Political Analysis (2012) 20:157–174
doi:10.1093/pan/mpr050

Modeling, Measuring, and Distinguishing Path Dependence, Outcome Dependence, and Outcome Independence

John E. Jackson and Ken Kollman
Department of Political Science, University of Michigan, Ann Arbor, MI 48109
e-mail: jjacksn@umich.edu (corresponding author)

Edited by R. Michael Alvarez, John R. Freeman, and John E. Jackson

Path dependence is commonly used to describe processes where "history matters," which encompasses many different kinds of temporal dynamics. This essay distinguishes path-, or equilibrium-, dependent processes where early conditions continue to matter, from outcome-dependent processes where recent history matters and from outcome-independent processes where history does not matter. Others have argued for a precise and restrictive definition of path dependence. We build on this and distinguish among different types of outcome-dependent processes when these conditions for path dependence are not fully satisfied.

1 Introduction

"Path dependence" is used to describe a range of phenomena. We build on previous research that identifies the precisely specified boundaries between actual path dependence and other kinds of dynamic processes. We locate substantively meaningful boundaries among several dynamic processes where "history matters" but that are not actually path dependent. Using examples, we discuss the empirical implications of these various processes and how one might distinguish among them.

Path dependence in Page (2006) and Jackson and Kollman (2010) is a dynamic process of a variable y with an equilibrium distribution that is a function of initial conditions, y_0, and possibly an early sequence of outcomes $y_1, y_2, y_3, \ldots, y_s$ in reaching that equilibrium. Informally, where one ends up depends on the starting point and the path taken to get there. Path dependence, which we think is actually quite rare, contrasts with most dynamic processes where previous values may affect a given outcome, y_t, but those previous values do not affect the equilibrium distribution. For such nonpath-dependent processes, only the values of the exogenous variables, say X (and not any previous value of y), determine the equilibrium. When previous values of y affect its current value along the path but not its equilibrium, we refer to such processes as equilibrium independent. As in a standard stationary time series, for equilibrium independence, the roads may differ but all roads lead to Rome (i.e., the same equilibrium). Our main purpose is to distinguish clearly, conceptually, and empirically among these different processes, some of which are often and casually lumped together as "path dependent." Concluding that a process is equilibrium independent can be important substantively, but it is not the same as concluding path dependence.

For intuition on the difference between path dependence and equilibrium independence, consider the debate about how to sequence reforms—and importantly, whether sequence matters—in countries transitioning from authoritarian regimes and controlled economies to democratic governments and free markets.

Author's note: Revision of a paper prepared for a Conference on Path Dependence, sponsored by the Society for Political Methodology and the Department of Political Science at the University of Minnesota, Minneapolis, MN, June 4 and 5, 2010. We are grateful to Jude Hayes and John Freeman in particular and the conference participants in general for their very helpful comments and suggestions. All errors are, of course, our own.

© The Author 2012. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oup.com

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

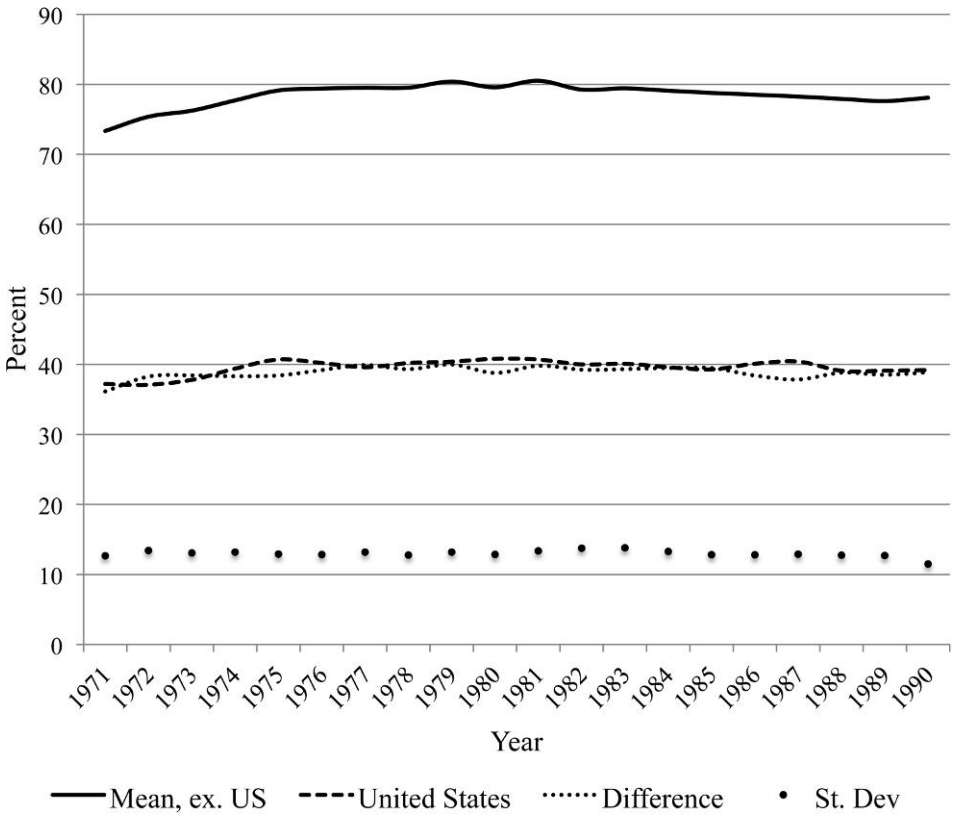


Fig. 1 Public health expenditures/total health expenditures.

Does it matter to the final outcome if economic reforms precede political reforms? If transitions are not path dependent, sequencing does not matter in achieving the end state, a democratic capitalist country. A country will find its equilibrium political institutions and economy regardless of which reforms are instituted first. Alternatively, one could argue that the economy must be reformed first, otherwise the country will revert to an authoritarian regime that will halt economic reforms in response to the dislocations in the economy (Nelson 1993; Hellman 1998). Or, governmental reforms, such as secure property rights and a strong legal and bureaucratic system to enforce them combined with democratic controls on the predatory propensities of strong states, are required first (Sachs and Pistor 1997, on property rights and Grzymala-Busse 2007, on democratic controls). These latter two arguments—that sequencing matters—imply path dependence. All three causal models—(1) the ordering does not matter, (2) that economic reforms must precede political reforms, or (3) vice versa—are stated simplistically here, but the contrast indicates the value of thinking seriously and precisely about path dependence.

Now consider another example. An often-cited example of path dependence is the absence of public health care funding in the United States in contrast to other industrial countries (Hacker 2002). In simple terms, the F. D. Roosevelt administration put the United States on the path to privately funded and provided health care in the 1930s. Once this system and its vested interests were in place, health care provision reached a distinctive U.S. equilibrium in contrast to that in Europe, the United Kingdom, and even in Canada. The explicit implication of this argument is that health care in the United States would remain essentially private, with the resulting consequences, such as limited coverage and provision tied to employment. Figure 1 plots the public share of total health care funding for the period 1971–1990 in the United States, the average in 16 other industrial (OECD) countries, their difference, and the standard deviation in the shares among all countries (OECD 2010).[1] Other than a slight increase in the public

[1]The non-U.S. countries are Australia, Austria, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Japan, New Zealand, Norway, Portugal, Spain, Sweden, and the United Kingdom (OECD 2010).

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

share in the non-U.S. countries in the early 1970s, these series appear to be constants, and the difference and standard deviation are particularly steady. Is there enough evidence in these data to conclude that the early decisions about funding health care in the United States put that country on a different path and a permanently different equilibrium than other industrial countries? We return to these data later to contrast different types of dynamic processes.

The distinction between path-dependent and equilibrium-independent processes has important substantive and methodological consequences. The example of sequencing economic and political reforms during transitions illustrates the substantive importance. Policy choices must be considered carefully in light of (ideally) good theory and documented empirical evidence about the dynamics of the process, not just the static equilibrium properties. The methodological implications are just as profound. It is commonly believed that cross-sectional analyses estimate the long-run equilibrium behavior given the values of the properly specified exogenous variables, X (Kuh 1959; Jackson and Mach 2011).[2] This statement assumes that the cross-sectional units' outcomes are at or near their equilibrium distribution means given X so the observed outcomes are these mean values plus a white noise term. If, however, the outcomes are generated by a path-dependent process, the observed outcomes, even if the units are close to their equilibrium means, will be functions of initial and early conditions in addition to the values of X. Omitting variables for these conditions in the cross-sectional study yields potentially biased coefficients. Jackson (1996) analyzes simulated data sets from a path-dependent process to show that cross-sectional regressions with only current exogenous variables do not reproduce the correct values of the underlying model. The implication is that if the dynamic process generating a set of cross-sectional observations is path dependent, one cannot reliably use the data to estimate long-run equilibrium relationships, at least not without including variables measuring the early historical outcomes.

We now proceed by matching different uses of the term path dependence with formal models to make explicit the processes that would produce the outcomes described. These formal models can then be used to discuss different estimation strategies for comparing models and ultimately for deciding if a process of substantive interest is consistent with one of these models and is or is not path dependent.

2 Equilibrating Processes

Discussions of path dependence must confront the crucial question of whether the process has an equilibrium and even what is meant by an equilibrium. An equilibrium is defined here as the limiting distribution for the outcomes of some process. We concentrate on models with outcomes measured with interval variables, such as vote shares or income as distinct from discrete outcomes. For such a process, an outcome is denoted by y, and the history of these outcomes from some initial period to time t, referred to as a series, is denoted by $y_s, s = 0, \ldots, t$. An equilibrium for this process is defined in terms of several criteria. The first is that the series must be stationary with expected value $E(y_t) = \mu_t$, finite variance $E[(y_t - \mu_t)^2] = \sigma_t^2$ and with lagged covariances, $E[(y_t - \mu_t)(y_{t-j} - \mu_{t-j})] = \gamma_j$, which are independent of time and decline with the length of the lag.

A second condition is that $E(y_t) = \mu_t \rightarrow \mu$ as t gets larger, and that

$$\lim_{T\rightarrow\infty} \frac{1}{T} \sum_{t=1}^{T} y_t = \bar{y} = \mu. \tag{1}$$

This condition means that as the length of the process is extended for an indefinite number of periods, the average of all the values of y_t in the series equals the long-run expected value of y. With finite variances, the σ_t^2s, this implies that in the limit, the time average of the series, $\bar{y}$, is collapsing about its mean.

The AR(1) linear model illustrates this definition of an equilibrium. Assume the process is initially at X_0 and y_0. There is a shift in X to X^*, and y adjusts such that,

$$y_1 = \rho y_0 + (1-\rho)X^*\beta + u_1 = X^*\beta + \rho(y_0 - X^*\beta) + u_1, \tag{2}$$

[2] We are restricting this discussion for now to single-valued functions of X.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

with $\rho < 1$.[3] If X remains at X^*, then at period t,

$$y_t = \rho y_{t-1} + (1-\rho)X^*\beta + u_t = X^*\beta + \rho^t(y_0 - X^*\beta) + \sum_{s=0}^{t-1}\rho^s u_{t-s}. \tag{3}$$

Assume that the u_ts are i.i.d. with $E(u_t) = 0$ and $E(u_t^2) = \sigma_u^2$. As t increases, the term involving ρ^t becomes ignorable and

$$\begin{aligned} E(y_t) &= X^*\beta + E(U_t) \\ &= X^*\beta = \mu. \end{aligned} \tag{4}$$

As T $\to \infty$, $\frac{1}{T}$ times the sum of the values for y in the series from $t = 0$ to T is

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} y_t = \lim_{T\to\infty}\frac{1}{T}\left[\sum_{t=1}^{T} X^*\beta + \left(\sum_{t=1}^{T}\rho^t\right)(y_0 - X^*\beta) + \sum_{t=1}^{T}\sum_{s=0}^{t-1}\rho^s u_{t-s}\right]. \tag{5}$$

Appendix A shows that the second and third terms of equation (5) go to zero, meaning $\bar{y} = X^*\beta = E(y_t)$ for large t and as $T \to \infty$, satisfying the equilibrium conditions.

3 Dynamic Processes

We have so far defined an equilibrium as the limiting distribution for some process. The remaining discussion focuses on processes with an equilibrium, recognizing that there are processes such as a random walk that do not have an equilibrium by our definition. Dynamic processes are separated into categories that distinguish how exogenous variables and shocks affect its equilibrium distribution. Page (2006) categorizes these processes as outcome independent and outcome dependent, depending upon whether previous outcomes do not or do affect current outcomes, respectively. Outcome-dependent processes are further divided into those that are equilibrium dependent and those that are equilibrium independent.[4] In equilibrium-dependent series, early historical events, possibly including their ordering, matter for the equilibrium distribution. This category is most consistent with common usage of the term path dependence. For equilibrium independence, some aspect of history is reflected in current outcomes along the path to an equilibrium, but the historical aspects do not affect the equilibrium distribution. This last category, which we suspect is by far the largest in terms of types and frequency, includes all processes where history is present in current outcomes but not in the mean value of the equilibrium distribution.

Our discussion of dynamic processes begins with the first two categories as these are extreme or pure cases. Then, we present examples of the third category, equilibrium-independent processes, and discuss them in relation to the other two categories. Freeman (2011) enumerates many dynamic processes and their properties. His account provides a way to place our discussion within the time series literature. (See also De Boef and Keele 2008.)

3.1 *Outcome-Independent Processes*

Consider outcome-independent processes where history matters not at all. In such processes, the realized values y_t depend only upon the current values for any exogenous variables. That is $y_t = F(X_t, \beta) + u_t$.

[3] Equation (3) can also be written as $y_t = \rho y_{t-1} + X^*B + u_t$. If ρ is constant over time and less than one, these forms are equivalent with the equilibrium of equation (3) being X^*B, whereas the equilibrium mean in this expression is $X^*B/(1-\rho)$. Each period's values for y_t are also equivalent using a similar transformation of the B's. If ρ varies temporally, these forms are not equivalent and the form in equation (3) is a more desirable expression as the equilibrium value remains X^*B. The expressions are also not equal if $\rho = 1$, in which case equation (3) becomes a random walk, whereas in the alternative expression, $y_t = y_0 + t \cdot X^*B + \sum_{s=0}^{t} u_s$. These differences are elaborated subsequently.

[4] Page (2006) does not use the term equilibrium independence. We introduce this term to define processes that are outcome dependent but not equilibrium dependent.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

In the linear model in equation (2), outcome independence occurs when $\rho = 0$. From our perspective, this anchors one end of the spectrum of dynamic processes. For outcome independence, what is being excluded is the possibility that y_t is related to any previous outcomes, denoted as y_s, for $s < t$. This is a very exacting condition as it requires the process to reach its equilibrium within one time period after any exogenous shock. It is hard to think of such a process that has much interest to social scientists who study time series data or data sets with comparable historical variables, or in fact any dynamic process.[5]

3.2 *Equilibrium-Dependent Processes*

At the other end of the spectrum is equilibrium, or path, dependence. With equilibrium dependence, the equilibrium distribution depends upon previous realized values for y and particularly its initial or early values. A formal definition requires use of the concept of a series introduced earlier to define an equilibrium. That discussion focused on a single realization of the entire process—one series. Now consider that conceptually there could be additional series, what Hamilton (1994) calls an ensemble. In other words, we could rerun history to get additional samples of the y_ts. Denote by y_s^1, $s = 1, \ldots, t$ the first series and by y_s^i, $i = 1, \ldots, I$ the set of all possible series. Each series will have its equilibrium mean value μ^i. In an equilibrium-dependent process, μ^i varies with initial and early conditions. Page (2006) distinguishes between path dependence and what he playfully calls phat dependence. In the former, the precise values for y_t and their order affect the equilibrium, where in the latter, the exact values for y_t matter but not their sequence. In this article, we are concerned with equilibrium dependence without reference to which of these two forms it takes though most examples and discussion relate to path dependence.

Page shows that a necessary condition for equilibrium dependence is that the values for ρ must vary over time, meaning ρ in equation (2) must have a t subscript, ρ_t. (See Beck 1983, for a discussion of time-varying parameters in time series models.) Jackson and Kollman (2010) show that two other conditions are required for equilibrium dependence. The first is that $\rho_t = 1$ for $t \geqslant t^*$, which means the equilibrium mean value for y_t at $t = t^*$ is "locked in" for future periods. If t^* is less than the time it takes for y_t to approach its equilibrium mean, this locked-in value for y_t will be a function of the initial conditions and early values for y_t. The second condition is that u_t must be a negatively autocorrelated stochastic term, $u_t = -\delta u_{t-1} + \epsilon_t$, with $0 < \delta < 1$. If $\delta = 0$, meaning there is no autocorrelation to u_t, the sequence for y_t becomes a random walk for the time periods after t^*. Jackson and Kollman (2010, 264) show that with $\delta > 0$, the process is stationary even with $\rho = 1$. This is a version of a stationary error correction model (ECM) where the correction occurs to the difference of y_{t-1} from its expected value at $t - 1$ rather than to its equilibrium mean at $t - 1$.[6] The latter model is a special case of the model here.

We elaborate on the first condition as it sets up later distinctions among types of dynamic processes. Alter equation (2) so that X follows the sequence $X_0, X_1, X_2, X^*, \ldots, X^*$ and so that ρ varies with time.[7] The change in the sequence of values for X means that the path by which the process reaches its equilibrium can be varied. These alterations change equation (3) so that,

$$y_t = X^*\beta + \left(\prod_{s=1}^{t} \rho_s\right)(y_0 - X_1\beta) + \left(\prod_{s=2}^{t} \rho_s\right)(X_1 - X_2)\beta + \left(\prod_{s=3}^{t} \rho_s\right)(X_2 - X_3)\beta + V_t, \qquad (6)$$

where

$$V_t = \left[\sum_{j=1}^{t-1}\left(\prod_{s=j+1}^{t} \rho_s\right) u_j + u_t\right].$$

[5] If observations on y_t are widely spaced, if ρ is small, and if X remains constant for an extended period, it could appear that equilibrium is reached in one period. See Freeman (1989) for a discussion of the problems created by temporal aggregation.

[6] We use the term stationary error correction model (ECM) to denote the functional form for a standard ECM but where the series are stationary, as shown in De Boef and Keele (2008) and Greene (2008, 689). This model is distinguished from the ECM discussed by Enders (2010), for example, which is defined for nonstationary but cointegrated variables.

[7] With temporal variations in ρ, the expression in equation (6) differs from the comparable expression for y_t when the dynamic model is written as $y_t = \rho_t y_{t-1} + X^*B + u_t$. In this latter case, $y_t = \left(\prod_{s=1}^{t} \rho_s\right) y_0 + \left(1 + \rho_t + \rho_t\rho_{t-1} + \cdots + \prod_{s=1}^{t} \rho_s\right) X^*B + v_t$. If $\rho^{\max}$ is the maximum value for ρ_s, the value of the sum in parentheses is less than $1/(1-\rho^{\max})$, which is a finite value with $\rho_s < 1$ for all s. The exact expected value of y_t depends on the recent values for ρ_s, and the equilibrium mean is not predictable based solely on X^*B.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

The condition that $\rho_t = 1$ for $t = t^* + \tau$, means that $y_t = y_{t-1} + u_t$, so that

$$y_t = X^*\beta + \left(\prod_{s=1}^{t^*} \rho_s\right)(y_0 - X_1\beta) + \left(\prod_{s=2}^{t^*} \rho_s\right)(X_1 - X_2)\beta + \left(\prod_{s=3}^{t^*} \rho_s\right)(X_2 - X_3)\beta + V_t. \quad (7)$$

Depending upon the values of t^* and the ρ_s, it may well be that $\prod_{s=1}^{t^*} \rho_s \neq 0$ so that y_t is "locked-in" to some equilibrium mean value μ that does not equal $X^*\beta$ but is $X^*\beta$ plus terms that include y_0, $X_1\beta$, and $X_2\beta$. Thus, the equilibrium depends upon the initial and early values of y. The lower the value for t^*, the number of periods prior to this lock-in, and/or the larger the values for the ρ_ss, the larger the weight given to initial and early conditions.

Now consider a second series with the same value for X^* but different initial and early values for y_0, X_1, and X_2. Using equation (7), we can show that for the same values for t^* and ρ_s, the equilibrium mean for this second series will be different than for the first series. If there are many series denoted by i and if we denote the equilibrium mean for each possible series i as μ^i, then μ^i will vary with the values for y_0^i, X_1^i, and X_2^i. This is the definition for equilibrium, or path, dependence.

The conditions for path dependence are general for this common dynamic model where y_t is a time-varying convex combination of its previous realization, y_{t-1}, and current exogenous conditions, $X_t B$, plus a stochastic term u_t. This equilibrium expression is written as a linear additive model but this need not be the case. The relationship could be any function of X_t, for example, $F(X_t, \beta)$, so that $y_t = \rho_t y_{t-1} + (1 - \rho_t)F(X_t, \beta,) + u_t$.

The requirements that $\rho_t = 1$ for periods subsequent to t^*, and that $\delta > 0$, are quite demanding conditions and suggest to us that equilibrium dependence, meaning path dependence, is a rare phenomena. But, when it does exist, path dependence is a powerful and important concept. To illustrate this, consider recent work on political and economic development (Acemoglu et al. 2001, 2002) that analyzes the variance in current economic performance across a set of former European colonies. They maintain that current institutions, such as rules about expropriation, property rights, and rule of law, play an important role in economic performance, and that the current institutions in their sample of countries are strongly related to the institutions put in place at the time of colonization. They do not use the term path dependence but their empirical analysis and discussion clearly imply what we are here calling equilibrium dependence.

Jackson and Kollman (2010) formalize the Acemoglu et al. model as two dynamic equations similar to equation (2), one modeling institutions and the other economic outcomes. Acemoglu et al. do not cast their discussion as a time series model but the structure is implicit in their argument. Further, they say that, "Europeans adopted very different colonization policies in different colonies, with different associated institutions.... These institutions persisted to the present" (Acemoglu et al. 2001, 1369). This clearly implies that in the implicit equation for institutions, ρ_t reached one early in the development process, the condition for path dependence, and that subsequent outcomes for these institutions are random fluctuations about this equilibrium mean. Current institutions then are functions of early forms of these institutions and of any exogenous variables creating them, up to the point that the institutions became locked in. Such an interpretation is consistent with Acemoglu et al.'s empirical results relating the protection against expropriation between 1985 and 1995 to political institutions in 1900 and settler mortality at the beginning of colonialization.

There are other explanations for the origins of the colonial institutions, such as the pattern and type of missionary settlements (Engerman and Sokoloff 2002; Woodberry et al. 2011), but the basic point is that these studies show the important and persistent role of early conditions on current outcomes. This interpretation of the Acemoglu et al. findings has implications for the earlier discussion of sequencing in transitions used to illustrate the importance of distinguishing path dependence from nonpath dependence. If institutions become locked in, as is likely given the constitutions establishing these institutions and the rules that are biased against change, then the transition process is likely to be path dependent, so sequencing matters. Second, their work argues that it was having the "right" institutions that enabled certain former colonies to take advantage of the major shock, the industrial revolution. Former colonies without the right institutions could not adapt and have grown less and continue to perform less well than countries with good institutions. Though they do not discuss path dependence and transitions and sequencing, their work implies path dependence. This means it is important to install good institutions prior to economic reforms in order to accommodate and build support for the economic shock. Roland

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

(2000, 49) seems to support this view when he says, "Note that, all over Eastern Europe, political reform has preceded economic reform. Even in China, ... The start of economic reforms in 1978 was preceded by a wave of depoliticization and relative liberalization of Chinese life." He goes on elaborate how Russia did the economic reforms first and has been the least successful transition country.

3.3 *Equilibrium Independence*

If outcome independence and equilibrium dependence are rare, then equilibrium independence, the "all others" category, describes most dynamic processes and is likely a heterogenous category inviting careful exploration. The classic example of an equilibrium-independent process is the standard autoregressive model shown in equation (3) with $\rho < 1$ and in equation (6) with varying values for ρ_s but with all $\rho_s < 1$ for $s \leqslant t$. In the latter equation, all the terms in parentheses go to zero as t increases. In this process, the realized values for y_t depend on its previous values but the equilibrium mean is $\mu = X^*\beta$, which is independent of initial conditions and values for y_t early in the sequence. Different series i with different values y_0^i, X_1^i, and X_2^i all have the same expected value for y_t for large values of t, meaning that $\mu^i = \mu$ for all series.[8] Thus, the equilibrium mean value for each series is the same and is independent of the initial and early outcomes, which is the definition of equilibrium independence.

Equilibrium-independent processes can have many different properties, some of which may resemble path dependence for varying lengths of their sequences. These different properties result from different patterns to the values for ρ and from the frequency of shocks to the process that change the values of X_t. We now explore some of these different systems and relate them to popular discussions of path dependence.

3.3.1 Sticky process

A common process that often appears to be equilibrium dependent is an equilibrium-independent process with a large value for ρ. A highly autoregressive process that is slow to adapt to changes in X and to reach its equilibrium will appear to be path dependent for most of its sequence. Again, using equation (7), we can show that when ρ is large, the effects of initial conditions and early events will be evident in current outcomes for large values of t. For example, after 100 time periods, the coefficients on y_0 in equation (7) will be 0.13 and 0.37 if $\rho = 0.98$ and 0.99, respectively. Such a process is also deceptive because the variations in y from one period to the next will be small and dominated by the stochastic term making it appear the process is in equilibrium.

High values of ρ may be characteristic of many political processes. In many policy areas, the influence of interest groups, the security of incumbents, the number of veto players, and bureaucratic inertia are likely to produce high values of ρ in any model of the policy or budgetary processes. Shortly, we consider whether the health expenditure data shown in Fig. 1 reflect a sticky process, path dependence, or something else.

3.3.2 Multiple and frequent shocks

In some processes, shocks to X may occur before y_t approaches its equilibrium. In the context of equation (7), assume that X moved from X_0 to X_1 in period one and a new value for X, denoted as X_2,

[8] Consider the situation with an ensemble with I series and let y_t^i be the value of y_t for the i^{th} series. Then, as $t \to \infty$,

$$\frac{1}{I}\sum_{i=1}^{I} y_t^i = X^*\beta = \mu.$$

It is also the case that

$$\lim_{T\to\infty} \bar{y}^i = \frac{1}{T}\sum_{t=1}^{T} y_t^i = X^*\beta = \mu,$$

for all series. This is the definition of ergodicity (Hamilton 1994) indicating that an equilibrium-independent process as shown in equation (6) is ergodic.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

occurs after period τ_1 rather than in period two such that $\prod_{s=1}^{\tau_1} \rho_s \neq 0$. This means the value for y_{τ_1}, immediately prior to the next shock, still reflects the initial conditions and early outcomes. We explore the time series implications of such a process both for equilibrium-independent and -dependent processes. We refer to such a process as near term path dependent as conditions that existed before X changed to the value X_1 may remain evident at and after time τ_1. With repeated shocks, however, even an equilibrium-dependent process will not appear so by casual observation.

The idea of near term path dependence is shown using the model in equation (7),

$$y_{\tau_1} = X_1\beta + \left(\prod_{s=1}^{\tau_1} \rho_s\right)(y_0 - X_1\beta) + v_{\tau_1}. \tag{8}$$

The higher the values of ρ_t and/or the shorter the number of time periods τ_1, the larger the influence of y_0 and early outcomes on y_t. Observations and analysis of y_t would lead to the conclusion that initial and early conditions still influence y. Observed values for y_t are consistent with both a sticky and an equilibrium-dependent process.

To continue, suppose that after τ_2 additional periods, there is another shock to X. The value for y at time $\tau_1 + \tau_2$, prior to this next shock, will be,

$$y_{\tau_1+\tau_2} = X_2\beta + \left(\prod_{s=1}^{\tau_2} \rho_s\right)(y_{\tau_1} - X_2\beta) + v_{\tau_1+\tau_2}. \tag{9}$$

Again, if the ρ_t are large and/or the number of periods τ_2 is short, the value for $y_{\tau_1+\tau_2}$ will incorporate its value at τ_1, not just the value for $X_2\beta$. If this is a process where X is continually being shocked at regular intervals and if the intervals are relatively short and/or the ρ_s are relative large, this will generate an interesting process where the values for y reflect the conditions each time X is shocked.

The value of the initial condition y_0 has a diminishing impact on y as the process continues. To see this substitute the expression for y_{τ_1} in equation (8) for $y_{\tau_1+\tau_2}$ in equation (9),

$$y_{\tau_1+\tau_2} = X_2\beta + \left(\prod_{s=1}^{\tau_1+\tau_2} \rho_s\right)(y_0 - X_1\beta) + \left(\prod_{s=1}^{\tau_2} \rho_s\right)(X_1 - X_2)\beta + u_{\tau_1+\tau_2}. \tag{10}$$

As this process continues for additional sequences of shock followed by periods of stability in X, the coefficients involving ρ_t on the terms $(y_0 - X_1\beta)$ and $(X_1 - X_2)\beta$ eventually approach zero unless $\rho_t = 1$ for all periods. The initial conditions, y_0, and early values of X have a decreasing effect on y_t. If, for example, there is a third shock to X after τ_3 more periods, the coefficient relating $y_{\tau_1+\tau_2+\tau_3}$ to $(y_0 - X_1\beta)$ is $\left(\prod_{s=1}^{\tau_1+\tau_2+\tau_3} \rho_s\right)$, which is clearly becoming smaller and smaller following each shock.

Equations (9) and (10) describe a process that is equilibrium independent and not equilibrium dependent with respect to initial conditions, y_0, and early values for X. Though this process is not equilibrium dependent, history matters, but ironically, it is recent and not early history that has the larger influence on current outcomes.

So far, there have been no conditions imposed on the values for ρ_t. What happens to the values for y_t implied by equation (10) if the requirements for equilibrium dependence are met, meaning there are periods where $\rho_t = 1$? Let $\rho_t = 1$ for $t_s^* \leqslant t \leqslant \tau_s$ so that after each shock, X is stable for τ_s periods and for the periods from t_s^* to τ_s, $\rho_t = 1$. This, along with a stochastic error correction term, means the process meets the requirements for equilibrium dependence for the period defined by τ_s.

It is easy to modify equation (8) to accommodate this condition,

$$y_{\tau_1} = X_1\beta + \left(\prod_{s=1}^{t_1^*} \rho_s\right)(y_0 - X_1\beta) + u_{\tau_1}. \tag{11}$$

With $t_1^* < \tau_1$, the values for y_0 are likely to be a large component of y_{τ_1}, particularly if t_1^* is small and/or the ρ_ts are large. As before, consider another shock to X, with X_2 lasting for τ_2 periods, with $\rho_t = 1$ for

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

the periods from t_2^* to τ_2. In this case, the equivalent version of equation (10) becomes,

$$y_{\tau_1+\tau_2} = X_2\beta + \left(\prod_{s=1}^{t_1^*}\rho_s\right)\left(\prod_{s=1}^{t_2^*}\rho_s\right)(y_0 - X_1\beta) + \left(\prod_{s=1}^{t_2^*}\rho_s\right)(X_1 - X_2)\beta + u_{\tau_1+\tau_2}. \quad (12)$$

With a third shock, the coefficient on $(y_0 - X_1\beta)$ is $\left(\prod_{s=1}^{t_1^*}\rho_s\right)\left(\prod_{s=1}^{t_2^*}\rho_s\right)\left(\prod_{s=1}^{t_3^*}\rho_s\right)$. Because $\rho_t < 1$ for each $t < t_s^*$, this coefficient decreases to zero as more shocks accumulate. This is exactly the outcome in equation (10) if all values of $\rho_t < 1$. The implication here is that a process that is equilibrium dependent following a single shock to X becomes only equilibrium independent as shocks accumulate.

3.4 *Contrasts among Dynamic Processes*

To reiterate, we think of processes as falling into two broad categories, outcome independent and outcome dependent. And then, within the latter (outcome dependent) category, we can consider path (equilibrium) dependent and equilibrium independent. Equilibrium independence includes what we have called sticky processes and near term path dependence. Let us now graphically illustrate the distinctions among the outcome-dependent categories by discussing and plotting the ensemble means, $\bar{y}_t = \sum_{i=1}^{I} y_t^i$ with $I = 40,000$, for four different processes. In all processes, there are three series with different starting values for X at t_0 of 0, 6, and 10, respectively. The processes are all variations on the basic model in equation (7), with the variations being in the values for ρ_t and the sequence for X. In two instances, $X = 5$ in period one and this value is retained for 50 periods. In two other processes, X follows the sequence $\{1, 2, 3, 4, 5\}$ retaining each value for 10 periods, meaning that all series end with the same values for X. Figure 2 shows the equilibrium-dependence process where ρ_t follows the sequence $\rho_t = \{0.6, 0.7, 0.8, 0.9, 0.95\}$ for $t \leqslant 5$ and then, $\rho_t = 1$ for all $t \geqslant 6$. X equals five for the whole series. As can be seen in the figure,

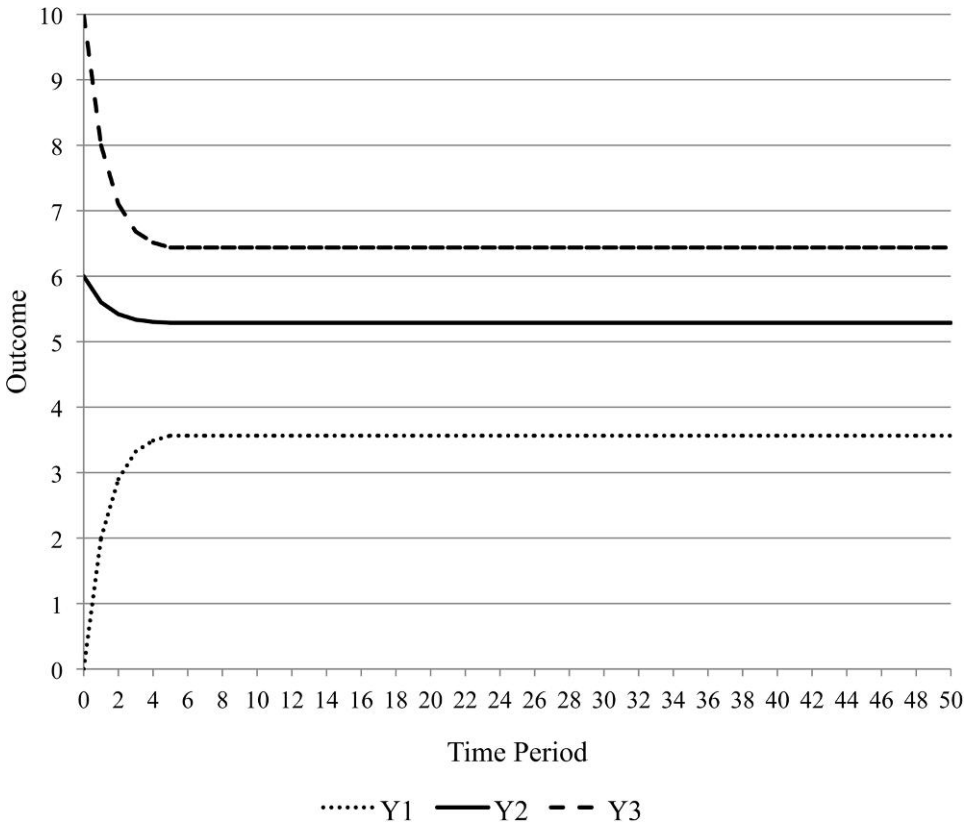


Fig. 2 Equilibrium-dependent process.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

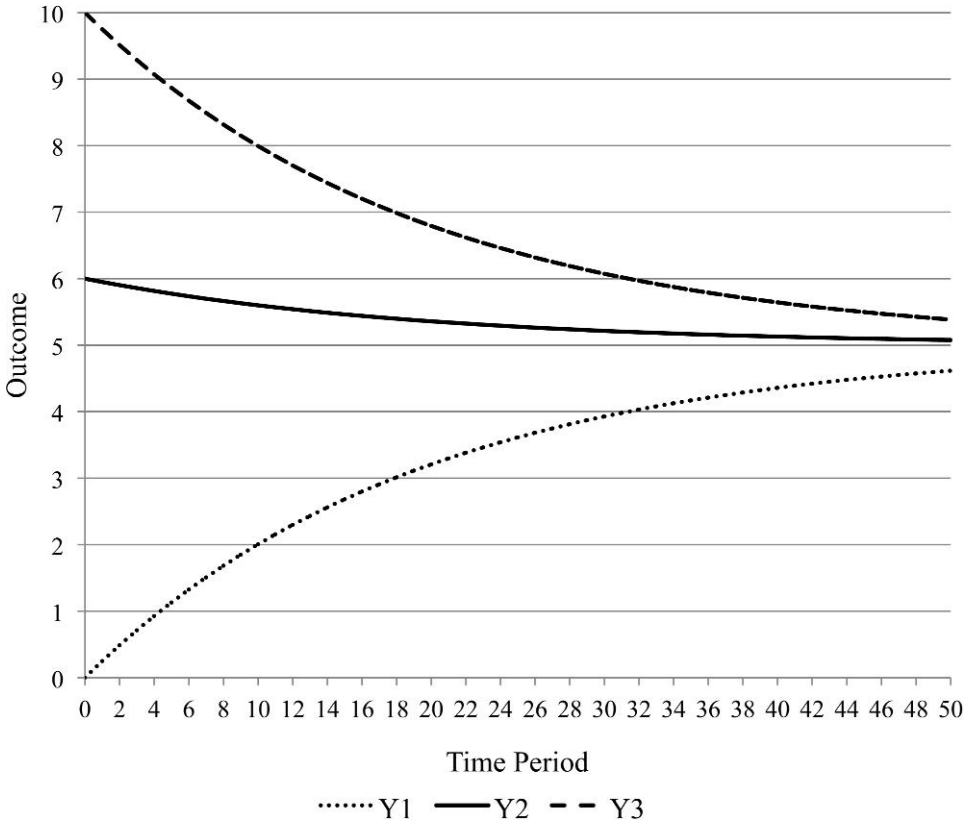


Fig. 3 Sticky process.

the differences in starting conditions remain evident in the values for $\bar{y}_{50}$ for each case. Figure 3 depicts what we have called a sticky process, with high values for ρ_t but always less than one. In this example, $\rho_t = 0.95$ for all t. Even after 50 periods, the effects of the different initial conditions are still evident, though the processes are converging to a value of five but slowly. After 50 periods, the values for $\bar{y}^i_{50}$ are 4.61, 5.08, and 5.38, and even after 100 periods, the values are 4.97, 5.01, and 5.03 for each respective $\bar{y}^i_{100}$.

Figure 4 depicts a process that is equilibrium dependent but with frequent shocks, as in equation (12). The process satisfies the conditions for equilibrium dependence for the 10 periods where X is stable, meaning that it is comparable to Fig. 2. In fact, the processes in Figs. 2 and 4 are identical for the first 10 periods. In Fig. 4, however, the value for X shifts from one to two and the values for ρ_t again follow the sequence $\rho_t = \{0.6, 0.7, 0.8, 0.9, 0.95\}$ for $11 \leqslant t \leqslant 15$ and $\rho_t = 1$ for $16 \leqslant t \leqslant 20$. This process is repeated for the sequence of values for X shown above. Within each 10-period interval, the process is equilibrium dependent but as Fig. 4 shows the series converge to the same value after about period 30. The differences in initial conditions are no longer related to the values for y for the three cases. These processes now resemble equilibrium independence and not equilibrium dependence. Compare Figs. 4 and 5. In Fig. 5, the processes are identical to those in Fig. 4, except that $\rho_t = 0.95$ for the last 5 periods of each cycle instead of $\rho_t = 1$ for these periods. Although the difference is subtle, it matters. The difference between the processes in the two figures is that if there are no more changes to X after period 50, the series in Fig. 4 will remain at the values of 4.60 shown in the figure regardless of how long the process is extended. For the processes in Fig. 5, the outcomes will eventually converge to the value of five, matching the value of X. With repeated shocks, the process will remain similar.

4 Empirical Challenges

The empirical challenge is to distinguish among these processes—is a process equilibrium, or path, dependent; is it near term path dependent; or equilibrium independent with a slow rate of adaptation? Each

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

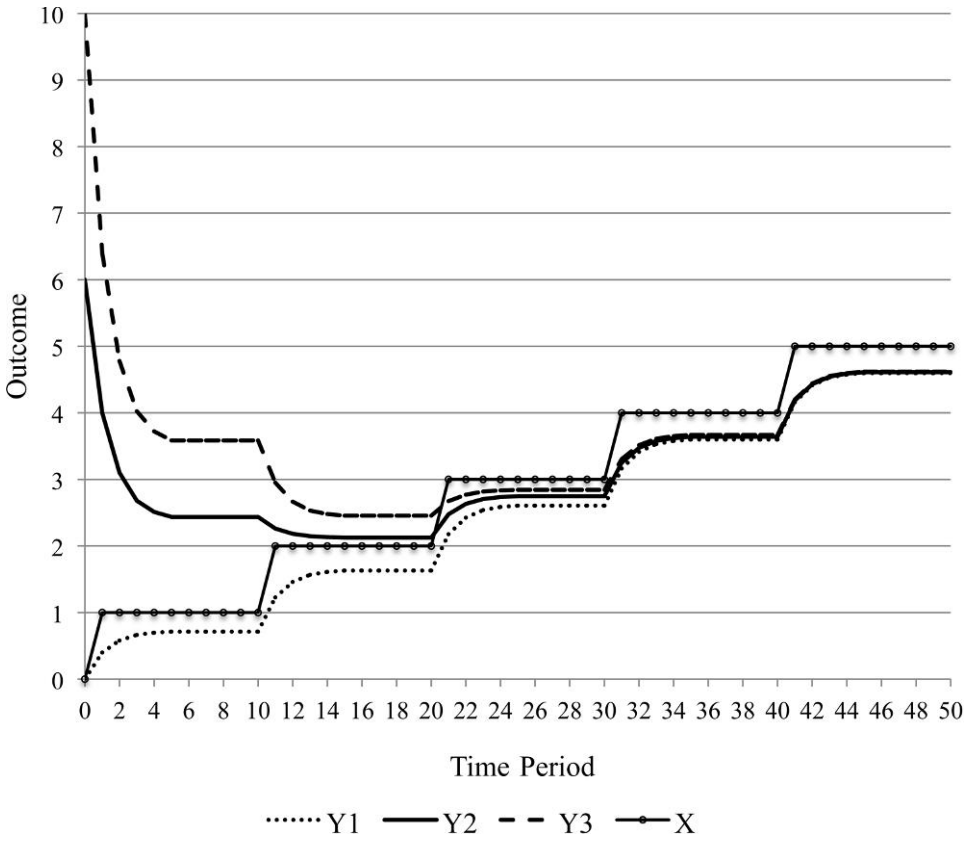


Fig. 4 Equilibrium-dependent with shocks.

has different substantive implications that can be quite important for prediction. Heckman (1991) raises the same question in the context of distinguishing the effects of initial conditions from heterogeneity in duration models. He concludes that externally imposed restrictions, such as a functional form, are required, which will be similar to our conclusion. The challenge resembles, at least on a conceptual level, a classic case of statistical inference. The analyst must decide to accept or reject the hypothesis that a process is path dependent. Rejection when a process is in fact equilibrium (path) dependent is akin to a type I error, whereas accepting this proposition when it is false is a type II error. Avoiding both types of errors may be extremely difficult for most commonly available data.

Observing a series of outcomes for y_t, even over an extended time period, may be insufficient to accept or reject the claim that they are generated by a path-dependent process. In fact, such a series may be misleading. Figure 3 depicts an equilibrium-independent series with high values of ρ_t, a sticky process. Many periods after X was shocked, there are clear differences in the values for y_t related to initial conditions for each series. Furthermore, for extended periods, the series may look like it is in equilibrium as there are small period-to-period changes in the expected value of y_t. The observed changes may be indistinguishable from the random fluctuations one expects in any outcome variable. Observing these two characteristics in a data set could lead an analyst to accept the proposition that this is an equilibrium-dependent process and make a type II error. Analysis with cross-sectional data, such as a sample of countries, states, individuals, etc. observed in a single time period, is particularly susceptible to this type of error. In a regression of values for y_{it} on the values for the exogenous X_{it} and the initial condition y_{i0}, the coefficient on y_{i0} will be different from zero unless t is quite large and/or ρ is small. This leads to the inference that initial conditions remain important, and thus, the process is equilibrium or path dependent.

Now consider Fig. 4. Here, each process is equilibrium dependent for a single shock to X, but when the outcomes after an extended time period and a sequence of shocks are compared, the effects of the initial differences at y_0 are no longer apparent. In addition, with frequent shocks, the series are not in equilibrium long enough to discern path dependence. Here, the analyst is most likely to observe these

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

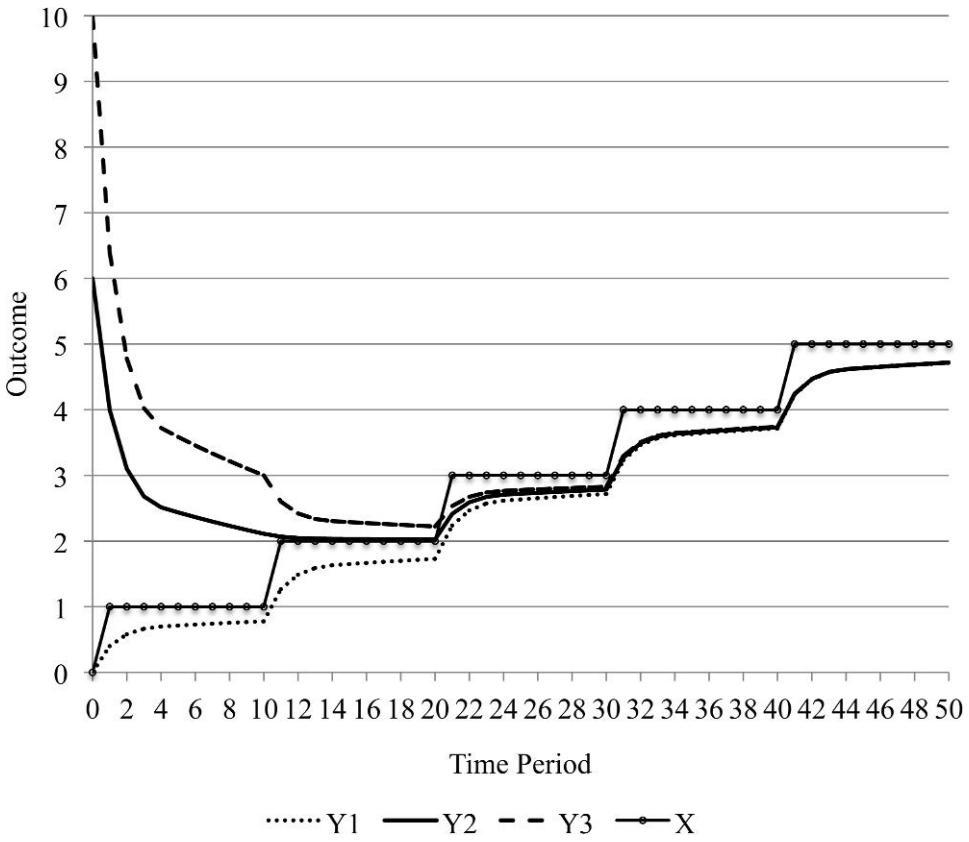


Fig. 5 Equilibrium-independent with shocks.

processes and infer they are not equilibrium dependent, committing a type I error. Again, cross-sectional data are particularly susceptible to this error if there have been many changes in X_{it} since the initial conditions were observed.

Two conditions may enable inferences about equilibrium dependence from analysis of cross-sectional data. First, one must have a defensible causal argument about why $\rho_t = 1$ for some process for an extended time period. Second, the empirical analysis should demonstrate that initial conditions still matter after a very long historical period. A sticky process can always be consistent with this result, but the longer the initial conditions remain important, the more credible the conclusion of equilibrium dependence. The economic performance of former European colonies, analyzed in Acemoglu et al.'s, Engerman and Sokoloff's, and Woodberry et al.'s cross-sectional studies discussed above, maybe such an example. Their implicit model for political institutions meets the criterion for ρ_t, and their empirical evidence is that differences in initial conditions remain evident in the cross-sectional comparisons centuries later. But, these conditions are at best reducing the possibility of a type I error.

A safer method for inferring the properties of a dynamic process is with time series data. We see two requirements for any analysis and assessment. First, one must have a well-articulated and defensible causal model. This model must be explicit about what variables are exogenous (the X); how these variables are related to the outcomes of interest (the y_t); and how outcomes at t are related to outcomes at $t-1$ (the ρ_t). The model must also specify a function for ρ_t, the parameters of which are defined so that specific values are consistent with equilibrium dependence. These parameters identify if and when $\rho_t = 1$. This model specification fits Freeman's (2011) definition of a nonlinear time series model, which he notes may be equilibrium dependent.

The second requirement is extensive time series data with which one can estimate the model's parameters. The causal model relating y_t to X is likely to be complex and nonlinear, and with a parametric expression for ρ_t, the full model is certain to be nonlinear in the parameters. This necessitates a nonlinear estimator such as nonlinear least squares. One must be able to argue for the eventual stability in the X

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

variables in order to distinguish true path dependence, where y_t has an equilibrium mean that includes the effects of initial and early conditions, from situations where continual shocks to the X's preclude attainment of this path-dependent equilibrium distribution. The existence of this stability condition might be defended ex ante with justifiable claims that X is and will be stable in the future, but of course, this is speculation. A sounder approach is to show that for an extended period within the data that X has been constant, or at least varying in a stationary pattern.

Jackson and Kollman (2010) illustrate this strategy and estimation procedure with both simulated data and a model of macropartisanship. Their model is,

$$y_t = \rho_t y_{t-1} + (1-\rho_t) X_t \beta - \delta u_{t-1} + \epsilon_t, \text{ where} \tag{13}$$

$$\rho_t = \alpha_1 e^{-\alpha_2[(\Delta X_t)^2 + (\Delta X_{t-1})^2]}. \tag{14}$$

Their model illustrates the empirical strategy recommended here. They present causal arguments for how X_t is defined and measured and for why it is related to y_t. The particular values for α_1 and α_2 distinguish between equilibrium dependence and equilibrium independence, with the former occurring when $\alpha_2 > 0, \alpha_1 = 1$, and $\Delta X_t = \Delta X_{t-1} = 0$. They demonstrate with the simulations that a nonlinear least squares estimator does well estimating the distribution of coefficients generated in the simulations.

Yet, the Jackson and Kollman simulations also demonstrate the inherent difficulty in distinguishing equilibrium and very sticky equilibrium-independent processes. In their simulations as α_1 gets closer to one, the proportion of type II errors increases, as one would expect, meaning it is becoming harder to distinguish the two processes. They also show that reducing the likelihood of these type II errors increases the chance of a type I error if the process is equilibrium dependent. This is the formal statistical version of the larger issue of type I and type II errors that began this section. In sum, a well-developed explicit dynamic model, extended time series data, and careful estimation are necessary conditions for distinguishing equilibrium dependence from equilibrium independence, but they are not themselves sufficient.

A better strategy would be to have a stationary time series for a set of cross-sectional units, as in pooled time series cross-section (TSCS) models. Such data provide the temporal variations necessary to model ρ_t and the stationary error correction term and to test whether the criteria for equilibrium dependence are met. They also provide the cross-sectional variations necessary to compare equilibrium outcomes and to relate them to values of the exogenous variables. There are parallels between the structure of pooled TSCS data and the notion of an ensemble used here in developing the definition of equilibrium dependence. These similarities are exploited in the methodological strategy proposed here to help distinguish between equilibrium-independent and -dependent processes. An analysis of such data requires combining the methods common to pooled TSCS studies with the nonlinear estimation used by Jackson and Kollman (2010). We are unaware of such a study but think it holds considerable promise for testing for path dependence.

The difficulty of distinguishing different processes is well illustrated with an extension of the health expenditure data shown in Fig. 1. Figure 6 extends these series to the period 1960–2007. What should be clear is that a process that might have appeared path dependent in Fig. 1 may appear less so when viewed over this longer period. The plots of public expenditures as a proportion of all expenditures now suggest a more convergent path, as indicated by the declining difference between the U.S. share and the mean of the other shares and by the diminishing standard deviation of all countries' shares. Is this an equilibrium dependent, sticky, or a near term path-dependent process? One might argue that these are simply very sticky processes, given the political institutions and societal pressures known to be in force in those countries. Alternatively, the process, in the United States at least, might be described as one of equilibrium dependence but with a series of shocks to the existing equilibrium, as depicted in Fig. 4. These shocks, for instance, would be the passage of Medicare in 1965, which dramatically boosted the public share from 1967 on, and the expansion of coverage and benefits to Medicaid in 1991 and 1992. Presumably, when the series is extended past 2013, there will be another shock to the equilibrium when the effects of Bush's prescription drug legislation and Obama's health care legislation take full effect.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

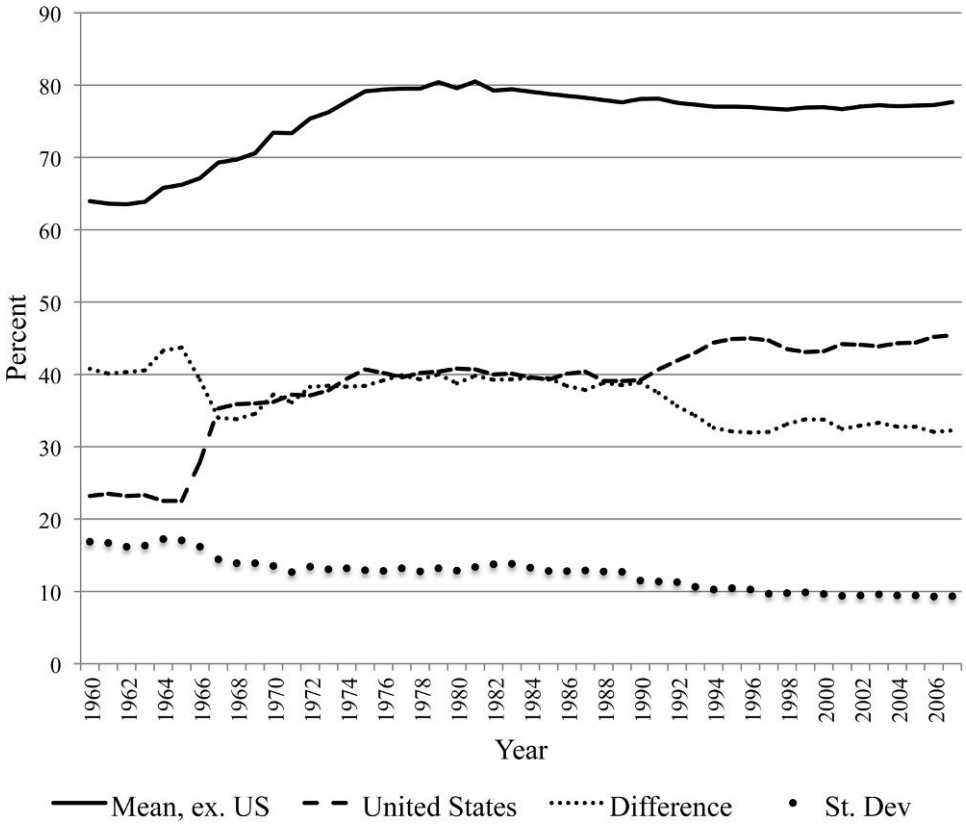


Fig. 6 Public health expenditures/total health expenditures.

4.1 *Empirical Illustration*

We now illustrate a methodological approach using pooled time series and cross-sectional data and show how it might offer some insight into the type of process represented by the data in Fig. 6. Twelve of the OECD countries in Fig. 6 have data on the public share of health expenditures that span the period 1960–2007.[9] These data are used to estimate an equation similar to equation (13) where the outcome variable, y_{it}, is country i's public share of all health expenditures in year t. What would be desirable would be exogenous variables, X_{it}, that are related to countries' public share of health expenditures, such as age and income distributions and income levels. These could then be used to model a country's equilibrium share. Without such data and a theory underlying them, we cannot engage in a credible analysis of why different countries have different public health care expenditures. But for the sake of demonstrating a proposed method and for illustrating different dynamic processes, each country's 1960 share, labeled y_{i0}, and its interaction with time, proxy these equilibrium shares. The interaction of initial conditions with time also estimates how persistent these early conditions are in subsequent years' shares.

The model for Y_{it}^*, a country's expected equilibrium share at time t, is

$$Y_{it}^* = \beta_1 + \beta_2 e^{-\beta_3(t-1961)/10} y_{i0}. \tag{15}$$

The coefficient β_1 indicates the long-term expected equilibrium share. The second part of the expression relates each country's deviation from this equilibrium value at time t to its share in 1960. The coefficient β_3 indicates if or how rapidly that association declines over time. If $\beta_3 = 0$, differences in initial shares persist in perpetuity. The larger the value of β_3, the more rapidly these effects dissipate and the more

[9]Data for Denmark, Germany, New Zealand, Portugal, and Sweden are missing for the 1960–1970 period. In addition, for three countries, Australia, Canada, and France, there are data for 1960 but there are gaps in the series that were filled with a linear interpolation.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

rapidly all countries' shares converge on the equilibrium-expected value of β_1. The coefficient β_2 indicates how much these initial conditions influence the equilibrium mean shares at t for small values of β_3 and/or for the period right after 1960.

The autoregressive coefficient, ρ_t, is specified to vary with time,

$$\rho_t = \alpha_1 + \alpha_2 e^{-\alpha_3(t-1961)/10}. \tag{16}$$

This is a very simplified version of a possible equilibrium-dependent process, where ρ must vary over time. Here, ρ can only increase or decrease with time or remain constant. If $\alpha_2 = 0$, then $\rho_t = \alpha_1$ for all t, which violates the conditions for equilibrium dependence. If $\alpha_2 > 0$, ρ_t decreases with time and its maximum value is $\alpha_1 + \alpha_2$, which is constrained to be less than or equal to one. If $\alpha_2 < 0$, ρ_t increases with time and its maximum value is α_1, which is also constrained to be less than or equal one.

The model estimated combines these two components along with the stationary error correction term denoted by $u_{i,t-1} = [y_{i,t-1} - E(y_{i,t-1})]$ in equation (13). The model also includes country fixed effects, denoted by the full set of country dummy variables C_i, which accommodate the pooled time series cross-sectional data. Continuing the notation in equations (15) and (16), the estimated model is,

$$\begin{aligned} y_{it} &= \rho_t y_{i,t-1} + (1-\rho_t)Y^*_{it} - \delta u_{i,t-1} + C_i D + \epsilon_{it} \\ &= \rho_t y_{i,t-1} + (1-\rho_t)Y^*_{it} - \delta[y_{i,t-1} - \rho_{t-1} y_{i,t-2} - (1-\rho_{t-1})Y^*_{i,t-1} - C_i D] + C_i D + \epsilon_{it} \\ &= \rho_t y_{i,t-1} + (1-\rho_t)Y^*_{it} - \delta[y_{i,t-1} - \rho_{t-1} y_{i,t-2} - (1-\rho_{t-1})Y^*_{i,t-1}] \\ &\quad + C_i D(1+\delta) + \epsilon_{it}. \end{aligned} \tag{17}$$

Substituting equations (15) and (16) into equation (17) gives a model that is nonlinear in the parameters, which is estimated using nonlinear least squares. (See Davidson and MacKinnon 1993, 2004.) In this application, the ϵ_{it}s are likely heteroskedastic suggesting a heteroskedastic robust method for calculating the estimated standard errors (Davidson and MacKinnon 2004, 243).

Before discussing the results, we stress that this is intended as an illustration of how one might analyze the processes shown in Fig. 6. The model and data fall short of what is required for a substantive analysis of how health care is funded in OECD countries. Besides the omission of variables related to countries' equilibrium public shares and how those values change over time, there are other problems with the model. For example, no consideration is given here to questions of functional form that would incorporate obvious constraints, such as that the public share cannot exceed one or that if ρ_t varies temporally, it is unlikely to do so in the monotonic way specified here. It is valuable, however, to see how one might use the type of pooled time series cross-sectional data available here and how to interpret results in the context of distinguishing between equilibrium independence and dependence.

Table 1 shows the estimated results, which satisfy two of the conditions for equilibrium dependence but not the most important one. The values for ρ_t do vary temporally, ranging from 1.00 in 1961 and 0.972 in 1962 to 0.862 in 1977 and thereafter. (Time in equations (15) and (16) is measured in decades.) The coefficient on the lagged error term is different from zero and implies that the type of error correction required for path dependence is present, though adjustments occur slowly.[10] The missing condition is that the value of ρ_t is less than one for every year but 1961.[11] These results are consistent with what we call a sticky but not equilibrium-dependent process. Further evidence of this stickiness is that the relationship between early conditions and the current share is decreasing but at a very slow rate. The coefficient on y_{i0} only decreases from 0.75 in 1961 to 0.62 in 2007. Lastly, the very long-run equilibrium value is 0.38, but with high values for ρ_t and the very slowly declining effect of initial shares, it will take an exceptionally long time for most countries to reach this share.

We extend this basic model to consider if the United States is different from other countries, other than its share in 1960. This is done by allowing β_1, the equilibrium share, and β_2, the effect of initial shares, to

[10]The coefficient on the lagged error term is only 1.72 times its standard error, but the likelihood ratio test with a p value less than .005 rejects the null that the coefficient is zero.

[11]Variations on the functional form that permitted $\rho_t = 1$ for additional years were explored but the fit degraded badly after only a few years.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

Table 1 Estimated public share

	Base model		*U.S. exception*			
Coefficient	*Value*	*Standard error*	*Value*	*Standard error*	*Value*	*Standard error*
ρ_t						
α_1	0.862[a]	–	0.844[a]	–	0.844[a]	–
α_2	0.138	0.030	0.156	0.039	0.156	0.039
α_3	−2.097	1.053	−1.754	0.870	−1.753	0.869
Y^*_{it}						
β_1	0.380	0.081	0.524	0.100	0.522	0.085
β_1*US			−0.007	0.217		
β_2	0.752	0.136	0.529	0.151	0.531	0.133
$\beta_2 *$ US			−0.975	0.877	−1.006	0.381
β_3	−0.043	0.015	−0.063	0.026	−0.063	0.023
Error correction						
δ	0.130	0.076	0.127	0.076	0.127	0.076
Log likelihood	1424.90		1426.36		1426.36	

[a]$\alpha_1 + \alpha_2$ constrained to sum to one.

vary for the United States. The results are shown in the second column of Table 1, and they improve the model is several ways. First though, there is no difference between the United States and other countries in their long-run equilibrium share. This value is now estimated to be 0.52, which seems more realistic than the value of 0.38 in the first equation. The U.S. difference from this value is less than 0.01 with a standard error over 0.2. Where the United States is different is in the relationship between current shares and initial conditions. For non-U.S. countries, the value of $\beta_2 = 0.53$ indicates that current shares are higher than the long-run equilibrium by an amount related to these initial shares. This amount ranges from 0.53 times the initial share in 1962 to 0.40 in 2007. In the United States, by contrast, the effective value for β_2 is -0.45, which indicates that current shares are less than the long-run equilibrium.[12] With an initial share of 0.23, this difference is 0.12 in 1962 and 0.08 in 2007.

The relationship between current shares and initial shares declines more rapidly in this model than in the previous one, but with $\beta_3 = -0.063$, it still takes a very long time for the effects of initial conditions to dissipate. The results suggest there is nothing particularly exceptional about the United States except its starting point in 1960 and how that value compares to the long-run equilibrium share.[13] If the process were to continue without any shock and if the results in Table 1 can be extrapolated into the future, each country in this model will converge to an equilibrium mean of 0.52 that is independent of initial conditions and of any historical path. But this convergence is very, very slow. After 200 years, the United States is 3% below and the average of the other countries is 10% above this equilibrium mean.

This exercise is an illustration of how one might use sophisticated estimation methods to distinguish among processes. The model estimated here exemplifies an equilibrium-independent sticky process, in which initial differences are apparent for a long time when one compares different countries. Given the limitations of the data and the model expressed above, we are not claiming this is a good representation of a model of national health care spending. Also, without exogenous variables that change over time, we cannot compare these results with those we called near path dependence as we cannot observe repeated shocks and what happens after those shocks. Furthermore, we need a more sophisticated model for ρ_t to examine if it equals one for shorter periods during the process, which could imply near path dependence.

The results illustrating a very sticky process provide an instructive contrast to the model estimated in Jackson and Kollman (2010), which they say fits the conditions for equilibrium dependence, but more

[12]The standard error of the U.S. difference in the value of β_2 is high due to the substantial collinearity between the two U.S. interaction variables. The last two columns in Table 1 shows the estimates with the U.S. adjustment to β_1 omitted.

[13]The country fixed effects coefficients do not explain a small public share for the United States. In fact, the mean fixed effect coefficient is −0.0044 and the value for the United States is 0.0009.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

likely is an equilibrium-dependent process with repeated shocks. As a methodological illustration, this example works well to demonstrate the potential power of pooled time series cross-sectional data and the type of models that are consistent with different forms of equilibrium independence. We want to use the example to encourage future researchers to think carefully about the conditions that distinguish outcome from equilibrium dependence and about the types of data and estimation strategies needed to make these distinctions.

5 Concluding Discussion

People often observe patterns in outcomes and call the process path dependent. We argue, as have others before (e.g., Page 2006), that for equilibrium, or path, dependence, these processes must meet very stringent conditions, which means that true path dependence may occur rarely. With careful and proper analysis, many processes that are called path dependent may actually be equilibrium independent. This is a very broad set of processes with critical distinctions among the different elements—sticky and near term path-dependence processes are qualitatively and substantively different. On closer examination, there are likely to be other distinct patterns that warrant detailed examination. These distinctions are not just about semantics. Whether a process is equilibrium dependent or equilibrium independent has important substantive and empirical implications. The political and economic transition example demonstrates that the determination of path dependence versus equilibrium independence could have profound policy consequences. Use of the terminology of path dependence without a well-grounded definition of an equilibrium and supporting empirical evidence leads to incoherent theories and poor predictions about future trajectories. Conversely, ignoring path dependence may lead to badly biased empirical results in empirical studies and wrong predictions. Future research could focus on exploring what we call equilibrium-independent processes, estimating parameters for (1) stickiness, ρ_t, which determine the trajectory and convergence rate for y_t, and (2) the ratio between this convergence rate and the frequency of shocks to X. Such parameterizations provide at least a sense of how far a process is from equilibrium dependence along the spectrum defined in this article. We recognize the strictness of the definition used here; it limits people to calling nearly everything we study to equilibrium independence. That may have to be the price for rigorous theorizing and appropriate empirical inferences.

Appendix A

This appendix demonstrates the conditions shown in equation (5). Begin with

$$y_t = X_1\beta + \rho^t(y_0 - X_1\beta) + \sum_{s=0}^{t-1} \rho^s u_{t-s}.$$

Summing y_t/T for $t = 1, \ldots, T$ and taking the limit as T $\to \infty$ gives

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} y_t = \lim_{T\to\infty} \frac{1}{T}\left[\sum_{t=1}^{T} X_1\beta + \left(\sum_{t=1}^{T}\rho^t\right)(y_0 - X_1\beta) + \sum_{t=1}^{T}\left(\sum_{s=0}^{t-1}\rho^s u_{t-s}\right)\right]$$

$$= X_1\beta + \lim_{T\to\infty}\left[\frac{1}{T}\cdot\frac{\rho}{1-\rho}\cdot(y_0 - X_1\beta)\right] + \lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{s=0}^{T-t}\rho^s\right)u_t\right].$$

The second term on the right side goes to zero as T $\to \infty$. The properties of the third term are not so obvious. The key is that as T $\to \infty$, this expression collapses to zero in probability, just as the sample mean becomes the population mean in probability as the sample size approaches infinity.

To simplify notation, let $a_t = \sum_{s=0}^{T-t}\rho^s$ for $t = 1, \ldots, T$ and with $1 \leqslant a_t < \frac{1}{1-\rho}$. The summation in brackets can be written as $S_T = \sum_{t=1}^{T} a_t u_t$. If the u_t are i.i.d. with mean zero and variance σ_u^2, then S_t is the sum of T independent random variables with mean zero and finite variance $a_t^2\sigma_u^2$. The strong law of large numbers (Feller 1968, 259) says as T $\to \infty$, this sum divided by T converges to zero with probability one.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press

With the second and third terms going to zero, we have

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} y_t = X_1\beta = E(y_T),$$

satisfying the conditions for a steady-state equilibrium.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5):1369–401.

———. 2002. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics* 117(4):1231–94.

Beck, Nathaniel. 1983. Time-varying parameter regression models. *American Journal of Political Science* 27(3):557–600.

Davidson, Russell, and James G. MacKinnon. 1993. *Estimation and inference in econometrics*. Oxford, UK: Oxford University Press.

———. 2004. *Econometric theory and methods*. Oxford, UK: Oxford University Press.

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

Enders, Walter. 2010. *Applied econometric time series*. 3rd ed. Hoboken, NJ: John Wiley & Sons.

Engerman, Stanley L., and Kenneth L. Sokoloff. 2002. Factor endowments, inequality, and paths of development among new world economies. *Economia* 3(1):41–88.

Feller, William. 1968. *An introduction to probability theory and its application.* Vol. 1. 3rd ed. Hoboken, NJ: John Wiley & Sons.

Freeman, John R. 1989. Systematic sampling, temporal aggregation, and the study of political relationships. *Political Analysis* 1:61–98.

———. 2011. Research note: Can time series methods be used to detect path dependence? Revised version of a paper presented at the Conference on Path Dependence. University of Minnesota, Minneapolis, MN. June 4 & 5, 2010. http://www.polisci.umn.edu/freeman/pathdependence.pdf.

Greene, William H. 2008. *Econometric analysis*. 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Grzymała-Busse, Anna. 2007. *Rebuilding Leviathan.* Cambridge, UK: Cambridge University Press.

Hacker, Jacob S. 2002. *The divided welfare state*. Cambridge, UK: Cambridge University Press.

Hamilton, James D. 1994. *Time series analysis.* Princeton, NJ: Princeton University Press.

Heckman, James J. 1991. Identifying the hand of the past: Distinguishing state dependence from heterogeneity. *American Economic Review* 81(2):75–9.

Hellman, Joel S. 1998. Winners take all: The politics of partial reforms in Post-communist transitions. *World Politics* 50(2):203–34.

Jackson, John E. 1996. Political methodology: An overview. In *New handbook of political science*, ed. Robert Goodin and Hans-Dieter Klingemann. Oxford, UK: Oxford University Press.

Jackson, John E., and Ken Kollman. 2010. A formulation of path dependence with an empirical example. *Quarterly Journal of Political Science* 5(3):257–89.

Jackson, John E., and Bogdan W. Mach. 2011. Statistics: Overview. In *International encyclopedia of political science*, ed. Bernard Badie, Dirk Berg-Schlosser, and Leonardo Morlino. Thousand Oaks, CA: Sage Publications.

Kuh, Edwin. 1959. The validity of cross-sectionally estimated behavior equations in time series applications. *Econometrica* 27(2):197–214.

Nelson, Joan M. 1993. The politics of economic transformation. *World Politics* 45(2):433–63.

Organization for Economic Cooperation and Development. 2010. *OECD health data 2010*. Paris, France.

Page, Scott E. 2006. Path dependence. *Quarterly Journal of Political Science* 1:87–115.

Roland, Gérard. 2000. *Transition and economics: Politics, markets, and firms*. Cambridge, MA: MIT Press.

Sachs, Jeffrey D., and Katharina Pistor. 1997. Introduction: Progress, pitfalls, scenarios, and lost opportunities. In *The rule of law and economic reform in Russia*, ed. Jeffrey D. Sachs and Katharina Pistor. Boulder, CO: Westview Press.

Woodberry, Robert, William Roberts Clark, and John A. Doces. 2011. *Protestant missionaries and economic growth*. Ann Arbor, MI: Department of Political Science, University of Michigan.

https://doi.org/10.1093/pan/mpr050 Published online by Cambridge University Press