11

Electoral institutions, cleavage structures, and the number of parties[1]

There does not exist a sustainable scientific proposition of high informative content concerning the effects of electoral systems that can be derived in complete isolation from social and political relations. The social, ethnic or religious homogeneity or heterogeneity of a society are very important for the structure of the party system.

Nohlen 1993:27

Thus far in this book I have considered two levels at which political actors may face incentives to coordinate their actions. First, within individual legislative constituencies, voters or parties may need to coalesce in order to convert their votes into legislative seats more efficiently. Second, within the nation as a whole, would-be legislators and would-be executives may need to cooperate in order to convert their resources into executive office more efficiently. I have also stressed two broad influences on the outcome of the electoral coordination game, whether at the district or the national level: First, the nature of the electoral procedure pertinent to each level (the local electoral system by which legislators are elected, the national electoral procedure by which executives are elected); second, the nature of the social actors and cleavages involved at each level.

Chapter 10 suggested that coordination at both levels would be pertinent to explaining the number of parties at the national level: coordination at the district level by affecting the number of parties within each district, coordination at the national level by affecting the degree to which the local party systems cumulated into a national party system. Chapter 10 also reasserted a point at the national level that had previously been developed at the district level: The number of parties should

[1]This chapter is based directly on my work with Octavio Amorim Neto (Cox and Amorim Neto N.d.).

be an interactive function of social diversity and electoral permissiveness. In this chapter, I develop a cross-sectional model of the effective number of parties (at the national level) in 54 polities that addresses these points.

The main purpose of this model is to investigate whether the hypothesized interaction between social and electoral structure appears in practice. The structure of the investigation is similar to that in a recent pair of works (Powell 1982; Ordeshook and Shvetsova 1994) that have included both sociological and institutional variables in regression analyses of cross-national variations in the number of parties. I put particular emphasis on testing Ordeshook and Shvetsova's main finding, that there is a significant *interaction* between social heterogeneity and electoral structure. In order to put this claim to a stringent test, I employ a substantially different dataset, one that includes about twice as many countries as have previous studies, including a large number of third-world democracies.

A second purpose of the model is to follow up on the suggestion made in Chapter 10 that the number of national parties in a system will depend interactively on (1) the degree of integration of executive and legislative elections; and (2) the strength of the executive choice procedure. The importance of variables affecting the degree of integration – in particular, whether presidential and legislative elections are held concurrently – has been stressed in recent work by Shugart and others (see below). My approach follows this earlier work but specifies the key relationships differently. In particular, I do not model presidential election rules as having a direct impact on the legislative party system. Instead, there is a two-step process: (1) Presidential election rules combine interactively with social diversity to produce an effective number of presidential candidates; (2) the effective number of presidential candidates affects the effective number of legislative competitors, with the size of the impact depending on the proximity of the presidential and legislative elections.

A final purpose of the model is to assess the importance of another national-level variable suggested in Chapter 10: The existence of upper tiers. Here too my operational approach differs somewhat from that in the previous literature, in that I have separate variables that reflect the nature of the upper tier (if any) and the lower tier of the electoral system, rather than trying to combine these two considerations into a single measure, such as Taagepera and Shugart's "effective magnitude."

The outline of the chapter is as follows. Section 11.1 reviews the previous work of Powell and of Ordeshook and Shvetsova. Section 11.2 explains how my data and methods differ from, and complement, these previous efforts. Section 11.3 presents the empirical results. Section 11.4 concludes.

11.1 THE PREVIOUS LITERATURE

Studies of the effective number of elective or legislative parties rarely investigate the impact of both social cleavages and electoral laws on party system fractionalization. Among quantitative studies I am aware of only two that do. The first of these, Powell (1982), looks only at legislative fractionalization while the second, Ordeshook and Shvetsova (1994), looks at both elective and legislative fractionalization.[2]

Powell's work focuses on a set of 84 elections held in 27 mostly European countries during the period 1965-1976. The dependent variable, legislative fractionalization, is measured by Rae's index (that is, $1 - \sum s_i^2$, where s_i is the seat share of the ith party). The independent variables of primary interest are three measures of social heterogeneity:

- ethnic fractionalization as measured by Rae's index (that is, $1 - \sum g_i^2$, where g_i is the proportion of the population in ethnic group i);
- an index of agricultural minorities (coded 3, 2, or 1 if the agricultural population comprises 20–49%, 50–80%, or 5–19% of the total population); and
- an index of Catholic minorities (coded similarly to the agricultural index);

and two measures of electoral structure:

- the "strength" of the electoral system for legislative elections (coded 3 for single-member plurality elections, 2 for the Japanese, German, and Irish systems, and 1 for proportional systems); and
- a dummy variable indicating whether or not the system is presidential (1 if yes, 0 if no).[3]

Regressing the independent variables just listed on the legislative fractionalization scores for each election, Powell (p. 101) finds that "fractionalization is encouraged above all by ... nonmajoritarian electoral laws, but also by all of the heterogeneity measures, and discouraged by presidential executives."

Ordeshook and Shvetsova (1994) consider several different datasets: Lijphart's (1990) sample of 20 Western democracies from 1945-85 (representing 32 distinct electoral systems); an extension of this

[2]Other studies that share the same basic conception, but do not run regressions with explicit measures of both electoral and social structure, include Nagel (1994) and Coppedge (1995).

[3]Two control variables, population (in millions, as of 1965) and GNP per capita (as of 1965), are also included.

dataset covering elections in 23 Western democracies from 1945-90 (representing 52 distinct electoral systems); and a further extension that includes Continental elections in the period 1918-39. The dependent variables that Ordeshook and Shvetsova investigate are four: the effective number of elective parties (ENPV = $1/\sum v_i^2$, where v_i is party i's vote share); the effective number of legislative parties (ENPS = $1/\sum s_i^2$, where s_i is party i's seat share); the number of parties that receive at least 1% of the vote in two or more successive elections; and the number of parties that secure one or more seats in two or more successive elections. They measure social structure chiefly in terms of ethnicity, calculating the effective number of ethnic groups (ENETH = $1/\sum g_i^2$, where g_i is the proportion of the population in ethnic group i); and measure electoral system properties by the average district magnitude and by Taagepera and Shugart's "effective magnitude" measure. They then use OLS regression to explain variations in their dependent variables (here I shall look just at ENPV), considering three basic specifications: (1) the institutionalist specification: ENPV as a function solely of the log of district magnitude, as in Taagepera and Shugart (1989); (2) the sociological specification: ENPV as a function solely of ethnic heterogeneity; and (3) the interactive specification: ENPV as a function of the *product* of ethnic heterogeneity and district magnitude. They find that the interactive model does best in explaining the data, summarizing their findings as follows:

> ... if the effective number of ethnic groups is large, political systems become especially sensitive to district magnitude. But if ethnic fractionalization is low, then only especially large average district magnitudes result in any "wholesale" increase in formally organized parties. Finally, if district magnitude equals one, then the party system is relatively "impervious" to ethnic and linguistic heterogeneity ... (Ordeshook and Shvetsova 1994:122).

Thus, whereas Powell (1982:81) had success with an additive specification, Ordeshook and Shvetsova find an interactive model to be superior.

Why should an interactive model work well? The answer suggested in previous chapters runs as follows. A polity will have many parties only if it *both* has many cleavages *and* has a permissive enough electoral system to allow political entrepreneurs to base separate parties on these cleavages. Or, to turn the formulation around, a polity can have few parties either because it has no need for many (few cleavages) or poor opportunities to create many (a strong electoral system). If these claims are true, they would rule out models in which the number of parties depends only on the cleavage structure, or only on the electoral

system, or only on an additive combination of these two considerations.[4]

Plausible though this formulation might be, it still leaves several questions unanswered. First, and most important, is the question of empirical evidence. Thus far we have one study in which an additive specification seems to work well (Powell) and one study in which an interactive specification proves superior (Ordeshook and Shvetsova). The latter study, moreover, is based largely on European evidence, and one might well ask what would happen if India (or other socially diverse third-world countries with strong electoral systems) were added. Since India appears to have lots of social cleavages and also to have lots of parties, would the addition of this (kind of) case to the analysis not bolster the importance of social heterogeneity and, perhaps, point more toward an additive rather than an interactive specification? Second, there is also the issue of what the form of the interaction between electoral and cleavage structure is. Perhaps the effective number of elective parties (ENPV) should equal the minimum of (1) the number of parties that the cleavage structure will support (loosely following Taagepera and Grofman [1985], we might say this number was $C + 1$, where C is the number of cleavages); and (2) the number of parties that the electoral system will support (following the "generalized Duverger's Law" of Taagepera and Shugart [1989], we might say this number was $2.5 + 1.25\log_{10}M$, where M is the district magnitude). That is, perhaps the equation should be something like ENPV = MIN[$2.5 + 1.25\log_{10}M, C + 1$]. Or, perhaps the form of the interaction is as Ordeshook and Shvetsova specify it, a simple product of factors reflecting electoral strength and number of cleavages. In Sections 11.2 and 11.3, I investigate both these questions, especially the first.

11.2 DATA AND METHODS

In considering the interaction between social heterogeneity and electoral permissiveness, my analytical strategy is to look at different data than did Ordeshook and Shvetsova (1994), using different operational measures of key variables. The notion is that, if their basic finding of a significant interaction is robust to these changes, then we can have more confidence in it. The most important differences between the present analysis and Ordeshook and Shvetsova's are as follows: I include a larg-

[4]An additive combination model, such as Powell's, allows the number of parties to be large either because there are many cleavages (regardless of how strong the electoral system is) or because the electoral system is very permissive (regardless of how few cleavages there are).

er number of countries, including many third-world democracies; I measure the strength of an electoral system by employing separate measures of lower-tier district magnitude and upper-tier characteristics, rather than combining these two factors (in an "effective magnitude") or ignoring upper tiers (by taking a simple average of the district magnitudes); and I include variables tapping the influence of presidential elections (if any) in the system. Let us consider each point in turn.[5]

Case selection: I have taken as a case every polity with an election in the 1980s (defined as 1980-1990 inclusive) that qualifies as "free" by Freedom House's score on political rights (either a 1 or a 2); if a polity has multiple such elections in the 1980s, I have taken the one closest to 1985.[6] These criteria of selection yield a substantially more diverse sample than that used by Ordeshook and Shvetsova (or Powell before them), one including India, Venezuela, Mauritius, and many other third-world countries (see Appendix C). The total number of countries included is 54. As there is only one observation per country, the sample can also be described as having observations on 54 electoral systems.

Measuring electoral structure. I differ from Ordeshook and Shvetsova and most of the previous literature in that I do not use average magnitude or Taagepera and Shugart's "effective magnitude" to indicate the strength of an electoral system. Instead, I use two variables, one to describe the magnitude of the lower-tier districts, one to describe the impact of the upper tier.

The lower-tier variable is based on the magnitude of the median legislator's district. An example may help to clarify why I use this variable rather than simply the average district magnitude. Suppose an electoral system has just two districts, one returning a single member and one returning 100 members. The average district magnitude in this system is (100+1)/2 = 50.5. But this process of averaging, in which each district counts equally, does not correspond to the usual way in which the effective number of parties is calculated. To see this, suppose that there are 100 voters in the 1-seat district, who split equally between two parties, while there are 10,000 voters in the 100-seat district, who split equally among ten parties. In this case, the effective number of parties in the 1-seat district, the 100-seat district, and the

[5]A copy of the full dataset, along with SAS code that reads and analyzes the data, can be found under the "publication-related datasets" heading of the Lijphart Elections Archive at http://dodgson.ucsd.edu/lij.

[6]The only exceptions to these rules are as follows. First, I have not included any of the Pacific Island states (e.g., Tuvalu) because I could not get complete data. Hungary (1990) is excluded for the same reason. Finally, I take the 1990 Brazilian election rather than the (unusual) 1986 election and the 1981 French election (held under the traditional runoff system) rather than the 1988 election (held under PR).

nation as a whole are respectively 2, 10, and almost 10. The national effective number of parties is much closer to the effective number of parties in the large district because the votes from both districts are simply added to arrive at the national vote totals, and there are 100 times more voters in the large district than in the small. The national effective number of parties, in other words, is a weighted average of the district figures, in which larger districts get more weight. Accordingly, it seems natural to use a similarly weighted measure of the central tendency in district magnitudes. I weight each district by the number of legislators from that district (which, if there is no malapportionment in the system, and turnout is equal across districts, will correspond to the weights used in calculating the national effective number of parties). I also use medians rather than means. In the example at hand, this yields a figure of 100: There are 101 legislators, of whom 100 are elected from a district of magnitude 100; the magnitude of the median legislator's district is thus 100. As it turns out, using the average of the legislators' district magnitudes, rather than the median, has virtually no impact on the results that follow. Finally, I follow Taagepera and Shugart (1989) and take the logarithm of the median legislator's district's magnitude, to produce a variable I call LML.

The upper tier variable that I use, denoted UPPER, equals the percentage of all assembly seats allocated in the upper tier(s) of the polity. It ranges from zero for polities without upper tiers to a maximum of 50% for Germany. The idea here is that instead of attempting to deduce how the existence of an upper tier affects the "effective magnitude" of a system, I simply let the upper tiers speak for themselves. Because all but one of the upper tiers in my sample are compensatory – designed specifically to increase the proportionality of the overall result – I can avoid some of the complexities of Taagepera and Shugart's "effective magnitude," which attempts to put the effects of compensatory and additional seats on a common metric (Taagepera and Shugart 1989, ch. 12).[7]

Presidentialism. Several previous studies, e.g., Powell (1982), Lijphart (1994), Jones (1994), Mainwaring and Shugart (1996), have included a code for presidential elections in investigations of legislative fractionalization. So do I. As my coding of this variable differs from these previous studies, however, I discuss it at some length.

The simplest way to code presidentialism is with a dummy variable (1 for presidential systems, 0 for parliamentary), as do Lijphart and Powell.

[7]The South Korean upper tier is designed to ensure that the largest party can secure a majority, or a near-majority, in the legislature, and thus in principle it *reduces* proportionality. The results do not change appreciably depending on how one codes South Korea. Nor do they change if South Korea is simply omitted from the analysis.

The problem with this approach is that there are different kinds of presidential elections (runoff, plurality), held at different times relative to the legislative elections (concurrently, non-concurrently), and these factors plausibly matter. Thus, other scholars, such as Shugart and Carey (1992), Jones (1994), Shugart (1995), and Mainwaring and Shugart (1996), have developed more elaborate schemes. My approach, which follows Shugart and Carey in general conception but differs in the details of implementation, takes the influence a presidential election exerts on a legislative election as depending on two factors: the proximity of the two elections; and the degree of fractionalization of the presidential election.

Proximity is a matter of degree. If the presidential and legislative elections are concurrent, then proximity is maximal. Here, I take the maximum value of proximity to be unity (so concurrent elections are "100% proximal," so to speak). At the other end of the scale are legislative elections held in complete isolation from presidential elections, i.e., in non-presidential systems.[8] Such legislative elections are not at all proximal to a presidential election, so they are coded as of zero proximity. In between these two extremes are presidential systems with nonconcurrent elections. If we denote the date of the legislative election by L_t, the date of the preceding presidential election by P_{t-1}, and the date of the succeeding presidential election by P_{t+1}, then the proximity value is

$$\text{PROXIMITY} = 2^* \left| \frac{L_t - P_{t-1}}{P_{t+1} - P_{t-1}} - \frac{1}{2} \right|.$$

This formula expresses the time elapsed between the preceding presidential election and the legislative election ($L_t - P_{t-1}$), as a fraction of the presidential term ($P_{t+1} - P_{t-1}$). Subtracting 1/2 from this elapsed time fraction, and then taking the absolute value, shows how far away from the midterm the legislative election was held. The logic of the formula is as follows: The least proximal legislative elections are those held at midterm. This particular formula gives a proximity value of zero to these elections, which equates them with the totally isolated elections of non-presidential systems.[9] The most proximal nonconcurrent elections are

[8] In deciding whether a system is presidential or not, I have followed Shugart and Carey (1992, ch. 8). Ireland, for example, in which the president has neither legislative nor governmental powers, is coded as non-presidential. All systems in which the president has either legislative, or governmental, or both kinds of powers are coded as presidential.

[9] It is possible to include an additional parameter to test whether midterm elections are significantly more affected by presidential politics than elections occurring in non-presidential systems. I have done so and found that one cannot reject the hypothesis that midterm and nonpresidential elections are equally unaffected by presidential elections.

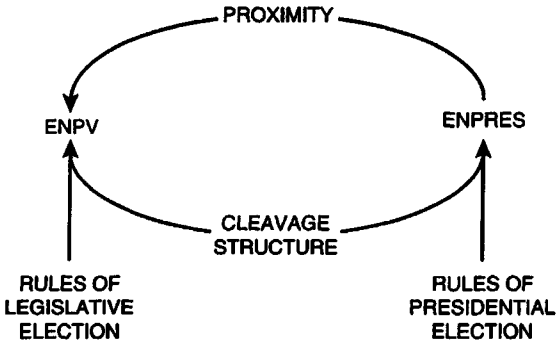

Figure 11.1. A schematic representation of the relationship between social cleavages, rules of election, and number of competitors in presidential and legislative elections

those held just before or just after a presidential election. The formula above gives them a proximity value that approaches one, the same value given to concurrent elections.[10]

The proximity of the presidential election to the legislative election is a necessary condition for the former to influence the latter. But the nature of that influence depends on the nature of the presidential election. One approach to coding the nature of the presidential election is institutional. Mainwaring and Shugart (1996), for example, introduce variables that distinguish three classes of presidential elections: concurrent plurality, majority runoff, and other. Although I report some results in a footnote that follow this route, my approach is different.

My point of departure is the notion that both presidential and legislative election results convey information about the impact of social cleavages and electoral laws. To put it another way, if one denotes the effective number of presidential candidates by ENPRES, and the effective number of elective parties in the legislative election by ENPV, then both ENPRES and ENPV may be thought of as dependent variables – products of social and electoral structure – along the lines of Figure 11.1.

There are three things to note about Figure 11.1. First, the picture assumes that the effect of the presidential election on the legislative elec-

[10]With the current dataset, it is difficult to test Shugart's (1995) hypothesis that there is a jump between nearly concurrent and exactly concurrent elections, with the depressive effect of presidential elections being much larger in the latter, since there are only five observations on concurrent elections. Some limited exploration – replacing the value "2" in the definition of PROXIMITY given in the text by "1.9" and other lower values – suggests that the main results of the paper do not depend on how one handles this issue.

tion dominates that of the legislative election on the presidential: Thus there is an arrow from ENPRES to ENPV but not one going in the reverse direction. In reality, there no doubt are reverse causal arrows of the kind omitted from Figure 11.1. But I believe that the direction of influence is primarily from executive to legislative elections, and making this assumption facilitates econometric estimation of the system of equations implied by Figure 11.1. In particular, one can first estimate an equation determining ENPRES and then estimate an equation, in which ENPRES appears as a regressor, determining ENPV (see below). (One can estimate a fully simultaneous pair of equations for ENPV and ENPRES but I do not believe that this is necessarily the best way to go econometrically and, in any event, the results do not change appreciably from those reported for the recursive model.[11])

The second thing to note is that the influence of presidential on legislative elections is mediated through the effective number of presidential candidates, ENPRES, and does not include a direct impact of presidential rules on legislative fractionalization, as does the Mainwaring and Shugart formulation. The justification for this runs as follows. Imagine a presidential election held under runoff rules that nonetheless – perhaps because the country is dominated by a single cleavage, perhaps for reasons idiosyncratic to the particular election – ends up as a two-way race. Given that there are just two candidates in the presidential race, I expect the same kind of influence as would be produced by a plurality race with two candidates. The nature of the coattail opportunities that face legislative candidates should be similar, the nature of the advertising economies of scale that might be exploited should be similar, and so forth. It is hard to see why the presidential rules themselves, having failed to produce the expected result in the presidential race, would nonethe-

[11]Estimating the equations using two-stage least squares rather than recursively substitutes a measurement error problem (because of the poor quality of the instruments that 2SLS produces in this case) for a simultaneity bias problem. It is true that asymptotically the measurement error problem goes away but in the present case I have only 51 observations, so it is not obvious that the tradeoff is favorable. The results, as noted in the text, do not change when two-stage least squares is used to estimate the main equation (in which the dependent variable is the effective number of elective parties). Probably this is so because there are only 16 presidential democracies in the sample of 51. This is hardly a standard simultaneous model in that the endogenous variables appear in interacted form, so that a substantial subset of the cases, those where PROXIMITY=0, are separably estimable. The results of the second equation, in which the effective number of presidential candidates is estimated, are changed substantially by using 2SLS. This is not too surprising, given that there are only 16 observations and one is adding two more variables to the specification. In any event, appealing to the asymptotic properties of the 2SLS estimator in this case seems even less justifiable, and so I prefer the OLS results.

less exert some direct influence on the legislative race. Thus, I prefer to include ENPRES as a regressor in the equation for ENPV, rather than including descriptors of presidential election rules (these rules, of course, do have an indirect impact via their influence on ENPRES). All told, my expectation is that legislative elections that are highly proximal to presidential elections should have a lower effective number of parties, but how much lower should depend on ENPRES. Thus I include both PROXIMITY and PROXIMITY*ENPRES in the analysis.[12]

A final point to note about Figure 11.1 is that it presupposes an interaction between electoral and social structure, both in the production of ENPV and in the production of ENPRES. If there is such an interaction in legislative elections, then there should also be an interaction in presidential elections, as argued in the last chapter.[13]

Specifying the equations. Having discussed the main differences of data and operationalization between my analysis and Ordeshook and Shvetsova's, I can turn to the issue of how I specify the relations of interest. I shall consider first the effective number of legislative parties (ENPS), then the effective number of elective parties (ENPV), and finally the effective number of presidential candidates (ENPRES).[14]

In investigating the first of these dependent variables (ENPS), I am interested in the purely mechanical features of how the legislative electoral system translates votes into seats. Accordingly, I include ENPV on the right-hand side (cf. Coppedge 1995). Indeed, the proper formulation is one in which ENPS would equal ENPV, were the electoral system perfectly proportional, with stronger electoral systems reducing ENPS below ENPV. Thus, I run the following regression:

$$\text{ENPS} = \alpha + \text{ENPV}^*(\beta_0 + \beta_1\text{LML} + \beta_2\text{UPPER}) + \varepsilon$$

If the electoral system employs single-member districts (so LML = 0) and has no upper tier (so UPPER = 0), then it is maximally strong, and only a fraction β_0 of ENPV is added to α to give the predicted effective number of

[12]My data on presidential structure are culled from Jones (1995), Mackie and Rose (1991), Nohlen (1993), and Santos (1990).

[13]In principle I would be happy to include the "effective number of prime ministerial candidates," if I could. In practice, I do not have the knowledge of each system that would be needed to code such a variable. Just taking the leader of each party as if he or she were an active candidate would not do; the whole point would be to see if there is widespread recognition that there are really only two (or three, four ...) viable contenders.

[14]My data on votes and seats (at the national level) were culled from Arms and Riley (1987), Brazil-Tribunal Superior Eleitoral (1990), Gorwin (1989), Mackie and Rose (1991), Nohlen (1993), Singh and Bose (1986), and Wightman (1990).

legislative parties in the system. As LML and UPPER increase, the system becomes more permissive and the fraction of ENPV that translates into seats should be greater. That is, the coefficients on ENPV*LML (i.e., β_1) and on ENPV*UPPER (i.e., β_2) should both be positive. One way to interpret this regression is simply as a check on the validity of the measures LML and UPPER. If LML properly measures the central tendency in lower-tier district magnitudes and UPPER really catches the impact of upper tiers, then the coefficients associated with both should be significant!

In the analysis of ENPV, I run five specifications: a pure institutionalist specification, with only variables pertaining to the legislative electoral system or the impact of presidential elections; a pure sociological model, with only a variable tapping into social heterogeneity (specifically, ENETH, the effective number of ethnic groups, used by Ordeshook and Shvetsova);[15] an additive model in which both sets of variables are included; an additive/interactive model in which an interaction term (between LML and ENETH) is added to the previous specification; and an interactive model in which the linear terms for LML and ENETH are omitted but the interaction term LML*ENETH is kept.

Finally, the analysis of ENPRES is as suggested in Figure 11.1. The main regressors are a dummy variable identifying runoff systems (RUNOFF), the effective number of ethnic groups (ENETH), and their interaction (RUNOFF*ENETH).

11.3 RESULTS

The main results are displayed in Tables 11.1, 11.2, and 11.3. Table 11.1 shows, not surprisingly, that a fair amount of the variance (93%) in ENPS can be explained by just ENPV and interactions between ENPV and two indicators of the strength of the electoral system – LML and UPPER. All variables have the expected sign and are statistically discernible from zero at about the .001 level or better. One way to explain the substantive impacts implied by the results in Table 11.1 is to compare two hypothetical systems, in neither of which there is an upper tier. System A has single-member districts, hence LML = 0. System B has ten-seat districts, hence LML = 2.3. Suppose that both systems have ENPV = 3 in a particular election. The stronger system (A) is predicted to reduce this number of elective parties by almost a full (effective) party, to

[15] I have also investigated the impact of the effective number of language groups and the effective number of religious groups, and various combinations of ethnic, religious, and linguistic heterogeneity, without finding significantly stronger results than those reported. My data on ethnic groups, language groups, and religious groups come from Olga Shvetsova (thank you!), the CIA *World Factbook* (1990, 1994), Vanhanen (1990), and the *Worldmark Encyclopedia of the Nations* (1984).

Table 11.1. The determinants of the effective number of legislative parties in 54 democracies, circa 1985

Dependent variable: ENPS *Independent variables*	*Estimated coefficients*	*Standard errors*
Constant	.582	.135
ENPV	.507	.048
ENPV*LML	.080	.012
ENPV*UPPER	.372	.111
Adjusted R^2 =	.921	
N of Obs =	54	

Table 11.2. The determinants of the effective number of electoral parties in 51 democracies, circa 1985

Dependent variable: ENPV					
			Model		
Independent variables	*1*	*2*	*3*	*4*	*5*
Constant	2.44	2.76	1.61	2.45	2.40
	(.25)	(.66)	(.47)	(.55)	(.21)
LML	.48	—	.52	−.23	—
	(.11)		(.11)	(.31)	
UPPER	3.64	—	3.95	3.46	3.51
	(1.52)		(1.48)	(1.40)	(1.37)
PROXIMITY	−5.98	—	−5.95	−6.01	−6.04
	(.97)		(.94)	(.89)	(.88)
PROXIMITY * ENPRES	2.18	—	2.14	2.05	2.09
	(.29)		(.28)	(.26)	(.26)
ENETH	—	.49	.51	.01	—
		(.40)	(.25)	(.30)	
LML*ENETH	—	—	—	.53	.39
				(.21)	(.07)
Adjusted R^2 =	.613	.01	.639	.679	.686
N =	51	51	51	51	51

Table 11.3. The determinants of the effective number of presidential candidates in 16 democracies, circa 1985

Dependent variable: ENPRES			
		Model	
Independent variables	*1*	*2*	*3*
Constant	2.26	4.30	2.68
	(.87)	(1.23)	(.36)
RUNOFF	.63	-2.49	—
	(.61)	(1.56)	
ENETH	.37	-.98	—
	(.50)	(.77)	
RUNOFF*ENETH	—	2.01	.58
		(.94)	(.29)
Adjusted R^2 =	-.015	.202	.171
N =	16	16	16

Table 11.4. Predicting the number of parties in Japan, 1960–90

Dependent variable: N OF PARTIES$_{it}$		
Independent variables	*Coefficient estimates (std err)*	*Coefficient estimates (std err)*
Constant	0.30 (.11)	1.02 (.29)
N OF PARTIES$_{i,t-1}$	0.55 (.02)	0.55 (.02)
MAGNITUDE$_{it}$	0.20 (.02)	0.01 (.07)
PURBE$_{it}$	.012 (.001)	0.15 (.41)
MAGNITUDE$_{it}$*PURBE$_{it}$	—	0.28 (.10)
ADJUSTED R^2	.60	.60
NUMBER OF OBS	1323	1323

Notes: N OF PARTIES$_{it}$ is the number of parties running candidates in district i, election t. MAGNITUDE$_{it}$ is the number of members elected from district i, election t. PURBE$_{it}$ is the percentage of all electors in district i, as of year t, who reside in urban areas. The estimation was performed using OLS and the unstandardized regression coefficients are reported in the table. Year dummies (as in Table 9.2) not reported. Results similar if EXCESS variable included.

Table 11.5. Predicting the effective number of parties in Japan, 1960–90

Dependent variable: $ENPV_{it}$ *Independent variables*	*Coefficient estimates (std err)*	*Coefficient estimates (std err)*
Constant	0.03 (.06)	0.41 (.17)
$ENPV_{i,t-1}$	0.76 (.02)	0.76 (.02)
$DISTRICT\ MAGNITUDE_{it}$	0.06 (.01)	-.04 (.04)
$PURBE_{it}$	.007 (.001)	0.15 (.24)
$MAGNITUDE_{it}$*$PURBE_{it}$	—	0.14 (.06)
ADJUSTED R^2	.80	.80
NUMBER OF OBS	1323	1323

Notes: $ENPV_{it}$ is the effective number of elective parties in district *i*, election *t*. The other variables are as defined in the notes to Table 11.4.

2.09 (shades of the United Kingdom in the 1980s!). The weaker system (B) is predicted to reduce the three effective parties competing in the election by much less, to 2.64 legislative parties. The substantive importance of this difference might vary from situation to situation, but it certainly suggests an important change from essentially a two-party legislative system with mostly single-party governments to a two-and-a-half or three-party legislative system with coalition governments as the norm.[16]

The results in Table 11.2 show the results for the five ENPV equations outlined in the Section 11.2. In running these regressions, I have omitted electoral systems with fused votes, that is, systems in which the voter casts a single vote for a slate that includes candidates for executive and legislative offices. The reason for omitting such systems is that they change the meaning of essentially all the institutional regressors. For example, do voters in such systems respond to the district magnitude at the legislative level or at the presidential level? Fused-vote systems really need to be analyzed separately (see Shugart [1985] for the case of Venezuela, which has a fused vote for senate and house races) but I do not attempt to do so here: I just omit the three cases of executive-leg-

[16]Because the translation of legislative votes into legislative seats is not affected by the existence (or not) of presidential elections in the system, nor by the number and character of social cleavages, these variables should not affect ENPS once ENPV is included. I have verified that they do not. I also note that a model that excludes the intercept term works slightly less well than the model with the intercept, in terms of the root mean squared error (.47 versus .40).

islative fused votes in the sample: Bolivia, Honduras, and Uruguay.[17] This reduces the number of observations to 51 for the regressions in Table 11.2. We shall discuss each briefly in turn.

The first model, with only institutional variables, explains about 61% of the variance in ENEP values. All coefficients are of the expected sign and significant at the .05 level or better. The second model, with only the effective number of ethnic groups (ENETH) as a regressor, produces a poor fit (an adjusted R^2 of .01) and an insignificant coefficient and regression. The third model, which combines the regressors from the first two, shows little change in the coefficients of the institutional variables but produces a coefficient on ENETH that is statistically significant at the .05 level. Apparently, proper controls for electoral structure are important in discerning any independent additive effect due to ethnic heterogeneity. The fourth model, which adds to the third an interaction term, LML*ENETH, reduces the coefficients on LML and ENETH to statistically insignificant values, while producing a substantial and statistically significant positive coefficient on the interaction term (LML*ENETH), together with little change in the coefficients of the remaining variables. Finally, the fifth model, in which the variables LML and ENETH are omitted, but their interaction is retained, produces a somewhat smaller interaction coefficient (but a substantially smaller interaction standard error), with other coefficients largely unchanged. If one chooses among specifications according to which produces the largest adjusted R^2 (not necessarily recommended; see the discussion in Kennedy [1994]), then the last specification, with an adjusted R^2 of .69, is the best.

I have also investigated a different formulation for the interactive term, using the minimum of LML and ENETH instead of their product. Substituting this minimum term for LML*ENETH in the last model produces little change in any of the other coefficients or in the overall fit of the equation. It is thus difficult on the basis of this study to say much one way or another about whether the form of the interaction should be thought of as a product or a minimum.[18]

[17]These cases did not need to be omitted in the first regressions because, once the votes are given, the translation to seats is via the legislative electoral system, so the variables LML and UPPER retain their meaning. Omitting these cases does not in any event change any of the previous results.

[18]Other variations in the model that I have explored include: introducing a dummy variable to identify the majoritarian systems in the sample (France and Australia) or, alternatively, coding them as $M = 2$ on the argument that they are similar to top-2 runoff systems; introducing a dummy variable to identify systems with primary elections (i.e., the United States); and introducing a population variable as another possible indicator of social diversity. None of these changes yields statistically significant results on the newly introduced variables and none of them changes the pattern of results described in the text.

Finally, Table 11.3 displays results for three regressions that take ENPRES as the dependent variable. The first model is additive, using RUNOFF and ENETH as regressors. As can be seen, neither regressor is statistically significant and the regression as a whole sports a negative adjusted R^2 (regressions with just RUNOFF and just ENETH are also insignificant). The second model adds the interaction term, RUNOFF*ENETH, to the first. The linear terms remain insignificant (albeit reversing sign) but the interaction term is appropriately signed and significant. The last model drops the linear terms, keeping only the interaction; the coefficient on the interaction term is again positive and statistically discernible from zero in a one-tailed test at the .05 level.

11.4 THE INTERACTION OF SOCIAL AND ELECTORAL STRUCTURE IN JAPAN

The evidence presented in the last section shows a substantial interaction between a national-level measure of electoral permissiveness (the magnitude of the median legislator's district) and a national-level measure of social diversity (the effective number of ethnic groups). Does this interaction work because the national-level variables reflect the typical situation in the districts, and the logic plays out at the district level (with especially large numbers of parties appearing within districts that combine high magnitudes with social diversity)? Does it work because the measure of electoral permissiveness in legislative elections covaries with the permissiveness of the executive choice procedure in each polity, and the logic plays out at the national level (with especially large numbers of executive candidates, hence especially low levels of cross-district legislative linkage, in countries that combine permissiveness in the executive choice procedure with social diversity)? Or is it some combination of effects at both levels that produces the result?

In this section, I address this issue by examining district-level evidence (on the number of parties, social diversity, and district magnitude) from Japan 1960–90. In Japan at this time, the big cities and countryside alike were carved up into a number of medium-magnitude districts (the range was from 1 to 6, with the vast bulk of districts between 3 and 5). Importantly for present purposes, there were districts of any given magnitude in both urban and rural areas. One can thus hope to tease out any interaction effects between social and electoral structure that there might be at the district level.

Operationally, I use two different measures of the number of parties: a simple count of the number of parties fielding candidates in each district; and the effective (or vote-weighted) number of candidates. The independent variables of primary interest are two. First, I use the district magnitude to assess the permissiveness of the local electoral system. Second, I use

the percentage of each district's registered electorate that lives in urban places to measure social heterogeneity, assuming (in accord with conventional wisdom in the literature) that the more urban districts are also the more heterogeneous ones. I also include, as controls, dummy variables for each year (except the first) and the lagged dependent variable (to deal with autocorrelation). I use multiple regression to estimate the specified relationship, using data from elections occurring in the 1960–90 period.

The results displayed in Table 11.4 take the scalar number of parties as the dependent variable. As can be seen, one can explain about 60% of the variance with the percent urban, district magnitude, lagged dependent variable, and year effects (Model 1). Increasing district magnitude by one seat increases the number of parties by .2.[19] Increasing the urbanness of a district by ten percent increases the number of parties by .12. As both effects are statistically significant, there is some evidence here for the importance of both electoral and social structure.

Model 2 in Table 11.4 adds the interaction of district magnitude and urbanness. As can be seen, the additive terms are no longer statistically significant but the interactive term is, while the overall statistical "fit" of the model is unaffected. That is, it appears to be the *product* of social heterogeneity (proxied by the urbanness of the registered electorate) and electoral permissiveness (proxied by the district magnitude) that produces a large number of parties, not the *sum* of these factors.

Similar results are obtained when the effective number of elective parties (ENEP) is used as the dependent variable. Model 1 in Table 11.5 shows that a model with district magnitude and urbanness modeled as having additive effects explains about 80% of the variance in ENEP, with both additive effects statistically significant. Model 2 shows that, when one allows for an interactive effect, both additive terms fall to insignificance and only the interactive term is significant.

11.5 CONCLUSION

In this chapter I have developed an econometric model that predicts the effective number of parties at the national level based on variables drawn from both the district and the national levels, and from both institutional and sociological perspectives. The discussion in Part II of the book

[19]One might ask why district magnitude does not have a bigger impact. After all, would not the $M + 1$ rule suggest that increasing district magnitude by one should increase the number of parties by one, i.e., that the coefficient on district magnitude should equal one, instead of about a fifth as reported in Table 11.4? There are two answers to this. First, the $M + 1$ rule refers to *candidates,* not parties. Second, the rule only imposes an upper bound in any event, so there is no necessary prediction of growth with M.

suggested district magnitude as the key indicator of the nature of the district-level electoral system. The discussion in Chapter 7 suggested that the district-level number of parties would be an interactive function of social diversity and electoral permissiveness. Both of these district-level considerations are embodied in the model. At the national level, the key institutional variables outlined in Chapter 10 pertained to the executive choice procedure and the existence of upper tiers. And, again, I suggested that there should be an interaction between social and electoral structure in determining the effective number of executive candidates. Both of these national considerations are also embodied in the model.

The results presented in the previous sections are remarkably similar to those generated by Ordeshook and Shvetsova (1994). Despite using a different data set – one that included many new and developing democracies rather than concentrating on the long-term democracies – and despite several differences in operationalization and specification, the basic result holds up: The effective number of parties appears to depend on the product of social heterogeneity and electoral permissiveness, rather than being an additive function of these two factors.[20] The intuitive formulation of this finding is that a polity can tend toward bipartism either because it has a strong electoral system or because it has few cleavages. Multipartism arises as the joint product of many exploitable cleavages and a permissive electoral system.

If this general conclusion is valid, it ought to hold, not just for elections to the lower house of the national legislature, but also for other elections. And there is a bit of evidence consistent with the notion that the effective number of presidential candidates is an interactive product of social and electoral structure. In particular, elections that are *both* held under more permissive rules (runoff rather than plurality) *and* occur in more diverse societies (with a larger effective number of ethnic groups) are those that tend to have the largest fields of contestants for the presidency.

The model also further documents the impact of two national-level variables on the legislative party system: the existence and nature of upper tiers; and the existence, nature, and timing of presidential elections. Here, I model the effect of executive choice rules as being indirect, rather than direct, as in the previous literature.

A second model that I explore in this chapter concerns district-level data from Japan. My results provide additional support for the hypothesis of this chapter, that multipartism arises as a joint product of social diversity and electoral permissiveness.

[20]Or, to take account of the results with the minimum of LML and ENETH just mentioned, perhaps one should say that the effective number of parties depends on an interaction between electoral and social structure.