Pergamon

Electoral Studies 19 (2000) 31–48

Electoral Studies

www.elsevier.com/locate/electstud

You must remember this: dealing with long memory in political analyses

Matthew J. Lebo *, Robert W. Walker, Harold D. Clarke

Washington State University, Department of Political Science, Criminal Justice Program, PO Box 644880, Pullman, WA 99164-4880, USA

Abstract

Recent research by Box-Steffensmeier and Smith (Box-Steffensmeier, J.M., Smith, R.M., 1996. The dynamics of aggregate partisanship. American Political Science Review 90, 567–580; Box-Steffensmeier, J.M., Smith, R.M., 1998. Investigating political dynamics using fractional integration methods. American Journal of Political Science 42, 661–689) has alerted scholars to the problems involved in the analysis of fractionally integrated time series. This paper pursues this line of inquiry by compiling evidence on the time series properties of a number of common variables used in political research, including macropartisanship, presidential approval, the monthly and quarterly index of consumer sentiment, percentage liberalism in Supreme Court decision making, and others. In applying a variety of formal tests to these series, we fail to reject hypotheses of random walk or fractionally integrated processes while commonly rejecting hypotheses of stationary behavior. Evidence obtained from point estimates of the fractional differencing parameter, *d*, supports these findings while providing a glimpse into the long-memory characteristics of many political time series. Finally, Monte Carlo studies are performed that demonstrate the likelihood of spurious regressions when researchers fail to account for the fractional dynamics of time series.

Keywords: Fractional integration; Long memorised processes; Time series; Fractional differencing; ARF/MA modeling

In efforts to model political phenomena over time, both past experiences and present conditions are said to influence future behavior (e.g. Clarke et al., 1996: 5–

* Corresponding author. Tel.: +1-509-335-4544; fax: +1-509-335-7990.
E-mail addresses: lebo@wsu.edu (M.J. Lebo), wlkr@troi.cc.rochester.edu (R.W. Walker)

PII: S0261-3794(99)00034-7

6). Political analysts have employed various techniques to study the impact of the past and the present on the future, but have encountered problems in including each of these distinct elements in their models of political relationships through time.

The ability to evaluate models of political time series that incorporate the influence of the distant past, the recent past and the present is profoundly limited when we are confined to the binary choice between stationary and unit-root processes. The recent introduction of fractional integration methods, and the fractional differencing parameter, d, to political scientists, however, provides the opportunity to overcome these limitations (Box-Steffensmeier and Smith 1996, 1998; Green et al., 1998).

Prior to this introduction, researchers' modeling choices were analogous to dealing with a fuel gauge constrained to read either full or empty. This constraint seems inappropriate for most situations. Thus, we allow for a continuum of possibilities (i.e. $1>d>0$) by empirically estimating the long-term relationships rather than assuming them to be full (d=1) or empty (d=0) based on either preliminary test evidence or a desire to maintain simplicity. This continuum improves the prospects for properly specifying models of political phenomena and thereby necessitates the re-examination of previous conclusions regarding the data generating processes (DGPs) of political variables.

While Box-Steffensmeier and Smith (1996, 1998) began this process with a re-examination of the macropartisanship series (and later the Index of Consumer Sentiment), we apply these methods to a broader range of variables. Our intentions are several. First, and most generally, we wish to further familiarize political scientists with the techniques for investigating the fractional dynamics of political time series. Second, consistent with the recommendations of Baillie et al. (1996): 12) and Box-Steffensmeier and Smith (1996)), we wish to demonstrate the importance of using multiple formal indicators when testing for the order of integration as we subject several popular political time series to repeated statistical testing.

Third, in the spirit of Nelson and Plosser's (1982) search for unit roots, we intend to alert researchers to the fact that fractional dynamics are by no means rare and need to be accounted for in their models. Fractional dynamics are found, or are strongly suspected, to characterize many of our series including presidential approval, the monthly and quarterly index of consumer sentiment, macropartisanship, and percentage liberalism in Supreme Court decision making. We initially suspected the commonality of fractional integration following our subjection of each series to tests with null hypotheses of stationary, trend stationary, and unit-root processes. In diagnosing each series based on these tests, we were left with contradictory evidence pointing to the need for estimation of the fractional differencing parameter, d.

Fourth, we intend to grapple with the question of why fractional dynamics are important to analyses of political time series. Using Monte Carlo techniques, these results should alert researchers that, through spurious regressions, ignoring the presence of fractional dynamics leads to an alarmingly high rate of type I errors — falsely rejecting true null hypotheses of no relationship.

This paper proceeds as follows. After a discussion of the characteristics of, and methods of dealing with, fractional dynamics, we test various series to diagnose their properties. We next discuss the relative merits of procedures for estimating the fractional differencing parameter, *d*, and estimate this parameter for each series. We conclude this section offering insights based on the statistical inventory performed. We next present our Monte Carlo evidence and discuss its implications for models that include fractionally integrated variables. We conclude by outlining some important questions for future research that arose during our investigation.

1. Stationary processes, fractional integration, and integer integration

Initially we must address the question of how fractional dynamics arise. Much research in political science and most of the series we examine are aggregate measures of individual-level responses or behaviors. Granger (1980) points out that the very act of aggregating data that measure heterogeneous individual-level behavior produces fractional dynamics in aggregate time series. The essence of Granger's argument is that the aggregate series is generated by different micro-level autoregressive and moving average processes among individuals, thereby introducing fractional dynamics.

Specifically, Granger (1980) examines independent series x_{jt} for j=1…,n individuals that are randomly generated using a first-order autoregressive model. This is shown in Eq. (1).

$$x_{j,t}=\alpha_j x_{j,t-2}+\varepsilon_t \text{ where } \alpha_j\sim\beta(0,1) \text{ and } \varepsilon_{j,t}\sim N(0,\sigma^2) \quad (1)$$

Granger shows that a fractionally integrated time series results from these individually unique autoregressive processes.

Granger (1980) further explains that fractional dynamics also may arise by aggregating variables containing dynamic relationships at the individual-level.[1] An example of this phenomenon can be extrapolated from V.O. Key's final thoughts on partisanship in mass electorates (Key with Cummings, 1966). We might expect macropartisanship to be fractionally integrated because the "stand-patters" and "switchers" represent distinct types of individual-level behavior with differing reactions to changing conditions over time (see also Converse, 1964). Similar patterns can be imagined for numerous other series.

Thus, through aggregating individual-level behavior we are likely to introduce fractional dynamics into time series and are left to identify these dynamics and deal with the consequences. The critical feature of these processes is the degree of memory. While shocks to stationary series die out rather quickly, the fractionally integrated series has longer memory, but not the infinite memory characteristic of a random walk process. Intuitively, the fractionally integrated model allows greater flexibility than the forced exponential decay of shocks in the stationary model or no decay at all in the integrated model. Recognizing the unique behavior of the frac-

[1] In addition to the aggregation results, a series may also contain fractional integration if its disturbance term is fractionally integrated.

tionally-integrated series leads us to examine the threats to statistical inference that arise when researchers fail to account for fractional dynamics.

Building on work begun by Yule (1926), Granger and Newbold (1974) highlight the threats to inference posed by untransformed nonstationary variables. Modeling independent simulated random walks in level form, their study shows that rejection rates of null hypotheses are alarmingly high. Similar problems occur with level-form treatments of near-integrated processes (see DeBoef and Granato, 1997a; DeBoef et al., 1996 and other collaborations of these authors).[2] Nonetheless, if a nonstationary series is differenced the spurious regression problem can be overcome (Granger and Joyeux, 1980: 15).[3] However, there are side effects to this transformation, the principal one being the loss of the long-term dynamics of the original series (Granger and Joyeux, 1980: 15–16). This consequence of first differencing motivates a need for careful investigation of non-integer values of the differencing parameter. Non-integer values imply fractional dynamics.

Formally, we can describe the data generating process for a series as,

$$\phi(L)(1-L)^d x_t=\theta(L)\varepsilon_t \tag{2}$$

The parameter $\phi(L)$ represents stationary autoregressive processes and $\theta(L)$ represents stationary moving average processes. The fractional differencing parameter is d. Different values of d represent distinct dynamics in the series of interest.

A weakly stationary series [d=0 in Eq. (2)] is characterized by mean reversion, finite variance and covariance stationarity. Thus, the data generating process can be modeled using combinations of stationary autoregressive and moving average terms [ARMA (p,0,q) models].[4] In contrast, the random walk, or first-order integrated series is characterized by mean nonstationarity, variance nonstationarity, and covariance nonstationarity. In such a model, where d=1 in Eq. (2), stationary autoregressive and moving average parameters combine with integer differencing to characterize the ARIMA (p,d,q) model.

The essence of our investigation lies in the realization that the space between d=0 and d=1 cannot be safely disregarded. Within this range, time series are mean reverting and fractionally integrated ($0<d<1$), but they have time-dependent variance and are covariance nonstationary in the range $0.5\leq d$ (Baillie, 1996: 22). Thus, the great deal of variability that exists when $0<d<1$ requires explanation. These

[2] Box-Steffensmeier and Smith (1998): fn. 9) cite Madalla to explain that near-integrated processes (local to unity asymptotics) and fractionally-integrated processes are not related. However, our preliminary analysis using Monte Carlo studies shows that it may be more difficult to distinguish between near integrated and fractionally-integrated processes using finite samples of data (d=0.45 and T=100 and 250). We are currently in the process of more rigorously investigating this distinction.

[3] We should always be mindful of the notion that differencing is not a trivial transformation of the data and that the threats to inference arising from overdifferencing data are largely unclear. Research on the implications of differencing in political time series can be found in Lebo (1999) and DeBoef and Granato (1997b).

[4] Time series analysts often characterize ARMA and ARIMA models using the notation (p,d,q). For notational consistency, the p's are equivalent to $\varphi(L)$ in the previous equation while the q's represent the $\theta(L)$.

long-term dynamics, *d*, complement the short-term dynamics of the ARMA structure and lead to more properly specified models of the data generating process.

2. Testing for *I*(*d*) processes

It seems reasonable to question the assumption that DGPs for complex variables would conveniently follow integer integration orders. Despite more complex possibilities, it is necessary to first examine the hypotheses of stationarity and unit-root behavior. If these integer integration orders can be rejected, it becomes necessary to consider the more complex issue of estimating the fractional differencing parameter. We now turn to the formal tests for our collection of political time series.

First, the question of stationarity must be addressed. Though unit root tests are an important preliminary step in the diagnosis of a series, relying on the results of only one such test to decide if a series needs to be differenced is insufficient. Thus, we implement four tests, each of which examines stationarity in a different fashion. We begin with the Dickey-Fuller unit root test (Dickey and Fuller 1979, 1981).

3. Dickey-Fuller test

The Dickey-Fuller test is based on the fully specified and augmented Dickey-Fuller Eq. (3).

$$\Delta\psi_t=\alpha_0+\gamma\psi_{t-1}+\alpha_1 t+\sum_{t=1}^{\pi}\Delta\psi_{t-\pi}+\varepsilon_t \tag{3}$$

The basic Dickey-Fuller equation consists of only the differenced series of interest, lagged values in level form, and the error term. Specifying a random walk with drift model adds the constant term α_0. The test with a deterministic time trend adds $\alpha_2 t$. Dickey and Fuller (1979) propose testing the null hypothesis of a pure random walk or random walk with drift, $\gamma=0$, against the hypothesis that $\gamma<0$. If the error term is autocorrelated, the augmented Dickey-Fuller test (ADF) involves the insertion of lagged values of $\Delta\psi_t$ such that the residuals from the Dickey-Fuller regression are white noise and, thus, inferences about γ are possible.

The Dickey-Fuller results indicate that our series represent a mixed bag of data generating processes.[5] The most striking finding is that in 11 out of 16 cases the results do not allow for a rejection of the random walk hypothesis. The domestic

[5] We used the Dickey-Fuller test with trend as well, but, as will become apparent later, it seems superior to use a test that employs a null hypothesis of trend stationarity rather than the joint F-test called for using the Dickey-Fuller. The results of the Dickey-Fuller tests with trend are available upon request. A version of this paper with more extensive results from the following tests can be found on the World Wide Web at http://www.cas.unt.edu/~rwalker/apsa98.pdf

policy mood and the quarterly and monthly Michigan Index of Consumer Sentiment (and most of the component series) all fail to reject this hypothesis. However, it is important to note that the Dickey-Fuller test has low relative power in the face of fractionally-integrated alternatives (Diebold and Rudebusch, 1991). It may be said that a more appropriate test involves a unit-root null hypothesis and fractionally-integrated processes as the alternative. To this end, we shall briefly discuss the variance ratio test.[6]

4. Variance ratio test

The variance ratio test is another test for unit-root processes but has the advantage of a null hypothesis of a random walk with drift tested against an alternative hypothesis of pure fractional noise (Diebold, 1989).[7] Remembering that the random walk never forgets a random shock, this test relies upon the assumption that the variance at time k should be k times the variance in period 1. Formally, the test statistic, $R(k)$ is given by:

$$R(k)=\frac{k\hat{\sigma}_1^2}{\hat{\sigma}_k^2},\ k=2,3,\ldots,K, \text{ where } \hat{\sigma}_k^2=\frac{1}{(T-k+1)}\sum_{t=k}^{T}(x_t-x_{t-k}-k\hat{\mu})^2, \text{ for } k =1,2,\ldots,K \text{ and } \hat{\mu}=\frac{1}{T}\sum_{t=1}^{T}(x_t-x_{t-1}) \tag{4}$$

In Eq. (4), k is the differencing interval.

The results of the variance ratio tests are similar to the Dickey-Fuller results. As with the Dickey-Fuller results, Stimson's quarterly mood variable and both indices of consumer sentiment fail to reject the null hypothesis of a random walk. Other series, including the two Supreme Court series have confirmed both the alternative hypotheses of stationary processes for the Dickey-Fuller test and pure fractional noise for the variance ratio test. *The variance ratio results and the Dickey-Fuller results suggest that unit roots are quite common among our series*.

As noted above, the low power of the Dickey-Fuller test in the face of fractional alternatives may help to explain this contradiction. Thus, it becomes necessary to employ a test that uses stationary processes as the null hypothesis tested against fractional alternatives. In implementing the tests proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992)KPSS hereafter), we are able to examine the two previous alternative hypotheses that can be contradictory.

---

[6] Diebold (1989) describes two tests, the *R* test and the *J* test (a joint test), and discusses the properties of both. Both because it is easily available and because it tests the one-sided alternative of unit-root versus fractional behavior, we chose the *R* test though the *J* test may be superior for other alternatives, i.e. explosive processes.

[7] Pure fractional noise can be formally described by: $(1\text{-}L)^d x_t=\varepsilon_t$.

5. KPSS test

The KPSS tests (1992) examine two distinct processes in the time series of interest. The test decomposes the series, by assumption, into a deterministic trend, a random walk, and stationary errors (KPSS, 1992: 162).[8] KPSS argue that a time series can be tested for strong mixing properties through a score test, shown in Eq. (5).

$$\eta_{\mu}=T^{-2}\sum_{t=1}^{T}\frac{S^{2}}{s^{2}(l)} \text{ where } S_{t}=\sum_{i=1}^{t}\varepsilon_{i} \tag{5}$$

In Eq. (5), T is the size of the sample, $s^2(l)$ is an estimate of the disturbance variance (see KPSS, 1992: 164), and S_t represents the partial sum of the residuals, ε_i. Using a base equation consisting solely of an intercept and the stationary errors, the first test, η_τ, tests for the strong mixing that is characteristic of stationary series against the alternative of either fractionally-integrated or unit-root processes.[9] Adding the presence of a linear trend to the intercept and errors, the second test, η_μ, tests the null hypothesis of strong mixing against the same alternative hypothesis. Lee and Schmidt (1996) show that the KPSS test is consistent against $I(d)$ alternatives and thus is useful for distinguishing among short- and long-term dynamics in the series of interest. Recently, political scientists have recognized the utility of the KPSS tests (Ostrom and Smith, 1993; Box-Steffensmeier and Smith 1996, 1998).

The KPSS tests (η_μ), using the recommended lag truncation parameter, show all but three of our series failing to reject the null hypothesis of strong mixing. These three series are the quarterly index of consumer sentiment, quarterly personal current economic conditions (PERNOW), and Supreme Court economics cases. All the other variables under examination allow for a rejection of the null hypothesis of stationary mixing processes (0.05 level). *From the KPSS results, stationary processes generate few, if any, of these time series*.

The KPSS test (η_τ) results can be summarized with a resounding rejection of all trend stationary null hypotheses using the recommended lag truncation parameter. In fact, only when the lag truncation equals 8 can a single variable (the policy mood variable) fail to reject the null hypothesis for the 0.05 level. *Trend stationary processes do not describe these time series*.

[8] The determination of the appropriate lag length (l) is discussed at length in KPSS (1992): 169–173). Given the size of most of our samples, we use (l)=4, though alternatives exist.

[9] Lo (1991): 1286) notes that the characteristic long-term dependence in the autocorrelation function is the defining characteristic of fractionally-integrated series (they decay at hyperbolic rather than exponential rates). This remains the case for integer integrated series as well because shocks infinitely persist due to the lack of mean reversion, the lack of decay in the theoretical autocorrelations, the time dependent variance, and the slow decay in the sample correlogram (Enders, 1995: 212). Thus, strong mixing processes are those which die-down quickly (stationary and invertible autoregressive and moving average parameters) rather than persisting over a large number of lags.

6. Formal Estimators of *d*

Results of the above tests, summarized in Table 1, suggest that many of the series have rejected null hypotheses of mean and trend stationary processes on the one hand and of unit-root processes on the other. Thus, the evidence suggests that most of these series are neither stationary nor integrated, though the possibility of $I(1)$ seems more popular. Confirmation of stationarity, integration, or fractional integration must be obtained by direct estimates of the order of integration from either semiparametric estimation (Geweke and Porter-Hudak, 1983; Robinson, 1995), approximate maximum likelihood estimation in the frequency domain (Fox and Taqqu, 1986), or exact maximum likelihood estimation in the time domain (Sowell, 1992). We will briefly detail the advantages and disadvantages of these three approaches to the estimation of the long-memory parameter, d, before turning to estimation issues and our results.[10]

Sowell (1992) explains that approximate maximum likelihood and semiparametric estimation methods have undesirable small sample properties. This is particularly unfortunate given the small size of most political time series. Baillie (1996): 35) documents the numerous studies which cast doubt on semiparametric estimation of d in the frequency domain, in terms of bias and mean square error (e.g. Hurvich and Ray, 1995; Agiakloglou et al., 1993). The only complaints about full maximum

Table 1
Summary of diagnostic results

Test variable	DF/ADF	Var. ratio	KPSS (η_τ)	KPSS (η_μ)	Prognosis
Civ. Rts. and Civ. Libs. Cases (a)	Reject d=1	Reject d=1	No trend	Reject d=0	$1>d>0$
Econ. Cases (a)	Reject d=1	Reject d=1	No trend	d=0	$1>d\geq 0$
Presidential App. (q)	Reject d=1	d=1	No trend	Reject d=0	$1\geq d>0$
Macropartisanship (q)	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
Policy Mood (q)	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
ICS Quarterly	d=1	d=1	No trend	d=0	$1\geq d\geq 0$
PERNOW (q)	d=1	Reject d=1	No trend	d=0	$1\geq d\geq$
PERFU (q)	d=1	Reject d=1	No trend	Reject d=0	$1\geq d>0$
BUSGB (q)	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
BUS5YR (q)	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
BUY (q)	Reject d=1	Reject d=1	No trend	Reject d=0	$1>d>0$
ICS Monthly	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
PERNOW (m)	d=1	Reject d=1	No trend	Reject d=0	$1\geq d>0$
PERFU (m)	d=1	Reject d=1	No trend	Reject d=0	$1\geq d>0$
BUSGB (m)	d=1	d=1	No trend	Reject d=0	$1\geq d>0$
BUS5YR (m)	Reject d=1	Reject d=1	No trend	Reject d=0	$1>d>0$
BUY (m)	d=1	Reject d=1	No trend	Reject d=0	$1\geq d>0$

[10] Baillie (1996) provides an excellent formal review of these techniques in addition to a description of the estimation procedure.

likelihood estimation concern its computationally intensive algorithms, but these pose few problems given recent advances in computing technology.[11]

When implementing this procedure, differencing is recommended prior to the estimation of d to avoid the estimation of the "troublesome intercept parameter" (Baillie, 1996: 39). Furthermore, due to bounded parameter space (d <0.5) and to ensure stationarity, Sowell's procedure is always applied to first differenced data (Box-Steffensmeier and Smith, 1996: fn.28, 1998: fn.14)).[12] Simulation evidence suggests the efficacy of Sowell's procedure (Smith et al. 1993, 1996).

We begin estimating d with models using neither stationary autoregressive nor moving average parameters before moving on to the more delicate specification of the "best" univariate noise model for each of our series. Owing to our interest in different procedures for applied researchers, we initially compare estimators available for OX and RATS.[13]

7. *d* Findings

The results, seen in Table 2, suggest that both Robinson's semiparametric estimator and Sowell's full maximum likelihood technique yield estimates that are quite close to one another. In fact, in most cases, the two estimates lie within a single standard error of each other. Furthermore, for all series the two estimators agree that d does not equal zero and in all but four cases, the two estimators agree on whether or not d equals one. Even in these cases where the estimators disagree, the point estimates do not substantially diverge from one another.

Hypotheses that d=0 can be rejected (one-tailed tests) for **all** of the series we analyze. In fact, the smallest t-ratio indicates more than five standard errors between the point estimate and zero. Using Robinson's estimator, all but three of the series (the quarterly ICS, quarterly long-term business expectations, and the mood variable

[11] The properties of some more recent approaches to semiparamteric estimation have yet to be examined (among these the procedure suggested by Robinson (1995) which we implement). Other approaches to estimation include bootstrap methods (Andersson and Gredenhoff, 1998) and Bayesian techniques (Koop et al., 1995). Semiparametric regression in the time domain may offer greater promise than previous semiparametric approaches (Baillie, 1996: 36), but insofar as the applied researcher is concerned, microcomputer implementations of this time domain estimator are not widely available in common econometrics packages such as S-Plus, RATS, or OX. Sowell's (1992) estimator receives the predominance of positive attention in the literature and is available for both RATS and the PCGive 9.0 companion OX. Smith et al. (1993, 1996) provide support for the conclusion that there are bias tradeoffs using both approximate maximum likelihood and Sowell's estimation procedure.

[12] Differencing prior to estimation of the d parameter is necessary for many estimators of the order of integration (Hamilton, 1994: 449). Both techniques that we employ require first differencing prior to estimation of the fractional differencing parameter, d.

[13] Sowell's (1992) maximum likelihood procedure has been implemented in RATS by David Montgomery (http://ww2.hawaii.edu/~dmontgom) for ARFIMA models of a limited number of autoregressive and moving average parameters. Because it allows for greater flexibility and is in many ways easier to use, we used Sowell's procedure in OX and Robinson's Semi-Parametric estimator in RATS and display the results for comparative purposes.

Table 2
Estimates of d obtained from (0,d,0) models

Point estimates of the order of integration (d)	Robinson's (1995) Gaussian semiparametric estimator			Sowell's (1992) exact maximum likelihood estimator		
Variable name	Estimate of d[a]	T value for d=0	T value for d=1	Estimate of d[b]	T value for d=0	T value for d=1
Civil Rights and Civil Liberties Cases (annual)	0.55 (0.079)	6.957	5.692	0.520 (0.069)	7.536	6.957
Economics Cases (annual)	0.52 (0.079)	6.578	6.072	0.480 (0.083)	5.783	6.265
Macropartisan ship (quarterly)	0.82 (0.066)	12.382	2.718	0.837 (0.066)	12.890	2.529
Presidential Approval (quarterly)	0.89 (0.066)	13.439	1.661	0.928 (0.085)	10.937	0.849
Domestic Policy Mood (quarterly)	1.04 (0.069)	15.143	0.582	1.032 (0.071)	14.535	0.451
Index of Consumer Sentiments (quarterly)	0.91 (0.069)	13.374	1.323	0.874 (0.070)	12.486	1.243
PERNOW	0.8 (0.069)	11.758	2.940	0.699 (0.063)	11.095	4.778
PERFU	0.71 (0.069)	10.435	4.262	0.651 (0.063)	10.333	5.540
BUSGB	0.87 (0.069)	12.786	1.911	0.917 (0.075)	12.226	1.107
BUS5YR	0.90 (0.069)	13.227	1.470	0.871 (0.071)	12.268	1.817
BUY	0.82 (0.069)	12.052	2.645	0.720 (0.067)	10.750	4.179
Index of Consumer Sentiments (monthly)	0.87 (0.057)	15.367	2.296	0.92 (0.059)	15.593	1.356
PERNOW	0.76 (0.057)	13.424	4.239	0.640 (0.046)	13.913	7.826
PERFU	0.67 (0.057)	11.835	5.829	0.595 (0.050)	11.9	8.1
BUSGB	0.86 (0.057)	15.191	2.473	0.889 (0.059)	15.068	1.88
BUS5YR	0.71 (0.057)	12.541	5.122	0.721 (0.053)	13.604	5.264
BUY	0.73 (.057)	12.894	4.770	0.773 (0.055)	14.055	4.127

[a] Estimates of d generated using Robinson's Gaussian Semi-Parametric estimation routine (Robinson, 1995). This procedure (RGSER.SRC) can be acquired for RATS from the Estima web site (http://www.estima.com). These estimates arise from the estimation of (0,1+d,0) on first differenced data because of the constrained parameter space ($-1.5<d<.5$). Thus, the results actually reflect the estimate of d +1.

[b] Estimates of d generated using Sowell's Exact Maximum Likelihood estimator for OX. OX is the matrix language that accompanies PCGive 9.0 and is also available in a stand-alone form through Jurgen Doornik's web site (http://www.nuff.ox.ac.uk/Users/Doornik). These estimates arise from the estimation of (0,1+d,0) on first differenced data because of the constrained parameter space ($-1.5<d<.5$). Thus, the results actually reflect the estimate of d +1.

(Stimson, 1991)) allow for a rejection of the hypothesis that d is equal to one. Using Sowell's estimator, four of the series (the monthly ICS, quarterly short-term business expectations, presidential approval, and again, the mood variable) fail to reject the hypothesis that d is equal to one. *Thus, Robinson's and Sowell's estimators agree that fractional integration is extremely common among political time series.*

One exception to this observation appears to be Stimson's policy mood variable. This variable is estimated to be an explosive process with d equal to 1.04 and 1.023, using Robinson's and Sowell's procedures, respectively. Nevertheless, the standard error bands do not allow for rejection of the hypothesis that *d* is equal to or less than unity.

The series estimated to contain the least memory is the Supreme Court economics cases, with an estimate of *d* approximately equal to 0.5. This highlights the finding that most of the series are covariance nonstationary because *d* is estimated to be greater than 0.5 in all cases but one. Taking the point estimates of *d* as a whole, our analyses demonstrate that fractional dynamics should be given much more attention in the study of political time series. Specifically, analysts make a perilous omission when failing to investigate the fractional dynamics of their time series.

As a final step, we investigated short-term dynamics (autoregressive and moving average parameters) in addition to the estimation of the fractional differencing parameter. In combining the estimation of these short-term and long-term dynamics, we chose to limit our investigation to models containing up to three autoregressive and three moving average parameters per series.[14] We also defined a method for choosing the "best" noise model.[15] We chose the Akaike Information Criteria (Akaike 1973, 1974).[16]

8. Univariate Noise Models (p,d,q)

These results were estimated on a slightly different selection of variables. We chose to analyze only the monthly and quarterly indices of consumer sentiment and not their components.[17] Also, we replicated the analyses of Box-Steffensmeier and Smith (1996) for their series of Republican and Democratic identifiers.[18] Using Sow-

[14] This results in the estimation of sixteen models per variable. Since most of the components in the Index of Consumer Sentiment are not as widely used as their combination, we have chosen to present only the monthly and quarterly indices.

[15] In some cases, models were estimated with standard errors in excess of 1.0. This seems odd and indicates some instability. However, we were comforted by the fact that all of these models were rejected by the Akaike Information Criterion. We note the arguments made by Box-Steffensmeier and Smith (1998): fn. 30) for using the Schwartz Criterion in that it favors more parsimonious models. Using the SC led Box-Steffensmeier and Smith to select the most parameterized model in every case they examined. Using the same data, we found that the AIC selected more parsimonious models. One potential explanation for this curiosity may be found in the sample size calculations in the SC formula. If one does not adjust the sample size, it may contribute to divergent findings, though as Enders (1995): 88 and fn.2) explains, neither adjustment nor failure to adjust can be said to be incorrect.

[16] The formula can be expressed as, AIC = - 2 * ln[maximized~likelihood] +2 *p*, where *p* is the number of independent parameters estimated. This formula is detailed in Chatfield (1996): 226).

[17] It might be interesting to analyze the components of the index and fractionally difference the components before constructing the index. We are not aware of systematic assessments of the consequences for mixing and averaging among variables with different time series properties, especially different degrees of persistence or long memory.

[18] These results are not reported because the estimated models are the same as those reported in Box-Steffensmeier and Smith (1996). The "best" fitting model according to the AIC for both series was a model containing only the fractional differencing parameter.

ell's full maximum likelihood estimator,[19] we found using the AIC as the criterion for selection that: *only the presidential approval series diverges from a (0,d,0) process*.[20] In other words, in all but one case, the fractional differencing parameter is the only parameter required in the "best" noise model. The one exception, presidential approval, is "best" approximated by a (1,d,0) model. In this model, the estimate of d is 0.395 with a standard error of 0.18. Thus, with a t-ratio greater than 2, we can safely reject the hypothesis that d=0. We also can reject the null hypothesis that d=1. The series appears to be best characterized as having long memory, though it is mean and variance stationary. Furthermore, the first-order autoregressive parameter is estimated to be 0.587 with a standard error of 0.17039. Thus, the selected model for presidential approval contains all statistically significant parameters and still allows for rejection of the hypotheses that d=0 or d=1.

9. Monte Carlo results

The above analyses suggest that fractional integration may indeed be a typical characteristic of political time series. Thus, we are interested in studying how the presence of fractional integration can disrupt the political analyst's task. Granger and Newbold (1974) warn analysts of the danger of spurious regressions when the problem of stationarity is not properly addressed. When non-stationary series are included in level form in a bivariate regression, null hypotheses of no relation between the variables will more often be rejected than chance would allow.[21] DeBoef and Granato (1997) used Monte Carlo methods to further demonstrate that the same problem of spurious regressions arises when data are near integrated. To advance this line of inquiry, we investigate whether the problem of spurious regressions is a likely outcome of fractionally integrated data.

There were several stages in our Monte Carlo analyses the first of which proceeded as follows: two series (Y_t, X_t) were randomly generated with the stipulation that d=0.1 for each series.[22] The dependent variable (Y_t) was regressed on a constant and the independent variable and, as with DeBoef and Granato (1997a), the following output was recorded:

[19] Robinson's procedure does not allow for the estimation of short- and long-term dynamics simultaneously. Thus, we are forced to rely on Sowell's estimator for the comparison between (0,d,0) and the ARFIMA models we tested.

[20] We do not report these results in tabular form because only the presidential approval series divereges from the (0,d,0) models reported in Table 2. Thus, these results increase our confidence in the estimates presented in that table.

[21] Granger and Newbold (1974) explained that spurious regressions arise between non-stationary variables because each strays from its own mean level. Similar tendencies are thus present in the data generating process of each of the otherwise unrelated variables and researchers may be lead to surmise that the DGP of one series is explicable by that of the other.

[22] An interesting theoretical treatment of fractionally integrated regressors in the linear regression model can be found in Gourieroux et al. (1989).

Table 3
% $|t|>1.96$ n=100

	*d*2 (X)									n=100
		1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
*d*1 (Y)	1	0.742	0.761	0.693	0.651	0.587	0.513	0.445	0.32	0.237	0.131
	0.9	0.704	0.715	0.704	0.629	0.584	0.52	0.409	0.333	0.231	0.108
	0.8	0.691	0.691	0.65	0.622	0.547	0.478	0.384	0.317	0.22	0.124
	0.7	0.672	0.642	0.623	0.569	0.529	0.458	0.331	0.268	0.18	0.099
	0.6	0.604	0.583	0.568	0.506	0.441	0.4	0.302	0.219	0.152	0.084
	0.5	0.558	0.493	0.453	0.416	0.387	0.301	0.26	0.182	0.134	0.084
	0.4	0.413	0.422	0.378	0.345	0.306	0.241	0.201	0.173	0.107	0.08
	0.3	0.337	0.328	0.302	0.287	0.22	0.202	0.16	0.135	0.071	0.054
	0.2	0.229	0.241	0.208	0.189	0.167	0.125	0.121	0.077	0.07	0.063
	0.1	0.137	0.116	0.112	0.124	0.109	0.097	0.067	0.067	0.065	0.056

Statistic Displayed
a) coefficient point estimates (combined as percentage rejections of two-
b) standard errors tailed t tests with $p<0.05$ shown in Table 3)
c) R-square (Extended document: via the Internet)
d) the Durbin-Watson statistic (Table 4)

This process was repeated 1000 times with a sample size of 100.[23] Then d_x was moved incrementally upward where d_x = (0.1, 0.2, 0.3,…,1.0) and another thousand

Table 4
Durbin-Watson d statistic n=100

	*d*2 (X)									n=100
		1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
*d*1 (Y)	1	0.18	0.18	0.19	0.19	0.19	0.19	0.18	0.16	0.15	0.13
	0.9	0.25	0.25	0.25	0.26	0.26	0.26	0.25	0.23	0.21	0.2
	0.8	0.35	0.37	0.36	0.36	0.37	0.34	0.33	0.33	0.31	0.3
	0.7	0.51	0.5	0.51	0.5	0.5	0.49	0.48	0.46	0.46	0.43
	0.6	0.7	0.69	0.69	0.2	0.68	0.67	0.67	0.64	0.63	0.63
	0.5	0.93	0.92	0.91	0.91	0.9	0.89	0.88	0.86	0.86	0.86
	0.4	1.17	1.16	1.16	1.15	1.14	1.12	1.12	1.12	1.11	1.11
	0.3	1.4	1.4	1.39	1.39	1.38	1.39	1.39	1.36	1.37	1.36
	0.2	1.64	1.63	1.64	1.62	1.62	1.62	1.6	1.6	1.6	1.6
	0.1	1.83	1.83	1.83	1.85	1.82	1.82	1.82	1.82	1.82	1.81

[23] Sample sizes of 100 were chosen because it is a typical size for a political variable. All of our analyses were repeated with n=250 with similar results. The results of the n=250 simulations are available upon request.

regressions were performed for each value of d_x. After 10 000 regressions, d_y was set to equal 0.2 (then 0.3, 0.4,…,1.0) and the first two loops were repeated.[24] The resulting data sets contains 100 000 observations for each of the four statistics.

Our results are reminiscent of those of DeBoef and Granato (1997a) and indicate the high likelihood that problems will arise when fractionally integrated data is not properly dealt with. As d moves away from zero on either side of the equation the false rejection rate of the null hypothesis ($\beta_x=0$) increases dramatically with a peak of 761 rejections for 1000 regressions when the independent variable is fractionally integrated ($d_x=0.9$) and the dependent variable is wholly integrated ($d_y=1.0$).[25] Even in cases where both variables are mean and covariance stationary, rejection of the null hypothesis ($\beta_x=0$) can be four times more likely than chance would allow ($p=0.05$ for $d_y=0.4$ and $d_x=0.4$ in Table 4). *Thus it is evident that the problem of spuriousness plagues regressions involving fractionally integrated data and researchers must address this problem before deriving generalizations from their models.*

10. Testing a first order autoregressive model

One suggested panacea for dealing with these problems is the inclusion of a lagged endogenous variable. To simplify their analyses researchers may be tempted to ignore the possibility of fractionally or wholly integrated data and revert to a simple first-order autoregressive [AR(1)] process. In order to assess the utility of this approach, we replicated our Monte Carlo analysis using Eq. (6).

$$y_t=\alpha+\beta_1x_t+\beta_2y_{t-1}+\varepsilon_t \tag{6}$$

From this equation, we performed another 100 000 regressions and saved the coefficients, standard errors, Durbin's h, and r-square. The rejection rate of $\beta_1=0$ at the 0.05 level of probability is shown in Table 5.

While less pronounced than those without the lagged endogenous variable, these results show that the inclusion of a lagged endogenous variable is an easy yet inadequate solution. In fact, when d_1 is equal to 0.6 and d_2 is equal to 1.0, the true null

[24] To perform the simulations we used RATS econometric software (Doan, 1992) and the data simulation procedure (ARFSIM.SRC) developed by Rob Schoen and made available through the Estima Corporation. ARFSIM.SRC is based on the method proposed by Davies and Harte (1987) and described in Beran (1994): 215–217). Randomization of simulations and the random number generators invoked are described at length in the manuals which accompany RATS (Doan, 1992). The simulations ran for approximately seventy-two hours each on a Pentium 120. The code necessary to invoke these simulations is available from the authors upon request, including the seed values for the experiments so that replication becomes possible. We also experimented with altering the seed values to ensure that this essentially random selection process did not itself influence the results.

[25] Overall, the results of our $n=250$ simulations are even more startling with rejection rates as high as 85% in some cases. Thus, as DeBoef and Granato (1997a) found for near-integrated series, larger sample sizes do not seem to alleviate the problem of spurious regressions for fractionally integrated data (see also Lebo, 1999).

Table 5
$\%|t| > 1.96$ n=100

	*d*2 (X)									n=100	
		1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
*d*1 (Y)	1	0.184	0.187	0.226	0.299	0.327	0.365	0.341	0.289	0.224	0.134
	0.9	0.156	0.162	0.239	0.28	0.328	0.346	0.292	0.26	0.198	0.114
	0.8	0.15	0.132	0.183	0.229	0.281	0.278	0.284	0.237	0.157	0.125
	0.7	0.111	0.089	0.131	0.177	0.207	0.232	0.242	0.205	0.145	0.11
	0.6	0.082	0.084	0.089	0.139	0.183	0.212	0.172	0.18	0.136	0.09
	0.5	0.089	0.065	0.07	0.092	0.13	0.155	0.138	0.155	0.107	0.086
	0.4	0.081	0.058	0.072	0.067	0.099	0.097	0.102	0.085	0.092	0.083
	0.3	0.054	0.065	0.063	0.054	0.054	0.085	0.076	0.071	0.067	0.063
	0.2	0.058	0.05	0.054	0.046	0.046	0.063	0.061	0.064	0.065	0.051
	0.1	0.044	0.049	0.039	0.063	0.061	0.075	0.055	0.065	0.053	0.053

hypothesis was rejected almost one-quarter of the time. Thus, even with the inclusion of an AR(1), rejection of the null hypothesis can be five times more likely when regressions are performed with two fractionally integrated variables than would occur if d=0 for both X and Y.

To ensure the accuracy of the AR(1) simulations, we implemented the following test.[26] At the 0.05 level of probability, random chance would suggest 50 rejections for each cell in Table 5. In 71 of the 100 cells, the number of rejections is at least two standard errors above 50. This result buttresses our simulation results and again highlights the serious threats to inference that arise from a failure to diagnose long-memory processes.

Certainly, our Monte Carlo evidence combined with evidence of the commonality of fractional integration in political time series does not bode well for models including these complex processes. It becomes clear that researchers must be careful to identify and properly deal with long and short term dependence in their time series of interest before including them in more complex multivariate models.

11. Conclusions

Scholars should employ a variety of formal tests that include estimation of the d parameter to investigate the data generating process of a variable before its inclusion in regression models. We set out to identify the univariate properties of several

[26] We applied two additional tests to confirm these findings. We applied a difference of means test between the rejection rate of our Monte Carlo evidence and the expected value of 0.05 given by statistical theory. We can safely reject the notion that our findings are within the realm of random chance. In fact, the mean of the rejection rates is a whopping 57.8 standard errors from 0.05 where n=100 000. Finally, a t-test was performed using the cells of Table 5 (the rejection rates). Again the rejection rates are substantially higher than 5%, with t=10.9 and n=100.

popular political time series. What becomes clear is that the choice is not a simple binary one between level stationarity and first-order integration. For none of the series we analyze can fractional integration be rejected. In fact, over this broad range of series it is clear that fractional dynamics are quite common. In the final section, we established the very serious threats to time series analysis that result from ignoring the presence of fractional integration.

Although we find evidence of fractional integration in all of the series we consider, rival methods for modeling long-memory processes are receiving attention among political methodologists. An alternative to the ARFIMA processes we describe are the near-integrated processes studied by DeBoef and Granato (1997a). Existing evidence suggests that near-integrated series and fractionally-integrated series may not be easy to differentiate with existing diagnostic tools. In fact, our preliminary simulation results suggest that it is very easy to misdiagnose the dynamic properties of a series. Thus, we must ask whether “near-integrated” series, and if so, of what length, can be clearly distinguished from fractionally-integrated alternatives using the methods we have employed here.

Until recently, long-term processes have been largely neglected because of an obsession with autoregressive and moving average parameters. The time has come to bring new rigor to analyses of the order of integration, and thereby achieve a better balance between investigations of short- and long-term dynamics. We have also demonstrated that long memory cannot be safely explained away with the simple insertion of past values of the dependent variable. From the evidence presented, it is clear that models of political relationships must more closely examine the temporal dynamics not only between their variables, but within them as well.

Acknowledgements

We thank Michael Ault for computer time and Eric L. Krueger for helpful comments, Jim Stimson, Don Green, Jan Box-Steffensmeier, Renée Smith, and colleagues at the University of North Texas for sharing their data. A previous version of this study was presented at the 1998 Annual Meeting of the Midwest Political Science Association (under the title “The Long and the Short of It: Stationarity and Memory in Political Analysis”). More detailed reports are available upon request.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19, 716–723.

Agiakloglou, C., Newbold, P., Wohar, M., 1993. Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis 14, 235–246.

Andersson, M.K., Gredenhoff, M.P., 1998. Robust testing for fractional integration using the bootstrap. Stockholm School of Economics Working Papers Series in Economics and Finance No. 218, January.
Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5–59.
Baillie, R.T., Chung, C.F., Tieslau, M.A., 1996. Analyzing inflation by the fractionally integrated ARFIMA-GARCH model. Journal of Applied Econometrics 11, 23–40.
Beran, J., 1994. Statistics for long memory processes. Chapman Hall, New York, NY.
Box-Steffensmeier, J.M., Smith, R.M., 1996. The dynamics of aggregate partisanship. American Political Science Review 90, 567–580.
Box-Steffensmeier, J.M., Smith, R.M., 1998. Investigating political dynamics using fractional integration methods. American Journal of Political Science 42, 661–689.
Chatfield, C., 1996. The Analysis of Time Series: An Introduction. 5th ed. Chapman and Hall Texts in Statistical Science, London, UK.
Clarke, H.D., Norpoth, H., Whiteley, P., 1996. Modelling Political and Social Dynamics: It's About Time. Typescript.
Converse, P.E., 1964. The nature of belief systems in mass publics. In: Apter, D. (Ed.), Ideology and Discontent. Free Press, New York, NY.
Davies, R.B., Harte, D.S., 1987. Tests for Hurst effect. Biometrika 74, 95–102.
DeBoef, S., Granato, J., 1997a. Near-integrated data and the analysis of political relationships. American Journal of Political Science 41, 619–640.
DeBoef, S., Granato, J., 1997b. Overdifferencing: implications for the time series study of political relationships. Typescript.
DeBoef, S., Baillie, R.T., Bange, M., Granato, J., Dasgupta, S., 1996. Long and short run memory. Presented at the annual meeting of the Midwest Political Science Association, Chicago, Illinois.
Dickey, D.A., Fuller, W., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 84, 427–431.
Dickey, D.A., Fuller, W., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49, 1057–1072.
Diebold, F.X., 1989. Random walks versus fractional integration: power comparisons of scalar and joint tests of the variance-time function. In: Raj, R. (Ed.), Advances in Econometric Modeling. Kluwer Academic Publishers, New York.
Diebold, F.X., Rudebusch, G.D., 1991. On the power of Dickey-Fuller tests against fractional alternatives. Economics Letters 35, 155–160.
Doan, T., 1992. RATS User's Manual, Version 4. Estima Corporation, Evanston, IL .
Enders, W., 1995. Applied Econometric Time Series. John Wiley and Sons, New York, NY.
Fox, R., Taqqu, M.S., 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. Annals of Statistics 14, 517–532.
Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4, 221–238.
Gourieroux, C., Maurel, F., Monfort, A., 1989. Least squares and fractionally integrated regressors. Working paper 8913:Institut National de la Statistique et des Etudes Economiques.
Granger, C.W.J., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227–238.
Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series 1, 15–29.
Granger, C.W.J., Newbold, P., 1974. Spurious regression in econometrics. Journal of Econometrics 2, 111–120.
Green, D.P., Palmquist, B.L., Schickler, E., 1998. Macropartisanship: a replication and critique. American Political Science Review 92 (4), 883–899.
Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.
Hurvich, C.M., Ray, B.K., 1995. Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. Journal of Time Series Analysis 16, 17–42.
Key, V.O., Cummings, M.C., 1966. The responsible electorate: rationality in presidential voting, 1936–1960. Vintage Books, New York, NY.

Koop, G., Ley, E., Osiewalski, J., Steel, M.F.J., 1995. Bayesian analysis of long memory and persistence using ARFIMA models. Journal of Econometrics: Forthcoming.
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54, 159–178.
Lebo, M.J., 1999. Fractional integration and political modeling. Doctoral dissertation, University of North Texas.
Lee, D., Schmidt, P., 1996. On the power of the KPSS test of stationarity against fractionally integrated alternatives. Journal of Econometrics 73, 285–303.
Lo, A.W., 1991. Long term memory in stock market prices. Econometrica 59, 1279–1313.
Nelson, C., Plosser, C., 1982. Trends and random walks in macroeconomic time series: some evidence and implications. Journal of Monetary Economics 10, 130–162.
Ostrom, C.W. Jr., Smith, R., 1993. Error correction, attitude persistence, and executive rewards and punishments: a behavioral theory of presidential approval. Political Analysis 4, 127–184.
Robinson, P.M., 1995. Gaussian semi-parametric estimation of long range dependence. Annals of Statistics 23, 1630–1661.
Sowell, F.B., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics 53, 165–188.
Smith, A.A. Jr, Sowell, R.B., Zin, S.E., 1993. Fractional integration with drift: estimation in small samples. Typescript.
Smith, A.A., Jr., Sowell, F.B., Zin, S.E., 1996. Fractional Integration with Drift: Estimation in Small Samples. Typescript (Revised July 1996).
Stimson, J.A., 1991. Public opinion in America: moods, cycles, and swings. Westview Press, Boulder, CO.
Yule, G.U., 1926. Why do we sometimes get nonsense correlations between time-series? A study in sampling and the nature of time series (with discussion). Journal of the Royal Statistical Society 89, 1–64.