death. An increase in the conditional survival probabilities of programs as time periods increase will in turn affect the autocorrelation function by building persistence into the memory of the process. Second, these findings support the evolutionary theories of APD, which recognizes the importance of temporal ordering in the development of policy. The more entrenched programs, or those with the time to build constituencies are the best positioned to survive. This indicates that policy and institutional development are well described by the concept of path dependence (Pierson 2000, 2004; Kay 2005). As noted by Freeman (2012), the presence of fractional integration implies path dependence. If the data on policy survival are fractionally integrated, we have an interesting opportunity to quantify the "stickiness" of federal policy development.

## 2.3 Error Duration Model

Parke (1999, p. 632) lays out the error duration model as follows. Assume a series of *iid* shocks $\{\epsilon_t, t = 1, 2, ...\}$ with mean zero and finite variance $\sigma^2$. Each shock has a stochastic survival duration, surviving from period $s$ to period $s + n_s$. Let $g_{s,t}$ be an indicator function for whether a shock that originates in time $s$ survives to time $t$. Thus $g_{s,t} = 1$ if $t \leq s + n_s$

and $g_{s,t} = 0$ if $t \geq s+n_s$. Shocks and their probability of survival are considered independent and the probability $p_k$ that the shock survives to period $s+k$ is $p_k = P(g_{s,s+k} = 1)$. All shocks are assumed to survive their initial inception, $p_0 = 1$, and the sequence of survival probabilities $(p_k)$ for $k > 0$ is monotonic non-increasing. The realization $y_t$ is the sum of all errors that survive until period $t$:

$$y_t = \sum_{s=-\infty}^{t} g_{s,t}\epsilon_s. \tag{2.1}$$

The link between the error duration model and other fractional representations is through the autocovariances. As discussed in Appendix A, a fractionally integrated process is one in which the autocorrelations and autocovariances decline slowly enough that they will not sum to a finite value. In the case of the error duration model, Parke demonstrates that the summation of the autocovariances is the sum of the survival probabilities.[5] If the survival probabilities decline slowly enough, defined as $p_k = k^{-2+2d}$, then the autocovariances are nonsummable, meaning that the process is fractionally integrated in the range $0 < d < 1$. More specifically, if the probabilities decline to zero more slowly than $k^{-2}$ the process is long memory and stationary $(0 < d < 1/2)$; if they decline more slowly than $k^{-1}$, then $y_t$ is long memory and non-stationary $(1/2 \leq d < 1)$.

Additionally, if the process is fractionally integrated this also means the conditional survival probabilities $(p_{k+1}/p_k)$ will eventually converge to unity given enough time.[6] If a

[5]The process is long memory if the summation of the survival probabilities $\sum_{k=1}^{n} kp_k$ is not finite. See Appendix A, Section, A.3 for more details on the summation.

[6]Parke derives the survival probabilities of this representation as

$$p_k = \frac{\Gamma(k+d)\Gamma(2-d)}{\Gamma(k+2-d)\Gamma(d)}. \tag{2.2}$$

and the autocovariances

$$\gamma_k = \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)}\sigma_0^2 \tag{2.3}$$

which generate an $I(d)$ process precisely.

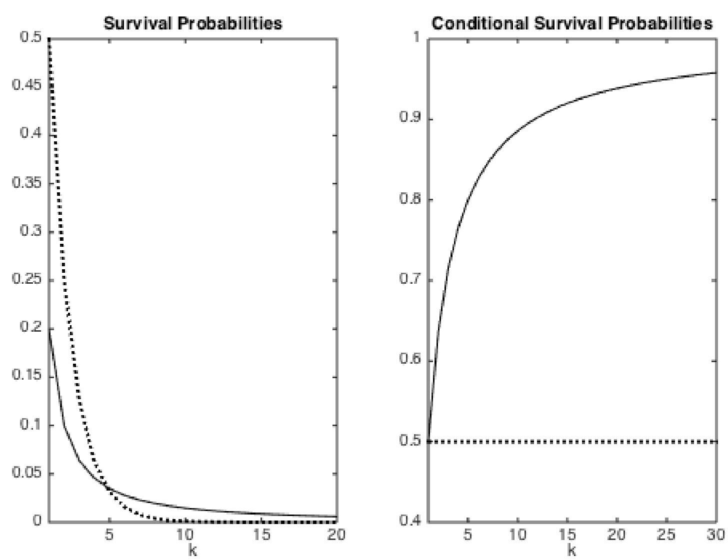
Key to the conditional survival probabilities, Parke offers the following recursion beginning at $p_0 = 1$:

$$p_{k+1} = \frac{k+d}{k+2-d}p_k \tag{2.4}$$

shock survives through time period $t$, it's chances of surviving period $t+1$ will increase, and the survival probability increases again in $t+2$ if it survives $t+1$, and on and on. This is clearly seen in the recursion provided by Parke highlighted in footnote 6; as $k$ increases, the ratio of each successive probability will increase as well. This is representative of the fact that those shocks that survive for long periods tend to continue to survive, and in the case of federal programs, is indicative of the declining hazard faced by programs as they age.

The increasing chance of survival found in a fractionally integrated process can be contrasted with the conditional survival probabilities of an AR process, which are expected to be constant at $\phi$. For comparison, the survival and conditional survival probabilities of an $I(d)$ and an AR series are presented in Figure 2.2. The fractional process is $d = .33$ and the AR, $\phi = 0.5$, thus both series have the same first-order autocorrelation of 0.5.

**Figure 2.2: Survival Probabilities of Error Duration Model**




Solid line is the survival and conditional survival probabilities of FI series, $I(d) = .33$. Dotted line is of an AR process $\phi = 0.50$. Note that the conditional survival of the FI series converges towards one whereas the conditional survival probabilities for the AR process remain constant.

The survival probabilities for the AR(1) model are $p_k = \frac{1}{2}^k$ and as is clear from the left hand panel, the probabilities for the AR(1) process are initially much greater than those of the fractionally integrated process. Despite beginning lower, the FI series exhibits a much slower decline, and by the fifth time period the probability of survival for the FI series is greater than that for the AR series. By the tenth period, the probability of survival for the AR process $p_{10} = 0.001$, whereas the FI series is $p_{10} = 0.015$.

The effect of the survival probabilities is also apparent in the autocorrelation functions of the two series. For the AR(1) process, the values $[\rho_1, \rho_{10}, \rho_{20}, \rho_{100}]$ are expected to be $[0.50, 10^{-3}, 10^{-6}, 10^{-30}]$. For the FI series, the same values are estimated [0.50, 0.234, 0.186, 0.109]. This disparity in the higher-order autocorrelations is due to the the high conditional probabilities following the survival of the first few periods.

### 2.3.1 Error Duration and Fractional Integration

Before applying Parke's error duration model to the federal program data, it is useful to consider how such a process generates fractional integration in other series. As the source of an example, consider the BBH data on the survival of federal domestic programs. At any given time, the federal policies that are in existence are comprised of programs of different ages. Some were enacted several decades ago and have survived to this point, while others were only recently enacted. Still other programs were either substantially modified or were not reauthorized. Members of Congress do not take over each successive Congress with a blank slate, but must legislate within the constraints of the programs already in place.

Now assume that we are interested in a time series capturing some level of governmental outlays, such as the level of non-defense spending on government consumption expenditures. The level of consumption expenditures at time point $t$ is partially dictated by the sum of all of the past and present grant programs ("shocks") to have survived to that time.[7] If the survival probabilities of federal programs decay at a fractional rate, then we might also expect that the level of consumption expenditures are fractionally integrated as well. These

[7]Federal consumption expenditures (non-defense) is estimated by the Department of Commerce through its Bureau of Economic Analysis. Government consumption expenditures accounts for the services produced by government and provided without charge to the public-at large. Services include public schools, universities, job training, transportation infrastructure, public health and a legal and judicial system among others. The expenditures are accounted for in the National Income and Product Accounts (NIPAs) which are quarterly measures of dollars spent.

downstream effects are potentially found in any political time series that is dependent on the survival rates of federal programs.

## 2.4 Estimating the EDM

In order to test Parke's error duration representation, I use the data set of Berry, Burden, and Howell (2010), which captures the survival of all federal programs from 1971 to 2004 at the yearly level. The BBH data come from the *Catalog of Federal Domestic Assistance* (CFDA), an annually updated compendium of federal programs. The CFDA is comprehensive and accounts for nearly all domestic programs of the federal government.[8]

Rather than use year-by-year survival rates, the data is aggregated by Congress (97th - 107th). The data set represents 2,038 unique federal programs. Despite BBH's finding of a significant amount of change occurring within the federal programs, actual program deaths are far less likely to occur. Because of the paucity of actual program deaths, I estimate the empirical survival rates of the number of programs that survive to time point $t$ without undergoing either death, or what BBH refer to as a "mutation." Mutations are defined and coded by BBH as substantive programmatic changes such as consolodations, splits and/or transfers. Simple renumberings or renamings are not counted.

Specifically, our dependent variable is the fraction of federal programs starting in Congress 1 that survive without substantial mutations to Congress $k$. The survival rates are found in Table 2.1. In this case, our first Congress is the 107th and the empirical survival rate $S_k$ of the federal programs created in the 107th Congress surviving through the 108th is 0.954. Upon investigation, the survival rates are declining slower than if it were a simple AR process. If the program survival rates were to fit an AR(1) model, then we could take the five-Congress survival rate $S_{103} = 0.555$ and square it in order to obtain an approximation

[8]The CFDA tracks "any function of a a Federal agency that provides assistance or benefits for a State or States, territorial possession, county, city, other political subdivision, grouping, or instrumentality thereof; any domestic profit or nonprofit corporation, institution, or individual other than an agency of the federal government." And "assistance" refers to the transfer of value from the federal government to a domestic beneficiary. Military spending and defense procurement are not included. (CFDA 2005, 1; excerpted from Berry, Burden, and Howell (2010, p.5).

**Table 2.1: Survival Rates and Conditional Survival Rates for Federal Programs**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Congress** $k$ | 107th | 106th | 105th | 104th | 103rd | 102nd | 101st | 100th |
| Survival $S_k$ | 0.954 | 0.804 | 0.743 | 0.600 | 0.555 | 0.534 | 0.512 | 0.436 |
| $S_k/S_{k-1}$ | 0.954 | 0.843 | 0.923 | 0.808 | 0.925 | 0.962 | 0.960 | 0.851 |
| | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| **Congress** $k$ | 99th | 98th | 97th | 96th | 95th | 94th | 93rd | 92nd |
| Survival $S_k$ | 0.417 | 0.406 | 0.390 | 0.373 | 0.329 | 0.320 | 0.293 | 0.280 |
| $S_k/S_{k-1}$ | 0.957 | 0.974 | 0.960 | 0.957 | 0.882 | 0.974 | 0.915 | 0.953 |

Survival $S_k$ indicates the empirical survival rate of the fraction of federal programs created in each respective Congress to survive through the 108th Congress. $S_k/S_{k-1}$ indicates the conditional survival rates.

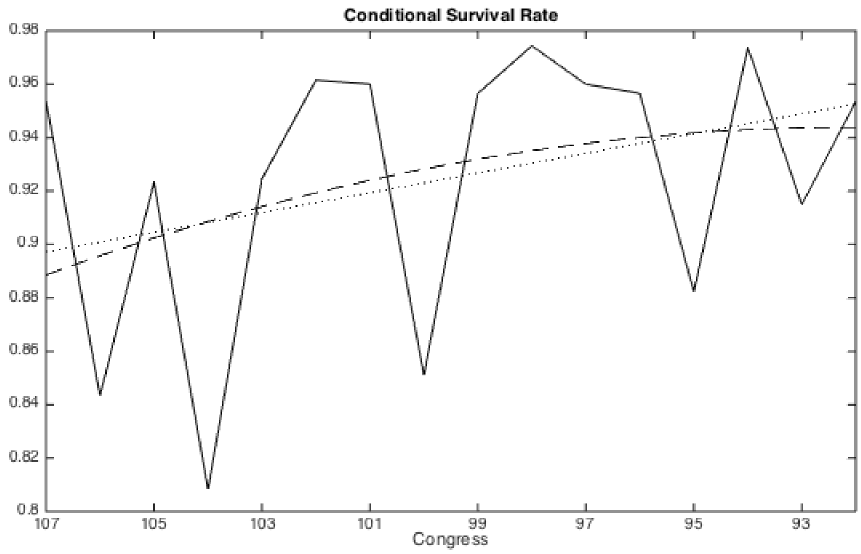
of the ten-Congress survival rate. The squared estimate is 0.308 while the actual survival rate is 0.406, 25% larger than the AR(1) expectation. The disparity in expected decay grows larger as lags increase, were we to do the same thing for the 100th and 92nd Congresses - $0.436^2 = 0.190$, or 30% below the actual survival rate of the 92nd Congress: $S_{92} = 0.280$.[9]

The conditional survival rate is not as clean in its approach to unity as would be preferred, however this might be expected provided research finding that political and economic considerations both at the time of, and subsequent to enactment, will affect whether laws are amended and programs altered (Maltzman and Shipan 2008; Berry, Burden, and Howell 2010). Considering that this data covers a wide period of time encompassing the Vietnam War, Watergate, the energy crisis and oil embargo, periods of stagflation, and a stock bubble it is not entirely inconsistent to find that the conditional survival rate is not monotonically increasing. That said, the trend is clearly one of increasing conditional probability, as is evidenced in Figure 2.3. Two trend lines, a linear and quadratic, are included, and each indicates a moderate, yet steady slope in the positive direction. The trends appear consistent with the hypotheses of Parke; if a survival process is fractionally integrated, the conditional survival probabilities will asymptotically tend towards unity.

Based on the comparatively slow decay of the survival rates and the upward trend of the conditional survival rates, it appears that that there is some evidence of long memory in

[9]In applying the EDM representation, the use of only 16 time points is not an issue. In fact, Parke (1999) provides an example of this process using only 10 time points.

**Figure 2.3: Conditional Survival Rate of Federal Program by Congress**




Solid line is the actual conditional survival rate of programs in the 108th Congress dependent on the Congress in which the programs were originally were created. Dotted line is the linear trend line. Dashed line is the quadratic trend line.

the survival rates of federal domestic programs. The next step is to fit the modeled survival rates to the approximation of fractional decay, $p_k = c_p k^{-2+2d}$. Parke provides us with a way to estimate $d$ by inverting the relation $p_n/p_m = (n/m)^{-2+2d}$ in order to obtain[10]

$$d = 1 + \frac{1}{2}\frac{log(p_n/p_m)}{log(n/m)} \tag{2.5}$$

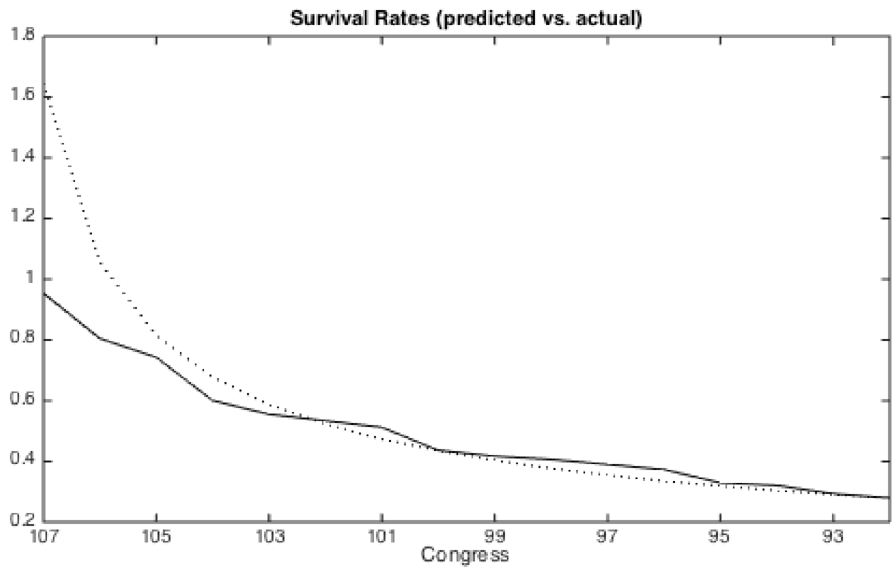
By substituting the empirical survival rates $S_{100}$ and $S_{92}$ in place of $p_8$ and $p_{16}$ in Equation 2.5 we estimate the order of fractional integration as $d = 0.68$.[11] From this estimate, it is then possible to predict the survival probabilities of federal programs, the results of which can be found in Figure 2.4. With constants greater than 1.0 there is an issue in the estimated survival rates of the first two congresses, but the fitted values settle pretty quickly thereafter. From the fifth lag, $S_{103}$ through to the end of the series $S_{92}$, the estimated function fits

[10]Parke notes that Equation 2.5 also reflects an important point about scaling. Were the scaling to change, for example from congressional terms to sessions of Congress, the ratios would remain the same. The result is a consistent and scaling invariant estimate of $d$, which is a prerequisite for fractionally integrated series. (see Chambers 1998, for more details).

[11]If we do this for a longer window, estimating over the range of all data, $p_2$ to $p_{16}$, our estimate of $d$ is slightly higher, $I(d) = 0.77$.

the actual survival rate to 0.01, indicating that the survival rates of federal programs are consistent, and fit well within the error duration model's estimation of long memory. These findings also fit nicely with the BBH results where programs enjoyed increasing safety after a period of 10 to 12 years of existence.

**Figure 2.4: Fitted Survival Rates**




Solid line is the actual survival rate of programs created in each Congress surviving through 108th Congress without substantial change. Dotted line is the predicted survival rate based on estimate of *d*.

Based on this model and its results, the survival rate of federal programs appears to be a fractionally integrated process. The survival probabilities decline at a rate sufficiently slow to be considered fractional, and the conditional survival probabilities tend towards unity, which indicates that the autocovariances (the sum of survival probabilities) are non-summable. This finding is theoretically consistent with recent analyses describing a declining hazard rate as federal programs age, becoming more entrenched within the system. That the process is fractionally integrated is also in line with the descriptions found in the field of American Policy Development. However, whereas most descriptions of APD are qualitative, the EDM representation provides an empirically grounded model that fits the expectation of government policy development as "sticky."

## 2.5 Downstream Effects

With evidence of an underlying process that is fractionally integrated, we can now test the expectation of downstream effects - series that are reliant on the rates of program survival should inherit the same fractional persistence. In this case, the series of interest is real federal non-defense consumption expenditures over the same duration of time as the BBH data set (1971-2004).[12]

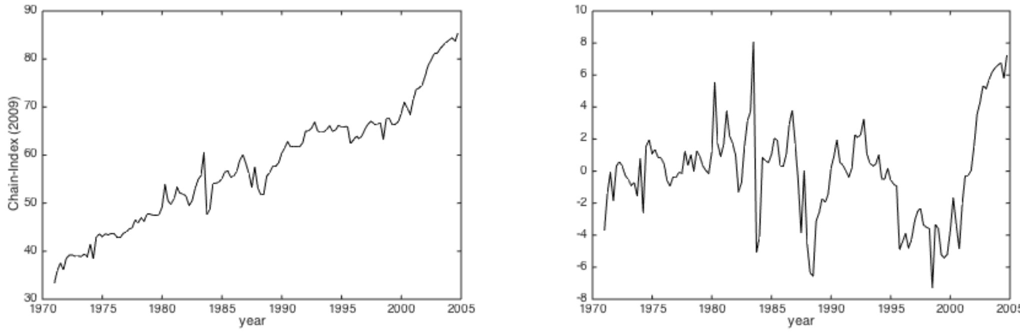
Non-defense government consumption expenditures (FNDCEX) consist of spending as well as services produced or provided by the federal government to the general public without charge; see footnote 7 for examples. Recall that the BBH data set captures any federal programs that are providing assistance in the form of non-defense spending. Thus, the level of federal outlays in the form of non-defense consumption expenditures are partially a result of the survival rates of federal programs. If the estimate of the EDM representation is proper then there is an expectation of finding that the expenditures series is fractionally integrated with roughly the same level of memory as estimated above. Note that this series is not an aggregate in the traditional sense as described by Granger (1980), but based upon our use of the EDM representation we have a strong theoretical justification for our expectations.

The visual presentation of the FNDCEX series is provided in Figure 2.5. The left panel is the series in levels. A linear time trend is evident, which can cause estimation problems if left unaddressed when using semiparametric estimators.[13] Because Robinson's (1995) local Whittle estimator is used as a robustness check of the results, the detrended series is used for the semiparametric estimates. The parametric estimates are the same regardless of which

[12] The series is calculated in a chain-type quantity index, chained to 2009. For more information on calculation of the series see here (http://www.bea.gov/national/pdf/mp5.pdf). For information on chain-indexing see here (http://www.bea.gov/scb/pdf/2003/11November/1103%20Chain-dollar.pdf.) The data can be downloaded from the Federal Reserve Bank of St. Louis here (https://research.stlouisfed.org/fred2/series/A542RA3Q086SBEA.)

[13] The problem with a trend comes about because semiparametric estimators are inconsistent and exhibit non-normal behavior if the series possesses a fractional order of integration, $d > \frac{3}{4}$. This is why the common method of estimation is to take the first difference and then add one back to the estimate. If a series is trend stationary however, such that $0 < d < 1/2$, taking the first difference puts the series' order $I(d) \in [-1, -1/2)$, at which point the estimates will be biased towards zero (Shimotsu and Phillips 2006; Shimotsu 2010).

**Figure 2.5: Real Federal Expenditures, Non-Defense**




Left plot real federal government consumption expenditures (non-defense 1971-2004 (chained to 2009 level). Right side is detrended series - the linear trend has been removed and the series forced to a mean of zero.

series, level or detrended, is used. The detrended series can be seen in the panel on the right of Figure 2.5.

The autocorrelation function of the level series is found in Figure 2.6. The series demonstrates a remarkable degree of memory, which is also consistent with our theory that the series is fractionally integrated.[14] Additionally, note that the autocorrelations are consistently positive as opposed to oscillating or sawtoothed. This indicates that if higher frequencies are present in the series, a $(1, d, 0)$ model is likely sufficient. It also leads us to expect that, if present, the AR parameter is positive. The sawtooth pattern is found with negative autocorrelations.

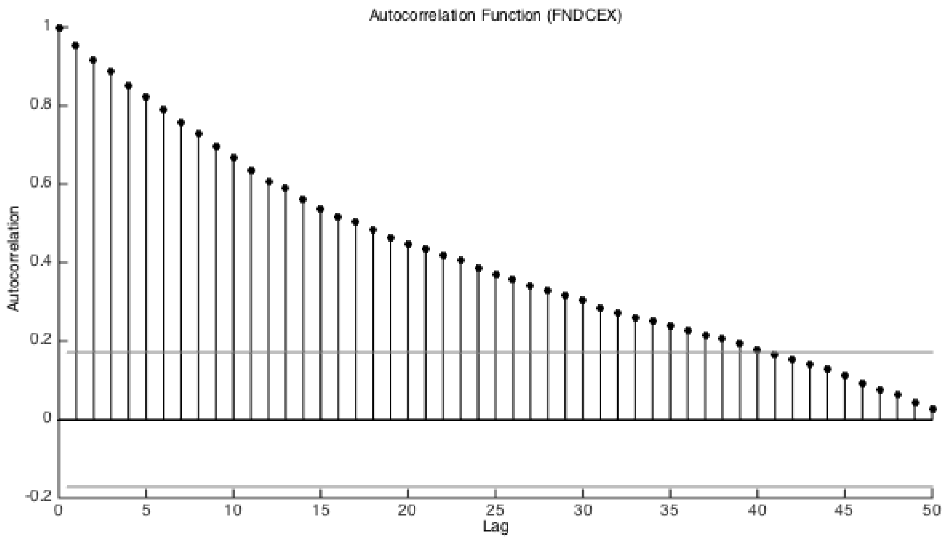
### 2.5.1 Testing Stationarity

At first glance, it appears that the FNDCEX might be trend stationary. A Dickey-Fuller test accounting for a linear trend rejects the null hypothesis of a unit root at the .05 level, however an augmented Dickey-Fuller fails to reject the null for any of the ten lags tested. The more powerful DF-GLS test (see Baum and Sperling 2000) is unable to reject the null of trend or mean stationarity.[15]

[14] For a demonstration in the difference in the ACF of an FI process versus an AR process see Figure A.2 in Appendix A.1.1.

[15] Test statistic (Crtitical Value) - DF notrend: -0.60 (-2.89); DF trend: -3.75 (-3.45); ADF notrend (1,5,10 lags): -0.18, 0.69, 0.53 (-2.89); ADF trend (1,5,10 lags): -2.92, -1.45, -2.01 (-3.45); DF-GLS notrend (SC chosen lag 2): 2.64 (-2.07); DF-GLS trend (SC chosen lag 1): -2.70 (-2.98).

**Figure 2.6: Autocorrelations of FNDCEX Series**




Autocorrelation plot of the FNDCEX series.

The stationarity of the series is next tested with the KPSS test of Kwiatkowski, Phillips, Schmidt, and Shin (1992). The null hypothesis that the FNDCEX series is trend stationary can be rejected at the .05 level for up to 4 lags. The null hypothesis that our series is level stationary is rejected for all lags. Lastly, the series is subjected to the variance-ratio test of Lo and MacKinlay (1988) for varying differencing parameters, $k$. Similar to the Dickey-Fuller, the variance ratio test has a null hypothesis of a unit root. The results of the KPSS and variance ratio tests are found in Tables 2.2 and 2.3 respectively.

**Table 2.2: KPSS Test for Strong Mixing**

| | Lags | | | | |
|---|---|---|---|---|---|
| | (0) | (2) | (4) | (6) | (8) |
| FNDCEX$^{\dagger}$ | 0.561$^{\star}$ | 0.232$^{\star}$ | 0.156$^{\star}$ | 0.123 | 0.105 |
| FNDCEX$^{\ddagger}$ | 12.5$^{\star}$ | 4.32$^{\star}$ | 2.66$^{\star}$ | 1.95$^{\star}$ | 1.56$^{\star}$ |

$^{\dagger}$ Null hypothesis is series is trend stationary.
$^{\ddagger}$ Null hypothesis is series is level stationary.
$^{\star}$ indicates significant at .05 level.

Because stationarity tests suffer from a lack of power when applied to fractionally integrated series, it is not surprising that we have contradictory findings for our series.[16] To summarize: the DF-GLS can't reject the null of a unit root when compared to the alternatives of either mean or trend stationarity, the DF test rejects the null in favor of trend stationarity; the KPSS test rejects the null hypothesis that our series is stationary at the fourth lag, which is the recommended lag given the number of observations (Kwiatkowski et al. 1992); and the variance ratio test rejects the null of a unit root at the fourth and eighth differencing interval, but not at the second or sixteenth. With the results of our stationarity tests inconclusive, it is a fair assumption that our series might be fractionally integrated.

**Table 2.3: Variance Ratio Test**

| | $k$ | | | |
|---|---|---|---|---|
| | VR(2) | VR(4) | VR(8) | VR(16) |
| FNDCEX$^{\dagger}$ | 0.832 | 0.556$^{\star}$ | 0.458$^{\star}$ | 0.479 |

$^{\dagger}$ Null hypothesis is series is a unit root.
$^{\star}$ indicates significant at .05 level.

### 2.5.2 Estimating Fractional Integration

We next estimate the order of fractional integration of the FNDCEX series using a variety of estimators - the time domain exact maximum likelihood estimator (EML), the frequency domain whittle likelihood estimator (FML), and the semiparametric local whittle estimator (LW). The main estimator of interest is the FML with the EML and LW estimates used as robustness checks. The Akaike Criterion (AIC) and Bayesian Information Criterion (BIC) from the FML were employed to determine model choice and the results are presented in Table 2.4. The AIC and BIC were both written following the conventions of Davidson (2014) and are coded such that "Larger = Better."[17]

[16] For more information on stationarity tests and their power in Monte Carlo simulations with fractionally integrated series, see Appendix A.1.3.

[17] All models were estimated in Matlab. I wrote both the EML and FML estimators and checked them versus estimators available in other programs. The EML is identical to Stata's "arfima" EML estimator. The FML provides almost identical estimates to the Whittle likelihood estimator used by Davidson (2014) in his proprietary statistics package "Time Series Modeling," which is an add on of OxMetrics (Doornik 2014). The semiparametric Whittle estimator is written by Katsumi Shimotsu and available on his home page at http://shimotsu.web.fc2.com/Site/Home.html

**Table 2.4: Whittle Likelihood Estimation: AIC and BIC values**

| | ARFIMA($p, d, q$) Model | | | | | |
|---|---|---|---|---|---|---|
| | 0,$d$,0 | 1,$d$,0 | 2,$d$,0 | 0,$d$,1 | 0,$d$,2 | 1,$d$,1 |
| FNDCEX (AIC) | -20.62⋆ | -21.53 | -21.18 | -21.42 | -20.63 | -20.95 |
| FNDCEX (BIC) | -22.08⋆ | -24.43 | -25.52 | -24.32 | -24.96 | -25.29 |

Computed from differenced data with 1 added back. ⋆ indicates the maximum value.

The AIC selects the $(0, d, 0)$ model by just .01 over the $(0, d, 2)$ model, however the BIC for the former is closer to zero than the BIC for the latter. Based on these criteria, we select the more parsimonious $(0, d, 0)$ model. That the best fitting model is the simplest, as opposed to requiring a higher-order is consistent with previous work finding that many political time series are well approximated as pure fractional noise (Byers, Davidson, and Peel 2000; Lebo, Walker, and Clarke 2000).[18]

**Table 2.5: Estimation of Long Memory: ARFIMA (0,$d$,0) Model**

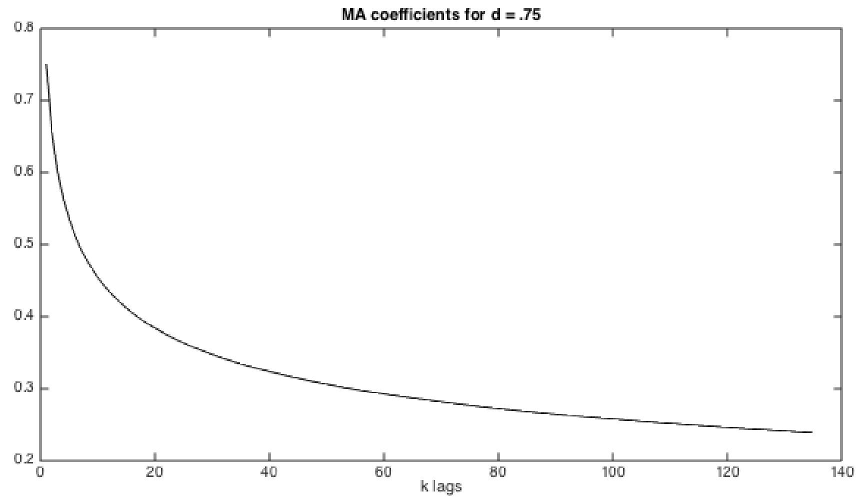
| Series | FML | EML | LW |
|---|---|---|---|
| FNDCEX | 0.748⋆<br>(0.066) | 0.692⋆<br>(0.072) | 0.76⋆<br>(0.09) |

⋆ indicates significantly different from 1.0. Standard errors in parentheses. Power of LW estimator is 0.70

The results of the three estimators are provided in Table 2.5. The FML estimates the FNDCEX series as fractionally integrated of order $I(d) = 0.75$, meaning that the series is considered non-stationary. Despite being non-stationary (or non-persistent), the process is still mean reverting over a long enough time horizon. The MA coefficients, which approximate the impulse response function of a unit shock in $\epsilon_t$, will eventually decline to 0. Figure 2.7 depicts the impulse response function of an ARFIMA $(0, d, 0)$ model with same level of persistence as estimated in the FNDCEX series. The gradual decay of the impulse response function is apparent. A one unit shock to the series at time point 0 is integrated into the process at time point 1, but even after 120 additional periods, the effects of the shock are still present in the series.

[18]For more information on pure fractional noise models versus higher order approximations as well as further evidence that many political time series are well estimated as (0,$d$,0) models, see Appendix C.8.1.

**Figure 2.7: Impulse Response Function**



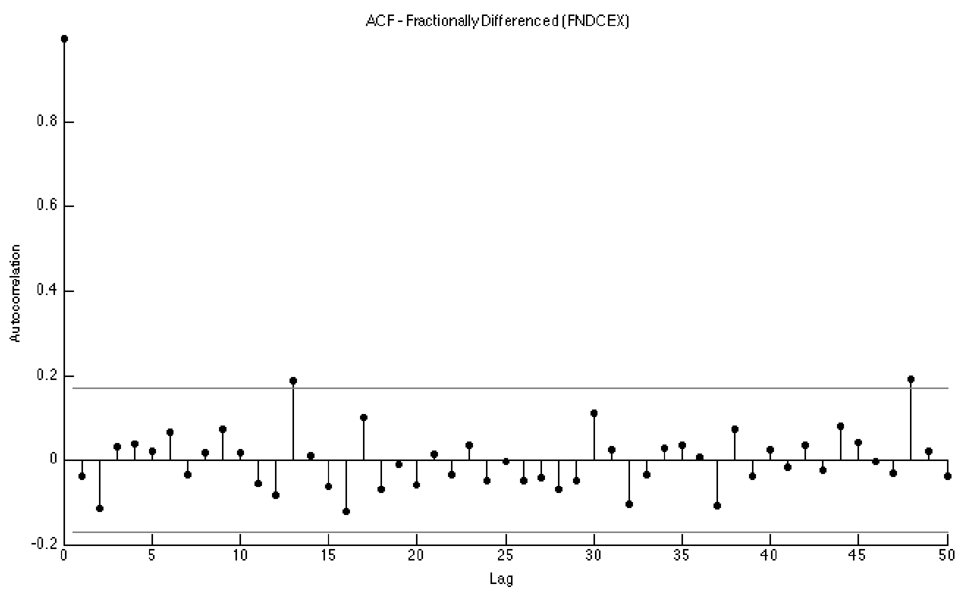

Impulse Response Function of a one unit shock to a stylized series $I(d) = .75$.

Returning to the estimates in Table 2.5, the use of additional estimators as a type of robustness check indicates that our model is well approximated - the estimates across the three models are consistent and the values are as expected. As noted in Chapter 3, the EML estimator has a negative bias when estimating a $(0, d, 0)$ model, and an estimate that is approximately .05 points lower than the FML falls into the demonstrated range of the EML in Monte Carlo simulations.

The same Monte Carlo simulations indicated that the LW estimator is unbiased when estimating pure fractional series, and this is apparent with this estimation - the LW and FML are in almost perfect agreement. The appeal of semiparametric estimators such as the LW comes from the fact that the models are agnostic as to the short-run dynamics of the process, meaning that the estimators are robust to researcher misspecification. This holds true for $(0, d, 0)$ models, however the semiparametric estimators are biased to varying degrees in the presence of short-run dynamics. In this case, the LW estimator acted as a secondary check on the AIC and BIC criteria. If higher-frequencies were present in the form of a significant and positive AR parameter, the LW estimator would most likely be biased upwards, producing an estimate of $d$ far greater than both parametric estimators. The fact that the LW estimator is in agreement with the FML supports the BIC model selection.

A second slight drawback to semiparametric estimators is that despite their ease of use and estimation, the LW estimator suffers from an inability to estimate model specific standard errors. Instead, we can only estimate an asymptotic standard error that is based on the bandwidth (number of frequencies) used in estimation. With only 135 usable observations, the LW standard errors are a little larger than the parametric estimators, but the estimate is still significant.

**Figure 2.8: ACF after Fractional Differencing**




Autocorrelation function of FNDCEX residuals after fractional differencing by $d = 0.75$

As a final check on whether our FNDCEX series is long memory, the autocorrelation plot of the residuals after fractional differencing serves as a useful tool. If the $(0, d, 0)$ model is appropriate then differencing the FNDCEX series by its order of fractional integration will be sufficient to significantly reduce the memory of the series. If fractional differencing was either inappropriate, or insufficient on its own, then we should find significant spikes in the ACF plot early and often. Figure 2.8 indicates that our initial hypothesis - the series tracking the government's federal consumption expenditures is a fractionally integrated process - is correct and that the parsimonious $(0, d, 0)$ model has adequately addressed the persistence in the series that was so evident in Figure 2.6.

## 2.6 Conclusion

This chapter began with a new theoretical justification for the expectation of finding fractional integration in political science. The error duration model of Parke (1999) offers a concise, empirically grounded theory that is easy to estimate and intellectually intuitive. The EDM process can apply to aggregate data, but unlike Granger's model, it is not a prerequisite. As a result, there is increased potential for finding and justifying the presence of fractionally integrated series.

This paper provides an example of how the EDM can be applied to political science data, but there are many other questions to which it can be applied. For instance, it has long been noted that the reelection rates of politicians follows a pattern similar to those of programs and policies, after an initial threat of electoral defeat, the incumbent grows more secure after the third term (Finocchiaro and Lin 2000). Unfortunately, the turnover rates in the House are low enough to where finding an appropriate election is difficult, however Ansolabehere and Snyder Jr (2002) demonstrated that state legislative offices also provide the same incumbent protections. Pooling statehouse elections together, one could form a larger data set allowing for the estimation of the error duration model in order to determine if the reelection of legislators is a fractionally integrated process.

To extend the use of the EDM further, one could also think of duration and survival in terms of the response to major shifts in policy. How long do individuals hold a policy position after their party has changed its position? Past work has focused on the idea of "conflict extension," the process by which strong party identifiers and party activists accept the party position and alter their own views in order to bring them into ideological alignment (Layman and Carsey 2002*a*,*b*; Layman, Carsey, Green, Herrera, and Cooperman 2010). That strong party identifiers and party activists shift their stances on issues ranging from social welfare, race relations, and abortion indicates that these positions are not permanent, integrated processes. But it is also true that these changes are not immediate. A recent example where we could expect to see this process occurred when President Obama shifted his views, and therefore the Democratic party's position, in favor of gay marriage. This process is currently ongoing with our newly open relations with Cuba, and will occur again in June assuming an agreement with Iran is reached over their nuclear program. While this type of public opinion could be considered an aggregate appropriate for Granger's theorem, applying the EDM to panel data on respondent positions allows one to estimate the diffusion process for a specific policy.[19]

This same idea of policy diffusion would likewise offer a much more detailed view of the public response to Supreme Court decision making. Ura (2014) argues that the public responds slowly to the ideological positions taken by the Supreme Court, with a response initially characterized by a short-term backlash before widespread acceptance. This is an interesting hypothesis that has ramifications for theories of Court legitimacy, however the conclusions are based on the variable *Public Mood* and the improper use of the general error correction model.[20] But questions of legitimacy and public response could be estimated with the EDM, and the data could focus much more specifically on the response to the issue decided. As an example, *Brown v. Board of Education* was handed down in 1954, but meaningful de-segregation in schools didn't occur until the mid-to-late 1960's. If one were to focus on school districts in the south, one could estimate the survival rate of segregated schools following the Court's decision. Compare this to the rate at which the states or

[19] I'm aware that this data likely does not exist. Parke's applied example used only 10 data points, so this process can be estimated with 10 waves of panel data, which I understand is a tall order.

[20] The *Public Mood* series is a bit of a black box - it is based on approximately 400 different policy questions and is assumed to capture the left-right spectrum of the policy debate in America. It is debatable whether this measure is a good approximation of the public's response to the Supreme Court decision-making, which theoretically occurs without regard for the current policy environment. Furthermore, considering that membership change on the bench plays such a large part in the ideological disposition of Supreme Court cases, and that membership change does not temporally match election results, one should expect that the Court's ideology lags behind that of mood considerably. For a discussion of the ways in which the GECM is commonly misused, see Chapter 4. The paper by Ura (2014) is not immune from the problems highlighted therein.

legislative districts in the south changed their voting rules following the Court's decision in *Shelby County v. Holder* (2013), which struck down the pre-clearance formula of the Voting Rights Act. By using the EDM, one should be able to achieve a much more nuanced view of the public's response to Supreme Court as well as how the Court's legitimacy may vary depending on the topic decided.

There are also political changes that sporadically occur that do not allow for immediate analysis, but will at some point in the future. Oregon just switched to automatic voter registration, California recently switched to a "top-two" primary, *Citizen's United* significantly altered campaign finance rules for corporations, voter ID laws have been proposed or enacted in a number of states. All of these changes represent innovations in the political world with consequences that can be tested, and the diffusion rate of those consequences can be tested via the EDM.

Moving beyond the estimation of the EDM, this paper also highlights the importance of downstream effects. If fractional series can impart persistence into other series, then the chances of encountering a FI series are increased further. In our estimation of the EDM representation, our initial series capturing the survival probabilities of federal programs was estimated as having an order of fractional integration of $I(d) = 0.68$. Similar to the idea that a shock to the price of some interconnected commodity such as oil will eventually affect prices in other sectors, it was hypothesized that series that are reliant on the original process will inherit its fractional characteristics. This is the case for the federal consumption expenditures series. Accounting for the standard error in the empirical estimates of the downstream series, the 95% confidence interval for the FNDCEX series is [0.61, 0.88]. The EDM estimation of federal program survival fits nicely within these bounds.

The concept of a fractionally integrated process fits well with current theories in a number of fields. This paper attempted to quantify the extent of path dependence that has been described theoretically in the field of American Political Development. For over a decade APD had described the development of government policies, programs, and institutions as "sticky." That term is sufficiently descriptive to provide the reader with a general understanding of the concept, but it is difficult to translate empirically. The concept of fractional integration allows us to bridge this gap. The term "sticky" implies a type of rigidity, or high viscosity, but it does not imply permanence. The current paradigm within time series - the $I(0)/I(1)$ dichotomy - cannot account for "sticky." And while political scientists have delved into the concept of near-integration, the memory process of policy development is simply too slow for an AR process to account for the rate of change. With a fractional process, we can quantify the concept of stickiness and move beyond a descriptive term to one that has empirical meaning.

## Bibliography

Ansolabehere, Stephen, and James M Snyder Jr. 2002. "The Incumbency Advantage in US Elections: An Analysis of State and Federal Offices, 1942-2000." *Election Law Journal* 1(3): 315–338.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(4): 1181–1206.

Baum, Christopher F, and Richard Sperling. 2000. "sts15. 1: Tests for Stationarity of a Time Series: Update." *Stata Technical Bulletin* 58: 35–36.

Baumgartner, Frank R., and Bryan D. Jones. 1993. *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.

Berry, Christopher R, Barry C Burden, and William G Howell. 2010. "After Enactment: The Lives and Deaths of Federal Programs." *American Journal of Political Science* 54(1): 1–17.

Binder, Sarah A. 1997. *Minority Rights, Majority Rule: Partisanship and the Development of Congress*. Cambridge: Cambridge University Press.

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

Byers, David, James Davidson, and David Peel. 1997. "Modelling Political Popularity: An Analysis of Long-range Dependence in Opinion Poll Series." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160(3): 471–490.

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

Chambers, Marcus J. 1998. "Long Memory and Aggregation in Macroeconomic Time Series." *International Economic Review* pp. 1053–1072.

Corder, J Kevin. 2004. "Are Federal Programs Immortal? Estimating the Hazard of Program Termination." *American Politics Research* 32(1): 3–25.

Davidson, James. 2014. "Time Series Modelling Version 4.44.".

Dickinson, Matthew J., and Matt J. Lebo. 2007. "Reexamining the Growth of the Institutional Presidency, 1940–2000." *Journal of Politics* 69(1): 206–219.

Doornik, Jurgen A. 2014. "An Object-Oriented Matrix Programming Language Ox 7.".

Finocchiaro, Charles J, and Tse-min Lin. 2000. *The Hazards of Incumbency: An Event History Analysis of Congressional Tenure*. Michigan State University, Institute for Public Policy and Social Research.

Freeman, John. 2012. "Research Note: Can Time Series Methods Be Used to Detect Path Dependence?" *Typescript. University of Minnesota. www. polisci. umn. edu/freeman* .

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

Jones, Bryan D. 1999. "Bounded Rationality." *Annual Review of Political Science* 2(1): 297–321.

Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *The American Political Science Review* 97(1): 151–169.

Kaufman, Herbert. 1976. *Are Government Organizations Immortal?* Brookings Inst Pr.

Kay, Adrian. 2005. "A Critique of the Use of Path Dependency in Policy Studies." *Public Administration* 83(3): 553–571.

Knight, Jack. 1992. *Institutions and Social Conflict.* Cambridge University Press.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of econometrics* 54(1): 159–178.

Layman, Geoffrey C., and Thomas M. Carsey. 2002*a*. "Party Polarization and "Conflict Extension" in the American Electorate." *American Journal of Political Science* 46(4): 786–802.

Layman, Geoffrey C, and Thomas M Carsey. 2002*b*. "Party Polarization and Party Structuring of Policy Attitudes: A Comparison of Three NES Panel Studies." *Political Behavior* 24(3): 199–236.

Layman, Geoffrey C., Thomas Carsey, John C. Green, Richard Herrera, and Rosalyn Cooperman. 2010. "Activists and Conflict Extension in American Party Politics." *American Political Science Review* 104(2): 324–346.

Lebo, Matthew J., and Daniel Cassino. 2007. "The Aggregated Consequences of Motivated Reasoning and the Dynamics of Partisan Presidential Approval." *Political Psychology* 28(6): 719–746.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

Lo, Andrew W, and Archie Craig MacKinlay. 1988. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *Review of financial studies* 1(1): 41–66.

Maltzman, Forrest, and Charles R Shipan. 2008. "Change, Continuity, and the Evolution of the Law." *American Journal of Political Science* 52(2): 252–267.

Parke, William R. 1999. "What is Fractional Integration?" *Review of Economics and Statistics* 81(4): 632–638.

Pierson, Paul. 2000. "Increasing Returns, Path Dependence, and the Study of Politics." *American political science review* 94(02): 251–267.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis.* Princeton University Press.

Robinson, Peter M. 1995. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of statistics* 23(5): 1630–1661.

Shimotsu, Katsumi. 2010. "Exact Local Whittle Estimation of Fractional Integration with Unknown Mean and Time Trend." *Econometric Theory* 26(02): 501–540.

Shimotsu, Katsumi, and Peter CB Phillips. 2006. "Local Whittle Estimation of Fractional Integration and Some of its Variants." *Journal of Econometrics* 130(2): 209–233.

Simon, Herbert A. 1985. "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review* 79: 293–304.

Ura, Joseph Daniel. 2014. "Backlash and Legitimation: Macro Political Responses to Supreme Court Decisions." *American Journal of Political Science* 58(1): 110–126.

Wildavsky, Aaron B. 1964. "Politics of the Budgetary Process.".

# Chapter 3

# Fractional Integration in Short Samples: Parametric versus Semiparametric Methods

**Abstract**

The finite sample properties of various parametric and semiparametric estimators of fractional integration are determined using Monte Carlo simulations. Previous work investigating the performance of fractional estimators relied on large sample sizes typical of economic and finance data sets. Here, the simulations are run with sample sizes representative of data sets commonly found in political science - between 40 and 100. Simulations are run on three different data generating processes $(0, d, 0)$, $(1, d, 0)$, and $(0, d, 1)$, with the AR and MA parameters and the order of fractional integration, $d$, varying with the order of fractional integration. The results indicate that semiparametric methods and the parametric frequency domain Whittle estimator are consistent across all ranges of observations with a purely fractional, $(0, d, 0)$ process, while the parametric time domain estimator exhibits a negative bias in its estimates. In the presence of a higher-order process the time domain estimator suffers dramatically and the semiparamteric estimators also exhibit bias that is potentially alleviated with proper choice of bandwidth. Throughout, the frequency domain maximum likelihood estimator outperforms other estimators, even in the presence of significant higher frequencies. The results indicate that fractional integration can be reliably estimated with lag lengths of 80 observations, but caution is still urged, particularly with the time domain estimator.

## 3.1 Introduction

Within the field of economics the ability to estimate the fractional integration (FI) of a series is an increasingly important tool for researchers. Empirical evidence of fractional integration is found in many economic series and FI methods are commonly found in macroeconomics and finance (for surveys of recent work, see Henry and Zaffaroni 2002; Gil-Alana and Hualde 2009). The same cannot be said for the use of fractional methods in political science research. After an initial foray into the concept of fractional integration and its estimation (Box-Steffensmeier and Smith 1996; Byers, Davidson, and Peel 1997; Box-Steffensmeier, Knight, and Sigelman 1998; Box-Steffensmeier and Tomlinson 2000; Byers, Davidson, and Peel 2000; Lebo, Walker, and Clarke 2000), this area of research has, for the most part, stalled. A few practitioners still rely on fractional methods, but for many the concept of fractional integration is either ignored entirely or simply deemed as a non-fruitful path for practitioners to head down.

The failure of fractional methods to gain wider acceptance stems from two primary concerns over the method. First, it is argued that political time series are too short to reliably estimate their degree of persistence. Unless and until we have more observations, fractional methods will have to wait. Second, it is argued that the estimators of fractional integration are significantly biased by the presence of short-memory dynamics. For instance, in the presence of persistent ARMA processes, semiparametric estimators are unable to differentiate between short-run autoregressive dynamics and the long-memory of a series (Baillie and Kapetanios 2009). As demonstrated below, the same problems also arise with the exact maximum-likelihood (EML) estimator (see also, Keele and Linn 2015).

Addressing the second argument first, concerns over biased estimators in the presence of persistent short-term memory are indeed valid. However, and without disregarding this concern entirely, one must ask how often we expect to find higher-order processes in political time series. For example, Byers, Davidson, and Peel (2000) examine 26 aggregated political opinion series from eight countries and finds that 22 of them are well approximated by a

(0,$d$,0) model.[1] Lebo, Walker, and Clarke (2000) find that of the four political time series they estimate to be fractionally integrated, three are best approximated by the simplest $(0, d, 0)$ model. Grant and Lebo (2015) estimate the order of fractional integration of 13 additional political time series and find that all 13 are best modeled as pure fractional processes (0,$d$,0). Whether short-memory dynamics are present is a question that must be addressed on a case by case basis, but their possible existence should not be enough to discount a method entirely.

Answering the first question of estimator reliability with smaller samples is the purpose of this paper. Many Monte Carlo studies of fractional integration estimators have been written, but none have investigated the estimators with smaller samples. The smallest sample size used in a Monte Carlo study was 64 by Robinson (1995$a$), however only one semiparametric estimator was studied in that case. The smallest number in a comprehensive study of multiple estimators was 128 (Nielsen and Frederiksen 2005).

In political science, part of the stated aim of a recent article by Keele and Linn (2015) is testing the reliability of FI estimates for series with a limited number of observations, with the authors ultimately concluding that estimates of fractional integration are too biased to be useful. There are two points worth noting about these simulations. First, they are conducted using only the time domain maximum likelihood estimator (EML) of Sowell (1992$a$), which as will be discussed, has a prominent and well-known negative bias in its estimates. Second, the results of this one estimator are then extrapolated to all other FI estimators in order to discount the possible existence of any consistent and unbiased fractional estimators. In reference to the bias of the EML estimator, even with greater than 100 observations, Keele and Linn (2015, p. 20) argue against the meaningful use of fractional methods stating, "[i]t is clearly inconsistent with the statement in [Grant and Lebo (2015)] that one needs at least 64 observations to reliably estimate the $d$ parameter." The interpretation of Grant and

[1]For those models in which a higher-order model was deemed appropriate, the authors note the difference in likelihood was minimal with a sizable increase the size of the standard errors. Byers, Davidson, and Peel (2000) were left to question whether the trade-off of efficiency for bias was worth it.

Lebo's claim is incorrect, and the conclusions Keele and Linn draw from it in relation to their own simulations are also wrong. Grant and Lebo (2015) reference the work of Robinson (1995*a*), which evaluated the semiparametric local Whittle estimator, not that of the EML. And as demonstrated below, Robinson was correct in finding the estimator unbiased. In fact, Robinson was quite cautious in his review, under the same circumstances tested by Robinson, the local Whittle estimator is unbiased with as few as 40 observations.

This paper considers several popular FI estimation techniques, two semiparametric and two parametric methods. These methods are compared via Monte Carlo simulations using multiple data generating processes with varying degrees of short-run dynamics and the performance of each is measured by estimator bias and root mean squared error (RMSE). Sample sizes were chosen to represent the size of data sets common to political science, ranging between 40 and 100 observations.

The results indicate that there is reason to be bullish as to the capability to reliably estimate fractional integration in shorter series. In the absence of short-run dynamics, both semiparametric estimators are unbiased with the local Whittle (LW) estimator of Robinson (1995*a*) outperforming the log periodogram regression (LPR) model of Geweke and Porter-Hudak (1983) in terms of RMSE. Using parametric methods under these circumstances, the frequency domain maximum likelihood (FML) estimator dominates the time domain exact maximum likelihood (EML) estimator of Sowell (1992*a*). In these smaller sample simulations the EML exhibits a negative bias that is prominent even when the series are simulated as pure fractional processes.

In the presence of short-run dynamics, particularly positive AR noise, the semiparametric estimators exhibit a positive bias that is alleviated somewhat by reducing the bandwidth used, however when the dynamics are especially persistent (for example, AR(1) with the autoregressive parameter equal to 0.80) estimates are severely biased upwards. When using parametric methods, the time-domain EML suffers from bias and a bloated RMSE in almost all circumstances. An increase in observations alleviates these problems somewhat, however not to the extent that EML can be recommended on its own. On the other hand, the frequency domain FML estimator does not suffer from the same problems to the same extent. With the exception of a few specific instances, the FML estimator performs admirably in the presence of short-run dynamics, generally returning unbiased estimates, particularly with series of at least 80 observations.

Another important concern is whether we can accurately and consistently estimate the short run dynamics when present. An inability to estimate the AR or MA parameters while also accounting for fractional integration severely limits the potential benefits and jeopardizes all inferences. This problem is well known with semiparametric estimators, and the bias in two step estimation has been demonstrated previously (Baillie and Kapetanios 2009). The results find that the EML performs terribly when estimating $(1, d, 0)$ series, drastically overestimating the extent of autoregression. The estimator performs better when confronted with moving average parameters and anti-persistent series, however bias increases with the level of memory of the process. In contrast, the FML (again, with limited exceptions) accurately estimates both the AR and MA parameters while simultaneously estimating the order of fractional integration.

In general, the conclusion of this paper stands in stark contrast to that of Keele and Linn (2015), and this disparity is due to their singular reliance on the EML estimator. Their estimator choice is particularly troubling for several reasons. First, the negative bias in the EML has been well documented for over 20 years (see, e.g., Li and McLeod 1986; Cheung and Diebold 1994; Hauser 1999).[2] Considering the very public recognition of bias in the EML, as well as the presence and use of other fractional estimators, any generalizations made as to the robustness of fractional methods based only on the performance of this one estimator are short-sighted. More problematic is the reason why Keele and Linn (2015) chose to use the EML in their simulations - it is the prepackaged fractional estimator in Stata, the most widely used software program in political science.

[2] Keele and Linn (2015) even note that the bias found in the EML is consistent with Hauser (1999).

Stata is presenting researchers with an estimator that is known to be biased and inefficient, meaning that practitioners using the program have no way of reliably estimating fractional integration. The result is a general assumption that FI methods are inappropriate or can't be reliably estimated. As a consequence, practitioners are pushed back into the $I(0)/I(1)$ dichotomy and proceed in general ignorance to the possibility that their series are fractionally integrated. As has been demonstrated, this ignorance can harm inferences and estimation results - assumptions as to the order of integration of series can lead to drastically different estimates, rates of spurious regressions can increase if FI is not accounted for, and short-run dynamics can be built into series through over differencing (Box-Steffensmeier and Smith 1998; Byers, Davidson, and Peel 2000; Lebo, Walker, and Clarke 2000; Tsay and Chung 2000). The inability of researchers to reliably and effectively estimate a series' order of fractional integration seriously precludes further development of fractional methods in political science, and potentially jeopardizes inferences as well. But as demonstrated by the findings herein, there are more reliable options that can be estimated just as easily as the EML. There is no reason why the EML should be the go-to estimator of fractional integration.

The rest of the paper proceeds as follows. The next section briefly discusses the concept of fractional integration and the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model. The estimators and their underlying assumptions are then described, followed by the results of the Monte Carlo simulations with finite sample bias and RMSE for each estimator. A brief discussion of the results follows. An appendix provides additional information on the frequency domain, estimates of AR and MA parameters when fractional integration is present and estimated, as well as Monte Carlo results indicating the cost of ignoring the presence of fractional integration in favor of a more simplistic ARMA representation. If a series is fractionally integrated and ignored, the ARMA estimates are biased, with the severity of the bias determined by the strength of the unaccounted for fractional memory.

## 3.2 What is Fractional Integration?

Assume a process that generates a series integrated of order $d$, in which $d$ is the number of differences that must be applied to the series for it to be stationary. In the standard case, in which $d$ is an integer, we have the $I(1)/I(0)$ distinction: if a series is an $I(1)$ unit root, it has perfect memory, meaning that a shock is fully integrated into the series. On the other hand, an $I(0)$ series has only short memory; it is weakly stationary with constant mean, finite variance, and constant covariance (see Enders 2004, p.54).

But if the assumption that $d$ must be an integer is relaxed and it is instead allowed to fall anywhere on the real number line, then the level of integration is considered fractional. The ARFIMA process, introduced by Granger and Joyeux (1980) and Hosking (1981) is labeled $(p, d, q)$ if its $d$'th difference yields a stationary and invertible ARMA $(p, q)$ process.

Assuming that $-1/2 < d < 1/2$ in order to ensure stationarity and invertibility, the properties of $y_t$ can be defined as an ARFIMA $(p, d, q)$ if

$$\phi(L)(1-L)^d(y_t-\mu)=\theta(L)\epsilon_t, \tag{3.1}$$

where $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + ... + \theta_q z^q$ are lag polynomials in the lag operator $L(Lx_t = x_{t-1})$ with roots strictly outside the unit circle, $\epsilon_t$ is $iid$ $N(0, \sigma^2)$ and $(1-L)^d$ is defined by its binomial expansion such that

$$(1-L)^d=\sum_{j=0}^{\infty}\psi_j(d)L^j,\quad \psi_j(d)=\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)},\ j=0,1,..., \tag{3.2}$$

where $\Gamma(z)=\int_0^\infty x^{z-1}e^{-x}dx$ is the gamma function.

The estimated $d$ parameter determines the memory of the series in question. If $-1/2 < d < 1/2$ the process is weakly stationary and invertible. If $d < 0$, the process is considered anti-persistent, in which case the autocorrelations are mostly negative and summable (sum to a finite constant). For $0 < d < 1/2$ the process is considered stationary but possesses "long memory;" it is characterized by a hyperbolically declining autocorrelation function, and the autocorrelations and autocovariances are nonsummable. Finally, for $1/2 \leq d < 1$, the series is no longer covariance stationary and is characterized as non-persistent. Even though the series is non-stationary, the MA coefficients, which approximate the impulse response function of a unit shock in $\epsilon_t$, will eventually decline to 0.

### 3.2.1 Equivalence of Time and Frequency Domain

All time series processes can be analyzed in either the time domain or in the frequency domain. Time domain analysis measures the variable and its evolution against time while the frequency domain approach is less concerned with how the series evolves over time and is instead concerned with the variance of the time series at certain frequencies. From the time domain, a series can be transformed via the Discrete Fourier Transform into the frequency domain. This transformation decomposes the original series into an infinite sum of periodic functions, each with a different frequency, $\lambda$, which ranges between 0 and $\pi$. In the frequency domain, it is the spectral density that is of primary concern, and it is the summation of this density across all possible frequencies that represent the total variance of the process.

Of course, a process in the frequency domain can be transformed into the time domain using the Inverse Discrete Fourier Transform, and any representation in the time domain has a mathematically equivalent representation in the frequency domain (Box and Pierce 1970). Despite the equivalency, the frequency domain does have several characteristics that lend itself to analysis of fractional time series. First, the visualization of the spectral density, through the sample periodogram, allows the researcher a visual inspection of the process and the frequencies which are the greatest overall contributors to its variance. As Granger (1966) demonstrated, most macroeconomic series tend to exhibit a typical spectral shape characterized by a concentration of the highest powers in the low frequencies.[3] This concentration indicates a persistence in the memory of a series that is consistent even after a series has

[3] Appendix B.1 provides figures and descriptions of the spectral density and how the frequency domain can serve as an excellent diagnostic tool if the researcher believes a series may be fractionally integrated.

been detrended, and remains the same regardless of the rate at which the series is sampled, both of which are characteristics now recognized as necessary prerequisites for the presence of fractional integration. Thus, a finding that the variance of the process is driven by the low frequencies is a strong indication that the series in question is fractionally integrated. Second, the concentration of power at the low frequencies is due to the the dramatic increase in the power of the spectral density at the zero frequency, which as will be discussed below, can dramatically bias estimates in the negative direction if left unaccounted for in estimation. Because estimators in the frequency domain explicitly exclude the zero frequency, this potential for bias is alleviated, which produces much more efficient and consistent estimates of long memory persistence. As a result, the frequency domain provides researchers a relatively easy and consistent method of estimating fractional integration. The equivalancy between the two domains is demonstrated below.

The autocorrelation function of a process in the time domain, such as in Equation (3.1), can be approximated as

$$\rho_k \sim c_\rho k^{2d-1},\ 0 < c_\rho < \infty, \quad \text{as}\, k \to \infty \tag{3.3}$$

in which $c_p$ is a constant and $\rho_k$ decays at a hyperbolic rate.[4]

The autocorrelations can also be described in the frequency domain. To do so, we first define the spectral density function of $y_t$, $f_y(\lambda)$, as

$$\gamma_k = \int_{-\pi}^{\pi} f_y(\lambda) e^{i\lambda k} d\lambda \tag{3.4}$$

in which $\gamma_k$ is the $k$'th autocovariance of $y_t$ with $\lambda$ representing the known frequency spectrum $[0,\ \pi]$.

[4] The process of obtaining the autocorrelation function from the autocovariances is outlined more specifically in Appendix A.1.1.

The spectral density of the ARFIMA $(p, d, q)$ process of Equation 3.1 can be given by

$$f_y(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-2d} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \tag{3.5}$$

$$= \frac{\sigma^2}{2\pi}(2\ \sin\lambda/2)^{-2d} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \tag{3.6}$$

in which the ARMA $(p, q)$ parameters are represented as

$$f_u(\lambda) = \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \frac{\sigma^2}{2\pi} \tag{3.7}$$

and the transfer function of the fractional filter $(1 - L)^d$ is approximated as

$$|1 - e^{i\lambda}|^{-2d} = (2\ \sin\lambda/2)^{-2d} \sim |\lambda|^{-2d} \ \text{ as } \lambda \to 0, \tag{3.8}$$

Thus, if we assume that $f_u(\lambda)$ is a constant as the frequencies, $\lambda \to 0$, then $f_u(\lambda) \to g$. Which means that the approximation of the autocorrelation function in Equation 3.3 can be restated in the frequency domain as the approximation of the spectral density function (Granger and Joyeux 1980; Beran 1994)

$$f_y(\lambda) \sim g|\lambda|^{-2d},\ 0 < g < \infty, \ \text{ as } \lambda \to 0. \tag{3.9}$$

The next sections describes the estimators used in the simulations. The parametric ARFIMA $(p, d, q)$ estimators are the EML - which is estimated in the time domain - and the FML which operates in the frequency domain. Our semiparametric estimators - the LPR and the LW - are both in the frequency domain.

## 3.3 Parametric Estimators

We use two different generic maximum likelihood procedures for our estimations: the frequency domain (Whittle) maximum likelihood and the exact Gaussian maximum likelihood procedures. Both methods have the same asymptotic properties, i.e., they are each $\sqrt{T}$-consistent and asymptotically normal.

### 3.3.1 Time Domain - Exact Maximum Likelihood (EML)

Assuming a process that is stationary and invertible, i.e., $-1/2 < d < 1/2$, the exact Gaussian log-likelihood is given by the objective function

$$\log\, L_E(d, \phi, \theta, \sigma^2, \mu) = -\frac{1}{2}[T\log(2\pi) + \log\, \det(\Sigma) + (Y - \mu l)'\Sigma^{-1}(Y - \mu l) \tag{3.10}$$

where $\Sigma$ is the covariance matrix of $Y$, $Y = (y_1, \ldots, y_n)$' and $l = (1, \ldots, 1)'$, $\phi$ and $\theta$ are the parameters $\phi(L)$ and $\theta(L)$, and $\mu$ is the mean of $Y$.

The EML investigates the reduced profile likelihood where $\mu$ is replaced with the sample mean $y$ and the solution of the maximization with respect to $\sigma^2$ is obtained. Gathering the parameters in the vector $\gamma = (d, \phi', \theta', \sigma^2, \mu)'$, the concentrated likelihood is then given as

$$\log\, L_E^*(\gamma) = -\frac{1}{2}[T\log(2\pi) + T\log(\sigma_u^{2*}) + \log\, \det(\overline{\Sigma}) + T] \tag{3.11}$$

with $\sigma^2 = \sigma_u^{2*}$ and $\sigma_u^{2*} = \frac{1}{T}(Y - \mu l)'\Sigma^{-1}(Y - \mu l)$ (Hauser 1999).

Sowell (1992*a*) demonstrated that the EML estimator of $d$ is asymptotically equivalent to the FML estimator, they are each asymptotically normal and $\sqrt{T}$-consistent. When using the estimator with finite samples in which the mean is known, the EML has better finite sample properties than the FML (Sowell 1992*b*). However, in most instances the mean is unknown and must be estimated instead. Under circumstances when the mean must first be estimated, the efficiency gains of the EML are lost (Cheung and Diebold 1994). Additionally, when the mean is unknown, the EML is known to bias the $d$ parameter in the negative direction (Li and McLeod 1986). This inherent negative bias inherent in estimation was demonstrated most recently within political science by Keele and Linn (2015) and is also evident in the results presented herein.

### 3.3.2 Frequency Domain - Whittle Approximate Maximum Likelihood (FML)

We can also approximate the ARFIMA$(p, d, q,)$ model in the frequency domain. Assuming Gaussian errors, the Whittle likelihood function is

$$L_F(d, \phi, \theta, \sigma^2) = T\left[\log(2\pi) - 1 - \log\frac{1}{T}\sum_{j=1}^{[T/2]}\frac{I(\lambda_j)}{f_y(\lambda_j)}\right] - \sum_{j=1}^{[T/2]}\log\, f_y(\lambda_j) \tag{3.12}$$

where $\lambda_j = 2\pi j/T$ are the Fourier frequencies, $I(\lambda_j) = \frac{1}{2\pi T}|\Sigma_{t=1}^{T} y_t e^{it\lambda}|^2$ is the periodogram of $y_t$, and $f_y(\lambda)$ is the spectral density of Equation (3.5).

The FML estimator has the same asymptotic properties as the EML, but when estimated with finite samples its properties are far superior to that of the EML. This is due to the way the two estimators treat the mean of the process, the zero frequency in the spectral representation, $j = 0$. The FML is invariant to the presence of a non-zero mean since $j = 0$ is explicitly excluded from the summation. On the other hand, the mean is incorporated into the EML estimation through the autocovariance function. In a zero-mean process, the periodogram is zero at frequency zero, but when $d > 0$ the spectral density approaches infinity as the frequency approaches zero. The $d$ parameter is essentially describing the slope of the spectrum as it nears the zero frequency, and the EML estimate of $d$ is trying to model this upward slope while still having to account for the null value of the mean at frequency zero. This causes the EML to suffer from a negative bias, which is made worse by the presence of additional ARMA parameters in ARFIMA$(1, d, 0)$ and $(0, d, 1)$ models (Cheung and Diebold 1994; Hauser 1999).

## 3.4 Semiparametric Estimators

Frequency domain semiparametric estimators use the approximation in Equation (3.3) as a semiparametric model for the spectral density at low frequencies. The low frequencies are characterized as the first $m$ Fourier frequencies, $\lambda_j = 2\pi j/T, j = 1, \ldots, m$, where

$$\frac{1}{m} + \frac{m}{T} \to 0 \text{ as } T \to \infty \tag{3.13}$$

such that $m$ increases with the sample size $T$, but at a slower rate. Only the spectrum around $\lambda = 0$ is modeled in terms of the long memory parameter $d$, so these estimators are referred to as local, or narrow band estimates.

In comparison to the parametric estimators, semiparametric estimators offer an easily estimated model that is generally robust to short-run dynamics because the presence of higher frequencies rarely contaminate the ordinates in the vicinity of the origin. These estimators ignore the presence of ARMA polynomials, instead focusing exclusively on the long memory parameter. While simpler to estimate, semiparametric estimators are less efficient than parametric methods, achieving only $\sqrt{m}$ consistency where $m = m(T)$ is the number of Fourier frequencies to be used. Compare this to the $\sqrt{T}$ consistency of both of our parametric estimators.

When specifying the model, the selection of $m$ (also known as the bandwidth) is important. In making this decision, the user must balance an interest in reducing the variance in the estimates with an increase in bias. Hurvich, Deo, and Brodsky (1998) and Henry and Robinson (1996) investigated the optimal bandwidth choice in terms of reduction of mean square error for the LPR and LW estimators respectively, and found that the ideal is $m = T^{0.8}$ for each. While this is the optimal, such a large bandwidth is rarely used because of the potential for contamination of short-run dynamics. The further out we move from the origin, the greater the chance our estimates may be biased by unmodeled ARMA dynamics. Semiparametric estimators attempt to measure only the low frequencies and the probability of capturing the high frequencies of short-run dynamics increases with the bandwidth. Consequently, when estimating the long memory of a series with no a priori information, it is best to begin with a conservative bandwith, $m = T^{0.5}$ or $T^{0.65}$ before increasing.

### 3.4.1 Log Periodogram Regression

The log periodogram regression (LPR) was introduced by Geweke and Porter-Hudak (1983) who demonstrated that by taking logs of both sides of Equation (3.3), we obtain

$$\log f_y(\lambda) = \log(g) - 2d \log(\lambda_j), \quad j = 1, \dots, m \tag{3.14}$$

If we substitute the periodogram of $y_t$, $I(\lambda_j)$ for $f_y(\lambda)$, then we get the linear regression model on the log-periodogram

$$\log I(\lambda_j) = a + dz_j + u_j, \quad j = 1, \dots, m \tag{3.15}$$

where $a$ is a constant, $z_j = 2 \log(\lambda_j)$ and $u_j$ is the error term. The LPR estimator is the OLS regression estimate of $d$, where $m$ is the bandwidth. Again, because we are working within the frequency domain, note that the estimator is invariant to a non-zero mean. Robinson (1995$b$) proved that the LPR has an asymptotically normal distribution for $d \in (-1/2,\ 1/2)$ and subsequent research has extended this such that the range of asymptotic normality is $d \in (-1/2, 3/4)$ and the estimator is consistent for all $d \in (-1/2, 1]$ (Velasco 1999$b$).

### 3.4.2 Local Whittle Estimator

The other semiparametric estimator used in our simulations is the local Whittle (LW) approach, which was first proposed by Kuensch (1987), and Robinson (1995$a$) showed the estimator's consistency and asymptotic normality for $d \in (-1/2,\ 1/2)$. Velasco (1999$a$) extended the results of Robinson to show that the LW estimator is consistent for $d \in (-1/2,\ 1)$ and asymptotically normally distributed for $d \in (-1/2,\ 3/4)$. Thus, both the LW and the LPR share the same asymptotic properties. The LW has been proven to have a smaller asymptotic variance than the LPR however, meaning that it is a more efficient estimator than the LPR. As a result, the LW estimator is expected to outperform the LPR.

The LW estimates $d$ by minimizing the objective function, defined in terms of the parameters $d$ and $G$, where $m$ is the chosen bandwidth

$$Q_m(d,G) = -\frac{1}{m}\sum_{j=1}^{m}\left[\log G\lambda^{-2d} + \frac{I(\lambda_j)}{G\lambda_j^{-2d}}\right] \tag{3.16}$$

where $\lambda_j = 2\pi j/T$ are the Fourier frequencies and $I(\lambda_j) = \frac{1}{2\pi T}|\Sigma_{t=1}^{T} y_t e^{it\lambda}|^2$ is the periodogram. The LW estimates $G$ and $d$ by minimizing $Q_m(G,d)$, such that

$$(\widehat{G}, \widehat{d}) = \underset{G\in(0,\infty), d\in[\nabla_1,\nabla_2]}{\arg\min} Q_m(G,d) \tag{3.17}$$

where $\nabla_1$ and $\nabla_2$ are numbers such that $-\infty < \nabla_1 < \nabla_2 < 1/2$. Concentrating Equation 3.16 with respect to $G$, $\widehat{d}$ satisfies

$$\widehat{d} = \underset{d\in[\nabla_1,\nabla_2]}{\arg\min} R(d), \tag{3.18}$$

where

$$R(d) = \log \widehat{G}(d) - 2d\,\frac{1}{m}\sum_{1}^{m} \log\lambda_j, \quad \widehat{G}(d) = \frac{1}{m}\sum_{1}^{m} \lambda_j^{2d} I(\lambda_j). \tag{3.19}$$

It should be noted that despite their relative ease of estimation, the LW and LPR estimators are constrained by the bounds of $d \in (-1/2,\ 3/4)$ and a researcher should be cognizant of these limits. Outside of these bounds, when $I(d) > 3/4$ the estimators take on non-normal limit distributions and achieve slower rates of convergence. When $d > 1$ the LW estimator is no longer consistent and has been shown to converge to unity rather than the true value of $d$ (Phillips and Shimotsu 2004). This is why common practice when estimating the order of fractional integration of a series is to first difference the data and then add one back. This can be problematic however if the series in question is actually trend stationary, $d \in (0,\ 1/2)$, such that differencing puts the series into the range of $d \in [-1,\ -1/2)$. When this occurs the estimator's behavior is again inconsistent and the estimates will converge, either to the true value or to zero depending on the number of frequencies chosen. When $m$ increases slowly, the estimate will converge to 0, however if $m$ increases quickly, the estimate will converge to the true value of $d$ (Shimotsu and Phillips 2006).

## 3.5 Monte Carlo Models

The Monte Carlo simulations are based on 1,000 simulations per data generating process (DGP) per time series of varying observations - 40 through 100 increasing by increments of 20. All simulations were generated in Matlab 2014a using the 'fracdiff' program of Shimotsu (2003) to impart fractional memory, and Matlab's own 'filter' command to add short-run dynamics in the form either AR or MA polynomials. The sample sizes are deliberately small, chosen to simulate the number of observations of series commonly found in political science.[5]

The simulations consider three types of DGP. The first is the fractional noise ARFIMA model

$$(1-L)^d(y_t-\mu)=\epsilon_t,\ \ \epsilon_t\sim N(0,\ 1) \tag{3.20}$$

The main parameter of interest is $d$, and series are generated for the values of $d$ $\{0, .25, .45, 0.75\}$. In this instance, $d=0$ can be considered a series that fits within the $I(0)/I(1)$ paradigm and is therefore not fractionally integrated. The value $d=.75$ is representative of a series that is fractionally integrated and non-persistent, but that has been first differenced such that we are estimating a series that is $d=-.25$ before adding 1 back. Because it is so common to first difference the series before estimation, the ability to reliably estimate the $d, \phi$, and $\theta$ parameters in this range is particularly important.

Also considered are ARFIMA$(1, d, 0)$ and ARFIMA$(0, d, 1)$ models. These models are represented as

$$(1-\phi L)(1-L)^d(y_t-\mu)=\epsilon_t,\quad \epsilon_t\sim N(0,\ 1) \tag{3.21}$$

$$(1-L)^d(y_t-\mu)=(1+\theta L)\epsilon_t,\quad \epsilon_t\sim N(0,\ 1) \tag{3.22}$$

[5]For example, the following are sample sizes from publications using time series data in the top three journals over the past several years: Keele (2007): 100; Jennings and John (2009): 41; Kelly and Enns (2010): 54; Layman, Carsey, Green, Herrera, and Cooperman (2010): 8; Casillas, Enns, and Wohlfarth (2011): 45; Faricy (2011): 37; Sanchez Urribarri, Schorpp, Randazzo, and Songer (2011): 35; Ura and Ellis (2012): 35; Enns (2014): 57; Enns, Kelly, Morgan, Volscho, and Witko (2014): 67; Ura (2014): 54

and in each DGP we assume that $\mu = 0$ and $\sigma^2 = 1$.

The values of $\phi$ and $\theta$ vary individually by increments of 0.4 such that the values take on one of $\{-0.40, 0, 0.40, 0.80\}$ where $\phi = \theta = 0$ represents the case in which the model is misspecified - no ARMA polynomials are present in the data, but the models have been overfit in an attempt to estimate short-run dynamics. Overfitting potentially introduces bias into the estimates, a serious concern not only for the estimates of $d$, but for the estimates of $\phi$ and $\theta$ as well.[6] Each model is estimated naively, meaning that no initial values are being input into the model to help our estimations.

## 3.6 Monte Carlo Results

The results for our Monte Carlo simulations are presented below.[7] We begin with the results of the ARFIMA$(0, d, 0)$ models; first the parametric results in Table 3.1, followed by the semiparametric estimates in Table 3.2. Because so much information is presented in each table the low values of bias and RMSE are bolded for each order of fractional integration ($d$) and number of observations ($T$).

### 3.6.1 Results - (0,$d$,0) Models

The first set of results to address are the parametric estimates of the $(0, d, 0)$ DGP. Past work in political science has found that a predominant number of fractionally integrated series are well estimated as $(0, d, 0)$ models, therefore the ability to estimate these series consistently is of the utmost importance.

The results are found in Table 3.1 and as expected the far left hand column demonstrates the negative bias of the EML estimator. Despite the fact that all of the DGPs are assumed to have a zero mean, the EML still accounts for the mean of the process through its autocovariance function, which introduces a negative bias in its estimates. The negative

[6] Note that one of the benefits of semiparametric estimators is that they are estimated without accounting for the short-run dynamics. Because the ARMA polynomials need not be specified, misspecification is not a concern. As a result, the $\phi = 0$ and $\theta = 0$ models are not presented for the semiparametric estimators.

[7] Results for the parametric estimates of the $\phi$ and $\theta$ parameters for the $(1, d, 0)$ and $(0, d, 1)$ models are presented in Appendix B.2.

bias increases as the level of persistence within the series increases, and while the bias is minimized somewhat with the inclusion of additional observations, it is a serious concern. That this negative bias is present even in the absence of short-run dynamics should give a researcher pause when considering using the EML, and as seen from the later simulations, the bias in the EML becomes more prominent with the increased complexity of the DGP.

The FML displays no such bias however. Again, this is due to the fact that all estimates in the frequency domain explicitly exclude the zero frequency. The result is that the FML estimator provides estimates of long memory that are unbiased and it does so with a low RMSE. The FML estimates of long memory are as efficient as those of a simple ARMA process.

**Table 3.1: Parametric Estimators - ARFIMA (0,*d*,0)**

| | | EML | | FML | |
|---|---|---|---|---|---|
| $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** |
| 0 | 40 | −.1128 | .2093 | **−.0202** | **.1838** |
| | 60 | −.0733 | .1496 | **−.0198** | **.1352** |
| | 80 | −.0618 | .1258 | **−.0111** | **.1133** |
| | 100 | −.0432 | .1028 | **−.0094** | **.0995** |
| 0.25 | 40 | −.1274 | .1990 | **−.0279** | **.1804** |
| | 60 | −.0821 | .1439 | **−.0145** | **.1331** |
| | 80 | −.0670 | .1223 | **−.0139** | **.1108** |
| | 100 | −.0465 | .1003 | **−.0056** | **.0962** |
| 0.45 | 40 | −.1613 | .2077 | **−.0552** | **.1462** |
| | 60 | −.1146 | .1502 | **−.0345** | **.1048** |
| | 80 | −.0931 | .1249 | **−.0198** | **.0840** |
| | 100 | −.0731 | .1022 | **−.0149** | **.0732** |
| 0.75 | 40 | −.1124 | .2110 | **−.0034** | **.1597** |
| | 60 | −.0689 | .1532 | **−.0052** | **.1228** |
| | 80 | −.0554 | .1231 | **−.0094** | **.1097** |
| | 100 | −.0406 | .1039 | **−.0080** | **.0960** |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each specified level of $d$ and $T$. $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

Table 3.2 provides the results of Monte Carlo simulations for the semiparametric estimators. As a reminder, the semiparametric estimators will vary depending on the bandwidth (number of Fourier frequencies) used in each model. For each $(0, d, 0)$ model four different bandwidths are utilized - $\{0.50, 0.65, 0.70, 0.75\}$. The higher the bandwidth the smaller the RMSE, however this reduction comes at a general cost of an increase in bias. Furthermore, increased bandwidth potentially introduces bias into the estimates if short-run dynamics are present in the underlying process.

In this instance, in which the estimates are of pure fractional series, induced bias isn't a particularly concerning problem and this is evident in the results for each estimator. The LPR estimates are found in the upper rows and the LW in the lower rows, and for each the estimates are consistently unbiased as the bandwidth increases. As expected, the increase in bandwidth reduces the overall RMSE of the estimator, with the lowest RMSE found when the highest number of Fourier frequencies are utilized in the estimation. Also as expected considering its lower asymptotic variance, the LW estimator outperforms the LPR estimator in terms of RMSE while the differences in bias amongst estimates are negligible.

Based on these results, it is clear that ARFIMA$(0, d, 0)$ models can be consistently and reliably estimated with a variety of estimators. Caution is needed when using the EML estimator however due to its negative bias. This bias is mitigated somewhat with additional observations, but it is still an issue. The fact that the bias is so prominent, even when estimating the simplest model, is especially concerning considering that the stock FI estimator in Stata, the most well known and widely used statistics program in political science (outside of R, of course) is the EML.

### 3.6.2 Results - (1,$d$,0) Models

The estimates for the parametric estimates of the $(1, d, 0)$ models are found in Table 3.3, and their results again indicate the potential bias problems for the EML estimator as well as the fact that the FML estimator is not perfect under all circumstances. Addressing the EML estimator first, it suffers from a persistent and significant negative bias in the presence of higher AR frequencies. This is expected given the treatment of the mean, however what is interesting is that when the autoregression becomes highly persistent, i.e., $\phi \geq 0.80$, the EML actually performs well to the point that it is consistent and more efficient than the FML. The performance of the EML in this case is actually helpful to the researcher by providing a larger coverage area than would otherwise be available due to the bias suffered

## Table 3.2: Semiparametric Estimators - ARFIMA (0,$d$,0)

| | | LPR $[m = T^{0.50}]$ | | LPR $[m = T^{0.65}]$ | | LPR $[m = T^{0.70}]$ | | LPR $[m = T^{0.75}]$ | |
|---|---|---|---|---|---|---|---|---|---|
| $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| 0 | 40 | .0321 | .4213 | .0132 | .2873 | .0136 | .2426 | .0109 | **.2222** |
| | 60 | −.0071 | .3768 | .0017 | .2394 | **−.0002** | .2101 | .0054 | **.1867** |
| | 80 | −.0069 | .3601 | −.0013 | .2129 | −.0015 | .1832 | **−.0011** | **.1609** |
| | 100 | −.0072 | .3055 | −.0065 | .1991 | **−.0002** | .1666 | .0011 | **.1498** |
| 0.25 | 40 | .0066 | .4380 | .0120 | .3063 | **.0043** | .2548 | .0067 | **.2398** |
| | 60 | **−.0001** | .3847 | .0013 | .2369 | .0028 | .2127 | .0019 | **.1856** |
| | 80 | −.0084 | .3593 | .0066 | .2146 | .0082 | .1869 | **.0037** | **.1651** |
| | 100 | **.0023** | .2894 | .0202 | .1916 | .0155 | .1646 | .0140 | **.1418** |
| 0.45 | 40 | .0188 | .4402 | .0211 | .3113 | **.0144** | .2612 | .0157 | **.2407** |
| | 60 | .0268 | .4042 | **−.0157** | .2453 | .0169 | .2134 | .0165 | **.1914** |
| | 80 | **.0058** | .3443 | .0073 | .2095 | .0096 | .0346 | .0128 | **.1651** |
| | 100 | **.0058** | .2934 | .0117 | .1891 | .0157 | .1636 | .0174 | **.1442** |
| 0.75 | 40 | −.0101 | .4413 | −.0131 | .3135 | −.0054 | .2546 | **−.0035** | **.2320** |
| | 60 | **−.0083** | .3977 | −.0111 | .2426 | −.0094 | .2101 | −.0127 | **.1915** |
| | 80 | **−.0035** | .3505 | .0055 | .1998 | .0048 | .1798 | .0089 | **.1609** |
| | 100 | .0232 | .2849 | .0159 | .1908 | .0136 | .1589 | **.0118** | **.1441** |

| | | LW $[m = T^{0.50}]$ | | LW $[m = T^{0.65}]$ | | LW $[m = T^{0.70}]$ | | LW $[m = T^{0.75}]$ | |
|---|---|---|---|---|---|---|---|---|---|
| $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| 0 | 40 | .0167 | .3309 | −.0135 | .2278 | −.0099 | .1911 | **−.0094** | **.1707** |
| | 60 | −.0314 | .2937 | −.0232 | .1941 | −.0210 | .1654 | **−.0136** | **.1416** |
| | 80 | −.0217 | .2881 | −.0141 | .1624 | −.0145 | .1412 | **−.0117** | **.1249** |
| | 100 | −.0332 | .2453 | −.0205 | .1579 | −.0114 | .1317 | **−.0078** | **.1155** |
| 0.25 | 40 | −.0356 | .3717 | **−.0243** | .2512 | −.0349 | .2044 | −.0359 | **.1858** |
| | 60 | −.0455 | .3188 | −.0337 | .1934 | **−.0292** | .1705 | −.0320 | **.1475** |
| | 80 | −.0391 | .3027 | −.0172 | .1746 | **−.0153** | .1473 | −.0202 | **.1262** |
| | 100 | −.0269 | .2499 | **−.0072** | .1552 | −.0115 | .1277 | −.0165 | **.1104** |
| 0.45 | 40 | −.0408 | .3622 | **−.0281** | .2580 | −.0388 | .2114 | −.0441 | **.1911** |
| | 60 | **−.0228** | .3412 | −.0232 | .2058 | −.0231 | .1759 | −.0315 | **.1521** |
| | 80 | −.0391 | .3027 | **−.0205** | .1716 | −.0209 | .1515 | −.0238 | **.1308** |
| | 100 | −.0268 | .2518 | −.0199 | .1591 | **−.0186** | .1318 | −.0219 | **.1147** |
| 0.75 | 40 | .0204 | .2867 | −.0042 | .2039 | **−.0003** | .1738 | .0053 | **.1626** |
| | 60 | .0165 | .2501 | −.0061 | .1681 | −.0037 | .1513 | **−.0010** | **.1351** |
| | 80 | .0140 | .2412 | **−.0002** | .1533 | .0010 | .1354 | .0090 | **.1198** |
| | 100 | .0153 | .2077 | **.0032** | .1501 | .0036 | .1243 | .0087 | **.1122** |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each order of fractional integration ($d$) and number of observations ($T$). $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

by the FML in the presence of such persistent autoregression.

In general, the FML performs well, but it does break down in two notable circumstances. First, in the presence of moderate autoregression ($\phi = 0.40$), the FML is negatively biased when the order of fractional integration is close to the border between stationary and non-persistent, $d = .45$ (and also when $d = 0.25$, but not to the same exent). By undervaluing the persistence of the series, the FML estimates will properly indicate the presence of fractional integration, but some autocorrelation in the series will be left unaddressed. This problem is alleviated somewhat by additional observations, but for this relatively small window, it doesn't appear that there is an easy fix, and extra caution is urged. A potential check of one's estimate is to fractionally difference the series by 0.50 and then re-estimate the series. By fractionally differencing by this value, the series is placed into a range of $d$ better estimated by the FML. If the estimates match, then one can be confident in the performance of the FML.

The FML also suffers when the series is non-persistent and faces severe autoregression ($d = .75, \phi = 0.80$). When a series is in this range and the extent of autoregression is very high, the FML will over estimate the level of fractional integration, likely leading the researcher to conclude that the series is a first-differenced unit root process. The estimate of a unit root will lead the researcher to first difference the data, thereby overdifferencing it. After integer differencing, the series will likely exhibit a significant MA parameter as a result (Box-Steffensmeier and Smith 1998; Lebo, Walker, and Clarke 2000). Note that in the same range, the the EML provides a reliable estimate of both the order of fractional integration as well as a consistent estimate of the AR parameter itself. Interestingly, this is the one instance where the EML is not biased in the negative direction.

The next question is how overfitting the model affects the estimates, which is an issue only with the parametric estimators. The results of misspecifying the model can be found in the lower left quadrant of Table 3.3, $\phi = 0$. Consistent with past results the EML generally suffers, providing negatively biased estimates as well as exhibiting an inflated

Table 3.3: **Parametric Estimators - ARFIMA (1,$d$,0)**

| | | | EML | | FML | | | | | EML | | FML | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | 0 | 40 | −.3000 | .4753 | **−.0746** | **.2668** | .40 | 0 | 40 | −.3329 | .4237 | **−.0055** | **.3082** |
| | | 60 | −.1551 | .2825 | **−.0489** | **.2008** | | | 60 | −.2834 | .3815 | **−.0167** | **.2916** |
| | | 80 | −.1119 | .2131 | **−.0363** | **.1609** | | | 80 | −.2516 | .3549 | **−.0427** | **.2623** |
| | | 100 | −.0836 | .1747 | **−.0262** | **.1421** | | | 100 | −.2466 | .3439 | **−.0454** | **.2559** |
| | 0.25 | 40 | −.3679 | .5576 | **−.1225** | **.3130** | | 0.25 | 40 | −.4321 | .4790 | **−.1013** | **.2891** |
| | | 60 | −.2034 | .3402 | **−.0554** | **.2048** | | | 60 | −.3823 | .4309 | **−.0871** | **.2689** |
| | | 80 | −.1375 | .2356 | **−.0340** | **.1615** | | | 80 | −.3443 | .4025 | **−.0961** | **.2674** |
| | | 100 | −.0960 | .1712 | **−.0273** | **.1408** | | | 100 | −.3280 | .3895 | **−.0860** | **.2522** |
| | 0.45 | 40 | −.4323 | .6075 | **−.1323** | **.2816** | | 0.45 | 40 | −.4961 | .5241 | **−.2447** | **.3456** |
| | | 60 | −.2428 | .3645 | **−.0857** | **.2084** | | | 60 | −.4711 | .5021 | **−.1717** | **.3306** |
| | | 80 | −.1632 | .2432 | **−.0533** | **.1369** | | | 80 | −.4334 | .4683 | **−.1976** | **.2971** |
| | | 100 | −.1245 | .1829 | **−.0339** | **.1082** | | | 100 | −.4052 | .4451 | **−.1876** | **.2903** |
| | 0.75 | 40 | −.2391 | .3778 | **−.0186** | **.1989** | | 0.75 | 40 | −.2690 | .3912 | **.0643** | **.2982** |
| | | 60 | −.1564 | .2837 | **−.0156** | **.1599** | | | 60 | −.2438 | .3651 | **.0341** | **.2627** |
| | | 80 | −.1140 | .2100 | **−.0204** | **.1435** | | | 80 | −.2394 | .3439 | **−.0021** | **.2293** |
| | | 100 | −.0887 | .1767 | **−.0177** | **.1236** | | | 100 | −.2145 | .3284 | **−.0174** | **.2038** |
| 0 | 0 | 40 | −.4778 | .6210 | **−.1031** | **.3283** | .80 | 0 | 40 | **−.0763** | **.1886** | .0852 | .2668 |
| | | 60 | −.3510 | .5148 | **−.0827** | **.2820** | | | 60 | −.0606 | **.1698** | **.0415** | .2421 |
| | | 80 | −.2913 | .4638 | **−.0805** | **.2504** | | | 80 | −.0372 | **.1572** | **.0481** | .2298 |
| | | 100 | −.2210 | .3916 | **−.0571** | **.2176** | | | 100 | **−.0268** | **.1549** | .0333 | .2139 |
| | 0.25 | 40 | −.5487 | .6710 | **−.1802** | **.3858** | | 0.25 | 40 | −.1395 | **.1981** | **.0137** | .2195 |
| | | 60 | −.4222 | .5679 | **−.1244** | **.3164** | | | 60 | −.1080 | **.1716** | **.0254** | .2094 |
| | | 80 | −.3188 | .4833 | **−.0840** | **.2681** | | | 80 | −.0896 | **.1538** | **.0155** | .1935 |
| | | 100 | −.2517 | .4083 | **−.0800** | **.2608** | | | 100 | −.0850 | **.1458** | **.0018** | .1816 |
| | 0.45 | 40 | −.6855 | .7536 | **−.2446** | **.4030** | | 0.45 | 40 | −.1970 | **.2302** | **−.1375** | .2298 |
| | | 60 | −.6015 | .6941 | **−.1860** | **.3515** | | | 60 | −.1619 | **.1909** | **−.1265** | .2084 |
| | | 80 | −.4799 | .6034 | **−.1528** | **.3117** | | | 80 | −.1520 | **.1779** | **−.1068** | .1853 |
| | | 100 | −.3821 | .5156 | **−.1038** | **.2494** | | | 100 | −.1360 | **.1607** | **−.1056** | .1791 |
| | 0.75 | 40 | −.3880 | .5226 | **.0217** | **.2352** | | 0.75 | 40 | **.0178** | **.2030** | .3715 | .4747 |
| | | 60 | −.3095 | .4607 | **−.0238** | **.2027** | | | 60 | **.0368** | **.2068** | .3496 | .4556 |
| | | 80 | −.2480 | .4074 | **−.0324** | **.1803** | | | 80 | **.0412** | **.2000** | .3216 | .4291 |
| | | 100 | −.1835 | .3403 | **−.0296** | **.1562** | | | 100 | **.0591** | **.2093** | .3076 | .4265 |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each specified level of $\phi$, $d$, and $T$. $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

## Table 3.4: Semiparametric Estimators - ARFIMA (1,$d$,0)

| | | | LPR $[m = T^{0.50}]$ | | LPR $[m = T^{0.65}]$ | | LW $[m = T^{0.50}]$ | | LW $[m = T^{0.65}]$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | 0 | 40 | **−.0309** | .4326 | −.1069 | .3149 | −.0495 | .3406 | −.1306 | **.2648** |
| | | 60 | **−.0276** | .3981 | −.0835 | .2577 | −.0484 | .3053 | −.1031 | **.2117** |
| | | 80 | **.0087** | .3459 | −.0479 | .2114 | −.0310 | .2731 | −.0723 | **.1747** |
| | | 100 | **−.0053** | .2988 | −.0442 | .1939 | −.0303 | .2469 | −.0631 | **.1654** |
| | 0.25 | 40 | **−.0296** | .4393 | −.0982 | .3164 | −.0790 | .3681 | −.1435 | **.2879** |
| | | 60 | **−.0282** | .3897 | −.0843 | .2462 | −.0676 | .3267 | −.1127 | **.2215** |
| | | 80 | **−.0316** | .3588 | −.0702 | .2229 | −.0660 | .3138 | −.0927 | **.1964** |
| | | 100 | **.0005** | .3039 | −.0407 | .1964 | −.0311 | .2580 | −.0624 | **.1688** |
| | 0.45 | 40 | **−.0280** | .4240 | −.0879 | .3115 | −.0825 | .3650 | −.1382 | **.2846** |
| | | 60 | **−.0056** | .3826 | −.0650 | .2452 | −.0592 | .3331 | −.1100 | **.2208** |
| | | 80 | **−.0188** | .3415 | −.0554 | .2089 | −.0515 | .3102 | −.0870 | **.1882** |
| | | 100 | **−.0077** | .3021 | −.0301 | .1994 | −.0398 | .2603 | −.0538 | **.1697** |
| | 0.75 | 40 | −.0534 | .4434 | −.0954 | .3245 | **.0014** | .2722 | −.0610 | **.1982** |
| | | 60 | **.0064** | .3523 | −.0482 | .2371 | .0097 | .2402 | −.0512 | **.1715** |
| | | 80 | −.0202 | .3624 | −.0594 | .2295 | **.0083** | .2393 | −.0514 | **.1618** |
| | | 100 | .0094 | .2953 | −.0324 | .1915 | **.0069** | .2113 | −.0394 | **.1432** |
| .40 | 0 | 40 | .1691 | .4581 | .2572 | .3927 | **.1311** | .3849 | .2206 | **.3350** |
| | | 60 | .1071 | .4157 | .2169 | .3281 | **.0748** | .3262 | .1960 | **.2766** |
| | | 80 | .1062 | .3654 | .1996 | .2879 | **.0635** | .2889 | .1828 | **.2457** |
| | | 100 | .0829 | .3095 | .1644 | .2517 | **.0559** | .2571 | .1478 | **.2142** |
| | 0.25 | 40 | .1772 | .4878 | .2696 | .4120 | **.1152** | .3829 | .2161 | **.3367** |
| | | 60 | .1018 | .4135 | .2153 | .3190 | **.0623** | .3329 | .1925 | **.2702** |
| | | 80 | .0676 | .3665 | .1839 | .2816 | **.0300** | .3118 | .1652 | **.2408** |
| | | 100 | .0913 | .3129 | .1698 | .2543 | **.0610** | .2662 | .1504 | **.2179** |
| | 0.45 | 40 | .1810 | .4671 | .2608 | .3988 | **.1069** | .3523 | .2059 | **.3101** |
| | | 60 | .1169 | .4070 | .2248 | .3304 | **.0623** | .3272 | .1859 | **.2648** |
| | | 80 | .0763 | .3551 | .1903 | .2783 | **.0408** | .3041 | .1639 | **.2352** |
| | | 100 | .0800 | .3131 | .1771 | .2670 | **.0465** | .2586 | .1565 | **.2238** |
| | 0.75 | 40 | **.1458** | .4526 | .2610 | .3964 | .1536 | .3576 | .2472 | **.3500** |
| | | 60 | .1307 | .3820 | .2471 | .3398 | **.1061** | **.2906** | .2231 | .3001 |
| | | 80 | **.0697** | .3716 | .1850 | .2869 | .0769 | .2728 | .1839 | **.2555** |
| | | 100 | .0953 | .3056 | .1726 | .2523 | **.0788** | .2404 | .1610 | **.2201** |
| .80 | 0 | 40 | .6608 | .7852 | .7475 | .8055 | **.5795** | **.6740** | .6860 | .7245 |
| | | 60 | .5595 | .6885 | .6907 | .7301 | **.5091** | **.6042** | .6678 | .6942 |
| | | 80 | .5046 | .6126 | .6642 | .6962 | **.4685** | **.5534** | .6561 | .6778 |
| | | 100 | .4697 | .5533 | .6100 | .6396 | **.4486** | **.5173** | .6096 | .6304 |
| | 0.25 | 40 | .6540 | .7866 | .7336 | .7967 | **.5103** | **.5853** | .6068 | .6350 |
| | | 60 | .5501 | .6743 | .6801 | .7180 | **.4679** | **.5433** | .6173 | .6334 |
| | | 80 | .4621 | .5961 | .6359 | .6723 | **.4143** | **.5006** | .6010 | .6174 |
| | | 100 | .4783 | .5619 | .6109 | .6386 | **.4439** | **.5032** | .5904 | .6060 |
| | 0.45 | 40 | .6258 | .7604 | .6780 | .7423 | **.4221** | **.4750** | .4868 | .5011 |
| | | 60 | .5300 | .6581 | .6408 | .6869 | **.3862** | **.4501** | .4991 | .5089 |
| | | 80 | .4574 | .5716 | .6112 | .6449 | **.3638** | **.4299** | .5017 | .5098 |
| | | 100 | .4391 | .5362 | .5861 | .6200 | **.3708** | **.4226** | .5001 | .5079 |
| | 0.75 | 40 | .6425 | .7788 | .7504 | .8108 | **.6047** | **.7083** | .7183 | .7617 |
| | | 60 | .5778 | .6810 | .7217 | .7595 | **.5321** | **.6180** | .6969 | .7264 |
| | | 80 | .4702 | .5933 | .6482 | .6824 | **.4582** | **.5533** | .6548 | .6806 |
| | | 100 | .4805 | .5618 | .6218 | .6495 | **.4636** | **.5303** | .6262 | .6459 |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each specified level of $\phi$, $d$, and $T$. $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

RMSE compared to properly specified models. The FML is also biased when attempting to fit a misspecified model, however not to the same extent. These results should hopefully push the researcher to rely on multiple information criteria to ensure model specification before estimating higher-order models.

Table 3.4 provides the results of the semiparametric estimators with the same series. Because the inclusion of a greater number of Fourier frequencies induces bias, the results presented represent only two choices of bandwidth, $m = T^{0.50}$ and $m = T^{0.65}$. The contamination caused by the presence of short-run dynamics is clearly evident when comparing the same estimator across different bandwidths. With $(0, d, 0)$ series, the researcher can generally increase bandwidth without penalty, but as seen here, the decrease in RMSE is paid for with greater bias in the estimates. Comparing both estimators across levels of $\phi$, the semiparametric models generally handle negative autoregression, but problems arise with positive AR noise, and the bias increases with the strength of autoregression. As expected, the LW outperforms the LPR in terms of both bias and RMSE.

### 3.6.3 Results - (0,$d$,1) Models

Tables 3.5 and 3.6 present the estimation results for when short-run dynamics are introduced in the form of a moving average parameter. For the parametric estimators, the EML is again dominated by the FML, however it doesn't appear that the presence of MA noise is nearly the burden on estimation that was present when an AR process was introduced. This holds true even when the models are misspecified - there is very little cost involved with overfitting the model in search of an MA parameter. Both the EML and FML can generally handle estimating the $(0, d, 1)$ model, and the point estimates for each become more efficient with the addition of observations. The EML does display some significant negative bias in its estimates when the MA parameter is negative and the fractional integration is positive however, and this bias should lead to Type II errors when the null is no fractional integration. As a result, the EML cannot be recommended. On the other hand, the FML is consistent and efficient and doesn't present any significant issues with estimation of the long memory parameter.

The semiparametric estimates also support the view that the presence of MA noise is not nearly as detrimental to the estimation of fractional integration as AR dynamics. The LPR and LW are fairly consistent in their estimates - lower bandwidth reduces the contamination from the high frequency MA parameter, and both the bias and RMSE decrease with an increase in the usable observations. As before, the LW dominates the LPR in terms of RMSE performance and when the MA parameter is positive also exhibits a lower overall bias in estimates. All told, both estimators function fairly well, and with the bandwidth set at $m = T^{0.5}$ both semiparametric estimators actually outperform the parametric EML when the MA parameter is positive.

## 3.7 How to Estimate Fractional Integration

The simulation results indicate that fractional integration can be reliably estimated, but as is made clear in the simulation roundup in Table 3.7 they also caution against simply treating the estimations of all series as one size fits all. This section briefly provides a how-to guide of best practices for the estimation of fractional integration.

**1. Stop using Stata.**

In general, Stata is a very powerful statistical program, but when it comes to the estimation of FI, it has some serious drawbacks. First, the stock estimator in the 'arfima' package is the EML, which is biased and unreliable. Unfortunately, the only efficient fractional differencing command in the program is reliant on the EML estimates. Both the LW and LPR estimators are available as .ado additions, but without the capacity for fractionally differencing the series by those estimates, they are only descriptive. Second, because Stata is proprietary, the program is a black-box. One types in a command, but often with very little knowledge of the underlying statistical process generating the results. And should one want to investigate what is happening under the hood, they are stuck, unless the program is an .ado file at which point, Stata's code is so clunky that divining what is happening is

Table 3.5: Parametric Estimators - ARFIMA (0,*d*,1)

| | | | EML | | FML | | | | | EML | | FML | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\theta$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | $\theta$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | 0 | 40 | −.2975 | .4517 | **−.1047** | **.3165** | .40 | 0 | 40 | −.1491 | .2656 | **−.0162** | **.2549** |
| | | 60 | −.2216 | .3506 | **−.0822** | **.2941** | | | 60 | −.1081 | **.1935** | **−.0125** | .1978 |
| | | 80 | −.2046 | .3141 | **−.0516** | **.2753** | | | 80 | −.0779 | .1625 | **−.0138** | **.1545** |
| | | 100 | −.1642 | .2677 | **−.0310** | **.2554** | | | 100 | −.0642 | .1406 | **−.0069** | **.1348** |
| | 0.25 | 40 | −.4072 | .5024 | **−.2004** | **.3838** | | 0.25 | 40 | −.1801 | .2626 | **−.0454** | **.2366** |
| | | 60 | −.3038 | .3850 | **−.1041** | **.3003** | | | 60 | −.1280 | .1971 | **−.0189** | **.1706** |
| | | 80 | −.2405 | .3075 | **−.0730** | **.2570** | | | 80 | −.0979 | .1586 | **−.0136** | **.1474** |
| | | 100 | −.2046 | .2754 | **−.0520** | **.2371** | | | 100 | −.0767 | .1360 | **−.0087** | **.1288** |
| | 0.45 | 40 | −.4465 | .5189 | **−.2210** | **.3597** | | 0.45 | 40 | −.2156 | .2707 | **−.0813** | **.1902** |
| | | 60 | −.3342 | .3969 | **−.1554** | **.2793** | | | 60 | −.1578 | .1955 | **−.0518** | **.1371** |
| | | 80 | −.2717 | .3178 | **−.1168** | **.2257** | | | 80 | −.1285 | .1643 | **−.0373** | **.1101** |
| | | 100 | −.2438 | .2921 | **−.0899** | **.1893** | | | 100 | −.1028 | .1329 | **−.0242** | **.0915** |
| | 0.75 | 40 | −.0785 | .3785 | **.0716** | **.2686** | | 0.75 | 40 | −.1392 | .2722 | **.0241** | **.2402** |
| | | 60 | −.0122 | .3489 | **−.0314** | **.2911** | | | 60 | −.0878 | .1927 | **.0062** | **.1778** |
| | | 80 | .0053 | .3190 | **.0388** | **..2477** | | | 80 | −.0720 | .1590 | **−.0047** | **.1479** |
| | | 100 | .0097 | .3048 | **.0288** | **.2303** | | | 100 | −.0597 | .1378 | **−.0081** | **.1221** |
| 0 | 0 | 40 | −.2124 | .3401 | **−.0144** | **.3161** | .80 | 0 | 40 | −.1184 | .2154 | **−.0100** | **.2077** |
| | | 60 | −.1518 | **.2530** | **.0059** | .2687 | | | 60 | −.0761 | .1608 | **−.0145** | **.1547** |
| | | 80 | −.1111 | **.2096** | **−.0050** | .2244 | | | 80 | −.0607 | .1289 | **−.0123** | **.1248** |
| | | 100 | −.1006 | **.1861** | **.0014** | .1938 | | | 100 | −.0528 | .1128 | **−.0077** | **.1104** |
| | 0.25 | 40 | −.2506 | .3405 | **−.0777** | **.2948** | | 0.25 | 40 | −.1387 | .2248 | **−.0246** | **.1986** |
| | | 60 | −.1791 | .2546 | **−.0283** | **.2171** | | | 60 | −.0972 | .1639 | **−.0130** | **.1418** |
| | | 80 | −.1307 | .2003 | **−.0194** | **.1852** | | | 80 | −.0744 | .1295 | **−.0105** | **.1237** |
| | | 100 | −.1100 | .1770 | **−.0112** | **.1673** | | | 100 | −.0553 | .1066 | **−.0063** | **.1080** |
| | 0.45 | 40 | −.2799 | .3408 | **−.1147** | **.2404** | | 0.45 | 40 | −.1770 | .2289 | **−.0603** | **.1590** |
| | | 60 | −.2001 | .2467 | **−.0759** | **.1788** | | | 60 | −.1236 | .1607 | **−.0392** | **.1140** |
| | | 80 | −.1694 | .2117 | **−.0577** | **.1431** | | | 80 | −.0953 | .1264 | **−.0254** | **.0925** |
| | | 100 | −.1416 | .1795 | **−.0404** | **.1178** | | | 100 | −.0831 | .1126 | **−.0176** | **.0794** |
| | 0.75 | 40 | −.1228 | .3334 | **.0663** | **.3076** | | 0.75 | 40 | −.1179 | .2248 | **.0126** | **.1856** |
| | | 60 | −.1181 | .2785 | **.0495** | **.2743** | | | 60 | −.0771 | .1648 | **−.0035** | **.1395** |
| | | 80 | −.0886 | **.2215** | **.0205** | .2267 | | | 80 | −.0578 | .1333 | **−.0098** | **.1217** |
| | | 100 | −.0850 | **.1900** | **.0108** | .1929 | | | 100 | −.0531 | .1169 | **−.0084** | **.1024** |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each specified level of $\theta$, $d$, and $T$. $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

## Table 3.6: Semiparametric Estimators - ARFIMA (0,$d$,1)

| | | | LPR [$m = T^{0.50}$] | | LPR [$m = T^{0.65}$] | | LW [$m = T^{0.50}$] | | LW [$m = T^{0.65}$] | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\theta$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | 0 | 40 | **−.1595** | .4624 | −.2438 | .3934 | −.1611 | .3430 | −.2638 | **.3314** |
| | | 60 | **−.1008** | .3961 | −.2194 | .3262 | −.1309 | .3120 | −.2267 | **.2881** |
| | | 80 | **−.0737** | .3644 | −.1930 | .2864 | −.0944 | .2801 | −.2016 | **.2572** |
| | | 100 | **−.0668** | .2950 | −.1470 | .2429 | −.0930 | .2572 | −.1722 | **.2301** |
| | 0.25 | 40 | **−.1655** | .4796 | −.2534 | .4012 | −.2168 | .4047 | −.2856 | **.3722** |
| | | 60 | **−.1307** | .4060 | −.2182 | .3211 | −.1493 | .3488 | −.2548 | **.3186** |
| | | 80 | **−.0851** | .3589 | −.1863 | .2828 | −.1191 | .3152 | −.2082 | **.2678** |
| | | 100 | **−.0769** | .2938 | −.1544 | .2430 | −.1033 | .2773 | −.1822 | **.2412** |
| | 0.45 | 40 | **−.1592** | .4470 | −.2363 | .3800 | −.2100 | .4197 | −.2785 | **.3733** |
| | | 60 | **−.0818** | .3990 | −.1957 | **.3099** | −.1194 | .3498 | −.2468 | .3155 |
| | | 80 | **−.0691** | .3688 | −.1655 | **.2634** | −.0937 | .3056 | −.2115 | .2724 |
| | | 100 | **−.0570** | .3059 | −.1525 | .2505 | −.0925 | .2824 | −.1700 | **.2284** |
| | 0.75 | 40 | −.1256 | .4565 | −.2356 | .3961 | **−.0564** | .2532 | −.1541 | **.2065** |
| | | 60 | −.0910 | .3934 | −.1894 | .3027 | **−.0483** | .2274 | −.1472 | **.1930** |
| | | 80 | −.0543 | .3558 | −.1696 | .2687 | **−.0361** | .2190 | −.1455 | **.1851** |
| | | 100 | −.0524 | .3073 | −.1394 | .2374 | **−.0463** | .1911 | −.1281 | **.1726** |
| .40 | 0 | 40 | **.0610** | .4425 | .1004 | .3171 | .0357 | .3540 | .0871 | .2618 |
| | | 60 | **.0283** | .3998 | .0700 | .2544 | −.0025 | .3060 | .0620 | **.2041** |
| | | 80 | **.0125** | .3411 | .0658 | .2138 | −.0235 | .2708 | .0496 | **.1722** |
| | | 100 | **.0251** | .2922 | .0507 | .1976 | −.0153 | .2384 | .0370 | **.1677** |
| | 0.25 | 40 | .0517 | .4411 | .1273 | .3233 | **−.0129** | .3649 | .0546 | **.2569** |
| | | 60 | **.0307** | .3820 | .0692 | .2545 | −.0319 | .3317 | .0559 | **.2065** |
| | | 80 | .0446 | .3389 | .0647 | .2117 | **−.0279** | .2963 | .0469 | **.1745** |
| | | 100 | .0251 | .3072 | .0498 | .1982 | **.0061** | .2486 | .0348 | **.1661** |
| | 0.45 | 40 | .0694 | .4415 | .1168 | .3241 | **−.0038** | .3494 | .0675 | **.2527** |
| | | 60 | .0466 | .3803 | .0902 | .2547 | **−.0103** | .3185 | .0519 | **.1981** |
| | | 80 | .0507 | .3393 | .0859 | .2263 | **−.0026** | .2892 | .0432 | **.1835** |
| | | 100 | .0322 | .3048 | .0626 | .1958 | **−.0149** | .2615 | .0388 | **.1616** |
| | 0.75 | 40 | **.0626** | .4271 | .1162 | .3209 | .0862 | .3284 | .1148 | **.2515** |
| | | 60 | **.0321** | .3855 | .0840 | .2561 | .0614 | .2749 | .0779 | **.2036** |
| | | 80 | **.0129** | .3511 | .0647 | .2179 | .0477 | .2571 | .0570 | **.1722** |
| | | 100 | **.0135** | .2903 | .0526 | .2000 | .0186 | .2177 | .0442 | **.1546** |
| .80 | 0 | 40 | .0546 | .4503 | .1309 | .3355 | **.0326** | .3432 | .1007 | **.2713** |
| | | 60 | .0351 | .3772 | .1000 | .2595 | **−.0063** | .3016 | .0858 | **.2249** |
| | | 80 | .0181 | .3489 | .0795 | .2193 | **−.0086** | .2776 | .0668 | **.1852** |
| | | 100 | .0114 | .3155 | .0608 | .2012 | **−.0040** | .2453 | .0464 | **.1601** |
| | 0.25 | 40 | .0534 | .4328 | .1384 | .3445 | **.0051** | .3647 | .0884 | **.2698** |
| | | 60 | .0446 | .3865 | .1008 | .2568 | **−.0145** | .3201 | .0720 | **.2244** |
| | | 80 | .0258 | .3373 | .0890 | .2273 | **−.0089** | .2938 | .0506 | **.1890** |
| | | 100 | .0236 | .2968 | .0587 | .2009 | **−.0094** | .2456 | .0476 | **.1663** |
| | 0.45 | 40 | .0673 | .4641 | .1376 | .3426 | **−.0026** | .3481 | .0915 | **.2556** |
| | | 60 | .0406 | .3953 | .1185 | .2766 | **.0017** | .3317 | .0736 | **.2075** |
| | | 80 | .0422 | .3459 | .0921 | .2269 | **−.0160** | .3032 | .0738 | **.1831** |
| | | 100 | .0409 | .2979 | .0798 | .2079 | **.0115** | .2475 | .0486 | **.1634** |
| | 0.75 | 40 | **.0647** | .4486 | .1591 | .3490 | .0799 | .3132 | .1293 | **.2638** |
| | | 60 | .0660 | .3847 | .1038 | .2649 | **.0415** | .2608 | .1031 | **.2090** |
| | | 80 | **.0231** | .3402 | .0764 | .2249 | .0247 | .2406 | .0682 | **.1821** |
| | | 100 | .0351 | .2963 | .0744 | .2039 | **.0251** | .2114 | .0577 | **.1590** |

*Note*: Results based on 1000 simulations for each model. Bolded entries are the low value for each specified level of $d$ and $\phi$. $d = 0.75$ indicates a series that was first differenced and then had 1 added back to its estimates.

**Table 3.7: Simulation Roundup**

| | EML | FML | LPR | LW |
|---|---|---|---|---|
| $(0,d,0)$ | Exhibits negative bias across all observations. Not recommended. | Unbiased & consistent. Outperforms all other estimators. Recommend $\geq$ 40 observations | Unbiased & consistent. Recommend $\geq$ 40 observations | Unbiased & consistent. Outperforms the LPR. Recommend $\geq$ 40 observations. |
| $(1,d,0)$ | Negatively biased estimates. Bias worsens as order $I(d)$ increases. Exception is strongly persistent $\phi = .80$ Not recommended. | Generally unbiased. Exceptions is $\phi = .40$ and $I(d)$ close to stationary border .50, and $\phi = .80, I(d) = .75$. Recommend $\geq$ 80 observations. | Unbiased estimates with negative $\phi$. Positive bias as $\phi$ increases. Lower bandwidth, lower bias. Ok if negative $\phi$, not positive $\phi$. | Minimal negative bias with negative $\phi$. Positive bias as $\phi$ increases. Lower bandwidth, lower bias. LPR outperforms with negative $\phi$. |
| $(0,d,1)$ | Negatively biased estimates. Bias worsens as $I(d)$ approaches stationary border $d = .50$ Not recommended. | Generally unbiased. Exception of negative $\theta$ and $I(d)$ close to stationary border. If positive $\theta$ recommend $\geq$ 60 observations. | Generally unbiased at lower bandwidths. Exception is negative $\theta$ & increasing $I(d)$. If positive $\theta$ recommend $\geq$ 60 observations. | Generally unbiased at lower bandwidths. Exception is negative $\theta$ & increasing $I(d)$. If positive $\theta$ recommend $\geq$ 60 observations. |

difficult. R and Matlab are both better programs and each have active user communities that share and refine programs and code. Most economists appear to have switched to Matlab and they are quite generous in sharing their code.

**2. Get to know your data.**

One should have a pretty good idea of whether their series are fractionally integrated before they ever estimate the order of fractional integration. The first step is to visualize the data. In the time domain this includes plotting the data as well as investigating the autocorrelation function of the series. The persistence of the ACF is a good indicator as to the potential for the long memory of a series. Visualizing the series in the frequency domain is also incredibly helpful, as demonstrated in Appendix B.1. If the low frequencies are the greatest contributors to the variance of the process and the higher frequencies have very little influence, the probability of the series being fractionally integrated is increased. Stationarity tests are also helpful despite their limited power when applied to fractionally integrated processes. If one uses the Dickey-Fuller, Augmented Dickey-Fuller, DF-GLS, KPSS test, and Variance Ratio tests and the estimators do not consistently find that the series is either stationary or a unit-root, it is likely that it is fractionally integrated.

**3. Begin with semiparametric estimators.**

Semiparametric estimators are simple and easy to estimate because they are agnostic as to the short-run dynamics of the process. As noted above, both the LW and LPR estimators are asymptotically normal for $d \in (-1/2, 3/4)$. Economists have also developed an extended local Whittle estimator (ELW) that is consistent and normally distributed for $d \in (-1/2, 2)$ and this estimator is freely available for use in Matlab (Shimotsu 2010). The capacity of the ELW to take on greater values of $d$ means that the researcher need not make assumptions about the order of integration of the series in its level form by first differencing the data. The estimator provides a nice first estimate of the fractional integration of the series. With this estimate, the researcher can then determine whether to first difference their series and can use the LW estimator, which generally outperforms the LPR. With all semiparametric estimators the choice of bandwidth is crucial and the estimation with various bandwidths can help diagnose the possible presence of short-run dynamics. Begin with a bandwidth of $T^{0.50}$ and then increase by increments of 0.05 or 0.10 up to $T^{0.80}$. If the estimates are consistent across bandwidths, higher frequencies are not contaminating the estimates, which indicates that short-run ARMA dynamics are not present. The ELW of Shimotsu (2010) automatically detrends the data, but don't forget to detrend the series before using the LW or LPR estimators.

**4. Use the FML estimator.**

As demonstrated, the FML is the best maximum likelihood estimator of fractional integration. It can reliably account for short-run ARMA dynamics and it is generally unbiased in their presence. Further, unlike semiparametric estimators that can only generate an asymptotic standard error based on the length of the series and the bandwidth, the FML produces series-specific standard errors. When using the FML, estimate the series as a $(0, d, 0)$ series, but also estimate for higher-order processes. Use the information criteria of the estimator (AIC, BIC, and log-likelihood) to help diagnose the possibility of ARMA dynamics. In addition to the information criteria, pay attention to the significance of the ARMA estimates

- higher-order estimations should only be made if the information criteria indicate a better model fit and the estimates are statistically significant. Between the semiparametric estimates made at varying bandwidths, the information criteria, and the estimates themselves, one should have a very good idea of the order of the process.

**5. Fractionally difference the data.**

Assuming that the process is fractionally integrated, fractionally difference the data by the estimated order of fractional integration, and then repeat step 2. In the time domain the series should look like white noise - it should be bouncing around the mean and the ACF should not indicate any significant spikes. In the frequency domain it should look like white noise - the contribution to the variance should be constant across all frequencies. All stationarity tests should indicate that it is a stationary process.

## 3.8 Conclusion

The purpose of this paper was to investigate whether fractional integration could reliably be estimated for series with a limited number of observations. Four estimators were compared, two parametric and two semiparametric, and each estimator's performance was evaluated in terms of bias and RMSE for three different DGPs. The results indicate that the answer is a qualified yes, depending on the data generating process and the estimator used. When the series is $(0, d, 0)$ pure fractional process, the frequency domain maximum likelihood (FML) and both semiparametric estimators render unbiased estimates of the $d$ parameter. The FML is clearly the most efficient estimator in this case, but the semiparametric local Whittle (LW) estimator is not far behind. In the presence of AR noise, the semiparametric estimators both suffer from contamination from the higher frequencies. As a result, their estimates of $d$ are biased, and the extent of the bias increases with the strength of the AR parameter. A reduction in the bandwidth serves to reduce some of this bias, and both the LW and LPR estimators are generally robust to short-run dynamics in the form of a negative AR parameter. The parametric FML estimator acquits itself well when attempting to estimate $(1, d, 0)$ series, exhibiting bias in only two instances - $(\phi = .80, d = .75)$ and $(\phi = .40, d = .45)$. In each of these instances, the direction of the bias will increase the probability of Type II errors.[8] The problems created by AR processes are generally not found when attempting to estimate fractional integration when MA noise is present. Both the semiparametric estimators and the FML are generally robust, with the FML again outperforming all other estimators.

The truly definitive take away from this exercise is to understand just how bad the time domain maximum likelihood estimator is across the full ranges of data types. Because the EML must account for the mean of the process, it exhibits a consistent and significant negative bias, which is made worse when short-run dynamics are present. The EML fails in almost all circumstances to the point that it is even outperformed by semiparametric estimators when moving average dynamics are present. The EML has nothing to offer it, and the FML should be the go to estimator for fractional integration. Unfortunately, the default estimator in Stata is the EML, which means that if a researcher is interested in fractional integration they are better served by using either R or Matlab.

[8] That the FML might produce Type II errors is the exact opposite conclusion reached by Keele and Linn (2015) and their simulations of the EML. By estimating fractional integration with the FML, the researcher will not find that all series are long memory.

## Bibliography


Baillie, Richard T, and George Kapetanios. 2009. Semi Parametric Estimation of Long Memory: Comparisons and Some Attractive Alternatives. In *conference in honor of Peter Phillip's Sixtieth Birthday at Singapore Management University*.

Beran, Jan. 1994. *Statistics for Long-Memory Processes*. Vol. 61 CRC Press.

Box, George EP, and David A Pierce. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." *Journal of the American Statistical Association* 65(332): 1509–1526.

Box-Steffensmeier, Janet, Kathleen Knight, and Lee Sigelman. 1998. "The Interplay of Macroideology and Macropartisanship: A Time Series Analysis." *Journal of Politics* 60: 131–149.

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

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

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

Byers, David, James Davidson, and David Peel. 1997. "Modelling Political Popularity: An Analysis of Long-range Dependence in Opinion Poll Series." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160(3): 471–490.

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

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

Cheung, Yin-Wong, and Francis X Diebold. 1994. "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally-Integrated Noise with Unknown Mean." *Journal of Econometrics* 62(2): 301–316.

Enders, Walter. 2004. *Applied Econometric Time Series*. 2nd ed. New York: Wiley.

Enns, Peter K. 2014. "The Public's Increasing Punitiveness and Its Influence on Mass Incarceration in the United States." *American Journal of Political Science* 58(4): 857–872.

Enns, Peter K, Nathan J Kelly, Janna Morgan, Thomas Volscho, and Christopher Witko. 2014. "Conditional Status Quo Bias and Top Income Shares: How U.S. Political institutions have Benefitted the Rich." *The Journal of Politics* 76: 289–303.

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

Geweke, John, and Susan Porter-Hudak. 1983. "The Estimation and Application of Long Memory Time Series Models." *Journal of time series analysis* 4(4): 221–238.

Gil-Alana, Luis A, and Javier Hualde. 2009. "Fractional Integration and Cointegration. An Overview and an Empirical Application." In *Palgrave Handbook of Econometrics*, ed. Terence C. Mills, and Kerry Patterson. Vol. 2 MacMillan Publishers London pp. 434–469.

Granger, Clive WJ. 1966. "The Typical Spectral Shape of an Economic Variable." *Econometrica: Journal of the Econometric Society* pp. 150–161.

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

Grant, Taylor, and Matthew J. Lebo. 2015. "Error Correction Methods with Political Time Series." *Political Analysis* .

Hauser, Michael A. 1999. "Maximum Likelihood Estimators for ARMA and ARFIMA Models: a Monte Carlo Study." *Journal of Statistical Planning and Inference* 80(1): 229–255.

Henry, M, and P. Zaffaroni. 2002. "The Long-Range Dependence Paradigm for Macroeconomics and Finance." In *Theory and Applications of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim, and M. Taqqu. Birkhäuser.

Henry, M, and PM Robinson. 1996. Bandwidth Choice in Gaussian Semiparametric Estimation of Long Range Dependence. In *Athens Conference on Applied Probability and Time Series Analysis*. Springer pp. 220–232.

Hosking, Jonathan RM. 1981. "Fractional Differencing." *Biometrika* 68(1): 165–176.

Hurvich, Clifford M, Rohit Deo, and Julia Brodsky. 1998. "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* 19(1): 19–46.

Jennings, Will, and Peter John. 2009. "The Dynamics of Political Attention: Public Opinion and the Queen's Speech in the United Kingdom." *American Journal of Political Science* 53(4): 838–854.

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

Keele, Luke, and Suzanna Linn. 2015. "Treating Time with All Due Seriousness." *Political Analysis* .

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

Kuensch, Hans R. 1987. Statistical Aspects of Self-Similar Processes. In *Proceedings of the first World Congress of the Bernoulli Society*. Vol. 1 VNU Science Press Utrecht, The Netherlands pp. 67–74.

Layman, Geoffrey C., Thomas Carsey, John C. Green, Richard Herrera, and Rosalyn Cooperman. 2010. "Activists and Conflict Extension in American Party Politics." *American Political Science Review* 104(2): 324–346.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

Li, WK, and A Ian McLeod. 1986. "Fractional Time Series Modelling." *Biometrika* 73(1): 217–221.

Nielsen, Morten Ørregaard, and Per Houmann Frederiksen. 2005. "Finite Sample Comparison of Parametric, Semiparametric, and Wavelet Estimators of Fractional Integration." *Econometric Reviews* 24(4): 405–443.

Phillips, Peter CB, and Katsumi Shimotsu. 2004. "Local Whittle Estimation in Nonstationary and Unit Root Cases." *The Annals of Statistics* 32(2): 656–692.

Robinson, Peter M. 1995*a*. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of statistics* 23(5): 1630–1661.

Robinson, Peter M. 1995*b*. "Log-Periodogram Regression of Time Series with Long Range Dependence." *The annals of Statistics* pp. 1048–1072.

Sanchez Urribarri, Raul A, Susanne Schorpp, Kirk A Randazzo, and Donald R Songer. 2011. "Explaining Changes to Rights Litigation: Testing a Multivariate Model in a Comparative Framework." *The Journal of Politics* 73(02): 391–405.

Shimotsu, Katsumi. 2010. "Exact Local Whittle Estimation of Fractional Integration with Unknown Mean and Time Trend." *Econometric Theory* 26(02): 501–540.

Shimotsu, Katsumi, and Peter CB Phillips. 2006. "Local Whittle Estimation of Fractional Integration and Some of its Variants." *Journal of Econometrics* 130(2): 209–233.

Sowell, Fallaw. 1992*a*. "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models." *Journal of econometrics* 53(1): 165–188.

Sowell, Fallaw. 1992*b*. "Modeling Long-Run Behavior with the Fractional ARIMA Model." *Journal of Monetary Economics* 29(2): 277–302.

Tsay, Wen-Jay, and Ching-Fan Chung. 2000. "The Spurious Regression of Fractional Integrated Processes." *Journal of Econometrics* 96: 155–182.

Ura, Joseph Daniel. 2014. "Backlash and Legitimation: Macro Political Responses to Supreme Court Decisions." *American Journal of Political Science* 58(1): 110–126.

Ura, Joseph Daniel, and Christopher R Ellis. 2012. "Partisan Moods: Polarization and the Dynamics of Mass Party Preferences." *The Journal of Politics* 74(1): 277–291.

Velasco, Carlos. 1999*a*. "Gaussian Semiparametric Estimation of Non-stationary Time Series." *Journal of Time Series Analysis* 20(1): 87–127.

Velasco, Carlos. 1999*b*. "Non-Stationary Log-Periodogram Regression." *Journal of Econometrics* 91(2): 325–371.

# Chapter 4

# Taking Auto-Correlation Seriously: Comparing Error Correction Techniques with Political Time Series

This chapter is an article that is co-authored with Matthew Lebo and is forthcoming in *Political Analysis*. All model estimates, Monte Carlo simulations, figures, and all material found in the appendix for this chapter were solely my own work product.


**Abstract**

While traditionally considered for non-stationary and cointegrated data, De Boef and Keele (2008) suggest applying a General Error Correction Model to stationary data with or without cointegration. The GECM has since become extremely popular in political science but practitioners have confused essential points. For one, the model is treated as perfectly flexible when, in fact, the opposite is true. Time series of various orders of integration – stationary, non-stationary, explosive, near- and fractionally-integrated – should not be analyzed together but researchers consistently make this mistake. That is, without *equation balance* the model is misspecified and hypothesis tests and long-run-multipliers are unreliable. Another problem is that the error correction term's sampling distribution moves dramatically depending upon the order of integration, sample size, number of covariates, and the *boundedness* of $Y_t$. This means that practitioners are likely to overstate evidence of error correction, especially when using a traditional $t$-test. We evaluate common GECM practices with six types of data, 746 simulations, and five paper replications.

## 4.1 Introduction

Nearly ninety years after Yule (1926) published "Why do we sometimes get nonsense correlations between time-series?" political applications are still plagued by spurious findings. Researchers often favor models that are easier to both implement and interpret and, when data are scarce, the desire for simplicity grows. Nevertheless, some models are particularly prone to give unreliable results.

Given the peculiarities of political data, one such model is the general (unrestricted) error correction model (GECM), first developed in econometrics (Davidson, Hendry, Srba, and Yeo 1978; Hendry and Mizon 1978) and introduced to political science over two decades ago (Beck 1992; Ostrom, Jr., and Smith 1993; Durr 1993).[1] More recently, De Boef and Keele (2008, D&K) re-introduced the GECM as an equivalent model to the autoregressive distributed lag (ADL) in an *AJPS* workshop piece, "Taking Time Seriously." D&K describe several promising aspects for stationary data: cointegration is unnecessary to look for error correction and one can estimate both short- and long-term impacts of covariates. Among political scientists, the method has since become extremely popular with D&K used as the go-to source. Google Scholar shows the paper cited 68 times in 2013 with many applications appearing in top journals.[2]

However, there are several problems with the GECM – some known and some explored here. For one, there are many types of political time series – integrated, near-integrated, fractionally-integrated, auto-regressive, and explosive – and researchers need to pay very close attention to the properties of their series. Each presents its own challenges in the GECM. Second, practitioners often fail to ensure that their equations are *balanced.* The order of integration needs to be consistent across all series in a model. Mixing together

[1]See: Banerjee, Dolado, Galbraith, and Hendry (1993) for the clearest exposition on the topic.

[2]As of October 2014 various versions of the paper have been cited 335 times. Recent examples of applied GECM papers include: Büthe and Milner (2014); Casillas, Enns, and Wohlfarth (2011); Enns (2014); Enns, Kelly, Morgan, Volscho, and Witko (2014); Faricy (2011); Jennings and John (2009); Kayser (2009); Kelly and Enns (2010); Kono (2008); Layman, Carsey, Green, Herrera, and Cooperman (2010); Ramirez (2009); Rickard (2012); Sanchez Urribarri, Schorpp, Randazzo, and Songer (2011); Ura (2014); Ura and Ellis (2012); Ura and Wohlfarth (2010); Volscho and Kelly (2012).

series of various orders on integration will mean a model is misspecified. Third, although the autoregressive distributed lag (ADL) model is algebraically equivalent to the GECM, the reorganization of parameters is not benign and easily leads to misinterpretation. Fourth, the interpretation of parameters should change across data types. With stationary data, the ECM's coefficient is misinterpreted so that mean reverting variables with little connection to each other are claimed to be in equilibrium relationships. Fifth, with unit-root data, the GECM's key hypothesis test is misunderstood: political science applications have incorrectly used a standard $t$-test for the ECM's coefficient. Sixth, using bounded series in the GECM has unexplored consequences. Given these and other issues, the GECM should only be used in rare instances, yet it is applied frequently and haphazardly. As a result, serious errors exist in the growing body of work that has relied upon the model.

We continue below with a time series primer, followed by a discussion of the GECM's performance with six types of dependent variables. Simulations demonstrate the issues with each case – typically an alarming rate of Type I errors but sometimes nonsensical inferences as well, e.g. error correction rates that can be misinterpreted to be above 100%. We also outline the problems presented by bounded political time series. Last, we replicate two recent GECM studies: Casillas, Enns, and Wohlfarth (2011) and Ura and Ellis (2012).[3] In each, standard practices greatly overstate the strength of relationships, especially concerning error correction. Without an appreciation of the method's limitations our understanding of dynamic political relationships will continue to be undermined.

## 4.2 Unit-Roots, Stationarity, and Error Correction

The univariate properties of time series $Y_t$ can be described as:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)(1 - L)^d Y_t = (1 + \sum_{i=1}^{q} \theta_i L^i)\epsilon_t \tag{4.1}$$

[3]Three additional replications are in the Supplement: Sanchez Urribarri et al. (2011), Kelly and Enns (2010), and Volscho and Kelly (2012). See Appendices C.5.1, C.5.2, and C.5.3 for respective details.

where $p$ refers to auto-regressive parameters, $q$ refers to moving average parameters, and $d$ is the (fractional) integration parameter. These can be summarized in a $(p, d, q)$ notation where $(0, 0, 0)$ is a white-noise, auto-correlation-free, process (Box and Jenkins 1970; Granger and Joyeux 1980; Hosking 1981; Enders 2004). $L$ is the lag operator such that $LY_t = Y_{t-1}$.

The parameter $d$ represents the memory of the series. If $d = 0$ the series is weakly stationary and has only short memory – it will tend towards a constant mean, has finite variance and constant covariance.[4] Autoregressive and moving average parameters may still exist where shocks will persist for finite periods as the series reverts back to its mean.[5] If $d = 1$ the series contains a unit-root (also known as integrated, $I(1)$, perfect-memoried, or a random-walk) with a non-stationary mean, variance, and covariances. It can wander in any direction with no expectation of returning to a long-term mean.

If we relax the assumption that the order of integration must be an integer, but can instead fall anywhere on a real number line, a series is considered fractionally integrated. If a series is $-1/2 < d < 1/2$ the process is invertible and possesses a linear representation. For $0 < d < 1$ the process is said to have long memory - it holds a mix of characteristics with long – not short or perfect – memory (Beran 1994; Box-Steffensmeier and Smith 1996; Granger and Joyeux 1980). If $1/2 < d < 1$ the series is variance and covariance non-stationary, however it is still mean reverting. (Baillie 1996). Fractional integration is widespread in political data (Box-Steffensmeier and Smith 1996; Box-Steffensmeier and Tomlinson 2000; Lebo, Walker, and Clarke 2000).

As any of the $(p, d, q)$ parameters deviate from zero, auto-correlation threatens hypothesis tests. Extensive work has concentrated on identifying auto-correlation and filtering to account for it (e.g.: Box and Jenkins 1970; Box-Steffensmeier and Smith 1996, 1998; Clarke and Stewart 1994; Clarke and Lebo 2003; Granger and Newbold 1974; Granger and Joyeux

[4] Enders (2004, p.54) describes a weakly stationary series as: mean constant ($E(Y_t) = E(Y_{t-s}) = \mu$), variance constant ($E[(Y_t - \mu)^2] = E[(Y_{\text{t-s}} - \mu)^2] = \sigma_y^2$) or ($var(Y_t) = var(Y_{\text{t-s}}) = \sigma_y^2$), and covariance constant ($E[(Y_t - \mu)(Y_{\text{t-s}} - \mu)] = E[(Y_{\text{t-j}} - \mu)(Y_{\text{t-j-s}} - \mu)] = \gamma_s$) or ($cov(Y_t, Y_{\text{t-s}}) = cov(Y_{\text{t-j}}, Y_{\text{t-j-s}}) = \gamma_s$).

[5] Near-integrated series have AR processes of almost 1 and are "long-memoried." The effects of shocks decrease at an exponential rate and the series are eventually stationary as well. (DeBoef and Granato 1997).

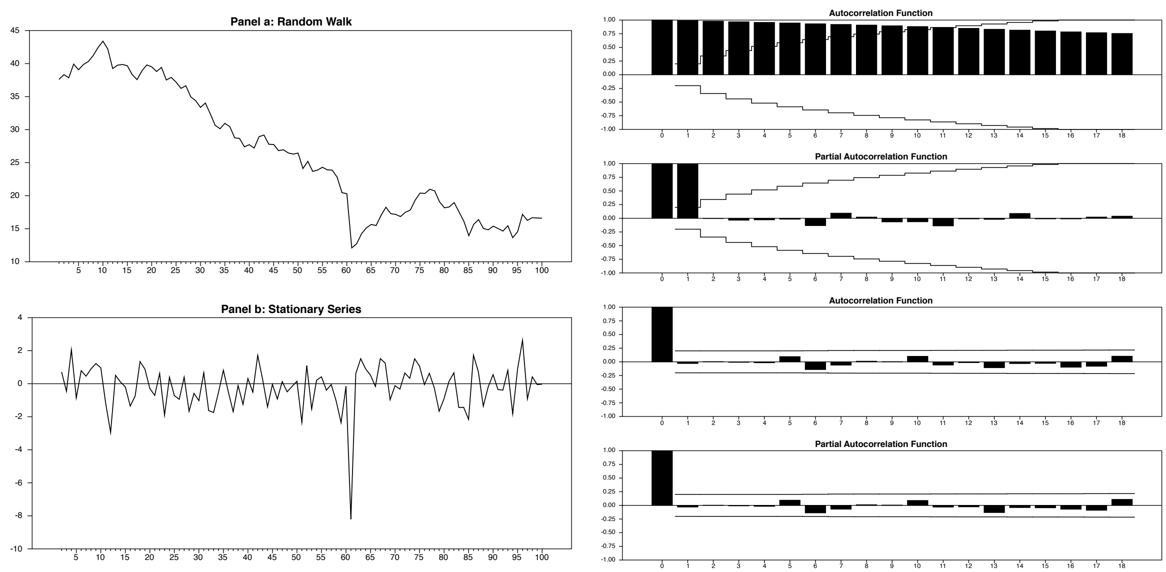
1980; Lebo, Walker, and Clarke 2000; Tsay and Chung 2000). This "pre-whitening," approach 1) identifies how a series depends upon its past values, then 2) models this behavior, and 3) uses the "white-noise" residuals so that data are *i.i.d.* and inferences are trustworthy. Pre-whitening approaches have developed from ARMA and ARIMA (Box and Pierce 1970; Box and Jenkins 1970) to ARFIMA models (Granger and Joyeux 1980; Hosking 1981). Researchers may be concerned about what is lost with pre-whitening – the common metaphor asks if one throws out the baby (interesting variation) with the bathwater (auto-correlation). But white noise does *not* mean that all useful information has been removed leaving only random noise. Rather, it means that the series' dependence on its own history has been filtered out and what is left – *seemingly* white noise – might be explained by independent variables. Thus, pre-whitening ensures that data hold the *i.i.d.* property.

The simplest example of pre-whitening begins with a unit-root, $(0, 1, 0)$: $Y_t = 1 * Y_{t-1} + \epsilon_t$. Figure 4.1, top-left, shows a unit-root series with its auto-correlation (ACF) and partial auto-correlation (PACF) plots to the right – significant spikes beyond the $0^{th}$ lag are deviations from *i.i.d.* and will cause problems.[6] First-differencing subtracts $Y_{t-1}$ from each side to create the white noise series, $Y_t^* = \epsilon_t$ with $\epsilon_t \sim N(0, \sigma^2)$ (bottom-left). As shown in $Y_t^*$'s ACF and PACF, it is free of auto-correlation. Regressions involving it will provide unbiased, efficient, and consistent estimates. Other filters can make other types of data safe as well: Box and Pierce (1970) use ARMA $(p, q)$ filters for auto-regressive and moving average processes, Box-Jenkins (1970) adapted it to unit-roots (ARIMA models), and Hosking (1981) introduced fractional differencing (ARFIMA) to filter a $(p, d, q)$ model into white noise, $(0, 0, 0)$.

Whether $(p, 0, q)$, $(p, 1, q)$, or $(p, d, q)$, the goal is the same: create white noise residuals for *all* variables to protect inferences (Clarke and Lebo 2003; Clarke and Stewart 1994; Granger and Newbold 1974; Tsay and Chung 2000). Granger and Newbold (1974) finds significant

[6]Techniques such as estimating and differencing by $d$ lose some precision with smaller samples - bias and RMSE will generally increase as the sample size decreases. The extent of the increase will depend on the complexity of the process as well as the type of estimator used; see Chapter 3 for more information. But with simple diagnostic tools like ACFs and PACFs, researchers can determine if they have properly modeled out auto-correlation. Fractional integration filters can also be approximated using several AR and MA parameters (Hosking 1981).

## Figure 4.1: An Integrated Series Pre- and Post-Differencing with ACF and PACF




*Note*: Spikes outside the solid lines on the auto-correlation and partial autocorrelation function indicate problematic correlations (not including the first spike at $t = 0$).

(.05 level) relationships about 75% of the time between two randomly generated unit-roots; Lebo, Walker, and Clarke (2000) finds similar problems with fractionally integrated data, and DeBoef and Granato (1997) do so with near-integrated data. In every case, pre-whitening solves these problems and minimizes Type I errors.

The inferential threats of auto-correlation are not debated, nor do researchers argue about whether pre-whitening works. No effort that we know of has sought to explain why under-differencing or over-differencing are not really problems. Yet, pre-whitening is often skipped. One worry is that information is lost by filtering – the “babies and bathwater” argument. Also, studying a series in differences means we cannot study long-term relationships (Banerjee et al. 1993; Beck 1992).

The desire to study both short- and long-term relationships motivates error correction methods. Classic cointegration posits equilibrium relationships so that shocks that separate series are short-lived and error correction mechanisms (ECMs) measure the rate of re-equilibration (Engle and Granger 1987). The two-step ECM approach begins with $I(1)$ variables and tests whether they are cointegrated - that is, is a linear combination of them $I(0)$? In the first step $Y_t$ is regressed on covariates and residuals are tested for stationarity. If the residuals are $I(0)$ then a second step uses the lagged residuals as the ECM in a differenced model:

$$\Delta Y_t = \beta_0 + \beta_1 \Delta X_t + \alpha_1 ECM_{t-1} + \epsilon_t \tag{4.2}$$

Both short- ($\beta_1$) and long-run ($\alpha_1$) effects are captured. The separation of $Y_{t-1}$ and $X_{t-1}$ (errors) are corrected at time $t$ as the series return to equilibrium. Many $Xs$ may affect $Y$ but few will prove to be in an equilibrium relationship.[7]

Many alternative models exist including one-step GECMs that skip testing specifically for cointegration. Single equation GECMs have quickly become the most popular ECM technique among political scientists who have seen it – correctly – as easier to implement and – incorrectly – as more flexible than competitors. Next we discuss the GECM and review it under various scenarios. We show that common practices among political scientists lead to large problems with Type I errors for independent variables and a sometimes massive tendency towards claiming error correction.

## 4.3 The General ECM - Origins and Explanation

A single equation general ECM was introduced to applied econometrics by Hendry and Anderson (1977) and Davidson et al. (1978) who each included level-form lags in differenced regressions to estimate equilibration.[8] Despite modeling individual $I(1)$ series, the relationship between them could take on a stationary equilibrium. With many economic and financial time series fitting the $I(0)/I(1)$ dichotomy – e.g. stock prices and exchange, interest, unemployment, and inflation rates – the model's assumptions were reasonable. Subsequent research by Engle and Granger (1987) formalized the concept as cointegration.

[7] Stock and Watson (2011, p.650) study one-year and three-month interest rates as an example of cointegration. Some political series may be fractionally cointegrated (e.g. Lebo and Young (2009) and Box-Steffensmeier and Tomlinson (2000)) but political data as close as Stock and Watson's are unlikely.

[8] Development of the method can be traced further back to Phillips (1954, 1957) and Sargan (1964).

In one form the bivariate ECM can be written as:

$$\Delta Y_t = \beta_0 \Delta X_t + \alpha_1 (Y_{t-1} - X_{t-1}) + \epsilon_t \tag{4.3}$$

$$\Delta X_t = \upsilon_t \tag{4.4}$$

in which $\epsilon_t$ and $\upsilon_t \sim N(0, \sigma^2)$.

In Equation (4.3) current changes of $\Delta Y_t$ can occur in response to $\Delta X_t$ or to correct disequilibrium between $Y_{t-1}$ and $X_{t-1}$. With $I(1)$ series, testing whether $\alpha_1 = 0$ is equivalent to a cointegration test such that when $\alpha_1 < 0$ cointegration exists, i.e. the series are in a long-run equilibrium (Banerjee et al. 1993). Further, with $I(1)$ data, $\Delta Y_t$ and $\Delta X_t$ are stationary and, if cointegration exists, then $Y_{t-1} - X_{t-1}$ is as well.[9] Thus, with cointegration, Equation (4.3) is *balanced* – a critical property for correct specification – and is safe from spurious regressions. Without cointegration, however, the model is unbalanced and the practitioner should set aside the estimates and choose a different specification.

Applying Equation (4.3) is somewhat cumbersome, however, since post-estimation calculations are needed to obtain the most useful parameters, including the error-correction parameter. Consequently, Bårdsen (1989) derived the reparamaterization that is now the commonly used form:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \epsilon_t \tag{4.5}$$

In Equation (4.5), both the ECM parameter ($\alpha_1$) and the short-run effects of $X_t$ are estimated directly while long-run multipliers are readily calculable. De Boef and Keele (2008) highlight these features and much of the second half of *Taking Time Seriously* is devoted to exploring

[9] Kremers et al. (1992) was the first to formally establish the relationship between the $t$-test on the disequilibrium term $Y_{t-1} - X_{t-1}$ and Engle & Granger's cointegration test-statistic; both are unit-root tests, but the GECM is a more powerful test for cointegration because it does not impose common factor restrictions. See Appendix C.1 in Supplementary Materials for further background on the links between the GECM model, the Dickey-Fuller test, and the Engle and Granger two-step cointegration method.

the potential of Equation (4.5) in political science.[10] Long after its discussion in a 1992 special cointegration issue of *Political Analysis's*, De Boef and Keele's re-introduction drew significant attention to the GECM. Equation (4.5) has since become extremely popular – perhaps the most popular time series model in political science.[11]

Yet, the GECM is consistently misused. Whereas D&K highlight the GECM when all series are stationary, researchers rarely pay close attention to the order of integration of their time series and often skip important tests. This has had serious inferential consequences. For example, political scientists have not recognized that if $Y_t$ is $I(1)$ then the GECM is effectively a cointegration test. With $Y_t$ as $I(1)$, *all* series in the model must be $I(1)$ and cointegration must be present. Otherwise, the model is unbalanced and misspecified.

Also missed is that the distribution of the ECM parameter's ($\alpha_1$) test-statistic is neither standard-normal nor dimension invariant - it shifts systematically depending on sample size as well as the number and stationarity properties of the variables (Ericsson and MacKinnon 2002; Hansen 1995; Kremers, Ericsson, and Dolado 1992). The usual practice erroneously uses a standard one-tail $t$-test for hypothesis testing resulting in over-blown findings. Seemingly, equilibrium relationships have been found everywhere.

With strictly $I(1)$ data one *can* calculate new critical values, however, and Ericsson and MacKinnon (2002) provide response surfaces for such calculations based upon the length of $T$ and the number of covariates. Table 4.1 provides our calculations of critical values for lower values of T and up to five independent variables. Yet these values are only applicable if *every* series is strictly $I(1)$. Any diversion for *any* variable in the model alters the test statistic's sampling distribution and critical values – often dramatically. Additionally, any loss of equation balance makes a cointegration test dubious so, again, if the dependent variable is $I(1)$ then the model should only include $I(1)$ independent variables.

[10] The equivalence of the ADL and the Bårdsen ECM is presented in De Boef and Keele (2008, pp.189-191).

[11] In economics the method has been largely supplanted by alternatives like the ARDL-bounds tests of Pesaran, Shin, and Smith (PSS, 2001). As explained below, the PSS test allows flexibility for regressors' orders of integration but the dependent variable must still be $I(1)$. The PSS paper has been cited over 3,000 times since publication in 2001 whereas Ericsson and MacKinnon (2002) – the source for the GECM's correct critical values – has been cited 243 times.

**Table 4.1: 5% MacKinnon Critical Values of ECM t-statistic for $I(1)$ Data†**

| $T$ | 1 IV | 2 IVs | 3 IVs | 4 IVs | 5 IVs |
|---|---|---|---|---|---|
| 35 | -3.316 | -3.613 | -3.867 | -4.082 | -4.268 |
| 40 | -3.300 | -3.598 | -3.850 | -4.066 | -4.255 |
| 45 | -3.290 | -3.587 | -3.838 | -4.055 | -4.246 |
| 50 | -3.283 | -3.578 | -3.829 | -4.047 | -4.239 |
| 55 | -3.277 | -3.570 | -3.822 | -4.040 | -4.233 |
| 60 | -3.270 | -3.565 | -3.816 | -4.035 | -4.229 |
| 65 | -3.267 | -3.560 | -3.811 | -4.030 | -4.226 |
| 70 | -3.263 | -3.556 | -3.807 | -4.027 | -4.223 |
| 75 | -3.259 | -3.552 | -3.803 | -4.023 | -4.220 |
| 80 | -3.256 | -3.549 | -3.800 | -4.021 | -4.218 |

*Note*: Critical values for each $T$ computed using response surfaces provided by Ericsson and MacKinnon (2002)
† Estimates of critical values based on model with intercept and no trend

Yet, recent research on near- and fractional integration demonstrates that true unit-roots are rare in political data (DeBoef and Granato 1997; Box-Steffensmeier and Smith 1996, 1998; Lebo, Walker, and Clarke 2000; Byers, Davidson, and Peel 2000).[12] Thus, although Table 4.1's values allow proper testing of the ECM with all $I(1)$ series, the complications of political data mean that it is rarely applicable.[13]

Below, we examine the accuracy of inferences using common practices and the GECM. When $\alpha_1$ is unknown and researchers estimate the model, how accurate are the conclusions? We study six types of dependent variables: 1) $Y_t$ as a unit-root ($d = 1$ in a $(p, d, q)$) model, 2) $Y_t$ as a bounded unit-root, 3) $Y_t$ as stationary ($d = 0$), 4) $Y_t$ as near-integrated, 5) $Y_t$ as fractionally integrated, $(0 < d < 1)$, and 6) $Y_t$ as explosive $(d > 1)$.

For each type we simulate various covariate types – $X_t$ as $I(1)$, $I(0)$, and fractionally integrated – and use 1-5 $X$s with T=60 and T=150. In all, 746 simulation exercises test the GECM's abilities to estimate error correction and make other inferences. Absent the model being used as a cointegration test for purely $I(1)$ series with proper critical values,

[12] There are fields in political science where one might find more unit root data, such as in international political economy, e.g. Bernhard and Leblang (2002), Leblang and Bernhard (2006).

[13] One caveat worth noting is that the model's estimates prove reliable if we specifically *generate* our data so that $\alpha_1 < 0$ and the ECM is used as a cointegration test. That is, in these specialized circumstances the ECM model does not make Type II errors, even with near-integrated data, and properly finds that cointegration is present (De Boef and Granato 1999; Kremers, Ericsson, and Dolado 1992; Hansen 1995; Zivot 2000). But in actual practice, we do not choose the value of $\alpha_1$ and, as Zivot (2000) shows with a single-equation ECM, $\alpha_1$ is far better at *detecting* cointegration than it is at measuring the *strength* of re-equilibration (i.e. as an ECM).

the GECM poorly measures the error correction parameter and finds it to be significant far too often – Type I errors occur at alarming rates. We begin with the case of $Y_t$ as $I(1)$.

## 4.4 The General ECM Under Six Scenarios

### 4.4.1 Case 1: The Dependent Variable is a Unit-Root, I(1)

What problems arise with the GECM and a unit-root $Y_t$? This scenario *can* work but there are four points that political scientists have missed: 1) $X_t$ must also be a unit-root to ensure equation balance (Banerjee et al. 1993, pp. 164-8), 2) $Y_t$ and $X_t$ must be cointegrated with each other (Banerjee, Dolado, and Mestre 1998; Ericsson and MacKinnon 2002), 3) non-standard critical values must be used (Ericsson and MacKinnon 2002) to evaluate cointegration, and 4) the ECM parameter is biased downwards as $X$s are added to the model.

The value of the first two points is more easily seen in Equation (4.3)'s version of the GECM: $\Delta Y_t = \beta_0 \Delta X_t + \alpha_1(Y_{t-1} - X_{t-1}) + \epsilon_t$. Here, if $Y_t$ and $X_t$ are unit-roots, then $\Delta Y_t$ and $\Delta X_t$ are stationary. And, if they are cointegrated, then $(Y_{t-1} - X_{t-1})$ is stationary as well.[14] OLS can be used here but it has been shown that the hypothesis test on $\alpha_1$ in Equation (4.5) is non-standard (Ericsson and MacKinnon 2002). The correct values can be found in common sources like Enders (2010, p. 493) or calculated for specific lengths of $T$ as we do in Table 4.1.

Unfortunately, practitioners are neither paying proper attention to their series' stationarity nor testing for cointegration.[15] In particular, one often-quoted point in *Taking Time Seriously* seems universally misinterpreted: "Alternately, as the ECM is useful for stationary and integrated data alike, analysts need not enter debates about unit roots and cointegration to discuss long-run equilibria and rates of reequilibration" (p.199). This statement must be carefully interpreted: the GECM can work for *all* stationary *or all* integrated data. If all

[14] D&K allude to this in their Footnote 11 guideline: "When using an ECM with integrated data, analysts must ensure that all terms on the right-hand side of the equation are stationary" (p. 190).

[15] The Granger Representation Theorem (Engle and Granger 1987) posits that if cointegration exists then Granger causality must be present in at least one direction. Failing to explicitly test for cointegration or Granger causality can lead to erroneous causal claims.

data are stationary then one does not have to test for cointegration to estimate the GECM. But researchers have assumed this means that they need *never* worry about stationarity or cointegration – that plugging a mix of data into the GECM is fine and that the model can always work with or without cointegration.[16] But if the right-hand-side variables are not truly $I(1)$ the equation is not balanced and, even if balanced, if the series are not truly cointegrated then problems arise.

**Table 4.2: Results of GECM Model with $I(1)$ Data, $T$=60[†]**

| Model | 1 IV | 2 IVs | 3 IVs | 4 IVs | 5 IVs |
|---|---|---|---|---|---|
| % ECM Significant – one tail $t$-distribution | 56.8 | 64.9 | 72.0 | 76.7 | 80.2 |
| % ECM Significant – MacKinnon Values | 5.3 | 5.5 | 5.3 | 5.0 | 5.1 |
| Mean of $\alpha_1$ | −0.12 | −0.15 | −0.18 | −0.21 | −0.24 |
| Mean of $\alpha_1^*$ | −0.32 | −0.37 | −0.42 | −0.47 | −0.52 |
| % ECM & $\geq 1\Delta X_t$ Significant | 3.8 | 8.6 | 13.3 | 18.4 | 24.3 |
| % ECM & $\geq 1X_{t-1}$ Significant | 14.1 | 25.0 | 34.2 | 41.1 | 47.1 |

[†]*Note*: Percentage results in each cell based on 10,000 simulations per ECM model with null hypothesis of no error correction. Finding of significant ECM and significant $X_{t-1}$ indicates presence of error correction.
Mean of $\alpha_1^*$ when $t-$statistic exceeds MacKinnon critical value, i.e., cointegration is present.
$\Delta X_t$ and $X_{t-1}$ significance (*p≤ 0.05, two-tail test). ECM significance (*p≤ 0.05, one-tail test)
DV and all IVs are unit-roots ($I(1)$). $T = 60$.
ECM Model: $\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \epsilon_t$.

In the absence of cointegration having unit-roots $Y_{t-1}$ and $X_{t-1}$ on the right-hand-side causes problems for hypothesis testing their coefficients. In common practice, a significant $\alpha_1$ is used to confirm cointegration and its value is given as the speed with which a long-run equilibrium between $Y_t$ and $X_t$ is restored following a shock that separates them. Since $\alpha_1$ is also used to calculate long-run multipliers for the covariates, bias and/or a Type I error gives a lot of false information.[17]

Table 4.2 shows GECM simulations using all $I(1)$ data (T=60). *All* political science applications of the GECM have used the $t$-distribution for $\alpha_1$'s hypothesis test. The first row shows Type I error rates when this mistake is made.[18] With a single $I(1)$ $X$, 56.8 percent

[16]For example, Volscho and Kelly (2012) say: "...ECMs accommodate stationary and integrated variables, which is useful because our analysis has a mix of both data types. In summary, the ECM is a very general model that is easy to implement and estimate, does not impose assumptions about cointegration, and can be applied to both stationary and nonstationary data (Banerjee et al. 1993; De Boef and Granato 1999; De Boef and Keele 2008)." See Appendix C.6 in the Supplementary Materials for other misinterpretations.

[17]For an example of this in action see Table C.16 in Supplementary Materials

[18]Simulations were run in RATS 8.0. Series were specified to have no relationship to each other. $I(1)$

of ECM coefficients are significant at the .05 level using a one-tailed $t$-test.[19] With correctly calculated critical values for $\alpha_1$ Type I errors on both the ECM parameter and the $X$s can be minimized. The second row shows just this – correct MacKinnon values put Type I errors where they should be for $\alpha_1$, around 5% with 1-5 covariates.

But using incorrect critical values makes one likely to falsely conclude that cointegration exists and that the equation is balanced. With $X_{t-1}$ as $I(1)$, unresolved auto-correlation leads to bias in the estimation of $\beta_1$, shrunken standard errors, and higher Type I error rates – $\beta_1$ is significant (.05-level) 14.1% of the time in a bivariate model. Additional $I(1)$ $X$s increase these problems and Type I errors creep higher. The row for $\Delta X$ shows that even though appropriately filtered, auto-correlation elsewhere gives us Type I errors on $\beta_0$ far too often. A model can begin to look quite good with both a significant ECM parameter and a long-run $X$. Using a standard distribution can lead one there far too easily.

Another issue is that additional $I(1)$ $X$s move the distribution of $\alpha_1$ further from zero. Even when the proper MacKinnon values indicate the presence of cointegration, the expected value of $\alpha_1$ decreases with each additional covariate. Ease of interpretability is a key selling point of the GECM but this is one of several complications ignored by practitioners.

In all, Table 4.2 shows problems with the GECM with unrelated unit-root series. Testing for cointegration and relying on MacKinnon values can address spuriousness but we might still misstate the rate of error correction in a multivariate model. And, in practice, this ideal scenario is exceedingly rare – not only would the data need to be truly $I(1)$ (rare enough) and actually cointegrated (rarer still), they would also need to be unbounded.

series were generated as $Y_t = Y_{t-1} + \epsilon_t$. $I(0)$ series were generated as $Y_t = \epsilon_t$. $I(d)$ series were generated to a specified order of $d$ using the RATS "arfsim" package. Near-integrated series were generated by specifying $\rho \sim (0.9 - 1.0)$ in $Y_t = \rho Y_{t-1} + \epsilon_t$. Bounded series were created following Nicolau (2002). A brief explanation of related versus unrelated series can be found in Appendix C.2.1. Details on procedures for bounded series are in Appendix C.2.2 of the Supplementary Materials. Complete tables of results are in Appendix C.7.1 (near-integration, one table each for T=60 with 1,2,3,4, and 5 IVs and one table each for T=150 with 1,2,3,4, and 5 IVs) and Appendix C.8.3 (fractional integration, one table each for T=60 with 1,2,3,4, and 5 IVs and one table each for T=150 with 1,2,3,4, and 5 IVs).

[19] Of the GECM articles we surveyed all used a standard $t$-test and averaged roughly four IVs per regression: Faricy (2011): 4, Jennings and John (2009): 2 plus interactions, Kayser (2009): 6-7 plus interactions (with a larger N), Kelly and Enns (2010): 3-4, Kono (2008): 9 (with a very large N), Ramirez (2009): 9, Rickard (2012): 5, Ura (2014): 4, Ura and Wohlfarth (2010): 4, and Ura and Ellis (2012): 5.

### 4.4.2 Case 2: The Dependent Variable is a *Bounded* Unit-Root

Time series are often *bounded* (sometimes called "limited" or "regulated") between an upper and lower limit.[20] Approval of leaders and parties, indices of economic evaluations like the ICS, Policy Mood, and percentages of political and economic phenomenon all fluctuate between upper and lower limits. Little is known about the consequences of bounded series – especially regarding their effects on error correction models.

A bounded time series is an odd case. For one, it cannot meet the textbook definition of a unit-root since its asymptotic properties include mean reversion and finite variance (Williams 1992). Series like presidential approval and Stimson's *Mood* (1991) cannot have infinite variance or break their bounds and will oscillate around a long-term mean. Still, over long periods a bounded series can exhibit the perfect-memory of integrated data or the long-memory of near- and fractionally integrated data. Indeed, recent work has shown that boundedness is compatible with unit-root properties and the idea of a *bounded unit-root*, $BI(1)$, exists in econometrics (Cavaliere 2005; Granger 2010).

Problems occur when the asymptotic properties of a bounded series are used to dismiss the possibility of a unit-root and treat a series as stationary.[21] Bounded series also over-reject the Dickey-Fuller test (1979), increasing the tendency to treat them as stationary (Cavaliere and Xu 2014). If one treats stationarity as a yes/no question and models a bounded series as simply stationary then auto-correlation remains and inferences can be threatened.

Here we are interested in the problems bounded series present to the GECM. Bias comes about simply: if $Y_{t-1}$ was near the series' bounds, $\Delta Y_t$ will have a strong tendency in the opposite direction. High (low) levels at $Y_{t-1}$ will be followed by negative (positive) changes at $t$, pushing $\alpha_1$ further into negative territory. This may tell us something about $Y_t$ but it

[20] Series bounded on one end are also problematic (Cavaliere 2005; Granger 2010).

[21] Durr (1992) pushes analysts to focus on the data in hand and rely on theory for whether series might be integrated or not. Researchers who are unaware or dismissive of fractional integration techniques and who also side with Williams' (1992) argument will assume bounded series like Congressional Approval cannot be $I(1)$ and will routinely treat them as $I(0)$. Keele (p.252 2007) is pragmatic on this point and tests for stationarity and fractional integration while noting that asymptotically the series cannot be integrated.

misinforms our inferences of re-equilibration between $Y_t$ and $X_t$. Movement away from the bounds and towards the long-term mean will move $\alpha_1$ but this is bias, *not* error correction.

**Table 4.3: GECM Estimation Problems with a Bounded Dependent Variable**

| | Bivariate | | | | | |
|---|---|---|---|---|---|---|
| **DV Bounds** | (1 – 100) | (49 – 71) | (1 – 100) | (49 – 71) | (1 – 100) | (49 – 71) |
| **Model Type** | **T=60, $\sigma$=1** | | **T=60, $\sigma$=2** | | **T=60, $\sigma$=3** | |
| % ECM Significant | 80.0 | 65.3 | 61.4 | 75.3 | 62.6 | 86.6 |
| % ECM Significant* | 6.0 | 7.7 | 6.1 | 9.9 | 6.1 | 12.4 |
| Mean of $\alpha_1$ | -0.12 | -0.14 | -0.13 | -0.17 | -0.13 | -0.20 |
| | **T=100, $\sigma$=1** | | **T=100, $\sigma$=2** | | **T=100, $\sigma$=3** | |
| % ECM Significant | 61.2 | 70.0 | 63.1 | 87.0 | 64.8 | 98.3 |
| % ECM Significant* | 7.0 | 10.5 | 7.6 | 14.8 | 8.1 | 23.1 |
| Mean of $\alpha_1$ | -0.08 | -0.10 | -0.08 | -0.13 | -0.09 | -0.16 |
| | **T=150, $\sigma$=1** | | **T=150, $\sigma$=2** | | **T=150, $\sigma$=3** | |
| % ECM Significant | 61.5 | 74.3 | 64.3 | 97.3 | 67.0 | 99.9 |
| % ECM Significant* | 6.5 | 12.2 | 7.6 | 20.5 | 8.4 | 45.6 |
| Mean of $\alpha_1$ | -0.05 | -0.07 | -0.06 | -0.10 | -0.06 | -0.14 |
| | **Multivariate (2 IVs)** | | | | | |
| **DV Bounds** | (1 – 100) | (49 – 71) | (1 – 100) | (49 – 71) | (1 – 100) | (49 – 71) |
| **Model Type** | **T=60, $\sigma$=1** | | **T=60, $\sigma$=2** | | **T=60, $\sigma$=3** | |
| % ECM Significant | 67.7 | 73.3 | 68.8 | 82.4 | 70.0 | 90.1 |
| % ECM Significant* | 6.2 | 7.8 | 6.4 | 10.6 | 6.6 | 12.6 |
| Mean of $\alpha_1$ | -0.16 | -0.18 | -0.16 | -0.21 | -0.17 | -0.24 |
| | **T=100, $\sigma$=1** | | **T=100, $\sigma$=2** | | **T=100, $\sigma$=3** | |
| % ECM Significant | 69.7 | 78.0 | 71.6 | 91.0 | 73.3 | 98.6 |
| % ECM Significant* | 6.2 | 9.6 | 6.6 | 13.9 | 7.0 | 20.8 |
| Mean of $\alpha_1$ | -0.10 | -0.12 | -0.10 | -0.15 | -0.11 | -0.18 |
| | **T=150, $\sigma$=1** | | **T=150, $\sigma$=2** | | **T=150, $\sigma$=3** | |
| % ECM Significant | 69.5 | 80.5 | 72.0 | 97.6 | 74.4 | 100 |
| % ECM Significant* | 6.8 | 11.5 | 7.7 | 18.5 | 8.4 | 38.5 |
| Mean of $\alpha_1$ | -0.07 | -0.09 | -0.07 | -0.12 | -0.07 | -0.15 |

*Note*: Percentage results are of 10,000 simulations for each specified model and $T$. All independent variables are integrated $I(1)$. *Significant using MacKinnon Values.

This is shown in Table 4.3. Following Nicolau (2002) we simulate unit-roots that bounce back as they near upper or lower thresholds.[22] We explored the effects of tighter bounds, longer series, series with more variance, and models with additional independent variables. In all, $BI(1)$ series are problematic for estimating the ECM coefficient: the average value

[22]The process is generated by: $X_t = X_{t-1} + e^k(e^{-\alpha_1(X_{t-1}-\tau)} - e^{\alpha_2(X_{t-1}-\tau)}) + \epsilon_t$ where $\alpha_1 \geq 0$; $\alpha_2 \geq 0$; $k < 0$; and $\epsilon_t$ is assumed to be *i.i.d.* with a mean of 0 and variance of 1. Appendix C.2.2 in Supplementary Materials provides additional detail and graphics on how these series were generated.

jumps and Type I errors rise. Also, the more the bounds come into play – tighter bounds, more variance, or a longer sample – the more Type I errors increase.

Next, we create series meant to mimic derivatives of Stimson's Mood (e.g. Kelly and Enns 2010) – $BI(1)$ with $\sigma = 3$ and bounds of 49 and 71 – and regress them on an $I(1)$ $X$ in the GECM. With $T = 150$ the ECM parameter is significant 99.9% of the time using a one-tail $t$-distribution and 45.6% of the time with MacKinnon values.[23] Boundedness does not seem to affect estimation of $\beta_1$ or $\beta_0$ but, if we use $\alpha_1$ to discuss equilibrium relationships or to calculate long-run multipliers, we will misstate the relationships between the variables. Note that Table 4.3 represents the GECM's *best* chance with a bounded $Y$. A bounded and fractionally integrated DV will be even more problematic than the $BI(1)$ series tested here. And, if $X$ were to deviate from $I(1)$, further problems would occur.

This adds an important caveat to the GECM model. Even if we find series that are strictly unit-roots and we use MacKinnon CVs, mistakes are still rampant if our dependent variable is one of the vast majority of political times series that is bounded. Just how prevalent are bounded series? For instance, we chose our five paper replications based on their journal prominence but, as it happens, the papers use a total of 13 series as dependent variables and *all* thirteen are bounded.

### 4.4.3 Case 3: The Dependent Variable and all Independent Variables are Stationary

Next we review the case where all series are stationary, precisely the type of data De Boef and Keele (2008) had in mind; their points that researchers may investigate long-run dynamics without worrying about cointegration is meant to apply to stationary data. Further, the use of a standard $t$-test for the error correction hypothesis is based on the assumption of stationarity. However, without adjusting how the GECM is interpreted with stationary data,

[23] Transformation of a bounded dependent variable does not seem to solve the Type I error problem. We transformed our series into log odds as $Y_t^* = log\frac{Y_t}{1-Y_t}$ which creates a series with an unbounded error process. But the Type I error rate actually worsens – with T=150, bounds of 49 to 71, and $\sigma$=3, Type I errors with MacKinnon values increase approximately 2%.

a number of problems remain.

First, a stationary dependent variable can usually guarantee a significant ECM parameter. Second, in many instances, neither the size of the ECM coefficient nor the strength of its significance implies a close relationship between the dependent variable and other series in the model. As a demonstration, Table 4.4 shows the simulated results of a bivariate GECM in which the DV is stationary but with varying degrees of auto-regression, i.e. ($\rho$,0,0), and the IV is a white noise process, i.e. (0,0,0).

**Table 4.4: ECM Significance and Coefficient Size by degree of DV Autoregression†**

| $\rho$ | 0 | .1 | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 |
|---|---|---|---|---|---|---|---|---|---|---|
| % ECM Significant | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 97.7 | 79.1 |
| Mean of $\alpha_1$ | -1.03 | -0.92 | -0.83 | -0.73 | -0.64 | -0.54 | -0.45 | -0.36 | -0.26 | -0.17 |

† *Note*: Cell entries are the result of 10,000 simulations for each bivariate model with the DV specified at $\rho$. IV integrated $I(0)$. ECM significance (*p$\leq$ 0.05, one-tail test).

The prevalence of significant ECMs indicates that we should rethink the meaning of the $\alpha_1$ coefficient when data are stationary. We are seeing reequilibration here, but it is the natural reversion of the dependent variable to its mean. That is, $\alpha_1$ is capturing the speed with which $Y$ returns to *its own* equilibrium (its mean) following a shock or an effect from $X$. What we see in Table 4.4 is not evidence of error correction – it is simply confirmation that $Y$ is stationary.[24] When working with stationary data, the larger the absolute value of the ECM, the *less* of a long-run relationship between $X$ and $Y$. If $Y$ has no long-run tendencies except mean reversion, it is an odd choice to be looking for its long-run relationship with $X$.

This is problematic when common practice is to use $\alpha_1$ to speak to error correction and re-equilibration *between* variables. This occurs frequently in practical applications – a statistically significant and substantively large $\alpha_1$ coefficient is claimed to indicate strong re-equilibrating behavior between $Y_t$ and some independent variables with weaker connections.[25] Using the raw ECM value to compute long-run multipliers would further solidify a

[24] This is not surprising given that all stationary series are in equilibrium, as recognized by De Boef and Keele (2008, p.189) and Banerjee et al. (1993, p.4).

[25] Casillas, Enns, and Wohlfarth (2011) investigate public opinion's effect on salient Supreme Court deci-

researcher's faith in the (falsely strong) relationship.

Problems extend beyond mistaken inferences regarding $\alpha_1$. To explain, we return to the linear transformation between the ADL and GECM. Despite the algebraic equivalence of the two models, as D&K note, "differing quantities are directly estimated in each model" (De Boef and Keele 2008, p.195). The autoregressive distributed lag (ADL) model is specified as:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t \tag{4.6}$$

And the Bårdsen GECM transforms this into:

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \epsilon_t \tag{4.7}$$

D&K (p.190) note the equivalence between parameters in the ADL and the GECM, specificially: $\alpha_1^* = (\alpha_1 - 1)$, $\beta_0^* = \beta_0$ and $\beta_1^* = \beta_0 + \beta_1$. Here we are interested in the long run parameter, $\beta_1^*$. It is the sum of two quantities from the ADL ($\beta_0 + \beta_1$) but, critically, its standard error is not.

For example, we estimate bivariate ADL and GECM models using data from Durr, Martin, and Wolbrecht (2000): the DV is *Supreme Court Support* and the IV is *Congressional Approval*, both of which are found to be stationary by the authors (fn.13, p.772). Our results are in Table 4.5.

Equation (4.6)'s ADL has two dynamic parameters, $\alpha_1$ and $\beta_1$. The short-term effect is captured as $\beta_0$ and the long-run effect can be calculated as $(\beta_0 + \beta_1)/(1 - \alpha_1)$. What we see in the first column of Table 4.5 is that the dynamic parameters fail to reject their null hypotheses – neither the lag of the dependent variable ($\alpha_1$=0.27, s.e.=0.16) nor the lagged

sions. Finding an ECM of -1.27 (s.e.=0.15) they say: "The magnitude of the error correction rate in this model suggests that, following just one term, the Court's behavior almost completely adjusts to changes in ideology and social forces at term *t*." And, (footnoted): "Because the error correction rate indicates the proportion of the long-term effect that occurs in each subsequent time period, an absolute value greater than 1 seems surprising." We find *Salient Cases* to be stationary ($d = 0.3$ (s.e.=0.08); DF=3.98$^\star$) in level-form which explains how the ECM value is possible. A second example, Jennings and John (2009), uses agenda items in Queen's Speeches 1960-2001 – close to stationary – as dependent variables and estimates many ECMs below -1.

**Table 4.5: The ADL vs. the GECM: Effects of Congressional Approval on Supreme Court Approval†**

| **Model** | ADL | GECM |
|---|---|---|
| Supreme Court Approval$_{t-1}$ | 0.27 (0.16) | |
| Error Correction$_{\text{(Supreme Court Approval}_{t-1})}$ | | -0.73* (0.16) |
| Congressional Approval | 0.49* (0.22) | |
| Congressional Approval$_{t-1}$ | -0.05 (0.24) | 0.43* (0.15) |
| $\Delta$ Congressional Approval | | 0.49* (0.22) |
| Constant | 41.83* (12.48) | 41.83* (12.48) |
| N | 41 | 41 |
| $R^2$ | .41 | .43 |
| Adjusted $R^2$ | .36 | .39 |

† *Note*: Data comes from Durr, Martin, and Wolbrecht (2000). ADL Model: DV *Supreme Court Approval* represents the authors' semi-annual measure of approval for the Court. ECM Model: DV $\Delta$ *Supreme Court Approval* represents changes to this measure of approval. Significance of ECM (*$p < .05$, one-tail test) and coeffcients (*$p < .05$, two-tail test)

IV ($\beta_1$=-0.05, s.e.=0.24) are significant. In short, there is no support for the use of a dynamic model.

But looking at the GECM, we see that the error correction parameter – $\alpha_1^*$ calculated as the ADL's $\alpha_1$ minus 1 – *is* significant as we would expect from Table 4.4. Additionally, because $\beta_1^*$ consists of the ADL's $\beta_0+\beta_1$ while the standard error is not additive, the long-run effect of $Congressional Approval$ is now significant at the $p < .05$ level. Whereas the results of the ADL conclude that we should estimate a static regression, the significant parameters in the GECM encourage the researcher in the opposite direction. Given publication bias towards significant findings, the GECM is again attractive. The ADL and the GECM may be algebraically equivalent, but the reorganization of parameters is not benign.

### 4.4.4 Case 4: The Dependent Variable is Strongly Autoregressive / Near-Intergrated

D&K argue that "the only situation where one would strongly prefer the ECM [as opposed to the ADL] is if the data are strongly autoregressive" and "[because the] variables are parameterized in terms of changes, helping us to avoid spurious findings if the stationarity of the series is in question due to strongly autoregressive or near-integrated data."(De Boef and Keele 2008, p.195). We find, however, that the GECM runs into serious problems here as well. In particular, rejection rates on the long-run parameter ($\beta_1^*$) reach unacceptable levels with strongly autoregressive, or near-integrated series.

De Boef (2001) and De Boef and Granato (1999) discuss the GECM in the case of near-integrated data and find it appropriate. Additionally, DeBoef and Granato (1997) investigate rates of spurious regressions with near-integrated data and find that the ADL solves the spurious regression problem. However, this is not the case for the GECM – spurious regressions appear too frequently with either near-integrated or other stationary data. With the DGP of $Y_t = \rho Y_{t-1} + \epsilon_t$, we find elevated rates of Type I errors on $\beta_1^*$ far below the range of $\rho$ that is typically considered near-integrated, i.e. 0.90 to 0.99.

Our findings are different from the previously mentioned studies for a number of reasons. First, the investigation of the power of the GECM in both De Boef (2001) and De Boef and Granato (1999) use models in which the data are pre-specified to be (near-)cointegrated. As we note above, the GECM is far more powerful at detecting cointegration (avoiding Type II errors) than it is at estimating its presence (avoiding Type I errors)(Zivot 2000). Second, as noted, the ADL and GECM do not estimate the same parameters and as a result, they produce different rejection rates. Whereas the ADL is suitable for near-integrated series, fitting the exact same data in a GECM produces spurious regressions. These problems are persistent, even with dependent variables that are clearly stationary.

To investigate the disparity in rejection rates between the ADL and the GECM we generate series with varying degrees of autoregression. The upper section of Table 4.6 contains the rejection rates of the null hypothesis that $\beta_1$ of the ADL model is equal to zero.[26] The lower section contains the rejection rates for the null hypothesis with respect to the GECM: that both the error correction parameter $\alpha_1^*$ and the long-run parameter $\beta_1^*$ of the GECM model are jointly equal to zero.[27]

Our findings for the ADL match those of DeBoef and Granato (1997) who find that the

[26] This is a recreation of Table 4 of DeBoef and Granato (1997), which tested the rejection rate of $\beta_0$. In our case, we are interested in the rejection rates of $\beta_1$, the coefficient on $X_{t-1}$ in each model. Additionally, we have expanded our criteria to include series with lesser degrees of autoregression. Neither of these changes affect the ultimate conclusion with respect to the ADL: the model is proper for use with near-integrated data.

[27] To reject the null hypothesis both $\alpha_1^*$ and $\beta_1^*$ must be significant.

**Table 4.6: Rejection Rates for the Null Hypothesis of ADL ($\beta_1$=0) and GECM ($\alpha_1^*$=0 & $\beta_1^*$=0) with Near-Integrated Data**

| | $\rho_x$ | | | | | |
|---|---|---|---|---|---|---|
| **ADL Model** | | | | | | |
| $\rho_y$ | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.99 |
| 0.75 | 5.7 | 5.8 | 5.6 | 5.6 | 5.9 | 6.0 |
| 0.80 | 5.7 | 5.9 | 5.7 | 5.7 | 5.9 | 6.0 |
| 0.85 | 5.7 | 5.9 | 5.8 | 5.8 | 6.0 | 6.0 |
| 0.90 | 5.7 | 5.9 | 5.8 | 5.8 | 6.0 | 6.1 |
| 0.95 | 5.7 | 5.9 | 6.0 | 5.8 | 6.1 | 6.3 |
| 0.99 | 6.0 | 5.9 | 6.0 | 5.8 | 5.9 | 6.2 |
| **GECM Model** | | | | | | |
| $\rho_y$ | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.99 |
| 0.75 | 7.6 | 8.1 | 8.3 | 8.8 | 9.1 | 9.3 |
| 0.80 | 7.9 | 8.4 | 8.9 | 9.4 | 9.9 | 10.2 |
| 0.85 | 8.2 | 8.8 | 9.2 | 10.2 | 10.7 | 11.4 |
| 0.90 | 8.1 | 8.7 | 9.6 | 10.8 | 11.8 | 13.0 |
| 0.95 | 7.3 | 8.0 | 9.2 | 10.6 | 12.3 | 14.3 |
| 0.99 | 6.4 | 7.1 | 8.1 | 9.7 | 11.4 | 13.7 |

*Note*: Percentage results based on 10,000 simulations.T=60. Each model contains one IV.
ADL model: $Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t$
GECM: $\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \epsilon_t$. Significance of $\alpha_1^*$ ($^\star p < .05$, one-tail test)
Significance of $\beta_1$, $\beta_1^*$ ($^\star p < .05$, two-tail test)

model has acceptable spurious regression rates with near-integrated data. But we also find that this does not translate for the same data in the GECM. Precisely the type of data D&K claim we should favor with the GECM proves susceptible to Type I errors. Excessive rejection rates are also found when data are less strongly autoregressive, e.g. $\rho = .75$.[28]

There are a number of important conclusions to draw here. First, despite being algebraically equivalent, the GECM and the ADL do not produce the same results. The GECM does not perform as well with strongly autoregressive, or near-integrated data. In terms of model preference, the ADL should be preferred over the GECM with either stationary, strongly autoregressive, or near-integrated data. With a stationary DV the error correction parameter is practically guaranteed significance and the substantive meaning of its coefficient is inscrutable.

[28]Since these results are of a joint hypothesis test, two parameters must be significant. The rejection rates of just the $\beta_1^*$ are higher with near-integrated data. Additionally, while these results are based on a bivariate model, in practice the GECM is fit with more IVs. The probability of finding a significant $\beta_1^*$ parameter increases with the inclusion of additional IVs. Further, the nature of the ECM with stationary data means the significance of the $\alpha_1^*$ parameter is not a sufficient gatekeeper.

Second, although using MacKinnon CVs with near-integrated data would limit the rate of spurious regressions (see: Section C.7.1 of Supplement), this cannot be recommended since the decision of when to switch to MacKinnon values with stationary data will be an arbitrary one. The MacKinnon values are recommended based on the unit-root or not distinction. Researchers cannot simultaneously argue that data are stationary while using unit-root critical values. Spurious regressions appear, for example, when $\rho = 0.75$ and the correct critical values in that case are derived from neither the MacKinnon nor the normal distribution – they are unique to the particular data being used. Of course, series that are not unit-roots may not be simply autoregressive. Many political time series have been shown to be fractionally integrated and these prove even more troublesome for the GECM.

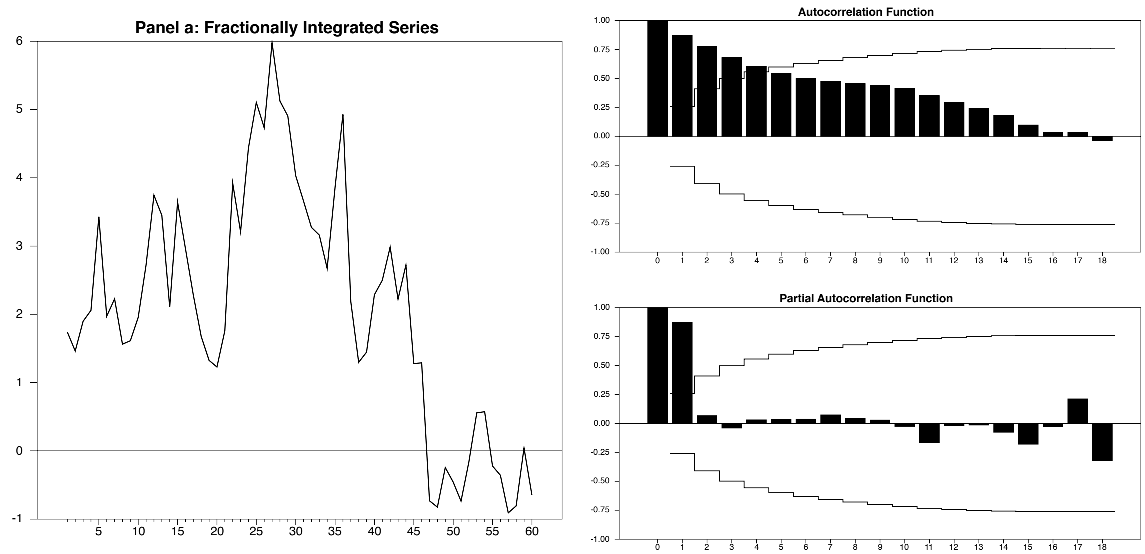
### 4.4.5 Case 5: The Dependent Variable is Fractionally Integrated, (0,$d$,0), and 0<$d$<1

Fractional integration (FI) is common among political time series (Box-Steffensmeier and Smith 1996, 1998; Byers, Davidson, and Peel 2000; Lebo, Walker, and Clarke 2000; Lebo and Moore 2003) so finding problems here should make political researchers especially wary of the GECM.[29] Figure 4.2 shows a FI series and its ACF and PACF. Auto-correlations persist for long periods and, without pre-whitening, cause problems in regression analyses (Lebo, Walker, and Clarke 2000; Tsay and Chung 2000). What are the consequences of using fractionally integrated series in the GECM?

Recall that the GECM model is $\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \epsilon_t$. We showed that when $Y_t$ is a unit-root the left-hand-side of the equation is properly treated and $\Delta Y_t$ is clear of auto-correlation. We also showed that if $Y_t$ begins as stationary, applying the differencing operator $\Delta$ will *over*-difference it. In the FI case, with $0 < d < 1$, the GECM will again over-difference the series to varying degrees depending on the initial value of $d$.

[29] Aggregating individuals with heterogeneous AR processes creates a long-memory series with a mix of stationary and non-stationary attributes (Granger 1980). Many political time series are created in this way. Pickup (2009) is the only attempt we know of that argues that FI does not apply to aggregate political data (but see Young and Lebo (2009) for a rebuttal).

**Figure 4.2: A Fractionally Integrated Series with ACF and PACF**




*Note*: Significant spikes on the ACF and PACF are problematic correlations.

For example, a series that in levels is $d$=0.7 will become $d$=-0.3.[30]

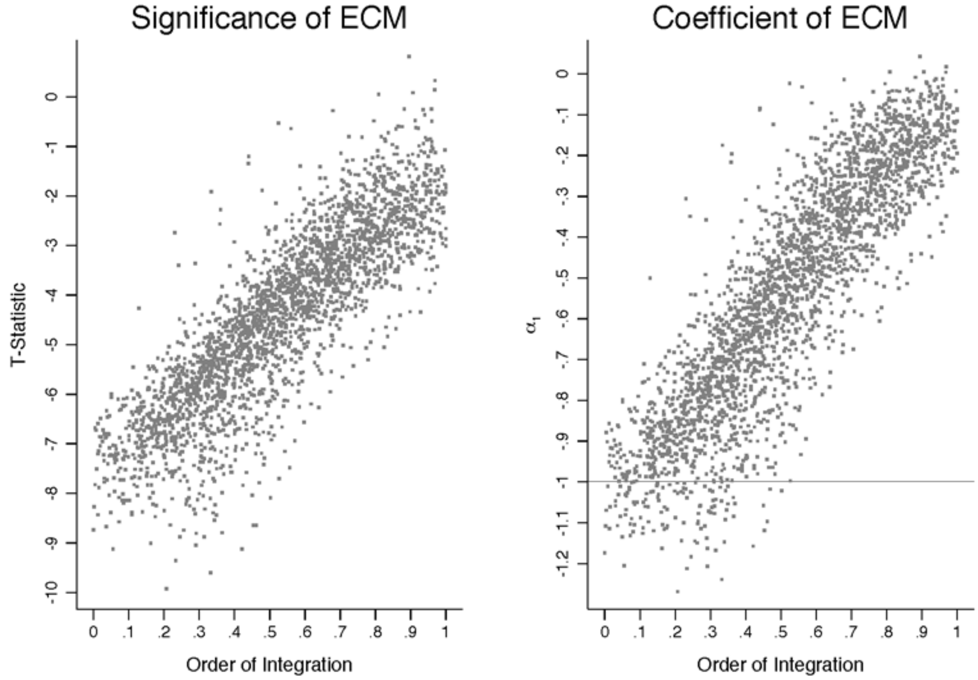
Over-differencing builds a moving average process into the transformed series (Dickinson and Lebo 2007). For example, if we (1) begin with a FI series: $(1-L)^d Y_t = \epsilon_t$ where $\epsilon_t \sim N(0, \sigma^2)$, (2) isolate $Y_t$, and (3) apply first differencing, we get: $\Delta Y_t = \frac{\epsilon_t}{(1-L)^d} - \frac{(\epsilon_{t-1})}{(1-L)^d}$. If $d = 0$ this leaves a non-invertible moving average process as discussed above. With smaller degrees of over-differencing, an MA process of $\frac{1}{(1-L)^d}$ becomes a part of $\Delta Y_t$. Over-differencing by just 0.4, for example, gives a 99.4% chance of creating a significant MA parameter (Dickinson and Lebo 2007).

Figure 4.3 shows the consequences of first-differencing FI data in the GECM. The X-axis shows $Y_t$'s original order of integration. On the left panel's Y-axis is the ECM's $t$-statistic and on the right panel's Y-axis is the ECM coefficient.[31] By chance, the probability of getting

[30] DeBoef and Granato (1997) discuss the near-integrated case and find that an over-differenced series will contain negative auto-correlation in its error term and will produce non-credible inferences.

[31] Here $T = 60$ and $X$ is $I(1)$. See Appendix C.8.3 and Tables C.48-C.57 in Supplementary Materials for detailed results where we vary sample size and the order of integration for both $X$ and $Y$. See Figure C.9 in Supplementary Materials for another view of bias in the ECM in this case.

**Figure 4.3: Significance and Size of ECM Parameter Based on Y's Order of Fractional Integration**



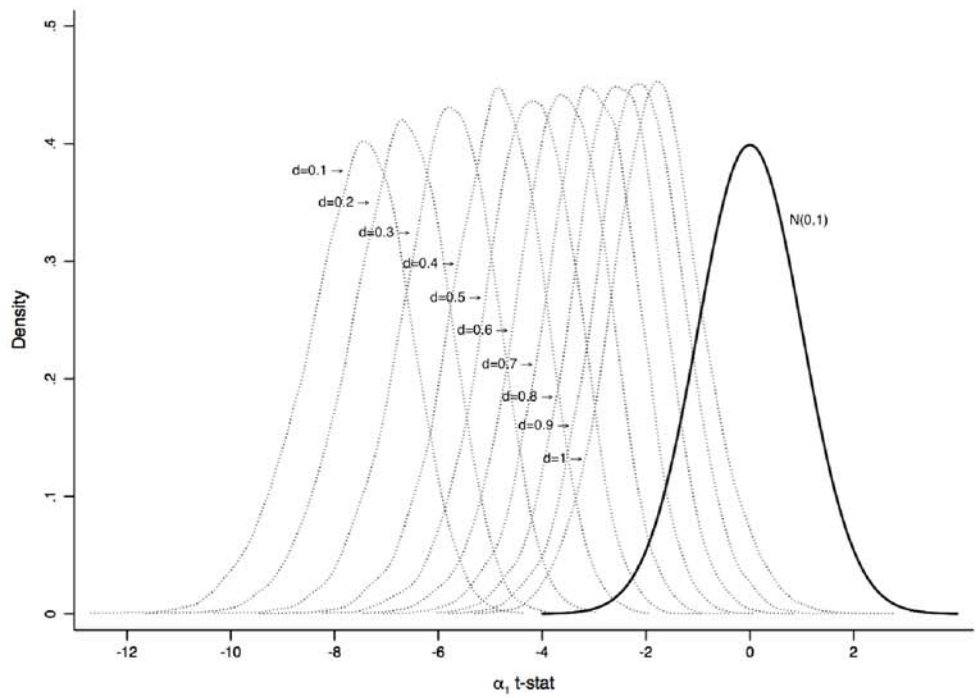

*Note*: Plots based on bivariate regressions with $T = 60$. Horizontal line on the right indicates the theoretical limit of an ECM.

a $t$-statistic of -4 or less is roughly 1 in 10,000 and -7 or less is roughly 1 in 100,000,000. But such $t$ values are routine when using unrelated FI data in the GECM. As $Y$ gets further into stationary territory, the coefficient for $Y_{t-1}$ declines and Type I errors become the norm.

The right-hand panel of Figure 4.3 shows the size of the ECM coefficient. Read as the (negative of the) proportion of the gap that is re-equilibrated in the next period, it should never be below -1. Here we see that it often is – a fact that might lead a researcher to conclude they have found an especially close relationship between $Y_t$ and the $X$s.

For a different view, Figure 4.4 shows the density plots of the ECM parameter's $t$-statistic as $Y_t$'s $d$ value decreases from one. The sampling distribution for the test-statistic moves along with the order of integration of $Y_t$. Even if we could pin down the correct critical values, the meaning of the ECM coefficient has been lost. We are basically sliding toward

**Figure 4.4: Sampling Distribution of the ECM $t$-statistic by Order of Fractional Integration**




*Note*: Density plots of $t$-distributions from simulations of bivariate GECMs when both DV and IV are FI, $I(d)$. $T = 60$. ECM Model: $\Delta^d Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta^d X_t + \beta_1 X_{t-1} + \epsilon_t$.

the ECM story told above for a $Y_t$ that is stationary. Ultimately, the value of $\alpha_1$ tells us more about the level of memory in $Y_t$ than about its relationship to independent variables.

### 4.4.6 Case 6: The Dependent Variable is Explosive, $d>1$

A time series where $d > 1$ is referred to as an explosive process so that increases (or decreases) are followed by larger increases (or decreases) (Fuller 2009). While not a common trait, explosive processes can be found in political science.[32] For our purposes here, the explosive case presents its own set of problems to the GECM the most basic of which is that the equation will not be balanced if one variable is explosive. When $d > 1$, using first-differencing for $Y$ will leave it under-differenced and some measure of auto-correlation

[32]For example, our estimate of $d$ for the *Policy Liberalism* variable of Kelly and Enns (2010) is 1.35 (s.e.=0.10) (See Table C.31 of Appendix C.5.2). This might indicate that the ideological tone of legislation passed by Congress shifts quickly following electoral gains and losses.

will remain. Additionally, having $Y_{t-1}$ on the right-hand side will introduce a great deal of auto-correlation to the model. In short, explosive data are in need of pre-whitening and, when possible, some fractional filter allowing $d > 1$ will be the best approach to remove auto-correlation and allow reliable inferences (Hosking 1981).

### 4.4.7 What Models Should We Use Instead?

Our findings indicate that if one can establish that all series in the model are stationary, the ADL model is preferable to the GECM. If all series contain unit-roots then one can look for cointegration and, if it is present, estimate a GECM while thinking about the consequences of bounds. Next, we discuss two more flexible ECM frameworks: the fractional cointegration and 3-step fractional ECM model of Clarke and Lebo (2003) and the auto-regressive distributive lag (ARDL) bounds approach of Pesaran, Shin, and Smith (2001, PSS).[33]

While the Engle and Granger (1987) approach begins with two $I(1)$ series combining to create a series of $I(0)$ residuals, fractional cointegration adds flexibility by relaxing the assumptions that a) stationarity should be dealt with as a dichotomy, b) the parent series need be $I(1)$, and c) the residuals between $Y$ and $X$ need be $I(0)$ (Baillie and Bollerslev 1994; Box-Steffensmeier and Tomlinson 2000; Cheung and Lai 1993; Dueker and Startz 1998). To find fractional cointegration, the researcher must first establish that the parent series are of the same order of fractional integration, and second, that the ECM is of a lower order of integration ($d$) than the parent series.

The first step follows Engle and Granger (1987) and regresses the level-form $Y$ on the level form $X$ (or $X$s) hypothesized to be error correcting with $Y$. The residual series is the ECM. The $d$ value is then identified for each series – $Y$, $X$, and the ECM – using an estimator such as that of Robinson (1995).[34] If the $d$ value for the ECM is less than the $d$ value for *both* $Y$

[33] This is not meant to be an exhaustive list of available techniques. For example, Brandt and Freeman (2006) evaluates several Bayesian and multi-equation approaches to time series analyses in political science. For a thorough evaluation of Bayesian approaches to time series see: Bauwens, Lubrano, and Richard (2000).

[34] Two concerns are commonly raised over the estimation of $d$ - the power of estimators with smaller samples and poor performance of estimators in the presence of significant short memory (Baillie and Kapetanios 2007).

and $X$, then one can conclude that error correction is occurring. As with Engle and Granger (1987), not finding any evidence of error correction should be enough evidence to drop an ECM specification at this point. However, if one wishes, the approach is capable of following D&K's advice and estimating error correction in the absence of (fractional) cointegration.

As Clarke and Lebo (2003) shows, to avoid bias and Type I errors, the ECM cannot be left as is – unlike Engle and Granger's model with a strictly stationary ECM, here the ECM may still have auto-correlation with $d > 0$. The next stage fractionally differences $Y$, $X$, *and* the ECM by each one's own $d$ value, creating $\Delta^{d_Y} Y_t$, $\Delta^{d_X} X_t$, and the FECM ($\Delta^{d_{ECM}} ECM_t$) before the final step. For example, if $Y$ and $X$ are both $d = 0.8$ and the cointegrating residuals are $d = 0.5$ then $Y$ and $X$ need to be differenced by 0.8 and the ECM needs to be differenced by 0.5. The last step estimates:

$$\Delta^{d_Y} Y_t = \alpha_0 + \alpha_1 \Delta^{d_{ECM}} ECM_{t-1} + \beta_1 \Delta^{d_X} X_t + \epsilon_t. \tag{4.8}$$

This follows the basic intuition of Engle and Granger's ECM framework and the logic of pre-whitening and balancing equations: both right- and left-hand-side variables must be free of auto-correlation to get trustworthy estimates (Hosking 1981; Tsay and Chung 2000; Lebo, Walker, and Clarke 2000; Clarke 2003). By transforming every component of the model to be (0,0,0), the three-step method establishes equation balance and protects inferences while being quite flexible.[35] Simulation results from FECM models with randomly generated FI

Byers, Davidson, and Peel (2000) investigate a wide range of public opinion series and find that the vast majority are well described as (0,$d$,0) models; political time series rarely demonstrate the need for higher-order approximations (see also: Lebo, Walker, and Clarke 2000)). Estimates of the data sets we replicated affirm the conclusion of Byers et. al, and we provide the Schwartz Bayesian criterion estimates for various ARFIMA ($p, d, q$) models in Appendix C.8.1. Regarding shorter time series, Robinson (1995) finds that the semi-parametric Whittle estimator is unbiased with as few as 64 observations. Grant (n.d.) investigates the finite sample properties of various estimators and finds that both semi-parametric and parametric estimators are unbiased with (0,$d$,0) series, even with N as small as 40. Further, when higher order approximations are necessary, the negative bias found in semiparametric estimators and the parametric exact MLE is not present in the frequency-domain approximate ML estimator. Appendix C.8.1 compares the long memory estimates of replicated data series using three estimators: the semi-parametric Whittle, the exact ML estimator, and the frequency-domain ML estimator. The estimates agree with each other closely.

[35] The FECM model can be used in a near-VAR framework so that one can measure how multiple endogenous variables re-equilibrate to each other. See, for example: Lebo, McGlynn, and Koger (2007).

series are shown in Table 4.7. The Type I error rate for error correction hypotheses stays below the 5 percent threshold regardless of the memory of the original series.

**Table 4.7: Results of FECM Model with FI Data, $T$=60†**

| Fractional Order of Integration (d=) | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 | 1.1 | 1.2 |
|---|---|---|---|---|---|---|---|---|
| % FECM Significant | 4.5 | 4.0 | 3.9 | 3.7 | 3.7 | 3.4 | 3.8 | 4.0 |
| % Significant X | 5.9 | 5.8 | 5.9 | 5.9 | 5.9 | 5.6 | 5.9 | 5.8 |
| % FECM Significant with $d < d(Y)$ and $d(X)$ | 1.7 | 1.9 | 1.9 | 1.9 | 2.1 | 1.9 | 2.3 | 2.4 |

†*Note*: Percentage results in each cell based on 10,000 simulations per bivariate FECM model with null hypothesis of no fractional cointegration. Significance of FECM based on one-tail test. Significance of X based on two-tail test. DV and IV are fractionally integrated at the same level of $d$. $T = 60$.
FECM Model: $\Delta^{d_Y} Y_t = \alpha_0 + \alpha_1 \Delta^{d_{ECM}} ECM_{t-1} + \beta_1 \Delta^{d_X} X_t + \epsilon_t$..

One disadvantage of the three-step model is interpretability. Explaining that "a one-unit change in fractionally differenced $X$ leads to a $\beta$ change in fractionally differenced $Y$" is difficult. And, while the FECM's $\alpha_1$ coefficient in Equation 4.8 has a similar interpretation to an ECM in Engle and Granger's two-step model – the speed of error correction – the coefficient will be slightly different from the value that relates to the level-form variables. In both cases, however, the extra difficulty in interpretability is a good trade-off for a trustworthy hypothesis test that avoids both Type I and Type II errors.

Pesaran et al.'s (2001) ARDL bounds test procedure falls between the GECM and fractional cointegration techniques in terms of flexibility:

$$\Delta Y_t = \beta_0 + \theta_1 Y_{t-1} + \theta_2 X_{t-1} + \sum_{i=1}^{p} \varpi_i \Delta Y_{t-i} + \sum_{j=0}^{p} \beta_j \Delta X_{t-j} + \epsilon_t. \tag{4.9}$$

The ARDL model uses the same $\Delta Y_t$ as the GECM for the dependent variable and the inclusion of the $\Delta Y_{t-i}$ and $\Delta X_{t-j}$ terms on the right-hand side of the equation allow for serial correlation and ensure the error term is white noise. 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. The ARDL model allows greater flexibility on the right hand side, but because it is a cointegration test, the dependent variable, $Y_t$, must still be $I(1)$. Further, the researcher must ensure that no series within the model is $I(2)$, the inclusion of which will invalidate the model's results.

The ARDL is a two-step procedure and the first step, bounds testing, refers to hypothesis testing the presence of a long-run relationship via an $F$-test, which tests the joint significance of the coefficients on the lagged level-form variables in Equation 4.9, i.e., $\mathrm{H}_0$: $\theta_1 = \theta_2 = 0$. The $F$-test has a non-standard distribution that depends on the number of regressors, the presence of an intercept or trend, and the sample size. The critical values for the bounds of the $F$-test were computed by PSS based upon sample sizes of 500 and 1000, however the method is appropriate for use with smaller samples. For smaller samples between 30 and 80, Narayan (2005) provides proper critical values.

Should the $F$-test exceed the upper bound, the researcher can reject the null of no cointegration and conclude that a long-run relationship is present, irrespective of the order(s) of integration of the regressors. If the $F$-statistic falls below the lower bound, the null hypothesis cannot be rejected. Note that the presence of cointegration implies the presence of Granger causality in at least one direction, however the ARDL does not indicate the direction of causal ordering (See: Dickinson and Lebo 2007, for a short summary and example of the method.). Among the model's disadvantages is the loss of degrees of freedom with longer lags. Additionally, the model's flexibility in terms of its allowance of uncertainty as to the regressors' orders of integration is helpful should the $F$-test exceed the bounds, however if the computed $F$-statistic fall within the critical value band, the order(s) of integration must be determined before a correct conclusion can be made.

## 4.5 Applications

Next, we replicate five recent GECM studies. Two are presented below and three (Sanchez Urribarri et al. (2011), Kelly and Enns (2010), and Volscho and Kelly (2012)) can be found in Appendices C.5.1, C.5.2, and C.5.3 of Supplementary Materials. In each study, key findings change when using alternative methods.

### 4.5.1 Example 1: Casillas, Enns, and Wohlfarth (*AJPS*, 2011)

Casillas, Enns, and Wohlfarth (2011, CEW) investigate the effect of public mood on the Supreme Court's reversal of cases in a liberal direction and find a strong positive relationship for non-salient cases; for salient cases, public mood has no effect. These are compelling findings given long debate on the relationship between public opinion and the Court's behavior. Using the GECM, CEW find Court responsiveness to be much stronger than previously shown. ECM parameters of -0.83, -0.77, and -1.27 are described as powerful equilibrium relationships where previous work had found tenuous ties at best.[36]

We first check the properties of the dependent variables.[37] As seen above, the further a series is from $I(1)$, the more negative the ECM coefficient and its $t$-statistic will be. As shown in Table 4.8, *All Reviews*, *Non-Salient Reviews*, and *Salient Reviews* test as fractionally integrated with the $d$ value for *Salient Reviews* noticeably lower. Error correction is very likely to appear strong in a GECM of *Salient Reviews*, regardless of the covariates.

**Table 4.8: Dickey-Fuller Results of CEW's Dependent Variables**$^{\dagger}$

| Model | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| Dickey-Fuller Coefficient | -0.18 | -0.18 | -0.51 |
| Dickey-Fuller $t$-statistic | -2.28 | -2.32 | -3.98$^{\star}$ |
| MacKinnon DF Critical Value | -2.94 | -2.94 | -2.94 |
| Estimated $d$ Value | 0.62 (0.11) | 0.62 (0.12) | 0.36 (0.08) |

$^{\dagger}$ *Note*: Dickey-Fuller critical values are from MacKinnon (1994). $d$ values estimated using Stata's exact ML estimator; Robinson's semi-parameteric estimator in RATS provides similar estimates: *All Reviews*($d = 0.63$); *Non-Salient Reviews* ($d = 0.65$); *Salient Reviews* ($d = 0.30$).

To demonstrate, we use CEW's DVs and simulate covariates to replicate their three models.[38] Table 4.9's Model 1 is based on 10,000 simulations per DV with three randomly generated (0,0,0) IVs. Following the authors' use of standard $t$ critical values, the models have 66.3%, 66.5%, and 97.7% Type I error rates for error correction, respectively. Switching

[36]ECM parameters of All Review, Non-Salient Review, and Salient Review models, respectively. For Non-Salient Reviews they report: "The error correction rate indicates that 77% of the long-run effect of mood will occur at term $t+1$ (0.68) and an additional 77% of the remaining effect will influence the Court at term $t+2$ (0.16). Therefore, 94% of the total long-run effect of public opinion at term $t$ will be manifested in the justices' behavior after just two terms."

[37]Replications of CEW's models and further details are in Appendix C.3.

[38]We follow CEW's use of 2SLS but our findings are roughly the same using OLS (see Appendix C.3.3).

to MacKinnon values provides reasonable results for *All Reviews* and *Non-Salient Reviews*: no errors at all. That $\Delta X$ and $X_{t-1}$ are significant far too often is still a problem considering the IVs were specifically generated as approximations of white noise processes. The *Salient Reviews* column highlights the problems of estimating the GECM with a DV with a low order of fractional integration. The MacKinnon values find over 25% of ECMs significant and the average size of the $\alpha_1$ parameter would be reported as strong re-equilibration.

**Table 4.9: Simulations of Results for CEW's Dependent Variables in Two Scenarios**

| | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| **Model 1**[‡] | | | |
| % ECM Significant - one tail $t$-distribution | 66.3 | 66.5 | 97.7 |
| % ECM Significant - MacKinnon Values | 0 | 0 | 26.7 |
| Mean of $\alpha_1$ | -0.16 | -0.16 | -0.51 |
| Mean $t$-statistic of $\alpha_1$ | -1.79 | -1.79 | -3.39 |
| % ECM & $\geq 1\Delta X_t$ Significant | 19.1 | 16.4 | 24.3 |
| % ECM & $\geq 1X_{t-1}$ Significant | 15.4 | 12.1 | 23.7 |
| **Model 2**[‡‡] | | | |
| % ECM Significant - one tail $t$-distribution | 94.7 | 95.2 | 99.9 |
| % ECM Significant - MacKinnon Values | 26.7 | 25.4 | 95.3 |
| Mean of $\alpha_1$ | -0.45 | -0.44 | -0.96 |
| Mean $t$-statistic of $\alpha_1$ | -3.23 | -3.21 | -5.87 |
| % ECM & $\geq 1\Delta X_t$ Significant | 26.6 | 23.5 | 31.3 |
| % ECM & $\geq 1X_{t-1}$ Significant | 53.9 | 51.8 | 78.5 |

*Note*: 10,000 simulations per model per DV. [‡]Model 1: IVs are level stationary $I(0)$. [‡‡]Model 2: IVs are unit-roots ($I(1)$). MacKinnon ECM critical values from Ericsson and MacKinnon (2002) for $T = 45$ with 3 IVs: -3.838.

Model 2 of Table 4.9 presents simulations for more realistic data – unit-root IVs. Even with MacKinnon values, we find a significant ECM over 25% of the time for *All Reviews* (26.7%) and *Non-Salient Reviews* (25.4%). The ECM coefficients, averaging -0.45 and -0.44, are far from zero and the hypothesis tests on $\Delta X_t$ and $X_{t-1}$ are much worse – at least one of the three lagged level-form IVs is significant over 50% of the time. The *Salient Reviews* model is as bad as we would expect: the average ECM coefficient is -0.96, what CEW would call almost complete re-equilibration, with random data. Comparing the results of Model 1 and Model 2 shows that bias in $\alpha_1$ as well as $\beta_0$ and $\beta_1$ increases markedly in relation to the memory of $X_t$, a reminder that pre-whitening is important for all variables in a model.

Model 2 also highlights inferential problems caused by the DV's order of integration. The

Dickey-Fuller test fails to reject the null of a unit-root for both *All Reviews* and *Non-Salient Reviews*, but when regressing these DVs on simulated $I(1)$ series we reject the null of no error correction far too often, even with MacKinnon values. In this case, the fractional integration of the dependent variables confounds the hypothesis test.

Next, Table 4.10 shows the results when we replace CEW's three IVs with: yearly *World-wide Shark Attacks*, *U.S. Tornado Fatalities*, and *U.S. Beef Consumption*.[39] Using CEW's interpretation and favored $t$-value we find strong evidence of error correction in all three models.[40] With an ECM of -1.04 (s.e.=.16), even using MacKinnon's critical values leads to finding extremely strong error correction in the *Salient Reviews* model.

The ECM problems are compounded by spurious regressions – our nonsense IVs are significant far too often. Across all three models, seven of our nine short-term variables and six long-term variables – along with their respective LRMs – are significant.[41]

Because each of the three dependent variables used by CEW are estimated to be fractionally integrated, we also re-estimate their data using fractional methods. We first pre-whiten by fractionally differencing each series by its estimated order of fractional integration. The $d$ estimates, standard errors, and confidence intervals for each series can be found in Appendix C.3, Table C.5. Each DV is fractionally integrated, however we cannot reject the null that the three IVs are unit-roots. Because no DV / IV combination shares the same order of fractional integration, no model is a candidate for fractional cointegration.

Table 4.11 presents the results of regression models after all variables have been fractionally differenced, and what is immediately apparent is that the results are consistent with

[39] Beef Consumption (million tons) data from USDA, *Production, Supply, and Distribution* database (www.fas.usda.gov/psdonline); Shark Attack data from Florida Museum of Natural History (http://www.flmnh.ufl.edu/fish/sharks/statistics/Trends2.htm); Tornado Fatalities from NOAA (http://www.nws.noaa.gov/om/hazstats/resources/weather_fatalities.pdf).

[40] We have no evidence of endogeneity problems, but to match CEW we use a 2SLS model and instrument for *Beef Consumption* with CEW's excluded "social forces," which are used as instruments of Martin-Quinn scores. Social forces such as the homicide rate, policy liberalism, inequality, etc., may predict mood, but not MQ scores. Instruments tests for proper inclusion in the first stage fail with both our data and CEW's data. See Appendix C.3.2 for CEW's first-stage results and Appendix C.3.1 for the full replication of CEW.

[41] The long-run multiplier (LRM) is calculated as the ratio of coefficients: $\beta_1/\alpha_1$. The formula for the standard errors of the LRM is: $((1/b^2)\text{Var}(a) + (a^2/b^4)\text{Var}(b) - 2(a/b^3)\text{Cov}(a,b))^{1/2}$ See De Boef and Keele (2008, pp.191-192).

**Table 4.10: Using the GECM to Explain the Court's Liberal Reversal Rate with Sharks, Tornadoes, and Beef Consumption**

| | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| **Long Run Multiplier** | | | |
| Shark Attacks | 0.56* | 0.53* | 0.51* |
| | (0.20) | (0.18) | (0.19) |
| Tornado Fatalities | | | |
| Beef Consumption | -1.79* | -1.69* | -2.04* |
| | (0.32) | (0.78) | (0.31) |
| **Long Run Effects** | | | |
| Shark Attacks$_{t-1}$ | 0.30* | 0.29* | 0.53* |
| | (0.14) | (0.13) | (0.21) |
| Tornado Fatalities$_{t-1}$ | 0.05 | 0.04 | 0.10 |
| | (0.04) | (0.04) | (0.08) |
| Beef Consumption (IV)$_{t-1}$ | -0.95* | -0.92* | -2.12* |
| | (0.36) | (0.33) | (0.46) |
| **Short Run Effects** | | | |
| ΔShark Attacks | 0.20* | 0.18 | 0.39* |
| | (0.12) | (0.11) | (0.22) |
| ΔTornado Fatalities | 0.06* | 0.04* | 0.10* |
| | (0.02) | (0.02) | (0.05) |
| ΔBeef Consumption (IV) | -7.10* | -6.88* | -2.90 |
| | (4.17) | (3.86) | (7.89) |
| **Error Correction and Constant** | | | |
| Percent Liberal$_{t-1}$ | -0.53* | -0.54* | -1.04* |
| | (0.17) | (0.17) | (0.16) |
| Constant | 61.71* | 61.52* | 129.11* |
| | (21.75) | 20.44 | (25.58) |
| **Fit and Diagnostics** | | | |
| Centered $R^2$ | 0.14 | 0.19 | 0.55 |
| Sargan ($\chi^2$) | 0.69 | 0.60 | 0.32 |
| N | 45 | 45 | 45 |

*Note*: Entries are two-stage least squares coefficients (s.e. in parentheses). ECM significance one-tail $t$-test. Coefficient significance (*p≤ 0.05, one-tail test).

the short-term results reported by CEW. Pre-whitening caused no harm to the findings. Changes in either public mood or the ideological makeup of the Court will affect reversal rates. However, by ensuring equation balance, we avoided the spurious conclusions that an equilibrium relationship exists between the Court's reversal rate and public mood.

As a final step in our replication of this data, we also estimate the first step of the ARDL bounds test of Pesaran, Shin, and Smith (2001) with CEW's *All Review* and *Non-Salient Review* models. After estimating an unrestricted model, the $F$-statistic of the lagged-level variables was compared to the bounds computed by Narayan (2005). Recall from our simulations in Table 4.9, Model 2 that the GECM was susceptible to Type I errors due to the fact that both DVs are fractionally integrated. Despite this, we rely on the fact that neither

**Table 4.11: Re-estimation with Fractional Methods (Casillas et. al (2011))**

| Review Type | $\Delta^d$ All Reviews | $\Delta^d$ Non-Salient Reviews | $\Delta^d$ Salient Reviews |
|---|---|---|---|
| **Short Run Effects** | | | |
| $\Delta^d$ Public Mood | 1.43* | 1.61* | 0.90 |
| | (0.73) | (0.72) | (1.76) |
| $\Delta^d$ Segal-Cover | 9.61* | 8.39* | 14.51 |
| | (4.67) | (4.58) | (11.30) |
| $\Delta^d$ Martin-Quinn | 0.11 | -0.30 | 1.79 |
| | (2.30) | (2.26) | (5.56) |
| Constant | -1.35 | -1.39 | 59.68* |
| | (1.25) | (1.23) | (3.02) |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.10 | 0.10 | -0.02 |
| Breusch-Pagan Test ($\chi^2$) | 0.06 | 0.47 | 0.93 |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.93 | 1.01 | 2.11 |

*Note*: Entries are OLS coefficients (s.e. in parentheses). All variables have been fractionally differenced by their estimate of *d*. Coefficient significance (*p≤ 0.05, two-tail test).

**Table 4.12: ARDL Bounds Test for Cointegration (Casillas et. al (2011))**

| | 5% Critical Bounds | |
|---|---|---|
| | $I(0)$ | $I(1)$ |
| | 3.535 | 4.733 |
| All Review Model | $F = 3.19$ | |
| Non-Salient Review Model | $F = 2.54$ | |

*Note*: $F$-statistic of unrestricted ARDL model with $K + 1$ lagged-level form variables. Critical values from Narayan (2005) based on $T = 45$, $K = 3$, unrestricted intercept and no trend. Neither model can reject null of no cointegration.

DV can reject the null of a unit-root as justification for using the ARDL to test for a long-run relationship. The results are found in Table 4.12, and neither model's $F$-statistic exceeds the bounds. We cannot reject the null hypothesis of no cointegration, which indicates that the GECM is misspecified.

### 4.5.2 Example 2: Ura and Ellis (*The Journal of Politics*, 2012)

Our second replication, Ura and Ellis (2012, U&E), posits that macro-economic and political variables have asymmetric effects on partisan groups' responses and that this is at least partially responsible for mass polarization. U&E use General Social Survey items to create a measure of macro-policy sentiment among partisan groups. With T=35, a large number of independent variables, and the use of the GECM, the paper's strong findings are suspect.

U&E's main results (their Table 2, replicated in Appendix C.4.1 of Supplementary Materials) include error correction rates of -0.40 (s.e.=0.09, $t$=-4.54) for Republicans and -0.69 (s.e.=0.17, $t$=-4.11) for Democrats. Our Table 4.1 gives the appropriate critical value from Ericsson and MacKinnon (2002) for $T = 35$ as -4.268 for $I(1)$ series. Thus, there is some evidence of error correction and the adjusted-$R^2$ values of 0.39 indicate well specified models.

But are the strong findings simply due to use of the GECM? We first replace their DVs – Republican and Democratic Mood – with randomly generated $I(1)$ series. The tendency to find significant error correction is strong – following U&E's use of a standard $t$-test the ECM parameter is significant (.05 level) in approximately 86% of the simulations.[42] Type I error rates on the other covariates are also troublesome. For example, in models where both the ECM and an $X_{t-1}$ are significant, we observe equilibrating relationships in almost 62% of all simulations. Using MacKinnon values we fare somewhat better but still make Type I errors on the ECM at over four times the rate we should (21.9% and 22.6%).

As discussed, MacKinnon values assume data are strictly $I(1)$ and become insufficient as any series in the model deviates from that. The correct 5% rejection region can only be found using the exact auto-correlation patterns of the data at hand.[43] In this case, 5% of the simulations have $t$-values below -5.81. Yet, even if it were feasible to ask researchers to run simulations to calculate idiosyncratic critical values, the bias in the ECM values and the Type I errors on the $X$s make the GECM problematic with mixed orders of integration.[44]

Next, we try our Shark-Tornado-Beef data as predictors of U&E's dependent variables. Since U&E have five independent variables we add yearly values of *U.S. Onion Acreage* and *U.S. Coal Emissions*.[45] Table 4.13 shows the results: the ECMs in both equations are

[42] Full simulation results are in Appendix C.4.1 of the Supplement.

[43] Estimates of $d$ using Stata's exact ML estimator: DemMood ($d$=1.15, s.e.=0.15), RepMood ($d$=1.05, s.e.=0.16), Domestic Spending ($d$=1.44, s.e.=0.08), Defense Spending ($d$=1.32, s.e.=0.11), Inflation ($d$=1.02, s.e.=0.25), Unemployment ($d$=0.94, s.e.=0.21), Inequality ($d$=0.89, s.e.=0.19).

[44] Note that both *Defense* and *Domestic Spending* series are estimated as explosive processes, biasing the ECM $t$-statistic in the negative direction making Type I errors more likely. See Appendix C.1 for further explanation. Additionally, the original DVs are bounded so our simulations of unbounded unit root DVs are not exactly equivalent.

[45] Onion Data (10,000s of acres) from USDA (http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1396); Coal Emission (million tons) data comes from DOE

**Table 4.13: What Else Moves Mood? (Re-estimating Ura and Ellis (2012))**

| | Republican Mood | Democrat Mood | \|Difference\| |
|---|---|---|---|
| **Long Run Multipliers** | | | |
| Onion Acreage | -0.29* (0.04) | -0.20* (0.05) | 0.09 |
| Coal Emissions | -0.13* (0.02) | -0.11* (0.02) | 0.01 |
| Beef Consumption | -0.40 (0.26) | -0.03 (0.05) | 0.36 |
| Shark Attacks | 0.01* (0.00) | 0.00 (0.00) | 0.01 |
| Tornado Fatalities | -0.05* (0.01) | -0.02* (0.01) | 0.03 |
| **Long Run Effects** | | | |
| Onion $\text{Acreage}_{t-1}$ | -0.18* (0.03) | -0.10* (0.02) | 0.08* |
| Coal $\text{Emissions}_{t-1}$ | 0.08* (0.01) | 0.06* (0.01) | 0.02 |
| Beef $\text{Consumption}_{t-1}$ | 0.25 (0.18) | 0.02 (0.17) | 0.23 |
| Shark $\text{Attacks}_{t-1}$ | -0.01* (0.00) | 0.00 (0.00) | 0.01 |
| Tornado $\text{Fatalities}_{t-1}$ | 0.03* (0.01) | 0.01 (0.01) | 0.02* |
| **Short Run Effects** | | | |
| $\Delta$Onion $\text{Acreage}_t$ | 0.01 (0.03) | 0.03 (0.02) | 0.02 |
| $\Delta$Coal $\text{Emissions}_t$ | 0.08* (0.02) | 0.05* (0.02) | 0.04 |
| $\Delta$Beef $\text{Consumption}_t$ | 0.83* (0.33) | 0.20 (0.20) | 0.63 |
| $\Delta$Shark $\text{Attacks}_t$ | 0.03* (0.02) | -0.00 (0.01) | 0.03 |
| $\Delta$Tornado $\text{Fatalities}_t$ | 0.02* (0.00) | 0.01 (0.01) | 0.01* |
| **Error Correction and Constant** | | | |
| Partisan $\text{Mood}_{t-1}$ | -0.62* (0.11) | -0.52* (0.14) | 0.11 |
| Constant | 12.08 (7.56) | 20.52* (5.25) | |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.57 | 0.29 | |
| Breusch-Pagan Test ($\chi^2$) | 0.00 | 1.86 | |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.29 | 0.11 | |

*Note*: Entries are seemingly unrelated regression coefficients (s.e. in parentheses). ECM significance one-tail $t$-test. Coefficient significance (*p$\leq$ 0.05, two-tail test). The difference column reports the absolute difference and $t$-test between coefficient estimates for the two models (*p$\leq$ 0.05).

substantively large, and even with MacKinnon values we would claim that Republic Mood is in equilibrium with four of our five nonsense series. Many of the short-term and long-run independent variables are significant as well. With an $R^2$ of 0.57 the model of Republican Mood is noticeably better than U&E's and the Democratic model is just a touch worse ($R^2$=0.29). The ECM values, hypothesis tests, model fit statistics, and long-run multipliers all give us erroneous inferences.

Importantly, these are series with lots of memory – $d$ varies between 0.89 and 1.44. Inattention to stationarity could potentially lead one to conclude that they are all simply $I(1)$ and that MacKinnon values will make for valid inferences. In fact, Dickey-Fuller tests on the *Domestic Spending* series indicates that it's $I(2)$, and inappropriate for use in either the GECM or ARDL models in its level form.

Finally, we fractionally difference U&E's variables and add a FECM between *Partisan*

(www.eia.gov/totalenergy/data/annual/showtext.cfm?t=ptb1101).

**Table 4.14: Re-estimation with Three-Step FECM (Ura and Ellis (2012))**

| | Republican Mood | | Democrat Mood | |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$ Domestic Spending \$10b | 0.02 | (0.04) | -0.00 | (0.04) |
| $\Delta^d$ Defense Spending \$10b | 0.31* | (0.13) | 0.11 | (0.12) |
| $\Delta^d$ Inflation | -0.31 | (0.20) | -0.05 | (0.18) |
| $\Delta^d$ Unemployment | -0.47 | (0.37) | -0.06 | (0.32) |
| $\Delta^d$ Top 1% Income Share | 1.22* | (0.42) | 0.55 | (0.37) |
| **Error Correction and Constant** | | | | |
| FECM | 0.12 | (0.17) | -0.16 | (0.19) |
| Constant | -1.30 | (0.68) | -0.18 | (0.59) |
| **Fit and Diagnostics** | | | | |
| $R^2$ | 0.38 | | 0.11 | |
| Breusch-Pagan Test ($\chi^2$) | 1.51 | | 0.07 | |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.62 | | 0.00 | |

*Note*: Entries are OLS coefficients (s.e. in parentheses). All variables have been fractionally differenced by their estimate of $d$. FECM significance (*p$\leq$ 0.05, one-tail test). Coefficient significance (*p$\leq$ 0.05, two-tail test).

*Mood* and *Top 1% Income Share*, the predictor with the smallest $p$-value in U&E's models.[46] The FECMs in both models are indistinguishable from zero as are the independent variables in the *Democratic Partisan Mood* model.[47] Republicans on the other hand do appear to shift in a liberal direction following increases to *Defense Spending* and *Top 1% Income Share*. Thus, some of U&E's findings stand, but many do not hold up to these alternative specifications, and we find no evidence of error correction. As a final check, we run FECM models on our Shark-Tornado-Beef and find none of the FECMs and only one of 48 hypothesis tests on the independent variables to pass the .05-level of significance.[48] With T=35 estimation of $d$ and fractional differencing are somewhat under-powered and some wariness is understandable. Still, the method is uncovering relationships we should expect (e.g. Defense Spending) and not setting off false alarms (e.g. Shark Attacks). With short series we advocate the use of multiple modeling techniques as checks for robustness.

[46] We estimate FECMs between *Partisan Mood* and each of the other covariates and include those results in Appendix C.4.3 of Supplementary Materials. All results are null.

[47] Running the model on first differences also returns null results.

[48] Table C.13 of Appendix C.4 uses the independent variable with the lowest $p$-value (Onion Acreage) in the FECM. Each of the 4 other series takes a turn in the FECM in the models presented in Tables C.14 (Republican Mood) and C.15 (Democratic Mood).

## 4.6 Replication Roundup

Table 4.15 summarizes the results of our replications. The general pattern is an overstatement of significant relationships in the original papers. The body of work using the GECM would make it seem that error correction is a common pattern among political time series but our reanalyses show otherwise. Error correction between variables is a very close relationship that should be obvious in a simple glance at the data. Non-intuitive findings of error correction should make political scientists highly suspicious.

One final lesson from these exercises is the value of using multiple approaches. Time series analysts do not have the luxury of being able to replicate studies with new data but robustness checks can come from finding similar results across diverse modeling choices.

**Table 4.15: Findings and Replications**

| **Authors** | Central Finding | Replication Finding |
|---|---|---|
| Casillas, Enns, & Wohlfarth (2011) | Public mood and Justice ideology exhibit significant short- and long-run effects on Court decision making. As public liberalism increases, so does the Court's rate of liberal decisions. | Public mood significant in short-term only. No evidence of cointegration for either *All Review* or *Non-Salient Review* models. *Salient Review* model inappropriate with mixed orders of integration. |
| Ura & Ellis (2012) | Partisan mood varies with political stimuli. Changes in mood and increases in polarization are asymmetric between parties with most movement coming from Republicans. | Short-term effects for defense spending and top 1% income share only. Three-step FECM better accounts for long-memoried series with no cointegration. |
| Sanchez Urribarri, et al. (2011) | Institutional conditions and changes in judicial ideology explain the rights agendas of multiple high courts. No expected support for theory of exogenous structural change hypothesis. | Change in judicial ideology only significant for UK model. Insignificant in US model. No significant institutional effects. No support for exogenous structural change hypothesis. GECM model either inappropriate with mixed orders of integration or unnecessary given stationarity of the data. |
| Kelly & Enns (2010) | Income inequality exhibits both short-term and long-run effects on public mood. Both high-income and low-income populations exhibit increased conservatism in response to inequality. | No support for short- or long-term effect of income inequality on public mood. No support for either high or low income conservatism with increasing inequality. No cointegration - GECM model inappropriate. |
| Volscho & Kelly (2012) | Political, policy, and economic inputs all significantly affect income distribution of top 1%. Political inequality not only a result of market forces, but also a product of political change. | No support for politics or policy indicators having significant effect on *Top 1% Income*. Some economic indicators significant. Significant FECM between *Trade Openness* and *S&P Composite* with *Top 1% Income*. GECM model inappropriate with mixed orders of integration. |

## 4.7 Conclusion

Political scientists' troubles with the GECM are not surprising when the method is traced back to the econometrics literature where its primary usage is as a cointegration test for $I(1)$ data (Banerjee et al. 1993; Ericsson and MacKinnon 2002). In that literature the GECM is not a one-step model, *it is the first step of a multi-step process.* The first step ignores the hypothesis tests on $X_{\text{t-1}}$ and $\Delta X_{\text{t}}$ and uses MacKinnon values to test whether $\alpha_1 = 0$. Failing to reject that hypothesis means that cointegration is not present. Rejecting the null means that cointegration is present. In either case, it is not a standard regression model. If cointegration is not found, no ECM is specified. If it is found, additional steps follow. Re-postulating the GECM as a single-equation regression puts too many conditions on the model.

In fact, we recommend the GECM in only one rare situation: when all of the variables are strictly unit-root series, $Y_t$ is unbounded, $Y_t$ and $X_t$ are cointegrated, and the MacKinnon critical values are used. We looked at many combinations of series' characteristics and in every case but that hypothetical one, the GECM ran into serious problems.

A careful look at the applied literature in political science will not find any examples that meet all those criteria. When data are bounded or diverge from $I(1)$ by being stationary, auto-regressive, explosive, fractionally, or near-integrated the ECM parameter begins to move dramatically. Adapting to MacKinnon critical values is not a sufficient fix and generating idiosyncratic critical values is untenable as well. Bias in the ECM ruins its ability to do what it should – tell us about equilibrium relationships. As for the GECM's use with stationary data, the ADL and GECM may be mathematically equivalent but the GECM adds complications without adding useful insights. If $Y$'s long-term tendency is simply to revert to its mean, we should not prefer a model whose purpose it to find $Y$'s long-run relationship with X.

As every parent knows, the solution to the babies and bathwater dilemma is not to just leave the baby sitting in the dirty water – the two can be carefully separated. With time series, the safe solution is to return to the logic of pre-whitening and use a multi-step process that removes auto-correlation from each variable prior to including it in a multivariate model. Time series data are complicated – too complicated to provide reliable results in a simple single step.

## Bibliography


Baillie, R, and George Kapetanios. 2007. Semiparametric Estimation of Long Memory: The Holy Grail or a Poisoned Chalice. In *International Symposium on Financial Econometrics*. pp. 1–2.

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

Baillie, Richard T, and Tim Bollerslev. 1994. "Cointegration, Fractional Cointegration, and Exchange Rate Dynamics." *The Journal of Finance* 49(2): 737–745.

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

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

Bårdsen, Gunnar. 1989. "Estimation of Long Run Coefficients in Error Correction Models." *Oxford Bulletin of Economics and Statistics* 51(3): 345–350.

Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard. 2000. *Bayesian inference in dynamic econometric models*. Oxford University Press.

Beck, Nathaniel. 1992. "Comparing Dynamic Specifications: The Caseof Presidential Approval." *Political Analysis* 3: 27–50.

Beran, Jan. 1994. *Statistics for Long-Memory Processes*. Vol. 61 CRC Press.

Bernhard, William, and David Leblang. 2002. "Democratic Processes, Political Risk, and Foreign Exchange Markets." *American Journal of Political Science* pp. 316–333.

Box, George E. P., and Gwilym M. Jenkins. 1970. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Box, George EP, and David A Pierce. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." *Journal of the American Statistical Association* 65(332): 1509–1526.

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

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

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

Brandt, Patrick T, and John R Freeman. 2006. "Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis." *Political Analysis* 14(1): 1–36.

Büthe, Tim, and Helen V Milner. 2014. "Foreign Direct Investment and Institutional Diversity in Trade Agreements: Credibility, Commitment, and Economic Flows in the Developing World, 1971–2007." *World Politics* 66(1): 88–122.

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

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

Cavaliere, Giuseppe. 2005. "Limited Time Series With a Unit Root." *Econometric Theory* 21: 907–945.

Cavaliere, Giuseppe, and Fang Xu. 2014. "Testing for Unit Roots in Bounded Time Series." *Journal of Econometrics* 178(2): 259–272.

Cheung, Yin-Wong, and Kon S Lai. 1993. "A Fractional Cointegration Analysis of Purchasing Power Parity." *Journal of Business & Economic Statistics* 11(1): 103–112.

Clarke, Harold D., and Marianne C. Stewart. 1994. "Prospections,Retrospections and Rationality: The 'Bankers' Model of Presidential ApprovalReconsidered." *American Journal of Political Science* 38: 1104–1123.

Clarke, Harold D., and Matthew Lebo. 2003. "Fractional (co)integration and Governing Party Support in Britain." *British Journal of Political Science* 33(02): 283–301.

Clarke, Kevin A. 2003. "Nonparametric Model Discrimination in International Relations." *Journal of Conflict Resolution* 47(1): 72–93.

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

De Boef, Suzanna. 2001. "Modeling equilibrium relationships: Error correction models with strongly autoregressive data." *Political Analysis* 9(1): 78–94.

De Boef, Suzanna, and Jim Granato. 1999. "Testing for Cointegrating Relationships with Near-Integrated Data." *Political Analysis* 8(1): 99–117.

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

DeBoef, Suzanna, and Jim Granato. 1997. "Near-integrated Data and the Analysis of Political Relationship." *American Journal of Political Science* 41: PP–PP.

Dickinson, Matthew J., and Matt J. Lebo. 2007. "Reexamining the Growth of the Institutional Presidency, 1940–2000." *Journal of Politics* 69(1): 206–219.

Dueker, Michael, and Richard Startz. 1998. "Maximum-Likelihood Estimation of Fractional Cointegration With an Application to US and Canadian Bond Rates." *Review of Economics and Statistics* 80(3): 420–426.

Durr, Robert H. 1992. "Of forests and trees." *Political Analysis* pp. 255–258.

Durr, Robert H. 1993. "What Moves Policy Sentiment?" *American Political Science Review* 87: 158–170.

Durr, Robert H, Andrew D Martin, and Christina Wolbrecht. 2000. "Ideological Divergence and Public Support for the Supreme Court." *American Journal of Political Science* pp. 768–776.

Enders, Walter. 2004. *Applied Econometric Time Series*. 2nd ed. New York: Wiley.

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

Enns, Peter K. 2014. "The Public's Increasing Punitiveness and Its Influence on Mass Incarceration in the United States." *American Journal of Political Science* 58(4): 857–872.

Enns, Peter K, Nathan J Kelly, Janna Morgan, Thomas Volscho, and Christopher Witko. 2014. "Conditional Status Quo Bias and Top Income Shares: How U.S. Political institutions have Benefitted the Rich." *The Journal of Politics* 76: 289–303.

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

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

Fuller, Wayne A. 2009. *Introduction to Statistical Time Series*. Vol. 428 Wiley. com.

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

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

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

Granger, C.W.J. 2010. "Some Thoughts on the Development of Cointegration." *Journal of Econometrics* 158: 3–6.

Hansen, Bruce E. 1995. "Rethinking the Univariate Approach to Unit Root Testing." *Econometric Theory* 11(5): 1148–71.

Hendry, David F., and G. J. Anderson. 1977. *Testing Dynamic Specification in Small Simultaneous Systems: An Application to a Model of Building Society Behaviour in the United Kingdom.* North Holland Publishing Company.

Hendry, David F, and Grayham E Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England." *The Economic Journal* 88(351): 549–563.

Hosking, Jonathan RM. 1981. "Fractional Differencing." *Biometrika* 68(1): 165–176.

Jennings, Will, and Peter John. 2009. "The Dynamics of Political Attention: Public Opinion and the Queen's Speech in the United Kingdom." *American Journal of Political Science* 53(4): 838–854.

Kayser, Mark Andreas. 2009. "Partisan Waves: International Business Cycles and Electoral Choice." *American Journal of Political Science* 53(4): 950–970.

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

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

Kono, Daniel Yuichi. 2008. "Democracy and Trade Discrimination." *Journal of Politics* 70(4): 942–955.

Kremers, Jeroen JM, Neil R Ericsson, and Juan J Dolado. 1992. "The Power of Cointegration Tests." *Oxford bulletin of economics and statistics* 54(3): 325–348.

Layman, Geoffrey C., Thomas Carsey, John C. Green, Richard Herrera, and Rosalyn Cooperman. 2010. "Activists and Conflict Extension in American Party Politics." *American Political Science Review* 104(2): 324–346.

Leblang, David, and William Bernhard. 2006. "Parliamentary Politics and Foreign Exchange Markets: The World According to GARCH." *International Studies Quarterly* 50(1): 69–92.

Lebo, Matthew J., Adam J. McGlynn, and Greg Koger. 2007. "Strategic Party Government: Party Influence in Congress, 1789–2000." *American Journal of Political Science* 51(3): 464–481.

Lebo, Matthew J, and Everett Young. 2009. "The Comparative Dynamics of Party Support in Great Britain: Conservatives, Labour and the Liberal Democrats." *Journal of Elections, Public Opinion and Parties* 19(1): 73–103.

Lebo, Matthew J, and Will H Moore. 2003. "Dynamic Foreign Policy Behavior." *Journal of Conflict Resolution* 47(1): 13–32.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

MacKinnon, James G. 1994. "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests." *Journal of Business & Economic Statistics* 12(2): 167–176.

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

Nicolau, Joao. 2002. "Stationary Processes That Look Like Random Walks - The Bounded Random Walk Process in Discrete and Continuous Time." *Econometric Theory* 18(1): 99–118.

Ostrom, Charles W., Jr., and Renée Smith. 1993. "Error Correction,Attitude Persistence, and Executive Rewards and Punishments: A BehavioralTheory of Presidential Approval." *Political Analysis* 4: 127–184.

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

Phillips, A William. 1954. "Stabilisation Policy in a Closed Economy." *The Economic Journal* 64(254): 290–323.

Phillips, AW. 1957. "Stabilisation Policy and the Time-Forms of Lagged Responses." *The Economic Journal* 67(266): 265–277.

Pickup, Mark. 2009. "Testing for Fractional Integration in Public Opinion in the Presence of Structural Breaks: A Comment on Lebo and Young." *Journal of Elections, Public Opinion and Parties* 19(1): 105–116.

Ramirez, Mark D. 2009. "The Dynamics of Partisan Conflict on Congressional Approval." *American Journal of Political Science* 53(3): 681–694.

Rickard, Stephanie J. 2012. "Welfare Versus Subsidies: Governmental Spending Decisions in an Era of Globalization." *Journal of Politics* 74(4): 1171–1183.

Robinson, Peter M. 1995. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of statistics* 23(5): 1630–1661.

Sanchez Urribarri, Raul A, Susanne Schorpp, Kirk A Randazzo, and Donald R Songer. 2011. "Explaining Changes to Rights Litigation: Testing a Multivariate Model in a Comparative Framework." *The Journal of Politics* 73(02): 391–405.

Sargan, John Denis. 1964. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology." *Econometric analysis for national economic planning* 16: 25–54.

Stimson, James A. 1991. *Public Opinion in America: Moods, Cycles, and Swings*. Boulder, CO: Westview Press.

Stock, James H, and Mark W Watson. 2011. "Introduction to Econometrics, 3/E." *Instructor* 10.

Tsay, Wen-Jay, and Ching-Fan Chung. 2000. "The Spurious Regression of Fractional Integrated Processes." *Journal of Econometrics* 96: 155–182.

Ura, Joseph Daniel. 2014. "Backlash and Legitimation: Macro Political Responses to Supreme Court Decisions." *American Journal of Political Science* 58(1): 110–126.

Ura, Joseph Daniel, and Christopher R Ellis. 2012. "Partisan Moods: Polarization and the Dynamics of Mass Party Preferences." *The Journal of Politics* 74(1): 277–291.

Ura, Joseph Daniel, and Patrick C Wohlfarth. 2010. ""An Appeal to the People": Public Opinion and Congressional Support for the Supreme Court." *Journal of Politics* 72(4): 939–956.

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.

Williams, John T. 1992. "What goes around comes around: Unit root tests and cointegration." *Political Analysis* pp. 229–235.

Young, Everett, and Matthew J Lebo. 2009. "Long Memory Methods and Structural Breaks in Public Opinion Time Series: A Reply to Pickup." *Journal of Elections, Public Opinion and Parties* 19(1): 117–124.

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): 1–63.

Zivot, Eric. 2000. "The Power of Single Equation Tests for Cointegration When the Cointegrating Vector is Prespecified." *Econometric Theory* 16(03): 407–439.

## Chapter 5

# Conclusion

The intent of this dissertation was to address some of the concerns which seem to have precluded the use of fractional methods within political time series. Initial work on fractional integration in political science focused on the aggregation theorem of Granger (1980), and this justification for the presence of long memory is useful considering how many series are aggregates. But additional work in economics has focused on other ways in which series may be fractionally integrated and these new justifications are varied and interesting. One of these is the error duration model of Parke (1999). Chapter 2 works through the error duration representation and finds that the survival of federal programs is a fractionally integrated process, as are series that are dependent on the original series. These findings are relevant because it opens up the possibility for a larger number of series to be long memory. Further, that fractional integration can be inherited by series reliant on the original process is a concept that is considered in economics, but it is rarely, if ever, tested in political science. It is hoped that these new justifications behind the expectation that a series may be fractionally integrated spurs researchers to test for fractional integration in more series, and more seriously consider the possibility that time series do not fit into the $I(0)/I(1)$ dichotomy.

One of the more prevalent arguments against the use of fractional integration is that the estimates are unreliable in small samples. One needs only look to Keele and Linn (2015) for the latest iteration of this argument. But while this argument against the use of fractional integration is common, the assumption has never actually been tested. Monte Carlo studies investigating the power of various estimators almost always involve larger sample sizes such as those that are commonly found in economics and finance. But it is difficult to extrapolate the results from these studies to performance with finite samples. Chapter 3 puts the unreliable estimation assumption to the test and finds that concerns are largely overstated. The results indicate that the frequency domain maximum likelihood is generally unbiased and efficient. Semiparametric estimators are also valuable considering their ability to easily and accurately estimate the memory parameter of $(0, d, 0)$ models. While these estimators, and especially the FML prove reliable, the exact maximum likelihood (EML) estimator is problematic.

If we return to the arguments levied against fractional integration in works like Keele and Linn (2015) we find that only one estimator is tested, and it is the one that performs the worst in our estimates: the EML. But the problems inherent in estimation with the EML have been known since the work of Li and McLeod (1986) and Cheung and Diebold (1994). Over the last 20 years, economists have developed a rapidly growing number of fractional estimators, to use the one estimator previously proven to be biased and inconsistent in order to argue against accounting for fractional integration is shortsighted. If we move away from the time domain exact maximum likelihood estimator we find that our fractional estimates are much improved.

A complementary argument that is used against the adoption of fractional methods is that the models are too complicated. As methods have grown more complex over the last two decades we have seen a growing desire for model simplicity. In turn, political scientists have adopted the general error correction model (GECM), which was first introduced by economists in the 1970s. Since being reintroduced to political scientists in 2008, the GECM has since become one of the more popular models in the field. As demonstrated in Chapter 4 however, the assumptions underlying this method have either been forgotten or misinterpreted. A multitude of papers have been published claiming to find strong error correction and the interpretation of these results is often in the tone of a cointegration analysis - referencing strong equilibrium relationships *between* variables. Yet as we demonstrate, if the model is not being used explicitly as a cointegration test between unit root series, the interpretation of the model results must change. If the dependent variable is stationary, or fractionally integrated, the error correction parameter is practically guaranteed significance and is demonstrating the rate of return to its own mean, it is not indicative of a relationship between variables. Furthermore, the current interpretation of the assumptions underlying the GECM have led political scientists to assume that the stationarity of the time series do not matter and therefore does not need to be tested. The empirical and theoretical consequences of this have been dreadful. In addition to demonstrating the pitfalls of the current use of the GECM and reviewing five separate articles using the method, it's propose that researchers consider fractional integration as an alternative specification. As argued in the chapter, FI methods are flexible, are able to address the spurious regression problems found in the GECM, and it also forces a consideration of stationarity, something that has been sorely missing from political time series analysis for the last several years.

## Bibliography

Cheung, Yin-Wong, and Francis X Diebold. 1994. "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally-Integrated Noise with Unknown Mean." *Journal of Econometrics* 62(2): 301–316.

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

Keele, Luke, and Suzanna Linn. 2015. "Treating Time with All Due Seriousness." *Political Analysis* .

Li, WK, and A Ian McLeod. 1986. "Fractional Time Series Modelling." *Biometrika* 73(1): 217–221.

Parke, William R. 1999. "What is Fractional Integration?" *Review of Economics and Statistics* 81(4): 632–638.

## Bibliography

Anderson, T. W. 1971. *The Statistical Analysis of Time Series.* New York: Wiley.

Ansolabehere, Stephen, and James M Snyder Jr. 2002. "The Incumbency Advantage in US Elections: An Analysis of State and Federal Offices, 1942-2000." *Election Law Journal* 1(3): 315–338.

Baillie, R, and George Kapetanios. 2007. Semiparametric Estimation of Long Memory: The Holy Grail or a Poisoned Chalice. In *International Symposium on Financial Econometrics.* pp. 1–2.

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

Baillie, Richard T, and George Kapetanios. 2009. Semi Parametric Estimation of Long Memory: Comparisons and Some Attractive Alternatives. In *conference in honor of Peter Phillip's Sixtieth Birthday at Singapore Management University.*

Baillie, Richard T, and Tim Bollerslev. 1994. "Cointegration, Fractional Cointegration, and Exchange Rate Dynamics." *The Journal of Finance* 49(2): 737–745.

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

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

Bårdsen, Gunnar. 1989. "Estimation of Long Run Coefficients in Error Correction Models." *Oxford Bulletin of Economics and Statistics* 51(3): 345–350.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(4): 1181–1206.

Baum, Christopher F, and Richard Sperling. 2000. "sts15. 1: Tests for Stationarity of a Time Series: Update." *Stata Technical Bulletin* 58: 35–36.

Baumgartner, Frank R., and Bryan D. Jones. 1993. *Agendas and Instability in American Politics.* Chicago: University of Chicago Press.

Bauwens, Luc, Michel Lubrano, and Jean-Francois Richard. 2000. *Bayesian inference in dynamic econometric models.* Oxford University Press.

Beck, Nathaniel. 1983. "Time-Varying Parameter Regression Models." *American Journal of Political Science* 27(3): 557–600.

Beck, Nathaniel. 1989. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1(1): 121–156.

Beck, Nathaniel. 1992. "Comparing Dynamic Specifications: The Caseof Presidential Approval." *Political Analysis* 3: 27–50.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* pp. 96–129.

Beran, Jan. 1994. *Statistics for Long-Memory Processes*. Vol. 61 CRC Press.

Bernhard, William, and David Leblang. 2002. "Democratic Processes, Political Risk, and Foreign Exchange Markets." *American Journal of Political Science* pp. 316–333.

Berry, Christopher R, Barry C Burden, and William G Howell. 2010. "After Enactment: The Lives and Deaths of Federal Programs." *American Journal of Political Science* 54(1): 1–17.

Binder, Sarah A. 1997. *Minority Rights, Majority Rule: Partisanship and the Development of Congress*. Cambridge: Cambridge University Press.

Box, George E. P., and Gwilym M. Jenkins. 1970. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Box, George EP, and David A Pierce. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." *Journal of the American Statistical Association* 65(332): 1509–1526.

Box-Steffensmeier, Janet, Kathleen Knight, and Lee Sigelman. 1998. "The Interplay of Macroideology and Macropartisanship: A Time Series Analysis." *Journal of Politics* 60: 131–149.

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

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

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

Brandt, Patrick T, and John R Freeman. 2006. "Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis." *Political Analysis* 14(1): 1–36.

Brehm, John, and Paul Gronke. 1994. "History, Heterogeneity, and Presidential Approval: a Modified ARCH Approach.".

Büthe, Tim, and Helen V Milner. 2014. "Foreign Direct Investment and Institutional Diversity in Trade Agreements: Credibility, Commitment, and Economic Flows in the Developing World, 1971–2007." *World Politics* 66(1): 88–122.

Byers, David, James Davidson, and David Peel. 1997. "Modelling Political Popularity: An Analysis of Long-range Dependence in Opinion Poll Series." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160(3): 471–490.

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

Caporale, Guglielmo Maria, and Luis A Gil-Alana. 2004. "Fractional Cointegration and Tests of Present Value Models." *Review of Financial Economics* 13(3): 245–258.

Caporale, Guglielmo Maria, and Nikitas Pittis. 1999. "Unit Root Testing Using Covariates: Some Theory and Evidence." *Oxford Bulletin of Economics and Statistics* 61(4): 583–595.

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

Cavaliere, Giuseppe. 2005. "Limited Time Series With a Unit Root." *Econometric Theory* 21: 907–945.

Cavaliere, Giuseppe, and Fang Xu. 2014. "Testing for Unit Roots in Bounded Time Series." *Journal of Econometrics* 178(2): 259–272.

Chambers, Marcus J. 1998. "Long Memory and Aggregation in Macroeconomic Time Series." *International Economic Review* pp. 1053–1072.

Cheung, Yin-Wong, and Francis X Diebold. 1994. "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally-Integrated Noise with Unknown Mean." *Journal of Econometrics* 62(2): 301–316.

Cheung, Yin-Wong, and Kon S Lai. 1993. "A Fractional Cointegration Analysis of Purchasing Power Parity." *Journal of Business & Economic Statistics* 11(1): 103–112.

Clarke, Harold D., and Marianne C. Stewart. 1994. "Prospections,Retrospections and Rationality: The 'Bankers' Model of Presidential ApprovalReconsidered." *American Journal of Political Science* 38: 1104–1123.

Clarke, Harold D., and Matthew Lebo. 2003. "Fractional (co)integration and Governing Party Support in Britain." *British Journal of Political Science* 33(02): 283–301.

Clarke, Kevin A. 2003. "Nonparametric Model Discrimination in International Relations." *Journal of Conflict Resolution* 47(1): 72–93.

Corder, J Kevin. 2004. "Are Federal Programs Immortal? Estimating the Hazard of Program Termination." *American Politics Research* 32(1): 3–25.

Davidson, James. 2014. "Time Series Modelling Version 4.44.".

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

De Boef, Suzanna. 2001. "Modeling equilibrium relationships: Error correction models with strongly autoregressive data." *Political Analysis* 9(1): 78–94.

De Boef, Suzanna, and Jim Granato. 1999. "Testing for Cointegrating Relationships with Near-Integrated Data." *Political Analysis* 8(1): 99–117.

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

DeBoef, Suzanna, and Jim Granato. 1997. "Near-integrated Data and the Analysis of Political Relationship." *American Journal of Political Science* 41: PP–PP.

Dickinson, Matthew J., and Matt J. Lebo. 2007. "Reexamining the Growth of the Institutional Presidency, 1940–2000." *Journal of Politics* 69(1): 206–219.

Diebold, Francis X, and Glenn D Rudebusch. 1991. "On the Power of Dickey-Fuller Tests Against Fractional Alternatives." *Economics letters* 35(2): 155–160.

Doornik, Jurgen A. 2014. "An Object-Oriented Matrix Programming Language Ox 7.".

Dueker, Michael, and Richard Startz. 1998. "Maximum-Likelihood Estimation of Fractional Cointegration With an Application to US and Canadian Bond Rates." *Review of Economics and Statistics* 80(3): 420–426.

Durr, Robert H. 1992. "Of forests and trees." *Political Analysis* pp. 255–258.

Durr, Robert H. 1993. "What Moves Policy Sentiment?" *American Political Science Review* 87: 158–170.

Durr, Robert H, Andrew D Martin, and Christina Wolbrecht. 2000. "Ideological Divergence and Public Support for the Supreme Court." *American Journal of Political Science* pp. 768–776.

Enders, Walter. 2004. *Applied Econometric Time Series*. 2nd ed. New York: Wiley.

Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica: Journal of the Econometric Society* pp. 987–1007.

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

Enns, Peter K. 2014. "The Public's Increasing Punitiveness and Its Influence on Mass Incarceration in the United States." *American Journal of Political Science* 58(4): 857–872.

Enns, Peter K, Nathan J Kelly, Janna Morgan, Thomas Volscho, and Christopher Witko. 2014. "Conditional Status Quo Bias and Top Income Shares: How U.S. Political institutions have Benefitted the Rich." *The Journal of Politics* 76: 289–303.

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

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

Finocchiaro, Charles J, and Tse-min Lin. 2000. *The Hazards of Incumbency: An Event History Analysis of Congressional Tenure*. Michigan State University, Institute for Public Policy and Social Research.

Freeman, John. 2012. "Research Note: Can Time Series Methods Be Used to Detect Path Dependence?" *Typescript. University of Minnesota. www. polisci. umn. edu/freeman* .

Fuller, Wayne A. 2009. *Introduction to Statistical Time Series*. Vol. 428 Wiley. com.

Geweke, John, and Susan Porter-Hudak. 1983. "The Estimation and Application of Long Memory Time Series Models." *Journal of time series analysis* 4(4): 221–238.

Gil-Alana, Luis A. 2003. "Testing of Fractional Cointegration in Macroeconomic Time Series*." *Oxford Bulletin of Economics and Statistics* 65(4): 517–529.

Gil-Alana, Luis A, and Javier Hualde. 2009. "Fractional Integration and Cointegration. An Overview and an Empirical Application." In *Palgrave Handbook of Econometrics*, ed. Terence C. Mills, and Kerry Patterson. Vol. 2 MacMillan Publishers London pp. 434–469.

Granger, Clive WJ. 1966. "The Typical Spectral Shape of an Economic Variable." *Econometrica: Journal of the Econometric Society* pp. 150–161.

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

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

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

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

Granger, C.W.J. 2010. "Some Thoughts on the Development of Cointegration." *Journal of Econometrics* 158: 3–6.

Grant, Taylor, and Matthew J. Lebo. 2015. "Error Correction Methods with Political Time Series." *Political Analysis* .

Hansen, Bruce E. 1995. "Rethinking the Univariate Approach to Unit Root Testing." *Econometric Theory* 11(5): 1148–71.

Hassler, Uwe, and Jürgen Wolters. 1994. "On the Power of Unit Root Tests Against Fractional Alternatives." *Economics Letters* 45(1): 1–5.

Hauser, Michael A. 1999. "Maximum Likelihood Estimators for ARMA and ARFIMA Models: a Monte Carlo Study." *Journal of Statistical Planning and Inference* 80(1): 229–255.

Hendry, David F., and G. J. Anderson. 1977. *Testing Dynamic Specification in Small Simultaneous Systems: An Application to a Model of Building Society Behaviour in the United Kingdom*. North Holland Publishing Company.

Hendry, David F, and Grayham E Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England." *The Economic Journal* 88(351): 549–563.

Henry, M, and P. Zaffaroni. 2002. "The Long-Range Dependence Paradigm for Macroeconomics and Finance." In *Theory and Applications of Long-Range Dependence*, ed. P. Doukhan, G. Oppenheim, and M. Taqqu. Birkhäuser.

Henry, M, and PM Robinson. 1996. Bandwidth Choice in Gaussian Semiparametric Estimation of Long Range Dependence. In *Athens Conference on Applied Probability and Time Series Analysis*. Springer pp. 220–232.

Hosking, Jonathan RM. 1981. "Fractional Differencing." *Biometrika* 68(1): 165–176.

Hurvich, Clifford M, Rohit Deo, and Julia Brodsky. 1998. "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* 19(1): 19–46.

Jennings, Will, and Peter John. 2009. "The Dynamics of Political Attention: Public Opinion and the Queen's Speech in the United Kingdom." *American Journal of Political Science* 53(4): 838–854.

Jones, Bryan D. 1999. "Bounded Rationality." *Annual Review of Political Science* 2(1): 297–321.

Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *The American Political Science Review* 97(1): 151–169.

Jones, Richard H. 1980. "Maximum Likelihood Fitting of ARMA Models to Time Series With Missing Observations." *Technometrics* 22(3): 389–395.

Kalman, Rudolph E, and Richard S Bucy. 1961. "New Results in Linear Filtering and Prediction Theory." *Journal of Fluids Engineering* 83(1): 95–108.

Kaufman, Herbert. 1976. *Are Government Organizations Immortal?* Brookings Inst Pr.

Kay, Adrian. 2005. "A Critique of the Use of Path Dependency in Policy Studies." *Public Administration* 83(3): 553–571.

Kayser, Mark Andreas. 2009. "Partisan Waves: International Business Cycles and Electoral Choice." *American Journal of Political Science* 53(4): 950–970.

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

Keele, Luke, and Suzanna Linn. 2015. "Treating Time with All Due Seriousness." *Political Analysis* .

Kellstedt, Paul, Gregory E. McAvoy, and James A. Stimson. 1993. "Dynamic Analysis with Latent Constructs: The Kalman DYMIMIC Specification." *Political Analysis* 5: 113–150.

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

Knight, Jack. 1992. *Institutions and Social Conflict*. Cambridge University Press.

Kono, Daniel Yuichi. 2008. "Democracy and Trade Discrimination." *Journal of Politics* 70(4): 942–955.

Kremers, Jeroen JM, Neil R Ericsson, and Juan J Dolado. 1992. "The Power of Cointegration Tests." *Oxford bulletin of economics and statistics* 54(3): 325–348.

Kuensch, Hans R. 1987. Statistical Aspects of Self-Similar Processes. In *Proceedings of the first World Congress of the Bernoulli Society*. Vol. 1 VNU Science Press Utrecht, The Netherlands pp. 67–74.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of econometrics* 54(1): 159–178.

Layman, Geoffrey C., and Thomas M. Carsey. 2002*a*. "Party Polarization and "Conflict Extension" in the American Electorate." *American Journal of Political Science* 46(4): 786–802.

Layman, Geoffrey C, and Thomas M Carsey. 2002*b*. "Party Polarization and Party Structuring of Policy Attitudes: A Comparison of Three NES Panel Studies." *Political Behavior* 24(3): 199–236.

Layman, Geoffrey C., Thomas Carsey, John C. Green, Richard Herrera, and Rosalyn Cooperman. 2010. "Activists and Conflict Extension in American Party Politics." *American Political Science Review* 104(2): 324–346.

Leblang, David, and William Bernhard. 2006. "Parliamentary Politics and Foreign Exchange Markets: The World According to GARCH." *International Studies Quarterly* 50(1): 69–92.

Lebo, Matthew J., Adam J. McGlynn, and Greg Koger. 2007. "Strategic Party Government: Party Influence in Congress, 1789–2000." *American Journal of Political Science* 51(3): 464–481.

Lebo, Matthew J., and Daniel Cassino. 2007. "The Aggregated Consequences of Motivated Reasoning and the Dynamics of Partisan Presidential Approval." *Political Psychology* 28(6): 719–746.

Lebo, Matthew J, and Everett Young. 2009. "The Comparative Dynamics of Party Support in Great Britain: Conservatives, Labour and the Liberal Democrats." *Journal of Elections, Public Opinion and Parties* 19(1): 73–103.

Lebo, Matthew J, and Will H Moore. 2003. "Dynamic Foreign Policy Behavior." *Journal of Conflict Resolution* 47(1): 13–32.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

Li, Richard P. Y. 1976. "A Dynamic Comparative Analysis of Presidential and House Elections." *American Journal of Political Science* 20(4): 671–691.

Li, WK, and A Ian McLeod. 1986. "Fractional Time Series Modelling." *Biometrika* 73(1): 217–221.

Lo, Andrew W, and Archie Craig MacKinlay. 1988. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *Review of financial studies* 1(1): 41–66.

MacKinnon, James G. 1994. "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests." *Journal of Business & Economic Statistics* 12(2): 167–176.

Maddala, G.S. 1998. "Rcent Developments in Dynamic Econometric Modelling: A Personal Viewpoint." *Political Analysis* pp. 59–87.

Maltzman, Forrest, and Charles R Shipan. 2008. "Change, Continuity, and the Evolution of the Law." *American Journal of Political Science* 52(2): 252–267.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *The Journal of Political Economy* 89(5): 914–927.

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

Nicolau, Joao. 2002. "Stationary Processes That Look Like Random Walks - The Bounded Random Walk Process in Discrete and Continuous Time." *Econometric Theory* 18(1): 99–118.

Nielsen, Morten Ørregaard, and Per Houmann Frederiksen. 2005. "Finite Sample Comparison of Parametric, Semiparametric, and Wavelet Estimators of Fractional Integration." *Econometric Reviews* 24(4): 405–443.

Ostrom, Charles W., Jr., and Renée Smith. 1993. "Error Correction,Attitude Persistence, and Executive Rewards and Punishments: A BehavioralTheory of Presidential Approval." *Political Analysis* 4: 127–184.

Parke, William R. 1999. "What is Fractional Integration?" *Review of Economics and Statistics* 81(4): 632–638.

Patterson, Kerry. 2012. *Unit Root Tests in Time Series Volume 2: Extensions and Developments*. Vol. 2 Palgrave Macmillan.

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

Phillips, A William. 1954. "Stabilisation Policy in a Closed Economy." *The Economic Journal* 64(254): 290–323.

Phillips, AW. 1957. "Stabilisation Policy and the Time-Forms of Lagged Responses." *The Economic Journal* 67(266): 265–277.

Phillips, Peter CB. 1991. "Optimal Inference in Cointegrated Systems." *Econometrica: Journal of the Econometric Society* pp. 283–306.

Phillips, Peter CB, and Katsumi Shimotsu. 2004. "Local Whittle Estimation in Nonstationary and Unit Root Cases." *The Annals of Statistics* 32(2): 656–692.

Phillips, Peter CB, and Sam Ouliaris. 1990. "Asymptotic Properties of Residual Based Tests for Cointegration." *Econometrica: Journal of the Econometric Society* pp. 165–193.

Pickup, Mark. 2009. "Testing for Fractional Integration in Public Opinion in the Presence of Structural Breaks: A Comment on Lebo and Young." *Journal of Elections, Public Opinion and Parties* 19(1): 105–116.

Pierson, Paul. 2000. "Increasing Returns, Path Dependence, and the Study of Politics." *American political science review* 94(02): 251–267.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton University Press.

Ramirez, Mark D. 2009. "The Dynamics of Partisan Conflict on Congressional Approval." *American Journal of Political Science* 53(3): 681–694.

Rickard, Stephanie J. 2012. "Welfare Versus Subsidies: Governmental Spending Decisions in an Era of Globalization." *Journal of Politics* 74(4): 1171–1183.

Robinson, Peter M. 1995*a*. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of statistics* 23(5): 1630–1661.

Robinson, Peter M. 1995*b*. "Log-Periodogram Regression of Time Series with Long Range Dependence." *The annals of Statistics* pp. 1048–1072.

Sanchez Urribarri, Raul A, Susanne Schorpp, Kirk A Randazzo, and Donald R Songer. 2011. "Explaining Changes to Rights Litigation: Testing a Multivariate Model in a Comparative Framework." *The Journal of Politics* 73(02): 391–405.

Sargan, John Denis. 1964. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology." *Econometric analysis for national economic planning* 16: 25–54.

Shimotsu, Katsumi. 2010. "Exact Local Whittle Estimation of Fractional Integration with Unknown Mean and Time Trend." *Econometric Theory* 26(02): 501–540.

Shimotsu, Katsumi, and Peter CB Phillips. 2006. "Local Whittle Estimation of Fractional Integration and Some of its Variants." *Journal of Econometrics* 130(2): 209–233.

Simon, Herbert A. 1985. "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review* 79: 293–304.

Sowell, Fallaw. 1992*a*. "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models." *Journal of econometrics* 53(1): 165–188.

Sowell, Fallaw. 1992*b*. "Modeling Long-Run Behavior with the Fractional ARIMA Model." *Journal of Monetary Economics* 29(2): 277–302.

Stimson, James A. 1991. *Public Opinion in America: Moods, Cycles, and Swings*. Boulder, CO: Westview Press.

Stock, James H, and Mark W Watson. 2011. "Introduction to Econometrics, 3/E." *Instructor* 10.

Tsay, Wen-Jay, and Ching-Fan Chung. 2000. "The Spurious Regression of Fractional Integrated Processes." *Journal of Econometrics* 96: 155–182.

Ura, Joseph Daniel. 2014. "Backlash and Legitimation: Macro Political Responses to Supreme Court Decisions." *American Journal of Political Science* 58(1): 110–126.

Ura, Joseph Daniel, and Christopher R Ellis. 2012. "Partisan Moods: Polarization and the Dynamics of Mass Party Preferences." *The Journal of Politics* 74(1): 277–291.

Ura, Joseph Daniel, and Patrick C Wohlfarth. 2010. ""An Appeal to the People": Public Opinion and Congressional Support for the Supreme Court." *Journal of Politics* 72(4): 939–956.

Velasco, Carlos. 1999*a*. "Gaussian Semiparametric Estimation of Non-stationary Time Series." *Journal of Time Series Analysis* 20(1): 87–127.

Velasco, Carlos. 1999*b*. "Non-Stationary Log-Periodogram Regression." *Journal of Econometrics* 91(2): 325–371.

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.

Wildavsky, Aaron B. 1964. "Politics of the Budgetary Process.".

Williams, John T. 1992. "What goes around comes around: Unit root tests and cointegration." *Political Analysis* pp. 229–235.

Young, Everett, and Matthew J Lebo. 2009. "Long Memory Methods and Structural Breaks in Public Opinion Time Series: A Reply to Pickup." *Journal of Elections, Public Opinion and Parties* 19(1): 117–124.

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): 1–63.

Zivot, Eric. 2000. "The Power of Single Equation Tests for Cointegration When the Cointegrating Vector is Prespecified." *Econometric Theory* 16(03): 407–439.

# Appendices

# Appendix A

# Appendix: Chapter 2

## A.1 What is Fractional Integration

Assume a process that generates a series integrated of order $d$, in which $d$ is the number of differences that must be applied to the series for it to be stationary. In the standard case, in which $d$ is an integer, we have the $I(1)/I(0)$ distinction: if a series is an $I(1)$ unit root, it has perfect memory, meaning that a shock is fully integrated into the series. On the other hand, an $I(0)$ series has only short memory; it is weakly stationary with constant mean, finite variance, and constant covariance (see Enders 2004, p.54).

But if we relax the assumption that $d$ must be an integer and instead allow it to fall anywhere on the real number line, then the level of integration is considered fractional.[1] The ARFIMA process, introduced by Granger and Joyeux (1980) and Hosking (1981) is labeled $(p, d, q)$ if its $d$'th difference yields a stationary and invertible ARMA $(p, q)$ process.

Assuming that $-1/2 < d < 1/2$ in order to ensure stationarity and invertibility, the properties of $y_t$ can be defined as an ARFIMA $(p, d, q)$ if

$$\phi(L)(1-L)^d(y_t - \mu) = \theta(L)\epsilon_t, \tag{A.1}$$

where $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + ... + \theta_q z^q$ are lag polynomials in the lag operator $L(Lx_t = x_{t-1})$ with roots strictly outside the unit circle, $\epsilon_t$ is $iid$ $N(0, \sigma^2)$ and $(1-L)^d$ is defined by its binomial expansion such that

$$(1-L)^d = \sum_{j=0}^{\infty} \psi_j(d) L^j, \;\; \psi_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \; j = 0, 1, ..., \tag{A.2}$$

where $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$ is the gamma function.

The estimated $d$ parameter determines the memory of the series in question. If $-1/2 < d < 1/2$ the process is weakly stationary and invertible. If $d < 0$, the process is considered anti-persistent, in which case the autocorrelations are mostly negative and summable (sum to a finite constant). For $0 < d < 1/2$ the process is considered stationary but possesses "long memory;" it is characterized by a hyperbolically declining autocorrelation function, and the autocorrelations and autocovariances are nonsummable. Finally, for $1/2 \leq d < 1$,

[1] Fractionally integrated processes can be analyzed in either the time domain or the frequency domain; this section focuses on the time domain, and the frequency domain is discussed in more depth in the following section.

the series is no longer covariance stationary and is characterized as non-persistent. Even though the series is non-stationary, the MA coefficients, which approximate the impulse response function of a unit shock in $\epsilon_t$, will eventually decline to 0.

### A.1.1 Autocovariance and Autocorrelation of a Stationary ARFIMA Process

To better describe these properties, we provide the autocovariances and autocorrelations of an ARFIMA process (Hosking 1981). The autocovariance is given by

$$E(y_t y_{t-k}) = \gamma_k = \frac{(-1)^k(-2d)!}{(k-d)!(-k-d)!} \tag{A.3}$$

setting $k = 0$ allows us to obtain the variance

$$\gamma_0 = \frac{1(-2d)!}{(-d)!(-d)!} = \frac{(-2d)!}{[(-d)!]^2} \tag{A.4}$$

We can then form the autocorrelation function $\rho_k$ as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{(-d)!(k+d-1)!}{(d-1)!(k-d)!} \quad (K = 0, \pm 1, ...) \tag{A.5}$$

$$= \frac{d(1+d)...(k-1+d)}{(1-d)(2-d)...(k-d)} \quad (K = 1, 2, ...) \tag{A.6}$$

$$= \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)} \tag{A.7}$$

As $k \to +\infty$ then $\frac{\Gamma(k+d)}{\Gamma(k+1-d)} \to k^{(d)-(1-d)} = k^{2d-1}$ which leaves us with

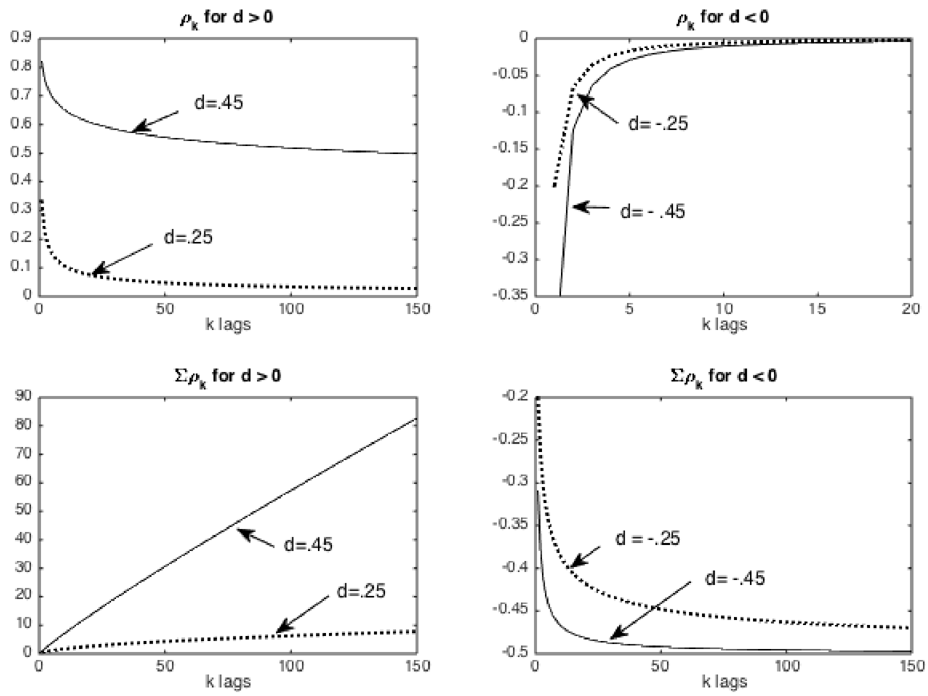
$$\rho_k \to \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1} \tag{A.8}$$

$$\rho_k \sim c_\rho k^{2d-1}, \; 0 < c_\rho < \infty, \quad \text{as}\, k \to \infty \tag{A.9}$$

in which $c_p$ is a constant and $\rho_k$ decays at a hyperbolic rate. From Equation A.9 it is apparent that the sum of the autocorrelations are driven by $k^{2d-1}$. If $d$ is greater than 1/2, then the exponent on $k$ is greater than 0 and the autocorrelations are not summable. Of course, it is also true that $d < 1/2$ is not a sufficient condition for the summability of the autocorrelations. In order for the autocorrelations to converge to a finite sum, $d$ must be less than 0. This is demonstrated with a $p$-series convergence test: if we set $k^{2d-1} = 1/k^{-(2d-1)}$, then in order for convergence to occur, the exponent $-(2d-1)$ must be greater than 1, which means that $d < 0$ (Patterson 2012).

The nonsummable nature of long memory processes is more clearly demonstrated in visual form. Below are a series of plots representing the autocorrelations of two long memory processes ($d = .25, .45$) and two anti-persistent series, ($d = -.25, -.45$). The upper left plot presents the autocorrelation functions for the two long-memory series and the slow decay for each is evident. The anti-persistent series (upper right) do not exhibit the same process, and as can be seen in the lower right panel, the autocorrelations for anti-persistent series are summable. The summed autocorrelations for long memory processes are in the lower left panel. The autocorrelations do not sum to a finite quantity, and this is clearly seen with the $d = .45$ series. As the value of $d$ gets closer to 0.5, the slope of the summed series steepens.

**Figure A.1: Autocorrelations of Stationary ARFIMA Model**




Autocorrelations depending on level of memory, $d$. Upper left ($d = .25, .45$); upper right ($d = -.25, -.45$). Bottom panels are the sums of each.

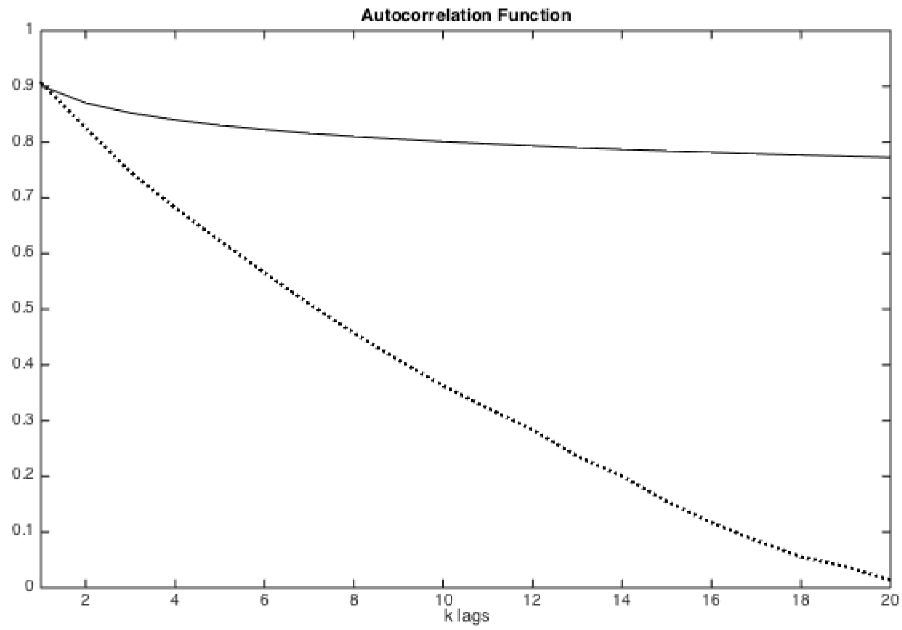
The slow decline of the autocorrelation function of the FI series is particularly noticeable when compared to the autocorrelation function of a simple AR process. Figure A.2 presents the autocorrelation functions of two series, each generated such that their first-order autocorrelations are equal to .90. The solid line is a fractionally integrated series, the dashed line is an AR(1) process sufficiently large to be considered near-integrated. Despite the substantial memory in the AR(1) process the disparity in the decay rates is obvious. It is not uncommon for researchers to confuse near-integrated and fractionally integrated series, or to assume that they are similar, but as we can see here, the two series have drastically different levels of memory.[2]

### A.1.2 AR and MA Coefficients of a Stationary ARFIMA Processes

The ARFIMA model separates the short-run and long-run dynamics by modeling the short-run through the ARMA lag polynomials $\phi(L)$ and $\theta(L)$ while capturing the long-run characteristics with the fractional differencing parameter. Referring back to Equation A.2, the

[2]Part of this confusion may stem from the rather loose usage of "long memory" in political science as a descriptor of both near-integrated and fractionally integrated series (see, e.g. De Boef 2001).

**Figure A.2: Autocorrelation: Near Integrated and Fractionally Integrated Series**




Autocorrelations for FI series (solid) and near integrated series (dashed).

autoregressive coefficients of $\psi_j(d)$ can be expressed as

$$\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} \tag{A.10}$$

Or, if using a moving average representation, in which case we are interested in $(1-L)^{-d}$, the moving average coefficients are given by

$$\frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} \tag{A.11}$$

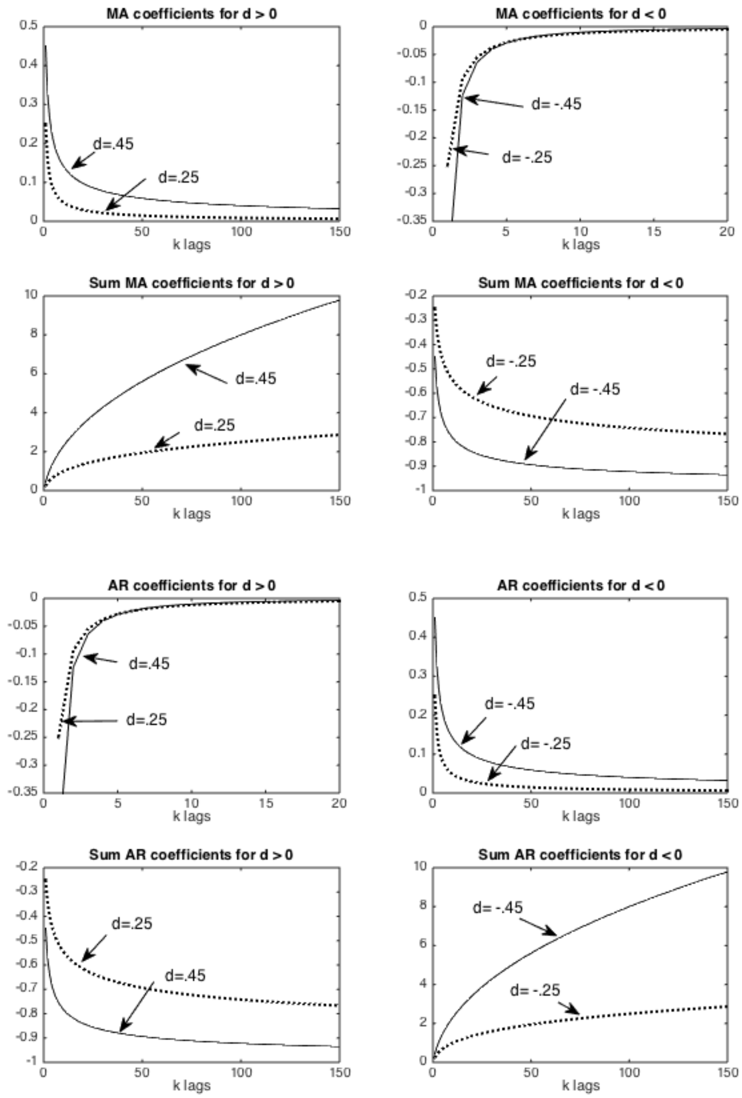
Using Stirling's theorem, as $j \to \infty$ the AR coefficients behave as $\Gamma(-d)^{-1}j^{-d-1}$ and the MA coefficients as $\Gamma(d)^{-1}j^{d-1}$. Note that the MA coefficients for a given $d$ are the same as the AR coefficients for $-d$.

Plot A.3 presents the AR and MA coefficients for the same $d$ series as used in the demonstration of the autocorrelation function. The MA coefficients are positive and declining when $d$ is positive, and will eventually decline so long as $0 < d < 1$. The AR coefficients are the negative of the MA coefficients.

### A.1.3 Hypothesis Testing and Fractionally Integrated Series

When working with series that may be fractionally integrated, it is important to recognize the long memory process described above will generally confound unit root tests. Further,

**Figure A.3: AR & MA coefficients**




given the already low power of these tests when used with small sample sizes, any test results are best viewed with a critical eye. In best practice, multiple types of unit root tests with different null hypotheses should be employed. If the Dickey-Fuller and the KPSS test both reject the null, that information is useful as a first step in determining the memory properties of a series (see Box-Steffensmeier and Smith 1996, for an example of this in practice).

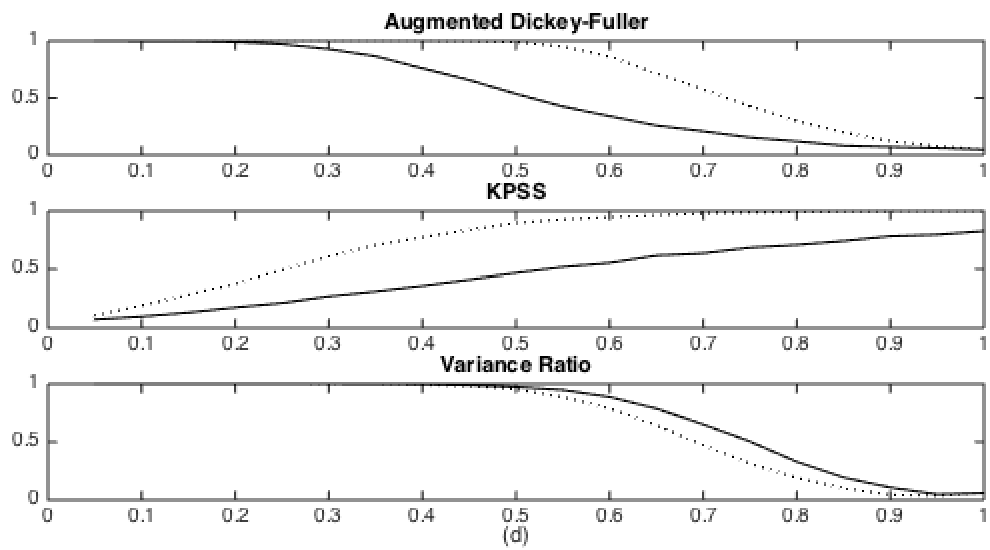
Previous research has demonstrated that the power of the Dickey-Fuller test decreases more quickly as the FI series moves away from one when compared to an AR process (Diebold and Rudebusch 1991). The augmented Dickey-Fuller also suffers from a loss of power, particularly if the number of lags increases too quickly given the number of observations (Hassler and Wolters 1994). And while the KPSS test of Kwiatkowski, Phillips, Schmidt, and Shin (1992) performs better, it loses power as the value of $d$ gets closer to the 0.5 threshold.

Figure A.4 provides results of Monte Carlo simulations for the classic and augmented

Dickey-Fuller, KPSS with and without lags, and the variance ratio test. The critical values used are for series with 100 observations and the results are based on 5,000 simulations for each fractional order of $d$ in increments of 0.05.

Beginning with the classic Dickey-Fuller test (dotted line), the test loses power as $d$ decreases, rejecting the null hypothesis of a unit root over 70% of the time when $d = .65$. The ADF test with four lags is slower to reject the null, but this leads to underrejection when the series is on the stationary side of 0.5. The range of $.4 < d < .65$ is particularly troublesome for the ADF and DF tests. The augmented KPSS suffers from some of the same issues - in the stationary, long memory region $0.3 < d < 0.5$, the rejection rate of the null hypothesis creeps upwards over 40% and is even higher were we to not include any lags. Interestingly, the KPSS with four lags demonstrates an unwillingness to reject the null of stationarity even when the fractional order of integration is higher. With an estimated $I(d)$ of 0.8, the KPSS will only reject the null of stationarity 70% of the time.[3] The variance ratio test performs similarly to the augmented Dickey-Fuller.

**Figure A.4: Power Curve**




Each panel illustrates the power function of the unit root test in terms of rejecting the null hypothesis, for each fractional order of integration $d$. Null hypothesis of ADF and Variance Ratio is of unit root. Null hypothesis of KPSS is stationarity. Solid lines for ADF and KPSS indicate tests with 4 lags; dotted lines indicate no lags used. Solid line in Variance Ratio is differencing interval of 4; dotted line represents differencing interval of 8. $T$=100.

The lack of power of unit root tests when applied to fractional series is all the more problematic because there are consequences on both sides of the outcome. With fractional series, unit root tests demonstrate a propensity to both under- and over-reject. If we are modeling a series that is fractional but we treat its level of integration as integer-valued only,

[3]The inclusion of four lags was an arbitrary decision, and is normally dictated by an information criterion. But, one must be aware of the potential threat of too many lags. Hassler and Wolters (1994) use an ADF on series of $T = 100$ with up to 12 lags and find that as the number of lags increase the power of the ADF diminishes considerably. With a fractional order of integration of $d = 0.3$, the ADF with ten lags will only reject the null of a unit root 22.9% of the time.

then failing to reject the null of an ADF or variance ratio test means we must difference the series. In so doing, we over-difference the series and build an MA parameter into the series (Box-Steffensmeier and Smith 1998; Dickinson and Lebo 2007). If we instead reject the null of the ADF or variance ratio test, especially when $d > 0.5$, then we leave a substantial amount autocorrelation unaddressed. Failing to either recognize or account for this autocorrelation invites spurious regressions (Lebo, Walker, and Clarke 2000; Tsay and Chung 2000). Given these choices - a significant MA parameter built into the process on the one hand, or spurious regressions on the other - it behooves the researcher to test for the fractional integration of the series in question. Even if only to serve as an additional unit root test, the estimation of the long memory parameter is useful and can assist in hypothesis testing.

## A.2 Generating Fractionally Integrated Series

We next turn to the question of how fractional integration in time series comes about. We discuss two processes which have been shown to generate fractionally integrated series. The first is from Granger (1980) and covers the micro-aggregation of heterogeneous AR series. The second is the error duration model of Parke (1999). Because Granger (1980) references fractional integration in the frequency domain, rather than the time domain, we briefly discuss their equivalence.

### A.2.1 Equivalence of Time and Frequency Domain

Recall that Equation A.9 defined the autocorrelation function of an ARFIMA process in the time domain. The autocorrelations can also be stated in the frequency domain. To do so, we first define the spectral density function of $y_t$, $f_y(\lambda)$, as

$$\gamma_k = \int_{-\pi}^{\pi} f_y(\lambda) e^{i\lambda k} d\lambda \tag{A.12}$$

in which $\gamma_k$ is the $k$'th autocovariance of $y_t$.

The spectral density of the ARFIMA $(p, d, q)$ process of Equation A.1 can be given by

$$f_y(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-2d} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \tag{A.13}$$

$$= \frac{\sigma^2}{2\pi} (2 \sin\lambda/2)^{-2d} \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \tag{A.14}$$

in which the ARMA $(p, q)$ parameters are represented as

$$f_u(\lambda) = \frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2} \frac{\sigma^2}{2\pi} \tag{A.15}$$

and the transfer function of the fractional filter $(1 - L)^d$ is approximated as

$$|1 - e^{i\lambda}|^{-2d} = (2 \sin\lambda/2)^{-2d} \sim |\lambda|^{-2d} \text{ as } \lambda \to 0, \tag{A.16}$$

Thus, if we assume that $f_u(\lambda)$ is a constant as $\lambda \to 0$, then $f_u(\lambda) \to g$. Which means that the approximation of the autocorrelation function in Equation A.9 can be restated in the frequency domain as the approximation of the spectral density function (Granger and Joyeux 1980; Beran 1994)

$$f_y(\lambda) \sim g|\lambda|^{-2d},\ 0 < g < \infty, \ \text{ as } \lambda \to 0. \tag{A.17}$$

### A.2.2 Aggregation

As previously mentioned, there are a number of examples in political science in which time series are generated as aggregates. Granger (1980) noted that the same was true of economics: income, consumption, unemployment, and profits are all examples of micro-components aggregated into macro series. From this observation, Granger (1980) proposed a model in which the stationary AR(1) components series were drawn from a beta distribution and aggregated into a series that exhibited long memory.

Granger offers an example of two time series

$$x_{jt} = \phi_j x_{j,t-1} + \epsilon_{jt}, \quad j = 1, 2, ... \tag{A.18}$$

and $\epsilon_{jt}$ are zero-mean, independent white noise shocks so that the aggregated variable is the sum of the two series

$$\hat{x} = x_{1t} + x_{2t}, \tag{A.19}$$

which obeys an ARMA(2,1) process, the AR lag polynomial is $(1 - \phi_1 L)(1 - \phi_2 L)$. As Granger notes, if there are $N$ independent AR(1) series, their sum will be ARMA$(N, N-1)$, however the order of the model could be reduced by cancellation of roots.

Assuming the component series forming the aggregate are independent of each other, the component AR(1) $x_{jt}$ has a spectral density function that is approximated by

$$f_j(\lambda) = \frac{1}{|(1 - e^{i\lambda})|^2} \frac{\sigma^2_{\epsilon_j}}{2\pi} \tag{A.20}$$

and the spectrum of $\hat{x}$ is

$$f(\lambda) = \sum_{j-1}^{N} f_j(\lambda) \tag{A.21}$$

Granger assumes that the component series are randomly drawn from a population of the form of a beta distribution on the range of (0,1). The distribution used by Granger is

$$dF(\phi) = \frac{2}{B(u,v)} \phi^{2u-1}(1 - \phi^2)^{v-1} d\phi,\ 0 < \phi < 1, \tag{A.22}$$

where $u > 0$, $v > 0$ and B(u,v) is the beta function, which is given as

$$\frac{2}{B(u,v)} \int_0^1 \phi^{2u+k-1}(1 - \phi)^{v-2} d\phi \tag{A.23}$$

From this, the $k$th autocovariance, $\gamma_k$ is drawn from the Fourier expansion of the spectrum. Assuming $v > 1$ is

$$\gamma_k = \frac{B(u + k/2, v - 1)}{B(u,v)} \tag{A.24}$$

$$= \frac{\Gamma(v-1)}{B(u,v)} \frac{\Gamma(u+k/2)}{\Gamma(u+k/2+v-1)} \tag{A.25}$$

Granger then applies Sterling's theorem such that as the number of lags, $k$, increase, we obtain

$$\gamma_k = ck^{1-v} \tag{A.26}$$

where $c$ is a constant.

If we compare this to our approximation in Equation A.9, we find that $1 - v = 2d - 1$ and that $d = 1 - (1/2)v$. Thus, if the range of long memory falls between $0 < d < 1$, then $0 < v < 2$.

A few caveats and clarifications have been added to Granger's aggregation theorem by Chambers (1998), who demonstrates that aggregation of AR(1) processes is not sufficient on its own to generate fractional integration. An aggregated series will only take on the properties of long memory if long memory is also present in one of the underlying micro series. Additionally, the estimated value of the long memory of the aggregate series will be the maximum of the long memory parameters present in the underlying micro series.

## A.3 Error Duration Model

### A.3.1 Autocovariances of FI Error Duration Model

If the process $y_t$ is long memory, then its autocovariances are nonsummable. Parke (1999) demonstrates that the summation of the autocovariances of the error duration model is the sum of the survival probabilities. The $k$-th autocovariance of $y_t$ is

$$\gamma_k = \sigma^2 \sum_{j=k}^{\infty} p_j. \tag{A.27}$$

After normalizing $\sigma^2$ to equal 1, the autocovariances are then a partial summation of the probabilities

$$\begin{aligned} \gamma(1) &= p_1 + p_2 + p_3 + ... \\ \gamma(2) &= p_2 + p_3 + p_4 + ... \\ &\vdots \\ \gamma(k) &= p_k + p_{k+1} + p_{k+1} + ... \end{aligned} \tag{A.28}$$

The sum of the autocovariances is then $\sum_{k=1}^{n} \gamma(k) = p_1 + 2p_2 + 3p_3 ...$

$$\sum_{k=1}^{n} kp_k \tag{A.29}$$

## Bibliography

Beran, Jan. 1994. *Statistics for Long-Memory Processes.* Vol. 61 CRC Press.

Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics ofAggregate Partisanship." *American Political Science Review* 90: 567–580.

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

Chambers, Marcus J. 1998. "Long Memory and Aggregation in Macroeconomic Time Series." *International Economic Review* pp. 1053–1072.

De Boef, Suzanna. 2001. "Modeling equilibrium relationships: Error correction models with strongly autoregressive data." *Political Analysis* 9(1): 78–94.

Dickinson, Matthew J., and Matt J. Lebo. 2007. "Reexamining the Growth of the Institutional Presidency, 1940–2000." *Journal of Politics* 69(1): 206–219.

Diebold, Francis X, and Glenn D Rudebusch. 1991. "On the Power of Dickey-Fuller Tests Against Fractional Alternatives." *Economics letters* 35(2): 155–160.

Enders, Walter. 2004. *Applied Econometric Time Series.* 2nd ed. New York: Wiley.

Granger, Clive WJ. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of econometrics* 14(2): 227–238.

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

Hassler, Uwe, and Jürgen Wolters. 1994. "On the Power of Unit Root Tests Against Fractional Alternatives." *Economics Letters* 45(1): 1–5.

Hosking, Jonathan RM. 1981. "Fractional Differencing." *Biometrika* 68(1): 165–176.

Kwiatkowski, Denis, Peter CB Phillips, Peter Schmidt, and Yongcheol Shin. 1992. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of econometrics* 54(1): 159–178.

Lebo, Matthew J, Robert W Walker, and Harold D. Clarke. 2000. "You Must Remember This: Dealing With Long Memory in Political Analyses." *Electoral Studies* 19(1): 31–48.

Parke, William R. 1999. "What is Fractional Integration?" *Review of Economics and Statistics* 81(4): 632–638.

Patterson, Kerry. 2012. *Unit Root Tests in Time Series Volume 2: Extensions and Developments.* Vol. 2 Palgrave Macmillan.

Tsay, Wen-Jay, and Ching-Fan Chung. 2000. "The Spurious Regression of Fractional Integrated Processes." *Journal of Econometrics* 96: 155–182.

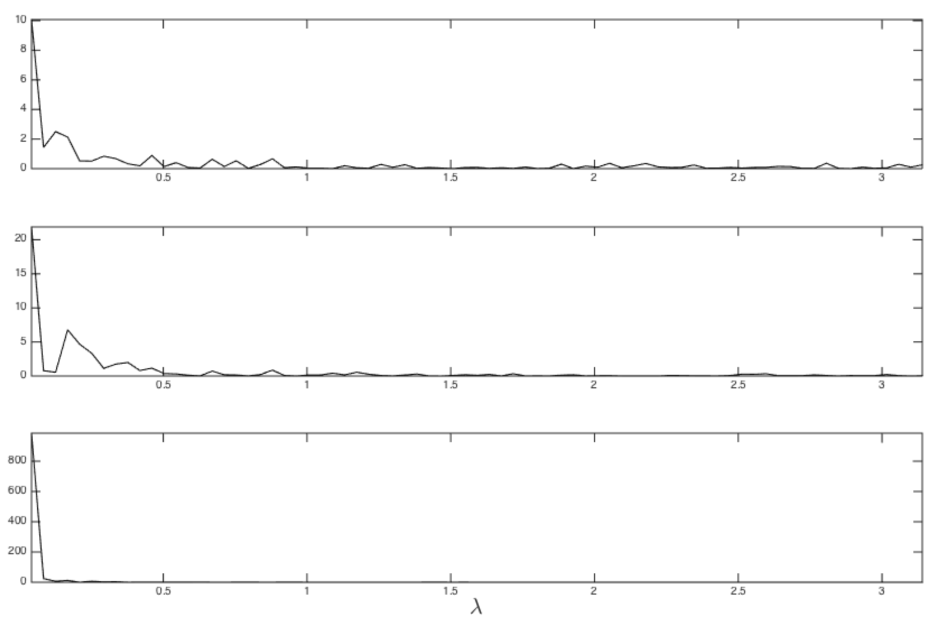
# Appendix B

# Appendix: Chapter 3

## B.1 The Sample Periodogram as a Diagnostic Tool

When using the frequency domain to investigate fractional integration, the focus is on the low-frequencies - those closest to the spectral pole at zero. Granger (1966) demonstrated that all macroeconomic variables tended to exhibit a typical spectral shape. Considering that the spectral density function, when summed across all frequencies, approximates the variance of the series in question, Granger noted that macroeconomic series provided the greatest contribution to the variance of the series in the frequency bands close to zero. If one plots the spectral density of a fractionally integrated series, or its approximation through a periodogram, the density will increase dramatically as the frequency approaches zero. Granger's finding suggested that most series in economics were persistent or long memory.

**Figure B.1: Periodogram of ARFIMA (0,*d*,0) Series**

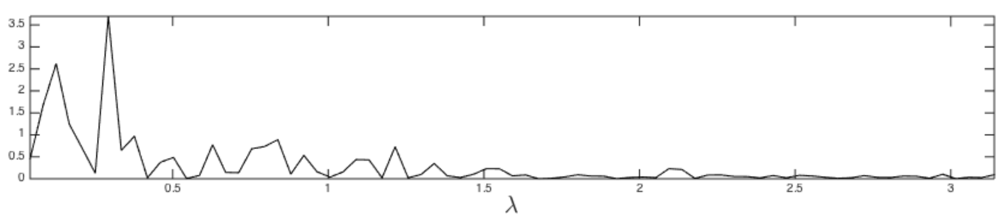



Periodogram, approximation of the spectral density for series of various orders of fractional integration. Top: $I(d) = 0.45$; Middle: $I(d) = 0.65$; Bottom: $I(d) = 0.85$. Frequencies ($\lambda$) range from 0 to $\pi$

The typical spectral shape of fractionally integrated series can be found in Figure B.1, which presents the sample periodograms, approximations of the spectral densities of the three series, with varying degrees of fractional memory - $I(d) = 0.45, 0.65, 0.85$. The increasing contribution to the variance as the spectral density approaches the pole is easy to see in each plot, but one can also see from the $y$-axis that the spectral power increases much faster as the order of fractional integration increases. All three plots indicate that the higher frequencies exert very little influence over the variance of the series. Both the top and middle plots exhibit a few spikes within the first few frequencies, but there is very little contribution in the same range for the heavily persistent series at bottom.

Transforming a time series and visualizing an estimate of the spectral density through the sample periodogram also provides a nice opportunity to evaluate the series' order of integration visually. Persistent memory in the form of fractional integration will exhibit the strongly increasing density described by Granger (1966) as the frequencies approach the poles, and assuming it's a $(0, d, 0)$ series, the periodogram should exhibit little contamination from the higher frequencies. This is in contrast to a stationary AR series - the autoregression in the series is a short-run dynamic that is evident in spikes in the higher frequencies of the periodogram in Figure B.2.

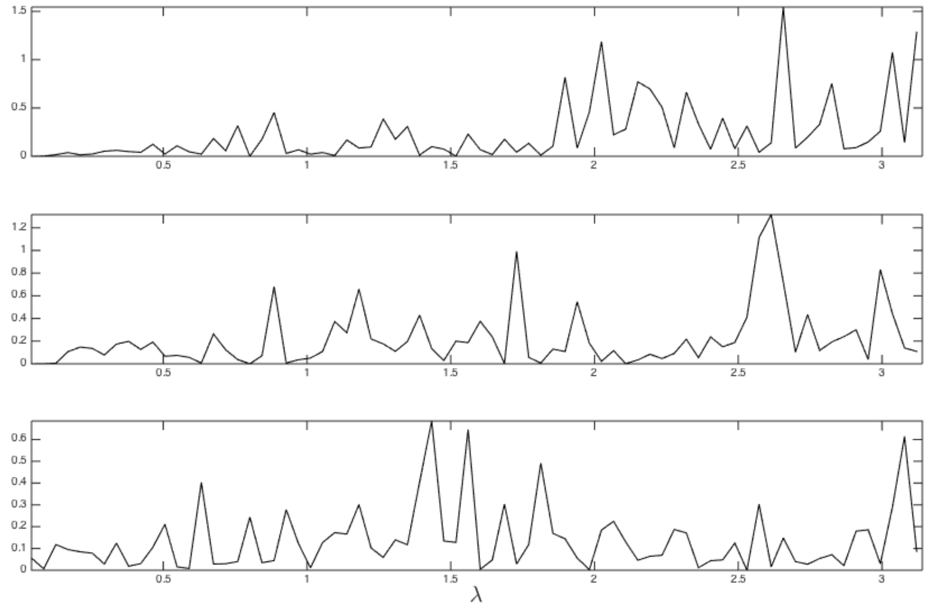
**Figure B.2: Periodogram of Stationary ARIMA (1,0,0) Series**




Periodogram, approximation of the spectral density of a stationary AR series: $y_t = .7x + \epsilon_t$, $N(0, 1)$. Frequencies ($\lambda$) range from 0 to $\pi$

The visual diagnostics can be taken a step further by viewing the series after they have been first differenced. If the series are fractionally integrated, integer differencing will over difference the series. The effects of this over differencing will be apparent visually in the periodogram. For example, if the original series in its level form has a fractional order of integration of $I(d) = 0.65$, then integer differencing will result in a series that is $I(d) = -0.35$. The series is anti-persistent, which means that the the spectrum will exhibit zero to low power at those lower frequencies while spiking in the higher. After differencing, the variance is completely driven by short-run dynamics that have been built into the process. A demonstration of the consequences of over differencing a fractionally integrated series can be found in Figure B.3. For this example, the level form series were the same as those from Figure B.1, from top to bottom $I(d) = 0.45, 0.65, 0.85$, respectively. The lack of power in the spectrum at the lower frequencies is clearly evident in all three plots, however the effects are most pronounced in the top and middle plots. For a fractionally integrated series that is closer to a unit root (bottom panel, $I(d) = 0.85$), the overall effect is less severe, but still clearly visible - the power of the spectrum is concentrated in the middle frequencies, with the lower frequencies exhibiting a minimal contribution to the overall variance of the series.

**Figure B.3: Periodogram of First Differenced ARFIMA (0,*d*,0) Series**

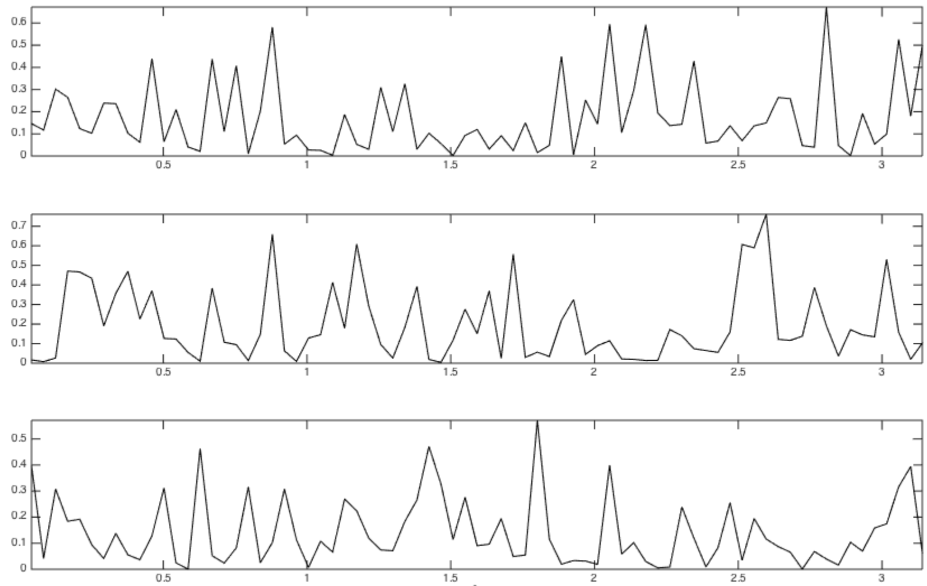


Periodogram, approximation of the spectral density for fractional series (Top-Bottom: $d = 0.45, 0.65, 0.85$) that have been integer differenced.

One of the benefits of viewing a time series in the frequency domain is that one can actually visualize the frequency at which the variance of the process is coming from. In order to achieve an estimation with reliable results, the goal of any researcher with any time series should be to pre-whiten the series such that it is a white-noise process - a process that is uncorrelated, has a zero mean, and possesses the same finite variance. In the frequency domain this means that the spectral density of the process is constant across all frequencies, and all frequencies contribute equally to the variance. By plotting the periodogram, the researcher can easily see if this goal has been achieved.

In the case of the fractionally integrated series however, the process of integer differencing clearly does not work; the low frequencies contribute nothing to the variance of the process. This is why the goal of fractional integration is to move away from integer differencing and to instead difference the series by its own order of fractional integration. If this estimation of fractional integration can be done consistently and efficiently, then the researcher is more likely to approximate the ideal of a white noise process than by simply differencing the series. Relying again on those three series above, the order of fractional integration for each was estimated and each was fractionally differenced by its estimate. Figure B.4 presents the spectral density approximations for the three fractionally differenced series. The overall contribution across frequencies is somewhat uneven given the limited number of observations used for these simulated series, but each exhibits a fairly consistent level of power across the spectrum. This consistency is quite clear when compared to the spectra of the integer differenced series in Figure B.3.

**Figure B.4: Periodogram of Fractionally Differenced ARFIMA (0,*d*,0) Series**




Periodogram, approximation of the spectral density for fractional series (Top-Bottom: $d = 0.45, 0.65, 0.85$) that have been differenced by their own order of fractional integration.

## B.2 Estimating the ARMA Parameters

While the main text focused on the estimation of the long memory parameter, $d$, it is also important for the parametric estimators to be able to estimate the ARMA $(p, q)$ parameters accurately. If the estimation of the short run dynamics is biased, then one cannot safely make inferences about the data generating process. Baillie and Kapetanios (2009) has demonstrated that when using semiparametric methods that the two-step process to first estimate $d$ and then the ARMA parameters suffers from bias. The results below in Tables B.1 and B.2 present the estimates of the EML and FML estimators, which estimate all parameters $(d, \phi, \theta)$ simultaneously.

As can be seen, the FML is a better estimate of $\phi$, which is consistent with the overall results in the main paper. The only time that the FML really struggles is when the AR noise is incredibly persistent and the order of fractional integration is $d = -.25$. The EML results are far better than those of its estimation of $d$, but the estimator still suffers in the presence of moderate AR dynamics. The EML estimates of $\phi$ are also biased when the model is misspecified. Overfitting the ARFIMA model will increase the probability of Type I errors when attempting to fit an AR parameter.

The estimation of the MA parameter is far more accurate. The EML is inconsistent when the number of observations is small, but the estimation is unbiased when at least 60 observations are in the series. The RMSE is wildly inconsistent in the EML estimations and this has to do with the presence of outliers drastically inflating the overall RMSE. The FML estimator is a consistent unbiased estimator of the MA parameter.

Table B.1: Parametric Estimation of AR Parameter - ARFIMA (1,$d$,0)

| | | | EML | | FML | | | | | EML | | FML | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | −0.25 | 40 | .1410 | .3157 | .0425 | .2089 | .40 | −0.25 | 40 | .1351 | .2874 | −.1094 | .2982 |
| | | 60 | .1091 | .2586 | .0290 | .1651 | | | 60 | .1472 | .2866 | −.0727 | .2708 |
| | | 80 | .0831 | .2044 | .0274 | .1453 | | | 80 | .1606 | .2807 | −.0353 | .2387 |
| | | 100 | .0536 | .1643 | .0237 | .1263 | | | 100 | .1468 | .2153 | −.0129 | .2114 |
| | 0 | 40 | .2190 | .4195 | .0785 | .2508 | | 0 | 40 | .1922 | .3073 | −.0544 | .2887 |
| | | 60 | .1076 | .2601 | .0435 | .1916 | | | 60 | .1775 | .3026 | −.0388 | .2758 |
| | | 80 | .0792 | .2014 | .0398 | .1532 | | | 80 | .1696 | .2896 | .0006 | .2480 |
| | | 100 | .0590 | .1741 | .0267 | .1397 | | | 100 | .1789 | .2880 | .0078 | .2420 |
| | 0.25 | 40 | .2735 | .4981 | .1145 | .2967 | | 0.25 | 40 | .2792 | .3318 | .0351 | .2595 |
| | | 60 | .1504 | .3152 | .0540 | .2018 | | | 60 | .2863 | .3341 | .0327 | .2546 |
| | | 80 | .1014 | .2214 | .0313 | .1642 | | | 80 | .2676 | .3279 | .0485 | .2460 |
| | | 100 | .0681 | .1699 | .0324 | .1394 | | | 100 | .2589 | .3227 | .0489 | .2395 |
| | 0.45 | 40 | .3426 | .5597 | .1182 | .2753 | | 0.45 | 40 | .3472 | .3691 | .1734 | .3032 |
| | | 60 | .1964 | .3588 | .0811 | .2245 | | | 60 | .3597 | .3820 | .1717 | .2875 |
| | | 80 | .1241 | .2498 | .0600 | .1576 | | | 80 | .3445 | .3720 | .1568 | .2671 |
| | | 100 | .0886 | .1821 | .0364 | .1299 | | | 100 | .3339 | .3649 | .1520 | .2652 |
| 0 | −0.25 | 40 | .2745 | .4428 | .0084 | .2503 | .80 | −0.25 | 40 | −.1144 | .2281 | −.3924 | .4973 |
| | | 60 | .2419 | .4207 | .0121 | .2216 | | | 60 | −.0979 | .2203 | −.3668 | .4744 |
| | | 80 | .2064 | .3849 | .0213 | .1967 | | | 80 | −.0933 | .2236 | −.3394 | .4469 |
| | | 100 | .1542 | .3311 | .0214 | .1730 | | | 100 | −.0268 | .2222 | −.3198 | .4265 |
| | 0 | 40 | .3571 | .5196 | .0686 | .3081 | | 0 | 40 | −.0400 | .1260 | −.1228 | .2633 |
| | | 60 | .2938 | .4727 | .0536 | .2785 | | | 60 | −.0226 | .1098 | −.0767 | .2104 |
| | | 80 | .2414 | .4289 | .0670 | .2510 | | | 80 | −.0235 | .1164 | −.0747 | .1979 |
| | | 100 | .1829 | .3682 | .0444 | .2235 | | | 100 | −.0261 | .1219 | −.0567 | .1737 |
| | 0.25 | 40 | .4363 | .5730 | .1321 | .3554 | | 0.25 | 40 | −.0080 | .0904 | −.0581 | .1844 |
| | | 60 | .3516 | .5108 | .0953 | .3164 | | | 60 | .0145 | .0784 | −.0317 | .1565 |
| | | 80 | .2697 | .4272 | .0633 | .2669 | | | 80 | .0105 | .0837 | −.0318 | .1469 |
| | | 100 | .2132 | .3913 | .0698 | .2623 | | | 100 | .0172 | .0767 | −.0363 | .1443 |
| | 0.45 | 40 | .5745 | .6666 | .1991 | .1465 | | 0.45 | 40 | .0285 | .0781 | −.0064 | .1450 |
| | | 60 | .5377 | .6551 | .1581 | .3473 | | | 60 | .0419 | .0763 | .0085 | .1264 |
| | | 80 | .4362 | .5877 | .1432 | .3189 | | | 80 | .0504 | .0769 | .0129 | .1137 |
| | | 100 | .3485 | .5091 | .0974 | .2582 | | | 100 | .0486 | .0740 | .0190 | .1043 |

*Note*: Results based on 1000 simulations for each model.

## Table B.2: Parametric Estimation of MA Parameter - ARFIMA $(0,d,1)$

| | | | EML | | FML | | | | | EML | | FML | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\theta$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** | $\theta$ | $d$ | $T$ | **Bias** | **RMSE** | **Bias** | **RMSE** |
| −.40 | −0.25 | 40 | −.3049 | 2.2903 | .0424 | .3277 | .40 | −0.25 | 40 | −.0202 | 1.7014 | .0207 | .2750 |
| | | 60 | −.0885 | .8287 | .0202 | .2701 | | | 60 | .0450 | .1501 | .0102 | .1990 |
| | | 80 | −.0522 | .4541 | .0188 | .2616 | | | 80 | .0387 | .1422 | .0056 | .1559 |
| | | 100 | −.0328 | .3231 | .0125 | .2403 | | | 100 | .0325 | .1254 | −.0012 | .1292 |
| | 0 | 40 | .0244 | 1.7604 | −.1179 | .3371 | | 0 | 40 | −.2069 | 3.6457 | −.0056 | .2632 |
| | | 60 | .1320 | .2601 | −.0881 | .2920 | | | 60 | .0664 | .1721 | .0055 | .2042 |
| | | 80 | .1819 | .2014 | −.0629 | .2732 | | | 80 | .0431 | .1445 | −.0046 | .1564 |
| | | 100 | .1327 | .1741 | −.0403 | ..2541 | | | 100 | .0318 | .1242 | .0042 | .1339 |
| | 0.25 | 40 | −.2446 | 3.5956 | −.1989 | .3811 | | 0.25 | 40 | −.0730 | 2.9927 | −.0309 | .2446 |
| | | 60 | .1535 | 1.2568 | −.1055 | .2956 | | | 60 | .0793 | .2555 | −.0046 | .1843 |
| | | 80 | .2017 | .4602 | −.0752 | .2640 | | | 80 | .0604 | .1386 | .0002 | .1544 |
| | | 100 | .2531 | 1.5526 | −.0595 | .2448 | | | 100 | .0404 | .1216 | −.0009 | .1299 |
| | 0.45 | 40 | −.3905 | 4.5139 | −.2206 | .3588 | | 0.45 | 40 | −.1672 | 4.8689 | −.0412 | .1998 |
| | | 60 | .3204 | 2.7377 | −.1605 | .2878 | | | 60 | .1240 | .3627 | −.0221 | .1617 |
| | | 80 | .3364 | 2.6146 | −.1292 | .2447 | | | 80 | .0853 | .2235 | −.0212 | .1293 |
| | | 100 | .2862 | 1.8104 | −.1012 | ..2098 | | | 100 | .0694 | .1966 | −.0097 | .1098 |
| 0 | −0.25 | 40 | −.3544 | 3.3413 | .0523 | .3512 | .80 | −0.25 | 40 | −.0290 | 1.9641 | −.0385 | .1913 |
| | | 60 | .0602 | .2971 | .0453 | .2988 | | | 60 | .0316 | .1113 | −.0172 | .1321 |
| | | 80 | .0594 | .2335 | .0230 | .2514 | | | 80 | .0863 | .0899 | −.0133 | .1045 |
| | | 100 | .0611 | .2080 | .0135 | .2155 | | | 100 | .0219 | .0789 | −.0109 | .0882 |
| | 0 | 40 | .4140 | 4.0157 | −.0153 | .3275 | | 0 | 40 | .0697 | .1536 | −.0462 | .1883 |
| | | 60 | .0238 | 1.2910 | .0186 | .2824 | | | 60 | .0387 | .1156 | −.0289 | .1432 |
| | | 80 | .0739 | .3341 | .0002 | .2391 | | | 80 | .0276 | .0956 | −.0186 | .1098 |
| | | 100 | .0804 | .1864 | .0086 | .2112 | | | 100 | .0268 | .0844 | −.0229 | .0929 |
| | 0.25 | 40 | −.3294 | 5.2329 | −.0723 | .3049 | | 0.25 | 40 | .0832 | .1733 | −.0512 | .1945 |
| | | 60 | .1163 | 1.9100 | −.0198 | .2383 | | | 60 | .0524 | .1328 | −.0415 | .1497 |
| | | 80 | .2020 | 1.7598 | −.0116 | .2075 | | | 80 | .0383 | .1031 | −.0353 | .1222 |
| | | 100 | .1451 | 1.3400 | −.0086 | .1842 | | | 100 | .0200 | .0776 | −.0317 | .1053 |
| | 0.45 | 40 | −.2818 | 5.4418 | −.1060 | .2578 | | 0.45 | 40 | .1213 | .2016 | −.0814 | .1932 |
| | | 60 | .2720 | 4.2916 | −.0719 | .2048 | | | 60 | .0928 | .1685 | −.0674 | .1636 |
| | | 80 | .7432 | 5.5890 | −.0632 | .1697 | | | 80 | .0595 | .1379 | −.0607 | .1351 |
| | | 100 | .3211 | 3.4475 | −.0430 | .1459 | | | 100 | .0431 | .1065 | −.0545 | .1181 |

*Note*: Results based on 1000 simulations for each model.

## B.3 Ignoring Fractional Integration - ARMA Estimation

Previous work by Box-Steffensmeier and Smith (1998) has demonstrated that assumptions made as to the stationarity properties of the data, and ignoring fractional integration, can have significant consequences on estimates as well as the conclusions drawn. In this instance, we assume that a researcher is interested in pre-whitening their data before estimation. In such a case, how reliable are the estimates of the AR and MA parameters if a series is, in fact, fractionally integrated but the long memory is ignored? The results indicate a significant bias in the estimates, determined by the level of memory in the series - upward bias if the series exhibits positive memory, negative if the series is anti-persistent. The estimates can be found in the Monte Carlo estimates for each model below. All results are based on 1000 simulations for each model. The orders of fractional integration present ranges between [-.25, 0, .25, .45] so all series are considered stationary despite being long memory.

Addressing Table B.3, which presents results of an autoregressive process, we find that a failure to account for the fractional integration of the series leads to biased estimates of the AR parameter. The greater the extent of fractional integration, the greater the positive bias, while an anti-persistent series will negatively bias our estimates of $\phi$. This bias is apparent even in series that do not suffer from autoregression - the unaccounted for fractional integration inflates the estimated AR parameter.

**Table B.3: ARMA Estimation - (1,$d$,0) Model**

| $\phi$ | $d$ | $T$ | **Bias** | **RMSE** | $\phi$ | $d$ | $T$ | **Bias** | **RMSE** |
|---|---|---|---|---|---|---|---|---|---|
| −.40 | −0.25 | 40 | −.1144 | .1635 | .40 | −0.25 | 40 | −.2661 | .3043 |
| | | 60 | −.1251 | .0254 | | | 60 | −.2574 | .2833 |
| | | 80 | −.1180 | .0216 | | | 80 | −.2577 | .2801 |
| | | 100 | −.1265 | .0218 | | | 100 | −.2504 | .2669 |
| | 0 | 40 | .0079 | .1414 | | 0 | 40 | −.0595 | .1611 |
| | | 60 | .0033 | .1139 | | | 60 | −.0376 | .1259 |
| | | 80 | .0090 | .0991 | | | 80 | −.0250 | .1035 |
| | | 100 | .0055 | .0917 | | | 100 | −.0215 | .0935 |
| | 0.25 | 40 | .1896 | .2584 | | 0.25 | 40 | .1575 | .2091 |
| | | 60 | .2044 | .2537 | | | 80 | .1920 | .2226 |
| | | 80 | .2164 | .2555 | | | 80 | .2074 | .2300 |
| | | 100 | .2180 | .2486 | | | 100 | .2152 | .2317 |
| | 0.45 | 40 | .4161 | .4667 | | 0.45 | 40 | .3050 | .3275 |
| | | 60 | .4619 | .4997 | | | 60 | .3455 | .3573 |
| | | 80 | .4936 | .5222 | | | 80 | .3659 | .3753 |
| | | 100 | .5143 | .5362 | | | 100 | .3841 | .3905 |
| 0 | −0.25 | 40 | −.2059 | .2481 | .80 | −0.25 | 40 | −.2807 | .3164 |
| | | 60 | −.1947 | .2242 | | | 60 | −.2538 | .2777 |
| | | 80 | −.1982 | .2227 | | | 80 | −.2482 | .2674 |
| | | 100 | −.1958 | .2146 | | | 100 | −.2353 | .2512 |
| | 0 | 40 | −.0247 | .1555 | | 0 | 40 | −.0899 | .1563 |
| | | 60 | −.0171 | .1270 | | | 60 | −.0611 | .1126 |
| | | 80 | −.0150 | .1145 | | | 80 | −.0498 | .0964 |
| | | 100 | −.0088 | .0960 | | | 100 | −.0319 | .0744 |
| | 0.25 | 40 | .2009 | .2634 | | 0.25 | 40 | .0411 | .0983 |
| | | 60 | .2316 | .2696 | | | 60 | .0717 | .0980 |
| | | 80 | .2468 | .2766 | | | 80 | .0869 | .1005 |
| | | 100 | .2629 | .2874 | | | 100 | .0923 | .1021 |
| | 0.45 | 40 | .4220 | .4582 | | 0.45 | 40 | .1099 | .1303 |
| | | 60 | .4682 | .4893 | | | 60 | .1316 | .1393 |
| | | 80 | .4964 | .5116 | | | 80 | .1416 | .1458 |
| | | 100 | .5149 | .5285 | | | 100 | .1498 | .1524 |

*Note*: Results based on 1000 simulations for each model. Estimated bias and RMSE are of estimates of the AR parameter in a $(1, d, 0)$ model. Processes are generated such fractional integration of order $d$ is present, but not estimated.

We find the exact same issues when attempting to estimate a series with an MA parameter. The persistence attributed to fractional integration will bias the MA parameter in the expected direction. Note that the bias dissipates as the short-term dynamics increase in strength. Positive values of fractional integration exert far less influence on the estimates when the AR or MA parameters are .80. One additional point worth noting is that for series with unaccounted for fractional memory, $0 < d < 1/2$, the bias in our estimates actually worsens with an increase in sample size.

## Table B.4: ARMA Estimation - (0,*d*,1) Model

| $\theta$ | $d$ | $T$ | Bias | RMSE | $\theta$ | $d$ | $T$ | Bias | RMSE |
|---|---|---|---|---|---|---|---|---|---|
| −.40 | −0.25 | 40 | −.3805 | .5169 | .40 | −0.25 | 40 | −.1979 | .3254 |
| | | 60 | −.2959 | .3485 | | | 60 | −.1932 | .2610 |
| | | 80 | −.2680 | .2966 | | | 80 | −.1929 | .2414 |
| | | 100 | −.2616 | .2807 | | | 100 | −.1887 | .2252 |
| | 0 | 40 | .0079 | .1414 | | 0 | 40 | −.0595 | .1611 |
| | | 60 | .0033 | .1139 | | | 60 | −.0376 | .1259 |
| | | 80 | .0090 | .0991 | | | 80 | −.0250 | .1035 |
| | | 100 | .0055 | .0917 | | | 100 | −.0215 | .0935 |
| | 0.25 | 40 | .1878 | .2718 | | 0.25 | 40 | .1491 | .2241 |
| | | 60 | .2191 | .2609 | | | 80 | .1567 | .1930 |
| | | 80 | .2407 | .2682 | | | 80 | .1575 | .1831 |
| | | 100 | .2376 | .2590 | | | 100 | .1602 | .1804 |
| | 0.45 | 40 | .3906 | .4448 | | 0.45 | 40 | .2587 | .3027 |
| | | 60 | .4298 | .4524 | | | 60 | .2649 | .2823 |
| | | 80 | .4479 | .4622 | | | 80 | .2724 | .2851 |
| | | 100 | .4573 | .4692 | | | 100 | .2781 | .2876 |
| 0 | −0.25 | 40 | −.2852 | .3860 | .80 | −0.25 | 40 | −.0756 | .2376 |
| | | 60 | −.2776 | .3317 | | | 60 | −.0936 | .1883 |
| | | 80 | −.2621 | .2981 | | | 80 | −.0820 | .1424 |
| | | 100 | −.2617 | .2892 | | | 100 | −.0850 | .1359 |
| | 0 | 40 | −.0247 | .1555 | | 0 | 40 | −.0899 | .1563 |
| | | 60 | −.0171 | .1270 | | | 60 | −.0611 | .1126 |
| | | 80 | −.0150 | .1145 | | | 80 | −.0498 | .0964 |
| | | 100 | −.0088 | .0960 | | | 100 | −.0319 | .0744 |
| | 0.25 | 40 | .1881 | .2593 | | 0.25 | 40 | .2241 | .2405 |
| | | 60 | .2077 | .2447 | | | 60 | .1257 | .1420 |
| | | 80 | .2135 | .2379 | | | 80 | .0834 | .0992 |
| | | 100 | .2264 | .2453 | | | 100 | .0715 | .0905 |
| | 0.45 | 40 | .3588 | .3920 | | 0.45 | 40 | .2357 | .2726 |
| | | 60 | .3792 | .3992 | | | 60 | .1344 | .1712 |
| | | 80 | .3883 | .4019 | | | 80 | .0984 | .1443 |
| | | 100 | .4000 | .4105 | | | 100 | .0722 | .1230 |

*Note*: Results based on 1000 simulations for each model. Estimated bias and RMSE are of estimates of the MA parameter in a $(0, d, 1)$ model. Processes are generated such fractional integration of order $d$ is present, but not estimated.

# Bibliography


Baillie, Richard T, and George Kapetanios. 2009. Semi Parametric Estimation of Long Memory: Comparisons and Some Attractive Alternatives. In *conference in honor of Peter Phillip's Sixtieth Birthday at Singapore Management University*.

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

Granger, Clive WJ. 1966. "The Typical Spectral Shape of an Economic Variable." *Econometrica: Journal of the Econometric Society* pp. 150–161.

# Appendix C

# Appendix: Chapter 4

## C.1 Additional Details on the Derivation of ECM Models

In this section we provide more information on the origins of the ECM model that DeBoef and Keele (2008) adapt from the econometrics literature.

Kremers, Ericsson, and Dolado (1992) were the first to derive the asymptotic distribution of a conditional ECM $t$-test with a pre-specified cointegrating vector. Importantly, they found that the distribution of the ECM was dynamic. Assuming an $I(1)$ dependent variable, it would shift based upon the properties of the covariate. The authors also noted the relationship between the ECM and Engle-Granger procedures, which we briefly discuss below.

If we assume a general ECM given as:

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 (y - \delta' z)_{t-1} + \upsilon_{1t} \tag{C.1}$$

$$\Delta z_t = \epsilon_t \tag{C.2}$$

where $y - \delta' z$ is the potential cointegrating relationship, then we can establish the relationship between the ECM and the Engle Granger cointegration test. By subtracting $\delta' \Delta z_t$ from both sides of (C.1) we get:

$$\Delta (y - \delta' z)_t = \gamma_1 (y - \delta' z)_{t-1} + \{(\gamma_o' - \delta') \Delta z_t + \upsilon_{1t}\} \tag{C.3}$$

Next, redefine the Engle-Granger residual, $(y - \delta' z_t)$ as $\omega_t$. From this point, (C.3) can be rewritten as

$$\Delta \omega_t = \gamma_1 \omega_{t-1} + e_t \tag{C.4}$$

where $e_t = (\gamma_0' - \delta') \Delta z_t + \upsilon_{1t}$.

Rewritten in this way, the $t$-test on $\gamma_1$ in (C.4) is equivalent to the test statistic of the Engle-Granger cointegration test. It is the Dickey-Fuller (DF) test of $\omega_t$, determining whether the cointegrating relationship of $y$ and $z$ indicates cointegration (i.e., stationarity).

Note in (C.4) however that the Dickey-Fuller test imposes a common factor restriction – ($\gamma_0 = \delta$) – the short and long run elasticities must be the same. In their work on the general ECM Kremers, Ericsson, and Dolado (1992) argue that, more often than not, such a restriction is invalid. In such instances the ECM $t$-test will be more powerful than the Engle-Granger because it does not impose the same restrictions.

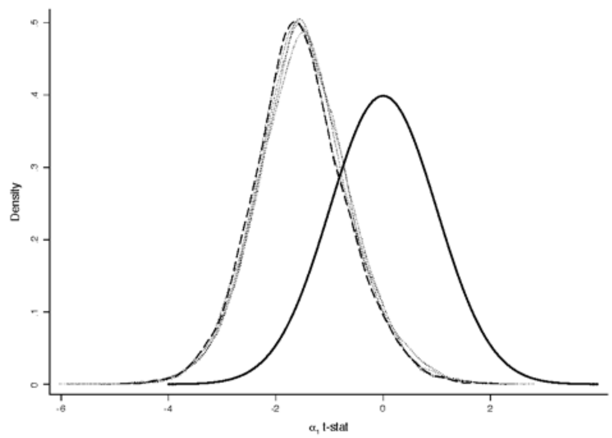
Kremers et al. (1992) derived the distribution of the ECM test under the null hypothesis of no cointegration and found the distribution to be dynamic – it is influenced by the covariate. They found that when the covariate was both (1) stationary and (2) a strong predictor of the dependent variable, the ECM distribution approached the standard normal. The extent to which the distribution would shift away from the Dickey-Fuller distribution was dictated by a nuisance parameter, which they referred to as the "signal-to-noise" ratio. This parameter, specified as $q$, captured the similarity of the error variances in $(e_t = (\gamma_0 - \delta)\Delta z_t + v_{1t})$. The larger the value of $\gamma_0$, the smaller the effect of $e_t$ and the closer the distribution will be to that of the Dickey-Fuller. Conversely, the more influential the stationary covariate, the larger the value of $e_t$ and the ECM distribution would approximate a normal distribution.

Kremers et al. (1992) gave their signal-to-noise ratio as:

$$q = -(\gamma_0 - \delta)s \tag{C.5}$$

where s is the ratio $\sigma_\epsilon/\sigma_v$.

**Figure C.1: Densities of ECM t-statistic With Valid Common Factor Restrictions**




*Note*: Dashed line is the density of the Dickey-Fuller $t$-statistic of the DV. Tight dots are the densities of the ECMs' $t$-statistics with up to 4 IVs included in the model. With $I(0)$ IVs in which $\Delta X_t$ =0, the distributions follows that of the DF. Solid line is the standard normal distribution. Each regression model contains a constant. DV is $I(1)$. $T = 60$

Building off of this work, Hansen (1995) developed the Covariate Augmented Dickey-Fuller (cADF) test which accounted for the information provided by additional stationary covariates. By dropping the common factor restrictions of the standard DF test and instead allowing the use of additional information, Hansen demonstrated a univariate unit-root test with better precision in the estimates, shorter confidence intervals, and increased power of the test statistic, sometimes substantially.[1] With Monte Carlo simulations, Hansen demonstrated that under the null hypothesis of a unit-root, $\delta = 0$, the asymptotic distribution of the cADF test statistic $t(\widehat{\delta}) = \widehat{\delta}/s(\widehat{\delta})$ is a weighted combination of the standard normal and

[1]Caporale and Pittis (1999) investigate the power gains of the cADF and find that the nature of the correlation between error terms of the series being tested and the covariates is determinant. Not only will the addition of a covariate increase the precision of standard errors, but under certain circumstances the coefficient on $Y_{t-1}$ increases in absolute value.

DF distributions, weighted by the included covariate's contribution to the long-run correlation in the error term. The covariate contribution is modeled as a nuisance parameter, $\rho^2$:

$$t(\widehat{\delta}) = \rho(DF) + (1-\rho^2)^{1/2}N(0,1) \tag{C.6}$$

which takes on values in the unit interval such that when when $\rho = 1$, the covariates are uncorrelated and the cADF is equivalent to the Dickey Fuller; as $\rho \Rightarrow 0$ the distribution approximates the standard normal. With the nuisance parameter set on a specific interval, Hansen was able to calculate specific critical values for hypothesis testing.

The asymptotic distribution and critical values of Hansen's cADF rely upon a very strong assumption of stationarity, however, and it will not hold if the assumption is violated. With an $I(1)$ dependent variable and a non-stationary $\Delta X_t$ that is not cointegrated, Hansen finds the distribution is biased away from the normal, approximating the distribution estimated by Phillips and Ouliaris (1990) for the Engle-Granger cointegration test. This distribution is even more biased away from the normal than the Dickey-Fuller distribution. The consequences of violating these stationarity assumptions extend to under-differenced data as well. If a covariate is $I(d)$ where $d > 1$, differencing will not account for all of the auto-correlation and the distribution will be biased away from the normal as a result.

The distribution derived by Hansen dealt with testing for univariate unit-roots, but it was a generalization of Kremers, Ericsson, and Dolado (1992). Whereas the nuisance parameter defined by Kremers et al. could take on any positive value making it difficult to use, Hansen's nuisance parameter was placed on a specific interval. In fact, Zivot (2000) derived the asymptotic distribution of Kremers, Ericsson, and Dolado's (1992) conditional ECM $t$-test for pre-specified cointegrating vectors under more general conditions and found the asymptotic distribution of Hansen's cADF for a unit-root was identical to the asymptotic distribution of the conditional ECM $t$-test.

### C.1.1 Limitations to Estimating the ECM Distribution

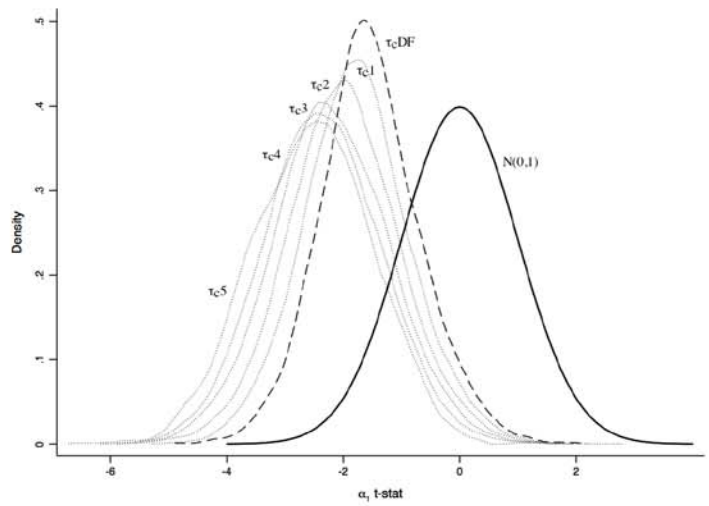
The past research of Kremers, Ericsson, and Dolado (1992), Hansen (1995), and Zivot (2000) determined the distributional properties of the ECM statistic in situations in which the cointegrating coefficients were known, but importantly, these properties only hold under situations of prior knowledge and proper pre-specification.

When the ECM model is used in applied research however, we often have no *a priori* knowledge of the cointegrating vector. We cannot pre-specify the nuisance parameter, dictate our covariate's level of influence, or set the direction of correlation, and these limitations have real consequences as to what assumptions we can make about the ECM $t$-statistic.

For instance, despite Kremers et al. (1992) proving the ECM distribution approximated the normal under certain circumstances, they include a footnote in their paper urging caution when using the ECM test in applied settings. Footnote 5 states that when a researcher has no prior information of the influence of $\Delta X_t$ that it is advisable to use the DF test statistics because they were substantively larger. In other words, meeting the assumption of stationarity was not enough. Unable to specify the signal-to-noise ratio in real life, a researcher cannot guarantee that the ECM's distribution approaches the normal. As a result, relying on the normal distribution for hypothesis testing potentially results in Type I errors. Caution behooves the researcher to use the more conservative distribution.

This distinction between the ability to pre-specify the cointegrating vector and having to estimate the cointegration coefficent is worth considering in greater depth because it has real consequences on how the ECM is used within the field of political science. With imperfect knowledge we are left to test for cointegration by estimation, and in order for our hypothesis tests to be valid we need a reliable set of test statistics. But from the work of multiple lines of research we know that the distribution of the ECM $t$-statistic is non-standard and is not dimension invariant. Kremers, Ericsson, and Dolado (1992) and Hansen (1995) established that when influential, stationary covariates are added, the distribution will shift towards the normal - however, in practice the extent of that shift is generally unknown. Hansen also found that when non-stationary covariates that did not cointegrate were added, the distribution was biased in the negative direction. Banerjee, Dolado, and Mestre (1998) and Ericsson and MacKinnon (2002) used Monte Carlo simulations to compute critical values for the ECM distribution and found that the distribution would shift in the negative with every additional $I(1)$ regressor. The MacKinnon values were computed for the widest number of situations and are widely used today for hypothesis testing for the ECM cointegration test.

**Figure C.2: Densities of ECM t-statistic by Number of Covariates**




*Note*: Density plots of the ECM $t$-statistic for up to 5 IVs included. Each additional IV shifts the distribution further away from the normal. DV and IVs all $I(1)$. $T = 60$

But the MacKinnon critical values are only meaningful if we can also guarantee that all series in the model are $I(1)$. *Thus, if we include a mix of stationary and $I(1)$ regressors, our ECM statistic will no longer fit the distribution, our critical values will be incorrect, and we won't know in which direction our test statistic is biased.* As the examination of Caporale and Pittis (1999) found, the reduction (and potential inflation) of standard errors as well as the possibility of inflated test-statistics depend on a mix of factors, including the direction of Granger causality as well as the contemporaneous correlation of the error terms. Alternatively, if we use an IV that is $I(d) > 1$, we can expect that the ECM statistic will be biased in the negative direction away from the MacKinnon critical values. The stationarity assumptions of Hansen (1995) are valid, and failing to take them into account will cause serious inferential problems.

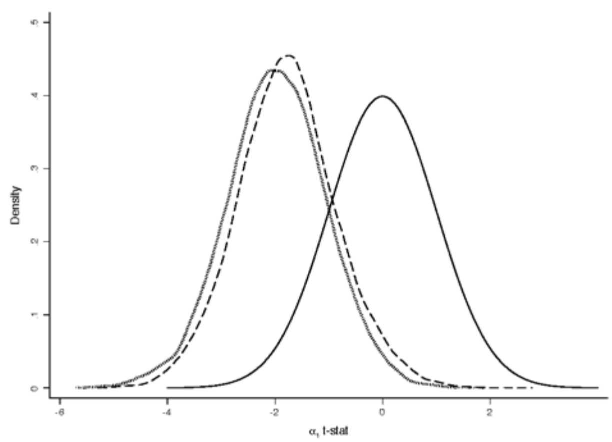
We demonstrate this with a simple bivariate regression using an independent variable from Ura and Ellis (2012). We use their variable, *Defense Spending*, which captures all federal defense spending in 2008 dollars. Estimates of its order of integration using Stata's ML estimator are: $d = 1.36$ (s.e.=0.11).

In its most commonly used form, the Bårdsen transformation, the ECM is given as:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 \Delta X_t + \beta_1 X_{t-1} + \epsilon_t \tag{C.7}$$

Recalling the stationarity concerns of Hansen - independent variables must be stationary, and if in differenced form they should not be under-differenced - it is clear that by including *Defense Spending* in the above model, we are going to bias our ECM $t$-statistic twice. The order of integration of *Defense Spending*$_{\text{t-1}}$, $I(d)$, is much larger than $I(1)$, and $\Delta$*Defense Spending*$_{\text{t}}$ will therefore be under-differenced. We should expect that the distribution of the ECM $t$-statistic is biased away from the normal.

**Figure C.3: Densities of ECM t-statistic with IV d > 1**




*Note*: Dashed line is density of ECM $t$-statistic of $I(1)$ bivariate regression. Mean $t$-statistic = -1.78. Tight dot is $t$-statistic of random $I(1)$ DV regressed on *Defense Spending*. Mean $t$-statistic of $\alpha_1$ = -1.99. Normal density included for comparison. $T = 36$

Figure C.3 presents the density plot of the ECM $t$-statistic after 10,000 simulations of this model. The density plot of the ideal $I(1)$ bivariate regression is included for comparison, and the bias is clearly evident. And while this bias may appear minimal, were we to use MacKinnon critical values we would reject the null of no cointegration in 7.1% of simulations, compared to the 4.7% rejection rate of the ideal regression. This is just one of the many ways the distribution of the test statistic is moved around.

### C.1.2 Testing the Null of No Cointegration

Research has previously investigated the strength of the ECM as a cointegration test. Crucially, these studies all evaluate the distributional properties of the ECM statistic when the cointegrating vectors *are already known* (De Boef and Granato 1999; De Boef 2001; Hansen 1995; Kremers, Ericsson, and Dolado 1992; Zivot 2000). Assuming that the known cointegrating vector is properly specified, such conditional ECM $t$-tests testing the null of no-cointegration are often far more powerful than cointegration tests in which the cointegrating vector must be estimated (Zivot 2000).

In the political science literature, studies that examines the ECM test under the null of no cointegration also pre-specifiy the cointegrating vectors (De Boef and Granato 1999; De Boef 2001). When the data do not contain a cointegrating relationship, the ECM test consistently performs as expected, even with near-integrated data. Similarly, under the alternative hypothesis of a cointegrating relationship, the ECM test performs admirably as the amount of available information increases. Under circumstances such as those tested by DeBoef and Granato, when the cointegrating vector is specifically set either to 0 (no cointegration) or -0.05 (cointegration), the ECM test does not suffer from either Type I or II errors. These findings are consistent with previous research (Kremers, Ericsson, and Dolado 1992; Zivot 2000).

But in actual practice, the cointegrating vector is unknown ahead of time and must be estimated. What is the power of the ECM test under these less than ideal circumstances? Were researchers simply using $I(1)$ data series, the work of Banerjee, Dolado, and Mestre (1998) and Ericsson and MacKinnon (2002) would be sufficient to show that the ECM model performs as it should. But in actual practice we have seen a proliferation of studies using the ECM model with series of all orders of integration, and these unbalanced models have led to all sorts of inferential problems. Additionally, little thought has been paid to the interpretations and conclusions when estimating an ECM equation using a stationary series as the dependent variable. And the models often have other complications that make interpretation even more difficult – e.g. interactions are sometimes included and the ECMs of small $T$ series have been estimated with up to 10 independent variables. And this has all occurred while using the standard one-tail $t$-statistic of -1.645 as the hypothesis test for error correction.

Given the large number of combinations of data series varying by length and order of integration, we set out to broadly simulate data common to political science. Such series tend to be short in duration and are rarely $I(1)$. Based on the work of De Boef and Granato (1999) and De Boef (2001) we also tested the ECM with near-integrated data under the null of no cointegration. When the series were comprised solely of $I(1)$ series, the ECM statistic performs as expected. As we deviate from $I(1)$ series, the power of the ECM model deteriorates rapidly. Type I errors abound and spurious regressions are the norm.

## C.2 Generating Series

All simulations were run in RATS 8.0 although almost all replications were also estimated in Stata to allow readers to more easily replicate our findings. Series were generated to fit a specified order of integration or level of autoregression.

For all simulations we relied on the RATS random number generator and set a specific seed value for replication. The starting value for each series was 0. For $I(1)$ series, the present values were computed as the series' value in the previous period, plus a random error term drawn from the normal distribution with mean of zero and a variance equal to 1:

$$Y_t = Y_{t-1} + \epsilon_t. \tag{C.8}$$

$I(0)$ series were generated simply as a white noise disturbance from a normal distribution with a mean of zero and variance equal to 1:

$$Y_t = \epsilon_t. \tag{C.9}$$

Stationary autoregressive series were generated with a value of $\rho$ that varied by increments of 0.01 between 0.90 and 0.99:

$$Y_t = \rho Y_{t-1} + \epsilon_t. \tag{C.10}$$

Fractionally differenced series were generated using the ARFSIM package in RATS, which generates $(0, d, 0)$ series where one can specify $d$ in the range $-1 < d < 1.5$. To ensure that each series was generated properly, we used Robinson's semi-parametric estimator (RGSE package) to estimate each series' order of $d$. If the mean $d$ estimate of 10,000 simulations was within 1% of our target integration level, we used the results of the simulation.[2]

### C.2.1 Related vs. Unrelated Series

Just a brief note on related versus unrelated series. When related series are modeled, the researcher is specifying that the two series share a common stochastic trend such that when the two series are cointegrated, their linear combination is stationary.

For example, in a bivariate system, let $\mathbf{Y}_t = (y_{1t}, y_{2t})' \sim I(1)$ and $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})' \sim I(0)$ in which $\mathbf{Y}_t$ is cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, -\beta)'$. The cointegration relationship can be represented by the following:

$$y_{1t} = \beta\mu_{1t} + \epsilon_{1t}$$
$$y_{2t} = \mu_{2t} + \epsilon_{2t}$$

In this case, if the linear combination of these two series is to be an $I(0)$ process, then it must be the case that the two series share a common stochastic trend: $\mu_{2t} = \beta\mu_{1t}$. Despite the fact that both series are independently unit roots, the cointegrating relationship annihilates the common trend and the equilibrium error term is stationary:

$$\boldsymbol{\beta}'\mathbf{Y}_t = \beta\mu_t + \epsilon_{1t} - \beta(\mu_t + \epsilon_{2t}) = \epsilon_{1t} - \beta\epsilon_{2t} \sim I(0)$$

[2] RATS packages, as well as examples and supporting documentation can be downloaded from the Estima website http://www.estima.com/procs_perl/mainproclistwrapper.shtml. The ARFSIM package has been updated so that its usage is more intuitive, however we found that it was creating a fatal error when combined with the RGSE package. As a result, we used an earlier version of the ARFSIM package, and our code is based on the old version. It is very easy to adapt our code to the new version of ARFSIM, however it may not be possible to simultaneously estimate the $d$ value of the series.

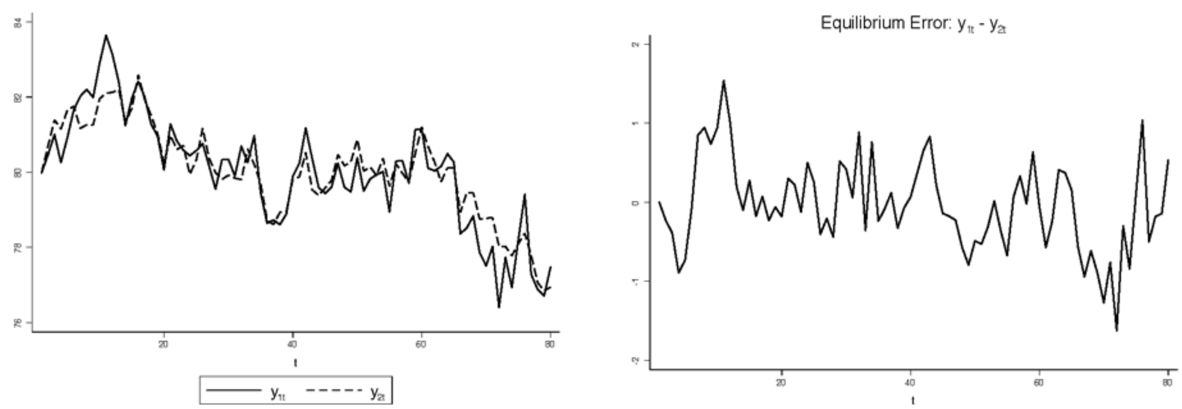
It is fairly simple to simulate cointegrated systems. Using Phillips (1991) triangular representation we can specify the relationship between $y_{1t}$ and $y_{2t}$ as:

$$y_{1t} = \beta y_{2t} + u_t \text{ where } u_t \sim I(0)$$
$$y_{2t} = y_{2t-1} + v_t \text{ where } v_t \sim I(0)$$

In this representation, $y_{1t}$ captures the long-run equilibrium relationship and $y_{2t}$ describes the common stochastic trend.

For example, Figure C.4 presents the plots of the series for $y_{1t}$ and $y_{2t}$ in which $u_t = .50 * u_{t-1} + \epsilon_t$, $\epsilon_t \sim$ iid $N(0,1)$. As expected, $y_{1t}$ and $y_{2t}$ follow each other closely and the plot of the cointegrating residual is clearly stationary.

**Figure C.4: Two Related, Pre-Specified Series**




*Note*: Left plot is of two series in level form. Plot on the right is of the cointegrating residual of the two series. Note that the residual series is stationary.

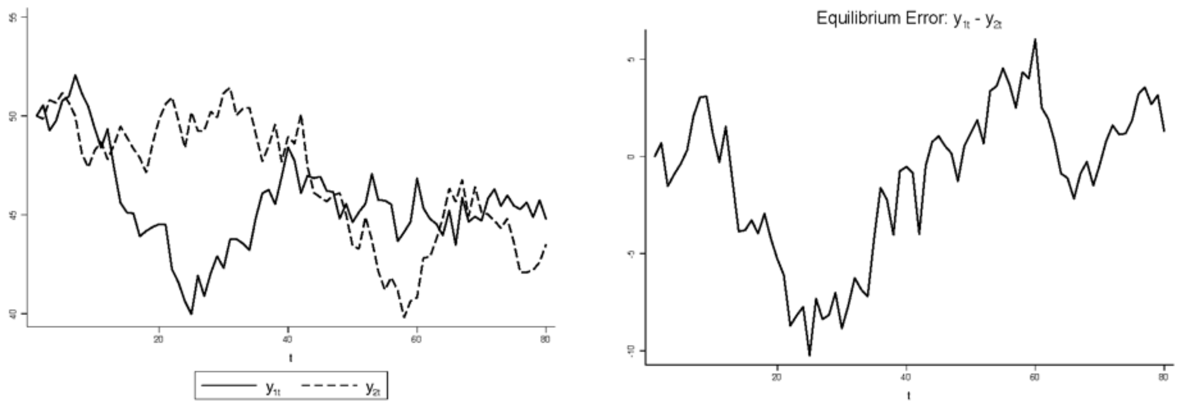
As we note in the main text, previous research has found that the GECM model adequately detects cointegration when the cointegrating vector has been pre-specified. That is, when two series have been created such that they share a common stochastic trend the GECM model has sufficient power.

On the other hand, this paper focuses on the Type I errors that come from *estimating* the cointegrating vector with the GECM model. In order to evaluate the model's performance, we randomly generate our data such that each series' error term is uncorrelated with all other series within the model. For example, if we simulate $y_{1t}$ and $y_{2t}$ as randomly generated unit-root processes and plot them together, we could get two series that look like those presented in Figure C.5.

In this case, $y_{1t}$ and $y_{2t}$ could potentially be related and an estimation of the GECM model between these two series, with $y_{1t}$ as the dependent variable, yields what appears to be a significant ECM parameter ($t = -2.34$). Such a conclusion would be in error however. As can be seen in the right hand plot of Figure C.5, the residuals of the level form regression of $y_{1t}$ on $y_{2t}$ is clearly not stationary. The two series do not share a common trend and there is no cointegration.

For further information on power of the GECM when the cointegrating vector is pre-specified, see Kremers, Ericsson, and Dolado (1992), De Boef and Granato (1999), and Zivot

**Figure C.5: Two Randomly Generated, Unrelated Series**




*Note*: Left plot is of two series in level form. Plot on the right is of the cointegrating residual of the two series. Note that the residual series is not stationary.

(2000). For further information on the power of other cointegration tests when the series are pre-specified as fractionally cointegrated, see Gil-Alana (2003) and Caporale and Gil-Alana (2004).

### C.2.2 Generating Bounded Series

We rely on the work of Nicolau (2002) and the author's discrete-time process to generate the bounded unit roots. The generated series approximate unit-roots when its values fall within a specified interval. However, as the series nears its upper and lower thresholds it tends towards mean reversion. The process is generated by:

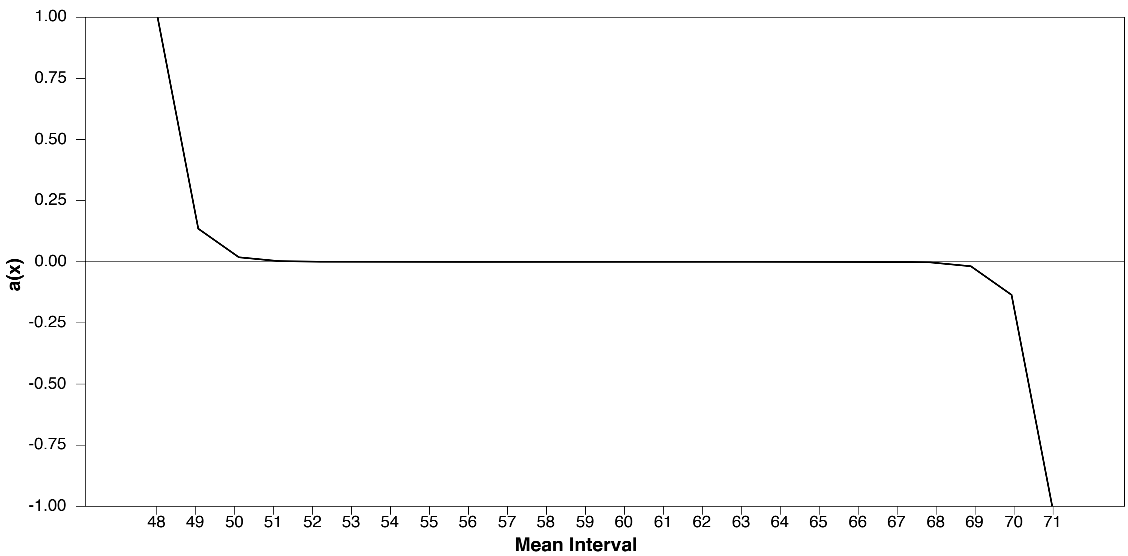
$$X_{\mathrm{t}} = X_{\mathrm{t-1}} + e^{k}(e^{-\alpha_1(X_{\mathrm{t-1}}-\tau)} - e^{\alpha_2(X_{\mathrm{t-1}}-\tau)}) + \epsilon_{\mathrm{t}} \tag{C.11}$$

where $\alpha_1 \geq 0$; $\alpha_2 \geq 0$; $k < 0$; and $\epsilon_{\mathrm{t}}$ is assumed to be *i.i.d.* with a mean of 0 and variance of 1.

Equation (C.11) adds an additional function to the basic unit-root. The function $a(x) = e^{k}(e^{-\alpha_1(X_{\mathrm{t-1}}-\tau)} - e^{\alpha_2(X_{\mathrm{t-1}}-\tau)})$ captures the mean reversion of the process. The range of the process under which $a(x)$ behaves as a unit-root is controlled by the $k$ parameter; as $|k|$ increases, so too does the interval over which the series will behave as an $I(1)$ process. The $\tau$ parameter is the central tendency of the process. The final parameters, $\alpha_1$ and $\alpha_2$, measure the reversion effect of the process as it approaches or exceeds the specified interval. The larger the value of $\alpha$ the stronger the reversion towards the mean. When $\alpha_1 = \alpha_2$, the parameter $\tau$ is the mean of the series and the interval is symmetric.

When $X_{\mathrm{t-1}} = \tau$, then $a(x) = 0$ and the generated process will behave as a unit-root. However we can widen the range over which the process varies by selecting the parameters $\alpha_1$, $\alpha_2$, and $k$, such that $a(x) \approx 0$ over a specific range of values surrounding $\tau$. In our case, we chose to create a series that approximates the range of Stimson's (1991) "Mood," a commonly used variable within political science. *Mood* is not strictly bounded, however as with many variables in the field, it has a rather limited range in practice. From 1952 to 2006

**Figure C.6: Value of *a(x)* Over Interval**




the low of *Mood* was 49 and its high was 69. Choosing parameter values of $\alpha_1 = \alpha_2 = 1.5$; $|k| = 16.5$; and $\tau = 59$, we created a variable of similar range. Figure C.6 plots the values of the function $a(x)$. When the series falls within the interval (49,70), the function $a(x)$ is near 0, but as the series shifts either above or below the interval, the strength of the mean reverting process increases.

An example of this process is shown in Figure C.7. We created two identical unit-root series, however the bounded series has been constrained by our chosen interval. As the two series reach the interval, either above or below the mean, the value of $a(x)$ increases and the bounded series is pushed back towards its mean. The two series are equal for the first 50 or so time points, but as they approach the upper threshold, the bounded unit-root is pushed back while the original unit-root continues to meander without bounds. The two series have diverged, however because their disturbance terms are equal, they still track each other over the range of observations.

Table 3 in the main text provides results of Monte Carlo simulations with a bounded dependent variable with various variances ($\sigma$ =1, 2, 3). To visualize what such series may look like, Figure B.3 shows plots of individual series with a $T$ of 60 and a specific value of $\sigma$.

**Figure C.7: Bounded Unit-Root Compared to a Unit-Root**

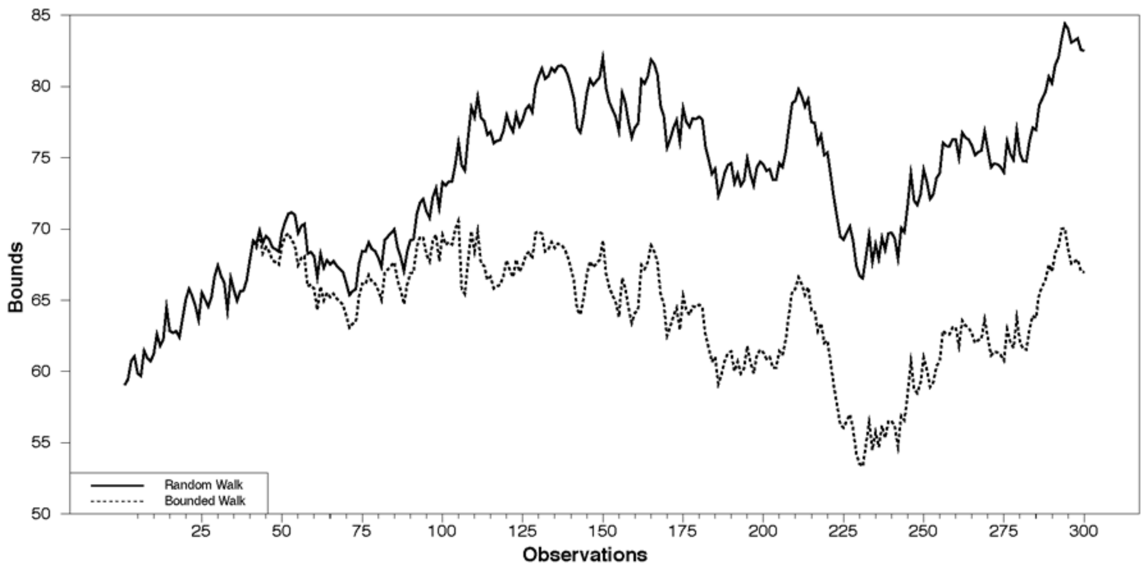
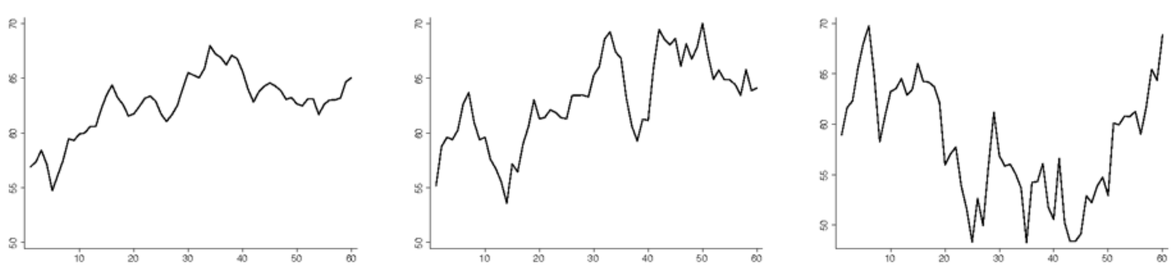



**Figure C.8: Series Plots as Variance Changes**



(a) $\sigma_t \sim N(0,1)$ (b) $\sigma_t \sim N(0,2)$ (c) $\sigma_t \sim N(0,3)$

## C.3 Additional Materials: Replication of Casillas, Enns, and Wohlfarth (2011)

We replicate Casillas, Enns, and Wohlfarth's (2011) Tables 1 and 2 in Table C.1. All results are comparable, with the exception of the standard error of the LRM for *Social Forces* in the *Salient Review* model.[3] We find the standard error is somewhat lower, but this has no effect on the substantive findings of the model.

### C.3.1 Full Paper Replication

**Table C.1: Replication of Casillas et al. (2011)**

| | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| **Long Run Multiplier** | | | |
| Public Mood | 1.05* | 1.15* | |
| | (0.51) | (0.56) | |
| Court Ideology | 11.67* | 10.90* | 13.31* |
| | (2.82) | (3.11) | (4.22) |
| Social Forces | | | 7.28* |
| | | | (3.83) |
| **Long Run Effects** | | | |
| Public Mood$_{t-1}$ | 0.87* | 0.88* | 0.71 |
| | (0.42) | (2.15) | (0.96) |
| Court Ideology$_{t-1}$ | 9.68* | 8.37* | 16.90* |
| | (3.10) | (3.09) | (5.73) |
| Social Forces (IV)$_{t-1}$ | 0.98 | -0.01 | 9.20* |
| | (2.14) | (2.15) | (5.09) |
| **Short Run Effects** | | | |
| $\Delta$Public Mood | 1.59* | 1.68* | 1.24 |
| | (0.78) | (0.79) | (1.81) |
| $\Delta$Court Ideology | 12.48* | 11.37* | 10.47 |
| | (4.29) | (4.41) | (10.03) |
| $\Delta$Social Forces (IV) | 2.78 | 2.56 | 7.49 |
| | (3.00) | (2.98) | (8.03) |
| **Error Correction and Constant** | | | |
| Percent Liberal$_{t-1}$ | -0.83* | -0.77* | -1.27* |
| | (0.15) | (0.15) | (0.15) |
| Constant | -6.03 | -11.05 | 31.02 |
| | (25.14) | (25.42) | (58.32) |
| **Fit and Diagnostics** | | | |
| Centered $R^2$ | 0.53 | 0.49 | 0.64 |
| Sargan ($\chi^2$) | 0.38 | 0.67 | 0.21 |
| N | 45 | 45 | 45 |

*Note*: Entries are two-stage least squares coefficients (standard errors in parentheses). ECM significance one-tail $t$-test. Coefficient significance (*p$\leq$ 0.05, one-tail test).

[3]The long-run multiplier (LRM) is calculated as the ratio of coefficients: $\beta_1/\alpha_1$. The formula for the standard errors of the LRM is: $((1/b^2)\text{Var}(a) + (a^2/b^4)\text{Var}(b) - 2(a/b^3)\text{Cov}(a,b))^{1/2}$ See De Boef and Keele (2008, pp.191-192).

### C.3.2 Replication of CEW - First Stage Results

Table C.2 presents the results of the first stage regressions in which Martin-Quinn scores were regressed on the "social forces" that predict the public mood. CEW note that these social forces are substantial predictors of the public mood, therefore when used as instruments for Martin-Quinn scores, they will capture that portion of justice ideology influenced by the same social forces that influence public mood.

Because the table is so large, the social forces which significantly predict our variables to be instrumented are bolded. Fit diagnostics are provided. The null hypothesis of each diagnostic is that the instrument is poor – thus we want to reject the null to show our instruments are valid and that the second stage estimates are reliable. With weak instruments, results of 2SLS are biased towards the results of OLS. This bias only gets worse as more weak instruments are added. With instruments as weak as CEW's, any inferences drawn from the model should be taken with extreme caution.

Note: The weak identification test does not provide a p-value, but instead the Stock-Yogo critical value to determine bias from weak identification of the instruments, the 5% critical value is 19.40. With values of 1.52, 1.59, and 0.97, respectively, the three models fail to reject the null of the Stock-Yogo. The instruments are not useful and the second stage is nearly identical to the OLS estimates. If endogeneity is indeed a problem with this data-set, it doesn't appear that the specified model provides an adequate solution.

## Table C.2: First Stage Results from 2SLS

| | All Reviews | | Non-Salient | | Salient | |
|---|---|---|---|---|---|---|
| **1st Stage Variables** | $\Delta$MQ | MQ$_{t-1}$ | $\Delta$MQ | MQ$_{t-1}$ | $\Delta$MQ | MQ$_{t-1}$ |
| Reviews$_{t-1}$ | **0.04**$^{\star}$ | -0.00 | **0.04**$^{\star}$ | -0.01 | 0.01 | 0.00 |
| | (0.01) | (0.04) | (0.01) | (0.02) | (0.01) | (0.01) |
| $\Delta$Mood$_t$ | **0.25**$^{\star}$ | **-0.24**$^{\star}$ | **0.25**$^{\star}$ | **-0.25**$^{\star}$ | **0.20**$^{\star}$ | **-0.24**$^{\star}$ |
| | (0.07) | (0.10) | (0.07) | (0.09) | (0.08) | (0.09) |
| Mood$_{t-1}$ | 0.13 | -0.13 | 0.13 | -0.13 | 0.15 | -0.14 |
| | (0.07) | (0.09) | (0.07) | (0.09) | 0.08 | (0.09) |
| $\Delta$Segal-Cover$_t$ | 0.65 | 0.95 | 0.65 | 0.94 | 0.66 | 0.96 |
| | (0.48) | (0.65) | (0.47) | (0.65) | (0.54) | (0.65) |
| Segal-Cover$_{t-1}$ | -0.52 | 1.48 | -0.58 | 1.51 | -0.22 | 1.47 |
| | (0.57) | (0.77) | (0.56) | (0.77) | 0.63 | (0.75) |
| **"Social Forces" Excluded in 2nd Stage** | | | | | | |
| $\Delta$% Change Inflation$_t$ | 6.41 | 0.60 | 4.49 | 1.15 | 10.95 | 0.55 |
| | (8.46) | (11.44) | (8.41) | (11.56) | (9.46) | (11.36) |
| % Change Inflation$_{t-1}$ | 13.67 | 12.16 | 13.47 | 12.28 | 14.73 | 12.10 |
| | (11.77) | (15.91) | (11.56) | (15.89) | (13.24) | (15.90) |
| $\Delta$Unemployment$_t$ | -0.18 | 0.07 | -0.16 | 0.09 | -0.03 | 0.06 |
| | (0.18) | (0.25) | (0.18) | (0.24) | (0.20) | (0.24) |
| Unemployment$_{t-1}$ | 0.21 | 0.19 | 0.20 | 0.19 | 0.23 | 0.19 |
| | (0.17) | (0.23) | (0.16) | (0.23) | (0.19) | (0.23) |
| $\Delta$Defense Budget$_t$ | **-0.15**$^{\star}$ | 0.09 | **-0.16**$^{\star}$ | 0.10 | -0.08 | 0.08 |
| | (0.06) | (0.08) | (0.06) | (0.08) | (0.06) | (0.07) |
| Defense Budget$_{t-1}$ | -0.08 | -0.04 | -0.08 | -0.04 | -0.05 | -0.04 |
| | (0.05) | (0.07) | (0.05) | (0.07) | (0.06) | (0.07) |
| $\Delta$Policy Liberalism$_t$ | -0.07 | **0.14**$^{\star}$ | -0.06 | **0.14**$^{\star}$ | -0.07 | **0.13**$^{\star}$ |
| | (0.04) | (0.05) | (0.04) | (0.37) | (0.05) | (0.06) |
| Policy Liberalism$_{t-1}$ | 0.02 | -0.03 | 0.02 | -0.03 | 0.03 | -0.03 |
| | (0.04) | (0.05) | (0.04) | (0.05) | (0.04) | (0.05) |
| $\Delta$Homicide Rate$_t$ | 0.18 | 0.14 | 0.18 | 0.12 | 0.04 | 0.16 |
| | (0.30) | (0.41) | (0.30) | (0.41) | (0.34) | (0.41) |
| Homicide Rate$_{t-1}$ | -0.31 | -0.02 | -0.34 | -0.03 | -0.39 | -0.01 |
| | (0.24) | (0.33) | (0.24) | (0.33) | (0.28) | (0.33) |
| $\Delta$Gini$_t$ | -27.94 | -3.91 | -34.11 | -2.28 | -15.22 | -4.01 |
| | (17.73) | (23.97) | (17.85) | (24.53) | (19.74) | (23.71) |
| Gini$_{t-1}$ | -8.21 | 3.39 | -10.83 | 4.49 | 1.20 | 2.91 |
| | (16.76) | (22.65) | (16.60) | (22.81) | (18.51) | (22.23) |
| Constant | -4.08 | 6.59 | -2.65 | 6.13 | 6.68 | 31.02 |
| | (9.72) | (13.13) | (9.60) | (13.19) | (13.05) | (58.32) |
| **Fit and Diagnostics** | | | | | | |
| Underidentification Test$^{\ddagger}$ | 0.08 | | 0.07 | | 0.26 | |
| Weak Identification Test$^{\dagger}$ | 1.52 | | 1.59 | | 0.97 | |
| Anderson-Rubin Weak Instrument$^{\ddagger}$ | 0.71 | | 0.57 | | 0.38 | |
| Stock-Wright Weak Instrument$^{\ddagger}$ | 0.53 | | 0.88 | | 0.24 | |

*Note*: First stage regressions results of Casillas, Enns, and Wohlfarth (2011). $^{\dagger}$5% Maximal IV relative bias = 19.40; 30% Maximal IV relative bias = 4.59. $^{\ddagger}$p-value ($^{\star}$p$\leq$ 0.05, two-tail test)

### C.3.3 Replications with OLS

In replicating with our series, we chose to adhere to Casillas, Enns, and Wohlfarth's (2011) estimation methods of 2SLS. Their use of 2SLS was intended to control for the effects of social forces on contemporaneous Justice ideology. As noted, we did not have concerns of the endogeneity of beef consumption with the ideological direction of the Court's decision making, however we wanted our estimations to be comparable. Despite CEW's concerns of potential endogeneity, their first stage instruments are so weak that their results are very strongly biased towards the results of OLS.

We also replicate their models, as well as ours, using OLS. With some very minor differences - for instance, the long term effect of Martin-Quinn justice ideology is significant in the *All Review* model, and *Mood* has a substantively larger effect - the results of the 2SLS and OLS replications of Casillas et al. are equivalent. As expected, so too are the results from our nonsense regression series.

**Table C.3: Replication of Casillas et al. (2011) with OLS**

| | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| **Long Run Effects** | | | |
| Public Mood$_{t-1}$ | 1.08* | 1.13* | 0.55 |
| | (0.41) | (0.42) | (0.95) |
| Court Ideology$_{t-1}$ | 8.07* | 6.00* | 18.16* |
| | (2.97) | (2.84) | (5.43) |
| Martin-Quinn$_{t-1}$ | 3.01* | 2.27 | 7.24* |
| | (1.57) | (1.56) | (3.73) |
| **Short Run Effects** | | | |
| $\Delta$Public Mood | 2.04* | 2.25* | 0.89 |
| | (0.74) | (0.74) | (1.71) |
| $\Delta$Court Ideology | 10.91* | 9.53* | 12.33 |
| | (4.39) | (4.44) | (10.06) |
| $\Delta$Martin-Quinn | 2.59 | 1.67 | 6.88 |
| | (3.00) | (2.16) | (4.95) |
| **Error Correction and Constant** | | | |
| Percent Liberal$_{t-1}$ | -0.84* | -0.75* | -1.25* |
| | (0.15) | (0.15) | (0.16) |
| Constant | -17.78 | -26.69 | 39.23 |
| | (24.65) | (24.63) | (57.84) |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.47 | 0.44 | 0.57 |
| Breusch-Pagan Test ($\chi^2$) | 2.22 | 3.13 | 7.02* |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.16 | 1.50 | 0.04 |
| N | 45 | 45 | 45 |

*Note*: Entries are OLS regression coefficients (standard errors in parentheses).
ECM significance one-tail *t*-test. Coefficient significance (*$p \leq 0.05$, one-tail test).

As for the ECMs, the estimates of Casillas, Enns, and Wohlfarth (2011) are nearly identical with OLS and their *Salient Review* model still has an estimated ECM of -1.25. The results of our nonsense data series are also similar, however our *Salient Review* ECM has shrunk to -0.96.

### C.3.4 Estimating Fractional Order of Integration

Table 11 of the main text provides the results of each fractionally differenced regression. Table C.5 presents the initial step in the process: investigating the order of integration of

**Table C.4: What Else Affects the Court's Liberal Reversal Rate (Re-estimation of Casillas et al. (2011)) Using OLS**

| | All Reviews | Non-Salient Reviews | Salient Reviews |
|---|---|---|---|
| **Long Run Effects** | | | |
| Shark Attacks$_{t-1}$ | 0.30* | 0.30* | 0.48* |
| | (0.12) | (0.12) | (0.21) |
| Tornado Fatalities$_{t-1}$ | 0.02 | 0.01 | 0.06 |
| | (0.03) | (0.03) | (0.06) |
| Beef Consumption (IV)$_{t-1}$ | -0.88* | -0.91* | -1.81* |
| | (0.29) | (0.29) | (0.44) |
| **Short Run Effects** | | | |
| ΔShark Attacks | 0.20* | 0.18* | 0.37 |
| | (0.11) | (0.10) | (0.22) |
| ΔTornado Fatalities | 0.04* | 0.03 | 0.09* |
| | (0.02) | (0.02) | (0.04) |
| ΔBeef Consumption (IV) | -0.80 | -1.81 | 4.49 |
| | (1.49) | (1.49) | (3.21) |
| **Error Correction and Constant** | | | |
| Percent Liberal$_{t-1}$ | -0.52* | -0.56* | -0.96* |
| | (0.14) | (0.15) | (0.16) |
| Constant | 56.62* | 60.37* | 110.90* |
| | (27.56) | (17.75) | 23.95 |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.31 | 0.27 | 0.53 |
| Breusch-Pagan Test ($\chi^2$) | 0.01 | 0.01 | 3.89* |
| Breusch-Godfrey LM Test ($\chi^2$) | 2.98 | 7.13* | 1.46 |
| N | 45 | 45 | 45 |

*Note*: Entries are OLS regression coefficients (standard errors in parentheses). ECM significance one-tail *t*-test. Coefficient significance (*p≤ 0.05, one-tail test).

each constituent series.[4] Recall that in order to estimate a three-step FECM the order of integration between a DV and IV must be equivalent. Absent this equivalence, the two (or more) series cannot share a common stochastic trend, and therefore cannot be fractionally cointegrated.

Our estimates of $d$ are consistent and reaffirm our decision to not fit an FECM with the data. The three DVs are fractionally integrated, which means that despite the fact that Dickey-Fuller tests cannot reject the null of a unit root for either *All Reviews* or *Non-Salient Reviews*, the two series are not integrated $I(1)$. Shocks are not permanent, but decrease at a hyperbolic rate. It is tempting to conclude that the DV *Salient Reviews* is long-memoried, but stationary ($d < 0.5$), however without more observations this can't be established with 95% confidence. While our results indicate the DVs are not pure unit roots, the confidence intervals on the estimates of the three IVs all overlap 1.0, meaning we cannot reject the null that the three IVs are all integrated.

[4] All estimates of $d$ were estimated using Stata's exact ML estimator. The exact ML estimator is consistent and asymptotically normal for series $-0.5 < d < 0.5$, which generally requires first differencing a series and then estimating $d$. The RATS Local Whittle estimator of Robinson (1995) is in general agreement with the estimates of the Stata program: *All Reviews* ($d = 0.62$); *Non-Salient Reviews* ($d = 0.63$); *Salient Reviews* ($d = 0.36$); *Public Mood* ($d = 1.14$); *Segal-Cover Scores* ($d = 1.15$); *Martin-Quinn Score* ($d = 1.09$). The asymptotic standard error of the Local Whittle estimator is based on $m$, the bandwidth, or number of Fourier frequencies used by the estimator. For all Local Whittle estimates, $m$ is calculated as $T^{4/5}$ and the asymptotic standard errors are calculated as $(1/\sqrt{4 * m})$: (s.e.=0.11).

### Table C.5: Fractional Order of Integration, Casillas et al (2011)

| | Individual Series $d$ | 95% Confidence Interval $d$ |
|---|---|---|
| All Reviews | 0.62 (0.11) | [0.404, 0.836] |
| Non-Salient Reviews | 0.62 (0.12) | [0.385, 0.858] |
| Salient Reviews | 0.36 (0.08) | [0.200, 0.529] |
| Public Mood | 1.09 (0.13) | [0.835, 1.345] |
| Segal-Cover Score | 1.07 (0.14) | [0.796, 1.344] |
| Martin-Quinn Score | 1.00 (0.15) | [0.706, 1.294] |

† *Note*: Entries are estimates of the order of integration of all series used by Casillas, Enns, and Wohlfarth (2011). Standard errors in parentheses. Estimates conducted in Stata using the exact ML estimator.

We also provide the estimates of fractional integration of our nonsense data: *Beef Consumption*: $d = 1.14$ (s.e.$=0.15$); *Shark Attacks*: $d = 0.56$ (s.e.$=0.13$); and *Tornado Fatalities*: $d = 0.05$ (s.e.$=0.12$). In the interests of completeness, we then estimate this data in both fractionally differenced regression models as well as (potentially misspecified) FECMs. As Tables C.6 C.7 demonstrate, when using fractional methods on our Beef-Shark-Tornado data we find only null results – a good sign for the method.

### Table C.6: FI Model of Nonsense Data and Supreme Court Reversals

| Review Type | All | Non-Salient | Salient |
|---|---|---|---|
| **Short Run Effects** | | | |
| $\Delta^d$ Shark Attacks | 0.18 | 0.16 | 0.28 |
| | (0.11) | (0.11) | (0.24) |
| $\Delta^d$ Tornado Fatalities | 0.03 | 0.02 | 0.07 |
| | (0.02) | (0.02) | (0.04) |
| $\Delta^d$ Beef Consumption | -1.01 | -1.96 | 4.37 |
| | (1.49) | (1.46) | (3.19) |
| Constant | -2.10 | -2.04 | -12.42* |
| | (1.69) | (1.66) | (3.62) |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.04 | 0.03 | 0.09 |
| Durbin-Watson | 2.13 | 2.17 | 2.35 |
| Breusch-Pagan Test ($\chi^2$) | 3.69 | 4.63 | 0.38 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.32 | 0.62 | 1.65 |

*Note*: Entries are OLS coefficients (standard errors in parentheses). All variables have been fractionally differenced by their estimate of $d$. Coefficient significance (*p$\leq$ 0.05, two-tail test).

## Table C.7: FECM Model of Nonsense Data and Supreme Court Reversals

| **Review Type** | **All** | **All** | **All** | **Non-Salient** | **Non-Salient** | **Non-Salient** | **Salient** |
|---|---|---|---|---|---|---|---|
| **Short Run Effects** | | | | | | | |
| $\Delta^d$ Shark Attacks | 0.16 | 0.17 | 0.16 | 0.15 | 0.15 | 0.14 | 0.28 |
| | (0.12) | (0.12) | (0.12) | (0.11) | (0.11) | (0.11) | (0.24) |
| $\Delta^d$ Tornado Fatalities | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.07 |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.04) |
| $\Delta^d$ Beef Consumption | -0.66 | -0.72 | -0.65 | -1.58 | -1.62 | -1.56 | 4.37 |
| | (1.55) | (1.55) | (1.55) | (1.53) | (1.53) | (1.52) | (3.19) |
| **Error Correction and Constant** | | | | | | | |
| $\text{FECM}_{\text{Sharks}}$ | -0.10 | | | -0.10 | | | |
| | (0.15) | | | (0.15) | | | |
| $\text{FECM}_{\text{Tornados}}$ | | -0.07 | | | -0.09 | | |
| | | (0.16) | | | (0.16) | | |
| $\text{FECM}_{\text{Beef}}$ | | | -0.14 | | | -0.14 | |
| | | | (0.16) | | | (0.16) | |
| Constant | -2.55 | -2.34 | -2.34 | -2.45 | -2.35 | -2.28 | -12.42* |
| | (1.89) | (1.87) | (1.75) | (1.86) | (1.83) | (1.73) | (3.62) |
| **Fit and Diagnostics** | | | | | | | |
| Adjusted $R^2$ | 0.10 | 0.09 | 0.11 | 0.09 | 0.08 | 0.09 | 0.09 |
| Durbin-Watson | 1.97 | 2.02 | 1.93 | 2.03 | 2.04 | 1.98 | 2.35 |
| Breusch-Pagan Test ($\chi^2$) | 6.69 | 5.84 | 7.15 | 7.86 | 9.27 | 8.57 | 0.38 |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.26 | 0.01 | 0.99 | 0.07 | 0.09 | 0.53 | 1.65 |

*Note*: Entries are OLS coefficients (s.e. in parentheses). All variables have been fractionally differenced by their estimate of $d$. FECM significance (*$p \leq 0.05$, one-tail test). Coefficient significance (*$p \leq 0.05$, two-tail test).

## C.4 Additional Materials: Replication of Ura and Ellis (2012)

### C.4.1 Full Paper Replication

Below is the complete replication of Table 2 from Ura and Ellis (2012). Because the authors used a Seemingly Unrelated Regression, all replications and Monte Carlo simulations were conducted in Stata so that we could exactly replicate their findings.

**Table C.8: Replication of Ura and Ellis - Table 2†**

| | Republican | Democrat | \|Difference\| |
|---|---|---|---|
| **Long Run Multipliers** | | | |
| Domestic Spending $10B | -0.28⋆ (0.08) | -0.11⋆ (0.05) | 0.16⋆ |
| Defense Spending $10B | 0.52⋆ (0.15) | 0.29⋆ (0.05) | 0.23 |
| Inflation | -0.57⋆ (0.28) | -0.47⋆ (0.19) | 0.10 |
| Unemployment | -0.17 (0.83) | 0.49 (0.37) | 0.66 |
| Top 1% Income Share | 2.45⋆ (0.94) | 1.54⋆ (0.27) | 0.91 |
| **Long Run Effects** | | | |
| Domestic Spending $10B$_{t-1}$ | -0.11⋆ (0.02) | -0.07⋆ (0.02) | 0.04 |
| Defense Spending $10B$_{t-1}$ | 0.21⋆ (0.05) | 0.20⋆ (0.06) | 0.01 |
| Inflation$_{t-1}$ | -0.23⋆ (0.12) | -0.33⋆ (0.16) | 0.10 |
| Unemployment$_{t-1}$ | -0.07 (0.33) | 0.34 (0.24) | 0.41 |
| Top 1% Income Share$_{t-1}$ | 0.99⋆ (0.35) | 1.07⋆ (0.29) | 0.08 |
| **Short Run Effects** | | | |
| ΔDomestic Spending $10B$_t$ | -0.04 (0.03) | -0.07⋆ (0.04) | 0.03 |
| ΔDefense Spending $10B$_t$ | 0.43⋆ (0.09) | 0.14⋆ (0.07) | 0.29⋆ |
| ΔInflation$_t$ | -0.12 (0.20) | -0.03 (0.14) | 0.09 |
| ΔUnemployment$_t$ | 0.12 (0.29) | 0.68⋆ (0.41) | 0.56 |
| ΔTop 1% Income Share$_t$ | 1.03⋆ (0.23) | 0.76⋆ (0.23) | 0.27 |
| **Error Correction and Constant** | | | |
| Partisan Mood$_{t-1}$ | -0.40⋆ (0.09) | -0.69⋆ (0.17) | 0.29 |
| Constant | 16.80⋆ (6.65) | 33.93⋆ (9.50) | 17.13 |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.39 | 0.39 | |
| Breusch-Pagan Test ($\chi^2$) | 0.01 | 0.40 | |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.42 | 0.70 | |

† *Note*: Entries are seemingly unrelated regression coefficients (standard errors in parentheses). The Dickey-Fuller Test statistic was not included as we couldn't replicate their estimates. N=35

### C.4.2 Full Monte Carlo Results

The result of 10,000 Monte Carlo simulations is presented in Table C.9. We have simulated $I(1)$ dependent variables and regressed them on the IVs used by Ura and Ellis (2012). In keeping with the methodological choice of Ura and Ellis, each iteration was estimated as a SUR, however because the set of independent variables is the same in the two equations, the results of the simulations are essentially equal for both models. This does not affect the conclusions that we draw from the results - the independent variables, two of which have estimated orders of fractional integration significantly greater than $I(1)$, are biasing the ECM $t-$distribution well away from that used to derive the MacKinnon critical values. The result is a finding that approximately 20% of all error correction parameters as significant. The rate of cointegration is overstated.

### C.4.3 Estimating FECM of Ura and Ellis (2012)

Because Ura and Ellis (2012) do not specify which of their variables they believe to be in an equilibrium relationship with the dependent variable we estimate a three-step fractional

**Table C.9: Summary Statistics Monte Carlo Simulation of Ura and Ellis (2012)**

| | Random DV 1 | Random DV 2 |
|---|---|---|
| % ECM Significant – one tail $t$-distribution | 86.4 | 86.5 |
| % ECM Significant – MacKinnon Values | 22.4 | 22.5 |
| Mean of $\alpha_1$ | -0.40 | -0.40 |
| % $\Delta X_t$ Significant | 57.2 | 57.5 |
| % $\geq 1 X_{t-1}$ Significant | 69.4 | 68.4 |
| % ECM and $\geq 1 \Delta X_t$ Significant | 50.0 | 50.1 |
| % ECM and $\geq 1 \Delta X_t$ Significant* | 15.4 | 14.9 |
| % ECM and $\geq 1 X_{t-1}$ Significant | 62.3 | 61.9 |
| % ECM and $\geq 1 X_{t-1}$ Significant* | 19.9 | 19.8 |

*Note*: Entries provide are the summary results of 10,000 simulations of a seemingly unrelated regression in which randomly generated $I(1)$ series are regressed upon the independent variables from Ura and Ellis (2012).
$\Delta X_t$ and $X_{t-1}$ significance (*p$\leq$ 0.05, two-tail test).
MacKinnon Value: $-4.268$

ECM for each independent variable in each model. After each bivariate regression we then estimated the order of integration of the residuals for comparison with the component series. Even if the parent series are $I(1)$, fractional cointegration may still exist if the fractional integration of the regression residuals are less than $I(1)$, but greater than $I(0)$. The results of these estimates, as well as estimates of the component series, can be found in Table C.10. Only the residuals from the regression of *Partisan Mood* on *Inflation* indicate any decrease in order of integration, however the reduction is marginal.

**Table C.10: FECM Order of Integration**

| | **Individual Series** $d$ | **Residuals Rep Mood & IV** $d$ | **Residuals Dem Mood & IV** $d$ |
|---|---|---|---|
| Republican Mood | 1.05 (0.16) | | |
| Democrat Mood | 1.15 (0.16) | | |
| Domestic Spending $10B | 1.44 (0.08) | 1.04 (0.16) | 1.14 (0.16) |
| Defense Spending $10B | 1.32 (0.11) | 1.03 (0.16) | 1.09 (0.16) |
| Inflation | 1.02 (0.24) | 0.91 (0.18) | 0.76 (0.21) |
| Unemployment | 0.94 (0.21) | 1.08 (0.17) | 0.99 (0.18) |
| Top 1% Income Share | 0.89 (0.19) | 1.06 (0.16) | 1.01 (0.16) |

† *Note*: Entries are estimates of the order of integration of each individual series as well as estimates of the residuals of bivariate regressions. Estimates based on results of the Stata exact ML estimator.

The estimates in Table C.10 argue strongly against any finding of fractional cointegration. To be sure, we estimate each individual FECM and present the results in Tables C.11 and C.12 for *Republican Mood* and *Democratic Mood*, respectively.

Comparing the results from the FECM models we see that *Republican Mood* is affected by both increases in *Defense Spending* and in the *Top 1% Income Share*. As each increases, *Republican Mood* shifts in the liberal direction. These results hold across all models. Despite the IVs' significance, we don't find any evidence of an equilibrium relationship between the DV and any of the covariates. For the *Democrat Mood* model, no variables achieve significance. These results are drastically different from those reported by Ura and Ellis with the single equation GECM.

## Table C.11: Three-Step FECM Results - Republican Mood

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^{d}\text{Domestic Spending \$10B}_{t}$ | 0.02 | 0.03 | 0.03 | 0.02 |
| | (0.04) | (0.04) | (0.04) | (0.04) |
| $\Delta^{d}\text{Defense Spending \$10B}_{t}$ | 0.31* | 0.31* | 0.30* | 0.30* |
| | (0.13) | (0.13) | (0.13) | (0.13) |
| $\Delta^{d}\text{Inflation}_{t}$ | -0.31 | -0.31 | -0.33 | -0.29 |
| | (0.20) | (0.21) | (0.20) | (0.21) |
| $\Delta^{d}\text{Unemployment}_{t}$ | -0.47 | -0.47 | -0.51 | -0.47 |
| | (0.37) | (0.37) | (0.38) | (0.37) |
| $\Delta^{d}\text{Top 1\% Income Share}_{t}$ | 1.21* | 1.21* | 1.19* | 1.22* |
| | (0.42) | (0.42) | (0.42) | (0.42) |
| **Error Correction and Constant** | | | | |
| $\text{Partisan Mood}_{t\text{-}1}{}^{\dagger}$ | 0.11 | 0.09 | 0.06 | 0.13 |
| | (0.16) | (0.16) | (0.15) | (0.17) |
| Constant | -1.30 | -1.31 | -1.34 | -1.30 |
| | (0.68) | (0.68) | (0.68) | (0.68) |
| **Fit and Diagnostics** | | | | |
| $R^2$ | 0.38 | 0.38 | 0.37 | 0.38 |
| Durbin-Watson | 2.22 | 2.18 | 2.09 | 2.23 |
| Breusch-Pagan Test ($\chi^2$) | 1.57 | 1.46 | 1.17 | 1.60 |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.27 | 0.93 | 0.20 | 1.38 |

*Note*: Entries are OLS coefficients. FECM significance (*p≤ 0.05, one-tail test). Coefficient significance (*p≤ 0.05, two-tail test).
† Model (1) - FECM of Mood & Domestic Spending; Model (2) - FECM of Mood & Defense Spending; Model (3) - FECM of Mood & Inflation; Model (4) - FECM of Mood & Unemployment

## Table C.12: Three-Step FECM Results - Democrat Mood

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^{d}\text{Domestic Spending \$10B}_{t}$ | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.04) | (0.04) | (0.04) | (0.04) |
| $\Delta^{d}\text{Defense Spending \$10B}_{t}$ | 0.09 | 0.10 | 0.10 | 0.10 |
| | (0.12) | (0.11) | (0.11) | (0.12) |
| $\Delta^{d}\text{Inflation}_{t}$ | -0.04 | -0.05 | -0.01 | -0.07 |
| | (0.18) | (0.18) | (0.17) | (0.21) |
| $\Delta^{d}\text{Unemployment}_{t}$ | -0.05 | -0.07 | 0.13 | -0.05 |
| | (0.33) | (0.33) | (0.32) | (0.32) |
| $\Delta^{d}\text{Top 1\% Income Share}_{t}$ | 0.51 | 0.55 | 0.62 | 0.51 |
| | (0.37) | (0.38) | (0.36) | (0.37) |
| **Error Correction and Constant** | | | | |
| $\text{Partisan Mood}_{t\text{-}1}{}^{\dagger}$ | -0.06 | -0.13 | -0.19 | -0.08 |
| | (0.19) | (0.20) | (0.13) | (0.17) |
| Constant | -0.13 | -0.16 | -0.19 | -0.12 |
| | (0.61) | (0.60) | (0.57) | (0.59) |
| **Fit and Diagnostics** | | | | |
| $R^2$ | 0.08 | 0.10 | 0.15 | 0.09 |
| Durbin-Watson | 2.15 | 2.08 | 1.99 | 2.14 |
| Breusch-Pagan Test ($\chi^2$) | 0.05 | 0.04 | 0.88 | 0.18 |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.83 | 0.52 | 0.00 | 0.44 |

*Note*: Entries are OLS coefficients. FECM significance (*p≤ 0.05, one-tail test). Coefficient significance (*p≤ 0.05, two-tail test).
† Model (1) - FECM of Mood & Domestic Spending; Model (2) - FECM of Mood & Defense Spending; Model (3) - FECM of Mood & Inflation; Model (4) - FECM of Mood & Unemployment

We also estimate three-step fractional ECMs for each independent variable in our nonsense model, beginning with onion acreage. The ECM results, presented in Tables C.13 - C.15, are null across all model specifications.

**Table C.13: Three-Step FECM Results - Partisan Mood & Onion Acreage**

| | Republican Mood | | Democrat Mood | |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$ Onion Acreage | 0.02 | (0.05) | 0.02 | (0.04) |
| $\Delta^d$ Coal Emissions | 0.02 | (0.03) | -0.00 | (0.02) |
| $\Delta^d$ Beef Consumption | 0.58 | (0.40) | -0.23 | (0.31) |
| $\Delta^d$ Shark Attacks | 0.02 | (0.03) | 0.00 | (0.02) |
| $\Delta^d$ Tornado Fatalities | 0.02 | (0.01) | 0.00 | (0.00) |
| **Error Correction and Constant** | | | | |
| FECM$^\dagger$ | -0.02 | (0.19) | 0.09 | (0.20) |
| Constant | -1.25 | (1.34) | 0.22 | (1.02) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.02 | | -0.13 | |
| Durbin-Watson | 2.04 | | 2.08 | |
| Breusch-Pagan Test ($\chi^2$) | 0.18 | | 0.46 | |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.13 | | 0.63 | |

*Note*: Entries are OLS coefficients (s.e. in parentheses). All variables have been fractionally differenced by their estimate of *d*.
$^\dagger$ FECM of Mood & Onion Acreage; FECM significance (*p$\leq$ 0.05, one-tail test).
Coefficient significance (*p$\leq$ 0.05, two-tail test).

## Table C.14: Three-Step FECM Nonsense Results - Republican Mood

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$Onion Acreage | 0.02 | 0.02 | 0.02 | 0.02 |
| | (0.05) | (0.05) | (0.05) | (0.05) |
| $\Delta^d$Coal Emissions | 0.02 | 0.02 | 0.02 | 0.02 |
| | (0.03) | (0.03) | (0.03) | (0.03) |
| $\Delta^d$Beef Consumption | 0.58 | 0.58 | 0.58 | 0.62 |
| | (0.39) | (0.39) | (0.39) | (0.39) |
| $\Delta^d$Shark Attacks | 0.02 | 0.02 | 0.02 | 0.01 |
| | (0.03) | (0.03) | (0.03) | (0.03) |
| $\Delta^d$Tornado Fatalities | 0.02 | 0.02 | 0.02 | 0.02* |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| **Error Correction and Constant** | | | | |
| FECM† | -0.04 | -0.02 | -0.03 | -0.11 |
| | (0.18) | (0.18) | (0.18) | (0.19) |
| Constant | -1.24 | -1.25 | -1.24 | -1.18 |
| | (1.32) | (1.32) | (1.33) | (1.32) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.03 | 0.02 | 0.03 | 0.04 |
| Durbin-Watson | 2.00 | 2.03 | 2.01 | 1.82 |
| Breusch-Pagan Test ($\chi^2$) | 0.15 | 0.16 | 0.16 | 0.04 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.02 | 0.09 | 0.03 | 1.99 |

*Note*: Entries are OLS coefficients. FECM significance (*p≤ 0.05, one-tail test). Coefficient significance (*p≤ 0.05, two-tail test).
† Model (1) - FECM of Mood & Coal Emissions; Model (2) - FECM of Mood & Beef Consumption; Model (3) - FECM of Mood & Shark Attacks; Model (4) - FECM of Mood & Tornado Fatalities

## Table C.15: Three-Step FECM Nonsense Results - Democrat Mood

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$Onion Acreage | 0.03 | 0.03 | 0.03 | 0.03 |
| | (0.04) | (0.04) | (0.04) | (0.04) |
| $\Delta^d$Coal Emissions | -0.00 | -0.00 | -0.00 | -0.00 |
| | (0.02) | (0.02) | (0.02) | (0.02) |
| $\Delta^d$Beef Consumption | -0.19 | -0.18 | -0.19 | -0.19 |
| | (0.32) | (0.31) | (0.31) | (0.31) |
| $\Delta^d$Shark Attacks | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.02) | (0.02) | (0.02) | (0.02) |
| $\Delta^d$Tornado Fatalities | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| **Error Correction and Constant** | | | | |
| FECM† | -0.00 | -0.05 | -0.02 | -0.02 |
| | (0.20) | (0.18) | (0.20) | (0.20) |
| Constant | 0.20 | 0.15 | 0.19 | 0.19 |
| | (1.04) | (1.04) | (1.03) | (0.20) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | -0.14 | -0.13 | -0.14 | -0.14 |
| Durbin-Watson | 1.91 | 1.80 | 1.88 | 1.87 |
| Breusch-Pagan Test ($\chi^2$) | 0.75 | 0.65 | 0.72 | 0.69 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.30 | 1.05 | 0.91 | 1.06 |

*Note*: Entries are OLS coefficients. FECM significance (*p≤ 0.05, one-tail test). Coefficient significance (*p≤ 0.05, two-tail test).
† Model (1) - FECM of Mood & Coal Emissions; Model (2) - FECM of Mood & Beef Consumption; Model (3) - FECM of Mood & Shark Attacks; Model (4) - FECM of Mood & Tornado Fatalities

### C.4.4 An Iterative Example of the ECM - Data from Ura and Ellis (2012)

We have previously discussed the fact that the ECM $t$-ratio is not dimension invariant, its distribution shifts with the number of regressors (Kremers, Ericsson, and Dolado 1992; Banerjee, Dolado, and Mestre 1998; Ericsson and MacKinnon 2002). In addition to the $t$-ratio however, the size of the $\alpha_1$ coefficient generally increases as covariates are added to the ECM equation. Monte Carlo simulations with randomly generated $I(1)$ data indicate that moving from a bivariate to a multivariate (5IV) model will double the ECM coefficient.[5] As we have noted, this bias is especially problematic because conclusions as to the strength of the equilibrium relationship are based in part on the size of the ECM coefficient.

To demonstrate this effect with real data, we use the *Democrat Mood* model from Ura and Ellis (2012) and present the results of an iterated ECM model. Adding one IV at a time helps to clarify the potential for inferential problems faced by researchers when using the ECM model. For purposes of this discussion, we treat ECM significance as defined by the standard $t$-test, which has been the common practice in political science.

**Table C.16: Iterative Estimation - Democrat Mood (Ura and Ellis 2012)**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Long Run Effects** | | | | | | |
| Unemployment$_{t-1}$ | | 0.12 | 0.30 | 0.29 | 0.65 | 0.34 |
| | | (0.22) | (0.23) | (0.30) | (0.43) | (0.37) |
| Domestic Spending \$10B$_{t-1}$ | | | 0.01* | 0.01 | -0.01 | -0.07* |
| | | | (0.01) | (0.01) | (0.02) | (0.02) |
| Inflation$_{t-1}$ | | | | -0.20 | -0.16 | -0.32* |
| | | | | (0.17) | (0.17) | (0.15) |
| Top 1% Income Share$_{t-1}$ | | | | | 0.48 | 1.07* |
| | | | | | (0.40) | (0.38) |
| Defense Spending \$10B$_{t-1}$ | | | | | | 0.20* |
| | | | | | | (0.06) |
| **Short Term Effects** | | | | | | |
| $\Delta$ Unemployment$_t$ | | -0.16 | -0.05 | 0.21 | 0.53 | 0.68 |
| | | (0.29) | (0.29) | (0.43) | (0.47) | (0.39) |
| $\Delta$ Domestic Spending \$10B$_t$ | | | -0.01 | -0.01 | -0.01 | -0.07 |
| | | | (0.04) | (0.04) | (0.04) | (0.04) |
| $\Delta$ Inflation$_t$ | | | | -0.05 | 0.03 | -0.03 |
| | | | | (0.18) | (0.20) | (0.17) |
| $\Delta$ Top 1% Income Share$_t$ | | | | | 0.78* | 0.76* |
| | | | | | (0.38) | (0.32) |
| $\Delta$ Defense Spending \$10B$_t$ | | | | | | 0.14 |
| | | | | | | (0.11) |
| **Error Correction and Constant** | | | | | | |
| DemMood$_{t-1}$ | -0.05 | -0.03 | -0.14 | -0.23 | -0.32* | -0.69* |
| | (0.06) | (0.08) | (0.09) | (0.13) | (0.14) | (0.15) |
| Constant | -3.26 | 1.12 | 6.36 | 13.62 | 12.59 | 33.93* |
| | (4.23) | (1.31) | (6.20) | (6.40) | (9.12) | (9.86) |
| **Fit and Diagnostics** | | | | | | |
| Adjusted $R^2$ | -0.01 | -0.04 | 0.04 | 0.02 | 0.10 | 0.39 |
| Breusch-Pagan Test ($\chi^2$) | | 7.03* | 0.20 | 0.73 | 2.76 | 0.40 |
| Breusch-Godfrey LM Test ($\chi^2$) | | 1.28 | 1.04 | 2.49 | 0.41 | 0.70 |

*Note*: Dependent Variable is the change in *Democrat Mood*. Entries are OLS coefficients (standard errors in parentheses). ECM significance (*p$\leq$ 0.05, one-tail test). Coefficient significance (*p$\leq$ 0.05, two-tail test). N=35

Column (1) is the result of a standard Dickey-Fuller test. *Democrat Mood* is not sta-

[5] See Table 2 of the main paper, or Appendices C.7&C.8 of the Supplement for simulation results.

tionary and in fact has a very long memory. *Unemployment* is insignificant across both models of Ura and Ellis (2012) and column (2) finds it insignificant as well when included in a bivariate model. By adding *Domestic Spending* in (3), we find our first significant long run variable, and we see that our ECM has now increased (in absolute terms) by 0.11, however it would still be considered insignificant were we using a standard $t$-test. Column (4) adds *Inflation* to our model, which is now completely insignificant. With the inclusion of *Inflation* our adjusted $R^2$ has decreased - we have been punished for its addition - yet our ECM has actually increased by an additional 0.09. We next add *Top 1% Income* and with its inclusion our model boasts one significant short-term IV. But the inclusion of a fourth IV has also increased the size of our ECM by yet another 0.09. The ECM coefficient has now increased by over 600% from the bivariate regression despite a lack of any significant long run variables. Finally, in column (6) the inclusion of *Defense Spending* has completely blown apart the consistency of our model. We now have 4 significant long run IVs - of the three that were previously insignificant in (5), their coefficients have more than doubled - and our adjusted $R^2$ has jumped substantially. More importantly, our ECM has also doubled in size and now sits at an impressive -0.69.

What conclusions should we draw from this exercise? First, we see why the ECM's $t$-distribution shifts with each added regressor. The absolute value of the coefficient increases - substantially with influential variables - but the standard errors increase much more slowly. Considering that the ECM is simply the dependent variable at $t-1$ predicting changes in itself at time $t$, such small increases in its standard errors should not come as a surprise. Second, the ECM's role as an identifier of a close, equilibrating relationship between an IV and DV has been rendered meaningless. The difference between columns (5) and (6) is the inclusion of one influential variable. Were we to stop at (5) we would draw the conclusion that our model was null. Only short-term changes in income inequality significantly affect Democrat mood. But with the inclusion of *Defense Spending* we gain 4 significant long-run variables, each of which supposedly share a close relationship with the dependent variable. Notably, the equilibrium relationship that *every significant regressor* supposedly shares with the DV has massively increased in strength. The inclusion of one influential variable has increased the absolute value of the ECM coefficient by 0.37.[6]

[6]Were we to first test and confirm that all variables in the model were $I(1)$, in practice we would then use the MacKinnon critical values in order to determine if cointegration was present. In this specific case, application of those values would lead to a finding that the ECM was insignificant. Without evidence of cointegration the model is misspecified and it would have to be run first-differences. The use of the model in this way would be just as Hendry and Mizon (1978) proposed. Running the regression in first differences would yield null results for the Democrat Mood model.

## C.5 Additional Replications

### C.5.1 Example 1: Sanchez Urribarri, Schorpp, Randazzo, and Songer (*The Journal of Politics, 2011)*, Rights Litigation

We next replicate a recent article by Sanchez Urribarri et al. (2011) in which the authors attempt to model the causes behind an increased propensity of the high courts of Canada, UK, and US to place "rights cases" on their dockets. We are agnostic as to causation, however we choose to replicate this article because it very clearly demonstrates the potential for inferential problems when using the ECM with a stationary, or short-memoried, dependent variable. The article has three models, for which the ECMs and their respective standard errors are: Canada ($\alpha_1 = -0.88$, s.e. $= 0.18$), UK ($\alpha_1 = -1.13$, s.e. $= 0.17$), and US ($\alpha_1 = -1.17$, s.e. $= 0.14$). These values draw our attention because it should be impossible to have an error correction rate that is greater than 100%. Our Monte Carlo simulations demonstrate how such results occur.[7]

To begin, we replicate the full results from Sanchez Urribarri et al. (2011) below. As seen in the results, a number of significant long-run effects are significant, but each model also has at least one significant short-run effect as well. Interestingly, the models that are fit by Sanchez Urribarri et al. include only one continuous independent variable, *Judicial Ideology.* The rest are coded either as ordinal (*Support Structure*), or as permanent interventions.

The stationarity of the dependent variables is of interest and Table C.18 provides the Dickey-Fuller test results for each of the three dependent variables used by the authors. From the results, it's clear that we have three different types of DV - *Canada* cannot reject the null of a unit-root and appears strongly autoregressive, the *UK* is very stationary, and the *US* is stationary, however it also indicates some level of autoregression.

We next regress the dependent variables from each model on randomly generated series of white noise independent variables. Recall from Appendix C.1 that regressing the DV on non-influential independent variables is equivalent to imposing the common factor restrictions of the Dickey-Fuller test (Kremers, Ericsson, and Dolado 1992). The Canada DV is regressed on four white noise series, the UK DV on three, and the US DV on five. The results are found in Table C.19

This exercise raises several points worth discussing. First, note the similarities between the $\alpha_1$ coefficient in Table C.19 and the Dickey-Fuller coefficient results in Table C.18. With stationary series, the baseline for the $\alpha_1$ parameter is its DF coefficient - the more stationary the DV, the larger the absolute size of error correction found. Second, our two stationary series are guaranteed a significant ECM when applying a one-tail test, and even a non-stationary series such as the Canada model finds over half of all ECM's as significant when applying a standard one-tail test. Do note the work that the MacKinnon values are doing however. With $I(0)$ covariates we should not expect to find evidence of cointegration, and in both the Canada and US models, the MacKinnon values properly exclude that possibility. On the other hand, because of the stationary properties of the dependent variable in the UK model, the DV is automatically in equilibrium - the model is simply capturing its tendency toward mean reversion. Third, even though we simulate white-noise independent variables,

[7] At 60 observations, the U.S. model is longer than our beef-shark-tornado data set. Rather than present an incomplete re-estimation of only the Canadian and UK series, we limit the replication to Monte Carlo simulations and FI estimations.

**Table C.17: Replication of Table 1 - Sanchez Urribarri et al. (2011)†**

| **Model‡** | Canada | UK | US |
|---|---|---|---|
| **Long Run Effects** | | | |
| Support Structure$_{t-1}$ | -0.00 (0.01) | 0.01 (0.01) | 0.00 (0.00) |
| Judicial Ideology$_{t-1}$ | -0.17 (0.18) | -0.23* (0.12) | .22* (0.09) |
| Docket Control$_{t-1}$ | 0.17* (0.06) | | |
| Charter$_{t-1}$ | 0.21* (0.07) | | |
| Human Rights Act$_{t-1}$ | | 0.05 (0.04) | |
| *Brown*$_{t-1}$ | | | 0.03 (0.03) |
| *Mapp*$_{t-1}$ | | | 0.10* (0.05) |
| Civil Rights Act$_{t-1}$ | | | 0.10* (.05) |
| **Short Run Effects** | | | |
| Δ Support Structure | 0.02 (0.03) | -0.04 (0.03) | 0.01 (0.02) |
| Δ Judicial Ideology | -0.21 (0.14) | -0.15* (.07) | 0.20* (0.08) |
| Δ Docket Control | -0.04 (0.10) | | |
| Δ Charter | 0.16* (0.09) | | |
| Δ Human Rights Act | | -0.09 (0.08) | |
| Δ *Brown* | | | -0.06 (0.06) |
| Δ *Mapp* | | | -0.05 (0.06) |
| Δ Civil Rights Act | | | -0.05 (0.07) |
| **Error Correction and Constant** | | | |
| Rights Agenda$_{t-1}$ | -0.88* (0.18) | -1.13* (0.17) | -1.17* (0.14) |
| Constant | 0.20* (0.06) | 0.40* (0.08) | 0.26* (0.05) |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.43 | 0.55 | 0.54 |
| Breusch-Pagan Test ($\chi^2$) | 1.70 | 0.20 | 1.94 |
| Breusch-Godfrey Test ($\chi^2$) | 2.78 | 0.46 | 0.16 |
| N | 35 | 36 | 60 |

† *Note*: Dependent variable Δ*Rights Agenda* represents changes to the rights agenda on high court dockets. Significance of ECM and coeffcients ($^{\star}p < .10$, two-tail test)

the UK and US models suffer from spurious regressions - approximately 13 percent of all models produce at least one significant long-run covariate along with significant error correction. These rates are much too high for randomly generated data, and certainly too high for data generated as white noise processes.

These results arise from the simplest models including only white-noise independent variables. The use of unit-root independent variables complicates matters further, and the results are found in Table C.20. As before, we replace the IVs used by Sanchez Urribarri et al. (2011), this time with series generated as $I(1)$.

The results from Table C.20 indicate the inferential risks from the GECM, even with MacKinnon values. First, both the UK and US models are unbalanced - each DV is stationary, and the inclusion of unit-root IVs has compromised the results. The $\alpha_1$ parameter for the UK model is below the theoretical limit of -1.00 and the US model is approaching it. But according to our Dickey-Fuller test results, the DV in the Canada model is a unit-root. This means that were we to apply a dichotomous $I(1)/I(0)$ value to our DV, the results in the first column would be considered as coming from a balanced model. Despite this fact, the model performs terribly. First, were we to assume that a one-tail test was appropriate, we'd be almost guaranteed a significant ECM, and a high rate of spurious regressions means at least one long-run variable is significant in over 68% of our models. But even were we to use MacKinnon values, a significance rate north of 35% for the ECM parameter means

### Table C.18: Dickey-Fuller Results of Three DVs†

| **Model** | Canada | UK | U.S. |
|---|---|---|---|
| Dickey-Fuller coefficient | -0.18 | -0.86 | -0.34 |
| Dickey-Fuller $t$-statistic | -1.99 | -5.07* | -3.68* |
| MacKinnon DF Critical Value | -2.97 | -2.97 | -2.92 |
| Estimated $d$ Value | 0.53 (0.13) | 0.13 (0.13) | 0.45 (0.05) |

† *Note*: Dickey-Fuller critical values are from MacKinnon (1994). $d$ values estimated by Stata's exact ML estimator.

### Table C.19: Monte Carlo Results with I(0) IVs and DVs of Sanchez Urribarri et al. (2011)†

| **Model**‡ | Canada | UK | US |
|---|---|---|---|
| % ECM Significant - one tail $t$-distribution | 55.9 | 100 | 100 |
| % ECM Significant - MacKinnon Values | 0 | 95.4 | 0.02 |
| Mean of $\alpha_1$ | -0.18 | -0.84 | -0.34 |
| Mean $t$-statistic of $\alpha_1$ | -1.71 | -4.63 | -3.39 |
| % ECM & $\geq 1\Delta X_t$ Significant | 10.3 | 15.2 | 23.4 |
| % ECM & $\geq 1X_{t-1}$ Significant | 4.8 | 14.4 | 13.3 |

† *Note*: Results based on 10,000 simulations of each model. ‡ All IVs are level stationary (integrated at $I(0)$). MacKinnon Values are the ECM critical values from Ericsson and MacKinnon (2002) - for Canada (4 IVs, T=35 CV=-4.082); for UK (3IVs, T=36 CV=-3.867); for US (5IVs, T=60 CV=-4.229)

the critical values offer little assistance. It appears that the Canada DV is confounding the model despite the Dickey-Fuller test results leading us to consider it a unit-root.

Tables C.19 and C.20 represent two extremes in that our IVs are generated as either $I(0)$ or $I(1)$ series. To what extent are these results being driven by the models being unbalanced?[8] To find out, we next simulate stationary series to act as our IVs - the DGP is a simple autoregressive process, $X_t = \rho X_{t-1} + \epsilon_t$ with $\rho$ set to 0.75. The results are in Table C.21.

Table C.21, columns 2 and 3, indicates that regression balance may help somewhat, but our issues are not completely solved. Because the US and UK DVs are both stationary, we apply the standard one-tail $t$-test and each model finds that 100% of ECMs are significant. As noted in the main text, this is to be expected with stationary dependent variables, but spurious regressions are still an issue. At least one long-run IV significant in 25-35% of the models.

As a final test in our replication, we re-estimate the three models after accounting for the fractional integration of the data. The orders of fractional integration of the dependent variables can be found in Table C.18. The estimated order of fractional integration of *Judicial Ideology* is as follows: Canada ($d = 0.32$, s.e.$= 0.11$), UK ($d = -0.13$, s.e.$= 0.16$), US ($d = 0.62$, s.e.$= 0.12$).

The results of the fractionally differenced models are drastically different from the original ECM models. Of the three re-estimated models only one IV - *Judicial Ideology* in the United Kingdom model - is significant. Its coefficient size is equivalent to that found in the ECM

[8]Recall that according to the DF results the Canada model in Table C.20 is balanced however.

**Table C.20: Monte Carlo Results with I(1) IVs and DVs of Sanchez Urribarri et al. (2011)†**

| **Model‡** | Canada | UK | US |
|---|---|---|---|
| % ECM Significant - one tail $t$-distribution | 97.9 | 100 | 100 |
| % ECM Significant - MacKinnon Values | 38.9 | 99.7 | 92.6 |
| Mean of $\alpha_1$ | -0.71 | -1.01 | -0.81 |
| Mean $t$-statistic of $\alpha_1$ | -3.82 | -5.50 | -5.78 |
| % ECM & $\geq 1\Delta X_t$ Significant | 32.3 | 17.4 | 34.6 |
| % ECM & $\geq 1X_{t-1}$ Significant | 68.4 | 31.7 | 82.1 |

† *Note*: Results based on 10,000 simulations of each model.
‡ All IVs are unit-roots (integrated at $I(1)$). MacKinnon Values are the ECM critical values from Ericsson and MacKinnon (2002) - for Canada (4 IVs, T=35 CV=-4.082); for UK (3IVs, T=36 CV=-3.867); for US (5IVs, T=60 CV=-4.229)

**Table C.21: Monte Carlo Results with Stationary IVs and DVs of Sanchez Urribarri et al. (2011)†**

| **Model‡** | Canada | UK | US |
|---|---|---|---|
| % ECM Significant - one tail $t$-distribution | 78.8 | 100 | 100 |
| % ECM Significant - MacKinnon Values | 0.3 | 99.0 | 33.1 |
| Mean of $\alpha_1$ | -0.29 | -0.94 | -0.47 |
| Mean $t$-statistic of $\alpha_1$ | -2.06 | -5.15 | -4.04 |
| % ECM & $\geq 1\Delta X_t$ Significant | 16.8 | 16.4 | 26.6 |
| % ECM & $\geq 1X_{t-1}$ Significant | 10.8 | 26.0 | 35.8 |

† *Note*: Results based on 10,000 simulations of each model.
‡ All IVs are stationary AR processes with $\rho = 0.75$. MacKinnon Values are the ECM critical values from Ericsson and MacKinnon (2002) - for Canada (4 IVs, T=35 CV=-4.082); for UK (3IVs, T=36 CV=-3.867); for US (5IVs, T=60 CV=-4.229)

model. Because the dependent variable in the UK model has an estimated order of fractional integration of $d = 0.13$ and the estimate of *Judicial Ideology* is $d = -0.13$, neither of which are statistically different from $I(0)$, the two series are not candidates for fractional cointegration.

## Table C.22: FI Model Results - Sanchez Urribarri et al. (2011)[†]

| **Model**[‡] | Canada | UK | US |
|---|---|---|---|
| $\Delta^d$ Judicial Ideology | -0.22 (0.15) | -0.14* (.068) | 0.05 (0.08) |
| $\Delta$ Support Structure | 0.05 (0.04) | -0.04 (0.03) | 0.01 (0.02) |
| $\Delta$ Docket Control | -0.14 (0.09) | | |
| $\Delta$ Charter | 0.13 (0.08) | | |
| $\Delta$Human Rights Act | | -0.10 (0.07) | |
| $\Delta$*Brown* | | | -0.10 (0.07) |
| $\Delta$*Mapp* | | | -0.04 (0.07) |
| $\Delta$ Civil Rights Act | | | -0.00 (0.07) |
| Constant | 0.05 (0.02) | 0.09 (0.02) | 0.04 (0.07) |
| **Fit and Diagnostics** | | | |
| Adjusted $R^2$ | 0.17 | 0.21 | 0.05 |
| Breusch-Godfrey Test ($\chi^2$) | 2.69 | 0.74 | 1.27 |
| Breusch-Pagan Test ($\chi^2$) | 6.02 | 7.40 | 3.19 |
| N | 35 | 36 | 59 |

[†] *Note*: Entries are OLS coefficients (standard errors in parentheses). DV is $\Delta^d$*Rights Agenda*. All variables have been fractionally differenced by their estimate of *d*. Coefficient significance (*p$\leq$ 0.05, two-tail test).

### C.5.2 Example 2: Kelly and Enns (*American Journal of Political Science*, 2010), Inequality and Mood

This replication concerns the article by Kelly and Enns (2010), which uses the indicator *Public Mood* to argue that economic inequality is self-reinforcing. Kelly and Enns (KE) use the GECM model to estimate the effects of economic inequality on public mood while controlling for both policy liberalism as well as a few objective economic indicators. In terms of mass preferences, there is a negative response to shifts in economic inequality. The public becomes less supportive of government intervention when income inequality rises.

Replication of their findings are found in Tables C.23 and C.24 below. The results from their Table 1 (our C.23) indicate consistently significant effects of both *Policy Liberalism* and *Income Inequality*. The ECM for *Liberal Mood* is consistently estimated with a value of -0.25, and the ECM of their DV, *Welfare Support*, indicates a very strong connection between the IVs.

**Table C.23: Replication of Kelly & Enns - Table 1**

| **Independent Variables** | (1) **Δ Liberal Mood** | (2) **Δ Liberal Mood** | (3) **Δ Liberal Mood** | (4) **Δ Support Welfare** |
|---|---|---|---|---|
| **Long Run Effects** | | | | |
| Policy Liberalism$_{t-1}$ | -0.07* | -0.09* | -0.07* | -0.65* |
| | (0.04) | (0.02) | (0.03) | (0.22) |
| Income Inequality$_{t-1}$ | | -16.22* | -18.00* | -152.17* |
| | | (8.92) | (9.44) | (65.71) |
| Unemployment$_{t-1}$ | 0.04 | | 0.10 | |
| | (0.25) | | (0.26) | |
| Inflation$_{t-1}$ | -0.08 | | -0.13 | |
| | (0.18) | | (0.17) | |
| **Short Run Effects** | | | | |
| Δ Policy Liberalism$_t$ | 0.16 | 0.10 | 0.10 | -0.24 |
| | (0.11) | (0.10) | (0.11) | (0.42) |
| Δ Income Inequality$_t$ | | 27.07 | -29.56 | -175.70 |
| | | (34.61) | (37.59) | (122.56) |
| Δ Unemployment$_t$ | -0.00 | | 0.01 | |
| | (-0.38) | | (0.38) | |
| Δ Inflation$_t$ | -0.12 | | -0.13 | |
| | (0.19) | | (0.19) | |
| **Error Correction and Constant** | | | | |
| Liberal Mood$_{t-1}$ | -0.25* | -0.25* | -0.26* | |
| | (0.07) | (0.07) | (0.07) | |
| Support Welfare$_{t-1}$ | | | | -0.55* |
| | | | | (0.16) |
| Constant | 15.41* | 21.70* | 22.99* | 80.03* |
| | (4.54) | (5.57) | (6.31) | (29.92) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.20 | 0.28 | 0.24 | 0.26 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.88 | 0.71 | 0.60 | 0.24 |
| Breusch-Pagan Test ($\chi^2$) | 0.75 | 0.49 | 0.30 | 3.44 |
| N | 54 | 54 | 54 | 33 |

*Note*: Entries are OLS regression coefficients (standard errors in parentheses). Two-tailed significance levels: *p<.10)

Replication of their second table (our C.24) provides the most interesting results of the paper. KE disaggregate public mood by income group and find that regardless of the level of income, the population responds in similar fashion to shifts in economic inequality. As income inequality rises, the public - whether rich or poor - responds in a conservative style. As they note, these findings contradict the theoretical predictions of both Benabou (2000) and Meltzer and Richard (1981).

### Table C.24: Replication of Kelly & Enns - Table 2

| | Δ Liberal Mood | | | |
|---|---|---|---|---|
| **Independent Variables** | **Low Income** | **High Income** | **Low Income** | **High Income** |
| **Long Run Effects** | | | | |
| Policy Liberalism$_{t-1}$ | -0.28* | -0.24* | -0.26* | -0.23* |
| | (0.08) | (0.07) | (0.08) | (0.07) |
| Income Inequality$_{t-1}$ | -48.65* | -44.26* | -64.57* | -61.98* |
| | (22.23) | (21.09) | (25.97) | (24.38) |
| Unemployment$_{t-1}$ | | | -0.33 | -0.45 |
| | | | (0.47) | (0.46) |
| Inflation$_{t-1}$ | | | -0.39 | -0.33 |
| | | | (0.34) | (0.31) |
| **Short Run Effects** | | | | |
| Δ Policy Liberalism$_t$ | -0.16 | 0.07 | -0.22 | 0.01 |
| | (0.15) | (0.13) | (0.17) | (0.15) |
| Δ Income Inequality$_t$ | 19.43 | -46.76 | -22.56 | -41.01 |
| | (53.05) | (42.99) | (55.18) | (44.14) |
| Δ Unemployment$_t$ | | | 0.14 | -0.13 |
| | | | (0.63) | (0.53) |
| Δ Inflation$_t$ | | | -0.46 | -0.19 |
| | | | (0.32) | (0.28) |
| **Error Correction and Constant** | | | | |
| Liberal Mood$_{t-1}$ | -0.48* | -0.46* | -0.58* | -0.57* |
| | (0.12) | (0.11) | (0.14) | (0.13) |
| Constant | 50.25* | 44.18 | 66.00 | 61.54 |
| | (14.91) | (12.93) | (19.56) | (17.03) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.20 | 0.25 | 0.18 | .023 |
| rho | 0.19 | 0.32 | 0.17 | 0.33 |
| N | 50 | 50 | 50 | 50 |

*Note*: Entries are Prais-Winsten regression coefficients (standard errors in parentheses). Two-tailed significance levels: *p<.10)

Before we begin our replication, we first check the order of integration of the dependent variables. From their Table 1, we have *Public Mood* and *Welfare Support*. From their Table 2, we test *Low Income Mood* and *High Income Mood*. None of the four reject the null hypothesis of a unit-root, although Mood is edging up to the 10% level of significance.

### Table C.25: Dickey-Fuller Results of K&E DVs†

| **Model** | **Liberal Mood** | **Welfare Support** | **Low Income Mood** | **High Income Mood** |
|---|---|---|---|---|
| Dickey-Fuller Coefficient | -0.18 | -0.22 | -0.11 | -0.09 |
| Dickey-Fuller *T*-statistic | -2.56 | -1.88 | -2.02 | -1.60 |
| MacKinnon DF Critical Value | -2.93 | -2.98 | -2.93 | 2.93 |
| Estimated *d* Value | 1.09 (0.11) | 0.88 (0.18) | 0.98 (0.13) | 1.04 (0.13) |

† *Note*: Dickey-Fuller critical values are from MacKinnon (1994). *d* values estimated by Stata's exact ML estimator. Standard errors in parentheses.

We next simulate our own independent variables in order to test the Type I error rates generated by the GECM models with the dependent variables used by KE. With randomly generated $I(1)$ series, we should find a 5% rejection rate on the significance of the error correction parameter. The results in Table C.26 indicate higher than expected rejection rates however. First, we clearly see the dangers of failing to account for ($\alpha_1$) $t$-statistic's non-standard distribution. Use of a one-tail test practically guarantees a significant ECM in all four models of Table 1. Fortunately, applying the MacKinnon values cuts down on the Type I errors, however with *Liberal Mood* as the DV, the model still tends towards significance - approximately 10% of all models claim an ECM as significant. On the other hand, Model 4, and its DV of *Welfare Support*, appears to understate the significance of the ECM when MacKinnon critical values are applied.

In Table C.26 the coefficient of the ECM is fairly consistent in Models 1 through 3 which is not unsurprising considering that the only variance between each model is the number of unrelated, simulated regressors. Model 4 sees a fairly substantial jump in coefficient size however. This is consistent with our simulations of an $I(d)$ model when $d = .9$.[9] Additional downward bias in Model 4's ($\alpha_1$) parameter is likely due to the limited number of observations. Finally, the ECM model suffers from substantial spurious regression problems. In Models 1-3, over half of all simulated models find both a significant error correction parameter and long-run regressor.

**Table C.26: Monte Carlo Results - K&E Table 1 - DVs on I(1) IVs†**

| **Model** | (1) **Δ Liberal Mood** | (2) **Δ Liberal Mood** | (3) **Δ Liberal Mood** | (4) **Δ Support Welfare** |
|---|---|---|---|---|
| % ECM Significant - one tail $t$-distribution | 94.9 | 96.9 | 92.9 | 85.3 |
| % ECM Significant - MacKinnon Values | 10.6 | 11.7 | 9.2 | 1.3 |
| Mean of $\alpha_1$ | -0.19 | -0.18 | -0.21 | -0.29 |
| % ECM & $\geq 1\Delta X_t$ Significant | 18.8 | 13.4 | 24.3 | 11.1 |
| % ECM & $\geq 1X_{t-1}$ Significant | 57.1 | 49.0 | 59.4 | 17.8 |

† *Note*: Entries provide the summary results of 10,000 simulations of an OLS regression in which the *Liberal Mood* and *Welfare Support* are regressed on randomly generated $I(1)$ series. Number of IVs match those of K&E. Models 1-3 regress *Mood* on IVs ($T = 54$). Model 4 DV is *Welfare Support* ($T = 35$). Significance of IVs based on two-tail test. MacKinnon CV Model 1: -3.822; Model 2: -3.57; Model 3: -4.040; Model 4: -3.621

But as we have seen, the independent variables also affect the ECM estimates. Of particular interest in this case is any potential bias of the ($\alpha_1$) $t$-statistic when the independent variables' order of fractional integration, $d$, exceeds 1. With *Policy Liberalism* estimated at $d = 1.35$ (s.e.$= 0.10$), the significance of error correction in these simulations is expected to be inflated. In Table C.27 we simulate our dependent variables and regress them on the IVs of Kelly and Enns. To better approximate a series such as *Mood* we use the bounded DGP of Nicolau (2002) to generate one DV and use it in Models 1 through 3. Because *Welfare Support* is less well defined in its boundaries, the DGP for the Model 4 dependent variable is left as a unit-root process.

The results in Table C.27 are as expected. Even when applying the MacKinnon values each model rejects the null of a significant ECM at a rate exceeding 5%. The potential for outsized influence of a particular IV is also on display. Compare the results of Models 2 and 3 to Model 1. Despite the same randomly generated DV across all three models, the

[9]See Section C.8.3 for simulation results based on varying orders of fractional integration.

**Table C.27: Monte Carlo Results - K&E Table 1 - BI(1) DVs on IVs†**

| Model | (1) BI(1) DV1 | (2) BI(1) DV1 | (3) BI(1) DV1 | (4) I(1) DV2 |
|---|---|---|---|---|
| % ECM Significant - one tail $t$-distribution | 65.6 | 86.8 | 89.4 | 75.8 |
| % ECM Significant - MacKinnon Values | 6.9 | 17.1 | 10.7 | 8.2 |
| Mean of $\alpha_1$ | -0.20 | -0.27 | -0.32 | -0.32 |
| % ECM & $\geq 1\Delta X_t$ Significant | 13.5 | 13.2 | 23.7 | 12.3 |
| % ECM & $\geq 1X_{t-1}$ Significant | 21.0 | 43.8 | 54.0 | 31.9 |

† *Note*: Entries provide the summary results of 10,000 simulations of an OLS regression in which randomly generated $I(1)$ DVs are regressed on the IVs from Table 1 of Kelly and Enns (2010). Significance of IVs based on two-tail test. MacKinnon CV Model 1: -3.822; Model 2: -3.57; Model 3: -4.040; Model 4: -3.621

inclusion of *Income Inequality* in Models 2 and 3 substantially increases the number of Type I errors and inflates the ECM coefficient of each model as well.

Replicating Table 2 of KE's paper, we chose to randomly generate two $I(1)$ variables for use as dependent variables and to run our MC simulation with the Prais-Winsten estimator.[10] The effect size due to estimator choice is striking. While not directly comparable, the number of significant ECMs using the Prais-Winsten estimator in Table C.28 is twice the size of those when using OLS.

**Table C.28: Monte Carlo Results - K&E Table 2 - I(1) DVs on IVs†**

| | Model 1 & 2 | Model 3 & 4 |
|---|---|---|
| % ECM Significant – one tail $t$-distribution | 86.1 | 84.7 |
| % ECM Significant – MacKinnon Values | 31.1 | 21.8 |
| Mean of $\alpha_1$ | -0.37 | -0.37 |
| % ECM & $\geq \Delta X_t$ Significant | 18.2 | 28.8 |
| % $\geq 1X_{t-1}$ Significant | 51.4 | 59.2 |

*Note*: Entries provide are the summary results of 10,000 simulations of a Prais-Winsten regression in which randomly generated $I(1)$ series are regressed upon the independent variables from Table 2 of Kelly and Enns (2010). Coefficient significance (*p≤ 0.05, two-tail test). MacKinnon CV Models 1&2: -3.578; Models 3&4: -4.047.

To put the ECM model's susceptibility to Type I errors into context, we use our nonsense data sets from the main paper in place of the IVs in the model. Unfortunately, because of the length of the series, we are unable to use our *Shark Attack* series. Regardless, as with all other replications, we find a high rate of significance. Fortunately, we finally find a model in which *Beef Consumption* is not a significant predictor. In Model 4, our two nonsense series do not predict *Welfare Support*, but note that had we been following standard practice in political science and using the normal $t$-test, we would conclude there was significant ECM despite no significant covariates.

For completeness, we also employ the same nonsense series with KE's dependent variables from their Table 2. Based on our past replications and simulations we expect spurious

[10] According to Engle and Granger (1987, p.264), the use of "serial correlation correction" such as a Cochrane Orcutt or Prais-Winsten estimator in the cointegrating regression will produce inconsistent estimates. Kelly and Enns use the Prais-Winsten estimator despite Breusch-Godfrey test $p > \chi^2$ values of .1748, .0634, .1828, and .1032 for OLS estimation of Table 2, Models 1-4 respectively.

## Table C.29: What Else Moves Mood? (Re-estimation of Table 1)

| Independent Variables | (1) Δ Liberal Mood | (2) Δ Liberal Mood | (3) Δ Liberal Mood | (4) Δ Support Welfare |
|---|---|---|---|---|
| **Long Run Effects** | | | | |
| Beef Consumption$_{t-1}$ | -0.10* | -0.19* | -0.36* | -0.94 |
| | (0.04) | (0.07) | 0.09 | 0.80 |
| Coal Emissions$_{t-1}$ | | 0.01* | 0.04* | 0.06 |
| | | (0.01) | 0.01 | (0.04) |
| Tornado Fatalities$_{t-1}$ | -0.00 | | -0.01 | |
| | (0.00) | | 0.00 | |
| Onion Acreage$_{t-1}$ | 0.17 | | -0.78* | |
| | (0.15) | | 0.32 | |
| **Short Run Effects** | | | | |
| Δ Beef Consumption$_t$ | 0.29 | 0.20 | 0.10 | 0.05 |
| | (0.30) | (0.30) | 0.29 | 1.01 |
| Δ Coal Emissions$_t$ | | 0.01 | 0.01 | -0.08 |
| | | 0.01 | (0.02) | (0.08) |
| Δ Tornado Fatalities$_t$ | 0.00 | | 0.00 | |
| | (0.00) | | 0.00 | |
| Δ Onion Acreage$_t$ | 0.28 | | -0.12 | |
| | (0.39) | | (0.38) | |
| **Error Correction and Constant** | | | | |
| Liberal Mood$_{t-1}$ | -0.18* | -0.17* | -0.20* | |
| | (0.06) | (0.06) | (0.05) | |
| Support Welfare$_{t-1}$ | | | | -0.28* |
| | | | | (0.13) |
| Constant | 12.99* | 12.66* | 20.91* | 25.56 |
| | (4.49) | (3.87) | (4.77) | (20.37) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.26 | 0.21 | 0.38 | 0.12 |
| Breusch-Pagan Test ($\chi^2$) | 0.74 | 0.15 | 1.07 | 3.40 |
| Breusch-Godfrey LM Test ($\chi^2$) | 1.25 | 0.00 | 0.27 | 0.11 |

*Note*: Entries are OLS regression coefficients (standard errors in parentheses). ECM significance (*p≤ 0.05, one-tail test). Coefficient significance: two-tailed: *p<.10).

regressions, and that is what we find.

Finally, we test the Kelly and Enns model for fractional cointegration and fractional error correction. As with all other replications, we first provide the order of integration of the IVs as well as the residuals of the static regressions in Table C.31. We do find some reduction in the residual's order of fractional integration, however a calculation of the confidence intervals surrounding these estimates indicates overlap between the residual estimates and the estimates of each constituent series.

In Tables C.32 and C.33 we report the results of the fractional ECM models. If two IVs were significant in a model, we included them both in the cointegrating equation. Whether we estimated the cointegration individually, or with both *Policy Liberalism* and *Income Inequality* as a part of the cointegrating vector, the results are the same. We fail to find any evidence of fractional cointegration in any model, and we also fail to find any significant independent variables.

## Table C.30: What Else Moves Mood? (Re-estimation of Table 2)

| Variables | (1) Δ Low Income Mood | (2) Δ High Income Mood | (3) Δ Low Income Mood | (4) Δ High Income Mood |
|---|---|---|---|---|
| **Long Run Effects** | | | | |
| Beef Consumption$_{t-1}$ | -0.51* | -0.46* | -0.53* | -0.70* |
| | (0.24) | (0.18) | 0.31 | 0.24 |
| Coal Emissions$_{t-1}$ | 0.04* | 0.04* | 0.05 | 0.07* |
| | (0.02) | (0.01) | (0.03) | (0.03) |
| Tornado Fatalities$_{t-1}$ | | | -0.00 | -0.00 |
| | | | 0.01 | (0.01) |
| Onion Acreage$_{t-1}$ | | | -0.49 | -0.87 |
| | | | (0.58) | (0.53) |
| **Short Run Effects** | | | | |
| Δ Beef Consumption$_t$ | -0.75 | -0.34 | -0.76 | -0.49 |
| | (0.46) | (0.41) | 0.50 | (0.45) |
| Δ Coal Emissions$_t$ | -0.02 | 0.01 | -0.02 | 0.02 |
| | (0.04) | 0.03 | (0.04) | (0.03) |
| Δ Tornado Fatalities$_t$ | | | 0.01 | 0.00 |
| | | | (0.01) | (0.00) |
| Δ Onion Acreage$_t$ | | | -1.01 | -0.36 |
| | | | (0.70) | (0.64) |
| **Error Correction and Constant** | | | | |
| Liberal Mood$_{t-1}$ | -0.31* | -0.25* | -0.30* | -0.30* |
| | (0.11) | (0.09) | (0.12) | (0.10) |
| Constant | 27.10* | 20.83* | 28.82* | 30.82* |
| | (10.18) | (7.85) | (13.19) | (10.54) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | 0.15 | 0.21 | 0.16 | 0.09 |
| Breusch-Pagan Test ($\chi^2$) | 0.03 | 0.51 | 0.09 | 0.37 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.83 | 0.92 | 0.96 | 0.28 |

*Note*: Entries are OLS regression coefficients (standard errors in parentheses). ECM significance (*p≤ 0.05, one-tail test). Coefficient significance: two-tailed: *p<.10).

## Table C.31: FECM Order of Integration - Kelly and Enns (2010)

| | Individual Series *d* | Residuals Liberal Mood *d* | Residuals Welfare *d* | Residuals Low Inc. Mood *d* | Residuals High Inc. Mood *d* |
|---|---|---|---|---|---|
| Liberal Mood | 1.08 (0.11) | | | | |
| Welfare Support | 0.88 (0.18) | | | | |
| Low Income Mood | 0.98 (0.13) | | | | |
| High Income Mood | 1.04 (0.13) | | | | |
| Policy Liberalism | 1.35 (0.09) | 1.06 (0.11) | 0.74 (0.19) | 0.65 (0.18) | 0.80 (0.16) |
| Income Inequality | 0.80 (0.10) | 1.09 (0.11) | 0.89 (0.18) | 1.00 (0.13) | 1.06 (0.13) |
| Unemployment | 0.83 (0.14) | | | | |
| Inflation | 0.95 (0.16) | | | | |
| Policy & Inequality | | 1.04 (0.11) | 0.66 (0.20) | 0.67 (0.18) | 0.79 (0.15) |

† *Note*: Entries are estimates of the order of integration of all significant long-run variables from Kelly and Enns (2010) as well as estimates of the residuals of static bivariate regressions of the DV and each significant variable. Estimates generated by Stata's exact ML estimator. Standard errors in parentheses

## Table C.32: Three-Step FECM Results - Kelly and Enns (2010)†

| **Table 1 Models** | **Liberal Mood** | **Liberal Mood** | **Liberal Mood** | **Welfare Support** |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$ Policy Liberalism | 0.09 | 0.09 | 0.09 | 0.10 |
| | (0.11) | (0.11) | (0.11) | 0.14 |
| $\Delta^d$ Income Inequality | | -29.58 | -33.91 | -58.80 |
| | | (38.99) | (39.97) | (43.63) |
| $\Delta^d$ Unemployment | -0.01 | | -0.01 | |
| | (0.29) | | (0.29) | |
| $\Delta^d$ Inflation | -0.11 | | -0.14 | |
| | (0.17) | | (0.18) | |
| **Error Correction and Constant** | | | | |
| FECM - Policy Liberalism$_{t-1}$ | -0.14 | | | |
| | (0.13) | | | |
| FECM - Policy Liberalism & Inequality$_{t-1}$ | | -0.15 | -0.16 | -0.01 |
| | | (0.13) | (0.13) | 0.06 |
| Constant | 0.02 | 0.05 | 0.07 | 0.25 |
| | (0.28) | (0.28) | (0.29) | (0.35) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | -0.03 | 0.01 | -0.02 | -0.03 |
| Durbin-Watson | 1.78 | 1.72 | 1.76 | 1.77 |
| Breusch-Pagan Test ($\chi^2$) | 1.01 | 0.93 | 1.74 | 0.89 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.83 | 2.16 | 1.37 | 0.40 |

† *Note*: Entries are OLS coefficients. Standard errors in parentheses. Models 1-3: DV is $\Delta^d$ Liberal Mood. Model 4: DV is $\Delta^d$ Welfare Support. FECM captures equilibrium relationship between *Policy Liberalism* and/or *Income Inequality* with DV. FECM significance (*p$\leq$ 0.05, one-tail test). Coefficient significance (*p$\leq$ 0.05, one-tail test).

## Table C.33: Three-Step FECM Results - Kelly and Enns (2010)†

| **Table 2 Models** | **Low Income Mood** | **High Income Mood** | **Low Income Mood** | **High Income Mood** |
|---|---|---|---|---|
| **Short Run Effects** | | | | |
| $\Delta^d$ Policy Liberalism | -0.10 | 0.10 | -0.09 | 0.11 |
| | (0.17) | (0.15) | (0.17) | 0.15 |
| $\Delta^d$ Income Inequality | 31.82 | -24.84 | 17.70 | -31.04 |
| | (67.45) | (59.74) | (67.62) | (59.98) |
| $\Delta^d$ Unemployment | | | -0.02 | -0.24 |
| | | | (0.51) | (0.44) |
| $\Delta^d$ Inflation | | | -0.44 | -0.23 |
| | | | (0.29) | (0.25) |
| **Error Correction and Constant** | | | | |
| FECM - Policy Liberalism & Inequality$_{t-1}$ | -0.14 | -0.14 | -0.17 | -0.14 |
| | (0.14) | (0.13) | (0.15) | 0.14 |
| Constant | -0.32 | -0.03 | -0.28 | 0.00 |
| | (0.49) | (0.42) | (0.49) | (0.42) |
| **Fit and Diagnostics** | | | | |
| Adjusted $R^2$ | -0.03 | -0.02 | -0.02 | -0.04 |
| Durbin-Watson | 1.79 | 1.72 | 1.86 | 1.77 |
| Breusch-Pagan Test ($\chi^2$) | 6.27 | 3.57 | 8.38 | 4.36 |
| Breusch-Godfrey LM Test ($\chi^2$) | 2.29 | 3.36 | 1.38 | 3.47 |

† *Note*: Entries are OLS coefficients. Standard errors in parentheses. Models 1& 3: DV is $\Delta^d$ Low Income Mood. Model 2 & 4: DV is $\Delta^d$ High Income Mood. FECM captures equilibrium relationship between *Policy Liberalism* and *Income Inequality* with DV. FECM significance (*p$\leq$ 0.05, one-tail test). Coefficient significance (*p$\leq$ 0.05, one-tail test).

### C.5.3 Example 3: Volscho & Kelly (*American Sociological Review*, 2012) Top 1% Income Share

What accounts for the rise of the super-rich? This is the question that Volscho and Kelly (2012) set out to answer in their recent paper, and they do so using the GECM model. We take note of this paper due to the sheer number of independent variables used in their models. Model 4 has 17 estimated IVs with a T of 60. In all of these models, VK rely on the standard $t$-statistic distribution and use a two-tail test of significance for their ECMs. Also, VK note in their Footnote 15 that they estimated the same models using a fractional ECM and received results that were largely similar. The FECM models were included in their supplement.

The ECM follows its own distribution, so the standard two-tail test is inappropriate for use with the ECM models. Additionally, the ECM distribution is biased by the inclusion of covariates that are not $I(1)$. As VK note, their data is a mix of data types (stationary and integrated), so any hypothesis tests will be based on unbalanced equations.

A main point of interest for this replication was the fractional ECM model that VK estimated. If, as they claim, a fractional model gives results very similar to the GECM then our arguments are significantly weakened. But VK's use of fractional methods is incorrect – they do not remove auto-correlation from each variable and from the ECM. Instead, what they do is fractionally difference their IVs, and then include them in the general ECM model with the level-form lags of all IVs and the lag of the dependent variable. With an $I(1)$ dependent variable, including non-stationary variables on the right hand side biases the model estimates and defeats the purpose of fractional differencing. Just as spurious regressions are prevalent in a static regression of non-stationary variables, the inclusion of some differenced - or fractionally differenced - independent variables does not provide a solution. Volscho and Kelly misapply fractional methods and should not be claiming to find corroboration in them.

We re-estimate their models with a proper fractional ECM model and find drastically different results from what they report. We initially estimate the order of fractional integration of all variables. As VK report, their variables are a mix of data types.

Note that we only estimate the residuals of two variables from Models 4 and 5. This is because these are the only two variables to achieve significance for which a fractional ECM could be estimated. With two significant variables, we included them in the same cointegrating regression. In this case the ECM captures the long-term equilibrium relationship of both *Trade Openness* and *Real S&P Composite Index* with the *Top 1% Income Share*. We find a significant ECM in Models 4 and 5.

With the exception of *Dem President* in Model 1, all other variables are insignificant and failed to approach significance. The results of the fractionally differenced models are provided below. Contrary to Volscho and Kelly's supplement, when the data is run through a proper fractional ECM model, the significance of the results are far different. Re-estimating their power resources model (Model 1), policy model (Model 2) and their politics and policy model (Model 3), we find no support for either politics or policy having significant effects on the concentration of income among the Top 1%. Furthermore, a look at the $R^2$ of these models indicate a rather poor fit. It is not until Models 4 and 5, which account for economic indicators, that we see an improvement. Within these fully specified models, we again fail to

## Table C.34: FECM Order of Integration, Volscho and Kelly (2012)

| | Individual Series *d* | Residuals Model 1 *d* | Residuals Model 2 *d* | Residuals Model 3 *d* | Residuals Model 4 | Residuals Model 5 |
|---|---|---|---|---|---|---|
| Top 1% Share | 0.93 | | | | | |
| % Congressional Democrat | 0.87 | | | | | |
| % Union Membership | 1.33 | | | | | |
| Top Marginal Tax Rate | 1.14 | | | | | |
| Cap. Gains Tax Rate | 1.26 | | | | | |
| 3 Month Treasury Bill Rate | 1.00 | | | | | |
| Trade Openness | 1.02 | | | | 0.92 | 0.92 |
| Unemployment Rate | 0.95 | | | | | |
| Log Real GDP | 0.96 | | | | | |
| Real S&P Composite Index | 1.37 | | | | 0.97 | 0.97 |
| Shiller Home Price Index | 1.58 | | | | | |

† *Note*: Entries are estimates of the order of integration of all significant long-run variables from Volscho and Kelly (2012) as well as estimates of the residuals of bivariate regressions of the DV and each significant variable. *Dem President* and *Divided Government* are dichotomous variables and are used in first differences. Estimates conducted in RATS using Robinson's (1995) semi-parametric estimator. Because $d$ was estimated using the Local Whittle, the asymptotic standard errors are provided - using the following formula $(1/\sqrt{4 * m})$ with $m = T^{4/5}$. (s.e. = 0.10).

find any significant effects of either politics or policy. Our re-estimation of their data finds that the only significant influences on the income share of the top 1% are market factors, specifically levels of trade openness and the valuation of the stock market.

**Table C.35: Three-Step FECM Results - Volscho and Kelly (2012)†**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| **Short Run Effects** | | | | | |
| $\Delta^d$ Democrat President | **1.30*** | | | | |
| | (0.55) | | | | |
| $\Delta^d$ % Congressional Democrat | -0.02 | | -0.02 | -0.00 | -0.01 |
| | (0.04) | | (0.04) | (0.04) | (0.04) |
| $\Delta^d$ Divided Government | 0.70 | | 0.03 | 0.13 | 0.15 |
| | (0.48) | | (0.42) | (0.36) | (0.36) |
| $\Delta^d$ % Union Membership | -0.05 | | 0.01 | -0.12 | -0.06 |
| | (0.30) | | (0.32) | (0.29) | (0.26) |
| $\Delta^d$ Top Marginal Tax Rate | | -0.02 | -0.02 | -0.02 | -0.02 |
| | | (0.04) | (0.04) | (0.03) | (0.03) |
| $\Delta^d$ Cap. Gains Tax Rate | | -0.07 | -0.06 | -0.08 | -0.07 |
| | | (0.06) | (0.06) | (0.05) | (0.05) |
| $\Delta^d$ 3 Month Treasury Bill Rate | | 0.12 | 0.11 | -0.15 | -0.13 |
| | | (0.11) | (0.12) | (0.15) | (0.14) |
| $\Delta^d$ Trade Openness | | | | **0.36*** | **0.38*** |
| | | | | (0.20) | (0.19) |
| $\Delta^d$ Unemployment Rate | | | | -0.15 | |
| | | | | (0.33) | |
| $\Delta^d$ Log Real GDP | | | | -1.67 | 3.45 |
| | | | | (13.39) | (6.89) |
| $\Delta^d$ Real S&P Composite Index | | | | **0.07*** | **0.07*** |
| | | | | (0.01) | (0.01) |
| $\Delta^d$ Shiller Home Price Index | | | | 0.32 | 0.32 |
| | | | | (0.27) | (0.27) |
| **Error Correction and Constant** | | | | | |
| $FECM_{t-1}$ | | | | **-0.38*** | **-0.35*** |
| | | | | (0.15) | (0.14) |
| Constant | 0.18 | 0.15 | 0.15 | 0.10 | 0.01 |
| | (0.15) | (0.15) | (0.16) | (0.26) | (0.18) |
| **Fit and Diagnostics** | | | | | |
| Adjusted $R^2$ | 0.03 | 0.00 | -0.06 | 0.35 | 0.36 |
| Durbin-Watson | 2.09 | 2.11 | 2.12 | 1.92 | 1.89 |
| Breusch-Pagan Test ($\chi^2$) | 2.85 | 8.33* | 11.18 | 14.53 | 13.81 |
| Breusch-Godfrey LM Test ($\chi^2$) | 0.37 | 0.27 | 0.40 | 1.71 | 0.97 |

† *Note*: Entries are OLS coefficients.Standard errors in parentheses. DV is $\Delta^d$ Top 1% Share. FECM captures equilibrium relationship between *Trade Openness* and *S&P Index* with DV. FECM significance (*p≤ 0.05, one-tail test). Coefficient significance (*p≤ 0.05, one-tail test).

## C.6 Additional Materials - Common Misinterpretation of De Boef and Keele (2008)

As we note in the main text, while De Boef and Keele are discussing the use of the GECM under the specific condition of stationarity, common practice is to assume that the model can be used under any circumstances. Most papers do not check the stationarity properties of their data and the reason for this oversight stems directly from a widespread belief that any and all types of data are appropriate for use in the model. Below we highlight a sample of this misunderstanding in practice.

- Volscho and Kelly (ASR 2012): "In summary, the ECM is a very general model that is easy to implement and estimate, does not impose assumptions about cointegration, and can be applied to both stationary and nonstationary data (Banerjee et a. 1993, DeBoef and Keele 2008)."
- Buthe and Milner (WP 2014): "These powerful dynamic models, which are equivalent to autoregressive distributed lag (ADL) models after a straightforward mathematical transformation (fn to D&K), also provide a safeguard against spurious correlation that might arise in time series analysis when variables are trending together."
- Casillas, Enns, Wohlfarth (AJPS 2011): "The ECM provides a conservative empirical test of our argument and a general model that is appropriate with both stationary and nonstationary data (De Boef and Keele 2008)."
- Enns (AJPS 2014): "Finally, because the ECM estimates the change in the dependent variable, we overcome the concern of estimating a spurious relationship among nonstationary time series (De Boef and Granato 1997, De Boef and Keele 2008)."
- Faricy (JOP 2011): "ECMs are appropriate when using both stationary and nonstationary data and offer a conservative test of the theory (DeBoef and Keele 2008)."
- Jennings and John (AJPS 2009): "While error-correction models tend to be isomorphic with the concept of cointegration for political scientists (see DeBoef and Keele 2008), this framework is appropriate for modeling feedback and equilibrium relationships in stationary as well as nonstationary data (Banerjee, Dolado, Galbraith 1993, Davidson and MacKinnon, DeBoef and Keele 2008)."
- Kayser (AJPS 2009): "I err on the side of caution and assume nonstationarity bearing in mind that many ECMs offer the same benefits of capturing long- and short-term dynamics in stationary data as in nonstationary but cointegrated data (DeBoef and Keele 2008)."
- Kelly and Enns (AJPS 2010): "While the use of an ECM is often motivated by the presence of a nonstationary time-series as a dependent variable, our application of this model is based on the fact that it is among the most general time-series models that imposes the fewest restrictions. (DeBoef and Keele 2008)."

- Kono (JOP 2008): "I employ an error-correction model both to guard against integration problems....Because the error-correction model thus imposes fewer restrictive assumptions than other time series models (DeBoef and Keele 2008), it is the most general and conservative one I can estimate."
- Rickard (JOP 2012): "Differencing the series also minimizes the potential for a spurious correlation between two series exhibiting a time trend (DeBoef and Keele 2008)."
- Sanchez-Urribari (JOP 2011): "Following the suggestion of DeBoef and Keele, we estimate an error correction model (ECM) to specifically control for the dynamic effect of the covariates."
- Ura (AJPS 2014): "Though ECMs were originally developed for investigating cointegrated time series, DeBoef and Keele (2008) note that they may also be applied in a variety of time-series contexts in the absence of cointegration with either stationary or nonstationary data."
- Ura and Ellis (JOP 2012): "The error-correction model is often used with integrated series, but is also appropriate for use with nonintegrated series as well (DeBoef and Keele 2008)."
- Ura and Wohlfarth (JOP 2010): "Though the model specification was originally developed for investigating cointegrated time series, DeBoef and Keele (2008) note that it may also be applied in a variety of time series contexts in the absence of cointegration with either stationary or nonstationary data."

## C.7 Additional Materials: Near-Integrated Data

### C.7.1 Tables - Varying AR - Monte Carlo Results

The results in the following tables are meant to provide as much information as possible in as little space as possible. Therefore, we have tried to include a wide a range of models. The top row of each table contains models run in which the level of autoregression in the DV varies, but the IV(s) all have unit roots. The middle row contain results of models in which all series within each model share the same level of autoregression. These equations are balanced. The bottom row also allows the DV to vary by levels of autoregression, but the IV(s) are all generated as $I(0)$ - the IV(s) contain very little, if any, information.

Table C.36: Summary Statistics of Bivariate ECM Model by DV and IV (T=60)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 8662 | 8408 | 8143 | 7872 | 7598 | 7031 | 7048 | 6780 | 6464 | 6132 | 5675 |
| ECM Significant* | 1281 | 1147 | 1062 | 955 | 862 | 789 | 708 | 669 | 605 | 564 | 533 |
| Mean of $\alpha_1$ | -0.21 | -0.20 | -0.19 | -0.18 | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 |
| $\geq 1\Delta X_t$ Significant | 592 | 585 | 587 | 587 | 594 | 605 | 600 | 605 | 603 | 607 | 610 |
| $\geq 1X_{t\text{-}1}$ Significant | 1369 | 1409 | 1478 | 1548 | 1597 | 1668 | 1714 | 1778 | 1765 | 1756 | 1716 |
| ECM and $\geq 1\Delta X_t$ Significant | 514 | 493 | 478 | 474 | 473 | 459 | 471 | 455 | 429 | 416 | 384 |
| ECM and $\geq 1\Delta X_t$ Significant* | 105 | 92 | 86 | 80 | 71 | 69 | 65 | 65 | 55 | 47 | 49 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1320 | 1343 | 1395 | 1443 | 1457 | 1498 | 1513 | 1554 | 1515 | 1467 | 1410 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 519 | 491 | 477 | 452 | 415 | 395 | 373 | 380 | 355 | 330 | 310 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 8295 | 8042 | 7764 | 7455 | 7176 | 6954 | 6731 | 6479 | 6257 | 6005 | |
| ECM Significant* | 1050 | 952 | 874 | 799 | 748 | 694 | 622 | 581 | 569 | 551 | |
| Mean of $\alpha_1$ | -0.19 | -0.19 | -0.18 | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.13 | |
| $\geq 1\Delta X_t$ Significant | 585 | 597 | 607 | 608 | 606 | 619 | 620 | 619 | 622 | 611 | |
| $\geq 1X_{t\text{-}1}$ Significant | 1136 | 1185 | 1242 | 1292 | 1331 | 1381 | 1467 | 1508 | 1563 | 1635 | |
| ECM and $\geq 1\Delta X_t$ Significant | 501 | 504 | 500 | 483 | 472 | 473 | 466 | 449 | 434 | 416 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 93 | 89 | 83 | 72 | 68 | 68 | 65 | 59 | 57 | 55 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1082 | 1118 | 1159 | 1190 | 1210 | 1227 | 1277 | 1297 | 1336 | 1367 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 363 | 353 | 346 | 334 | 330 | 313 | 303 | 303 | 302 | 307 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 7906 | 7552 | 7181 | 6799 | 6422 | 6075 | 5745 | 5422 | 5160 | 4909 | 4482 |
| ECM Significant* | 597 | 529 | 469 | 416 | 375 | 312 | 276 | 251 | 242 | 240 | 219 |
| Mean of $\alpha_1$ | -0.17 | -0.17 | -0.16 | -0.15 | -0.14 | -0.13 | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 561 | 560 | 562 | 563 | 557 | 561 | 566 | 560 | 548 | 551 | 542 |
| $\geq 1X_{t\text{-}1}$ Significant | 595 | 601 | 601 | 599 | 601 | 601 | 597 | 606 | 606 | 602 | 593 |
| ECM and $\geq 1\Delta X_t$ Significant | 449 | 436 | 427 | 414 | 388 | 372 | 361 | 338 | 317 | 308 | 257 |
| ECM and $\geq 1\Delta X_t$ Significant* | 66 | 59 | 53 | 44 | 38 | 34 | 35 | 30 | 27 | 25 | 21 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 497 | 485 | 475 | 449 | 428 | 405 | 390 | 384 | 364 | 345 | 308 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 89 | 81 | 74 | 64 | 57 | 52 | 42 | 39 | 38 | 36 | 33 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.27.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IV and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IV is integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.37: Summary Statistics of Multivariate (2 IVs) ECM Model by DV and IV (T=60)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^\star$ | | | | | | | | | | | |
| ECM Significant | 8925 | 8744 | 8562 | 8353 | 8140 | 7905 | 7673 | 7432 | 7177 | 6926 | 6488 |
| ECM Significant* | 1133 | 1053 | 970 | 896 | 842 | 781 | 718 | 677 | 632 | 593 | 552 |
| Mean of $\alpha_1$ | -0.24 | -0.23 | -0.22 | -0.22 | -0.20 | -0.19 | -0.19 | -0.18 | -0.17 | -0.16 | -0.15 |
| $\geq 1\Delta X_t$ Significant | 1213 | 1214 | 1227 | 1228 | 1223 | 1228 | 1225 | 1217 | 1208 | 1212 | 1227 |
| $\geq 1X_{t\text{-}1}$ Significant | 2541 | 2592 | 2664 | 2716 | 2580 | 2849 | 2933 | 2947 | 3019 | 3002 | 2945 |
| ECM and $\geq 1\Delta X_t$ Significant | 1109 | 1099 | 1092 | 1058 | 1032 | 1014 | 985 | 963 | 914 | 895 | 862 |
| ECM and $\geq 1\Delta X_t$ Significant* | 216 | 204 | 185 | 174 | 161 | 154 | 148 | 137 | 125 | 116 | 110 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 2463 | 2491 | 2538 | 2563 | 2580 | 2619 | 2669 | 2643 | 2643 | 2594 | 2497 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 731 | 699 | 658 | 614 | 584 | 552 | 513 | 491 | 477 | 451 | 430 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 8459 | 8278 | 8056 | 7856 | 7645 | 7456 | 7250 | 7101 | 6946 | 6769 | |
| ECM Significant* | 871 | 816 | 761 | 718 | 667 | 632 | 652 | 579 | 553 | 547 | |
| Mean of $\alpha_1$ | -0.22 | -0.21 | -0.20 | -0.20 | -0.19 | -0.18 | -0.17 | -0.17 | -0.16 | -0.15 | |
| $\geq 1\Delta X_t$ Significant | 1193 | 1197 | 1210 | 1206 | 1205 | 1214 | 1228 | 1243 | 1235 | 1241 | |
| $\geq 1X_{t\text{-}1}$ Significant | 2159 | 2227 | 2304 | 2380 | 2470 | 2574 | 2657 | 2718 | 2801 | 2864 | |
| ECM and $\geq 1\Delta X_t$ Significant | 1049 | 1042 | 1034 | 1004 | 984 | 969 | 950 | 950 | 931 | 912 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 154 | 144 | 136 | 131 | 124 | 123 | 114 | 111 | 106 | 109 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 2037 | 2081 | 2137 | 2184 | 2229 | 2284 | 2328 | 2350 | 2379 | 2438 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 526 | 505 | 491 | 474 | 448 | 436 | 428 | 415 | 403 | 412 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 7661 | 7321 | 6964 | 6633 | 6257 | 5895 | 5581 | 5296 | 4974 | 4760 | 4340 |
| ECM Significant* | 260 | 229 | 213 | 177 | 157 | 139 | 127 | 115 | 98 | 88 | 94 |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 1092 | 1092 | 1096 | 1096 | 1101 | 1105 | 1103 | 1030 | 1101 | 1110 | 1099 |
| $\geq 1X_{t\text{-}1}$ Significant | 1216 | 1216 | 1207 | 1212 | 1214 | 1212 | 1226 | 1222 | 1224 | 1215 | 1195 |
| ECM and $\geq 1\Delta X_t$ Significant | 878 | 845 | 800 | 774 | 739 | 715 | 665 | 645 | 617 | 602 | 546 |
| ECM and $\geq 1\Delta X_t$ Significant* | 45 | 41 | 36 | 28 | 23 | 22 | 21 | 17 | 15 | 13 | 20 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 990 | 960 | 915 | 885 | 855 | 822 | 794 | 763 | 733 | 703 | 634 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 78 | 69 | 63 | 48 | 42 | 39 | 33 | 27 | 23 | 20 | 22 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p\leq 0.05$, one-tail test). Coefficient significance ($p\leq 0.05$, two-tail test) *ECM significance (*$p\leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.565.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.38: Summary Statistics of Multivariate (3 IVs) ECM Model by DV and IV (T=60)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9127 | 8986 | 8834 | 8675 | 8503 | 8320 | 8150 | 7954 | 7731 | 7489 | 7197 |
| ECM Significant* | 1071 | 990 | 923 | 852 | 809 | 761 | 712 | 647 | 636 | 593 | 527 |
| Mean of $\alpha_1$ | -0.27 | -0.26 | -0.25 | -0.24 | -0.23 | -0.23 | -0.22 | -0.21 | -0.20 | -0.19 | -0.18 |
| $\geq 1\Delta X_{\text{t}}$ Significant | 1707 | 1712 | 1706 | 1710 | 1710 | 1712 | 1715 | 1735 | 1738 | 1762 | 1733 |
| $\geq 1X_{\text{t-1}}$ Significant | 3496 | 3554 | 3641 | 3705 | 3770 | 3814 | 3858 | 3892 | 3909 | 3921 | 3897 |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant | 1585 | 1566 | 1544 | 1525 | 1501 | 1488 | 1461 | 1455 | 1420 | 1404 | 1329 |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant* | 266 | 253 | 246 | 223 | 216 | 205 | 191 | 158 | 178 | 170 | 152 |
| ECM and $\geq 1X_{\text{t-1}}$ Significant | 3387 | 3422 | 3476 | 3506 | 3539 | 3541 | 3538 | 3544 | 3521 | 3488 | 3415 |
| ECM and $\geq 1X_{\text{t-1}}$ Significant* | 802 | 761 | 726 | 679 | 650 | 615 | 586 | 542 | 525 | 507 | 452 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 8600 | 8453 | 8286 | 8108 | 7954 | 7815 | 7685 | 7544 | 7419 | 7321 | |
| ECM Significant* | 785 | 730 | 685 | 642 | 620 | 606 | 595 | 578 | 557 | 541 | |
| Mean of $\alpha_1$ | -0.24 | -0.23 | -0.23 | -0.22 | -0.21 | -0.21 | -0.20 | -0.20 | -0.19 | -0.18 | |
| $\geq 1\Delta X_{\text{t}}$ Significant | 1737 | 1746 | 1776 | 1789 | 1803 | 1820 | 1824 | 1827 | 1809 | 1776 | |
| $\geq 1X_{\text{t-1}}$ Significant | 3159 | 3230 | 3356 | 3429 | 3532 | 3636 | 3685 | 3726 | 3783 | 3846 | |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant | 1532 | 1517 | 1521 | 1510 | 1503 | 1504 | 1487 | 1465 | 1433 | 1388 | |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant* | 214 | 199 | 196 | 188 | 191 | 186 | 189 | 183 | 171 | 162 | |
| ECM and $\geq 1X_{\text{t-1}}$ Significant | 3003 | 3042 | 3117 | 3179 | 3243 | 3296 | 3308 | 3325 | 3349 | 3394 | |
| ECM and $\geq 1X_{\text{t-1}}$ Significant* | 589 | 556 | 537 | 511 | 504 | 498 | 497 | 479 | 460 | 463 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 7330 | 6990 | 6640 | 6294 | 5933 | 5615 | 5313 | 5053 | 4801 | 4544 | 4206 |
| ECM Significant* | 137 | 115 | 102 | 87 | 80 | 66 | 67 | 57 | 523 | 49 | 53 |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_{\text{t}}$ Significant | 1513 | 1512 | 1521 | 1515 | 1508 | 1500 | 1501 | 1506 | 1504 | 1495 | 1530 |
| $\geq 1X_{\text{t-1}}$ Significant | 1715 | 1717 | 1713 | 1714 | 1710 | 1722 | 1705 | 1695 | 1691 | 1700 | 1683 |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant | 1133 | 1089 | 1040 | 998 | 944 | 896 | 846 | 818 | 772 | 730 | 701 |
| ECM and $\geq 1\Delta X_{\text{t}}$ Significant* | 35 | 30 | 27 | 23 | 25 | 22 | 21 | 15 | 16 | 19 | 21 |
| ECM and $\geq 1X_{\text{t-1}}$ Significant | 1341 | 1304 | 1245 | 1197 | 1139 | 1099 | 1031 | 990 | 955 | 900 | 846 |
| ECM and $\geq 1X_{\text{t-1}}$ Significant* | 71 | 58 | 53 | 47 | 40 | 32 | 29 | 26 | 22 | 19 | 25 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.816.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.39: Summary Statistics of Multivariate (4 IVs) ECM Model by DV and IV (T=60)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9270 | 9187 | 9069 | 8964 | 8827 | 8693 | 8553 | 8354 | 8165 | 7954 | 7666 |
| ECM Significant* | 1054 | 989 | 911 | 840 | 785 | 734 | 693 | 665 | 605 | 555 | 496 |
| Mean of $\alpha_1$ | -0.30 | -0.29 | -0.28 | -0.27 | -0.27 | -0.26 | -0.25 | -0.24 | -0.23 | -0.22 | -0.21 |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2301 | 2304 | 2307 | 2300 | 2298 | 2308 | 2302 | 2295 | 2300 | 2300 | 2292 |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 4261 | 4329 | 4382 | 4447 | 4518 | 4559 | 4604 | 4599 | 4617 | 4624 | 4599 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2174 | 2159 | 2138 | 2113 | 2094 | 2065 | 2033 | 1986 | 1954 | 1914 | 1838 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 349 | 334 | 314 | 295 | 281 | 269 | 255 | 244 | 227 | 227 | 200 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 4135 | 4184 | 4208 | 4243 | 4275 | 4282 | 4302 | 4252 | 4226 | 4189 | 4108 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 869 | 826 | 769 | 713 | 677 | 640 | 605 | 585 | 535 | 505 | 447 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 8733 | 8621 | 8476 | 8370 | 8242 | 8125 | 8053 | 7950 | 7846 | 7760 | |
| ECM Significant* | 699 | 667 | 640 | 628 | 595 | 579 | 563 | 538 | 526 | 506 | |
| Mean of $\alpha_1$ | -0.27 | -0.26 | -0.25 | -0.25 | -0.24 | -0.24 | -0.23 | -0.23 | -0.22 | -0.22 | |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2301 | 2309 | 2318 | 2325 | 2326 | 2335 | 2347 | 2350 | 2342 | 2318 | |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3898 | 3974 | 4060 | 4171 | 4268 | 4337 | 4421 | 4510 | 4545 | 4589 | |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2041 | 2027 | 2010 | 1999 | 1982 | 1973 | 1983 | 1962 | 1930 | 1886 | |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 254 | 243 | 241 | 242 | 240 | 241 | 238 | 229 | 223 | 214 | |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3692 | 3739 | 3791 | 3859 | 3925 | 3969 | 4020 | 4067 | 4090 | 4121 | |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 580 | 562 | 542 | 544 | 522 | 512 | 501 | 481 | 476 | 465 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 7168 | 6883 | 6499 | 6140 | 5799 | 5489 | 5149 | 4900 | 4658 | 4339 | 3998 |
| ECM Significant* | 72 | 65 | 55 | 46 | 40 | 31 | 23 | 19 | 21 | 19 | 16 |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2006 | 1997 | 1980 | 1978 | 1976 | 1992 | 1978 | 1989 | 1999 | 2014 | 2012 |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 2204 | 2209 | 2219 | 2219 | 2223 | 2222 | 2224 | 2206 | 2203 | 2216 | 2201 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 1466 | 1400 | 1339 | 1276 | 1204 | 1138 | 1075 | 1033 | 1016 | 959 | 881 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 25 | 24 | 20 | 16 | 15 | 12 | 8 | 9 | 10 | 8 | 9 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 1684 | 1630 | 1560 | 1487 | 1415 | 1343 | 1294 | 1221 | 1174 | 1124 | 1055 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 45 | 40 | 30 | 24 | 21 | 18 | 15 | 13 | 13 | 13 | 8 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.035.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.40: Summary Statistics of Multivariate (5 IVs) ECM Model by DV and IV (T=60)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9318 | 9237 | 9137 | 9047 | 8921 | 8785 | 8668 | 8544 | 8422 | 8271 | 8024 |
| ECM Significant* | 992 | 937 | 883 | 828 | 788 | 728 | 686 | 641 | 597 | 569 | 514 |
| Mean of $\alpha_1$ | -0.33 | -0.32 | -0.31 | -0.31 | -0.30 | -0.29 | -0.28 | -0.27 | -0.26 | -0.25 | -0.24 |
| $\geq 1\Delta X_t$ Significant | 2844 | 2848 | 2839 | 2850 | 2851 | 2855 | 2869 | 2892 | 2886 | 2900 | 2875 |
| $\geq 1X_{t\text{-}1}$ Significant | 4967 | 5044 | 5083 | 5113 | 5146 | 5148 | 5198 | 5210 | 5218 | 5217 | 5228 |
| ECM and $\geq 1\Delta X_t$ Significant | 2688 | 2675 | 2642 | 2631 | 2607 | 2579 | 2572 | 2565 | 2533 | 2511 | 2428 |
| ECM and $\geq 1\Delta X_t$ Significant* | 408 | 383 | 362 | 344 | 337 | 318 | 306 | 288 | 272 | 260 | 250 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 4806 | 4857 | 4868 | 4878 | 4883 | 4833 | 4851 | 4821 | 4795 | 4761 | 4710 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 860 | 824 | 778 | 728 | 696 | 642 | 611 | 576 | 539 | 519 | 481 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 8801 | 8718 | 8636 | 8559 | 8483 | 8373 | 8284 | 8214 | 8153 | 8076 | |
| ECM Significant* | 646 | 626 | 610 | 604 | 597 | 586 | 568 | 555 | 557 | 540 | |
| Mean of $\alpha_1$ | -0.29 | -0.29 | -0.28 | -0.28 | -0.27 | -0.27 | -0.26 | -0.26 | -0.25 | -0.25 | |
| $\geq 1\Delta X_t$ Significant | 2819 | 2832 | 2853 | 2862 | 2868 | 2885 | 2892 | 2897 | 2893 | 2892 | |
| $\geq 1X_{t\text{-}1}$ Significant | 4583 | 4683 | 4770 | 4833 | 4918 | 4994 | 5038 | 5087 | 5127 | 5192 | |
| ECM and $\geq 1\Delta X_t$ Significant | 2553 | 2552 | 2544 | 2535 | 2514 | 2507 | 2501 | 2494 | 2468 | 2451 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 276 | 268 | 259 | 259 | 259 | 255 | 252 | 254 | 263 | 259 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 4328 | 4408 | 4466 | 4509 | 4559 | 4604 | 4617 | 4650 | 4665 | 4688 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 564 | 558 | 548 | 540 | 538 | 528 | 519 | 510 | 514 | 501 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 6895 | 6590 | 6290 | 5982 | 5629 | 5318 | 5050 | 4769 | 4537 | 4298 | 3959 |
| ECM Significant* | 36 | 32 | 30 | 26 | 25 | 21 | 16 | 18 | 12 | 14 | 12 |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 2480 | 2477 | 2479 | 2477 | 2480 | 2479 | 2474 | 2464 | 2449 | 2452 | 2467 |
| $\geq 1X_{t\text{-}1}$ Significant | 2647 | 2657 | 2660 | 2663 | 2663 | 2659 | 2652 | 2653 | 2646 | 2645 | 2633 |
| ECM and $\geq 1\Delta X_t$ Significant | 1767 | 1702 | 1642 | 1559 | 1484 | 1424 | 1366 | 1288 | 1210 | 1149 | 1057 |
| ECM and $\geq 1\Delta X_t$ Significant* | 12 | 13 | 13 | 9 | 7 | 8 | 7 | 10 | 6 | 8 | 6 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1976 | 1915 | 1836 | 1766 | 1682 | 1607 | 1536 | 1476 | 1406 | 1351 | 1251 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 22 | 23 | 22 | 17 | 18 | 15 | 11 | 12 | 8 | 9 | 7 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance (p$\leq$ 0.05, one-tail test). Coefficient significance (p$\leq$ 0.05, two-tail test) *ECM significance (*p$\leq$ 0.05, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.229.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.41: Summary Statistics of Bivariate ECM Model by DV and IV (T=150)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^\star$ | | | | | | | | | | | |
| ECM Significant | 9971 | 9941 | 9877 | 9755 | 9559 | 9261 | 8791 | 8141 | 7427 | 6708 | 5769 |
| ECM Significant* | 4768 | 3973 | 3263 | 2677 | 2133 | 1686 | 1318 | 1014 | 790 | 662 | 501 |
| Mean of $\alpha_1$ | -0.14 | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.08 | -0.07 | -0.06 | -0.05 |
| $\geq 1\Delta X_t$ Significant | 550 | 550 | 550 | 555 | 554 | 553 | 554 | 558 | 562 | 559 | 562 |
| $\geq 1X_{t\text{-}1}$ Significant | 943 | 973 | 1044 | 1109 | 1164 | 1264 | 1391 | 1492 | 1635 | 1774 | 1639 |
| ECM and $\geq 1\Delta X_t$ Significant | 550 | 548 | 545 | 545 | 531 | 510 | 487 | 457 | 414 | 386 | 338 |
| ECM and $\geq 1\Delta X_t$ Significant* | 275 | 227 | 186 | 156 | 127 | 104 | 82 | 68 | 46 | 41 | 36 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 942 | 972 | 1043 | 1108 | 1156 | 1247 | 1348 | 1424 | 1492 | 1526 | 1343 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 809 | 761 | 737 | 685 | 618 | 580 | 513 | 442 | 381 | 342 | 270 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9944 | 9894 | 9816 | 9672 | 9402 | 9011 | 8454 | 7759 | 7056 | 6473 | |
| ECM Significant* | 4286 | 3478 | 2820 | 2231 | 1741 | 1366 | 1074 | 844 | 685 | 598 | |
| Mean of $\alpha_1$ | -0.13 | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.06 | |
| $\geq 1\Delta X_t$ Significant | 549 | 552 | 559 | 560 | 570 | 579 | 574 | 578 | 581 | 560 | |
| $\geq 1X_{t\text{-}1}$ Significant | 733 | 755 | 795 | 832 | 900 | 972 | 1077 | 1191 | 1335 | 1489 | |
| ECM and $\geq 1\Delta X_t$ Significant | 547 | 547 | 549 | 543 | 542 | 519 | 487 | 454 | 424 | 378 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 234 | 206 | 155 | 124 | 101 | 81 | 70 | 52 | 48 | 35 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 732 | 751 | 790 | 826 | 891 | 942 | 1022 | 1098 | 1178 | 1247 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 542 | 506 | 469 | 422 | 392 | 363 | 330 | 314 | 298 | 280 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9954 | 9910 | 9836 | 9639 | 9350 | 8837 | 8128 | 7207 | 6243 | 5452 | 4542 |
| ECM Significant* | 3604 | 2856 | 2206 | 1667 | 1234 | 906 | 631 | 461 | 332 | 274 | 216 |
| Mean of $\alpha_1$ | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.05 | -0.05 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 548 | 550 | 547 | 554 | 552 | 554 | 550 | 546 | 541 | 534 | 536 |
| $\geq 1X_{t\text{-}1}$ Significant | 559 | 560 | 563 | 563 | 563 | 563 | 564 | 564 | 566 | 571 | 561 |
| ECM and $\geq 1\Delta X_t$ Significant | 546 | 547 | 539 | 535 | 520 | 489 | 447 | 398 | 340 | 299 | 270 |
| ECM and $\geq 1\Delta X_t$ Significant* | 206 | 170 | 134 | 101 | 77 | 53 | 37 | 27 | 19 | 12 | 13 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 556 | 555 | 552 | 541 | 522 | 489 | 453 | 404 | 353 | 333 | 294 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 231 | 195 | 154 | 117 | 89 | 73 | 53 | 40 | 27 | 17 | 14 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.236.
$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.
$\star\star$ Model 2: Both IV and DV data generating process based on same level of $\rho$.
$\star\star\star$ Model 3: IV is integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.42: Summary Statistics of Multivariate (2 IVs) ECM Model by DV and IV (T=150)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9963 | 9935 | 9884 | 9788 | 9632 | 9382 | 9027 | 8588 | 8096 | 7511 | 6695 |
| ECM Significant* | 3861 | 3277 | 2724 | 2216 | 1789 | 1429 | 1111 | 900 | 741 | 618 | 467 |
| Mean of $\alpha_1$ | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 |
| $\geq 1\Delta X_t$ Significant | 1036 | 1035 | 1038 | 1036 | 1034 | 1034 | 1032 | 1033 | 1031 | 1060 | 1047 |
| $\geq 1X_{t\text{-}1}$ Significant | 1936 | 2010 | 2085 | 2189 | 2293 | 2442 | 2584 | 2750 | 2929 | 3070 | 2958 |
| ECM and $\geq 1\Delta X_t$ Significant | 1032 | 1027 | 1028 | 1012 | 993 | 968 | 929 | 897 | 857 | 823 | 728 |
| ECM and $\geq 1\Delta X_t$ Significant* | 446 | 386 | 326 | 254 | 206 | 155 | 128 | 108 | 100 | 91 | 59 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1936 | 2009 | 2080 | 2176 | 2269 | 2395 | 2505 | 2622 | 2716 | 2727 | 2525 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1347 | 1257 | 1150 | 1049 | 920 | 807 | 685 | 604 | 526 | 470 | 365 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9936 | 9870 | 9780 | 9618 | 9375 | 9053 | 8625 | 8113 | 7603 | 7154 | |
| ECM Significant* | 3047 | 2497 | 2040 | 1642 | 1342 | 1064 | 850 | 693 | 585 | 549 | |
| Mean of $\alpha_1$ | -0.14 | -0.13 | -0.12 | -0.11 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.07 | |
| $\geq 1\Delta X_t$ Significant | 1015 | 1010 | 1016 | 1012 | 1024 | 1028 | 1030 | 1029 | 1037 | 1046 | |
| $\geq 1X_{t\text{-}1}$ Significant | 1536 | 1589 | 1669 | 1754 | 1865 | 1972 | 2154 | 2330 | 2547 | 2808 | |
| ECM and $\geq 1\Delta X_t$ Significant | 1008 | 996 | 994 | 962 | 956 | 930 | 894 | 854 | 815 | 787 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 351 | 280 | 229 | 183 | 147 | 117 | 93 | 78 | 68 | 65 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1533 | 1580 | 1653 | 1731 | 1819 | 1904 | 2042 | 2153 | 2297 | 2456 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 886 | 817 | 757 | 676 | 616 | 531 | 473 | 419 | 386 | 362 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9932 | 9864 | 9747 | 9525 | 9197 | 8674 | 7964 | 7067 | 6114 | 5333 | 4475 |
| ECM Significant* | 1944 | 1470 | 1058 | 760 | 542 | 385 | 271 | 186 | 134 | 112 | 85 |
| Mean of $\alpha_1$ | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.05 | -0.05 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 994 | 992 | 984 | 988 | 985 | 977 | 980 | 985 | 973 | 966 | 969 |
| $\geq 1X_{t\text{-}1}$ Significant | 1034 | 1034 | 1035 | 1039 | 1037 | 1044 | 1051 | 1052 | 1048 | 1049 | 1018 |
| ECM and $\geq 1\Delta X_t$ Significant | 988 | 980 | 962 | 944 | 906 | 846 | 784 | 714 | 619 | 541 | 447 |
| ECM and $\geq 1\Delta X_t$ Significant* | 214 | 162 | 120 | 89 | 64 | 44 | 28 | 16 | 11 | 9 | 7 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1029 | 1022 | 1008 | 989 | 958 | 909 | 862 | 782 | 672 | 607 | 492 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 249 | 193 | 157 | 114 | 88 | 58 | 40 | 29 | 19 | 25 | 14 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.528.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.43: Summary Statistics of Multivariate (3 IVs) ECM Model by DV and IV (T=150)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^\star$ | | | | | | | | | | | |
| ECM Significant | 9970 | 9947 | 9913 | 9846 | 9712 | 9553 | 9300 | 8977 | 8556 | 8111 | 7348 |
| ECM Significant* | 3405 | 2878 | 2411 | 2012 | 1657 | 1381 | 1115 | 916 | 754 | 612 | 479 |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 |
| $\geq 1\Delta X_t$ Significant | 1514 | 1527 | 1526 | 1525 | 1535 | 1534 | 1549 | 1558 | 1565 | 1563 | 1559 |
| $\geq 1X_{t-1}$ Significant | 2898 | 2878 | 3088 | 3218 | 3339 | 3494 | 3646 | 3751 | 3871 | 3966 | 3903 |
| ECM and $\geq 1\Delta X_t$ Significant | 1508 | 1519 | 1508 | 1497 | 1495 | 1470 | 1449 | 1426 | 1360 | 1296 | 1184 |
| ECM and $\geq 1\Delta X_t$ Significant* | 562 | 494 | 425 | 351 | 290 | 246 | 201 | 175 | 145 | 116 | 92 |
| ECM and $\geq 1X_{t-1}$ Significant | 2894 | 2983 | 3079 | 3204 | 3310 | 3438 | 3532 | 3577 | 3603 | 3609 | 3428 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1705 | 1570 | 1405 | 1278 | 1113 | 988 | 843 | 712 | 603 | 500 | 396 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9934 | 9885 | 9808 | 9667 | 9472 | 9190 | 8888 | 8481 | 8091 | 7707 | |
| ECM Significant* | 2381 | 1983 | 1650 | 1355 | 1094 | 909 | 753 | 664 | 592 | 521 | |
| Mean of $\alpha_1$ | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 | -0.09 | -0.08 | |
| $\geq 1\Delta X_t$ Significant | 1518 | 1518 | 1526 | 1538 | 1541 | 1552 | 1569 | 1562 | 1573 | 1568 | |
| $\geq 1X_{t-1}$ Significant | 2359 | 2445 | 2569 | 2660 | 2790 | 2937 | 3137 | 3342 | 3557 | 3827 | |
| ECM and $\geq 1\Delta X_t$ Significant | 1504 | 1499 | 1500 | 1492 | 1463 | 1437 | 1414 | 1355 | 1301 | 1241 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 379 | 330 | 265 | 234 | 195 | 162 | 138 | 117 | 109 | 92 | |
| ECM and $\geq 1X_{t-1}$ Significant | 2353 | 2436 | 2548 | 2626 | 2731 | 2852 | 2992 | 3115 | 3255 | 3416 | |
| ECM and $\geq 1X_{t-1}$ Significant* | 1093 | 984 | 886 | 770 | 681 | 607 | 538 | 494 | 465 | 426 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9933 | 9855 | 9727 | 9508 | 9146 | 8591 | 7856 | 6979 | 6080 | 5286 | 4420 |
| ECM Significant* | 1067 | 792 | 573 | 403 | 269 | 190 | 128 | 92 | 64 | 45 | 32 |
| Mean of $\alpha_1$ | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.05 | -0.05 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 1469 | 1466 | 1472 | 1474 | 1473 | 1478 | 1477 | 1472 | 1474 | 1467 | 1467 |
| $\geq 1X_{t-1}$ Significant | 1540 | 1550 | 1557 | 1558 | 1551 | 1556 | 1562 | 1559 | 1569 | 1599 | 1581 |
| ECM and $\geq 1\Delta X_t$ Significant | 1457 | 1439 | 1423 | 1391 | 1340 | 1266 | 1147 | 1029 | 908 | 794 | 659 |
| ECM and $\geq 1\Delta X_t$ Significant* | 162 | 122 | 94 | 70 | 47 | 32 | 21 | 20 | 10 | 10 | 9 |
| ECM and $\geq 1X_{t-1}$ Significant | 1531 | 1529 | 1515 | 1478 | 1424 | 1351 | 1258 | 1136 | 996 | 912 | 746 |
| ECM and $\geq 1X_{t-1}$ Significant* | 242 | 187 | 141 | 103 | 66 | 37 | 24 | 22 | 18 | 8 | 6 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance (p$\leq$ 0.05, one-tail test). Coefficient significance (p$\leq$ 0.05, two-tail test) *ECM significance (*p$\leq$ 0.05, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.78.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.44: Summary Statistics of Multivariate (4 IVs) ECM Model by DV and IV (T=150)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9963 | 9937 | 9901 | 9838 | 9745 | 9630 | 9442 | 9189 | 8897 | 8521 | 7931 |
| ECM Significant* | 2960 | 2507 | 2145 | 1820 | 1509 | 1264 | 1058 | 875 | 741 | 620 | 503 |
| Mean of $\alpha_1$ | -0.18 | -0.17 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | -0.09 |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2020 | 2027 | 2032 | 2040 | 2053 | 2059 | 2034 | 2032 | 2034 | 2057 | 2066 |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3562 | 3665 | 3789 | 3921 | 4053 | 4193 | 4347 | 4492 | 4625 | 4708 | 4678 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2014 | 2015 | 2015 | 2008 | 2003 | 1983 | 1911 | 1853 | 1801 | 1773 | 1659 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 689 | 603 | 506 | 442 | 387 | 325 | 277 | 223 | 177 | 152 | 129 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3557 | 3656 | 3775 | 3898 | 4020 | 4142 | 4265 | 4361 | 4418 | 4404 | 4258 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 1776 | 1597 | 1463 | 1303 | 1141 | 995 | 865 | 739 | 643 | 540 | 449 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9937 | 9891 | 9825 | 9701 | 9503 | 9300 | 9014 | 8731 | 8449 | 8201 | |
| ECM Significant* | 1983 | 1642 | 1403 | 1206 | 1015 | 861 | 737 | 663 | 590 | 545 | |
| Mean of $\alpha_1$ | -0.16 | -0.15 | -0.14 | -0.13 | -0.13 | -0.12 | -0.11 | -0.10 | -0.10 | -0.09 | |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 2010 | 2017 | 2021 | 2032 | 2052 | 2051 | 2061 | 2061 | 2077 | 2045 | |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3101 | 3191 | 3324 | 3461 | 3623 | 3771 | 3973 | 4183 | 4379 | 4573 | |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 1994 | 1994 | 1985 | 1970 | 1944 | 1898 | 1850 | 1800 | 1773 | 1704 | |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 446 | 377 | 310 | 274 | 232 | 199 | 189 | 169 | 150 | 141 | |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 3093 | 3179 | 3305 | 3425 | 3552 | 3667 | 3818 | 3957 | 4081 | 4194 | |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 1181 | 1039 | 934 | 844 | 750 | 667 | 588 | 559 | 510 | 480 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9912 | 9831 | 9664 | 9422 | 9030 | 8478 | 7771 | 6887 | 5998 | 5171 | 4371 |
| ECM Significant* | 592 | 422 | 288 | 203 | 126 | 80 | 52 | 36 | 28 | 18 | 15 |
| Mean of $\alpha_1$ | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.05 | -0.05 | -0.04 |
| $\geq 1\Delta X_{\mathrm{t}}$ Significant | 1922 | 1912 | 1914 | 1912 | 1914 | 1922 | 1935 | 1936 | 1932 | 1929 | 1917 |
| $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 2032 | 2034 | 2033 | 2037 | 2042 | 2040 | 2030 | 2035 | 2020 | 2009 | 2018 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant | 1900 | 1873 | 1831 | 1785 | 1712 | 1619 | 1474 | 1306 | 1154 | 990 | 865 |
| ECM and $\geq 1\Delta X_{\mathrm{t}}$ Significant* | 140 | 103 | 76 | 58 | 37 | 22 | 15 | 9 | 6 | 7 | 5 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant | 2008 | 1996 | 1964 | 1913 | 1846 | 1743 | 1607 | 1436 | 1268 | 1094 | 961 |
| ECM and $\geq 1X_{\mathrm{t\text{-}1}}$ Significant* | 195 | 141 | 98 | 58 | 42 | 27 | 19 | 14 | 10 | 7 | 7 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance ($^*p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.003.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

Table C.45: Summary Statistics of Multivariate (5 IVs) ECM Model by DV and IV (T=150)

| $\rho$ | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 | 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 9982 | 9971 | 9947 | 9899 | 9830 | 9733 | 9560 | 9367 | 9148 | 8891 | 8414 |
| ECM Significant* | 2812 | 2419 | 2092 | 1779 | 1494 | 1276 | 1082 | 896 | 741 | 626 | 505 |
| Mean of $\alpha_1$ | -0.19 | -0.18 | -0.17 | -0.16 | -0.16 | -0.15 | -0.14 | -0.13 | -0.12 | -0.11 | -0.10 |
| $\geq 1\Delta X_t$ Significant | 2479 | 2489 | 2499 | 2505 | 2502 | 2528 | 2535 | 2542 | 2549 | 2537 | 2541 |
| $\geq 1X_{t\text{-}1}$ Significant | 4474 | 4586 | 4697 | 4825 | 4959 | 5044 | 5159 | 5266 | 5386 | 5448 | 5427 |
| ECM and $\geq 1\Delta X_t$ Significant | 2477 | 2484 | 2487 | 2483 | 2460 | 2460 | 2421 | 2382 | 2342 | 2273 | 2171 |
| ECM and $\geq 1\Delta X_t$ Significant* | 797 | 703 | 615 | 524 | 440 | 392 | 339 | 282 | 246 | 206 | 173 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 4471 | 4580 | 4687 | 4806 | 4925 | 4987 | 5059 | 5116 | 5171 | 5161 | 5008 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1963 | 1764 | 1592 | 1403 | 1213 | 1057 | 929 | 801 | 668 | 573 | 470 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9925 | 9873 | 9809 | 9705 | 9559 | 9390 | 9177 | 8976 | 8770 | 8567 | |
| ECM Significant* | 1648 | 1422 | 1189 | 1014 | 886 | 773 | 677 | 592 | 541 | 510 | |
| Mean of $\alpha_1$ | -0.17 | -0.16 | -0.15 | -0.14 | -0.14 | -0.13 | -0.12 | -0.12 | -0.11 | -0.10 | |
| $\geq 1\Delta X_t$ Significant | 2402 | 2421 | 2432 | 2452 | 2456 | 2467 | 2492 | 2513 | 2543 | 2543 | |
| $\geq 1X_{t\text{-}1}$ Significant | 3772 | 3895 | 4023 | 4176 | 4356 | 4552 | 4763 | 4976 | 5182 | 5321 | |
| ECM and $\geq 1\Delta X_t$ Significant | 2386 | 2392 | 2382 | 2378 | 2349 | 2325 | 2297 | 2274 | 2252 | 2213 | |
| ECM and $\geq 1\Delta X_t$ Significant* | 478 | 422 | 356 | 300 | 274 | 250 | 224 | 196 | 183 | 174 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 3759 | 3877 | 3995 | 4132 | 4284 | 4445 | 4601 | 4749 | 4893 | 4960 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1115 | 1008 | 887 | 784 | 699 | 636 | 576 | 512 | 481 | 463 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 9879 | 9784 | 9638 | 9388 | 9001 | 8465 | 7709 | 6808 | 5920 | 5143 | 4342 |
| ECM Significant* | 359 | 266 | 181 | 127 | 75 | 48 | 30 | 22 | 19 | 9 | 14 |
| Mean of $\alpha_1$ | -0.13 | -0.12 | -0.11 | -0.10 | -0.09 | -0.08 | -0.07 | -0.06 | -0.05 | -0.05 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 2303 | 2310 | 2313 | 2308 | 2304 | 2310 | 2310 | 2319 | 2324 | 2318 | 2313 |
| $\geq 1X_{t\text{-}1}$ Significant | 2438 | 2433 | 2435 | 2438 | 2444 | 2452 | 2462 | 2484 | 2466 | 2460 | 2418 |
| ECM and $\geq 1\Delta X_t$ Significant | 2270 | 2257 | 2228 | 2162 | 2064 | 1941 | 1769 | 1581 | 1379 | 1216 | 1043 |
| ECM and $\geq 1\Delta X_t$ Significant* | 108 | 76 | 58 | 44 | 26 | 19 | 13 | 8 | 7 | 3 | 5 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 2407 | 2383 | 2351 | 2291 | 2221 | 2107 | 1953 | 1770 | 1549 | 1355 | 1136 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 150 | 107 | 79 | 53 | 33 | 22 | 16 | 12 | 9 | 5 | 6 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified level of $\rho$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.205.

$\star$ Model 1: IV is integrated $I(1)$ and DV varies by $\rho$.

$\star\star$ Model 2: Both IVs and DV data generating process based on same level of $\rho$.

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV varies by level of $\rho$ specified.

## C.8 Additional Materials: Fractionally Integrated Series

### C.8.1 Addressing Concerns as to Estimation of Long Memory

When estimating the long-memory of a series, semiparametric estimators such as Robinson's Local Whittle are appealing for a number of reasons. First, they are agnostic as to the short-run dynamics of the process, and are therefore robust to misspecification. The practitioner does not need to know the true DGP of the process to consistently estimate the long-memory parameter. Additionally, the estimators are generally robust in the presence of short-run dynamics. Semiparametric estimators require the practitioner to choose the bandwidth of the estimator, a decision which involves a trade-off between minimization of bias and variance. With a (0,$d$,0) series, increasing the number of frequencies (increasing the bandwidth) reduces the variance of the estimate at the expense of a slight increase in bias. This increase in bias is generally negligible with pure fractional processes. Should it become apparent that short-run dynamics are present however, the practitioner can simply reduce the number of frequencies, and focus more specifically on the low frequency components of the periodogram.

While this is generally true, semiparametric estimators are not perfect. In the presence of significant short-run dynamics the low frequencies are contaminated by the higher frequencies of the spectral density. Significant ARMA coefficients, especially positive AR noise, will bias estimates, sometimes substantially so. Baillie and Kapetanios (2007) investigates the bias of semiparametric estimators under these circumstances and find that the bias increases as the short-run noise becomes more persistent. Thus, in the presence of short-run dynamics, a practitioner is better served by reducing the bandwidth of the estimator, however the extent of the reduction in bias is dependent upon the degree of AR persistence.

Short-run dynamics will also bias the estimates of the exact ML (EML) estimator (Nielsen and Frederiksen 2005).[11] This bias occurs even though the EML can simultaneously estimate the short-run dynamics and the long-memory of a series. When ARMA dynamics are present, the EML estimates are negatively biased, the variance of the estimates balloon, and the estimator is generally unreliable. With persistent AR noise in the range of 0.40, the EML actually performs worse than most semiparametric estimators even when the DGP is properly specified. The problems associated with the EML are alleviated somewhat with longer time series, but not to the point that using the EML in the presence of AR dynamics is recommended.

The possibility of finite sample bias of semiparametric estimators and the outright failure of the EML estimator when AR noise is present does not mean that long-memory can't be reliably estimated. First, in terms of estimators, the problems highlighted above are not encountered with the approximate frequency domain ML (FML) estimator. Nielsen and Frederiksen (2005) compares the performance of the FML in larger samples and finds that the estimator is unbiased, robust to AR and MA noise, and doesn't suffer from bias when the true DGP is overfit. Grant (n.d.) compares the FML to various semiparametric and parametric estimators in small samples and finds the FML dominates all other estimators in terms of bias and RMSE, even when AR and MA noise is present. Thus, reliable estimates are possible even when a series exhibits significant short-run dynamics.

Another important consideration when it comes to estimating the long-memory of a series

[11] The EML is the standard estimator in Stata's ARFIMA package.

is whether we should expect political time series to exhibit significant short-run dynamics. In other words, do we routinely estimate political time series as higher-order processes? Byers, Davidson, and Peel (2000) estimates the long-memory of a number of political opinion series and finds that a large majority of the series were best approximated by (0,$d$,0) models. Of those series for which the information criteria chose a higher-order model, the gains were often minimal. This is expected to be the case for most political time series, which are generally shorter in duration and the product of an aggregation of heterogeneous autoregressive processes (Granger and Joyeux 1980).[12] To demonstrate that most series are fractional noise - they are uncorrelated after differencing by only the $d$ parameter - and that long-memory estimates can be consistent, we provide Tables C.46 and C.47. Table C.46 shows the BIC values after our fitting multiple $(p, d, q)$ models - the BIC values are from the Stata EML estimator.[13] Table C.47 presents long-memory estimates of each series from three different estimators - the semiparametric Local Whittle of Robinson (1995), the EML estimator used in Stata's ARFIMA package, and the FML estimator, which is estimated using the TSM add-on of Davidson (2014) to Ox (Doornik 2014). A fourth column also reports the results of the Box-Pierce test of the residuals of each series after fractional differencing. The results are consistent across all three estimator types.

**Table C.46: Exact Maximum-Likelihood Estimation - BIC Values**

| | (0,$d$,0) | (1,$d$,0) | (0,$d$,1) | (1,$d$,1) |
|---|---|---|---|---|
| **CEW** | | | | |
| All Reviews | 334.19⋆ | 335.44 | 335.17 | 338.41 |
| Non-Salient Reviews | 332.48⋆ | 334.57 | 334.42 | 338.01 |
| Salient Reviews | 419.34⋆ | 420.89 | 421.54 | 424.71 |
| **UE** | | | | |
| Democrat Mood | 129.98⋆ | 136.59 | 146.85 | 140.01 |
| Republican Mood | 151.57⋆ | 156.15 | 155.04 | 158.00 |
| **SSRS** | | | | |
| Canada Rights | -65.85⋆ | -65.76 | -65.80 | -65.68 |
| UK Rights | -79.11⋆ | -75.51 | -75.50 | *** |
| US Rights | -141.89⋆ | -138.59 | -138.39 | -134.52 |
| **KE** | | | | |
| Liberal Mood | 245.23⋆ | 257.14 | 247.51 | 250.86 |
| Welfare Mood | 215.18⋆ | 221.72 | 217.21 | 215.99 |
| Low Income Mood | 266.57⋆ | 269.10 | 268.78 | 272.69 |
| High Income Mood | 253.14⋆ | 260.67 | 255.88 | 264.00 |
| **VK** | | | | |
| Top 1% Income | 344.99⋆ | 348.41 | 348.26 | 352.80 |

*Note*: ⋆ Marks the minimum value in each row. *** indicates the model did not converge in Stata.

As seen in Table C.46 of the thirteen dependent variables estimated, the BIC chose a (0,$d$,0) model for each. This is not to say that we will never encounter a higher-order process, but quite often the simple model that estimates only the long memory will be sufficient. Further, absent the presence of significant short-memory dynamics many of the

[12] Previous work on fractional integration of political time series has reported higher-order models after relying on information criteria (see, e.g., Box-Steffensmeier and Tomlinson 2000). This is most certainly due to an error in the formula used by Ox at the time. Re-estimating these series with both the EML and FML estimators using Schwartz Information Criteria indicates that (0,$d$,0) models are preferred.

[13] Because the series are generally short, we only go so far as to estimate (1,$d$,1) models.

estimation problems discussed above are not at issue with our estimates. The estimates do not suffer from bias nor is the variance of our estimators inflated. In fact, our estimates are remarkably consistent across the three different estimators and the Box-Pierce residual test is insignificant for the residuals of each series. Fractionally differencing our series with its estimate of $d$ is sufficient to address any autocorrelation in the series. Additionally, note that the 95% confidence intervals of many of the series would overlap with 1, indicating a unit-root. Considering the low power of many unit-root tests, the ability to estimate the order of fractional integration provides an additional tool when a variable's order of integration is in question. This can be seen in the disparity in results for *All Reviews* and *Non-Salient Reviews*, both of which failed to reject the null of the Dickey-Fuller, but which all three of our estimators found to be fractionally integrated.

**Table C.47: Various Estimates of Long Memory - (0,d,0) Models**

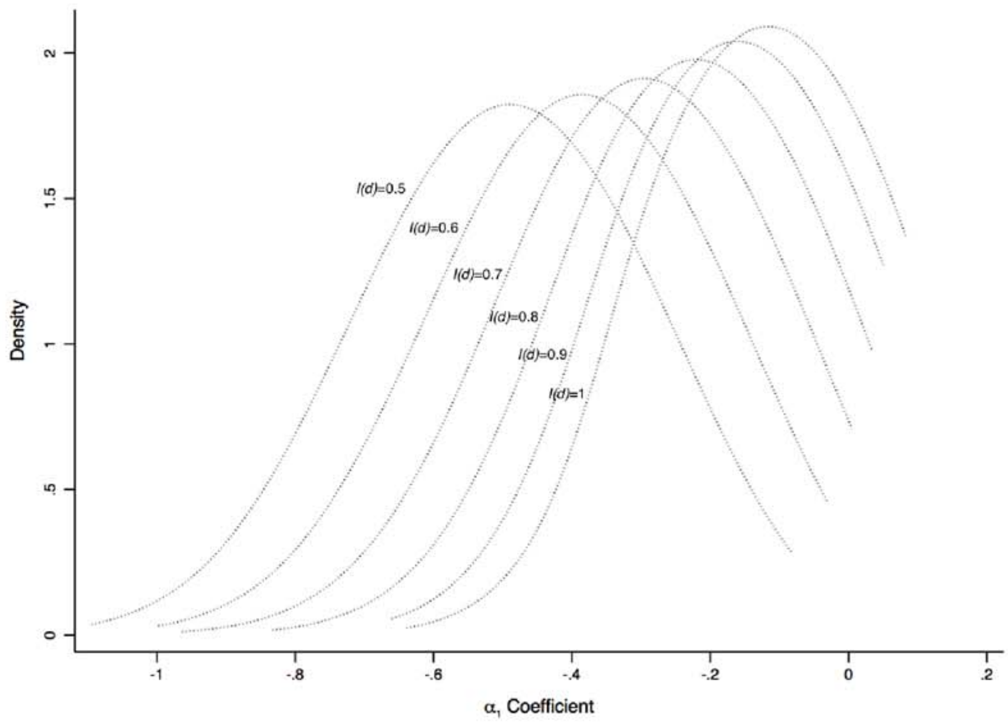
| | LW | EML | FML | Box-Pierce (12 lags) |
|---|---|---|---|---|
| **CEW** | | | | |
| All Reviews | 0.63 (0.11) | 0.62 (0.11) | 0.63 (0.11) | 18.51 |
| Non-Salient Reviews | 0.65 (0.11) | 0.62 (0.11) | 0.65 (0.13) | 15.32 |
| Salient Reviews | 0.30 (0.11) | 0.36 (0.08) | 0.31 (0.12) | 11.24 |
| **UE** | | | | |
| Democrat Mood | 1.28 (0.12) | 1.15 (0.15) | 1.27 (0.18) | 5.18 |
| Republican Mood | 1.17 (0.12) | 1.04 (0.16) | 1.17 (0.18) | 6.34 |
| **SSRS** | | | | |
| Canada Rights | 0.62 (0.10) | 0.52 (0.12) | 0.57 (0.13) | 9.38 |
| UK Rights | 0.08 (0.12) | 0.13 (0.13) | 0.08 (0.17) | 4.98 |
| US Rights | 0.40 (0.10) | 0.45 (0.05) | 0.41 (0.10) | 11.79 |
| **KE** | | | | |
| Liberal Mood | 1.12 (0.10) | 1.08 (0.11) | 1.14 (0.11) | 12.92 |
| Welfare Mood | 1.01 (0.13) | 0.88 (0.18) | 1.00 (0.21) | 8.09 |
| Low Income Mood | 1.03 (0.11) | 0.98 (0.13) | 1.05 (0.14) | 11.33 |
| High Income Mood | 1.11 (0.11) | 1.04 (0.13) | 1.15 (0.14) | 5.15 |
| **VK** | | | | |
| Top 1% Income | 0.93 (0.10) | 0.95 (0.09) | 0.91 (0.11) | 7.30 |

*Note*: LW: semiparametric Local Whittle; EML: exact maximum likelihood; FML: frequency maximum likelihood. All estimates based on the first-differenced series and then adding 1 back to the estimate. Standard errors in parentheses. Standard errors for the LW are computed as $(1/sqrt4 * m)$ with the bandwidth $m$ calculated as $T^{4/5}$. Box-Pierce test is of the residuals following fractional differencing with the FML estimator.

### C.8.2 Shifting ECM Coefficient

Any deviation from strictly $I(1)$ series will lead to substantial issues for the interpretation of the ECM coefficient and its $t$-statistic. Figures 3 and 4 in the main paper depict both the increasing significance of the ECM as well as the increasing (absolute) size of the coefficient as the fractional integration of the dependent variable moves further below 1. Figure C.9 presents another look at the size of the ECM coefficient. The bias in the ECM coefficient is troubling. If our DV series is fractionally integrated, we are likely to overstate the strength of whatever equilibrium relationship is being studied. Further, because the Long Run Multipliers are are calculated as a ratio in which the denominator is the $\alpha_1$ coefficient, our estimation and interpretation of the Long Run Multipliers will also be subject to this bias.

**Figure C.9: Mean ECM Coefficient by Order of Fractional Integration**




*Note*: Density plots from 10,000 Monte Carlo simulations of bivariate ECM model. DV and IV both fractionally integrated of order $I(d)$. $T = 60$

### C.8.3 Tables - I(d) - Monte Carlo Simulations

The results in the following tables are meant to provide as much information as possible in as little space as possible. Therefore, we have tried to include a wide a range of models. The top row of each table contains models run in which the order of fractional integration in the DV varies, but the IV(s) all have unit roots. The middle row contain results of models in which all series within each model share the same order of fractional integration. These equations are balanced. The bottom row also allows the DV to vary by its order of fractional integration, but the IV(s) are all generated as $I(0)$ - the IV(s) contain very little, if any, information.

Table C.48: Summary Statistics of Bivariate ECM Model by Fractionally Integrated DV and IV (T=60)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9998 | 9959 | 9686 | 8853 | 7356 | 5675 |
| ECM Significant* | 10000 | 10000 | 10000 | 9997 | 9862 | 9099 | 7233 | 4627 | 2425 | 1070 | 533 |
| Mean of $\alpha_1$ | -1.03 | -1.03 | -0.93 | -0.79 | -0.64 | -0.53 | -0.42 | -0.32 | -0.24 | -0.17 | -0.12 |
| $\geq 1\Delta X_t$ Significant | 574 | 588 | 596 | 609 | 640 | 670 | 661 | 659 | 643 | 624 | 610 |
| $\geq 1X_{t-1}$ Significant | 603 | 579 | 932 | 1371 | 1744 | 1922 | 1954 | 1888 | 1763 | 1630 | 1716 |
| ECM and $\geq 1\Delta X_t$ Significant | 574 | 588 | 596 | 609 | 640 | 670 | 660 | 642 | 590 | 491 | 384 |
| ECM and $\geq 1\Delta X_t$ Significant* | 574 | 588 | 596 | 609 | 634 | 620 | 516 | 338 | 188 | 82 | 49 |
| ECM and $\geq 1X_{t-1}$ Significant | 603 | 579 | 932 | 1371 | 1744 | 1922 | 1954 | 1886 | 1744 | 1541 | 1410 |
| ECM and $\geq 1X_{t-1}$ Significant* | 603 | 579 | 932 | 1371 | 1744 | 1900 | 1808 | 1446 | 924 | 499 | 310 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 9998 | 9931 | 9607 | 8651 | 7160 | |
| ECM Significant* | | 10000 | 10000 | 9990 | 9752 | 8685 | 6548 | 4042 | 2092 | 1014 | |
| Mean of $\alpha_1$ | | -1.02 | -0.91 | -0.76 | -0.61 | -0.50 | -0.40 | -0.30 | -0.22 | -0.16 | |
| $\geq 1\Delta X_t$ Significant | | 569 | 582 | 610 | 667 | 687 | 685 | 677 | 646 | 610 | |
| $\geq 1X_{t-1}$ Significant | | 540 | 570 | 711 | 879 | 1005 | 1152 | 1234 | 1328 | 1471 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 569 | 582 | 610 | 667 | 687 | 679 | 644 | 561 | 458 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 569 | 582 | 609 | 652 | 599 | 452 | 299 | 172 | 91 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 540 | 570 | 711 | 879 | 1004 | 1150 | 1230 | 1303 | 1367 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 540 | 570 | 710 | 876 | 978 | 1041 | 904 | 678 | 484 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9996 | 9909 | 9390 | 8107 | 6140 | 4482 |
| ECM Significant* | 10000 | 10000 | 10000 | 9989 | 9692 | 8353 | 5804 | 3096 | 1364 | 501 | 219 |
| Mean of $\alpha_1$ | -1.01 | -1.02 | -0.91 | -0.76 | -0.59 | -0.48 | -0.37 | -0.27 | -0.19 | -0.13 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 560 | 578 | 581 | 564 | 562 | 579 | 576 | 580 | 607 | 592 | 542 |
| $\geq 1X_{t-1}$ Significant | 541 | 559 | 567 | 537 | 512 | 510 | 503 | 500 | 519 | 557 | 593 |
| ECM and $\geq 1\Delta X_t$ Significant | 560 | 578 | 581 | 564 | 562 | 579 | 572 | 535 | 477 | 369 | 257 |
| ECM and $\geq 1\Delta X_t$ Significant* | 560 | 578 | 581 | 563 | 543 | 329 | 363 | 177 | 88 | 29 | 21 |
| ECM and $\geq 1X_{t-1}$ Significant | 541 | 559 | 567 | 537 | 512 | 510 | 501 | 481 | 450 | 386 | 308 |
| ECM and $\geq 1X_{t-1}$ Significant* | 541 | 559 | 567 | 537 | 506 | 472 | 369 | 226 | 132 | 59 | 33 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.27.

$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified

$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

**Table C.49: Summary Statistics of Multivariate (2IVs) ECM Model by Fractionally Integrated DV and IV (T=60)**

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9997 | 9984 | 9837 | 9261 | 8079 | 6488 |
| ECM Significant* | 10000 | 10000 | 9998 | 9989 | 9723 | 8802 | 6834 | 4368 | 2379 | 1129 | 552 |
| Mean of $\alpha_1$ | -1.05 | -1.05 | -0.95 | -0.82 | -0.68 | -0.57 | -0.47 | -0.37 | -0.28 | -0.21 | -0.15 |
| $\geq 1\Delta X_t$ Significant | 1163 | 1033 | 1078 | 1131 | 1189 | 1158 | 1187 | 1175 | 1172 | 1187 | 1227 |
| $\geq 1X_{t-1}$ Significant | 1112 | 1067 | 1586 | 2226 | 2725 | 3098 | 3177 | 3133 | 3020 | 2955 | 2945 |
| ECM and $\geq 1\Delta X_t$ Significant | 1163 | 1033 | 1078 | 1131 | 1189 | 1157 | 1186 | 1161 | 1115 | 1013 | 862 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1163 | 1033 | 1078 | 1130 | 1158 | 1058 | 888 | 627 | 384 | 206 | 110 |
| ECM and $\geq 1X_{t-1}$ Significant | 1112 | 1067 | 1586 | 2226 | 2725 | 3098 | 3176 | 3132 | 2993 | 2797 | 2497 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1112 | 1067 | 1586 | 2226 | 2715 | 3007 | 2813 | 2150 | 1380 | 762 | 430 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 9999 | 9992 | 9931 | 9650 | 8930 | 7824 | |
| ECM Significant* | | 10000 | 9998 | 9951 | 9366 | 7940 | 5552 | 3369 | 1783 | 933 | |
| Mean of $\alpha_1$ | | -1.02 | -0.91 | -0.77 | -0.62 | -0.52 | -0.42 | -0.33 | -0.26 | -0.20 | |
| $\geq 1\Delta X_t$ Significant | | 1013 | 1044 | 1103 | 1179 | 1208 | 1221 | 1236 | 1208 | 1181 | |
| $\geq 1X_{t-1}$ Significant | | 1064 | 1137 | 1329 | 1594 | 1787 | 1999 | 2156 | 2304 | 2511 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1013 | 1044 | 1103 | 1179 | 1208 | 1218 | 1203 | 1114 | 975 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1013 | 1043 | 1097 | 1107 | 1006 | 728 | 495 | 296 | 168 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 1064 | 1137 | 1329 | 1594 | 1787 | 1998 | 2145 | 2259 | 2369 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 1064 | 1137 | 1328 | 1571 | 1676 | 1593 | 1309 | 936 | 600 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9995 | 9873 | 9329 | 7976 | 6033 | 4340 |
| ECM Significant* | 10000 | 10000 | 9998 | 9947 | 9100 | 7005 | 4209 | 1988 | 755 | 245 | 94 |
| Mean of $\alpha_1$ | -1.02 | -1.02 | -0.90 | -0.75 | -0.59 | -0.48 | -0.37 | -0.27 | -0.19 | -0.13 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 1124 | 1050 | 1059 | 1066 | 1065 | 1045 | 1027 | 1032 | 1021 | 1030 | 1099 |
| $\geq 1X_{t-1}$ Significant | 1088 | 1116 | 1123 | 1124 | 1053 | 950 | 946 | 962 | 980 | 1072 | 1195 |
| ECM and $\geq 1\Delta X_t$ Significant | 1124 | 1050 | 1059 | 1066 | 1065 | 1043 | 1004 | 963 | 826 | 650 | 546 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1124 | 1050 | 1059 | 1059 | 960 | 728 | 472 | 273 | 126 | 51 | 20 |
| ECM and $\geq 1X_{t-1}$ Significant | 1088 | 1116 | 1123 | 1124 | 1053 | 950 | 938 | 911 | 830 | 707 | 634 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1088 | 1116 | 1123 | 1121 | 1010 | 807 | 583 | 348 | 162 | 64 | 22 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.565.
$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

**Table C.50: Summary Statistics of Multivariate (3IVs) ECM Model by Fractionally Integrated DV and IV (T=60)**

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9988 | 9882 | 9467 | 8586 | 7197 |
| ECM Significant* | 10000 | 10000 | 10000 | 9972 | 9556 | 8552 | 6548 | 4234 | 2275 | 1067 | 527 |
| Mean of $\alpha_1$ | -1.07 | -1.07 | -0.97 | -0.85 | -0.72 | -0.61 | -0.51 | -0.42 | -0.32 | -0.25 | -0.18 |
| $\geq 1\Delta X_t$ Significant | 1542 | 1540 | 1572 | 1646 | 1714 | 1723 | 1733 | 1763 | 1754 | 1762 | 1733 |
| $\geq 1X_{t\text{-}1}$ Significant | 1563 | 1561 | 2163 | 2883 | 3509 | 3828 | 4004 | 3979 | 3904 | 3864 | 3897 |
| ECM and $\geq 1\Delta X_t$ Significant | 1542 | 1540 | 1572 | 1646 | 1714 | 1723 | 1730 | 1738 | 1675 | 1563 | 1329 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1542 | 1540 | 1572 | 1642 | 1653 | 1536 | 1266 | 885 | 542 | 303 | 152 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1563 | 1561 | 2163 | 2883 | 3509 | 3828 | 4004 | 3974 | 3873 | 3712 | 3415 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1563 | 1561 | 2163 | 2883 | 3490 | 3646 | 3304 | 2480 | 1566 | 835 | 452 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 9999 | 9994 | 9941 | 9724 | 9161 | 8268 | |
| ECM Significant* | | 9999 | 9993 | 9870 | 8801 | 7070 | 4790 | 2912 | 1586 | 874 | |
| Mean of $\alpha_1$ | | -1.02 | -0.91 | -0.78 | -0.64 | -0.54 | -0.44 | -0.36 | -0.28 | -0.23 | |
| $\geq 1\Delta X_t$ Significant | | 1589 | 1620 | 1680 | 1801 | 1836 | 1847 | 1828 | 1797 | 1781 | |
| $\geq 1X_{t\text{-}1}$ Significant | | 1499 | 1611 | 1861 | 2240 | 2488 | 2764 | 2997 | 3256 | 3501 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1589 | 1620 | 1680 | 1800 | 1833 | 1838 | 1778 | 1680 | 1539 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1589 | 1619 | 1651 | 1588 | 1331 | 965 | 632 | 408 | 255 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | | 1499 | 1611 | 1861 | 2240 | 2487 | 2762 | 2980 | 3195 | 3306 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | | 1499 | 1611 | 1861 | 2167 | 2201 | 1988 | 1550 | 1069 | 693 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9991 | 9853 | 9201 | 7776 | 5880 | 4206 |
| ECM Significant* | 10000 | 10000 | 9995 | 9830 | 8194 | 5586 | 2937 | 1207 | 415 | 125 | 53 |
| Mean of $\alpha_1$ | -1.02 | -1.02 | -0.91 | -0.76 | -0.59 | -0.48 | -0.37 | -0.27 | -0.19 | -0.13 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 1500 | 1515 | 1503 | 1507 | 1504 | 1545 | 1535 | 1532 | 1531 | 1514 | 1530 |
| $\geq 1X_{t\text{-}1}$ Significant | 1546 | 1515 | 1526 | 1478 | 1424 | 1408 | 1371 | 1369 | 1403 | 1501 | 1683 |
| ECM and $\geq 1\Delta X_t$ Significant | 1500 | 1515 | 1503 | 1507 | 1504 | 1541 | 1501 | 1390 | 1196 | 925 | 701 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1500 | 1515 | 1503 | 1468 | 1218 | 909 | 515 | 238 | 97 | 40 | 21 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1546 | 1515 | 1526 | 1478 | 1424 | 1407 | 1352 | 1298 | 1170 | 1007 | 846 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1546 | 1515 | 1526 | 1466 | 1299 | 1033 | 641 | 329 | 132 | 50 | 25 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.816.

$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified

$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

**Table C.51: Summary Statistics of Multivariate (4IVs) ECM Model by Fractionally Integrated DV and IV (T=60)**

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9991 | 9927 | 9640 | 8926 | 7666 |
| ECM Significant* | 10000 | 9999 | 9998 | 9934 | 9407 | 8238 | 6201 | 3997 | 2127 | 1030 | 496 |
| Mean of $\alpha_1$ | -1.09 | -1.09 | -1.00 | -0.88 | -0.75 | -0.65 | -0.56 | -0.46 | -0.37 | -0.29 | -0.21 |
| $\geq 1\Delta X_t$ Significant | 2073 | 2016 | 2057 | 2161 | 2243 | 2242 | 2284 | 2312 | 2282 | 2314 | 2292 |
| $\geq 1X_{t\text{-}1}$ Significant | 2069 | 2074 | 2738 | 3445 | 4047 | 4324 | 4509 | 4585 | 4575 | 4534 | 4599 |
| ECM and $\geq 1\Delta X_t$ Significant | 2073 | 2016 | 2057 | 2161 | 2243 | 2242 | 2282 | 2296 | 2214 | 2102 | 1838 |
| ECM and $\geq 1\Delta X_t$ Significant* | 2073 | 2015 | 2053 | 2146 | 2151 | 1916 | 1549 | 1105 | 639 | 334 | 200 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 2069 | 2074 | 2738 | 3445 | 4047 | 4324 | 4508 | 4577 | 4530 | 4369 | 4108 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 2069 | 2074 | 2737 | 3440 | 3983 | 4005 | 3524 | 2654 | 1593 | 870 | 447 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 9995 | 9948 | 9774 | 9297 | 8553 | |
| ECM Significant* | | 9998 | 9978 | 9659 | 8210 | 6183 | 4158 | 2482 | 1425 | 802 | |
| Mean of $\alpha_1$ | | -1.02 | -0.92 | -0.79 | -0.65 | -0.56 | -0.47 | -0.39 | -0.32 | -0.26 | |
| $\geq 1\Delta X_t$ Significant | | 1987 | 2041 | 2125 | 2247 | 2246 | 2281 | 2265 | 2264 | 2282 | |
| $\geq 1X_{t\text{-}1}$ Significant | | 1918 | 2045 | 2354 | 2846 | 3104 | 3394 | 3654 | 3917 | 4286 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1987 | 2041 | 2125 | 2247 | 2245 | 2268 | 2215 | 2141 | 2026 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1985 | 2035 | 2044 | 1891 | 1514 | 1084 | 719 | 438 | 284 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | | 1918 | 2045 | 2354 | 2846 | 3104 | 3390 | 3632 | 3844 | 4067 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | | 1918 | 2045 | 2333 | 2657 | 2505 | 2161 | 1568 | 1031 | 672 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 9998 | 9977 | 9789 | 9008 | 7544 | 5682 | 3998 |
| ECM Significant* | 9996 | 9995 | 9970 | 9497 | 7150 | 4376 | 2023 | 742 | 246 | 66 | 16 |
| Mean of $\alpha_1$ | -1.02 | -1.02 | -0.91 | -0.76 | -0.60 | -0.48 | -0.37 | -0.27 | -0.20 | -0.13 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 1982 | 2028 | 2048 | 2055 | 2038 | 1984 | 2007 | 2000 | 1979 | 1979 | 2012 |
| $\geq 1X_{t\text{-}1}$ Significant | 2007 | 2086 | 2088 | 2028 | 1928 | 1839 | 1762 | 1799 | 1872 | 2038 | 2201 |
| ECM and $\geq 1\Delta X_t$ Significant | 1982 | 2028 | 2048 | 2055 | 2038 | 1975 | 1938 | 1777 | 1490 | 1159 | 881 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1980 | 2027 | 2039 | 1942 | 1485 | 902 | 435 | 155 | 63 | 23 | 9 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 2007 | 2086 | 2088 | 2028 | 1928 | 1834 | 1739 | 1667 | 1517 | 1159 | 1055 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 2005 | 2085 | 2087 | 1984 | 1638 | 1114 | 588 | 248 | 115 | 23 | 8 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: $5\% \leq -4.035$.
$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.52: Summary Statistics of Multivariate (5IVs) ECM Model by Fractionally Integrated DV and IV (T=60)

| | $I(0)$ | $d = .1$ | $d = .2$ | $d = .3$ | $d = .4$ | $d = .5$ | $d = .6$ | $d = .7$ | $d = .8$ | $d = .9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 9998 | 9992 | 9947 | 9726 | 9108 | 8024 |
| ECM Significant* | 9998 | 9998 | 9978 | 9847 | 9192 | 7913 | 5979 | 3813 | 2143 | 1078 | 514 |
| Mean of $\alpha_1$ | -1.10 | -1.10 | -1.02 | -0.91 | -0.79 | -0.69 | -0.60 | -0.50 | -0.41 | -0.32 | -0.24 |
| $\geq 1\Delta X_t$ Significant | 2454 | 2517 | 2598 | 2700 | 2802 | 2766 | 2801 | 2822 | 2851 | 2867 | 2875 |
| $\geq 1X_{t-1}$ Significant | 2492 | 2534 | 3129 | 3837 | 4436 | 4842 | 5042 | 5156 | 5166 | 5180 | 5228 |
| ECM and $\geq 1\Delta X_t$ Significant | 2454 | 2517 | 2598 | 2700 | 2802 | 2766 | 2799 | 2812 | 2793 | 2672 | 2428 |
| ECM and $\geq 1\Delta X_t$ Significant* | 2453 | 2517 | 2597 | 2679 | 2642 | 2323 | 1881 | 1315 | 868 | 476 | 250 |
| ECM and $\geq 1X_{t-1}$ Significant | 2492 | 2534 | 3129 | 3837 | 4436 | 4842 | 5041 | 5149 | 5129 | 5011 | 4710 |
| ECM and $\geq 1X_{t-1}$ Significant* | 2492 | 2534 | 3128 | 3823 | 4311 | 4352 | 3808 | 2798 | 1727 | 942 | 481 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 9999 | 9990 | 9932 | 9757 | 9405 | 8809 | |
| ECM Significant* | | 9982 | 9908 | 9322 | 7451 | 5417 | 3556 | 2137 | 1303 | 756 | |
| Mean of $\alpha_1$ | | -1.02 | -0.92 | -0.79 | -0.66 | -0.57 | -0.49 | -0.41 | -0.35 | -0.29 | |
| $\geq 1\Delta X_t$ Significant | | 2409 | 2419 | 2511 | 2655 | 2689 | 2711 | 2765 | 2779 | 2818 | |
| $\geq 1X_{t-1}$ Significant | | 2359 | 2414 | 2772 | 3222 | 3573 | 3915 | 4195 | 4493 | 4833 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 2409 | 2419 | 2511 | 2654 | 2684 | 2686 | 2696 | 2633 | 2563 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 2401 | 2391 | 2338 | 2044 | 1609 | 1144 | 755 | 522 | 337 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 2359 | 2414 | 2772 | 3221 | 3571 | 3911 | 4171 | 4413 | 4598 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 2358 | 2406 | 2709 | 2883 | 2637 | 2189 | 1563 | 1065 | 676 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 9998 | 9976 | 9751 | 8945 | 7396 | 5425 | 3959 |
| ECM Significant* | 9986 | 9989 | 9883 | 8970 | 5928 | 3178 | 1308 | 440 | 147 | 37 | 12 |
| Mean of $\alpha_1$ | -1.02 | -1.02 | -0.91 | -0.76 | -0.60 | -0.48 | -0.37 | -0.28 | -0.20 | -0.13 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 2466 | 2451 | 2468 | 2498 | 2486 | 2361 | 2389 | 2422 | 2411 | 2425 | 2467 |
| $\geq 1X_{t-1}$ Significant | 2406 | 2434 | 2438 | 2367 | 2260 | 2161 | 2124 | 2126 | 2254 | 2396 | 2633 |
| ECM and $\geq 1\Delta X_t$ Significant | 2466 | 2451 | 2468 | 2498 | 2486 | 2354 | 2323 | 2155 | 1803 | 1361 | 1057 |
| ECM and $\geq 1\Delta X_t$ Significant* | 2460 | 2447 | 2431 | 2251 | 1577 | 787 | 357 | 133 | 52 | 17 | 6 |
| ECM and $\geq 1X_{t-1}$ Significant | 2406 | 2434 | 2438 | 2367 | 2259 | 2157 | 2093 | 1965 | 1804 | 1481 | 1251 |
| ECM and $\geq 1X_{t-1}$ Significant* | 2403 | 2433 | 2432 | 2264 | 1736 | 1023 | 497 | 198 | 81 | 21 | 7 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.229.

$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified

$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified

$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.53: Summary Statistics of Bivariate ECM Model by Fractionally Integrated DV and IV (T=150)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9980 | 9604 | 8091 | 5769 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9998 | 9770 | 8056 | 4497 | 1624 | 501 |
| Mean of $\alpha_1$ | -1.01 | -1.01 | -0.89 | -0.73 | -0.55 | -0.42 | -0.30 | -0.21 | -0.13 | -0.08 | -0.05 |
| $\geq 1\Delta X_t$ Significant | 536 | 514 | 530 | 539 | 564 | 527 | 529 | 522 | 515 | 524 | 562 |
| $\geq 1X_{t\text{-}1}$ Significant | 530 | 535 | 1173 | 2057 | 2709 | 2925 | 2846 | 2565 | 2092 | 1721 | 1639 |
| ECM and $\geq 1\Delta X_t$ Significant | 536 | 514 | 530 | 539 | 564 | 527 | 529 | 522 | 491 | 438 | 338 |
| ECM and $\geq 1\Delta X_t$ Significant* | 536 | 514 | 530 | 539 | 564 | 527 | 517 | 425 | 240 | 83 | 36 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 530 | 535 | 1173 | 2057 | 2709 | 2925 | 2846 | 2565 | 2090 | 1671 | 1343 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 530 | 535 | 1173 | 2057 | 2709 | 2925 | 2843 | 2438 | 1551 | 717 | 270 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9959 | 8476 | 7938 | |
| ECM Significant* | | 10000 | 10000 | 10000 | 10000 | 9990 | 9615 | 7492 | 4009 | 1466 | |
| Mean of $\alpha_1$ | | -1.01 | -0.89 | -0.72 | -0.53 | -0.40 | -0.29 | -0.19 | -0.12 | -0.08 | |
| $\geq 1\Delta X_t$ Significant | | 508 | 522 | 567 | 639 | 700 | 689 | 641 | 602 | 539 | |
| $\geq 1X_{t\text{-}1}$ Significant | | 516 | 554 | 717 | 1036 | 1296 | 1490 | 1544 | 1515 | 1511 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 508 | 522 | 567 | 639 | 700 | 688 | 637 | 565 | 425 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 508 | 522 | 567 | 639 | 699 | 660 | 487 | 270 | 92 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | | 516 | 554 | 717 | 1036 | 1296 | 1490 | 1544 | 1511 | 1449 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | | 516 | 554 | 717 | 1036 | 1296 | 1484 | 1448 | 1090 | 613 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9941 | 9187 | 7045 | 4542 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9991 | 9448 | 6710 | 2901 | 802 | 216 |
| Mean of $\alpha_1$ | -1.01 | -1.01 | -0.88 | -0.71 | -0.52 | -0.39 | -0.27 | -0.18 | -0.11 | -0.06 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 532 | 498 | 491 | 490 | 516 | 477 | 473 | 491 | 494 | 493 | 536 |
| $\geq 1X_{t\text{-}1}$ Significant | 530 | 521 | 525 | 501 | 443 | 393 | 363 | 339 | 385 | 444 | 561 |
| ECM and $\geq 1\Delta X_t$ Significant | 532 | 498 | 491 | 490 | 516 | 477 | 473 | 489 | 453 | 342 | 270 |
| ECM and $\geq 1\Delta X_t$ Significant* | 532 | 498 | 491 | 490 | 516 | 477 | 446 | 327 | 136 | 43 | 13 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 530 | 521 | 525 | 501 | 443 | 393 | 363 | 338 | 354 | 337 | 294 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 530 | 521 | 525 | 501 | 443 | 393 | 350 | 257 | 141 | 47 | 14 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.236.
$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.54: Summary Statistics of Multivariate (2IVs) ECM Model by Fractionally Integrated DV and IV (T=150)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9991 | 9794 | 8721 | 6695 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9998 | 9777 | 8086 | 4480 | 1667 | 467 |
| Mean of $\alpha_1$ | -1.02 | -1.02 | -0.91 | -0.75 | -0.58 | -0.45 | -0.34 | -0.24 | -0.16 | -0.10 | -0.06 |
| $\geq 1\Delta X_t$ Significant | 994 | 1005 | 1045 | 1079 | 1109 | 1094 | 1113 | 1103 | 1073 | 1071 | 1047 |
| $\geq 1X_{t-1}$ Significant | 1027 | 998 | 1954 | 3189 | 4053 | 4380 | 4365 | 4047 | 3553 | 3119 | 2958 |
| ECM and $\geq 1\Delta X_t$ Significant | 994 | 1005 | 1045 | 1079 | 1109 | 1094 | 1113 | 1100 | 1054 | 937 | 728 |
| ECM and $\geq 1\Delta X_t$ Significant* | 994 | 1005 | 1045 | 1079 | 1109 | 1094 | 1081 | 905 | 526 | 208 | 59 |
| ECM and $\geq 1X_{t-1}$ Significant | 1027 | 998 | 1954 | 3189 | 4053 | 4380 | 4365 | 4047 | 3552 | 3051 | 2525 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1027 | 998 | 1954 | 3189 | 4053 | 4380 | 4351 | 3782 | 2459 | 1064 | 365 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9972 | 9669 | 8481 | |
| ECM Significant* | | 10000 | 10000 | 10000 | 10000 | 9972 | 9443 | 7013 | 3629 | 1321 | |
| Mean of $\alpha_1$ | | -1.01 | -0.89 | -0.72 | -0.54 | -0.41 | -0.30 | -0.21 | -0.14 | -0.09 | |
| $\geq 1\Delta X_t$ Significant | | 1033 | 1036 | 1159 | 1320 | 1332 | 1330 | 1246 | 1143 | 1088 | |
| $\geq 1X_{t-1}$ Significant | | 1024 | 1071 | 1395 | 1876 | 2275 | 2562 | 2664 | 2640 | 2660 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1033 | 1036 | 1159 | 1320 | 1332 | 1329 | 1243 | 1106 | 919 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1033 | 1036 | 1159 | 1320 | 1330 | 1259 | 864 | 454 | 166 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 1024 | 1071 | 1395 | 1876 | 2275 | 2562 | 2664 | 2637 | 2567 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 1024 | 1071 | 1395 | 1876 | 2275 | 2542 | 2372 | 1651 | 809 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9940 | 9208 | 7002 | 4475 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9944 | 8967 | 5379 | 1830 | 404 | 85 |
| Mean of $\alpha_1$ | -1.00 | -1.00 | -0.88 | -0.71 | -0.52 | -0.39 | -0.27 | -0.18 | -0.11 | -0.06 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 984 | 995 | 1016 | 990 | 1011 | 1012 | 1011 | 1006 | 997 | 982 | 969 |
| $\geq 1X_{t-1}$ Significant | 984 | 998 | 967 | 891 | 832 | 819 | 761 | 750 | 798 | 888 | 1018 |
| ECM and $\geq 1\Delta X_t$ Significant | 984 | 995 | 1016 | 990 | 1011 | 1012 | 1011 | 1002 | 925 | 697 | 447 |
| ECM and $\geq 1\Delta X_t$ Significant* | 984 | 998 | 1016 | 990 | 1011 | 1008 | 906 | 569 | 199 | 49 | 7 |
| ECM and $\geq 1X_{t-1}$ Significant | 984 | 998 | 967 | 891 | 832 | 819 | 761 | 748 | 739 | 660 | 492 |
| ECM and $\geq 1X_{t-1}$ Significant* | 984 | 998 | 967 | 891 | 832 | 816 | 701 | 455 | 205 | 66 | 14 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.528.
$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.55: Summary Statistics of Multivariate (3IVs) ECM Model by Fractionally Integrated DV and IV (T=150)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**⋆ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9994 | 9891 | 9174 | 7348 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9994 | 9777 | 8042 | 4496 | 1653 | 479 |
| Mean of $\alpha_1$ | -1.03 | -1.02 | -0.92 | -0.77 | -0.60 | -0.48 | -0.36 | -0.26 | -0.18 | -0.12 | -0.07 |
| $\geq 1\Delta X_t$ Significant | 1451 | 1424 | 1470 | 1535 | 1580 | 1564 | 1591 | 1564 | 1546 | 1523 | 1559 |
| $\geq 1X_{t\text{-}1}$ Significant | 1454 | 1388 | 2454 | 3902 | 4877 | 5165 | 5230 | 4962 | 4541 | 4078 | 3903 |
| ECM and $\geq 1\Delta X_t$ Significant | 1451 | 1424 | 1470 | 1535 | 1580 | 1564 | 1591 | 1564 | 1529 | 1406 | 1184 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1451 | 1424 | 1470 | 1535 | 1580 | 1564 | 1560 | 1276 | 730 | 306 | 92 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1454 | 1388 | 2454 | 3902 | 4877 | 5165 | 5230 | 4962 | 4538 | 4012 | 3428 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1454 | 1388 | 2454 | 3902 | 4877 | 5165 | 5201 | 4540 | 2919 | 1270 | 396 |
| **Model 2**⋆⋆ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9979 | 9757 | 8897 | |
| ECM Significant* | | 10000 | 10000 | 10000 | 10000 | 10000 | 9222 | 6749 | 3466 | 1318 | |
| Mean of $\alpha_1$ | | -1.01 | -0.89 | -0.73 | -0.55 | -0.43 | -0.32 | -0.23 | -0.16 | -0.11 | |
| $\geq 1\Delta X_t$ Significant | | 1516 | 1578 | 1731 | 1905 | 1984 | 1936 | 1840 | 1725 | 1627 | |
| $\geq 1X_{t\text{-}1}$ Significant | | 1495 | 1574 | 1989 | 2673 | 3222 | 3572 | 3694 | 3673 | 3666 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1516 | 1578 | 1731 | 1905 | 1984 | 1936 | 1837 | 1685 | 1460 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1516 | 1578 | 1731 | 1905 | 1973 | 1796 | 1260 | 605 | 227 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | | 1495 | 1574 | 1989 | 2673 | 3222 | 3572 | 3694 | 3668 | 3572 | |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | | 1495 | 1574 | 1989 | 2673 | 3220 | 3508 | 3164 | 2025 | 962 | |
| **Model 3**⋆⋆⋆ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9931 | 9086 | 6874 | 4420 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9877 | 8177 | 4171 | 1177 | 208 | 32 |
| Mean of $\alpha_1$ | -1.01 | -1.01 | -0.88 | -0.71 | -0.53 | -0.39 | -0.27 | -0.18 | -0.11 | -0.06 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 1452 | 1441 | 1445 | 1456 | 1458 | 1437 | 1444 | 1441 | 1458 | 1455 | 1467 |
| $\geq 1X_{t\text{-}1}$ Significant | 1480 | 1414 | 1417 | 1322 | 1198 | 1076 | 1006 | 1049 | 1114 | 1256 | 1581 |
| ECM and $\geq 1\Delta X_t$ Significant | 1452 | 1441 | 1445 | 1456 | 1458 | 1437 | 1444 | 1425 | 1304 | 1019 | 659 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1452 | 1441 | 1445 | 1456 | 1458 | 1413 | 1172 | 618 | 189 | 42 | 9 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant | 1480 | 1414 | 1417 | 1322 | 1198 | 1076 | 1006 | 1040 | 1020 | 902 | 746 |
| ECM and $\geq 1X_{t\text{-}1}$ Significant* | 1480 | 1414 | 1417 | 1322 | 1198 | 1068 | 890 | 539 | 178 | 51 | 6 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -3.78.
⋆ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
⋆⋆ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
⋆⋆⋆ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.56: Summary Statistics of Multivariate (4IVs) ECM Model by Fractionally Integrated DV and IV (T=150)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**⋆ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9996 | 9930 | 9412 | 7931 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9988 | 9731 | 8029 | 4497 | 1722 | 503 |
| Mean of $\alpha_1$ | -1.03 | -1.03 | -0.93 | -0.78 | -0.62 | -0.50 | -0.39 | -0.29 | -0.20 | -0.14 | -0.09 |
| $\geq 1\Delta X_t$ Significant | 1910 | 1998 | 2024 | 2055 | 2101 | 2144 | 2163 | 2135 | 2152 | 2148 | 2066 |
| $\geq 1X_{t-1}$ Significant | 1751 | 1779 | 2860 | 4233 | 5292 | 5724 | 5795 | 5597 | 5240 | 4840 | 4678 |
| ECM and $\geq 1\Delta X_t$ Significant | 1910 | 1998 | 2055 | 2101 | 2144 | 2163 | 2134 | 2134 | 2137 | 2035 | 1659 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1910 | 1998 | 2055 | 2101 | 2144 | 2119 | 1781 | 1781 | 1062 | 429 | 129 |
| ECM and $\geq 1X_{t-1}$ Significant | 1751 | 1779 | 2860 | 4233 | 5292 | 5724 | 5795 | 5597 | 5239 | 4779 | 4258 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1751 | 1779 | 2860 | 4233 | 5292 | 5724 | 5754 | 5095 | 3287 | 1438 | 449 |
| **Model 2**⋆⋆ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9987 | 9819 | 9198 | |
| ECM Significant* | | 10000 | 10000 | 10000 | 9998 | 9917 | 8976 | 6400 | 3304 | 1346 | |
| Mean of $\alpha_1$ | | -1.00 | -0.89 | -0.73 | -0.56 | -0.44 | -0.33 | -0.24 | -0.18 | -0.12 | |
| $\geq 1\Delta X_t$ Significant | | 1930 | 2003 | 2176 | 2396 | 2505 | 2441 | 2326 | 2206 | 2093 | |
| $\geq 1X_{t-1}$ Significant | | 1882 | 1981 | 2465 | 3291 | 3866 | 4268 | 4385 | 4361 | 4421 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 1930 | 2003 | 2176 | 2396 | 2505 | 2441 | 2322 | 2167 | 1935 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 1930 | 2003 | 2176 | 2395 | 2483 | 2213 | 1537 | 762 | 325 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 1882 | 1981 | 2465 | 3291 | 3866 | 4268 | 4385 | 4352 | 4325 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 1882 | 1981 | 2465 | 3291 | 3861 | 4129 | 3549 | 2194 | 1065 | |
| **Model 3**⋆⋆⋆ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9925 | 9035 | 6806 | 4371 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 9996 | 9694 | 7287 | 3234 | 771 | 135 | 15 |
| Mean of $\alpha_1$ | -1.01 | -1.01 | -0.89 | -0.71 | -0.52 | -0.39 | -0.27 | -0.18 | -0.11 | -0.06 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 1946 | 1964 | 1955 | 1929 | 1895 | 1888 | 1906 | 1934 | 1924 | 1955 | 1917 |
| $\geq 1X_{t-1}$ Significant | 1956 | 1894 | 1862 | 1754 | 1592 | 1492 | 1445 | 1446 | 1527 | 1678 | 2018 |
| ECM and $\geq 1\Delta X_t$ Significant | 1946 | 1964 | 1955 | 1929 | 1895 | 1888 | 1906 | 1918 | 1729 | 1351 | 865 |
| ECM and $\geq 1\Delta X_t$ Significant* | 1946 | 1964 | 1955 | 1929 | 1895 | 1827 | 1397 | 657 | 174 | 34 | 5 |
| ECM and $\geq 1X_{t-1}$ Significant | 1956 | 1894 | 1862 | 1754 | 1592 | 1492 | 1445 | 1437 | 1405 | 1211 | 961 |
| ECM and $\geq 1X_{t-1}$ Significant* | 1956 | 1894 | 1862 | 1754 | 1592 | 1469 | 1187 | 612 | 172 | 41 | 7 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.003.
⋆ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
⋆⋆ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
⋆⋆⋆ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

Table C.57: Summary Statistics of Multivariate (5IVs) ECM Model by Fractionally Integrated DV and IV (T=150)

| | $I(0)$ | $d=.1$ | $d=.2$ | $d=.3$ | $d=.4$ | $d=.5$ | $d=.6$ | $d=.7$ | $d=.8$ | $d=.9$ | $I(1)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model 1**$^{\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9965 | 9630 | 8414 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 10000 | 9993 | 9767 | 8077 | 4502 | 1696 | 505 |
| Mean of $\alpha_1$ | -1.04 | -1.04 | -0.94 | -0.80 | -0.64 | -0.52 | -0.41 | -0.31 | -0.22 | -0.15 | -0.10 |
| $\geq 1\Delta X_t$ Significant | 2361 | 2342 | 2371 | 2474 | 2495 | 2599 | 2603 | 2563 | 2537 | 2515 | 2541 |
| $\geq 1X_{t-1}$ Significant | 2246 | 2239 | 3388 | 4771 | 5787 | 6204 | 6322 | 6186 | 5868 | 5517 | 5427 |
| ECM and $\geq 1\Delta X_t$ Significant | 2361 | 2342 | 2371 | 2474 | 2495 | 2599 | 2603 | 2563 | 2529 | 2433 | 2171 |
| ECM and $\geq 1\Delta X_t$ Significant* | 2361 | 2342 | 2371 | 2474 | 2495 | 2596 | 2539 | 2142 | 1226 | 458 | 173 |
| ECM and $\geq 1X_{t-1}$ Significant | 2246 | 2239 | 3388 | 4771 | 5787 | 6204 | 6322 | 6186 | 5865 | 5467 | 5008 |
| ECM and $\geq 1X_{t-1}$ Significant* | 2246 | 2239 | 3388 | 4771 | 5787 | 6203 | 6283 | 5538 | 3457 | 1443 | 470 |
| **Model 2**$^{\star\star}$ | | | | | | | | | | | |
| ECM Significant | | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9988 | 9868 | 9387 | |
| ECM Significant* | | 10000 | 10000 | 10000 | 9999 | 9867 | 8817 | 6042 | 3079 | 1276 | |
| Mean of $\alpha_1$ | | -1.01 | -0.89 | -0.73 | -0.56 | -0.45 | -0.34 | -0.26 | -0.19 | -0.14 | |
| $\geq 1\Delta X_t$ Significant | | 2290 | 2374 | 2573 | 2798 | 2893 | 2870 | 2734 | 2647 | 2515 | |
| $\geq 1X_{t-1}$ Significant | | 2351 | 2434 | 2980 | 3860 | 4454 | 4900 | 5056 | 5074 | 5121 | |
| ECM and $\geq 1\Delta X_t$ Significant | | 2290 | 2374 | 2573 | 2798 | 2893 | 2869 | 2730 | 2069 | 2363 | |
| ECM and $\geq 1\Delta X_t$ Significant* | | 2290 | 2374 | 2573 | 2797 | 2855 | 2531 | 1678 | 859 | 382 | |
| ECM and $\geq 1X_{t-1}$ Significant | | 2351 | 2434 | 2980 | 3860 | 4454 | 4900 | 5056 | 5067 | 5056 | |
| ECM and $\geq 1X_{t-1}$ Significant* | | 2351 | 2434 | 2980 | 3860 | 4448 | 4690 | 3863 | 2340 | 1098 | |
| **Model 3**$^{\star\star\star}$ | | | | | | | | | | | |
| ECM Significant | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9997 | 9805 | 9015 | 6843 | 4342 |
| ECM Significant* | 10000 | 10000 | 10000 | 10000 | 9992 | 9486 | 6483 | 2479 | 487 | 77 | 14 |
| Mean of $\alpha_1$ | -1.01 | -1.01 | -0.88 | -0.71 | -0.52 | -0.39 | -0.27 | -0.18 | -0.11 | -0.06 | -0.04 |
| $\geq 1\Delta X_t$ Significant | 2278 | 2324 | 2311 | 2293 | 2305 | 2293 | 2275 | 2268 | 2260 | 2294 | 2313 |
| $\geq 1X_{t-1}$ Significant | 2363 | 2267 | 2246 | 2144 | 1959 | 1842 | 1729 | 1711 | 1854 | 2071 | 2418 |
| ECM and $\geq 1\Delta X_t$ Significant | 2278 | 2324 | 2311 | 2293 | 2305 | 2293 | 2274 | 2249 | 2029 | 1580 | 1043 |
| ECM and $\geq 1\Delta X_t$ Significant* | 2278 | 2324 | 2311 | 2293 | 2302 | 2169 | 1485 | 578 | 122 | 25 | 5 |
| ECM and $\geq 1X_{t-1}$ Significant | 2363 | 2267 | 2246 | 2144 | 1959 | 1842 | 1728 | 1700 | 1696 | 1472 | 1136 |
| ECM and $\geq 1X_{t-1}$ Significant* | 2363 | 2267 | 2246 | 2144 | 1959 | 1787 | 1284 | 552 | 157 | 27 | 6 |

*Note*: Cell entries are the result of 10,000 simulations for each model at each specified order of $d$. ECM significance ($p \leq 0.05$, one-tail test). Coefficient significance ($p \leq 0.05$, two-tail test) *ECM significance (*$p \leq 0.05$, one-tail test) with MacKinnon CVs: 5% $\leq$ -4.205.
$\star$ Model 1: IVs are integrated $I(1)$ and DV is integrated at $I(d)$ order specified
$\star\star$ Model 2: Both IVs and DV are integrated at the same $I(d)$ order specified
$\star\star\star$ Model 3: IVs are integrated $I(0)$ and DV is integrated at $I(d)$ order specified

## Bibliography

Baillie, R, and George Kapetanios. 2007. Semiparametric Estimation of Long Memory: The Holy Grail or a Poisoned Chalice. In *International Symposium on Financial Econometrics*. pp. 1–2.

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

Bårdsen, Gunnar. 1989. "Estimation of Long Run Coefficients in Error Correction Models." *Oxford Bulletin of Economics and Statistics* 51(3): 345–350.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* pp. 96–129.

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

Byers, David, James Davidson, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19(1): 49–62.

Caporale, Guglielmo Maria, and Luis A Gil-Alana. 2004. "Fractional Cointegration and Tests of Present Value Models." *Review of Financial Economics* 13(3): 245–258.

Caporale, Guglielmo Maria, and Nikitas Pittis. 1999. "Unit Root Testing Using Covariates: Some Theory and Evidence." *Oxford Bulletin of Economics and Statistics* 61(4): 583–595.

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

Davidson, James. 2014. "Time Series Modelling Version 4.44.".

De Boef, Suzanna. 2001. "Modeling equilibrium relationships: Error correction models with strongly autoregressive data." *Political Analysis* 9(1): 78–94.

De Boef, Suzanna, and Jim Granato. 1999. "Testing for Cointegrating Relationships with Near-Integrated Data." *Political Analysis* 8(1): 99–117.

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

Doornik, Jurgen A. 2014. "An Object-Oriented Matrix Programming Language Ox 7.".

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

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

Gil-Alana, Luis A. 2003. "Testing of Fractional Cointegration in Macroeconomic Time Series*." *Oxford Bulletin of Economics and Statistics* 65(4): 517–529.

Granger, Clive WJ, and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

Hansen, Bruce E. 1995. "Rethinking the Univariate Approach to Unit Root Testing." *Econometric Theory* 11(5): 1148–71.

Hendry, David F, and Grayham E Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England." *The Economic Journal* 88(351): 549–563.

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

Kremers, Jeroen JM, Neil R Ericsson, and Juan J Dolado. 1992. "The Power of Cointegration Tests." *Oxford bulletin of economics and statistics* 54(3): 325–348.

MacKinnon, James G. 1994. "Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests." *Journal of Business & Economic Statistics* 12(2): 167–176.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *The Journal of Political Economy* 89(5): 914–927.

Nicolau, Joao. 2002. "Stationary Processes That Look Like Random Walks - The Bounded Random Walk Process in Discrete and Continuous Time." *Econometric Theory* 18(1): 99–118.

Nielsen, Morten Ørregaard, and Per Houmann Frederiksen. 2005. "Finite Sample Comparison of Parametric, Semiparametric, and Wavelet Estimators of Fractional Integration." *Econometric Reviews* 24(4): 405–443.

Phillips, Peter CB. 1991. "Optimal Inference in Cointegrated Systems." *Econometrica: Journal of the Econometric Society* pp. 283–306.

Phillips, Peter CB, and Sam Ouliaris. 1990. "Asymptotic Properties of Residual Based Tests for Cointegration." *Econometrica: Journal of the Econometric Society* pp. 165–193.

Robinson, Peter M. 1995. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of statistics* 23(5): 1630–1661.

Sanchez Urribarri, Raul A, Susanne Schorpp, Kirk A Randazzo, and Donald R Songer. 2011. "Explaining Changes to Rights Litigation: Testing a Multivariate Model in a Comparative Framework." *The Journal of Politics* 73(02): 391–405.

Ura, Joseph Daniel, and Christopher R Ellis. 2012. "Partisan Moods: Polarization and the Dynamics of Mass Party Preferences." *The Journal of Politics* 74(1): 277–291.

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.

Zivot, Eric. 2000. "The Power of Single Equation Tests for Cointegration When the Cointegrating Vector is Prespecified." *Econometric Theory* 16(03): 407–439.