Research Note: Can Time Series Methods Be Used to Detect Path Dependence? *

John Freeman

June 14, 2011

That path dependence is a key feature of human systems now is well-recognized by students of politics as well as by other social scientists. Progress has been made in clarifying the concept of path dependence and related ideas like equilibrium dependence (Page 2006). Thanks to the work of historical sociologists the importance of initial conditions–"contingent events"–is clearer now as well (Mahoney 2000). But, with some notable exceptions (Jackson and Kollman 2010), we don't know how (if) path dependence is manifest in data. The empirics in much of this genre amount to analyses of ball-urn models and to narratives about historical episodes. These studies provide little guidance about how (if) certain statistical results connote path dependence. For example, a well-established argument in political science is that macropartisanship is a "running tally" of political shocks. Leading scholars like Erikson, MacKuen and Stimson (1998: 904-5) argue that an individual's equilibrium partisanship is a random walk. But,even if this can be established statistically, is a random walk evidence of path dependence in macropartisanship? How so? [1]

Time series analysis expressly focuses on historical dependence–on "unpacking historical causality." It is rooted in a dynamic systems framework.[2] What does time series analysis teach us about path dependency? Do some time series methods include tests for path dependency? Do the unit root tests performed by Erikson, MacKuen and Stimson, for instance, constitute such a test? Or is the concept of path dependence more complex? Are conventional time series methods inadequate for the study of path dependency in macropartisanship and related features of American political dynamics?

This note answers these questions. It shows that *linear* time series models illuminate early and outcome path (and phat) dependence but not equilibrium dependence. Moreover, familiar tools like unit root tests reveal distinct data generating processes, processes that

---

*Prepared for the NSF sponsored conference on path dependence, University of Minnesota, June 4-5, 2010, Second draft. Comments welcome. Please do not quote without author's permission

[1]Illustrative of the debate about the nature of partisanship at the individual and macro level is the exchange between Green et al 1998 and Erikson et al 1998. The former (1998: 886-7) argue that, in fact, at the macro level the statistical evidence is at best equivocal. More generally, Page (2006, 97, fn. 8; 98, 104) notes that different econometric methods are needed to distinguish phat from path dependency and also that time series regression models imply historical dependence. But he does not explain which econometric methods are most useful or how (if) time series regression is best used to establish path dependency.

[2]The phrase "unpacking historical causality" is from Page (2006). Page investigates path dependency both in dynamic systems and decision theoretic frameworks. For brevity, I focus here only on the former framework.

embody phat as opposed to alternative kinds of dependence. At the same time, there are concepts in the study of linear time series models that have no clear parallel in the path dependency literature, in particular, the idea of a (correction to) moving equilibrium in phat dependent processes (cointegration). *Nonlinear* time series analysis also is useful. For example, threshold autoregressive (TAR) models connote both path dependence and equilibrium dependence. The notion of switches between different paths of adjustment to a moving equilibrium in phat dependent processes is suggested by nonlinear error correction models. In sum, existing time series methods give us the tools we need to study a wide range of concepts associated with the idea of path dependence. Analysts simply must be clear about which concepts are embodied in each model.

The discussion is organized in terms of the linear-nonlinear distinction. For simplicity, I focus on the likelihoodist tradition in time series analysis.[3]

1 Conceptualization

At least three conceptual distinctions are important in assessing the applicability of time series methods to the study of path dependence. The first is the idea of "initial conditions." Historical sociologists such as Mahoney (2000) and Goldstone (1998) define path dependence in terms of the impact of initial conditions on current outcomes; "[path dependent] outcomes are related stochastically to initial conditions" (*Ibid. p. 834*). Page (2006: 103ff) distinguishes early path dependence from late path dependence. His Founder Process–an ball-urn procedure in which one of two balls is chosen, replaced, and then the other ball is removed–illustrates early path dependence. In time series analysis one usually assumes the initial condition is known. As I show below, this initial condition is important in assessing the early path and equilibrium dependence of integrated processes. [4]

Second are the concepts of phat, path, and equilibrium dependence. Roughly speaking, phat dependence connotes dependence on the *set* of previous events as distinct from path dependence which connotes dependence on the *order* of past events. The Polya Process and Page's (2006: 103) Burden of History process illustrate these two kinds of dependence, respectively. Equilibrium dependence is the idea that the long-run distribution of over possible outcomes depends on the history of past outcomes. Path dependent processes may or may not be equilibrium dependent (*Ibid.*). A related, important, concept here is ergodicity. In the dynamic systems framework this idea has to do with the possibility that, through a series of states, one can move from one state to any other state. In time series analysis, ergodicity is conceived in terms of the convergence of time and ensemble averages. Time averages are calculated for the single observed sequence of observations of a variable whereas ensemble averages correspond to the idea of "rerunning history" multiple times and finding the average value of a variable at each time point. For example, a covariance stationary process is said to be ergodic for the mean if the sample estimate for the mean converges in probability to the expected value of that variable as $T \to \infty$. [5]

[3] I thank Jeff Gill for explaining the relative virtues of this term rather than "frequentist".

[4] Enders (2010: 11) shows how to solve a stochastic difference equation when the initial condition is not known.

[5] Page (2006: 95) defines ergodicity in terms of state dependence, the possibility of writing a mapping

Model	Impact of (Initial) "Contingent" Events	Phat Dependence	Path Dependence	Equilibrium Dependence
Univariate Linear Models:				
Stationary ARMA	no	no	yes	no
Fractionally Integrated	no	no	yes	no
Nonstationary (Random Walk)	yes	yes	no	no
Multivariate Linear Models:				
Single Equation Time Series Regression	no	no	yes	yes
Singe Equation ECM	yes	yes	no	yes(coint)
Stationary VAR	no	no	yes	no
VECM (Less than full rank)	yes	yes	yes	yes(coint)
Nonlinear Time Series Models:				
Threshold Autoregressive	no	no	yes	yes
Nonlinear Error Correction	yes	yes	no	yes(coint)
NLS Regression	no	no	yes	yes

Table 1: Time Series Models, Path Dependency and Related Concepts. Note. Coint denotes cointegration.

2 Linear Time Series Models

With one exception, linear time series models embody all the concepts above except for equilibrium dependence. Linear time series models also illuminate an idea that apparently is not captured by the writing on dynamic systems or historical sociology: common trends in

of each history into one of N states. In a state dependent process the outcome in any period depends only upon the state of the process at that time. Its state transition rule is the same in every time period. A state dependent process is ergodic if through some series of states it is possible to get from one state to any other. The Ergodic Theorem stipulates that a stationary, ergodic state-dependent process generates a unique equilibrium distribution over outcomes. In the time series literature (Hamilton 1994: Chapter 3) ergodicity is defined in terms of ensemble and time averages. An ensemble is a set of realizations of a time series denoted by $[y_1^{(i)}, y_2^{(i)}, \ldots, y_T^{(i)}]$ where i denotes each realization of the *entire* series. The time average, for the first of these realizations is $\bar{y} \equiv (\frac{1}{T}) \sum_{t=1}^{T} y_t^{(1)}$. Let the expected value of the ensemble conception of the series be denoted $E(Y_t)$ A covariance stationary process is ergodic for the mean if this equation for $\bar{y}$ converges in probability to $E(Y_t)$ as $T \rightarrow \infty$. (More detailed conditions for ergodicity are derived in Hamilton's chapter on asymptotics for time series.)

phat dependent processes (cointegration). Vector error correction models can have multiple moving equilibria (cointegating vectors) composed of phat dependent processes (variables).

2.1 Univariate Linear Time Series Models

Consider the simple first order autoregressive model with constant coefficients. This model can be written

$$y_t = a_0 + a_1 y_{t-1} + \epsilon_t \tag{1}$$

where a_0, a_1 are constants, and ϵ_t is a white noise process. This is a stochastic difference equation. It can be solved in several ways. The solution is

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \epsilon_{t-i} \tag{2}$$

.

This solution manifests path dependence insofar as the initial condition has an impact on the current value of y_t as does the the *sequence* of shocks which are *weighted* by a_1 raised to different powers of t. But does this simple linear univariate model connote equilibrium dependence? For the stationary case in which $|a_1| < 1$ the answer is no. In this case, as $t \rightarrow \infty$ we have

$$\lim y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \epsilon_{t-i} \tag{3}$$

Taking expectations of both sides of this equation, we obtain $E(y_t) = \frac{a_0}{1-a_1}$, a finite and time independent value. It is easy to show that the variance also is finite and time independent (Enders, 2010: 55-56). So the equilibrium outcome does not depend on initial condition, set, or order of shocks. For $|a_1| < 1$ this process therefore is path dependent but not equilibrium dependent. This conception of macropartisanship is closest to those advanced by Green et al (1998) and by Box-Steffensmeier and Smith (1996). Conventional Box-Jenkins methods could be used to identify and estimate a process that is path dependent in this sense. Tests for fractional integration can be used for the same purpose. [6]

The random walk model in contrast is early path dependent and phat dependent. But it too is not equilibrium dependent. This model–which is the most simple version of the "running tally" thesis for macropartisanship–can be written as

$$y_t = y_{t-1} + \epsilon_t. \tag{4}$$

Its solution is simply

$$y_t = y_0 + \sum_{i=1}^{t} \epsilon_{t-i}. \tag{5}$$

[6] Box-Steffensemeier and Smith (1996) actually argue that macropartisanship is fractionally integrated. This too implies path dependency in the sense that the sequence of shocks affects the current value of the series. But due to the stationarity of fractionally integrated systems, they too embody path but not equilibrium dependence.

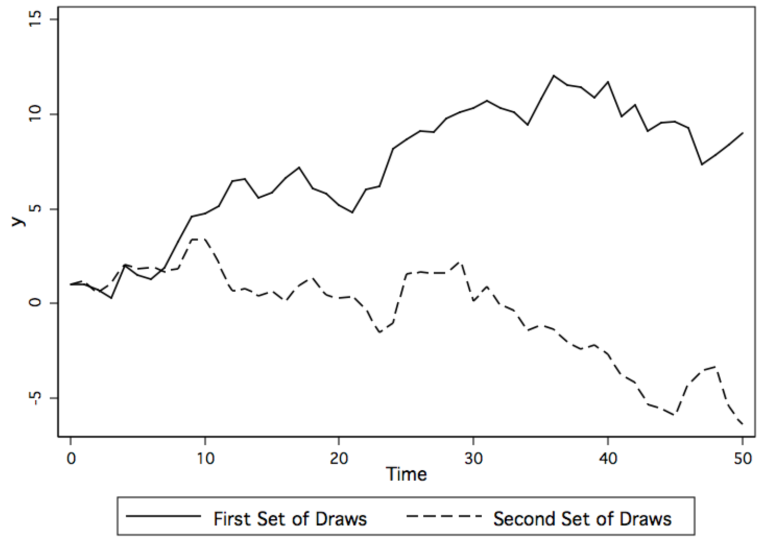

Figure 1: Two Realizations of a Random Walk Procees (A two realization ensemble)

This solution is different from the solution for the simple autoregressive model above. First, note that the initial condition does not disappear as t grows. In fact, $E(y_t) = y_0$. So, in this sense, the initial condition has a lasting impact on the series. Put another way, although the series ensemble can display many different kinds of behavior (cf. Figure 1), the equilibrium value is always the same (y_0). Second, because the shocks now are all equally weighted, their impact is more akin to phat than path dependence. In this sense, the "running tally" idea as operationalized in some of the work by Erikson et al (1998) does not connote either path or equilibrium dependence. The running tally idea is closer to a ball-urn model of the Balancing Process type; it is not a Polya process because running tally implies phat but not equilibrium dependence.[7] It follows that tests for unit roots for univariate series, in effect, are tests of early path dependence and of phat dependence. In this regard,they indeed do constitute tests for distinct kinds of data generating processes (cf. Page 2006: 97).

2.2 Multivariate Linear Time Series Models

These models can be divided into strongly and weakly restricted varieties (Freeman et al 1989). Among other things, strongly restricted models are based on investigator imposed exact restrictions for exogeneity and lag length.

2.2.1 Strongly Restricted Multivariate Time Series Models

There are at least two categories here. The first is the familiar, single equation time series regression model. Page argues that "if a regression equation does not include any time lags

[7]Page's (2006:99) Balancing Process is a follows: Initially an urn contains one maroon ball and one brown ball. In any period, if a brown (resp. a maroon) ball is selected then it is put back in the urn together with an additional ball of the opposite color. This process is outcome phat dependent and it has a unique equilibrium.

it captures only phat dependence." He contends that retrospective voting models do capture recent path dependence (2006: 98, 104).[8]

One of the most common models of this kind is:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{i=0}^{p} \beta_i x_{t-i} + \epsilon_t \tag{6}$$

This model is strongly restricted insofar as only one lag of the endogeneous variable is included on the right hand side of the equation, the x variable is assumed to be exogenous, and p lags of x are stipluated to be causally related to y. According to Page (2006) path dependence is embodied in the fact that the sequence of the realization of the x variable (*not* the shocks) affect the values of y_t. Whether this path dependence is of the "recent" type supposedly depends on the magnitudes of the β coefficients.

But what does this equation tell us about equilibrium dependence? The answer is that if y_t always depends on past x_t the equilibrium value of y_t constantly varies. So equation (6) connotes both path dependence and equilibrium dependence. The initial condition for the exogenous variable, x_0, probably is not important since its weight decreases as time increases.[9] The idea of a Polya type equation that manifest phat and equilibrium dependence is embodied in a simpler version of equation (6):

$$y_t = \alpha y_{t-1} + \beta x_t + \epsilon_t \tag{7}$$

Here the x variable has no lags so, in Page's framework, it connotes phat dependence. But the multipliers associated with this equation connote equilibrium dependence insofar as different, one-time increases in the values of x produce different long term values in y_t. [10]

The second type of model specifies a relationship between two random walks or, to be more precise, two processes of the same order of integration. This is the single equation, error correction model (ECM). It is based on the idea that even though two processes may each be nonstationary, a weighted sum of them may be stationary. This weighted sum connotes the idea of long-run equilibrium. When the sum is zero, long-run equilibrium is achieved. In the short term, in the single equation case, changes in one variable are a function of weighted lagged changes in itself, weighted lagged changes in an exogenous variable, and *error correction*. An example of such a single equation model is

$$\Delta r_t = a_{10} + \alpha[r_{t-1} - \beta s_{t-1}] + \sum_{i=1}^{p} a_{11}(i)\Delta r_{t-i} + \sum_{i=1}^{p} a_{12}(i)\Delta s_{t-i} + \epsilon_{1t} \tag{8}$$

[8] Page presents a useful example in which the dependent variable is vote for an incumbent congressperson and the independent variable is a measure of that individual's ideology, a measure based on the *set* not order of the congressperson's roll call votes. With no lag of this measure on the right hand side of the respective regression equation, the model captures phat dependence. Presumably, lagging this same measure means that the vote depends on the *sequence* of ideology scores and therefore comes closer to capturing path dependence. No example of a retrospective voting model that captures path dependence is provided by Page.

[9] Technically, the initial value of the dependent variable, y_0, also must be considered. But, again, if $|a_1| < 1$, this initial value should decay.

[10] For a useful study of the nature and problems of estimating equation (6), see Keele and Kelly 2005.

Here $a_{10}$is a constant, s_t is assumed to be exogenous, p lags of the short term changes in each variable are stipulated, the second term on the right hand side of the equation is the cointegrating vector or equilibrium relationship, and α is the rate of error correction. So when the system is in long term equilibrium, $r_{t-1} = \beta s_{t-1}$; there is no error correction in r_t. The Granger Representation Theorem holds that when processes of are integrated of the same order, such an error correction model exists. Erikson et al (1998) posit a model of this kind to explain macropartisanship in terms of (purged) approval and consumer sentiment.

Recall that the expected value–unconditional mean– of random walks are equal to their initial conditions; in this sense, each has a unique equilibrium. Beyond this random walks are phat dependent. What then does it mean to say such processes are in long term equilibrium? Or, where in the historical sociology or dynamic systems framework is there a concept that corresponds to two or more phat dependent systems "trending together toward a long-term equilibrium relationship." This idea seems to have no parallel in the literature on path dependence. [11]

2.2.2 Weakly Restricted Multivariate Time Series Models

The parallels in this tradition are the vector autoregression (VAR) and vector error correction (VECM) models. The restrictions in these models are relatively weaker insofar as analysts let the data choose lag lengths, exogeneity assumptions are avoided, etc. [12] The general form of the VAR model with p lags, VAR(p), is:

$$y_t = AY_{t-1} + B_0 x_t + u_t \tag{9}$$

where y_t is a K x 1 vector of endogenous variables, A is a K x Kp matrix of coefficients, B_0 is a K x M matrix of coefficients, x_t is a M x 1 vector of (presumed) exogenous variables, u_t is a K x 1 vector of white noise shocks, and Y_t is a the Kp x 1 matrix denoted by $Y_t = \begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}$.

If the modulus of each eigenvalue of the matrix A is strictly less than one, the estimated VAR is stable; certain conditions also apply to the initial conditions for the dynamics system. [13] Enders (2010: 295ff) explains the stability conditions for a simple, two variable model with one lag. He calls this the VAR model in standard form:

$$x_t = A_0 + A_1 x_{t-1} + \epsilon_t \tag{10}$$

where x_t is the 2 x 1 vector of variables, A_0 is a 2 x 1 vector of constants, A_1 is a 2 x 2 matrix of constant coefficients, and ϵ_t is a nx1 vector of white noise shocks. He shows that

[11] In fact, Enders (2010: 359) argues that the idea cointegration connotes equilibrium is problematic. He contends that in econometrics the term can be interpreted causally, behaviorally, or simply as a reduced form relationship.

[12] Note the word "relatively." As equation (9) shows, even VAR models contain restrictions including, sometimes, presumed exogenous variables. Once more, the focus here is on likelihoodist models.

[13] The notation for the general version of the VAR(p) is taken from the STATA9 time series manual (pps. 293ff, 345); the original source is Lütkepohl(1993). STATA now provides a test of this stability condition under the rubric varstable.

the solution of this equation can be written

$$x_t = \mu + \sum_{i=0}^{\infty} A_1^i \epsilon_{t-i} \tag{11}$$

where μ is a 2 x 1 vector of the means of the two variables. So once again, if the relevant coefficients are less than one in absolute value, the expected value of this dynamic process is the mean of each series, μ. In this sense, the stable VAR(p) model also implies path dependence but not equilibrium dependence.

The vector error correction model, VECM, is designed to analyze a system of variables, a system which may contain *multiple* cointegrating vectors.[14] Consider the system of equations:

$$x_t = A_1 x_{t-1} + \epsilon_t \tag{12}$$

where x_t is a n x 1 vector of variables, A_1 is an n x n matrix of parameters, ϵ_t is a n x 1 vector of shocks. Subtract x_{t-1} from each side of equation 12 and define I as the n x n identity matrix. The result is:

$$\begin{aligned} x_t &= -(I - A_1)x_{t-1} + \epsilon_t \\ &= \pi x_{t-1} + \epsilon_t \end{aligned}$$

where π is the n x n matrix $-(I - A_1)$. If the rank of π is zero, the system amounts to a set of independent, first order integrated variables. In other words, we have an independent set of phat dependent processes for which the respective initial values of the variables do not decay. If the rank of π, r, greater than zero but less than n, there are r cointegrating vectors. That is, there are r moving equilibria between the phat dependent variables. [If the rank of π equals n, all the variables in the system are stationary. Hence they are path dependent but they are not equilibrium dependent.]

To my knowledge, neither Erikson et al or other scholars have explored the possibility of multiple moving equilibria in the system that explains macropartisanship. [15]

3 Nonlinear Time Series Models

Most of these models exhibit equilibrium dependence as well as the properties of phat or path dependence. One such model allows for switches in the rate (paths) of adjustment to a moving equilibrium in phat dependent processes.

3.1 Univariate Nonlinear Time Series Models

There are many types of these models. One collection is based on the idea of "regime switches." An example is the threshold autoregressive model, TAR:

$$y_t = \begin{cases} a_1 y_{t-1} + \epsilon_{1t} & \text{if } y_{t-1} > 0 \\ a_2 y_{t-1} + \epsilon_{2t} & \text{if } y_{t-1} \leq 0 \end{cases}$$

[14]The following presentation of the VECM model is summarized from Enders (2010: 371ff)

[15]A study in American politics more sensitive to this possibility is Ostrom and Smith (1993). For a Bayesian approach this is more consistent with this idea see Brandt and Freeman (2009).

This data generating process is a combination of two simple AR(1) processes, depending on the sign of its previous value, y_{t-1}. So it will exhibit two kinds of path dependence because the coefficients in each process are different. But, in both cases the process has the same expected value, namely, zero.

If (one of) the AR processes contain (a) constants, this TAR model will exhibit both path and equilibrium dependence. An example of such a model is:

$$y_t = \begin{cases} a_{10} + a_1 y_{t-1} + \epsilon_{1t} & \text{if } y_{t-1} > 0 \\ a_{20} + a_2 y_{t-1} + \epsilon_{2t} & \text{if } y_{t-1} \leq 0 \end{cases}$$

Each AR process will have a different expected value, either $\frac{a_{10}}{1-a_1}$ or $\frac{a_{20}}{1-a_2}$. So it will exhibit path dependence and equilibrium dependence.

For macropartisanship, say m_t, we might have a data generating process in which the switch occurs when m_{t-1}, exceeds a level such as .60. In other words, when the data generating process is in the first regime and the level of Democratic partisan identification exceeds sixty percent the American polity gravitates to one equilibrium. But when, because of shocks embodied in ϵ_{1t}, m_t drops below .60, the American polity gravitates toward a different equilibrium. With one exception (Jackman 1987) this kind of path and equilibrium dependence appears not to have been explored in the literature.

3.2 Multivariate, Nonlinear Time Series Models

Jackson and Kollman (2010) analyze strongly restricted, nonlinear, multivariate time series regression models in which one variable is posited to be exogenous. They show how such models can exhibit path and near-path dependence and, concomitantly, equilibrium dependence.

As regards the idea of random walks as phat dependent processes, one model embodies the idea of nonlinear (switching) error correction.[16] Assume two processes, r_{Lt}, r_{St} are both first order integrated and also cointegrated (a linear combination of the two processes is stationary). Then a threshold model of the momentum type, M-TAR, might be used to represent switching between *multiple* error correction processes. This model would have the form:

$$\Delta r_{Lt} = \alpha_{11} I_t [s_{t-1} - \beta] + \alpha_{12}(1 - I_t)[s_{t-1} - \beta] + A_{11}(L)\Delta r_{L,t-1} + A_{12}\Delta r_{S,t-1} + \epsilon_{1t}$$
$$\Delta r_{St} = \alpha_{21} I_t [s_{t-1} - \beta] + \alpha_{22}(1 - I_t)[s_{t-1} - \beta] + A_{21}(L)\Delta r_{L,t-1} + A_{22}\Delta r_{S,t-1} + \epsilon_{2t}$$

where the α terms are adjustment coefficients, $s_t = r_{Lt} - r_{St}$, the $[s_{t-1} - \beta]$ terms are cointegrating vectors, the $A(L)$ terms are lag operators, and the I_t variable is an indicator function defined as

[16]This example is a simplified version of an example in Enders (2010: 481. In his model r_{LT}, r_{St} represent the interest rate on ten year government securities and the federal fund rate, respectively. Each series is I(1). The model explains regime shifts in terms of how changes in the interest rate spread, $s_t = r_{Lt} - r_{St}$, increasing vs. decreasing, translate into different rates of error correction. In this case, there is no error correction when $s_{t-1} = \beta$.

$$I_t = \begin{cases} 1 & \text{if } \Delta s_{t-1} > 0 \\ 0 & \text{if } \Delta s_{t-1} \leq 0 \end{cases}$$

Thus, the rate of adjustment to the moving equilibrium between the two phat processes varies depending on whether in the previous period s_t was increasing or decreasing.

Suppose we think of r_{Lt} and r_{St} as macropartisanship and presidential approval, respectively. Suppose further than these two variables, the former representing long term macropolitical disposition and the latter short term disposition, are cointegrated. Then this model suggests that the American polity switches between two regimes each with phat dependence and also with different rates of error correction. While scholars like Jackman (1987) have explored the possibility of switching in American macropolitics, no one, to my knowledge, has analyzed this possibility of nonlinear error correction. [17]

4 Conclusion

Conventional time series methods give us tools to identify and analyze data generating processes that embody most of the key concepts associated with the idea of path dependency. We simply need to be clear about the nature of each model, how (if) each model embodies the impact of initial condition, the set or sequence of shocks that a data generating process experiences, and multiple equilibria. As I have shown, doing this illuminates new and potentially useful ideas about the nature of American macropolitical dynamics. It also suggests the need for tests for nonlinearity in macropartisanship and in (macropartisanship's relationship to) other theoretically important series. [18]

5 References

Brandt, Patrick T. and John R. Freeman. 2009 "Modeling Macropolitical Dynamics" *Political Analysis* 17(2):113-142.

Enders, Walter 2010. *Applied Econometric Time Series* Third Edition. John Wiley And Sons.

Freeman, John, John Williams, and Tse-min Lin 1989. "Vector Autogression and the Study of Politics" *American Journal of Political Science* 33(4): 842-877.

Box-Steffensmeier, Janet and Renee Smith. 1996. "The Dynamics of Aggregate Partisanship" *American Political Science Review*90(3): 567-580.

[17] Of course, the first step is to test for evidence of nonlinearity in the respective relationship. Various tools are available for this purpose, e.g., the McLeod-Li, RESET, and LM tests.

[18] The deeper challenge is to develop theories that predict mulitiple, moving equilibria in American political dynamics. See, for example, Mebane (2000).

Erikson, Robert, Michael MacKuen, and James Stimson. 1998 What Moves Macropartisanship? A Reply to Green, Palmquist, and Schickler" *American Political Science Review* 92(4): 901-912

Goldstone, Jack A. 1998 "Initial Conditions, General Laws, Path Dependence and Explanation in Historical Sociology" *American Journal of Sociology* 104: xx

Green, Donald, Bradley Palmquist and Eric Schickler.1998 "Macropartisanship:Replication and Critique" *American Political Science Review* 92(4): 883-900.

Hamilton, James D. 1994 *Time Series Analysis* Princeton, N.J. Princeton University Press.

Jackman, Simon. 1987 "Public Opinion Over the Long Run: Presidential Approval, the Macroeconomy, and Rallies." Paper Presented at the Annual Meeting of the APSA, Washington, D.C.

Jackson, John and Kenneth Kollman. 2010. "Modeling, Distinguishing and Measuring Path Dependence, Near Path Dependence, and Outcome Dependence." Paper prepared for a conference on Path Dependence, University of Minnesota, Minneapolis. June.

Keele, Luke and Nathan J. Kelly 2005 "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables" *Political Analysis* 23:186-205.

Lütkepohl, Helmut. 1993. *Introduction to Multiple Time Series Analysis* Berlin: Springer-Verlag.

Mahoney, James. 2000. "Path Dependence in Historical Sociology" *Theory and Society* 29:507-548.

Mebane, Walter R. Jr. 2000. "Congressional Campaign Contributions, District Service, and Electoral Outcomes in the United States: Statistical Tests of a Formal Game Theoretic Model with Nonlinear Dynamics." In *Political Complexity: Nonlinear Dynamics of Politics* Diana Richards Editor. Ann Arbor: University of Michigan Press.

Ostrom, Charles W. Jr. and Renee Smith 1993 "Error Correction, Attitude Persistence, and Executive Rewards and Punishments: A Behavioral Theory of Presidential Approval." In *Political Analysis* Vol. 4 John R. Freeman Editor, Ann Arbor: University of Michigan Press.

Page, Scott. 2006. "Path Dependence" *Quarterly Journal of Political Science* 1:87-105.