Web Appendix for "Was Ralph Nader a Spoiler? A Study of Green and Reform Party Voters in the 2000 Presidential Election"

Michael C. Herron[1] Jeffrey B. Lewis[2]

September 20, 2006

[1]Department of Government, Dartmouth College, and Department of Political Science, Northwestern University. HB 6108, Hanover NH 03755 (Michael.C.Herron@Dartmouth.edu).

[2]Department of Political Science, University of California, Los Angeles. 4289 Bunche Hall, Los Angeles, CA 90095-1472 (jblewis@ucla.edu).

A Ballot counts by County

Table A lists the ten counties whose ballots are recorded in the NES ballot image archive, and the table also provides various summary statistics for each county based on presidential vote totals (based on the ballot images). We do not assume that election day and absentee voters have similar partisanship patterns, and Table A therefore breaks down each county's vote totals into election day and absentee totals.

Valid Presidential Votes among Ten Florida Counties

	Election Day				Absentee				
County	Bush	Gore	Nader	Buchanan	Bush	Gore	Nader	Buchanan	Total
Broward	156876	359255	6512	706	20447	27306	589	82	571773
Highlands	12379	8709	359	84					21531
Hillsborough	157367	155327	6898	774	11161	7390	301	43	339821
Lee	92665	67188	3270	258	13462	6346	314	48	183551
Marion	47324	40652	1632	507	7329	3712	170	55	101381
Miami-Dade	265211	311879	5054	516	20882	14138	251	41	617972
Pasco	59881	64096	3107	514	8701	5468	286	56	142109
Palm Beach	130688	241806	5014	3277	16819	19052	390	81	417127
Pinellas	158380	183138	9162	884	26445	17493	860	129	396491
Sarasota	70726	64363	3669	265	12374	8491	400	40	160328
Total	1151497	1496413	44665	7785	137620	109396	3561	575	2952084

Table 1: *Data extracted from the NES Ballot Image Archive (*`http://www.umich.edu/~nes/florida2000/data/data_files.htm`*).* Marginal totals do not include votes for seven other minor candidates, undervotes, or overvotes. The grand total number of cast ballots in data set is 3,066,843.

The NES Florida ballot image archive is broken down by county, and each county has what is called a *balfile* for each of its precincts. For instance, Pasco County has 148 balfiles, one per election day precinct (total of 131) and one per absentee precinct (17). The balfiles for the ten counties in Table A were assembled by Dan Keating of *The Washington Post*, but some of the balfiles are audit files, i.e., they represent a set of Votomatic punchcards that were counted in order to verify ballot counter accuracy. Audit files can override real election balfiles, and we do not want to include audit balfiles in our analysis. Keating's documentation offers suggestions on how to do this for Palm Beach, Miami-Dade, Hillsborough, Pasco, and Highlands Counties; the complement of this group is not known to have audit files in the archive. We followed Keating's rules on eliminating audit balfiles except forthose rules on Hillsborough County. For Hillsborough, and based on a phone conversation with an official in the Hillsborough Supervisor of Elections office, we treated as audit balfiles all precinct balfiles that had exactly 59 ballots in them. We also dropped all precincts to which Keating affixed an "X" and dropped all absentee precincts that did not correspond exactly with official Hillsborough results. Finally, we dropped Hillsborough precincts labeled R433C and R433C1. We have not uncovered any reasons to think that missing balfiles are systematically different than those that are not missing.

B The voting model and estimator

Consider an electorate comprised of voters with quadratic preferences or ideal points over a single partisan dimension. In an election each voter casts a vote in contests $k = 1, 2, \ldots, K$ where each contest includes J_k candidates. Index the candidates in each contest by $j = 1, 2, \ldots, J_k$, and let c_{jk} be the location on the partisan dimension of candidate j in race k. The utility of candidate j to voter i with ideal point θ_i in race k is

$$U(\theta_i, c) = v_{jk} - (\theta_i - c_{jk})^2 + \epsilon_{ijk}$$

where v_{jk} reflects non-spatial utility provided by candidate j (the value of incumbency, for example) and ϵ_{ijk} is an idiosyncratic utility shock which is assumed to follow a type I extreme value distribution. Voter i selects candidate j in race k if

$$U(\theta_i, c_{jk}) \geq U(\theta_i, c_{j'k}) \text{ for } j' = 1, 2, \ldots, J_k.$$

Note that the continuity of the distribution of idiosyncratic shocks (ϵ) insures that $U(\theta_i, c_{jk}) = U(\theta_i, c_{j'k})$ for $j \neq j'$ is a zero probability event; thus, the possibility that a voter is indifferent between two candidates can therefore be ignored.

As shown by McFadden (1974), the assumption of independent type I extreme value shocks means that the probability a voter with ideal point θ supports candidate j in race k is

$$\Pr(Y_k = j|\theta) = \frac{\exp(v_{jk} - (c_{jk} - \theta)^2)}{\sum_{j'=1}^{J} \exp(v_{j'k} - (c_{j'k} - \theta)^2)} \text{ for } j = 1, 2, \ldots, J.$$

Rearranging shows that

$$\Pr(Y_k = j|\theta) = \frac{\exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta) exp(-\theta^2)}{\exp(-\theta^2) \sum_{j'=1}^{J} \exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)} = \frac{\exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)}{\sum_{j'=1}^{J} \exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)}.$$

Letting $\alpha_{jk} = v_{jk} - c_{jk}^2$ and $\beta = 2c_{jk}$ produces the familiar multinomial logit model,

$$\Pr(Y_k = j|\theta, \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k) = \frac{\exp(\alpha_{jk} + \beta_{jk}\theta)}{\sum_{j'=1}^{J} \exp(\alpha_{jk} + \beta_{jk}\theta)}.$$

Conditional on θ, votes for candidates across races are independent. Thus, the probability of voting for a particular set of candidates can be written as

$$P(j_1, j_2, \ldots, j_K|\theta, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \Pr(Y_1 = j_1, Y_2 = j_2, \ldots, Y_K = j_K) = \prod_{k=1}^{K} \Pr(Y_k = j|\theta, \boldsymbol{v}_k, \alpha_k, \boldsymbol{\beta}_k).$$

These probabilities could be evaluated if θ were observed. However, we treat θ as a random variable. In particular, we partition our approximately three million voters into 240 county, presidential-vote, voting-place (election day or absentee) triples. For example, one such group is (Gore, Broward, Election Day). The distribution of ideal points θ within each of these groups $g = 1, 2, \ldots, 240$ is assumed *a priori* to be normal with mean μ_g and standard deviation one. Thus, *a priori* we assume the distribution of preferences within each voter group differs only by a shift

parameter μ. Fixing the standard deviation to one results in no further loss of generality as this restriction only serves to identify the otherwise unidentified units of the underlying dimension.

We marginalize the distribution of vote choices of voters from each group g with respect to θ so that

$$P(j_1, j_2, \ldots, j_K|\boldsymbol{\alpha}, \boldsymbol{\beta}, \mu_g) = \int P(j_1, j_2, \ldots, j_K|\theta, \boldsymbol{\alpha}, \boldsymbol{\beta})\phi(\theta|\mu_g)d\theta.$$

The parameter matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and the parameter vector μ are estimated by marginal maximum likelihood. We group voters into common patterns of votes cast across the k offices within each group. Let n_{pg} be the number of voters casting vote pattern p in group g, and let $p(g, k)$ represent the candidate in the kth race chosen by a voter in the gth group casting the pth voter pattern. Then, the loglikelihood is

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}) = \sum_g \sum_p n_{pg} \ln P\left(p(g, 1), p(g, 2), \ldots, p(g, K)|\boldsymbol{\alpha}, \boldsymbol{\beta}\right).$$

This likelihood can be maximized by standard numerical techniques (using Gauss-Hermite quadrature to approximate the normal integral) or by an EM approach.

To identify the model, the valance and position of the first candidate in each race is normalized to zero. That is, we assume $v_{1k} = 0$ and $c_{1k} = 0$ (and consequently $\alpha_{1k} = 0$ and $\beta_{1k} = 0$) for $k = 1, 2, \ldots, K$. This is a standard normalization used in multinomial logit models. Candidate valances and locations can then be thought of as relative to the valance and location of the first candidate in each race. It should be noted that this precludes the possibility of directly comparing valances and locations of candidates across races. For such comparisons additional identifying restricts are required. For example, the average valance might be assumed to be zero in each race. Finally, and without loss of generality, we linearly transform our estimates of μ so that they range from -1 to 1 where $\mu = -1$ is the most Democratic group and $\mu = 1$ for the most Republican group.

Standard errors for our likelihood problem are arrived at through a non-parametric bootstrap in which individual voters are sampled with replacement from the full dataset. The sampling is stratified by voter type so that each voter type appears as many times in each bootstrap sample as in our original data. Standard error estimates are based on 70 bootstrap samples. Confidence ellipses presented in the paper are based the assumption of normal sampling distributions and the covariance matrix of the estimates derived from the bootstrap.

Given estimates of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\mu}$, the posterior distribution of θ for a given voter group conditional on observed vote choices is calculated by Bayes' rule. The allocation of voters to Bush and Gore are made by finding the estimated probability that a voter located at each position voted for Gore conditional on voting for either Bush or Gore and then taking the expectation of these conditional probabilities over various posterior distributions.

C Allocating minor-party voters to Bush and Gore

Allocating minor party votes, overvotes, and undervotes

	Palm Beach		Broward		Highlands	
Choice	Pct. Dem.	Swing	Pct. Dem.	Swing	Pct. Dem.	Swing
Absentee						
Green	0.57	58	0.63	155		
Reform	0.51	1	0.42	-13		
Constitution			0.49	0		
Libertarian	0.42	-12	0.52	4		
Natural Law	0.63	3	0.66	5		
Socialist	0.84	8	0.85	2		
Socialist Workers			0.82	5		
Workers World	0.70	3	0.60	3		
Undervote	0.62	241	0.62	315		
Overvote	0.73	108	0.83	257		
Total	0.61	410	0.64	733		
Election Day						
Green	0.62	1234	0.64	1826	0.52	11
Reform	0.83	2175	0.48	-34	0.34	-27
Constitution	0.45	-16	0.46	-6	0.05	-5
Libertarian	0.52	21	0.57	164	0.50	-0
Natural Law	0.63	35	0.66	37	0.65	3
Socialist	0.86	205	0.71	13	0.70	1
Socialist Workers	0.72	20	0.78	25	0.31	-1
Workers World	0.62	22	0.69	41	0.65	2
Undervote	0.63	2319	0.74	2578	0.51	6
Overvote	0.81	11501	0.86	5433	0.58	54
Total	0.73	17515	0.73	10086	0.51	44

Table 2: *Allocating election day and absentee presidential undervotes, overvotes, and votes for minor party candidate Bush and Gore. The "Pct. Dem." column is the fraction of voters allocated to Gore, and the "Swing" column shows the net difference in votes allocated to Gore over Bush.*

Allocating minor party votes, overvotes, and undervotes (cont.)

	Hillsborough		Lee		Marion	
Choice	Pct. Dem.	Swing	Pct. Dem.	Swing	Pct. Dem.	Swing
Absentee						
Green	0.59	57	0.60	61	0.52	7
Reform	0.34	-14	0.54	4	0.32	-20
Constitution	0.10	-7	0.61	0	0.28	-0
Libertarian	0.42	-9	0.57	10	0.36	-11
Natural Law	0.49	-0	0.62	3	0.09	-2
Socialist	0.37	-1			0.98	1
Socialist Workers			0.81	3	0.72	1
Workers World	0.50	-0	0.66	2	0.74	2
Undervote	0.53	45	0.43	-77	0.41	-106
Overvote	0.72	28	0.38	-21	0.58	4
Total	0.54	99	0.49	-15	0.42	-124
Election Day						
Green	0.59	1202	0.57	448	0.53	97
Reform	0.37	-214	0.41	-45	0.41	-88
Constitution	0.37	-15	0.31	-12	0.13	-15
Libertarian	0.39	-224	0.58	78	0.40	-66
Natural Law	0.60	42	0.72	30	0.53	2
Socialist	0.75	22	0.77	2	0.58	1
Socialist Workers	0.52	2	0.54	2	0.64	3
Workers World	0.58	22	0.61	21	0.47	-2
Undervote	0.57	564	0.51	39	0.48	-70
Overvote	0.71	1475	0.45	-234	0.40	-297
Total	0.58	2885	0.52	326	0.46	-435

Table 3: *Continued from Table 2.*

Allocating minor party votes, overvotes, and undervotes (cont.)

	Miami–Dade		Pasco		Pinellas	
Choice	Pct. Dem.	Swing	Pct. Dem.	Swing	Pct. Dem.	Swing
Absentee						
Green	0.71	107	0.57	38	0.63	220
Reform	0.42	-6	0.42	-8	0.39	-29
Constitution	0.18	-2			0.19	-6
Libertarian	0.44	-5	0.46	-3	0.43	-16
Natural Law	0.36	-3	0.40	-1	0.59	9
Socialist	0.34	-1	0.62	0	0.52	0
Socialist Workers	0.60	0			0.80	3
Workers World	0.67	2	0.85	3	0.60	4
Undervote	0.56	136	0.44	-36	0.47	-73
Overvote	0.69	105	0.59	18	0.63	53
Total	0.59	334	0.50	11	0.52	165
Election Day						
Green	0.66	1664	0.58	490	0.63	2456
Reform	0.43	-73	0.46	-46	0.42	-145
Constitution	0.41	-11	0.40	-3	0.29	-27
Libertarian	0.54	57	0.49	-7	0.47	-63
Natural Law	0.61	25	0.59	13	0.67	134
Socialist	0.70	13	0.70	5	0.71	10
Socialist Workers	0.75	43	0.55	1	0.64	10
Workers World	0.71	49	0.60	14	0.59	27
Overvote	0.72	7681	0.68	734	0.71	1741
Undervote	0.65	2836	0.57	198	0.60	569
Total	0.68	12292	0.59	1399	0.62	4686

Table 4: *Continued from Table 2.*

Allocating minor party votes, overvotes, and undervotes (cont.)

	Sarasota		Total	
Choice	Pct. Dem.	Swing	Pct. Dem.	Swing
Absentee				
Green	0.62	97	0.61	800
Reform	0.25	-20	0.40	-105
Constitution	0.15	-1	0.25	-16
Libertarian	0.47	-3	0.46	-45
Natural Law	0.51	0	0.56	12
Socialist	0.35	-0	0.66	9
Socialist Workers			0.80	12
Workers World	0.52	0	0.61	194
Undervote	0.42	-80	0.52	365
Overvote	0.45	-6	0.69	546
Total	0.49	-13	0.55	1766
Election Day				
Green	0.60	739	0.61	10173
Reform	0.44	-30	0.59	1474
Constitution	0.18	-8	0.38	-118
Libertarian	0.44	-46	0.49	-82
Natural Law	0.63	21	0.64	343
Socialist	0.91	3	0.82	275
Socialist Workers	0.64	3	0.67	109
Workers World	0.49	-1	0.61	195
Undervote	0.53	66	0.52	9109
Overvote	0.55	100	0.74	28197
Total	0.56	847	0.57	51275

Table 5: *Continued from Table 2.*

References

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand." *Journal of Public Economics* 3:308–328.