Higher-Dimension Markov Models

David L. Epstein and Sharyn O'Halloran
Columbia University*

Abstract

Markov transition models are becoming a popular tool for exploring the dynamics of systems that can take on a finite number of states. However, their application in political science has thus far been mostly limited to the two-state case. This paper explains the techniques necessary to estimate and interpret higher-dimension Markov models. We then apply them to the study of democratic transitions, where we find that a three-state model including an intermediary "partial democracy" category outperforms the previous two-state model of Przeworski, et. al. (2000).

1 Introduction

Markov models are used in situations where a system can exist in any of a finite number of states in each period, and we wish to estimate the probabilities of its transitioning from state a at time $t-1$ to state b at time t. These models are well-known in both biometrics and econometrics (for example, Amemiya 1985; Ware, Lipsitz and Speizer 1988), and are becoming a popular method within political science for investigating time-varying processes with qualitative dependent variables. Jackman (2000), for example, provides a number of examples of their application, including delegation, war, and trade pacts.

Up until now, though, the use of Markov models in political science has been mostly limited to the simple, dichotomous case.[1] These are sufficient for many applications, but one can imagine others where the system being studied has more than two possible states. Relations between countries, for example, might be friendly, tense, or violent; members of Congress might face no challenger, a weak challenger, or a strong challenger; parliaments might be controlled by right-wing coalitions, left-wing coalitions, or minority governments; or (the example on which this paper will concentrate) a country may be autocratic, partially democratic, or fully democratic.

*Paper prepared for presentation at the Political Methodology Summer Meetings, Tallahassee, Florida, July 21-23, 2005. Preliminary draft; comments welcome. Thanks to the Russell Sage Foundation and the National Science Foundation for financial support.

[1]Exceptions include Dean and Moran (1977); Jones, Kim and Starz (2005); and Walker (2005).

The treatment of these higher-dimension Markov settings differs in many respects from the simpler 2-state case. This paper has as its primary aim the explanation of these differences, and it supplies some techniques for estimating higher-dimension processes. We then apply these models to a subject area of particular interest: democratic transitions. We first show that the currently most sophisticated analysis of these transitions, a dichotomous Markov model by Przeworski, Alvarez, Cheibub and Limongi (2000, hereafter referenced "PACL"), is riddled with errors. We then show that, even once those errors have been corrected, a three-by-three Markov model which incorporates a middle category of "partial democracies" performs better than their two-by-two model of autocracy and democracy. The analysis not only offers a nice illustration of the power of higher-dimension Markov models, it also provides important substantive insights into a scholarly debate with significant real-world implications; namely, how countries can move along the path to democracy, and when such transitions will be stable.

2 The Mechanics of Higher-Dimension Markov Models

Markov processes model transitions of a system from one state to another. Assume that there are C ordered categories of the dependent variable, labeled $0, 1, \ldots, C-1$.[2] The first-order Markov assumption is that, conditional on the state of the system at time $t-1$, the transition events are uncorrelated, so that ordinary logistic regression can be used to estimate regression coefficients and their standard errors.

If the transition probabilities from state a to state b are labeled π_{ab}, where $a, b \leq C-1$ and $\sum_b \pi_{ab} = 1$, then the simplest Markov process consists of a two-state system:

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} \end{array},$$

where the rows give the state of the system at time $t-1$ and the columns are the state at time t.

Using a logit link, this two-state case could be estimated by a single regression:

$$\Pr(Y_t = 1) = \text{Logit}(X_{t-1}\beta). \tag{1}$$

This formulation, though, rests on the assumption that the factors moving the state from 0 to 1 are equal and opposite from those that move it from 1 to 0. Clearly, this need not be the case; religious factionalism may start ethnic wars, for instance, while international intervention is required to stop them. Or good economic conditions might foster transitions out of autocracy, while group-based politics and violence trigger reversals to dictatorship. In any event we would want to test whether this assumption of equal and opposite effects is true, rather than assume that it holds *a priori*.

The Markov approach is to estimate the system by a pair of logit regressions, each depending explicitly on the prior state of the system:

$$\Pr(Y_t = 1 | Y_{t-1} = 0) = \text{Logit}(X_{t-1}\beta) \tag{2}$$
$$\Pr(Y_t = 1 | Y_{t-1} = 1) = \text{Logit}(X_{t-1}\alpha) \tag{3}$$

[2]Markov models can also be applied to unordered dependent variables; we abstract from these considerations in the present paper.

which can be written more compactly as

$$\Pr(Y_t = 1) = \text{Logit}(X_{t-1}\beta + Y_{t-1}X_{t-1}\gamma). \tag{4}$$

where

$$\gamma = \alpha - \beta. \tag{5}$$

It is clear that the original α coefficients can be recovered from this estimation equation as $\beta + \gamma$. The advantage of combining the two equations into one is that the γ terms provide information on whether a given independent variable does in fact have a different effect when moving from 0 to 1 as opposed to moving from 1 to 0. In particular, if all the γ terms are insignificant, then one can run the system using Equation 1.

For a specific example, let us think of relations between the dyad of countries i and j. At time t these relations can be friendly ($Y_{ijt} = 0$) or unfriendly ($Y_{ijt} = 1$). Our one explanatory variable will be the amount of bilateral trade (BT) between the countries. The estimated Markov equation would then be:

$$\Pr(Y_t = 1) = \text{Logit}(\beta_0 + \beta_1 BT_{t-1} + \beta_2 Y_{t-1} + \beta_3 Y_{t-1} BT_{t-1}). \tag{6}$$

Here, β_1 gives the impact of trade on the probability of developing hostile relations (i.e., relations are currently friendly, so $Y_{t-1} = 0$); β_2 is the impact of current hostile relations on the probability of hostile relations next period, and β_3 is the *difference* of the impact of trade on relations when relations are already hostile, as compared to when they are peaceful. If we want to test the theory that trade has a different effect on relations depending on the starting state, then we are interested in the significance of β_3. But if we want to know whether good trade relations help end a conflict, then we are interested in the significance of $\beta_1 + \beta_3$.[3]

Let us now move to the more general case in which the dependent variable can take on C different values. For the 3-state case we want to estimate the probabilities π_{ab} in the transition matrix

$$\begin{array}{c c} & \begin{array}{ccc} Y_0 & Y_1 & Y_2 \end{array} \\ \begin{array}{c} Y_0 \\ Y_1 \\ Y_2 \end{array} & \begin{pmatrix} \pi_{00} & \pi_{01} & \pi_{02} \\ \pi_{10} & \pi_{11} & \pi_{12} \\ \pi_{20} & \pi_{21} & \pi_{22} \end{pmatrix} \end{array}, \tag{7}$$

where the rows represent the state of the system at time $t-1$, and the columns are the state at time t. Similar to Equations 2 and 3 above, we could run nine regular logits, one for each entry in the matrix; this is known as the "fully saturated model."[4] But there are some improvements we can make.

First, following Clayton (1992), it is convenient to work with cumulative transition probabilities. That is, we will express the estimation equations in terms of Y^* variables, where $Y^*_a = 1$ if $Y \leq a$. In the 3-state case, for example, the translation from Y to Y^* is given in Table 1.

[3] This last point becomes important in our subsequent discussion of PACL's analysis of democratic transitions.

[4] Note that in the 2-state case we could get away with running only two individual logits, since if the system is not in state 0 it must be in state 1. That is, the coefficients one obtains from estimating $\Pr(Y_{it} = 0|Y_{it-1} = 0)$ are equal and opposite to the coefficients from estimating $Pr(Y_{it} = 1|Y_{it-1} = 0)$. In the 3-state case, though, we must estimate $Pr(Y_{it} = b|Y_{it-1} = a)$ for all three values of both a and b.

Table 1: Definition of Y^* Variables

Y :	0	1	2
Y_0^*:	1	0	0
Y_1^*:	1	1	0

Note that $Y_2^* = 1$. As by definition, $Pr(Y \leq a) = Pr(Y \leq a-1) + Pr(Y = a)$, we can recover the individual transition probabilities from the set of cumulative probabilities.

As a simple example, the log-odds model of cumulative probabilities is:

$$\operatorname{logit} \Pr(Y \leq a) = \log \frac{\Pr(Y \leq a)}{\Pr(Y > a)} = \theta_a + X\beta.$$

If $X = 0$, then $\Pr(Y \leq a) = e^{\theta_a}/(1 + e^{\theta_a})$, which is non-decreasing in a, so $\theta_0 \leq \theta_1 \leq \ldots \leq \theta_{C-2}$. Further, if $\theta_a = \theta_{a+1}$, then $\Pr(Y \leq a) = \Pr(Y \leq a+1)$, and categories a and $a+1$ can therefore be collapsed.

The second improvement is to run each column of (7) as a single estimation equation, as we did with dichotomous Markov regressions in Equation 4. Assume that for any given a, the model to be estimated is:

$$\Pr(Y_t = b | Y_{t-1} = a) = \operatorname{logit}(\theta_{ab} + X\beta_a)$$

Then we can write:

$$\Pr(Y_t = b) = \operatorname{logit}\left(X_{t-1\beta} + \sum_a Y^*_{at-1} X_{t-1}\gamma_a\right).$$

As a specific example, say we have three states and one independent variable. Then we estimate:

$$\Pr(Y_{it} = b) = \beta_0 + \beta_1 y_0^* + \beta_2 y_1^* + \gamma_0 X + \gamma_1 X y_0^* + \gamma_2 X y_1^*$$

Now if $X = 0$, then

$$\begin{aligned}
\Pr(Y_t = b | Y_{t-1} = 2) &= \beta_0 \\
\Pr(Y_t = b | Y_{t-1} = 1) &= \beta_0 + \beta_2 \\
\Pr(Y_t = b | Y_{t-1} = 0) &= \beta_0 + \beta_2 + \beta_1.
\end{aligned}$$

Similarly, for general values of X:

$$\begin{aligned}
\Pr(Y_t = b | Y_{t-1} = 2, X) &= \gamma_0 \\
\Pr(Y_t = b | Y_{t-1} = 1, X) &= \gamma_0 + \gamma_2 \\
\Pr(Y_t = b | Y_{t-1} = 0, X) &= \gamma_0 + \gamma_2 + \gamma_1.
\end{aligned}$$

Equation 2 can be estimated separately for each value of b or with an ordered logit, where the dependent variable is the ordered category $Y_0, Y_1, \ldots, Y_{C-1}$. Diggle, Liang, and

Zeger (2002) suggest that one begin with a "fully saturated" model, with right-hand side variables consisting of the lagged regressors, the lagged values of the indicator variables Y_0^* and Y_1^*, and all interactions between the regressors and indicators.[5] From this initial model, with its profusion of interactive terms, one tests down, eliminating insignificant interactions to arrive at a more parsimonious specification.

To illustrate, start with the international relations example given above, but assume that there are now three ordered states of the dependent variable: relations can be friendly, tense, or outright hostile.[6] Keeping the single independent variable of bilateral trade, and suppressing the subscripts for the moment, we would estimate:

$$P(Y_t = y_t | Y_{t-1}^* = y_{t-1}^*) = \text{Logit}(\beta_0 + \beta_1 BT + \beta_2 y_0^* + \beta_3 y_0^* BT + \beta_4 y_1^* + \beta_5 y_1^* BT). \qquad (8)$$

Now β_1 measures the impact of trade on relations if relations in the previous period were hostile ($Y_{t-1} = 2$). Further, β_5 measures the difference between this effect and the impact of trade when relations are tense (this is where the cumulative probability measures come in handy), and $\beta_1 + \beta_5$ measures the total impact of trade when relations are tense. Finally, β_3 gives the difference between the impact of trade when relations are friendly as opposed to tense, and $\beta_1 + \beta_3 + \beta_5$ gives the overall significance of trade when relations are friendly.

To summarize, for any given independent variable X in a Markov model with $C > 2$ dimensions:

- The γ_0 coefficient on the un-interacted X term gives the impact of X_{t-1} on Y_t when $Y_{t-1} = Y_{C-1}$, the "last" category of Y.
- The γ_a coefficients on the interaction terms Xy_a^* give the *differential* impact of X_{t-1} on Y_t between $Y_{t-1} = Y_{a+1}$ and $Y_{t-1} = Y_a$.
- The cell probabilities giving the impact of X on $\Pr(Y_t = b | Y_{t-1} = a)$ when $a < C - 1$ can be recovered as the sums of the γ coefficients, in the order $\gamma_0 + \gamma_{C-1} + \gamma_{C-2} + \ldots + \gamma_{a+1}$.

3 Democratic Transitions: Review of PACL's Results

We illustrate these techniques through the analysis of democratic transitions. One of the first empirical regularities discovered in political science is the relation between a country's level of economic development and its level of democracy. Known generally as "modernization theory," this relationship has always been assumed to be causal: development leads to democracy.[7]

PACL (2000) challenged this long-standing hypothesis, providing analysis indicating that, although the observed correlations between development and democracy certainly did exist, they did not add up to a cause-and-effect relationship. PACL note that countries may

[5] Note that $Y_2^* = 1$. The saturated model gives the same results, of course, as running the nine individual regressions mentioned above.

[6] The middle category of tense relations might be associated with a lack of full diplomatic relations, or countries currently observing a cease-fire agreement, as in Fortna (2004).

[7] The exact mechanism for this translation of economic conditions into political regimes has, however, been hotly disputed. See Lerner (1958), Lipset (1959), and Barrington Moore (1966) for classic expositions, and Londregan and Poole (1996) for an especially careful test of the relation between income and democracy.

become democratic due to reasons unrelated to their level of economic development. Once prosperous, however, if democracies with higher levels of GDP per capita were to avoid slipping back into autocracy, then over time the relationship between GDP and democracy would emerge, even though economic growth does *not* directly cause democratization.

This line of research has had significant impact on both the scholarly and policy communities. We agree with PACL that a true test of modernization theory should examine both the impact of GDP on democratization and its ability to promote the consolidation of established democracies. However, we take issue with their conclusion that economic development does not play a significant role in transitions away from autocracy. This section reviews PACL's results, indicating a number of errors in their analysis. The next section presents our view of democratic transitions.

3.1 PACL's Simple Transition Model

As mentioned above, PACL employ a dichotomous regime classification. If (i) the chief executive is elected; (ii) the legislature is elected; (iii) there is more than one political party; and (iv) an incumbent regime has lost power, then the country is deemed democratic; otherwise, it is classified authoritarian. They apply this definition to a comprehensive set of 156 countries from 1950 to 1990 to construct their dependent variable.

PACL claim that increases in per capita GDP do not influence transitions from autocracy to democracy; rather, they help countries that are already democratic to remain so. They apparently base these conclusions on Tables 2.12 and 2.17 from Chapter 2 of their book. The former, reproduced as the first two columns of Table 2, performs a Markov probit regression of regime type on lagged values of per capita GDP, its square, and year-to-year GDP growth:

$$\begin{aligned} P(D_{it}) = \Phi\{\, & \beta_0 + \beta_1 GDP + \beta_2 GDP^2 + \beta_3 Growth + \\ & \beta_4 I_D + \beta_5 I_D GDP + \beta_6 I_D GDP^2 + \beta_7 I_D Growth\}, \end{aligned} \tag{9}$$

where $P(D_{it})$ signifies the probability that country i is a dictatorship in year t, $\Phi(\cdot)$ is the cumulative normal distribution, and I_D is an indicator variable for dictatorship in the previous period.[8] As indicated in the first two columns of Table 2, PACL report the coefficients on GDP and GDP^2 in this regression as insignificant, when predicting transitions both to and from democracy. PACL take this as evidence that the level of GDP per capita does not influence democratic transitions.

Note that when $I_D = 1$, however, the coefficient on GDP will be $\beta_1 + \beta_5$, the coefficient on GDP^2 will be $\beta_2 + \beta_6$, and likewise for the constant ($\beta_0 + \beta_4$) and Growth ($\beta_3 + \beta_7$). Table 2.12 correctly reports these summed coefficients in the columns labeled "Transitions to democracy" (the second column of our Table 2), but the reported P-values are those for β_4 through β_7 alone, rather than for the summed coefficients.

To calculate the P-values for transitions to democracy, one must perform a Wald test on the hypothesis that the sum of the appropriate coefficients is 0.[9] For example, the coefficient on β_1 in Equation 9 is -0.201, with a P-value of 0.162, and the coefficient on β_5

[8] Relative to PACL's Table 2.12, the coefficients on GDP and GDP^2 in Table 2 are multiplied by 1000.

[9] All Wald tests were performed using the post-estimation `test` command in Stata 9.0. Note that these same P-values can also be calculated simply by running two probits, one when the regime at time $t-1$ is democratic and another when it is a dictatorship.

Table 2: Results from PACL Table 2.12

Indep. Var.	*Democ. → Autoc.*	*Autoc. → Democ.* (Original)	*Autoc. → Democ.* (Corrected)
Constant	-1.144**	-2.524**	-2.524**
	(0.000)	(0.000)	(0.000)
GDP	-0.201	0.329	0.329**
	(0.162)	(0.484)	(0.004)
GDP^2	-0.003	-0.029	-0.029
	(0.874)	(0.191)	(0.069)
GDP Growth	-0.042**	-0.021**	-0.021*
	(0.003)	(0.000)	(0.015)
N	1584	2407	2407
Pseudo R^2	0.19	0.05	0.05

Note: P-values in parentheses. * = 0.05; ** = 0.01.

is -0.128 with a P-value of 0.484. The sum of the coefficients is -0.329, and PACL then correctly reverse the sign to indicate the impact of GDP on transitions from dictatorship to democracy.[10]

What these results tell us is that the impact of GDP on transitions to dictatorships is not significantly different from 0, and that the impact of GDP on transitions to democracy is not significantly different from *its impact on transitions to dictatorship*; that is, -0.329 is not significantly different from -0.201. But in this context we are interested in whether the sum of these coefficients is different from 0: that is, whether GDP is a significant predictor of transitions to democracy. And a Wald test of the hypothesis that $\beta_1 + \beta_5 = 0$ shows that it can be rejected with a P-value of 0.004.

Substituting the corrected standard errors into the analysis yields the results reported in the last column of Table 2. As shown, these results actually run *counter* to PACL's central hypothesis: GDP influences transitions to democracy but not transitions to autocracy.

On the other hand, both the GDP and GDP^2 terms contribute to the total impact of GDP on transitions. To evaluate this impact, we employ the delta method, which involves evaluating the derivative $\partial P/\partial GDP$. For Equation 9, the derivative is:

$$\Phi'\left(\beta_0 + \beta_1 GDP + \beta_2 GDP^2 + \beta_3 Growth\right) \cdot (\beta_1 + 2\beta_2 GDP) \tag{10}$$

when $I_D = 0$, and

$$\begin{aligned}\Phi'\left[(\beta_0 + \beta_4) + (\beta_1 + \beta_5)GDP + (\beta_2 + \beta_6)GDP^2 + (\beta_3 + \beta_7)Growth\right]\\ \cdot\left[(\beta_1 + \beta_5) + 2(\beta_2 + \beta_6)GDP\right]\end{aligned} \tag{11}$$

[10]The -0.329 coefficient indicates the impact of GDP on transitions from dictatorship to *dictatorship*, which is equal and opposite to its impact on transitions to democracy.

when $I_D = 1$. Performing these calculations, we find that the overall coefficient on GDP for transitions to autocracy is -0.0034 with a standard error of 0.0015, and for transitions to democracy the coefficient is -0.011 with a standard error of 0.0034. The total impact of GDP on regime change according to this specification is thus *significant* in both directions, rather than insignificant both ways as reported by PACL.

3.2 PACL's Extended Transition Model

PACL's Table 2.17, reproduced as the first two columns of Table 3, reports the results from another Markov regression, this time without GDP^2 but with a host of other covariates. The authors acknowledge that the coefficient on GDP is now significant in both directions, but discount this result, saying that "it is orders of magnitude larger for democracies." (p. 123) They do not indicate the basis for this statement.

As with Table 2.12, however, PACL fail to report the significance level of the sum of the relevant coefficients. The corrected version of these results is shown in the third column of Table 3.

This time the revised results are more favorable to their central hypothesis: GDP is a significant predictor of transitions to autocracy but not to democracy. These results, however, are far from dispositive, as they are highly sensitive to model specification. For example, in most specifications the inclusion of the Previous Transitions variable (labeled "$STRA$" in PACL) makes the coefficient on GDP insignificant. But an examination of the data patterns indicates that the greater the number of previous transitions, the less of an effect GDP has on the outcome. This in turn suggests including an interactive term, and indeed when this term is added all three variables (GDP, $STRA$, and $GDP * STRA$) are significant.

3.3 Summary

PACL's central claim is that "wealthy countries tend to be democratic not because democracies emerge as a consequence of economic development under dictatorships but because, however they emerge, democracies are much more likely to survive in affluent societies." It is hard to see how they draw this conclusion from the evidence they analyze. They estimate only two Markov regressions: one shows that GDP is insignificant both ways by their reckoning, the other shows it to be significant both ways. They offer no convincing explanation as to why their verbal description of their results is so at odds with their estimation equations.

On the other hand, their conclusions are based on a misinterpretation of their own findings, and the actual results are more favorable to their central hypothesis. In fact, their first regression, correctly analyzed, shows GDP to be significant in both directions, while the second shows it to be significant only in transitions away from democracy. They also make a number of other unsupported claims about transitions; these will be detailed in the following section. For such a well-cited, important work, these lapses are surprising, to say the least. PACL's results thus leave open the central issue: the significance of GDP in transitions to democracy.

Table 3: Results from PACL Table 2.17

Indep. Var.	*Democ. → Autoc.*	*Autoc. → Democ.* (Original)	*Autoc. → Democ.* (Corrected)
Constant	0.114 (0.899)	3.414** (0.002)	3.414** (0.000)
GDP	-0.547** (0.000)	-0.033** (0.000)	-0.033 (0.445)
GDP Growth	-0.022 (0.181)	0.018* (0.027)	0.018 (0.079)
Leadership Turnover	0.975** (0.001)	-0.527** (0.000)	-0.527** (0.007)
Religious Fractionalization	0.026** (0.010)	-0.001* (0.014)	-0.001 (0.816)
% Catholic	3.937* (0.048)	-0.369 (0.105)	-0.369 (0.707)
% Protestant	2.626* (0.039)	0.038 (0.118)	0.038 (0.965)
% Moslem	5.087* (0.016)	-0.147 (0.932)	-0.147 (0.890)
New Country	-0.012 (0.978)	0.432 (0.365)	0.432* (0.039)
British Colony	-0.842* (0.048)	-0.164 (0.153)	-0.164 (0.423)
Previous Transitions	0.897** (0.000)	-0.362** (0.000)	-0.362** (0.000)
% World Democracies	-3.735* (0.047)	-3.040 (0.750)	-3.040* (0.011)
N	1584	2407	2407
Pseudo R^2	0.19	0.05	0.05

Note: P-values in parentheses. * = 0.05; ** = 0.01.

4 A Three-Category Approach to Democratic Transitions

In this section we redefine the democratic transitions problem as a three-by-three Markov process. We first present our data and classification of the dependent variable; we then review the results from our 3-state Markov regression; finally, we explicitly compare our specification with PACL's to demonstrate the problems inherent in defining democratic transitions as merely a binary choice between autocracy and democracy.[11]

4.1 Partial Democracies as an Intermediate Regime Category

PACL employ a dichotomous regime classification: a country is either democratic or else it is authoritarian. Consider, however, the 85 authoritarian regimes that Geddes (1999, pp. 115-16) records as having collapsed during the "third wave." Of these, 34 re-emerged as authoritarian regimes, and 30 as stable democracies; 21 others, however, remained contested and unstable, she notes, and of these, four descended into "warlordism." Geddes' discussion thus reminds us of the significance of partial democracies, a category that dichotomous measures fail to — indeed, cannot — capture.

Using the Polity IV scaling of regimes from -10 to +10, we therefore categorize regimes as belonging to one of three categories: Autocracies (Polity value -10 to 0), Partial Democracies (+1 to +7), or (Full) Democracies (+8 to +10). Partial democracies emerge as an intermediate category relative to PACL's regime classification formula: in the country-years for which our data sets overlap, 97% of regimes that we code as autocratic PACL also code as autocracies, and 92% of our full democracies are democracies in their data too. But of our partial democracies, 52% are PACL democracies and 48% are PACL autocracies.

Autocracies in our coding scheme share a lack of political competition: all countries with non-elected executives and less than substantial limits on executive authority, for instance, are classified as autocratic. Partial democracies, on the other hand, are countries with an elected executive, but with incomplete constraints on executive authority due to either weak institutional checks or a weak process of executive recruitment. And full democracies are countries that enjoy binding constraints on executive authority, and an open and competitive electoral process.

Thus defined, partial democracies comprise 14.3% of country-years in our sample, which includes 169 countries from 1960 to 2000. As Figure 1 shows, the percentage of partial democracies among the world's societies has grown markedly in recent years: it had a minimum value of 3.6% in 1976 and rose to its maximum of 26.1% in 2000, with a notable increase after the fall of the Soviet Union. Huntington's "third wave" peopled the globe with *partial* democracies.

Table 4 examines the dynamics of change from one regime category to another. It shows the distribution of autocracies, partial democracies, and democracies, conditioning on the previous year's category. The table reveals that both autocracies and full democracies are stable in the short run: an average of 97.3% of all autocracies remain autocratic the next year, while an average of 98.2% of all democracies remain democratic; thus around 2% of countries in these categories change in a given year. Partial democracies are over four times less stable, however, with 9.6% of them changing into an autocracy or full democracy the

[11] For a more comprehensive analysis of the 3-state transitions model, which includes Tobit and duration analysis as well, see our companion paper, Epstein, Bates, Goldstone, Kristensen and O'Halloran (2005).

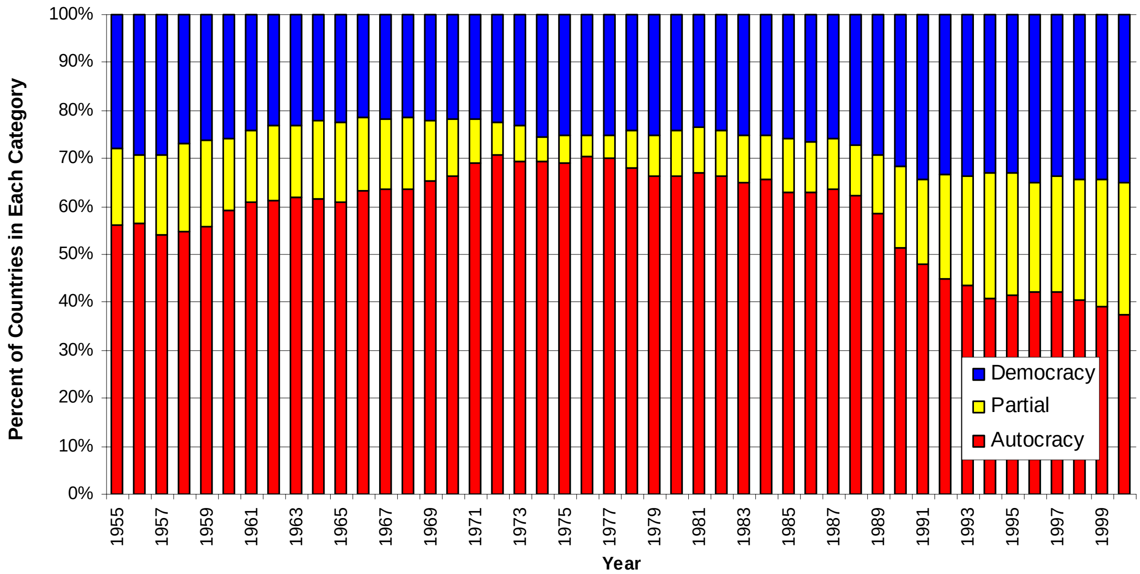

Figure 1: World Democratization Trends, 1960-2000

following year.

Table 4: Regime Category Transitions— One Year Lag

	Current Year		
Previous Year	Autocracy	Partial Democracy	Democracy
Autocracy	97.3%	2.1%	0.7%
	(3,121)	(66)	(22)
Partial Democracy	6.4%	90.4%	3.3%
	(49)	(695)	(25)
Democracy	1.1%	0.8%	98.2%
	(16)	(12)	(1,496)
Total	3,186	773	1,543

Note: Numbers in parentheses are cell counts.

These differences become even more pronounced when we expand the time horizon to five years. About 11% of all autocracies change into partial or full democracies after five years, and 7% of democracies change category five years later. The most volatile group, again by a large margin, is partial democracy: almost 40% of these change category after five years. Movements in or out of the category of partial democracy account for 80% of the transitions in our sample.

For our key independent variables, we employ the standard set of modernization indicators: log of GDP per capita, year-to-year GDP growth, the percent of the population living in cities, and log of population density.[12] As controls, we use a measure of prior experiences with democratization, defined as the cumulative sum at any given time of a country's negative changes in Polity score since 1960; log of trade openness, defined as exports plus imports over GDP; and a variable indicating whether over 75% of national income is derived from sales of minerals or petroleum. This latter variable captures the "resource curse" hypothesis (as in Ross 1999 and Boix and Stokes 2003), which argues that countries that derive a large share of national income from easily extractable natural resources tend to be unstable. Table 5 provides descriptive statistics for all variables.[13]

Table 5: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Polity Score	-0.45	7.58	-10	10	5671
Regime Category	0.70	0.88	0	2	5671
Log of Per Capita GDP	8.14	1.04	5.64	10.21	4417
Percent Change in GDP	0.02	0.06	-0.52	1.01	4475
Percent Urban Pop.	44.94	24.29	2.3	100	5245
Log of Population Density	3.61	1.46	-0.49	8.77	5600
Log of Trade Openness	3.98	0.62	0.43	6.16	4902
Previous Transitions	3.96	6.41	0	31	5671
Resource Curse	0.23	0.42	0	1	5671

4.2 Results from Three-State Markov Regressions

As a first look at the data, consider Figure 2, which shows a local regression (lowess) plot of logged GDP per capita and the probability of transitions in or out of full democracy. The most obvious pattern is that GDP does seem to have a significant impact on the probability of transition both ways, and with roughly similar magnitudes. This initial view of the data induces skepticism regarding PACL's claim that GDP impacts transitions down from democracy but not up out of autocracy.

PACL make much of the fact that no democracy has ever fallen with a GDP per capita greater than $6,055, the prevailing level of income in Argentina when it transitioned to autocracy in 1975.[14] They thus imply that the probability of transitioning to autocracy falls sharply once a country passes this key income level. As shown in Figure 2, though, no sharp dropoff is evident; the probability of leaving democracy declines smoothly as GDP

[12] One might also add percent of GDP originating from agriculture to this list, but its correlation with urbanization is over 90%. Thus we use only urbanization in our analysis.

[13] Data Sources: Polity Score—Polity V, IRIS, University of Maryland; GDP—Penn World Tables; Urban Population—Population Division of the Department of Economic and Social Affairs of the United Nations; Population Density—Hybrid of UN Population Division, World Development Indicators, and Banks population density series (WDI is used if UND is not available; BNK is used if WDI is not available); Trade Openness—Hybrid data series of World Development Indicators and Penn World Tables trade openness (WDI is the primary source; PWT is used if WDI is missing); Resource Curse—United Nations: Trade and Development Statistics.

[14] One might respond that all this proves is that every function has a maximum; they might as well argue that $293 (Burma's per capita GDP in 1959) is the key lock-in number for autocracy, since no country has ever transitioned out of autocracy with a GDP below this level.

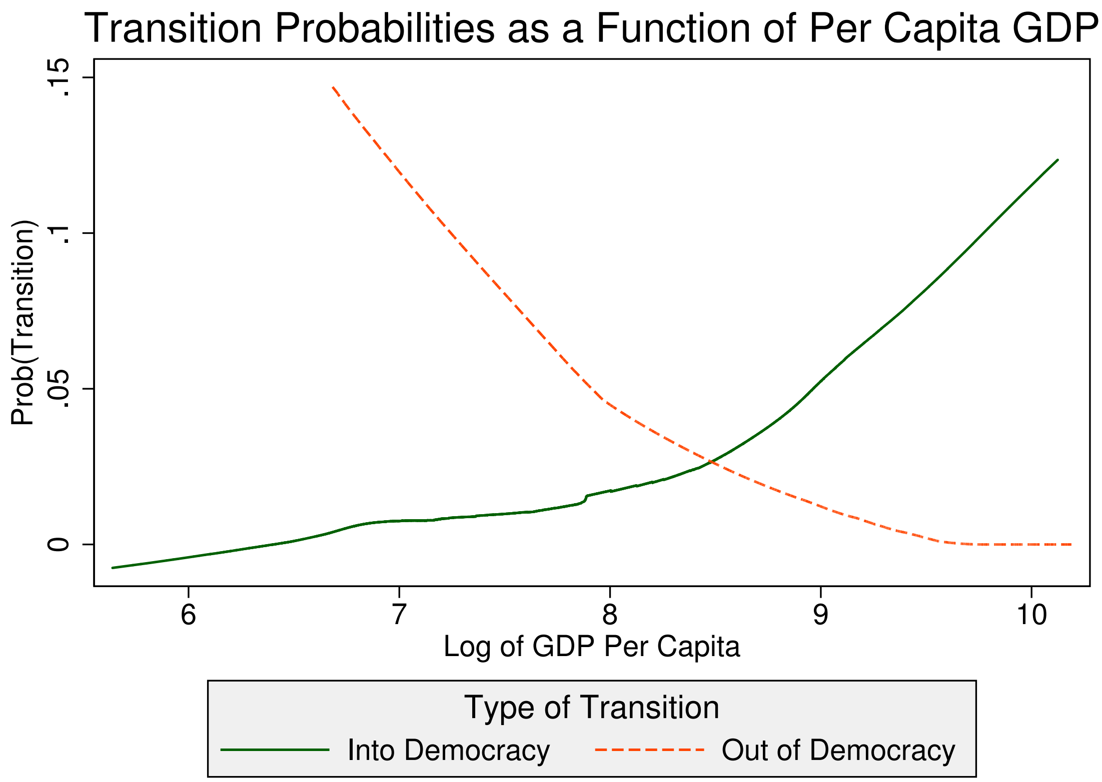

Figure 2: Impact of GDP on Transition Probabilities

increases, without any indication that one level of wealth is more critical than another along the way.

PACL also claim that the income levels at which countries transition out of autocracy show significantly more variation than the levels at which countries transition out of democracy. Figure 3 shows that the data do not support PACL's claim: the distribution of GDP values for transitions to democracy actually has a slightly *smaller* variance than the distribution of income for transitions to autocracy (0.712 vs. 0.742).

We begin our Markov analysis with all possible interactions between the regressors and lagged values of Y_0^* and Y_1^*, and then test down to a more parsimonious model.[15] Recall that if, for example, the interaction between GDP and Y_0^* ($GDP * Y_0^*$) is significant, this means that GDP has a different effect on the level of democracy if the regime is autocratic in the previous period, as opposed to partially or fully democratic. Similarly, if $GDP*Y_1^*$ is significant, GDP has a different impact when the regime is fully democratic in the previous period, as opposed to the other two alternatives. Consequently, if both $GDP * Y_0^*$ and $GDP * Y_1^*$ are significant, GDP has a significantly different effect for all three possible

[15]We also tested to see if lags of two or three periods added anything to the analysis. In no case were these higher-order coefficients significant. This is not too surprising in our context; countries in a given regime at time $t-1$ are almost certain to be in that same regime at times $t-2$ and $t-3$ as well.

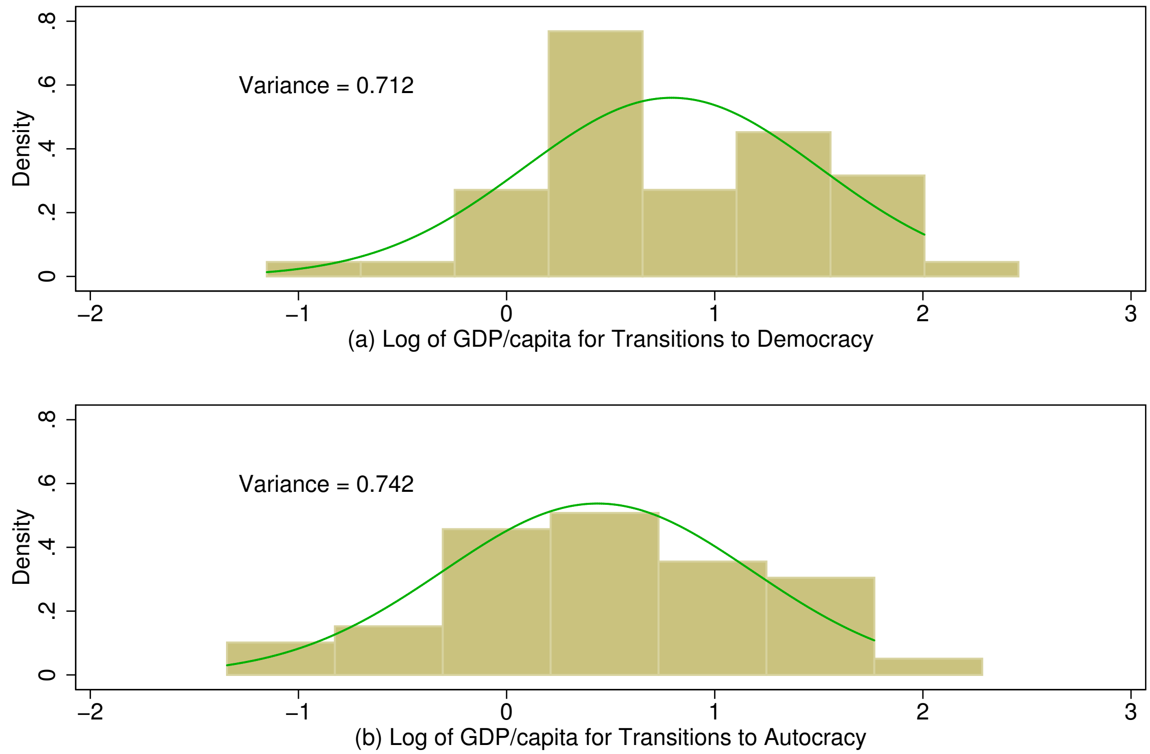

Figure 3: Distribution of GDP for Transitions To and From Democracy

lagged regime types. If the coefficients are insignificant, then adjacent categories can be combined.

The results of this analysis, after winnowing out insignificant interactive terms, are illustrated in Table 6 in raw form, and in Table 7 in a more easily interpretable format. Beginning with the former, we see that the significance levels of the coefficients on the modernization variables are similar with and without the political variables. The only exception is population density, for which the sum of the coefficients is marginally significant in Model 2, but which is dropped from the analysis in Model 3 as neither the direct effect nor any interaction terms were significant.

Table 7 distills the results from the analysis, showing only the relevant (sums of) coefficients from the direct and interactive effects. Coefficients that straddle table rows have similar effects for the adjacent categories. In all three models, for example, GDP has a similar impact on democratization when the country in question was autocratic or partially democratic in the previous period, as opposed to fully democratic. If the country was fully democratic, then the coefficient on GDP in Model 1 is 0.80 (the direct effect from Table 6); if the country was autocratic or partially democratic, the coefficient is 0.18 (the sum of GDP and $GDP * Y_1^*$ in Table 6). Both are significant, as they are in Models 2 and 3 as well, indicating that higher GDP does produce more democratic regimes, no matter what the starting point, and no matter which sets of covariates are added to the estimation equation.

The other findings in the table are also interesting. The coefficient on growth, for

Table 6: Markov regression analysis

	Model		
	(1)	(2)	(3)
Lagged Y_0^*	-2.68 (.073)***	-2.97 (.75)***	-2.86 (.10)***
Lagged Y_1^*	2.22 (.85)***	4.13 (1.49)***	3.83 (1.53)**
GDP Per Capita	.80 (.096)***	1.11 (.20)***	.99 (.21)***
GDP per capita * Y_1^*	-.62 (.10)***	-.97 (.21)***	-.80 (.23)***
GDP Growth		-.12 (.97)	-.38 (.98)
GDP Growth * Y_0^*		-2.23 (1.30)*	-1.844 (1.30)
Pct. Urban Pop.		-.016 (.007)**	-.012 (.008)
Pct. Urban Pop. * Y_1^*		.019 (.008)**	.013 (.008)
Population Density		.017 (.034)	
Population Density * Y_0^*		.075 (.049)	
Trade Openness			.22 (.12)*
Trade Openness * Y_1^*			-.23 (.14)*
Previous Transitions			-.023 (.013)*
Previous Transitions * Y_0^*			.033 (.009)***
Previous Transitions * Y_1^*			.026 (.015)*
Resource Curse			-.18 (.098)*
N	4299	3789	3789
Pseudo-R^2	.773	.776	.78

Note: Standard errors in parentheses. All independent variables lagged one year. *=.10; **=.05; ***=.01.

Table 7: Summary of Markov Results

		Model		
		(1)	(2)	(3)
GDP Per Capita	A, P	0.18***	0.15**	0.19**
	D	0.80***	1.04***	1.00***
GDP Growth	A		-2.34^{***}	-2.23^{**}
	P, D		-0.189	-0.386
Percent Urban Pop.	A, P		0.002	-0.0001
	D		-0.015^{**}	-0.013^{*}
Population Density	A		0.095***	
	P, D		0.021	
Trade Openness	A, P			-0.010
	D			0.227*
Previous Transitions	A			0.035***
	P			0.002
	D			-0.024^{*}
Resource Curse	A, P, D			-0.186^{*}

Note: Coefficients refer to the relevant sums of direct and interactive effects. $* < 0.10$; $** < 0.05$; $*** < 0.01$.

example, is significant only for countries starting as autocracies, in which case it inhibits democratic transitions; otherwise, growth is not a significant factor. Urbanization, on the other hand, appears to undermine democracies but has no effect on other regime categories. And population density, significant in Model 2 only, promotes transitions out of autocracy but has no impact on partially or fully democratic regimes.

Turning to the political variables, trade openness helps stabilize full democracies, but it does not help autocratic or partially democratic regimes move up the ladder. The results for the Previous Transitions variable illustrate the power of the Markov approach. Previous transitions destabilize autocracies, have no impact on partial democracies, and make full democracies more likely to backslide. Thus a single variable can have different impacts (in fact, opposite signs) depending on the starting point in the previous period. Finally, the resource curse tends to make all regime categories more autocratic.

We performed two diagnostic checks on the analysis. First, a Brant test did not reject the null hypothesis that the parallel regressions assumption holds for all independent variables in our ordered logistic regression. Second, the standard errors in a Markov regression will be consistent only if the first-order Markov assumption holds. Even if it does not hold, though, the robust standard errors will be asymptotically consistent. Thus one way to check the Markov assumption is to compare the regular and robust standard errors. In all cases, these were close to each other, never changing the significance level of a variable or interactive term.

4.3 Comparison with PACL's Two-State Model

Why do our results from the Markov analysis vary so markedly from PACL's? They, after all, test a similar model to ours. Perhaps the difference comes from our coding of the dependent variable: we use Polity scores, while PACL employ their own measure of autocracy and democracy. If we substitute our Polity measure into their regressions, though, combining partial and full democracies into a single democratic category, the estimation results from PACL's model specification still hold. In particular, even with a Polity version of the dependent variable, lagged GDP is shown to be a significant predictor of transitions out of democracy, but not to democracy.

Conversely, we dropped the "partial" category in our data set by using only the interactions of the regressors with Y_0^*, and eliminating all interactions with Y_1^*. With this specification, the coefficient on GDP is, as PACL concluded, significant for transitions to autocracy, but not to democracy. In both data sets, then, one can reproduce PACL's results using a dichotomous regime classification.

Let us then investigate further whether our three-way regime division is correct. We estimated a model similar to those illustrated in Table 6 but with no independent variables other than the lagged values of Y_0^* and Y_1^*. These terms were both significant, indicating that none of the adjacent regime categories should be combined. We then added a fourth category—a "partial autocracy" category consisting of regimes with Polity scores between -6 and 0—and defined Z_0^*, Z_1^*, and Z_2^* as cumulative state indicators, parallel to the Y^* variables we constructed for the three-state case. Estimating an ordered probit of the new four-state regime against lagged values of these indicators revealed only the coefficient on Z_0^* to be insignificant, indicating that the full and partial autocracies should indeed be combined.

Turning directly to our differences with PACL, our hypothesis is that their two-state dependent variable masks a good deal of transition activity taking place within the single category that they term "autocracy." If we are correct, we should observe relatively more transitions out of our autocracy category than theirs, and these transitions should correspond more closely with a country's level of economic development.

We therefore subdivide the two PACL regime categories into four: 1) PACL autocracies that we did not list as partial democracies; 2) PACL autocracies that we list as partial democracies; 3) PACL democracies that we list as partial democracies; and 4) PACL democracies that we did not list as partial democracies. PACL combine categories 1 and 2 versus 3 and 4, while we combine 2 and 3 together, but leave 1 and 4 as distinct categories. If our hypothesis is correct, then we should see relatively more transitions out of category 1 into categories 2, 3, or 4 than we would see from categories 1 and 2 to categories 3 or 4. And, in fact, 2.63% of regimes transition out of category 1, which is a 49% increase over the 1.76% that transition out of categories 1 or 2.

Moreover, when we run a Markov regression with GDP as an independent variable, we find that the coefficients separating categories 2 and 3 are uniformly insignificant, while those separating category 1 from category 2 and category 3 from category 4 are uniformly significant. Thus the data support our tripartite regime classification over PACL's dichotomous specification.

Finally, note the elusive nature of partial democracies. Although we can gain some understanding of the factors that make autocracies (or full democracies) become partially democratic, we have little information as to the factors that would lead partial democracies to either slide down to autocracy or to move up to full democracy. In fact, examining the saturated regression with all direct and interactive effects, we find that *none* of the coefficients on partial democracy are significant on their own.

5 Conclusion

This paper introduced the techniques necessary to estimate Markov transition models when the dependent variable can take on more than two values. The interpretation of the coefficients in these models, with their profusion of interactive terms, can be daunting. But working with cumulative probabilities of events can simplify the analysis and aid in the comparison of competing models.

We then applied this technique to the particular area of democratic transitions. We showed, first, that the example usually cited as the examplar of this field, Przeworski, et. al. (2000) contains significant flaws in its analysis. Even when these are corrected, though, their two-state model has less explanatory power than a three-state model incorporating "partial democracies."

References

[1] Amemiya, Takeshi. 1985. *Advanced Econometrics*. Cambridge: Harvard University Press.

[2] Boix, Carles and Susan Stokes. 2003. "Endogenous Democratization." *World Politics* 55: 517-549.

[3] Clayton, D.G. 1992. "Repeated Ordinal Measurements: A Generalised Estimating Equation Approach." *Medical Research Council Biostatistics Unit Technical Report*. Cambridge, England.

[4] Diggle, Peter, Kung-Yee Liang and Scott Zeger. 2002. *Analysis of Longitudinal Data*. Oxford: Oxford University Press.

[5] Epstein, David, Robert Bates, Jack Goldstone, Ida Kristensen, and Sharyn O'Halloran. 2005. "Democratic Transitions." Manuscript, Columbia University.

[6] Fortna, Page. 2004. *Peace Time*. Princeton: Princeton University Press.

[7] Geddes, B. 1999. "What Do We Know About Democratization After Twenty Years?" *Annual Review of Political Science* 2: 115-44.

[8] Jackman, Simon. 2000. "In and Out of War and Peace: transitional models of international conflict." Working Paper, Stanford University.

[9] Lerner, D. 1958. *The Passing of Traditional Society*. New York: Free Press of Glencoe.

[10] Lipset, S. M. 1959. "Some Social Requisites of Democracy." *American Political Science Review* 53: 69-105.

[11] Londregan, John B. and Keith T. Poole. 1996. "Does High Income Promote Democracy?" *World Politics* 49: 1-30.

[12] Moore, Barrington. 1966. *Social Origins of Dictatorship and Democracy*. Boston: Beacon Press.

[13] Przeworski, Adam, Michael E. Alvarez, Jose Cheibub and Fernando Limongi. 2000. *Democracy and Development*. New York: Cambridge University Press.

[14] Ross, Michael L. 1999. "The Political Economy of the Resouce Curse." *World Politics* 51: 325–61.

[15] Ware, J. H., S. Lipsitz and F. E. Speizer. 1988. "Issues in the Analysis of Repeated Categorical Outcomes." *Statistics in Medicine* 7:95107.