Appendix 8e – Multinomial Probit and Multinomial Logit Analysis of Vote Choice in 1996

I utilize a multinomial probit (MNP) model to examine what factors shaped voter choice in the 1996 HR election. I divide the dependent variable into three categories: Old Left, LDP, and New Parties. I focus in particular on eight independent variables. The first two are based on the characteristics of the voters. (1) *Union.* Is the voter a member of a union? I expect union members to be more likely to support new parties over the LDP, but to support the old Left most of all. (2) *Age*. Older voters ought to be more likely to support the traditional parties.

I also expect voters' opinions on issues will provide a clue into the parties they are most likely to support. (3) *Liberalization*. Within the JEDS data set, this variable is based on a question in which respondents were asked to what extent they supported the liberalization or protection of agricultural markets. Within the data set, it is the most explicit variable addressing voters' support for liberalization. Responses run from 1 to 4, with 1 indicating greater protectionism and 4 indicating greater support for liberalization. I expect that as voters grow more supportive of liberalization, they will be more likely to support the new parties. Therefore, the regression coefficient for views on liberalization ought to be positive for new parties and negative for Japan's long-existing parties. (4) *Subsidy*. Respondents were asked to name candidates in their district and offer a description of them. My subsidy variable indicates that the respondent described the first candidate they could think of as being good at delivering subsidies. The variable takes the form of a 0-1 dummy variable. Although technically the answer indicates only what respondents notice about their candidates and not what their own opinion of the candidate's position is, if we assume that respondents notice what they care about, there is a good chance that Subsidy provides a good proxy for the importance of subsidies to the respondent. I therefore expect that voters who cite its importance will be more likely to support the LDP. (5) *Decentralization*. Did the respondent mention favoring an increase in regional or local autonomy? For reasons mentioned earlier in Chapter 6, I expect voters favoring decentralization to be more likely to vote for the new parties.

(6) *Economy*. Did the respondent think the economy worsened in the previous few years? I expect that voters who felt that the economy was getting worse were likely to blame the LDP and therefore vote for a new party. *Economy* increases in value as the respondent's impression of the economy grows worse. Therefore, I expect a negative sign for the coefficient in models predicting support for the LDP. (7) *Standard of Living (SOL)*. Did the respondent feel that her standard of living had worsened in the previous few years? Those who did ought to be more likely to support the new parties. SOL increases in value as the respondent grows more unsatisfied with his or her standard of living, so we ought to see a negative sign in the model predicting support for the LDP. (8) *Ideology*. This survey item is a feeling thermometer score chosen by the respondent, with 0 being the most progressive and 10 being the most conservative. I expect that the most progressive respondents will vote for the old Left, and the most conservative will vote for the LDP.[1] I also include three 0-1 dummy variables, with each indicating the presence or absence of an incumbent from the LDP, a new party, or an old Left party running in the respondent's SMD. All else being equal, voters ought to be more likely to cast ballots for incumbents (in SMD races) and the incumbent's party (in PR races).

---

[1] The question was phrased: "The words conservative and progressive are used to express some positions in politics. If progressive is a '0' and conservative is a '10' what number do you think you might best indicate your own position on these matters?"

Results for the model of SMD vote choice are listed in Table 8.6 and the results for the PR model are in Table 8.7. In the tables, support for new parties (NFP or DPJ) is the base category. A positive sign, therefore, indicates a greater likelihood of supporting the party in question ("Old Left"—Socialist or Communist—or LDP) than of supporting a new party. With the exception of Economy (and SOL for the Old Left)—which is positive, but non-significant—the sign on every variable's coefficient in Table 8.6 is in the expected direction and most of the key variables are statistically significant.

As predicted, Table 8.6 indicates that the presence of incumbents clearly influences one's SMD vote choice. As expected, and as the coefficient and significance of Ideology indicates, more conservative voters cast ballots for LDP candidates and more progressive voters turned to old Left candidates. Union is positive and significant for Old Left voters. Although Economy is not significant, SOL (standard of living) is negative and significant for the LDP half of the model.

The results of the analysis for the remaining variables are probably the most interesting, as they indicate a clear set of differences between people who vote for Japan's long-existing parties and those who support the country's new parties. The significant and positive sign on Age indicates that older voters were clearly more likely to support the old Left and LDP. The significant and negative sign on Liberalization in each section of the model indicates that voters who were more supportive of liberalizing agricultural markets were also more likely to cast votes for new party candidates. The significant and positive sign on Subsidy in each section of the model indicates that supporters of subsidies were also more likely to cast ballots for old Left and LDP candidates. The negative sign on Decentralization (which is significant in the LDP portion of the model) indicates that voters supporting decentralization were more likely to vote for new party candidates.

Multinomial probit analysis results of PR voting are listed in Table 8.7. The overall results are quite close to those in Table 8.6. By and large, similar patterns underlie SMD candidate and PR party voting. The main differences are: (1) Union is significant in the LDP portion of the PR model, but not in the SMD model. (2) While non-significant in the model used to analyze SMD races, Decentralization is significantly and (as in the SMD model) negatively related to support for the old Left in the PR voting. (3) With the exception of Socialist/Communist incumbency in the old Left portion of the model, incumbency is non-significant in the PR model.

[Tables 8.6 and 8.7 about here]

Identically specified multinomial *logit* (MNL) models (see Tables 8.8 and 8.9) obtain results very similar (especially in terms of significance levels) to the results of the multinomial probit models reported in Tables 8.6 and 8.7. Generating predicted probabilities based on multinomial probit results can be cumbersome, but simulating probabilities based on multinomial logit results using Clarify simulation software (King, et al. 1998; Tomz, et al. 1999) is quite simple. Of the two models, multinomial probit is often considered the more appropriate statistical method for vote choice models because it does not assume the independence of irrelevant alternatives (IIA) (see. e.g., Alvarez and Nagler 1995). Nevertheless, Dow and Endersby (2004) demonstrate that when used to study voter choice in multiparty/candidate elections MNL performs at least as well as MNP. For this reason, given the similarities in the levels of statistical significance in the two different methods here, I feel comfortable using the

results of the multinomial logit models as an approximation for simulating the predicted probabilities.

Using Clarify, I take the identically specified multinomial logit results, hold constant the other variables while adjusting the "clientelism" variables (Subsidy, Liberalization, and Decentralization) to simulate the probability of a voter casting a ballot for SMD candidates from the LDP, an old Left party, or a new party in 1996. I should note that in all the predicted probability simulations, no party has an incumbent running (i.e., incumbency for party is held at 0). The presence of incumbent candidates of course has a substantial impact on the probability of a voter casting a ballot for any party, but this appears to help all parties.

In generating the predicted probabilities in Figure 8.5, I defined a clientelistic voter as one who scores 1 on the subsidy variable, does not support liberalizing the agricultural market at all, and does not support decentralization. I define anti-clientelist voters as ones with the opposite positions.

In generating predicted probabilities (utilizing the multinomial logit results) for voter support for SMD candidates at different levels of urban-ness (Figure 8.6), I computed the predicted probabilities for any given region by using the region's mean for each variable (with incumbency held at 0). That is, I used the mean value for each variable in rural areas to compute the probability of voting for each party in rural districts; I used the mean in mixed areas to compute the probabilities for mixed districts; and I used the mean value for urban areas to compute the probabilities for urban districts.

Finally, I ran similar, but more complex, multinomial logit models in which I divided the new parties into two categories in order to examine the NFP and DPJ each separately. These models offer similar results to those obtained in the above 3-category (i.e., where I assume voters choose between three types of party options: Old Left, LDP, and new parties) model. However, to deal with issues surrounding candidate coordination and competition between the new parties, it was necessary to add a variable for each party to indicate whether it ran a candidate at all. When I simulated only one new-party candidate running in a district, the 4-category model (where the choices are Old Left, LDP, NFP and DPJ) obtained results similar to the 3-category model. But when the simulation included both the NFP and DPJ running candidates, the new parties almost invariably were likely to lose to the LDP, if for no other reason than they split the new-party vote. As I note in Chapters 2 and 5, this coordination dilemma was far less severe in 2000, as the NFP had died off and the DPJ was the primary challenger in the opposition.

Table 8.6 – Multinomial Probit Estimates for Vote Choice in SMDs in 1996

Number of Observations	744	
Log Likelihood	-214.538	
Variable	**Old Left/NFP**	**LDP/NFP**
Union	0.753***	-0.098
	(0.256)	(0.166)
Age	0.009*	0.005**
	(0.006)	(0.002)
Liberalization	-0.230***	-0.625**
	(0.069)	(0.310)
Subsidy	0.064	0.282***
	(0.120)	(0.121)
Decentralization	-0.110	-0.254**
	(0.212)	(0.137)
LDPinc	-0.605***	1.056***
	(0.170)	(0.309)
Newinc	-0.246**	-1.096***
	(0.132)	(0.246)
Leftinc	0.954***	-0.230
	(0.306)	(0.215)
Economy	0.104	0.564
	(0.096)	(0.613)
SOL	0.103	-0.130***
	(0.081)	(0.051)
Ideology	-0.390***	0.188***
	(0.042)	(0.023)
Constant	0.097	-0.104***
	(0.437)	(0.037)
σ Left/New Parties	-1.028**	
	(0.510)	
σ LDP/New Parties	1.8246***	
	(0.479)	

Results are coefficients with standard errors in parentheses.

*p<.05 (one-tail); **p<.05 (two-tail), ***p<.01 (two-tail)

Table 8.7 – Multinomial Probit Estimates for Vote Choice in PR in 1996

Number of Observations	824	
Log Likelihood	-1319.2	
Variable	**Old Left/NFP**	**LDP/NFP**
Union	0.870***	-0.868***
	(0.126)	(0.225)
Age	0.991***	0.986***
	(0.009)	(0.007)
Liberalization	-1.034***	-0.958***
	(0.056)	(0.134)
Subsidy	1.069***	0.932***
	(0.158)	(0.099)
Decentralization	-0.851***	-1.149***
	(0.171)	(0.221)
LDPinc	-0.072	0.153
	(0.133)	(0.120)
Newinc	-0.005	0.086
	(0.106)	(0.114)
Leftinc	1.127***	-0.126
	(0.206)	(0.311)
Economy	-0.212	-0.191*
	(0.173)	(0.125)
SOL	-0.204	-0.468***
	(0.202)	(0.099)
Ideology	-0.280***	0.284***
	(0.031)	(0.091)
Constant	-0.533	-1.459***
	(0.526)	(0.436)
σ Left/New Parties	-0.844***	
	(0.125)	
σ LDP/New Parties	0.748***	
	(0.084)	

Results are coefficients with standard errors in parentheses.

*p<.05 (one-tail); **p<.05 (two-tail), ***p<.01 (two-tail)

Table 8.8 – Multinomial Logit Estimates for Vote Choice in SMDs in 1996

Number of Obs.	744	
LR Chi-Squared (22)	261.44	
Prob>Chi-Sq.	0.000	
Pseudo R-Sq.	0.1759	
Variable	**Old Left/NFP**	**LDP/NFP**
Union	0.979***	0.057
	(0.363)	(0.317)
Age	0.022**	0.015**
	(0.009)	(0.007)
Liberalization	-0.462***	-0.224***
	(0.115)	(0.081)
Subsidy	0.525*	0.594***
	(0.318)	(0.225)
Decentralization	-0.315	-0.472*
	(0.335)	(0.243)
LDPinc	-0.651**	0.403**
	(0.272)	(0.189)
Newinc	-0.729***	-0.743***
	(0.214)	(0.147)
Leftinc	1.095**	-0.208
	(0.463)	(0.415)
Economy	0.153	0.117
	(0.172)	(0.114)
SOL	-0.014	-0.239**
	(0.133)	(0.094)
Ideology	-0.414***	0.226***
	(0.063)	(0.045)
Constant	0.972	-0.856
	(1.003)	(0.693)

Results are coefficients with standard errors in parentheses.

*p<.05 (one-tail); **p<.05 (two-tail), ***p<.01 (two-tail)

Table 8.9 – Multinomial Logit Estimates for Vote Choice in PR in 1996

Number of Obs.	824	
LR Chi-Squared (22)	229.82	
Prob>Chi-Sq.	0.000	
Pseudo R-Sq.	0.1355	
Variable	**Old Left/NFP**	**LDP/NFP**
Union	0.821***	-0.635**
	(0.298)	(0.318)
Age	0.026***	0.019***
	(0.008)	(0.006)
Liberalization	-0.271***	-0.166**
	(0.096)	(0.076)
Subsidy	0.521**	0.360*
	(0.262)	(0.212)
Decentralization	-0.825***	-0.877***
	(0.305)	(0.244)
LDPinc	-0.112	0.231
	(0.230)	(0.181)
Newinc	-0.212	-0.146
	(0.178)	(0.138)
Leftinc	0.976**	-0.241
	(0.401)	(0.413)
Economy	-0.036	-0.048
	(0.140)	(0.107)
SOL	-0.007	-0.310***
	(0.112)	(0.090)
Ideology	-0.299***	0.278***
	(0.052)	(0.043)
Constant	0.035	-0.972
	(0.850)	(0.648)

Results are coefficients with standard errors in parentheses.

*p<.05 (one-tail); **p<.05 (two-tail), ***p<.01 (two-tail)