Coordination, Moderation, and Institutional Balancing in American Presidential and House Elections

by

Walter R. Mebane, Jr.

Cornell University

forthcoming, *American Political Science Review*

Abstract

Voters have been coordinating their choices for President and House of Representatives in recent presidential election years, with each voter using a strategy that features policy moderation. Coordination is defined as a noncooperative rational expectations equilibrium among voters, in which each voter has both common knowledge and private information about the election outcome. Stochastic choice models estimated using individual-level NES data from 1976–96 support coordination versus a model in which voters act nonstrategically to moderate policy. The empirical coordinating model satisfies the fixed-point condition that defines the common knowledge expectation voters have about the outcome in the theoretical equilibrium. The proportion of voters who split their tickets in order to balance the House with the President has been small but large enough to affect election outcomes. Moderation has usually been based on voters' expectations that the President will be at least an equal of the House in determining postelection policy.

Do Americans coordinate their votes for President and for Congress? Suppose a voter knows that policy outcomes are compromises between the positions taken by the President and Congress and believes the two political parties push for distinct policy alternatives. Then the voter believes that different combinations of party control of the presidency and Congress will produce different policy outcomes. If the voter cares about the policy outcomes, then the voter's choices among candidates should depend on the combination of party control the choices will help bring about. Some voters would do best to vote a party-line, straight ticket, while others would do best to split their tickets to try to bring about a "moderate" policy outcome, that is, an outcome in some sense between the parties' positions (Alesina and Rosenthal 1989, 1995, 1996; Fiorina 1988, 1992, 73–81; Ingberman and Villani 1993). To cast the most effective vote, each voter should consider how likely it is that the election will produce each possible combination of party control. Coordination occurs when each voter makes the best possible assessment of what the election outcome will be and then uses that assessment to choose candidates in the way most likely to produce the best possible result for the voter.

Moderation does not necessarily imply coordination. Moderation refers to the relationship between the policy outcome intended by the voter and the parties' policy positions. There is moderation if the intended policy outcome is an intermediate combination of the parties' positions. Coordination refers to a relationship among different voters' choices among candidates.[1] There is coordination if each voter's choice is in a strategic sense in equilibrium with every other voter's choice. I interpret this to mean three things: Each voter's strategy for choosing among candidates is in equilibrium with every other voter's strategy for choosing among candidates; each voter's beliefs about every other voter's preferences and strategy are compatible with the voter's own strategy; and each voter's beliefs are compatible with every other voter's beliefs.

Coordination among voters is not the same as cooperation among them. In a financial market, economic agents who possess private information and act strategically may coordinate their decisions to buy or sell based on the system of prices that all agents know, without the need for sets of agents to reach binding prior agreements about how to act. The information that the agents have in common may be enough to let the market operate in a state of strategic equilibrium among agents (e.g., Grossman 1989). In the theoretical model I develop in this article, voters do not cooperate; each voter is able to make an equilibrium strategic choice among candidates, in accurate anticipation of the aggregate result of the choices all other voters intend to make, by using information that all voters know. Alesina and Rosenthal's (1989, 1995, 1996) theory of moderation features coordination—each voter's choice depends on an expected election result that everyone knows and that is an equilibrium—but Fiorina's (1988, 1992, 73–81) theory does not (Alesina and Rosenthal 1995, 66–71). In Fiorina's theory, voters do not act strategically at all. Voters act strategically in Alesina and Rosenthal's treatment, but no one has any private information about the election outcome. In the model developed here, different voters have beliefs about the upcoming election results that are very similar but not exactly the same in equilibrium. The similarity is a result of the common knowledge they have, while the differences that remain are due to private information each voter has about concerns that affect the voter's preferences among candidates.

I use National Election Study (NES) data from presidential election years 1976 through 1996 to estimate a stochastic choice model that closely matches the theoretical model. Based on the empirical results, I argue that in recent years American voters' choices for President and the House of Representatives have been coordinated, with each voter using a strategy that features policy moderation. In both the theoretical model and the empirical model, each voter forms preferences about the candidates based on a personally distinct idea about what the parties' policy positions are. The coordination takes the form of a rational expectations equilibrium among voters. The notion of voter equilibrium here is similar to the concept of a "fulfilled expectations" equilibrium

(Kreps 1977) developed for elections by McKelvey and Ordeshook (1984, 1985a, 1985b).[2]

The models of moderation-with-coordination developed by Alesina and Rosenthal (1989, 1995, 1996) have implications for the midterm loss phenomenon and for election-related fluctuations in economic phenomena such as growth. The latter have been tested using aggregate time-series data, with mostly positive results (Alesina, Londregan, and Rosenthal 1993; Alesina and Rosenthal 1995; Alesina, Roubini, and Cohen 1997).

Evidence that individuals coordinate has been lacking. Several authors report individual-level (Alvarez and Schousen 1993; Born 1994a) or aggregate-level (Burden and Kimbal 1998; Frymer 1994) tests aimed at Fiorina's (1988; 1992, 73–81) theory of moderation, with controversially (Born 1994b; Fiorina 1994) negative results. As previously noted, Fiorina's theory does not posit coordination among voters. The tests of his theory have not examined whether voters' expectations about the election outcomes affect their choices. I use the NES data to estimate a stochastic choice model that implements Fiorina's theory and test it against the coordinating model.

The heart of the formal model of coordination that I develop and test empirically is a fixed-point theorem that defines the common knowledge belief that all voters have about the upcoming election results. The election results are summarized in the values of two aggregate statistics: (i) the proportion of the two-party vote to be cast nationally for Republican candidates for the House and (ii) the probability that the Republican presidential candidate will defeat the Democrat. Each voter cares about those values because they affect the loss each voter expects to experience from policies the government will adopt after the election. Each voter chooses among candidates so as to minimize that loss. Therefore, each voter's belief about the aggregate values affects the voter's choices. In general, if there is a single pair of values that every voter believes the aggregate statistics to have and every voter chooses based on that belief, then the aggregate result of all their choices may be expected to differ from the values they all believed. In that case, beliefs and expected actions are inconsistent with one another. The fixed-point result demonstrates that a pair of aggregate values exists for which there is no such inconsistency: When every voter chooses based on belief in the fixed-point aggregate values, the aggregate result to be expected from all their choices is the very same pair of aggregate values. In that case, beliefs match expected actions. Common knowledge of the fixed-point values is not quite enough to establish an equilibrium, because the common knowledge does not include private information each voter has about the voter's actual choice. If each voter's belief equals the common knowledge values adjusted by amounts that correspond to the voter's choices, then the belief each voter has about the aggregate result is consistent with the choices the voter knows it will make. Despite knowing that every other voter is also acting on the basis of beliefs that differ slightly from the common knowledge values, no voter knows any other voter's private information, and so no voter can do any better than to use the common knowledge values as its aggregate expectation for what everyone else is going to do. Because all voters are situated similarly, it is common knowledge that every voter is forming beliefs in that way. Thus there is an equilibrium in which every voter has a slightly different expectation regarding the upcoming election results.

The fixed-point result that determines what voters' beliefs are in equilibrium in the theoretical model imposes a constraint on the statistical model to be used to estimate the parameters of the model with survey data. The empirical model is defined to correspond as closely as possible to the theoretical model. Both models use the same functional forms for each voter's loss function as well as the same specifications for the statistical distribution of the random disturbances that affect the candidate choices each voter makes. Hence, both models specify the same stochastic choice model for each voter. The special constraint is that the estimates of the parameters of the empirical model must be such that the aggregate election summary statistics estimated from the survey data satisfy the fixed-point condition. Stochastic choice models of electoral behavior have not previously imposed such a constraint, and its introduction here is an important innovation in the empirical model specification. The empirically determined fixed-point values are the estimates for the aggregate values that are common knowledge in equilibrium in the theoretical model. A distinction between the empirical and theoretical models is that, because of limitations on what we can observe with survey data, the empirical model treats the fixed-point values as the belief that every voter has regarding the upcoming election outcome. The empirical model does not refine each voter's belief to take the voter's private information into account.

The theoretical and empirical models use assumptions about what individual voters know that are unquestionably unrealistic and unbelievable. It may help to put both the theory and the empirical findings in what I think is the most productive perspective if I say something about this at the outset. In the Conclusion, I briefly outline how in reality a coordinating voting equilibrium must depend on a collection of institutions to aggregate and broadcast information that the current model treats as common knowledge among voters. My analysis is therefore implicitly conditional on the existence of an appropriate institutional context. In a way that I touch on briefly in the Conclusion,[3] I think voters possess a variety of beliefs about things the current model treats as common knowledge; the values that are common knowledge in the current model are more like the mean of the distribution of individuals' beliefs than they are values that every individual actually possesses. Even if the current model correctly characterizes the mean, it understates the variability across individuals.

A Model of Coordinating Voting with Moderation

The model treats the election as a game among all voters, assumed to be a large number. Voters act noncooperatively. The model focuses on the choices each voter makes between two candidates for President and two candidates for a House seat. The votes for President occur in the context of the Electoral College. There are many House districts, but each voter votes in only one of them. There are two parties, Democratic and Republican, and each race has one candidate running for each party. Both whether an incumbent is running and the incumbent's party vary over districts. Each voter's preferences regarding pairs of candidates depend on the expected election outcome and, therefore, on the choice strategy every other voter is using. Equilibrium occurs when, given everything each voter knows—including the voter's own intended choice and accurate expectations regarding other voters' strategies—no voter expects to gain by using a different strategy.

The preferences each voter has regarding the candidates are based in part on spatial comparisons between the voter's ideal point and the policies the voter believes will result from various election outcomes. Those policies are functions of the policy position the voter expects each party will act on after the election. The expected post-election policy position for each party is a combination of two positions that may differ: the position the voter associates with the party on the basis of previous turns in office and previous campaigns, and the position the voter thinks the party's current presidential candidate has adopted. All presidential candidates do not have the same ability to move their parties to the position the candidate supports. For each election, the expected party position is a weighted average of the candidate's position and the prior party position. The more influential the candidate, the greater is the weight of the candidate's position.

I use ϑ_{Di}, ϑ_{Ri}, ϑ_{PDi}, and ϑ_{PRi} to denote values in the interval $[0,1]$ that voter i, $i=1,\ldots,N$, has in mind at election time for the prior position of the Democratic party (ϑ_{Di}) and Republican party (ϑ_{Ri}) and for the position of the Democratic presidential candidate (ϑ_{PDi}) and Republican presidential candidate (ϑ_{PRi}). The Democratic positions need not be to the "left" of (i.e., numerically less than) the Republican positions. The policy positions voter i expects the Democratic party and the Republican party to act on after the election are, respectively,

$$\theta_{Di} = \rho_D \vartheta_{PDi} + (1 - \rho_D)\vartheta_{Di} \ , \qquad 0 \leq \rho_D \leq 1 \ , \tag{1a}$$

$$\theta_{Ri} = \rho_R \vartheta_{PRi} + (1 - \rho_R)\vartheta_{Ri} \ , \qquad 0 \leq \rho_R \leq 1 \ . \tag{1b}$$

If the Democratic party's position is expected to be very close to that of its presidential candidate, then the weight ρ_D is near one. If the Democratic party's position is expected to remain close to its previous value, then ρ_D is near zero. The parameter ρ_R has an analogous relationship to the position voter i expects the Republican party to act on after the election.

The policy expected to result from each possible election outcome depends on four factors: the parties' expected policy positions (just defined); the position expected to be supported in Congress; the President's strength in comparison to the House; and which party's candidate is elected President. To represent the expected position of Congress, I simplify by ignoring both the Senate and all internal structure in the House, such as seats, committees, bills, and rules; the expected position of the House is simply a weighted average of the expected positions of the two parties. Each party's weight is equal to the proportion of the vote that i expects to be cast nationally for the party's House candidates. Using $\bar{H}_i$ to denote the expected Republican proportion, the expected position of the House is $\bar{H}_i\theta_{Ri} + (1 - \bar{H}_i)\theta_{Di}$.

The postelection policy expected given any particular value of $\bar{H}_i$ is then a weighted average of the expected position of the House and the expected position of the President's party. The weight of the President represents the President's strength in comparison to the House. For each value of $\bar{H}_i$ there are two expectations for postelection policy, depending on the President's party:

$$\tilde{\theta}_{Di} = \alpha_D \theta_{Di} + (1 - \alpha_D)[\bar{H}_i\theta_{Ri} + (1 - \bar{H}_i)\theta_{Di}] \ , \qquad 0 \leq \alpha_D \leq 1 \ , \tag{2a}$$

$$\tilde{\theta}_{Ri} = \alpha_R \theta_{Ri} + (1 - \alpha_R)[\bar{H}_i\theta_{Ri} + (1 - \bar{H}_i)\theta_{Di}] \ , \qquad 0 \leq \alpha_R \leq 1 \ . \tag{2b}$$

Policy $\tilde{\theta}_{Di}$ is expected to occur if a Democrat is President, and policy $\tilde{\theta}_{Ri}$ is expected to occur if a Republican is President. The weights, α_D and α_R, represent the strength a President from each party is expected to have. The value $\alpha_D = 1$ means that a Democratic President is expected to dictate policy, so that the legislature would play no role, while $\alpha_D = 0$ means that the legislature is expected to determine policy, with the President being irrelevant. The interpretation of α_R is analogous for a Republican President. The functional forms of $\tilde{\theta}_{Di}$ and $\tilde{\theta}_{Ri}$ are essentially the same as the simplest policymaking formalism considered by Alesina and Rosenthal (1995, 47–8).

The preference each voter has for each possible election outcome is measured by the loss the voter expects given that outcome. The loss depends on the absolute discrepancy between the voter's ideal point, denoted $\theta_i \in [0, 1]$, and the policy expected given the election outcome. The absolute discrepancy is $|\theta_i - \tilde{\theta}_{Di}|$ with a Democratic President and $|\theta_i - \tilde{\theta}_{Ri}|$ with a Republican President. A voter's expected loss always increases as the absolute discrepancy between the voter's ideal point and the expected policy gets larger.

I use a functional form for the loss that is flexible in two ways. First, an exponent $q > 0$ allows the loss to be a concave ($0 < q < 1$), linear ($q = 1$), or convex ($q > 1$) function of the absolute discrepancy between θ_i and each expected policy. Second, a variable β_i, $0 < \beta_i < 1$, represents the possibility that voters do not treat the two discrepancies equally. If $\beta_i < \frac{1}{2}$ ($\beta_i = \frac{1}{2}$, $\beta_i > \frac{1}{2}$), voter i weights a discrepancy associated with a Republican President more than (the same as, less than) a discrepancy of the same magnitude associated with a Democratic President. The voter's expected losses with, respectively, a Democratic and a Republican President are

$$\lambda_{Di} = \beta_i |\theta_i - \tilde{\theta}_{Di}|^q + \xi_{Di} ,$$

$$\lambda_{Ri} = (1 - \beta_i) |\theta_i - \tilde{\theta}_{Ri}|^q + \xi_{Ri} ,$$

where the continuous random variables ξ_{Di} and ξ_{Ri} represent other things besides the policy-related discrepancies that affect the voter's expected loss.

At the time of the election, the voter does not know for sure which candidate will be elected

President. I assume that the voter computes an expectation for the expected loss, by using the probability that each candidate will win in a standard expected-value formula. $\bar{P}_i$ denotes voter i's expectation for the probability that the Republican wins. The voter's expected loss is

$$\lambda_i = (1 - \bar{P}_i)\beta_i|\theta_i - \tilde{\theta}_{Di}|^q + \bar{P}_i(1 - \beta_i)|\theta_i - \tilde{\theta}_{Ri}|^q + \xi_i \ , \tag{3}$$

where $\xi_i = (1 - \bar{P}_i)\xi_{Di} + \bar{P}_i\xi_{Ri}$.

The specifications of equations 2a, 2b and 3 have functional forms that are quite similar to those used by Alesina and Rosenthal (1995), but there is an important difference in what the current model says about the information that voters have. The specifications of equations 2a, 2b and 3 differ from Alesina and Rosenthal's (1995) theory in allowing the expected policies $\tilde{\theta}_{Di}$ and $\tilde{\theta}_{Ri}$ and the expected election outcomes $\bar{H}_i$ and $\bar{P}_i$ to vary over voters. In the theory of Alesina and Rosenthal (1995, 1996), the expected policies and expected election outcomes have values that all voters have identically in common. The variations in the values over voters in the current model mean that the current model endows voters with private information in a way that the theory of Alesina and Rosenthal (1995, 1996) essentially does not do. The current model therefore in an important way generalizes the approach taken by Alesina and Rosenthal (1995, 1996).

A voter chooses the pair of candidates—one candidate for President and one for the House—that minimizes the voter's expected loss, λ_i. I measure the effect on λ_i of choosing any one pair, as follows. Let $\bar{P}_{i,R}$ denote the expected probability that the Republican presidential candidate wins if the voter chooses the Republican, and let $\bar{P}_{i,D}$ denote the probability that the Republican wins if the voter chooses the Democrat. Because the effect of a single voter's choice on $\bar{P}_i$ is very small, the effect on λ_i of the voter's choosing the Republican presidential candidate rather than the Democrat is well approximated by $(\bar{P}_{i,R} - \bar{P}_{i,D})d\lambda_i/d\bar{P}_i$.[4] If $(\bar{P}_{i,R} - \bar{P}_{i,D})d\lambda_i/d\bar{P}_i$ is positive, then voting for the Republican increases the voter's expected loss. Likewise, let $\bar{H}_{i,R}$ ($\bar{H}_{i,D}$) be the proportion of the national two-party vote received by Republican House candidates if the voter chooses the Republican (Democrat) running in the voter's district. Then $(\bar{H}_{i,R} - \bar{H}_{i,D})d\lambda_i/d\bar{H}_i$ closely approximates the effect on λ_i of the voter's choosing a Republican House candidate. Sums and differences of such quantities measure the effect on λ_i of each of the four possible pairs of candidate choices. I use $\lambda_{i,ph}$ to denote such a measure of the effect of voter i's choosing the presidential candidate of party p and the House candidate of party h. For simplicity, I ignore the differences in voter population sizes across States. Instead, I assume that $\bar{P}_{i,R} - \bar{P}_{i,D}$ and $\bar{H}_{i,R} - \bar{H}_{i,D}$ are constants that are the same for all voters: $b_P = \bar{P}_{i,R} - \bar{P}_{i,D}$ and $b_H = \bar{H}_{i,R} - \bar{H}_{i,D}$. The effects are

$$\lambda_{i,RR} = b_P \frac{d\lambda_i}{d\bar{P}_i} + b_H \frac{d\lambda_i}{d\bar{H}_i}, \qquad \lambda_{i,RD} = b_P \frac{d\lambda_i}{d\bar{P}_i} - b_H \frac{d\lambda_i}{d\bar{H}_i},$$
$$\lambda_{i,DR} = -b_P \frac{d\lambda_i}{d\bar{P}_i} + b_H \frac{d\lambda_i}{d\bar{H}_i}, \qquad \lambda_{i,DD} = -b_P \frac{d\lambda_i}{d\bar{P}_i} - b_H \frac{d\lambda_i}{d\bar{H}_i}.$$

Choosing the candidate of party p for President and the candidate of party h for the House increases the voter's expected loss if $\lambda_{i,ph} > 0$ and decreases the voter's expected loss if $\lambda_{i,ph} < 0$.

To compute an explicit rule for the voter's choice, I make two more simplifying assumptions. I assume that β_i is not a function of either $\bar{P}_i$ or $\bar{H}_i$, so that $d\beta_i/d\bar{P}_i = d\beta_i/d\bar{H}_i = 0$. And I assume there are constants b_{HP} and b_{PH} such that $b_{HP} = d\bar{H}_i/d\bar{P}_i$ and $b_{PH} = d\bar{P}_i/d\bar{H}_i$. The effects of the possible choices on the voter's expected losses are then

$$\lambda_{i,RR} = (b_P + b_H b_{PH})w_{Pi} + (b_H + b_P b_{HP})w_{Hi} + \xi_{RRi}\ ; \tag{4a}$$

$$\lambda_{i,RD} = (b_P - b_H b_{PH})w_{Pi} + (b_P b_{HP} - b_H)w_{Hi} + \xi_{RDi}\ ; \tag{4b}$$

$$\lambda_{i,DR} = (b_H b_{PH} - b_P)w_{Pi} + (b_H - b_P b_{HP})w_{Hi} + \xi_{DRi}\ ; \tag{4c}$$

$$\lambda_{i,DD} = -(b_P + b_H b_{PH})w_{Pi} - (b_H + b_P b_{HP})w_{Hi} + \xi_{DDi}\ , \tag{4d}$$

with

$$w_{Pi} = (1-\beta_i)|\theta_i - \tilde{\theta}_{Ri}|^q - \beta_i|\theta_i - \tilde{\theta}_{Di}|^q \ ;$$

$$w_{Hi} = q(\theta_{Di} - \theta_{Ri})[(1-\alpha_R)\bar{P}_i(1-\beta_i)|\theta_i - \tilde{\theta}_{Ri}|^{q-1}\operatorname{sgn}(\theta_i - \tilde{\theta}_{Ri})$$

$$+ (1-\alpha_D)(1-\bar{P}_i)\beta_i|\theta_i - \tilde{\theta}_{Di}|^{q-1}\operatorname{sgn}(\theta_i - \tilde{\theta}_{Di})] \ ,$$

where $\operatorname{sgn}(x) = -1$ if $x < 0$, $\operatorname{sgn}(x) = 0$ if $x = 0$, and $\operatorname{sgn}(x) = 1$ if $x > 0$. The variables ξ_{RRi}, ξ_{RDi}, ξ_{DRi}, and ξ_{DDi} are defined in Appendix A. The voter's choice rule is: choose the candidate combination that corresponds to the smallest of the four values, $\lambda_{i,RR}$, $\lambda_{i,RD}$, $\lambda_{i,DR}$, and $\lambda_{i,DD}$. If we represent the two candidates the voter chooses by a random variable Y_i that takes values in the choice set $K = \{RR, RD, DR, DD\}$, then we may write the voter's choice rule as

$$Y_i = \operatorname*{argmin}_{ph \in K} \lambda_{i,ph} \ . \tag{5}$$

Equation 5 defines a strategy for each voter in a large-scale game in which all voters participate. The moves available to each voter are the four possible pairs of candidate choices. I assume that all voters move simultaneously, that each voter plays the game noncooperatively, and that each voter knows that every other voter is playing the game the same way. Both w_{Pi} and w_{Hi} depend directly on the parties' shares of the national House vote ($\bar{H}_i$), and w_{Hi} depends directly on the parties' chances of winning the presidency ($\bar{P}_i$). The voter's best choice therefore depends on what every other voter is going to do. The strategy defined by equation 5 is an equilibrium if it is the rule that minimizes each voter's expected loss when each voter assumes that everyone else is using the same rule.

For such an equilibrium to exist, each voter must assign values to $\bar{H}_i$ and $\bar{P}_i$ in a way that accurately corresponds to the choices every other voter is likely to make according to equation 5. To be able to do that, each voter must know something about the losses other voters expect to incur from the possible election outcomes. With sufficient information about those losses, each voter can anticipate what other voters will do in response to each electoral circumstance that may arise. Knowledge of that response pattern allows each i to determine the mutually consistent values of $\bar{H}_i$ and $\bar{P}_i$. The pair $(\bar{H}_i, \bar{P}_i)$ is mutually consistent if, given everything i knows, i chooses in such a way that—taking i's own choice into account—the proportion Republican that i expects among votes for the House is $\bar{H}_i$ and the probability with which i expects the Republican presidential candidate to win is $\bar{P}_i$. In equilibrium, all pairs $(\bar{H}_i, \bar{P}_i)$, $i = 1, \ldots, N$, are mutually consistent.

I assume it is common knowledge (Fudenberg and Tirole 1991, 541–6) that every voter i has an expected loss λ_i as defined by equation 3, and that every i is choosing candidates so as to minimize λ_i. I assume that the common knowledge includes the values of all parameters, including ρ_D, ρ_R, α_D, α_R, q, b_P, b_H, b_{PH}, b_{HP}, and any parameters in β_i, ξ_{RRi}, ξ_{RDi}, ξ_{DRi}, or ξ_{DDi}. It is then common knowledge that, for some values of the policy position and other variables, equation 5 is every voter's choice rule.

No voter knows the values of the ideal point, the party or presidential candidate policy positions, or the other variables on the basis of which any other voter is choosing via equation 5. Each voter does know the probability distribution of those values. Let Z_i be an ordered set that includes all the observable variables in $\lambda_{i,RR}$, $\lambda_{i,RD}$, $\lambda_{i,DR}$, and $\lambda_{i,DD}$. That includes θ_i, ϑ_{Di}, ϑ_{Ri}, ϑ_{PDi}, and ϑ_{PRi}; any variables that make up β_i; and the component z_{phi} of each ξ_{phi}, $ph \in K$, that may be observed using some generally known technology (e.g., an opinion survey). There are $M \ll N$ mutually exclusive and exhaustive groups of voters, denoted V_k, $k = 1, \ldots, M$. For every voter $i \in V_k$, the election-time value of Z_i is generated, independently across individuals, by a process that takes values in a set $\tilde{Z}$ and has unimodal probability measure f_k, $\int_{\tilde{Z}} df_k(Z_i) = 1$. Let ϵ_{phi} denote the component of ξ_{phi} not covered by Z_i, $ph \in K$. These variables—collectively, the disturbance $\epsilon_i = (\epsilon_{RRi}, \epsilon_{RDi}, \epsilon_{DRi}, \epsilon_{DDi})$—have a joint probability distribution that is the same for every voter. The joint cumulative distribution function is

$$\Pr(\epsilon_{RRi} < x_{RR}, \epsilon_{RDi} < x_{RD}, \epsilon_{DRi} < x_{DR}, \epsilon_{DDi} < x_{DD}) = F_{PH}(x_{RR}, x_{RD}, x_{DR}, x_{DD}) ,$$

where F_{PH} is a generalized extreme value (GEV) distribution, independent of each f_k. I assume that M, $\tilde{Z}$, f_k, $k = 1, \ldots, M$, the number of voters in each group (M_k),[5] and the fact that ϵ_i is identically and independently distributed as F_{PH} are all common knowledge.

A GEV distribution arises in a natural way as a model for the disturbance vector ϵ_i.[6] Suppose that, as time goes by, voter i randomly encounters a large number of pieces of information about the alternative candidate pairs but remembers only the best piece of information about each pair. That is, the voter remembers the information about each pair of candidates that makes the loss from that pair seem the smallest. For a wide range of random processes that may generate the information each voter encounters, the distribution of the information the voter remembers about each pair is a GEV distribution (Galambos 1987, 286–314; Resnick 1987, 263–90).

I assume that the pieces of information each voter randomly encounters have a special form of dependence. Information about the two-Republican candidate pair (RR) arrives independently of information about the two-Democrat pair (DD), and the information about those two pairs arrives independently of information about either of the two mixed pairs (RD or DR); but information about one mixed pair does not arrive independently of information about the other mixed pair. In this case, ϵ_{RRi} is independent of ϵ_{DDi}, and ϵ_{RRi} and ϵ_{DDi} are both independent of ϵ_{RDi} and ϵ_{DRi}, but ϵ_{RDi} and ϵ_{DRi} vary together.[7] The most natural way to think about this pattern is that a voter accumulates information about each party and also about the possibility of casting a split-ticket vote, but in so doing the voter tends not to distinguish sharply one split-ticket alternative from the other.

To specify the GEV distribution that such a pattern of dependence produces, start by defining

$$x_{RRi} = (b_P + b_H b_{PH}) w_{Pi} + (b_H + b_P b_{HP}) w_{Hi} + z_{RRi} \; ; \tag{6a}$$

$$x_{RDi} = (b_P - b_H b_{PH}) w_{Pi} + (b_P b_{HP} - b_H) w_{Hi} + z_{RDi} \; ; \tag{6b}$$

$$x_{DRi} = (b_H b_{PH} - b_P) w_{Pi} + (b_H - b_P b_{HP}) w_{Hi} + z_{DRi} \; ; \tag{6c}$$

$$x_{DDi} = -(b_P + b_H b_{PH}) w_{Pi} - (b_H + b_P b_{HP}) w_{Hi} + z_{DDi} \; . \tag{6d}$$

Using $v_{phi} = \exp\{-x_{phi}\}$, $ph \in K$, define

$$G_i = v_{RRi} + \left(v_{RDi}^{1/1-\tau} + v_{DRi}^{1/1-\tau} \right)^{1-\tau} + v_{DDi} \; , \qquad 0 \le \tau < 1 \; .$$

Parameter τ measures the dependence between ϵ_{RDi} and ϵ_{DRi}. If ϵ_{RDi} and ϵ_{DRi} are independent, then $\tau = 0$. The GEV distribution function is $F_{PH}(-x_{RRi}, -x_{RDi}, -x_{DRi}, -x_{DDi}) = \exp\{-G_i\}$.

Given Z_i but ϵ_i known only to have the distribution F_{PH}, the choice rule of equation 5 implies that the probability that voter i chooses each candidate pair is

$$\Pr(Y_i = ph \mid Z_i) = \frac{v_{phi}}{G_i} \frac{\partial G_i}{\partial v_{phi}} \; , \quad ph \in K \tag{7}$$

(Börsch-Supan 1990; McFadden 1978; Resnick and Roy 1990). Define $\mu_{phi} = \Pr(Y_i = ph \mid Z_i)$, $ph \in K$. The probability that i votes Republican for the House is $\eta_i = \mu_{RRi} + \mu_{DRi}$, and for President, $\pi_i = \mu_{RRi} + \mu_{RDi}$.

To characterize the conditions under which mutually consistent pairs $(\bar{H}_i, \bar{P}_i)$ exist when each voter i knows all its own attributes, that is, knows Z_i and ϵ_i, I first consider what happens when each voter knows only to which group the voter belongs. Because M, M_k, $\tilde{Z}$, f_k, and F_{PH} are common knowledge, no voter then has any information that would produce an expectation about the election outcome different from what the common knowledge would imply. No voter has any relevant private information. Therefore, if a set of mutually consistent pairs exists, the expectations $(\bar{H}_i, \bar{P}_i)$ will be the same for all voters (Aumann 1976; Nielsen et al. 1990). Let $(\bar{H}, \bar{P})$ denote the common value that all the pairs have in this case.

Knowing only the group V_k to which voter i belongs, and therefore knowing only the range $\tilde{Z}$ and probability measure f_k of the variables in Z_i, every voter determines the same probability for i to choose each candidate pair, by using f_k to integrate over the unknown data:[8]

$$\Pr(Y_i = ph \mid i \in V_k) = \int_{\tilde{Z}} \Pr(Y_i = ph \mid Z_i) df_k(Z_i) \ , \quad ph \in K \ . \tag{8}$$

Define $\bar{\mu}_{phk} = \Pr(Y_i = ph \mid i \in V_k)$, $ph \in K$, and $\bar{\eta}_k = \bar{\mu}_{RRk} + \bar{\mu}_{DRk}$. Using f_{k_i} to denote the measure for the group to which i belongs, the expected House outcome is

$$\begin{aligned}
\bar{H} &= \underbrace{\int_{\tilde{Z}} \cdots \int_{\tilde{Z}}}_{N \text{ times}} \left(N^{-1} \sum_{i=1}^{N} \eta_i \right) \prod_{i=1}^{N} df_{k_i}(Z_i), && \text{by common knowledge,} \\
&= N^{-1} \sum_{k=1}^{M} \sum_{i \in V_k} \int_{\tilde{Z}} \eta_i df_k(Z_i), && \text{by independence,} \\
&= \sum_{k=1}^{M} \frac{M_k}{N} \bar{\eta}_k \ . &&
\end{aligned} \tag{9}$$

The expected probability that the Republican candidate wins the presidency is more complicated, because $\bar{P}$ is the probability that the Republican wins a majority of the Electoral College. Let C_j denote the number of Electoral College votes for State j, $j = 1, \ldots, S$, with $\sum_{j=1}^{S} C_j = C = 538$. Let $C^{[r]}$ denote the set of all subsets s of States such that $\sum_{j \in s} C_j = r$. The expected presidential outcome[9] is

$$\bar{P} = \frac{\sum_{r=C/2}^{C} \sum_{s \in C^{[r]}} \left(\prod_{j \in s} \bar{y}_{Pj} \right) \left[\prod_{j \notin s} (1 - \bar{y}_{Pj}) \right]}{\sum_{r=0}^{C} \sum_{s \in C^{[r]}} \left(\prod_{j \in s} \bar{y}_{Pj} \right) \left[\prod_{j \notin s} (1 - \bar{y}_{Pj}) \right]} \ , \tag{10}$$

where $\bar{y}_{Pj}$, defined in Appendix A as a function of $\bar{\pi}_k = \bar{\mu}_{RRk} + \bar{\mu}_{RDk}$, is the probability that the Republican candidate wins a majority of the vote in State j.

I assume that the functional forms of equations 9 and 10 are common knowledge. It then follows that the set of possible values of $\bar{H}$ and $\bar{P}$ is common knowledge. If only the group membership of each voter is known, then the mutual consistency condition requires that the $\bar{H}$ and $\bar{P}$ values must reproduce themselves when they are used to compute the functions $\bar{\mu}_{phk}$, $ph \in K$, and hence equations 9 and 10. In technical terms, the pair of computed values must be a fixed point of the mapping $[0,1] \times [0,1] \to [0,1] \times [0,1]$ that equations 9 and 10 specify. In Appendix A (Theorem 2) I show that such a pair $(\bar{H}, \bar{P})$ always exists (except on a set of measure zero). Necessary or sufficient conditions for $\bar{H}$ and $\bar{P}$ to be uniquely determined by the parameters of the model and the measures f_k are not clear.

The result when each voter i knows Z_i and ϵ_i may be expressed as a collection of small deviations from a fixed point $(\bar{H}, \bar{P})$. Let $\bar{\eta}_{k_i}$ and $\bar{\pi}_{k_i}$ denote the group-specific probabilities for the group V_k to which i belongs. Let $\tilde{\mu}_{phi}$ indicate the value of Y_i of equation 5 when voter i knows Z_i and ϵ_i but for other voters has only the common knowledge: $\tilde{\mu}_{phi} = 1$ if $Y_i = ph$, $\tilde{\mu}_{phi} = 0$ if $Y_i \neq ph$, $ph \in K$. Define $\tilde{\eta}_i = \tilde{\mu}_{RRi} + \tilde{\mu}_{DRi}$ and $\tilde{\pi}_i = \tilde{\mu}_{RRi} + \tilde{\mu}_{RDi}$. The values of $\tilde{\mu}_{phi}$, $ph \in K$, depend on $\bar{H}_i = \bar{H}_{i\tilde{\eta}_i}$ and $\bar{P}_i = \bar{P}_{i\tilde{\pi}_i}$, where $\bar{H}_{i\tilde{\eta}_i}$ and $\bar{P}_{i\tilde{\pi}_i}$ are obtained by replacing one instance of $\bar{\eta}_{k_i}$ in equation 9 with $\tilde{\eta}_i$ and one instance of $\bar{\pi}_{k_i}$ in equation 10, via equation (see below) 18, with $\tilde{\pi}_i$:

$$\bar{H}_{i\tilde{\eta}_i} = \bar{H} + (\tilde{\eta}_i - \bar{\eta}_{k_i})/N \ , \tag{11a}$$

$$\bar{P}_{i\tilde{\pi}_i} = \bar{P} + g_i(\tilde{\pi}_i - \bar{\pi}_{k_i}) \ , \tag{11b}$$

where $g_i(\cdot)$ is the function that makes the adjustment from $\bar{P}$, $g_i(0) = 0$. Notice that, because i knows Z_i and ϵ_i, the vote choice is not random as far as i is concerned: Given $\bar{H}_i$ and $\bar{P}_i$, one of the four choices in K will certainly be best according to equation 5.[10] An equilibrium set of choices Y_i and expectations $(\bar{H}_i, \bar{P}_i)$, $i = 1, \ldots, N$, is therefore given by the following.

Theorem 1 *There is a coordinating voter equilibrium if, with all voters using the same fixed point $(\bar{H}, \bar{P})$ computed from common knowledge, each voter i has $(\bar{H}_i, \bar{P}_i) = (\bar{H}_{i\tilde{\eta}_i}, \bar{P}_{i\tilde{\pi}_i})$ and $Y_i = ph$ for whichever of the four possible pairs of values $(\bar{H}_{i\tilde{\eta}_i}, \bar{P}_{i\tilde{\pi}_i})$ produces the smallest value of $\lambda_{i,ph}$, $ph \in K$, where $(\bar{H}_{i1}, \bar{P}_{i1})$ sets $\lambda_{i,RR}$, $(\bar{H}_{i0}, \bar{P}_{i1})$ sets $\lambda_{i,RD}$, $(\bar{H}_{i1}, \bar{P}_{i0})$ sets $\lambda_{i,DR}$, and $(\bar{H}_{i0}, \bar{P}_{i0})$ sets $\lambda_{i,DD}$.*

Plainly the expectations $(\bar{H}_i, \bar{P}_i) = (\bar{H}_{i\tilde{\eta}_i}, \bar{P}_{i\tilde{\pi}_i})$ that minimize $\lambda_{i,ph}$, $ph \in K$, match the choices voter i makes according to equation 5. And while each voter i knows that every other voter i' also has expectations $\bar{H}_{i'}$ and $\bar{P}_{i'}$ as defined by equations 11a and 11b, $\bar{H}$ and $\bar{P}$ are the best estimates of the election outcome—*not* taking into account the choices i will make—available to i in the absence of knowledge of the particular values of $Z_{i'}$ and $\epsilon_{i'}$: based on the common knowledge, the expectation of $N^{-1}\sum_{i=1}^{N}(\tilde{\eta}_i - \bar{\eta}_{k_i})/N$ is zero, and $N^{-1}\sum_{i=1}^{N} g_i(\tilde{\pi}_i - \bar{\pi}_{k_i})$ converges in probability to zero as the size of the population of voters increases;[11] there is no information beyond the common knowledge with which to produce estimates that have smaller variance. Since it is common knowledge that all voters are similarly situated (i.e., exchangeable), it is common knowledge that $(\bar{H}, \bar{P})$ is every voter's best estimate of the election outcome, not counting the voter's own choices. *Q.E.D.*

A Model for Estimation with Survey Data

For the coordinating voting model to be empirically estimable, the fixed point $(\bar{H}, \bar{P})$ that is the basis for the voter equilibrium must satisfy a condition of local stability. To understand what is needed, imagine that every voter's belief about $(\bar{H}, \bar{P})$ undergoes a small perturbation from the fixed-point values, say to $(\bar{H}^{(1)}, \bar{P}^{(1)})$. If each voter uses the perturbed values in equation 8 and hence evaluates equations 9 and 10, the result will in general be values $(\bar{H}^{(2)}, \bar{P}^{(2)}) \neq (\bar{H}^{(1)}, \bar{P}^{(1)})$. The fixed point $(\bar{H}, \bar{P})$ is locally stable if a sequence $(\bar{H}^{(1)}, \bar{P}^{(1)}), (\bar{H}^{(2)}, \bar{P}^{(2)}), \ldots$ thus produced by repeated cycling through equations 8, 9 and 10 converges to $(\bar{H}, \bar{P})$. Perturbations being inevitable, it is easy to see that any equilibrium that occurs in reality must have such a property: even the smallest change would set voters on a path leading rapidly away from any unstable pair $(\bar{H}, \bar{P})$.

Assuming that the equilibrium does exist, that the likelihood function that defines the empirical model specification is correct, and that the variables that affect voters' choices in that specification are measured sufficiently well, the stability condition allows an iterative estimation algorithm such as the one described in Appendix A to converge to the parameter estimates that characterize the choices voters make in equilibrium. The need for the computed pair $(\hat{\bar{H}}, \hat{\bar{P}})$ to be a fixed point imposes an additional condition for the estimation algorithm to be judged to have converged. In addition to the usual conditions for having found parameter estimates that maximize the likelihood function, the estimated values $(\hat{\bar{H}}, \hat{\bar{P}})$ must be constant over successive iterations of the estimation algorithm. The requirement from the theoretical model that $(\bar{H}, \bar{P})$ be a fixed point imposes a fixed-point constraint on the maximum likelihood solution.

With NES survey data we can observe the candidate choices Y_i reported by each sampled voter i and a number of variables that affect vote choices—a set of variables Z_i. We can use the observed Z_i values and a set of parameter estimates (not necessarily the maximum likelihood estimates [MLEs]) in equation 7 to compute estimated choice probabilities $\hat{\mu}_{phi}$, $ph \in K$, for each sampled voter, and hence $\hat{\eta}_i = \hat{\mu}_{RRi} + \hat{\mu}_{DRi}$ and $\hat{\pi}_i = \hat{\mu}_{RRi} + \hat{\mu}_{RDi}$. I use such estimated probabilities to compute $(\hat{\bar{H}}, \hat{\bar{P}})$ for each set of parameter estimates without having to specify any particular set of groups V_k. The method is to use the sampling weight associated with each voter in each survey to estimate the totals in the formulas for $\bar{H}$ and $\bar{P}$. The sampling weight for voter i is proportional to $1/\omega_i$, where ω_i is the probability, determined by the sampling design, that i is included in the sample. Ratio estimates based on such weighted totals implicitly average over groups, without requiring explicit definition of the groups. To compute $\hat{\bar{H}}$ for a sample of n voters, I use

$$\hat{\bar{H}} = \frac{\sum_{i=1}^{n} \hat{\eta}_i/\omega_i}{\sum_{i=1}^{n} 1/\omega_i} \,, \tag{12}$$

which is a design-consistent estimator that does not require any group information.[12]

I use four different methods to compute $\hat{\bar{P}}$, three of which take the Electoral College into account. Those three methods are based on a ratio estimator of the proportion y_{Pj} of the vote the Republican presidential candidate is expected to win in each State j for which we observe data (NES samples do not have observations from every State). To define the ratio estimator, let $\bar{v}_j$

denote the set of voters in the survey sample from State j. If $\bar{v}_j$ is not empty, then the ratio estimator is

$$\hat{y}_{Pj} = \frac{\sum_{i\in\bar{v}_j} \hat{\pi}_i/\omega_i}{\sum_{i\in\bar{v}_j} 1/\omega_i} \ . \tag{13}$$

For the first three methods to compute $\hat{\bar{P}}$, I use functions of the $\hat{y}_{Pj}$ values to estimate $\bar{y}_{Pj}$, the expected probability that the Republican wins a majority of the vote in j and therefore wins all the electoral votes from j. For States from which there are no sample observations ($\bar{v}_j = \emptyset$), I assign all the electoral votes to the candidate who actually won the State: $\hat{\bar{y}}_{Pj} = 1$ if the Republican won, and $\hat{\bar{y}}_{Pj} = 0$ if the Democrat won. Let ζ_R denote the proportion of electoral votes the Republican will win. Given probabilities $\hat{\bar{y}}_{Pj}$, its expectation is $\bar{\zeta}_R = \sum_{j=1}^{51} \hat{\bar{y}}_{Pj} C_j/538$.

Each of the first three methods uses a different approach to estimate $\Pr(\zeta_R > .5)$. Method one treats $\hat{y}_{Pj}$ directly as if it were a probability rather than a proportion, setting $\hat{\bar{y}}_{Pj} = \hat{y}_{Pj}$, then uses $\bar{\zeta}_R$ directly as the estimate for $\bar{P}$: With $\hat{\bar{y}}_{Pj} = \hat{y}_{Pj}$ for States with sample data, $\hat{\bar{P}} = \bar{\zeta}_R$. Method two also uses $\hat{\bar{y}}_{Pj} = \hat{y}_{Pj}$ but defines $\hat{\bar{P}}$ as the upper tail of a beta distribution defined to match the variance of ζ_R given $\hat{\bar{y}}_{Pj}$; details are in Appendix A. Method three uses the sampling error of $\hat{y}_{Pj}$ to estimate $\hat{\bar{y}}_{Pj} = \Pr(y_{Pj} > .5)$ via a normal approximation. Treating the parameter estimates as fixed numbers, the standard error of $\hat{y}_{Pj}$ is approximately $\sigma_{\hat{y}_{Pj}} = [\hat{y}_{Pj}(1-\hat{y}_{Pj})/(\sum_{i\in\bar{v}_j} 1/\omega_i)]^{1/2}$. Method three sets $\hat{\bar{y}}_{Pj} = 1-\Phi((.5-\hat{y}_{Pj})/\sigma_{\hat{y}_{Pj}})$, where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, then uses $\hat{\bar{y}}_{Pj}$ in a normal approximation for $\Pr(\zeta_R > .5)$; details are in Appendix A.

The fourth method ignores the Electoral College and estimates $\bar{P}$ as the simple proportion of votes expected to be received by the Republican: $\hat{\bar{P}} = (\sum_{i=1}^{n} \hat{\pi}_i/\omega_i)/(\sum_{i=1}^{n} 1/\omega_i)$. This is tantamount to using the Republican share of the (two-party) popular vote.

For two reasons, $(\hat{\bar{H}}, \hat{\bar{P}})$ is the best we can do to estimate $(\bar{H}_i, \bar{P}_i)$ with survey data. Using $\hat{\eta}_i$ and $\hat{\pi}_i$ to estimate $\bar{\eta}_{k_i}$ and $\bar{\pi}_{k_i}$ in equations 11a and 11b would underestimate the magnitudes of

$(\tilde{\eta}_i - \bar{\eta}_{k_i})/N$ and $g_i(\tilde{\pi}_i - \bar{\pi}_{k_i})$. And with $N > 10^8$, survey samples are too small to detect the effects of the deviations from $(\bar{H}, \bar{P})$. For every sampled voter I therefore set $(\hat{\bar{H}}_i, \hat{\bar{P}}_i) = (\hat{\bar{H}}, \hat{\bar{P}})$.

Given $T \geq 1$ samples each containing n_t observations of the choices Y_i and variables Z_i, with each observation being chosen with probability ω_i from a large population ($\omega_i > 0$, $i = 1, \ldots, N_t$) for each election year $t = 1, \ldots, T$, the parameters of the model may be estimated by maximum likelihood. Let $y_{phi} = 1$ if $Y_i = ph$ and $y_{phi} = 0$ if $Y_i \neq ph$, $ph \in K$. The log-likelihood is

$$L = \sum_{t=1}^{T} \sum_{i=1}^{n_t} \sum_{ph \in K} y_{phi} \log \mu_{phi} \ . \tag{14}$$

Iterations to estimate the parameter values by maximizing the log-likelihood recompute $(\hat{\bar{H}}, \hat{\bar{P}})$ separately for each year at each iteration, as described in Appendix A.

Table 1 displays the principal points of correspondence between the theoretical and empirical models. Both assume the same choice set, K, for each voter. The choices that each voter makes in the theoretical model are assumed to be the choices observed in the survey data. The models use the same definition, equations 6a–6d, for the observable effects that the choices voter i makes have on voter i's loss. Detailed specifications using NES data for equations 6a–6d's variables z_{phi}, $ph \in K$, appear in equations 17a–17d below. The definition of each voter's probability of making each choice, equation 7, is the same in both models. The functional forms for the expected election outcomes differ, due to the need to estimate $\bar{H}$ and $\bar{P}$ from survey samples, but in both models the aggregate statistics satisfy the fixed-point condition that defines the statistics' values as common knowledge in equilibrium in the theoretical model.

Tests of Coordination

I use two kinds of tests of whether voters coordinate. I check whether the estimated values of particular model parameters satisfy certain conditions necessary for coordination to exist, and I compare the performance of the coordinating voting model to a model that does not have coordination among voters but is otherwise similar.

If the President solely determines policy ($\alpha_D = \alpha_R = 1$), then there is no coordination among voters because $\tilde{\theta}_{Di} = \theta_{Di}$, $\tilde{\theta}_{Ri} = \theta_{Ri}$ and $w_{Hi} = 0$, so that voters' strategies do not depend on $\bar{P}_i$ or $\bar{H}_i$. A necessary condition for coordination is therefore that at least one of $\alpha_D < 1$ or $\alpha_R < 1$ is true. The test against the specification with $\alpha_D = \alpha_R = 1$ is important also because the constrained model is the familiar kind of model that includes a unidimensional spatial comparison along with nonspatial characteristics: In equations 4a–4d, $w_{Pi} = (1 - \beta_i)|\theta_i - \theta_{Ri}|^q - \beta_i|\theta_i - \theta_{Di}|^q$ and $w_{Hi} = 0$. I use confidence interval estimates and likelihood-ratio (LR) tests to check whether the estimates for α_D and α_R differ significantly from the value that annihilates the possibility of coordination.

Other conditions necessary for the choice models to describe coordination are that $q > 0$ and that both b_P and b_H are positive. If $q = 0$, then we have the degenerate values $w_{Hi} = 0$ and $w_{Pi} = 1 - 2\beta_i$. If $b_P = b_H = 0$, then whatever the discrepancies $|\theta_i - \tilde{\theta}_{Di}|^q$ and $|\theta_i - \tilde{\theta}_{Ri}|^q$ may be, they have no effect on voters' choices. The estimated parameters b_P and b_H do not equal the theoretical differences $\bar{P}_{i,R} - \bar{P}_{i,D}$ and $\bar{H}_{i,R} - \bar{H}_{i,D}$, but they are proportional to those values.[13] So for compatibility with the derivation of $\lambda_{i,ph}$, $ph \in K$, we must have $b_P > 0$ and $b_H > 0$.

The alternative model to which I compare the coordinating voting model implements Fiorina's (1988; 1992, 73–81) noncoordinating (indeed, nonstrategic) theory of ticket splitting. According to that theory, each voter chooses the mix of party control of the presidency and the legislature—either unified or divided government—that would produce a policy outcome nearest the voter's ideal point, but those choices are not affected by the anticipated election results.

To focus the test as powerfully as possible on the existence or nonexistence of coordination, I formulate the noncoordinating vote choice model to resemble the coordinating model as closely as possible. I replace the terms involving w_{Pi} and w_{Hi} in the coordinating model with terms that are suitable for the noncoordinating theory but otherwise leave the model unchanged. In particular, for the noncoordinating model, policy outcomes under unified government are θ_{Di} or θ_{Ri} of equations 1a and 1b, while policy outcomes under divided government are

$$\tilde{\theta}_{DRi} = \alpha_D \theta_{Di} + (1 - \alpha_D)\theta_{Ri} \ , \qquad 0 \leq \alpha_D \leq 1 \ ,$$

$$\tilde{\theta}_{RDi} = \alpha_R \theta_{Ri} + (1 - \alpha_R)\theta_{Di} \ , \qquad 0 \leq \alpha_R \leq 1 \ .$$

$\tilde{\theta}_{DRi}$ is the policy that voter i believes will occur with a Democratic President and Republican-controlled House, and $\tilde{\theta}_{RDi}$ is the policy with a Republican President and the House controlled by the Democrats. For the observed attributes of each choice define, instead of equations 6a–6d,

$$x_{RRi} = b_{NC}\beta_{RRi}|\theta_i - \theta_{Ri}|^q + z_{RRi} \ ;$$

$$x_{RDi} = b_{NC}\beta_{RDi}|\theta_i - \tilde{\theta}_{RDi}|^q + z_{RDi} \ ;$$

$$x_{DRi} = b_{NC}\beta_{DRi}|\theta_i - \tilde{\theta}_{DRi}|^q + z_{DRi} \ ;$$

$$x_{DDi} = b_{NC}\beta_{DDi}|\theta_i - \theta_{Di}|^q + z_{DDi} \ ,$$

where $b_{NC} > 0$, and β_{RRi}, β_{RDi}, β_{DRi} and β_{DDi} are defined in Appendix A.[14] The key feature of the model is that x_{RRi}, x_{RDi}, x_{DRi}, and x_{DDi} are functions of neither the expected Republican House vote proportion ($\bar{H}_i$) nor the expected probability that the Republican presidential candidate wins ($\bar{P}_i$). Nothing in the noncoordinating model makes the choice of voter i depend on the choice or likely choice of any other voter; each voter responds solely to what each believes about the parties and the candidates.

The noncoordinating model has one more free parameter than the coordinating model does, but it does not nest the coordinating model. I use a likelihood-based method for non-nested hypotheses (Dastoor 1985) to determine which model is superior. Let L_{C} and L_{NC} denote, respectively, the log-likelihood functions for the coordinating and noncoordinating models, both having the form of equation 14, and let $\hat{L}_{\mathrm{C}}$ and $\hat{L}_{\mathrm{NC}}$ denote, respectively, the values of L_{C} and L_{NC} computed using the

MLEs for each model, estimated separately. If $\hat{L}_{\mathrm{C}} > \hat{L}_{\mathrm{NC}}$, a test of whether the noncoordinating model may be rejected in favor of the coordinating model may be based on the MLE of ψ in the log-likelihood function $L_{\mathrm{TEST}} = (1-\psi)L_{\mathrm{C}} + \psi\hat{L}_{\mathrm{NC}}$, $0 \leq \psi \leq 1$. If $\hat{\psi}$ is significantly different from one, the noncoordinating model is rejected. If $\hat{L}_{\mathrm{C}} < \hat{L}_{\mathrm{NC}}$, a test of whether the coordinating model may be rejected in favor of the noncoordinating model may be performed by reversing the roles of L_{C} and L_{NC} in L_{TEST}. Inability to reject either model may indicate that some voters are coordinating, with behavior best described by L_{C}, while others are moderating without coordinating, their behavior being best described by L_{NC}.[15]

Detailed Choice Model Specifications

To estimate the models I use data from the NES surveys of 1976, 1980, 1984, 1988, 1992, and 1996 (Miller and Miller 1977; Miller and the National Election Studies 1982, 1986, 1989; Miller et al. 1993; Rosenstone et al. 1997).[16] I pool the data over all years. Some parameters are constant over all years while others vary over years.

I use the observed responses to several seven-point (and one four-point) NES survey scales to measure θ_i, ϑ_{Di}, ϑ_{PDi}, ϑ_{Ri}, and ϑ_{PRi}. I use the variables' empirical cumulative distributions to code the responses in the $[0,1]$ interval. The idea is to make the responses comparable across the substantively different scales by using relative units of measurement. In formulating their theory in terms of a uniform distribution of voters' ideal points, Alesina and Rosenthal (1989; 1995, 22, 86) use relative units: The value of each ideal point and policy position corresponds to the cumulative proportion of support for that position among voters. McKelvey and Ordeshook (1985a) similarly use relative measurement of positions.

The codes I use for the NES scales are to be interpreted as measuring the proportion of all survey respondents who support a position as liberal as or more liberal than the indicated position. Each scale either refers to liberal-conservative ideological labels or pertains to a policy issue. Each of the values θ_i, ϑ_{Di}, ϑ_{Ri}, ϑ_{PDi}, and ϑ_{PRi} averages the values for the named referent—self, Democratic or Republican party, Democratic or Republican presidential candidate—over only the scales for which voter i placed all five referents on the scale. The values for different voters may therefore be computed using different subsets of the substantive scales in each survey. There is no assumption that every voter is using the same substantive policy dimension. Appendix A lists the scales used from each of the surveys and describes in detail the method used to compute codes for each scale.

I specify β_i to be a function of each voter's retrospective evaluation of the national economy. To measure economic evaluations I use a variable, EC_i, that is based on responses to a question asking whether the national economy has gotten worse or better over the past year. EC_i ranges from -1 to 1; $\mathrm{EC}_i = -1$ corresponds to a judgment that the economy is "much worse," and $\mathrm{EC}_i = 1$ corresponds to a judgment that the economy is "much better" (see Appendix B, 1). The definition of β_i is

$$\beta_i = (1 + \exp\{-b_{E0} - b_{E1}\mathrm{EC}_i\})^{-1} , \tag{16}$$

where b_{E0} and b_{E1} are constants. One reason economic evaluations may affect β_i is that voters may believe a worsening economy calls for increased government intervention to protect or create jobs, or in other ways to help the vulnerable. A voter with such beliefs who thinks economic conditions are worsening may give more of the benefit of the doubt to the party with the stronger reputation for choosing such policies. During 1976–96, the Democratic party had a stronger reputation for choosing interventionist policies than did the Republicans (e.g., Palmer and Sawhill 1984). A voter with the indicated beliefs who thinks economic conditions are worsening may therefore weigh the policy discrepancy associated with a Republican President more heavily than the discrepancy with a Democrat ($\beta_i < .5$). When economic conditions are improving, presumably it is the Democratic party that receives closer scrutiny ($\beta_i > .5$). Another possibility is that voters believe declining economic conditions call for *reduced* government intervention, or at least for a reduced rate of expansion, because that is when the country can least afford it. In this case, we should have $\beta_i < .5$ ($\beta_i > .5$) for voters who think economic conditions are improving (getting worse).

The other observed attributes (z_{phi}, $ph \in K$) that pertain to the vote choices include retrospective economic evaluations, individual partisanship and House incumbent advantage.

Many have examined the effects that retrospective economic evaluations may have on vote choices (e.g., Fiorina 1981; Markus 1988). According to Alesina and Rosenthal (1995), such evaluations may be relevant to a voter's choice as an expression of two considerations: the voter's taste for macroeconomic outcomes (pp. 167–71); and the voter's judgment of the competence of the incumbent administration (pp. 191–5).[17] Alesina and Rosenthal treat those considerations in the context of particular models that they specify to describe the relationship between output and inflation in the actual economy. The current analysis does not rely on any such model. Rather, I assume that each voter reports its evaluation when responding to the question used to measure EC_i. I multiply EC_i by a variable (PP_i) that changes sign depending on the incumbent President's party: $\mathrm{PP}_i = 1$ if Republican; $\mathrm{PP}_i = -1$ if Democrat.

Party identification has long been known to affect vote choices, particularly ticket-splitting (e.g., Campbell and Miller 1957). Party identification is in part an index of tastes for policy outcomes and of judgments accumulated over time about the competence of the parties' successive administrations (Fiorina 1981; Franklin 1984; Franklin and Jackson 1983; Jackson 1975; Markus and Converse 1979; Page and Jones 1979). People who support different parties also tend to have different policy preferences and perceptions (Brady and Sniderman 1985). I measure party identification with six dummy variables that correspond to the levels of the NES seven-point scale measure of partisanship, using "Strong Democrat" as the reference category: $\mathrm{PID}_{\mathrm{D}i}$, $\mathrm{PID}_{\mathrm{ID}i}$, $\mathrm{PID}_{\mathrm{I}i}$, $\mathrm{PID}_{\mathrm{IR}i}$, $\mathrm{PID}_{\mathrm{R}i}$, and $\mathrm{PID}_{\mathrm{SR}i}$ (see Appendix B, 2).

To take incumbent advantage into account, I use a pair of dummy variables that indicate whether a Democrat or Republican is running for reelection or whether there is an open seat: $\text{DEM}_i = 1$ if a Democrat is running for reelection in individual i's congressional district, otherwise $\text{DEM}_i = 0$; $\text{REP}_i = 1$ if a Republican incumbent is running, otherwise $\text{REP}_i = 0$ (see Appendix B, 3). If $\text{DEM}_i = \text{REP}_i = 0$, then the district has an open seat. Alesina and Rosenthal (1995) use district-level data for 1950–86 to show that midterm cycle effects predicted by their theory operate pretty much independently of incumbency effects.[18]

To understand the functional form I use for z_{phi}, $ph \in K$, recall that in equations 6a–6d an increase in x_{phi} represents an increase in the loss voter i expects from choosing the pair of candidates $ph \in K$. G_i is specified to decrease as x_{phi} increases, via $v_{phi} = \exp\{-x_{phi}\}$, $ph \in K$, so that, by equation 7, the probability that voter i chooses a candidate pair decreases if the loss expected from choosing that pair increases. In equations 6a–6d, an increase in z_{phi} implies an increase in x_{phi}, $ph \in K$. So any variable that should increase the probability of choosing the candidate pair $ph \in K$ and that is included with an additive effect in z_{phi} should have a negative coefficient.

The functional form for z_{phi}, $ph \in K$, is

$$\begin{aligned} z_{RRi} = {} & -c_{P0} - c_{H0} - c_{REP}\text{REP}_i - (c_{P1} + c_{H1})\text{PP}_i\text{EC}_i \\ & - c_D\text{PID}_{\text{D}i} - c_{ID}\text{PID}_{\text{ID}i} - c_I\text{PID}_{\text{I}i} - c_{IR}\text{PID}_{\text{IR}i} - c_R\text{PID}_{\text{R}i} - c_{SR}\text{PID}_{\text{SR}i}\ ; \end{aligned} \tag{17a}$$

$$z_{RDi} = -c_{P0} + c_{H0} - c_{DEM}\text{DEM}_i - (c_{P1} - c_{H1})\text{PP}_i\text{EC}_i\ ; \tag{17b}$$

$$z_{DRi} = c_{P0} - c_{H0} - c_{REP}\text{REP}_i + (c_{P1} - c_{H1})\text{PP}_i\text{EC}_i\ ; \tag{17c}$$

$$\begin{aligned} z_{DDi} = {} & c_{P0} + c_{H0} - c_{DEM}\text{DEM}_i + (c_{P1} + c_{H1})\text{PP}_i\text{EC}_i \\ & + c_D\text{PID}_{\text{D}i} + c_{ID}\text{PID}_{\text{ID}i} + c_I\text{PID}_{\text{I}i} + c_{IR}\text{PID}_{\text{IR}i} + c_R\text{PID}_{\text{R}i} + c_{SR}\text{PID}_{\text{SR}i}\ , \end{aligned} \tag{17d}$$

where c_{P0}, c_{P1}, c_{H0}, and c_{H1} are coefficients constant in each year, and c_D, c_{ID}, c_I, c_{IR}, c_R, c_{SR}, c_{DEM}, and c_{REP} are coefficients constant over all years. Coefficient signs should be c_{P0},

$c_{H0} < 0$ and c_{P1}, c_{H1}, c_{DEM}, c_{REP}, c_D, c_{ID}, c_I, c_{IR}, c_R, $c_{SR} > 0$. The following examples help understand the signs expected for the coefficients. First, consider the effects of party identification. Other things equal, we expect a Strong Democrat to suffer a smaller loss from choosing Democratic candidates than from choosing Republican candidates. Ignoring the effects of incumbency and economic evaluations (that is, suppose $c_{REP} = c_{DEM} = c_{P1} = c_{H1} = 0$), that is so in equations 17a–17d for presidential candidates if $c_{P0} < 0$; in that case $z_{DDi} < z_{RDi}$ and $z_{DRi} < z_{RRi}$. Likewise, it is so for House candidates if $c_{H0} < 0$; in that case $z_{DDi} < z_{DRi}$ and $z_{RDi} < z_{RRi}$. Other things equal, a Strong Republican should experience a bigger loss from choosing a Democratic candidate than a Strong Democrat would experience; hence, $c_{SR} > 0$. In omitting separate effects for the levels of party identification in the specifications for z_{RDi} and z_{DRi}, I am assuming that the partisan effects for presidential candidates are equal to those for House candidates, so that they cancel one another when a voter is evaluating the split-ticket alternatives.

Next consider the effects of incumbency status. Other things equal, the loss from choosing a Democratic House candidate should be smaller when that candidate is an incumbent; hence, $c_{DEM} > 0$. The motivation for $c_{REP} > 0$ is analogous.

Finally, consider the effects of retrospective economic evaluations. If there is a Republican President, one voter who thinks the economy has been going well and another voter who thinks the economy has been going badly, then a simple retrospective voting hypothesis says two things. First, other things equal, the voter who views the economy favorably should experience a smaller loss from choosing Republican candidates than does the voter with the unfavorable view. Second, other things equal, the voter with the more favorable view should experience a larger loss from choosing Democratic candidates than does the other voter. If there is a Republican President, we have $\mathrm{PP}_i = 1$. A voter who thinks the economy has been getting better has $\mathrm{EC}_i > 0$, while a voter who thinks the economy has been getting worse has $\mathrm{EC}_i < 0$. If $c_{P1} > 0$, then the voter with the positive view of the economy has values of z_{RRi} and z_{RDi} that are more negative than those of the voter with the negative view. The voter with the positive view has values of z_{DRi} and z_{DDi} that are more positive than those of the other voter. If $c_{H1} > 0$, then the values of z_{RRi} and z_{DRi} are more negative for the voter with the positive view of the economy than they are for the other voter, while the values of z_{RDi} and z_{DDi} are more positive for the voter with the positive view than they are for the voter who views the economy unfavorably. With a Democratic President we have $\mathrm{PP}_i = -1$, so that the orderings of the expected losses, given the voters' opinions about the economy, are reversed.

To measure vote choices Y_i (equivalently y_{phi}, $ph \in K$) I use the postelection choices reported by individuals who said they voted (see Appendix B, 4). The sample size of voters used, pooled over the six NES surveys, is 4,859 (by year, 1976–96, the sizes are 683, 627, 976, 719, 980, 874). Only those who voted for either a Democrat or a Republican for both President and House seat are included. Seven percent of such voters (386/5,245 cases) are omitted because they have data missing due to failure to measure the policy position, economic evaluation, or party identification variables (see Appendix B, 7).

Model Estimates and Results of Tests of Coordination

The coordinating and noncoordinating models produce similar results, but the former is superior. MLEs and standard errors (SEs) for the parameters of both models appear in Table 2.[19] The coordinating model estimates in Table 2 use method two (beta approximation) to compute $\hat{\bar{P}}$. Most of the parameters that have the same interpretation in both models have statistically indistinguishable estimates. The MLEs for c_{P1} and c_{H1} in each year have the correct signs for the usual kind of retrospective voting effect in which candidates of the President's party lose votes among those who think economic conditions have gotten worse, although not all the estimates are statistically significant. The MLEs for c_{P0}, c_{H0}, c_D, c_{ID}, c_I, c_{IR}, c_R, and c_{SR} are appropriate for the usual effects of party identification—slightly larger in the noncoordinating model. The MLEs for c_{DEM} and c_{REP} point to a substantial incumbent advantage, slightly larger for Republicans than for Democrats. The function β_i in the coordinating model and the corresponding functions in the noncoordinating model imply similar patterns of sensitivity to retrospective economic evaluations (discussed below).[20] But the coordinating model log-likelihood ($\hat{L}_{\mathrm{C}} = -3186$) is much greater than that of the noncoordinating model ($\hat{L}_{\mathrm{NC}} = -3225$). Using L_{TEST}, the MLE for ψ is $\hat{\psi} = 0$, with a profile-likelihood 95% confidence interval (Barndorff-Nielsen and Cox 1994, 90) of (0, .05). The non-nested hypothesis test clearly rejects the noncoordinating model and does not reject the coordinating model.

For every year except 1976, the coordinating voting model passes the parameter-based tests of the conditions necessary for it to describe coordinating behavior. This entails rejecting the spatial model that arises when $\alpha_D = \alpha_R = 1$. Table 3 reports the LR test statistics for the constraint $\alpha_D = \alpha_R = 1$, imposed separately for each year. It is rejected in every year except 1976. The 95% confidence intervals shown in Table 4 support the same conclusions. Both the LR tests and the confidence intervals in Table 4 have been adjusted in light of the fact that four parameters have MLEs on a boundary of the parameter space; see the discussion in Appendix A. Regarding the other necessary conditions, 95% confidence intervals computed as in Table 4 show q (.9, 1.1), b_P (3.1, 4.7) and b_H (2.4, 8.3) to be positive and bounded well away from zero.

As the values of $\hat{L}_{\mathrm{C}}$ reported in Table 5 suggest, the coordinating model is superior to the noncoordinating model for all four methods of computing $\hat{\bar{P}}$. The $\hat{\bar{P}}$ values from the beta approximation method are prima facie acceptable in the sense that $\hat{\bar{P}} < .5$ whenever the Democratic candidate won the presidential race and $\hat{\bar{P}} > .5$ whenever the Republican candidate won. The normal approximation method values for $\hat{\bar{P}}$ appear intuitively to be more plausible, in that they are more extreme—closer to zero or one—for years such as 1984 when prospects for one of the candidates were widely believed to be hopeless.[21] The normal approximation gives $\hat{\bar{P}} > .5$ for 1976, however, and using the normal approximation produces a substantially more negative value for $\hat{L}_{\mathrm{C}}$, indicating a much worse fit to the data, than does the beta approximation method. Both the method that directly uses the expected proportion of electoral votes as $\hat{\bar{P}}$ and the method that simply uses the expected proportion of individuals' votes perform virtually as well as the beta approximation method in terms of their $\hat{L}_{\mathrm{C}}$ values. Both have $\hat{\bar{P}} > .5$ for 1976, however.[22]

The results in Table 5 may suggest that voters use proportions instead of probabilities in evaluating weighted averages such as equation 3. Perhaps it would be better to think of equation 3 as representing some kind of expected bargaining outcome, which would parallel the interpretation being applied to $\bar{H}_i\theta_{Ri} + (1 - \bar{H}_i)\theta_{Di}$ as the expected position of the House, rather than as an expected value. Or it may be, as many studies have suggested, that people do not use probability numbers quite as formal probability theory suggests they should. Note that during the period from September 15 to election day, the average value respondents to the 1984 NES Continuous Monitoring Survey gave for Reagan's chance of winning was .80, a value greater than the electoral vote proportion value ($\hat{\bar{P}} = .70$) but smaller than the normal approximation value ($\hat{\bar{P}} = .93$) (see Appendix B, 8).

Moderation and Institutional Balancing

With expected postelection policies $\tilde{\theta}_{Di}$ and $\tilde{\theta}_{Ri}$ defined as in equation 2, moderation is almost always a feature of every voter's choices in the coordinating voting model. Unless $\alpha_D = \alpha_R = 1$, every voter intends to produce a policy outcome that is an intermediate combination of the parties' positions. The estimates for $\bar{H}$, in Table 5, show the expected position of the House, $\bar{H}\theta_{Ri} + (1 - \bar{H})\theta_{Di}$, always to have been close to the midpoint between the parties' positions. The

House position was expected to be somewhat closer to the Democratic position in 1976, 1984, 1988, and 1992, somewhat closer to the Republican position in 1996, and almost exactly at midpoint in 1980.

Such moderation, which is essentially built into the definitions of $\tilde{\theta}_{Di}$ and $\tilde{\theta}_{Ri}$, has no direct implication for the number of voters who may have been splitting their tickets to try to balance the House position with that of the future President. One point that bears on the question of such institutional balancing is the fact that the values of $\hat{\bar{P}}$ and $\hat{\bar{H}}$ in Table 5 are positively, not negatively, correlated. The product-moment correlation is .09.[23] One reason for the positive correlation is suggested by the MLEs for b_{HP} and b_{PH}. The latter is near zero and statistically insignificant, but $\hat{b}_{HP} = 1.9$.[24] Recall that, theoretically, $b_{HP} = d\bar{H}_i/d\bar{P}_i$, which should be interpreted as a relation among mutually consistent, equilibrium pairs $(\bar{H}_i, \bar{P}_i)$, $i = 1, \ldots, N$. Indeed, given equations 11a and 11b we have $b_{HP} = d\bar{H}/d\bar{P}$, a function of the common knowledge. The positive estimate for the derivative suggests that when the equilibrium value of $\bar{P}$ increases, the value of $\bar{H}$ that is mutually consistent with it tends to increase as well: An increase in the equilibrium expected probability that the Republican will win the presidency induces an increase in the equilibrium expected Republican proportion of the House vote.[25] In contrast, the estimate of virtually zero for $b_{PH} = d\bar{P}/d\bar{H}$ suggests that an increase in the expected Republican proportion of the House vote has no effect on the expected probability of a Republican presidential victory. What we have in this asymmetric relationship between expected presidential and House voting outcomes is a presidential coattail effect (Alesina and Rosenthal 1995, 104; Calvert and Ferejohn 1984; Ferejohn and Calvert 1983), most notably a coattail effect that characterizes the relationship between different equilibrium election outcomes.

The positive relationship between $\hat{\bar{P}}$ and $\hat{\bar{H}}$ reflects the patterns in which changes have occurred across election periods in ideal points, party positions, partisanship, economic evaluations, and incumbency. To assess the prevalence of balancing behavior, we need to focus on the motivation each voter has that arises solely from the policy-related components of $\lambda_{i,ph}$, $ph \in K$. In the notation of equations 6a–6d, those components are $(x_{phi} - z_{phi})$. Voter i has a policy-based incentive to balance the House with the President (or, equivalently, the President with the House) if one of the two split-ticket alternatives minimizes the policy-related component of the effect the voter's choice has on the voter's expected loss, that is, if $\text{argmin}_{ph \in K}(x_{phi} - z_{phi}) \in \{RD, DR\}$.

A necessary condition for voter i to have a policy-related incentive to balance is that the voter's ideal point falls between the policy positions the voter expects the parties to act on after the election; that is, either $\theta_{Ri} > \theta_i > \theta_{Di}$ or $\theta_{Di} > \theta_i > \theta_{Ri}$. Table 6 shows the percentage of voters who have such an ordering in each year in the NES data. The figure jumps from 29% in 1976 to 45% in 1980, then varies between 43% and 52% in subsequent years. The number of voters having a policy-related incentive to balance is therefore potentially quite large.

But only a minority of those voters who could have such an incentive, given θ_i, θ_{Di} and θ_{Ri}, actually do have such an incentive once the presidential candidates' anticipated strengths in relation to the House (α_D and α_R) and the expected election outcome ($\bar{H}$ and $\bar{P}$) are taken into account. Table 7 shows the distribution of vote choices that minimize $x_{phi} - z_{phi}$, $ph \in K$, in each year. The voters who have a policy-related incentive to balance are those for whom either $x_{RDi} - z_{RDi}$ or $x_{DRi} - z_{DRi}$ is the smallest of the four values of $x_{phi} - z_{phi}$, $ph \in K$. The percentage of such voters is smaller than 20% in every year and is less than or equal to 10% in three of the six years (1980, 1988, and 1996). Such percentages match the predictions of Alesina and Rosenthal's (1995, 103) theoretical model, in which only a small proportion of the voters who have ideal points between the positions of the parties are predicted to split their tickets in the presidential election year.

Partisanship, economic evaluations, and incumbency frequently outweigh policy-related balancing considerations in the overall determination of whether a voter casts a split-ticket vote. Moreover, the estimate for τ (95% confidence interval (.31, .47)) is large enough to suggest that a fair number of split-ticket votes are significantly motivated by other considerations that either enhance or reduce the appeal of both split-ticket alternatives without sharply distinguishing between them. Of those found to have a policy-related balancing incentive in the NES data, over all years, only 31% actually split their tickets. Only 22% of those who did *not* have a policy-related balancing incentive split their tickets. The Pearson chi-squared statistic for independence in the cross-classification of split-ticket versus straight-ticket votes by split-ticket versus straight-ticket policy-related incentives (measured by $\operatorname{argmin}_{ph \in K}(x_{phi} - z_{phi})$) is 18.8 with 1 degree of freedom, which indicates a statistically significant association at any conventional test level.[26] Factors that may have little or nothing to do with a voter's policy preferences substantially affect the voter's choices. But both the number of voters who have a policy-related balancing incentive and the degree to which that incentive affects their choices are large enough to support a conclusion that policy-related balancing has often been an important determinant of election outcomes.

Conclusion

Recent American elections have featured coordination among voters, based on voters' intentions to produce postelection policy moderation between the President and the House. Inspired by policy concerns that have been mediated through a coordination-with-moderation mechanism, a small but significant proportion of voters have been motivated to vote a split ticket in order to increase the chances of institutional balance. Perhaps not in 1976, but in each presidential-year election from 1980 through 1996, each voter's choices between the major party candidates for President and the House have been constrained by the anticipated election result, that is, by the aggregate result of the choices all other voters were about to make. In recent elections the strategies and beliefs of different voters have been strongly tied together.

For the most part, moderation has been based on voters' expectations that the President will be at least an equal of the House in determining postelection policy. Most of the α_D and α_R point estimates are greater than .5 (see Table 2), which suggests that voters usually believe the President to have more weight in policy outcomes than does the House. Jimmy Carter running for reelection in 1980 is an exception ($\hat{\alpha}_{D,80} = .4$), but the most striking case of anticipating a weak President is Bob Dole in 1996 ($\hat{\alpha}_{R,96} = .1$). One interpretation is that voters expected a Dole presidency to be driven by the centralized and confrontational House of Speaker Newt Gingrich.

In at least one respect, the moderating mechanism has worked to the disadvantage of Democratic candidates. The estimates for the weight variable β_i (see equation 16) suggest that the Democratic party cannot choose policies that are more extreme than those of the Republican party, without suffering electoral disadvantage, but the Republican party has often not been similarly restricted. Voters who think economic conditions have worsened treat the discrepancies between their ideal points and the policies they expect will occur with each party's President about equally: someone who says the economy is "worse" has $\beta_i = .52$, and someone who says "much worse" has $\beta_i = .46$.[27] But voters who think economic conditions have improved put much more weight on the discrepancy with a Democratic President: someone who says the economy is "better" has $\beta_i = .63$, and someone who says "much better" has $\beta_i = .69$. Such a pattern is compatible with voters treating the Democratic party more skeptically in times of prosperity, because of its interventionist reputation, but not giving it any particular advantage when times are tough. Whatever its origin, the asymmetric treatment of the parties means that the Democrats have had less leeway than the Republicans to choose policy positions that are not at the center of the distribution of voters' ideal points.

The coordinating voting model is underdeveloped in at least one crucial respect: the model says nothing about how the equilibrium described by the fixed-point pair $(\bar{H}, \bar{P})$ and other common knowledge might come about. It is unreasonable to think that voters spontaneously agree on the equilibrium values. Indeed, it is unreasonable to believe that each voter independently keeps track of all the information needed to define and sustain an equilibrium. A coordinating voting equilibrium can exist only if institutions exist that can aggregate and broadcast the information everyone needs to know.

Such institutions exist. McKelvey and Ordeshook show that polls (1984, 1985a) or interest groups and history (1985b) may support rational expectations voter equilibria even when many voters are poorly informed. It should be possible to build such institutions directly into a model much like the current one, with similar empirical implications but with weaker assumptions about what each voter knows about the distribution of voters.[28] More difficult to model would be an institution that brings everyone to agreement about the current values of the model's parameters. Informally it is reasonable to say that this may happen during the electoral campaign, but it is difficult to say anything specific without invoking assumptions at least as strong as the direct assumption that the parameters are common knowledge.[29]

On the whole, it seems reasonable to say that the equilibria here modeled with virtually no institutional foundations are in reality sustained by a collection of institutions that may informally be summarized as "the campaign," "polls," and "pundits." The campaign carries the processes through which the expectations here modeled as $(\bar{H}, \bar{P})$ and the party and governmental struggles here modeled using parameters such as ρ_D, ρ_R, α_D, and α_R converge to equilibrium values. Polls and pundits allow each voter to monitor the expectations and struggles for power as they evolve, without the voter having to exert much personal effort. To learn the candidates' chances in the presidential race and the expected outcome in House races across the country, a voter need only invest a little time listening to pollsters say what they think will happen. Brief attention to media commentators may let a voter know how strong each presidential candidate may be expected to be, if elected, relative to the candidate's party and to the legislature. Having so easily learned the commonly accepted expectations for the election outcome and for the power relationships, all a voter then needs in order to coordinate with all other voters is knowledge of the components of the voter's own personal preferences: the voter's ideal point and what the voter thinks are the relevant policy positions of the parties and candidates; the voter's partisanship and evaluation of the economy; the presence or absence of an incumbent in the House race in the voter's district; and other factors here treated as random. If sufficient institutional support exists, it is not difficult for voters to coordinate.

Appendix A. Proofs and Technical Measurement Issues

Effects of Candidate Choices on a Voter's Expected Loss

Using $d\beta_i/d\bar{P}_i = d\beta_i/d\bar{H}_i = 0$, we have $d\lambda_i/d\bar{P}_i = w_{Pi} + w_{Hi}d\bar{H}_i/d\bar{P}_i + d\xi_i/d\bar{P}_i$ and $d\lambda_i/d\bar{H}_i = w_{Hi} + w_{Pi}d\bar{P}_i/d\bar{H}_i + d\xi_i/d\bar{H}_i$. The other variables used in equations equations 4a–4d are $\xi_{RRi} = b_P d\xi_i/d\bar{P}_i + b_H d\xi_i/d\bar{H}_i$, $\xi_{RDi} = b_P d\xi_i/d\bar{P}_i - b_H d\xi_i/d\bar{H}_i$, $\xi_{DRi} = -b_P d\xi_i/d\bar{P}_i + b_H d\xi_i/d\bar{H}_i$, and $\xi_{DDi} = -b_P d\xi_i/d\bar{P}_i - b_H d\xi_i/d\bar{H}_i$.

Probability That the Republican Wins the Presidency (Using Only Group Information)

Let V_{jk} be the set and $N_{jk} = |V_{jk}|$ the number of voters in State j who are members of group V_k. $\bar{V}_j = \bigcup_{k=1}^{M} V_{jk}$ is the set of all voters in State j, and the total number of voters in the State is $N_j = \sum_{k=1}^{M} N_{jk}$. Let $\bar{V}_j^{(r)}$ denote the set of all subsets of $\bar{V}_j$ of size r. If Z_i is known for all i, then the probability that the Republican gets a majority of the votes in State j is

$$\bar{y}^*_{Pj} = \frac{\sum_{r=\lceil N_j/2\rceil}^{N_j} \sum_{s\in\bar{V}_j^{(r)}} \left(\prod_{i\in s} \pi_i\right) \left[\prod_{i\in V_j\backslash s}(1-\pi_i)\right]}{\sum_{r=0}^{N_j} \sum_{s\in\bar{V}_j^{(r)}} \left(\prod_{i\in s} \pi_i\right) \left[\prod_{i\in V_j\backslash s}(1-\pi_i)\right]}$$

$$= \frac{\sum_{r=\lceil N_j/2\rceil}^{N_j} \sum_{s\in \bar{V}_j^{(r)}} \prod_{k=1}^{M} \left(\prod_{i\in s_{jk}} \pi_i\right) \left[\prod_{i\in V_{jk}\setminus s_{jk}} (1-\pi_i)\right]}{\sum_{r=0}^{N_j} \sum_{s\in \bar{V}_j^{(r)}} \prod_{k=1}^{M} \left(\prod_{i\in s_{jk}} \pi_i\right) \left[\prod_{i\in V_{jk}\setminus s_{jk}} (1-\pi_i)\right]}\,, \quad s_{jk} = s\cap V_{jk}\ .$$

The probability that the Republican candidate gets half or more of the Electoral College votes is

$$\bar{P}^* = \frac{\sum_{r=C/2}^{C} \sum_{s\in C^{[r]}} \left(\prod_{j\in s} \bar{y}^*_{Pj}\right) \left[\prod_{j\notin s}(1-\bar{y}^*_{Pj})\right]}{\sum_{r=0}^{C} \sum_{s\in C^{[r]}} \left(\prod_{j\in s} \bar{y}^*_{Pj}\right) \left[\prod_{j\notin s}(1-\bar{y}^*_{Pj})\right]}\ .$$

Replacing each π_i in $\bar{y}^*_{Pj}$ with the group-specific probability $\bar{\pi}_k$, $i\in V_k$, gives $\bar{P}$ of equation 10 with

$$\bar{y}_{Pj} = \frac{\sum_{r=\lceil N_j/2\rceil}^{N_j} \sum_{s\in \bar{V}_j^{(r)}} \prod_{k=1}^{M} \bar{\pi}_k^{|s_{jk}|}(1-\bar{\pi}_k)^{|V_{jk}\setminus s_{jk}|}}{\sum_{r=0}^{N_j} \sum_{s\in \bar{V}_j^{(r)}} \prod_{k=1}^{M} \bar{\pi}_k^{|s_{jk}|}(1-\bar{\pi}_k)^{|V_{jk}\setminus s_{jk}|}}\,, \quad s_{jk} = s\cap V_{jk}\ . \tag{18}$$

Since each Z_i is independent and $M \ll N$, Slutzky's theorem implies that $\int_{\tilde{Z}} \cdots \int_{\tilde{Z}} \bar{P}^* \prod_{i=1}^{N} df_{k_i}(Z_i)$ converges in probability to $\bar{P}$ as $N\to\infty$ in any way such that, for every j and k, $N_{jk}\to\infty$ or $N_{jk}=0$. In terms of asymptotic approximation error,

$$\underbrace{\int_{\tilde{Z}} \cdots \int_{\tilde{Z}}}_{N \text{ times}} \bar{P}^* \prod_{i=1}^{N} df_{k_i}(Z_i) = \bar{P} + O_p\left(\max_{j,k:\, N_{jk}\neq 0} N_{jk}^{-1/2}\right)\ .$$

$\bar{P}$ therefore differs negligibly from $\int_{\tilde{Z}} \cdots \int_{\tilde{Z}} \bar{P}^* \prod_{i=1}^{N} df_{k_i}(Z_i)$ if N_{jk} is large for all j and k.

Existence of Fixed-Point Pairs $(\bar{H}, \bar{P})$

Let Ψ be a bounded set of positive Lebesgue measure that is the parameter space for $\lambda_{i,ph}$, $ph\in K$.[30] Let $\bar{\pi}_{k\ell\jmath}$ and $\bar{\eta}_{k\ell\jmath}$ denote the values of $\bar{\pi}_k$ and $\bar{\eta}_k$ when evaluated using $\bar{P}=\ell$ and $\bar{H}=\jmath$ for some $\ell\in[0,1]$ and some $\jmath\in[0,1]$. Let $\bar{P}_{\ell\jmath}$ denote the value of $\bar{P}$ as defined by equation 10 when $\bar{\pi}_k=\bar{\pi}_{k\ell\jmath}$, and let $\bar{H}_{\ell\jmath}$ denote the value of $\bar{H}$ as defined by equation 9 when $\bar{\eta}_k=\bar{\eta}_{k\ell\jmath}$.

Theorem 2 *For almost every $\psi\in\Psi$, there exist values $\ell,\jmath\in(0,1)$ such that, simultaneously, $\bar{P}_{\ell\jmath}=\ell$ and $\bar{H}_{\ell\jmath}=\jmath$.*

Because $\bar{\pi}_{k\ell\jmath}$ varies continuously as a function of ℓ for each k and $\bar{P}_{\ell\jmath}$ is a continuous function of the $\bar{\pi}_{k\ell\jmath}$ values, there is at least one fixed point of $\bar{P}_{\ell\jmath}$ as a function of ℓ for each fixed value of $\jmath$, that is, there is at least one value ℓ such that $\bar{P}_{\ell\jmath} = \ell$. Because $\bar{P}_{0\jmath} > 0$ and $\bar{P}_{1\jmath} < 1$, the number of fixed points is odd, except on a subset of Ψ of measure zero. For $q > 1$ there is similarly an odd number of fixed points $\bar{H}_{\ell\jmath} = \jmath$ for each ℓ. If $q = 1$, the step functions in w_{Hi} imply that $\bar{H}_{\ell\jmath}$ does not vary continuously as a function of $\jmath$, so if $q = 1$ there is not necessarily a fixed point $\bar{H}_{\ell\jmath} = \jmath$. If $0 < q < 1$, then there is a discontinuity at infinity in w_{Hi} whenever $\theta_i = \tilde{\theta}_{Di}$ or $\theta_i = \tilde{\theta}_{Ri}$. But for a finite number N of voters, the number of points of discontinuity in both cases is finite.[31] So the values of ℓ for which there is no fixed point $\bar{H}_{\ell\jmath} = \jmath$ form a set of measure zero: for any $\epsilon > 0$ there is a value ℓ', $|\ell - \ell'| < \epsilon$, such that there is a fixed point $\bar{H}_{\ell'\jmath} = \jmath$ for some $\jmath \in [0, 1]$. Now consider the graph of the set of fixed points $\bar{P}_{\ell\jmath} = \ell$ as $\jmath$ varies over $[0, 1]$ and the graph of the set of fixed points $\bar{H}_{\ell\jmath} = \jmath$ as ℓ varies over $[0, 1]$. When $q > 1$ both graphs are continuous and so necessarily intersect. That is, there is necessarily at least one pair of values $(\ell, \jmath)$ such that $\bar{P}_{\ell\jmath} = \ell$ and $\bar{H}_{\ell\jmath} = \jmath$. If $0 < q \leq 1$, then each graph is in general not continuous, because of the discontinuities in w_{Hi}, but each graph does have a finite number of continuous components. In this case, the graphs need not intersect. But because the number of points of discontinuity is finite, the set of parameter values for which there is no intersection has measure zero. For finite $N > 0$, the argument holds for arbitrary finite sets of probability measures f_k. *Q.E.D.*

Survey-based Probabilities That the Republican Wins the Presidency ($\hat{\hat{P}}$)

Let C_{NR} denote the number of electoral votes assigned to the Republican according to the $\bar{v}_j = \emptyset$ convention, and let C_{ND} denote the number thus assigned to the Democrat. The proportion of the remaining electoral votes that the Republican needs to win is $\zeta_\pi = (269 - C_{NR})/(538 - C_{NR} - C_{ND})$. *Beta approximation method.* For States with sample data, set $\hat{\hat{y}}_{Pj} = \hat{y}_{Pj}$. Compute the variance of ζ_R as $\sigma^2_{\zeta_R} = (\sum_{j=1}^{S} C_j^2)^{-1} \sum_{j=1}^{S} \hat{\hat{y}}_{Pj}(1 - \hat{\hat{y}}_{Pj})C_j^2$. To evaluate $\Pr(\zeta_R > .5)$, use the beta distribution on $[0,1]$ obtained by matching moments:

$$\hat{\hat{P}} = \int_{\zeta_\pi}^{1} \frac{\zeta_R^{p_1-1}(1-\zeta_R)^{p_2-1}}{B(p_1,p_2)} d\zeta_R \,, \tag{19}$$

where $p_1 = \bar{\zeta}_R^2(1-\bar{\zeta}_R)/\sigma_{\zeta_R}^2 - \bar{\zeta}_R$ and $p_2 = \bar{\zeta}_R(1-\bar{\zeta}_R)^2/\sigma_{\zeta_R}^2 - (1-\bar{\zeta}_R)$ match the mean $\bar{\zeta}_R$ and variance $\sigma_{\zeta_R}^2$ (Johnson, Kotz, and Balakrishnan 1995, 222), and $B(p_1,p_2)$ denotes the beta function. If $p_1 \leq 0$ or $p_2 \leq 0$, then the beta distribution is inappropriate. *Normal approximation method.* Set $\hat{\hat{y}}_{Pj} = 1 - \Phi((.5 - \hat{y}_{Pj})/\sigma_{\hat{y}_{Pj}})$ and evaluate $\Pr(\zeta_R > .5)$ as $\hat{\hat{P}} = 1 - \Phi((\zeta_\pi - \bar{\zeta}_R)/\sigma_{\zeta_R})$.

Iterative Computation of MLEs and Expectations $(\hat{\hat{H}}, \hat{\hat{P}})$

Given estimates for the parameters and estimates $\hat{\hat{H}}^{(\imath)}$ and $\hat{\hat{P}}^{(\imath)}$ from iteration step $\imath$, use equation 7 to compute probability values $\hat{\mu}_{phi}^{(\imath)}$, $ph \in K$, and then $\hat{\pi}_i^{(\imath)}$ and $\hat{\eta}_i^{(\imath)}$ for each voter. For each State $\jmath$ with $\bar{v}_\jmath \neq \emptyset$ in the referent year, use equation 13 to compute $\hat{y}_{Pj}^{(\imath)}$ and hence compute an updated estimate $\hat{\hat{P}}^{(\imath+1)}$ (e.g., using mean $\bar{\zeta}_R^{(\imath)}$, variance $\sigma_{\zeta_R}^{2(\imath)}$ and equation 19). Use equation 12 to estimate $\hat{\hat{H}}^{(\imath+1)}$, with n being the total sample size of voters for the referent year. Use $\hat{\hat{H}}^{(\imath+1)}$ and $\hat{\hat{P}}^{(\imath+1)}$ and the parameter values of step $\imath$ to apply one step of Gauss-Newton maximization to the log-likelihood, in order to get new parameter values for use in iteration $\imath+1$. Iterate until, for successive iterations $\imath$ and $\imath+1$, $(\hat{\hat{H}}^{(\imath)}, \hat{\hat{P}}^{(\imath)}) = (\hat{\hat{H}}^{(\imath+1)}, \hat{\hat{P}}^{(\imath+1)})$ and the parameter estimates satisfy the first- and second-order conditions for a local maximum of the log-likelihood. In each Gauss-Newton step, each row of the Jacobian matrix of the parameters is weighted by $1/\omega_i$.[32] $1/\omega_i$ is scaled to sum to the total sample size $\tilde{n}$ in each year, where $\tilde{n}$ includes all survey respondents, both voters and nonvoters (see Appendix B, 5).

For a fixed set of parameter estimates, the sequence $(\hat{\hat{H}}^{(1)}, \hat{\hat{P}}^{(1)}), (\hat{\hat{H}}^{(2)}, \hat{\hat{P}}^{(2)}), \ldots$ converges if, to extend the notation of the paragraph on the existence of fixed-point pairs, $\hat{\hat{P}}_{\ell\jmath}$ is decreasing in $\jmath$ and $\hat{\hat{H}}_{\ell\jmath}$ is decreasing in ℓ for $\jmath, \ell \in (0,1)$.

Estimates on a Boundary of the Parameter Space

Table 2 shows $\alpha_{D,76}$, $\rho_{D,92}$, $\rho_{R,80}$, and $\rho_{R,96}$ to have MLEs equal to 1.0, on the conceptual boundary of the parameter space. The boundary is not a "natural" boundary (Moran 1971b) for the probability model: L_C is a proper (although meaningless) likelihood for values of α_D, ρ_D, or ρ_R outside $[0, 1]$. The on-boundary values do not imply problems with the consistency of the MLEs (Moran 1971a): The model is identified and the MLEs globally maximize L_C subject to $0 \leq \alpha_D, \alpha_R, \rho_D, \rho_R \leq 1$. But the MLEs' asymptotic distribution is complicated (Moran 1971b; Self and Liang 1987). To regularize the model for the LR tests, in all the constrained models I fix equal to 1.0 the parameters with on-boundary MLEs. With this method the test statistic for $\alpha_{D,76} = \alpha_{R,76} = 1$ has only one degree of freedom. None of the statistics in Table 3 are close to the critical value for a test based on χ^2_2 (χ^2_1 for 1976), so slight variations from χ^2_2 in the statistics' distributions would not change any conclusions. Unfortunately, it is computationally infeasible to bootstrap the LR test statistics.

The asymptotic distribution of the MLEs when parameters fall on a boundary of the parameter space is a mixture of censored multivariate normal distributions (Moran 1971b; Self and Liang 1987). For the coordinating voting model, the hypothesis $\alpha_{D,76} = \rho_{D,92} = \rho_{R,80} = \rho_{R,96} = 1$ implies a distribution that is a mixture of 16 censored distributions. I use a bootstrap (20,000 resamples) of the score vectors associated with the MLEs of Table 2 to tabulate that mixture distribution and estimate the confidence intervals of Table 4. With a few substantial exceptions (e.g., $\alpha_{R,88}$) the confidence intervals do not differ greatly from the intervals obtained by using normal theory with the MLEs and SEs in Table 2 (i.e., MLE $\pm$ 1.96 SE).

Comparison Weights in the Noncoordinating Vote Choice Model

The weight of the spatial comparison associated with each form of government is

$$\beta_{ji} = \frac{\exp\{-b_{jE0} - b_{jE1}\mathrm{EC}_i\}}{1 + \sum_{j' \in \{RR,RD,DR\}} \exp\{-b_{j'E0} - b_{j'E1}\mathrm{EC}_i\}}, \qquad j \in \{RR, RD, DR\}$$

$$\beta_{DDi} = 1 - \beta_{RRi} - \beta_{RDi} - \beta_{DRi}$$

for coefficients b_{jE0}, b_{jE1}, $j \in \{RR, RD, DR\}$. See the text preceding equation 16 for the definition of EC_i. If all forms of government are treated symmetrically, $\beta_{RRi} = \beta_{RDi} = \beta_{DRi} = \beta_{DDi} = \frac{1}{4}$.

Measuring Ideal Points and Policy Positions

To measure θ_i, ϑ_{Di}, ϑ_{PDi}, ϑ_{Ri}, and ϑ_{PRi}, I use a collection of sets of five seven-point (and one set of five four-point) placement scales. Each set includes one scale for self, one for each party, and one for each presidential candidate, all referring to the same substantive policy description (see Appendix B, 6). All scale items are oriented so that the "liberal" position during the given time period has the lower number.

The general motivation for the response codes is the observation that any continuous random variable X can be mapped onto a scale such that the values of the scale have probability mass spread uniformly over $[0, 1]$. Simply map each value x of the variable to the value of the variable's distribution function, $F_X(x) = \Pr(X < x)$. If X is multivariate with components $X_1, \ldots, X_k$, then the mapping from vectors $(x_1, \ldots, x_k)$ to a uniform distribution on $[0, 1]$ may be based on the distribution function $F_X(x) = \Pr(X_1 < x_1, \ldots, X_k < x_k)$. I do not try to estimate the joint, five-dimensional distribution function for the five scales of each substantive description. Rather I crudely approximate the joint distribution by computing the simple average of the observed one-dimensional marginal distributions.

To assign a code in the $[0, 1]$ interval to the response levels of each scale, I first compute the cumulative response proportions for the five scales in the set for each substantive description. Let $0 = r_{sj0} \leq r_{sj1} \leq \cdots \leq r_{sj6} \leq r_{sj7} = 1$ denote the successive cumulative proportions for substantive description s (e.g., s = "Liberal/Conservative") and scale $j \in \{S, D, R, PD, PR\}$, S = self, D = Democratic party, R = Republican party, PD = Democratic candidate, and PR = Republican candidate. For each $m \in \{0, \ldots, 7\}$, I compute $r_{sm} = (r_{sSm} + r_{sDm} + r_{sRm} + r_{sPDm} + r_{sPRm})/5$. For all five scales of type s, the numerical code used for original response category $m \in \{1, \ldots, 7\}$ is $\bar{r}_{sm} = (r_{s,m-1} + r_{sm})/2$. To determine values for θ_i, ϑ_{Di}, ϑ_{PDi}, ϑ_{Ri}, and ϑ_{PRi} for each voter, I average the score for the corresponding scale j over all (and only) the substantive descriptions for which the person responded to all five scales for that substantive description.

Appendix B. Data Notes

1. *Economic Evaluations* (EC_i). For 1976 the question wording is: "Would you say that at the present time business conditions are better or worse than they were a year ago?" For 1980, 1984, and 1988 the question wording is: "What about the economy? Would you say that over the past year the nation's economy has gotten better, stayed about the same, or gotten worse?" For 1992 the initial part of the question was changed to read: "How about the economy." For 1996 the initial part was: "Now thinking about the economy in the country as a whole." For all years except 1976, the responses are coded "much worse" (-1), "somewhat worse" ($-.5$), "same" (0), "somewhat better" (.5), and "better" (1). For 1976 only three levels of response were recorded, coded here as "worse now" ($-.5$), "about the same" (0), and "better now" (.5). The NES variable numbers for each year are 3139 (1976), 150 (1980), 228 (1984), 244 (1988), 3532 (1992), and 960386 (1996).

2. *Party Identification* ($\mathrm{PID}_{\mathrm{D}i}$, $\mathrm{PID}_{\mathrm{ID}i}$, $\mathrm{PID}_{\mathrm{I}i}$, $\mathrm{PID}_{\mathrm{IR}i}$, $\mathrm{PID}_{\mathrm{R}i}$, $\mathrm{PID}_{\mathrm{SR}i}$). The levels of the party identification scale are Strong Democrat, Democrat, Independent Democratic, Independent, Independent Republican, Republican, and Strong Republican. The variable numbers are 3174

(1976), 266 (1980), 866 (1984), 274 (1988), 3634 (1992), and 960420 (1996).

3. *Incumbency Status* (DEM_i, REP_i). For 1976, incumbency status variables were built from note 17 in the codebook file nes1976.cbk. The variable numbers for the other years are 740 (1980), 59 (1984), 50 (1988), 3021 (1992; with errors corrected as indicated in the codebook file nes92int.cbk), and 960097 (1996).

4. *Vote Choices* (y_{phi}). The variable numbers are 3665 and 3673 (1976), 994 and 998 (1980), 788 and 793 (1984), 763 and 768 (1988), 5609 and 5623 (1992), and 961082 and 961089 (1996).

5. *Sampling Weights* ($1/\omega_i$). In the NES data, $1/\omega_i$ is the number of adults in each household, which is multiplied by a time-series weight in years (1976, 1992, 1996) when part of the sample belonged to a multiyear panel cohort. I rescaled the number of adults and time-series weight variables to give each a mean of 1.0 over the whole of each survey sample. The variable numbers are 3003 and 3021 (1976); 43 (1980); 70 (1984); 91 (1988); 29, 3076, and 7000 (1992); and 960046A and 960005 (1996).

6. *Placement Scales* (θ_i, ϑ_{Di}, ϑ_{PDi}, ϑ_{Ri}, ϑ_{PRi}). Here are the brief substantive description and variable numbers for each set of scales for each year. The label "reversed" indicates an item whose categories were reordered to reverse the original ordering of 1 to 7 (or for Abortion in 1996, 1 to 4). **1976**: Government Guaranteed Job and Living Standard, 3241–5; Rights of the Accused, 3248–52; School Busing to Achieve Integration, 3257–61; Government Aid to Minorities, 3264–8; Government Medical Insurance Plan, 3273–7; Liberal/Conservative Views, 3286–90; Government Guaranteed Job and Living Standard, 3758–62; Urban Unrest, 3767–71; Legalization of Marijuana, 3772–6; Change in Tax Rate, 3779–83; and Equal Rights for Women, 3787–91. **1980**: Liberal/Conservative, 267–9, 278, 279; Defense Spending, 281–3, 286, 287; Government Services/Spending (reversed), 291–3, 296, 297; Reduce Inflation/Reduce Unemployment (reversed), 301–3, 306, 307; Liberal/Conservative Views, 1037–9, 1053, 1054; Government Aid to Minorities,

1062–4, 1073, 1074; Getting Along with Russia, 1078–80, 1089, 1090; Equal Rights for Women, 1094–6, 1105, 1106; and Government Guaranteed Job and Living Standard, 1110–2, 1121, 1122. **1984**: Liberal/Conservative Placement, 119–21, 123–5, 127–9, 131–3, 135–7; Liberal/Conservative, 369, 371–4; Government Services/Spending (reversed), 375–9; Minority Aid/No Aid, 382–6; Involvement in Central America (reversed), 388–92; Defense Spending, 395–9; Social/Economic Status of Women, 401–5; Cooperation with Russia, 408–12; and Guaranteed Standard of Living/Job, 414–8. **1988**: Liberal/Conservative, 228, 231, 232, 234, 235; Government Services/Spending (reversed), 302–4, 307, 308; Defense Spending, 310–2, 315, 316; Government-Funded Insurance, 318–22; Guaranteed Standard of Living/Job, 323–5, 328, 329; Social/Economic Status of Blacks, 332–4, 337, 338; Social/Economic Status of Minorities, 340–2, 345, 346; Cooperation with Russia, 368–70, 373, 374; and Women's Rights, 387–91. **1992**: Ideological Placement, 3509, 3514, 3515, 3517, 3518; Government Services/Spending (reversed), 3701–5; Defense Spending, 3707–11; and Job Assurance, 3718–22. **1996** (omitting the variable number prefix 960): Liberal/Conservative, 365, 369, 371, 379, 380; Government Services/Spending (reversed), 450, 453, 455, 461, 462; Defense Spending, 463, 466, 469, 477, 478; Abortion (reversed), 503, 506, 509, 517, 518; Jobs/Environment, 523, 526, 529, 535, 536; and Environmental Regulation, 537, 538, 539, 541, 542.

7. *Missing Data.* The loss of cases from ignoring third-party votes substantially affects only presidential votes. Among the 7,198 cases in which any vote cast was for a Democrat or a Republican, 1,426 cast a vote for only one office (either President or House member). Among the 5,772 who said they voted for Democratic or Republican candidates for both offices, 527 were excluded because the incumbent in their district was unopposed. Of the remaining cases, 379 were missing data for the ideal point, policy positions or economic evaluation (295 missing only one or more position variables, 58 missing only EC_i, 26 missing both); seven more cases were missing data for party identification.

8. *Reagan's Chance of Winning*. The Continuous Monitoring Survey is Miller and the National Election Studies (1985). The item wording (var. 1001) is: "As you know, Ronald Reagan has been nominated by the Republican party. What chance do you think he has of winning the 1984 Presidential election? We will use a scale that goes from 0 to 100 where 0 represents NO chance of winning the Presidency, 50 represents an even chance, and 100 represents certain victory." In computing the mean I weighted observations in proportion to the number of adults (var. 16) and in inverse proportion to the number of nonbusiness telephones (vars. 1129, 1131, 1133) in the household. The text reports the mean divided by 100.

Notes

Walter R. Mebane, Jr., is Associate Professor of Government, Cornell University, Ithaca, NY 14853-4601 (email wrm1@cornell.edu).

Earlier versions of this paper were presented at the 1998 Summer Methods Conference, University of California, San Diego, July 23–26; 1998 annual meeting of the Midwest Political Science Association, April 23–25, Chicago, IL; and 1997 Summer Methods Conference, Ohio State University, Columbus, July 24–27. I worked on the paper while visiting during fall 1997 at the Department of Social and Decision Sciences at Carnegie Mellon University, at the invitation of Bill Keech. I thank Jasjeet Sekhon, Jonathan Wand, Jonathan Cowden, and Matthew Wexler for helpful comments, Jonathan Cowden and Jasjeet Sekhon for letting tron and lapo help macht with the computing, and Jonathan Hoffman for assistance in preparing the final copy of the manuscript. Data were made available in part by the Cornell Institute for Social and Economic Research and the Inter-University Consortium for Political and Social Research. All errors are solely the responsibility of the author.

1. Studies of electoral coordination often focus on the "wasted vote" logic of Duverger's Law, which is said to keep the number of parties receiving votes small (e.g., Cox 1997; Fey 1997; Palfrey 1989). I examine coordination that relates not to the number of parties—two are assumed—but to possible moderation in voting for President and the House of Representatives.

2. McKelvey and Ordeshook's (1984, 1985a, 1985b) equilibrium concepts are similar in that they use rational expectations, but they differ technically because their models feature purely spatial utility rather than probabilistic voting, assume the existence of a single policy dimension (although some voters are uninformed about candidates' positions), and have each voter voting for only one office rather than two. Enelow and Hinich's (1982, 1983a, 1983b) probabilistic vote choice models also emphasize voters' expectations regarding other voters' votes.

3. See especially note 28.

4. Let $\lambda_{i,R}$ denote voter i's loss if i chooses the Republican candidate, and let $\lambda_{i,D}$ denote the loss if i chooses the Democrat. Obviously, $\lambda_{i,R} - \lambda_{i,D} = (\bar{P}_{i,R} - \bar{P}_{i,D})(\lambda_{i,R} - \lambda_{i,D})/(\bar{P}_{i,R} - \bar{P}_{i,D})$, if $\bar{P}_{i,R} - \bar{P}_{i,D} > 0$. But if $\bar{P}_{i,R} - \bar{P}_{i,D}$ is very small, then $(\lambda_{i,R} - \lambda_{i,D})/(\bar{P}_{i,R} - \bar{P}_{i,D}) \approx d\lambda_i/d\bar{P}_i$.

5. Strictly speaking, to define $\bar{P}_i$ with reference to the Electoral College, I will assume that the number of voters in each of the M groups in each State is common knowledge.

6. See Maddala 1983 for an introductory discussion of GEV choice models.

7. It is possible to specify GEV distributions in which there is also dependence among all the pairs that include a candidate from one party. That is, ϵ_{RRi}, ϵ_{RDi}, and ϵ_{DRi} are dependent, or ϵ_{RDi}, ϵ_{DRi}, and ϵ_{DDi} are dependent. In an analysis reported in Mebane 1998, using the same NES data, tests show neither party-based pattern of dependence to be present, as long as party identification is included among the variables in $(z_{RRi}, z_{RDi}, z_{DRi}, z_{DDi})$. Given those empirical results, I simplify the current model by ignoring the possibility of party-based dependence.

8. To accommodate the combination of continuous and discrete variables that Z_i may contain, the integral in $\Pr(Y_i = ph \mid i \in V_k)$ may be thought of as a Riemann-Stieltjes, or if necessary Lebesgue-Stieltjes, integral (Billingsley 1986, 230).

9. $\bar{P}$ is the expected probability that the Republican wins or ties in the Electoral College. I ignore the complications involved in breaking a tie by votes in the House.

10. An exact tie between two $\lambda_{i,ph}$ values, $ph \in K$, is a measure zero event that may be ignored.

11. Let f_ϵ denote the density of ϵ_i on $\mathbb{R}^4$ (the real line four times). The basic result is

$$\begin{aligned}
&\int_{\tilde{Z}} \int_{\mathbb{R}^4} (\tilde{\mu}_{phi} - \bar{\mu}_{phk_i}) df_\epsilon(\epsilon_i) df_{k_i}(Z_i) \\
&\quad = \int_{\tilde{Z}} \int_{\mathbb{R}^4} \tilde{\mu}_{phi} df_\epsilon(\epsilon_i) df_{k_i}(Z_i) - \bar{\mu}_{phk_i} \int_{\tilde{Z}} \int_{\mathbb{R}^4} df_\epsilon(\epsilon_i) df_{k_i}(Z_i) \\
&\quad = \int_{\tilde{Z}} \mu_{phi} df_{k_i}(Z_i) - \bar{\mu}_{phk_i} = 0 \, , \quad ph \in K, \quad i = 1, \ldots, N \, .
\end{aligned}$$

We have immediately that $N^{-1} \sum_{i=1}^{N} \int_{\tilde{Z}} \int_{\mathbb{R}^4} (\tilde{\eta}_i - \bar{\eta}_{k_i})/N \, df_\epsilon(\epsilon_i) df_k(Z_i) = 0$ and, by Slutzsky's theorem based on the continuity of g_i and using $M \ll N$, $N^{-1} \sum_{i=1}^{N} \int_{\tilde{Z}} \int_{\mathbb{R}^4} g_i(\tilde{\pi}_i - \bar{\pi}_{k_i}) df_\epsilon(\epsilon_i) df_k(Z_i) \xrightarrow{\Pr} 0$.

12. Isaki and Fuller (1982) rigorously develop the relevant concept of consistency. See Särndal, Swensson, and Wretman (1992, 166–85) for an overview.

13. If $\mathrm{var}(\epsilon_{phi}) = \frac{1}{6}\pi^2\sigma^2$, $ph \in K$ (see Johnson, Kotz and Balakrishnan 1995, 12), then the estimated parameters correspond to $b_P = \bar{P}_{i,R} - \bar{P}_{i,D}$ and $b_H = \bar{H}_{i,R} - \bar{H}_{i,D}$ divided by σ.

14. Born's (1994a) implementation of Fiorina's theory is similar to the noncoordinating model, the differences being that Born constrains $\alpha_D = \alpha_R$; assumes uncorrelated disturbances ($\tau = 0$); fixes $q = 1$; uses a single, constant parameter for each of the coefficients here allowed to vary over voters via $b_{NC}\beta_{phi}$, $ph \in K$; measures ideal points and policy positions differently; and includes in z_{phi} variables to measure "Evaluation of President's Job Performance" but not retrospective economic evaluations. Born does not estimate the $\alpha_R = \alpha_D$ parameter simultaneously with the rest of his model but instead repeatedly reestimates the model for several values in the range .5 to .85.

15. Ingberman and Villani (1993) study a model in which there are both strategic and nonstrategic voters.

16. Data and documentation for 1976–92 are from the "ANES 1948–1994 CD-ROM" of May 1995 (ICPSR CD0010). Data and documentation for 1996 were obtained by ftp from ICPSR in June 1997.

17. Alesina, Roubini, and Cohen (1997) further review "rational retrospective voting" models.

18. The district-level analysis is subject to ecological inference biases.

19. Estimates were computed using SAS, `PROC NLIN` (SAS Institute 1989–95), with numerical derivatives. Over all years, the percentage correctly classified by "predicting" for each observation the pair of vote choices that has the highest probability using the parameter MLEs is 73.6% (by year: 68.7%, 67.0%, 74.4%, 75.9%, 74.6%, 78.3%). Over all years the average probability of the

pair of choices actually made is .64 (by year: .58, .58, .65, .66, .65, .69).

20. Indeed, using the MLEs for each model, $\hat{\beta}_i = \hat{\beta}_{DRi} + \hat{\beta}_{DDi}$.

21. The estimates from the normal approximation method are numerically the closest to the more extreme values that Alesina, Roubini, and Cohen (1997, 114–20, 135–7) obtain by using the statistics of a Brownian motion (running in reverse) to model sequences of Gallup poll results. A key assumption of the Brownian motion model, that the differences between successive polls are independent, is incompatible with the idea that the sequence of polls informs voters about the aggregate distribution of voters' preferences, and therefore it is incompatible with convergence results such as McKelvey and Ordeshook (1985a) and Nielsen et al. (1990) discuss.

22. Except for 1976, the simple vote proportions in Table 5 have values somewhat more extreme than the actual Republican proportion of the two-party popular vote, reflecting the overreporting of votes for the winner that occurs in NES postelection data. In order by year, the actual proportions are .489, .553, .592, .539, .465, and .453 (Miller 1997).

23. With 1996 omitted, the correlation is .85. These correlations are for the beta approximation value of $\hat{\bar{P}}$. Results for the other methods are similar.

24. The 95% confidence interval for b_{HP}, computed as in Table 4, is (.9, 2.8).

25. It is important to keep clear that b_{HP} refers to local variations among equilibria. Numerical computations using the NES data and the parameter MLEs of Table 2 show that $\partial\hat{\bar{P}}_{\ell\jmath}/\partial\jmath < 0$ and $\partial\hat{\bar{H}}_{\ell\jmath}/\partial\ell < 0$ for $\ell, \jmath \in (0,1)$. Moreover, the values $\hat{\bar{P}}_{\ell\jmath} = \ell$ are decreasing in $\jmath$, and the values $\hat{\bar{H}}_{\ell\jmath} = \jmath$ are decreasing in ℓ for $\ell, \jmath \in (0,1)$ (see Appendix A for the $\ell, \jmath$ notation). The assumption of global linearity that $b_{HP} = d\bar{H}/d\bar{P}$ reflects is restrictive.

26. Overall, the coordinating voting model does not do especially well in classifying split-ticket voters separately from straight-ticket voters. Of the 597 voters "predicted" to split their tickets, 48% did so; of the 1,101 voters who split their tickets, only 26% were "predicted" to do so.

27. The values for β_i are computed using the MLEs for b_{E0} and b_{E1}.

28. For instance, suppose the decision rule of equation 5 is common knowledge, but no voter knows the distribution of other voters' preferences, that is, no one knows the probability measures f_k or the distribution F_{PH}. If the unknown distributions are such that $\partial \bar{P}_{\ell\jmath}/\partial\jmath < 0$ and $\partial \bar{H}_{\ell\jmath}/\partial\ell < 0$ for $\ell, \jmath \in (0,1)$, then a sequence of preelection polls will converge to a fixed point $(\bar{H}, \bar{P})$. Nielsen et al. (1990) discuss convergence for a scalar aggregate statistic. To achieve convergence to a neighborhood of $(\bar{H}, \bar{P})$, it is not necessary for every voter to learn the result of every poll, nor for every voter to have the same probability of learning about each poll. If under the common knowledge assumptions of the current model the fixed point $(\bar{H}, \bar{P})$ would be unique and stable for the given population of voters, and if each voter has a positive probability of learning about each poll and updates consistently, then Bowden's (1987) results imply that aggregate pre-election poll results evolve according to a Markov process that converges to a stationary distribution with mean $(\bar{H}, \bar{P})$. It seems obvious that if each voter uses Bayesian updating based on a suitable likelihood, then the mode of the distribution of each voter's beliefs converges over time to $(\bar{H}, \bar{P})$ (cf. McKelvey and Page 1990). The process may not converge by the time of any particular election, however.

29. For instance, at least as strong would be an assumption that all voters have a common knowledge prior distribution for the parameters, from which they then update based on public campaign information.

30. Confine $\alpha_D, \alpha_R \in [0,1)$. As discussed in the text, $\alpha_D = \alpha_R = 1$ is incompatible with coordination.

31. If the ideal point and policy position variables are continuous, then the set of discontinuities has measure zero.

32. In `PROC NLIN` I used `_WEIGHT_` to weight the Jacobian matrix (SAS Institute 1990, 1148).

References

Alesina, Alberto, John Londregan and Howard Rosenthal. 1993. "A Model of the Political Economy of the United States." *American Political Science Review* 87 (March): 12–33.

Alesina, Alberto, and Howard Rosenthal. 1989. "Partisan Cycles in Congressional Elections and the Macroeconomy." *American Political Science Review* 83 (June): 373–98.

Alesina, Alberto, and Howard Rosenthal. 1995. *Partisan Politics, Divided Government, and the Economy.* New York: Cambridge University Press.

Alesina, Alberto, and Howard Rosenthal. 1996. "A Theory of Divided Government." *Econometrica* 64 (November): 1311–41.

Alesina, Alberto, Nouriel Roubini and Gerald D. Cohen. 1997. *Political Cycles and the Macroeconomy.* Cambridge, MA: MIT Press.

Alvarez, R. Michael, and Matthew M. Schousen. 1993. "Policy Moderation of Conflicting Expectations: Testing the Intentional Models of Split-ticket Voting." *American Politics Quarterly* 21 (4): 410–38.

Aumann, Robert J. 1976. "Agreeing to Disagree." *Annals of Statistics* 4 (November): 1236–9.

Barndorff-Nielsen, O. E., and D. R. Cox. 1994. *Inference and Asymptotics.* London: Chapman and Hall.

Billingsley, Patrick. 1986. *Probability and Measure.* 2d ed. New York: Wiley.

Born, Richard. 1994a. "Split-ticket Voters, Divided Government and Fiorina's Policy-balancing Model." *Legislative Studies Quarterly* 19 (1): 95–115.

Born, Richard. 1994b. "Rejoinder." *Legislative Studies Quarterly* 19 (1): 126–9.

Börsch-Supan, Axel. 1990. "On the Compatibility of Nested Logit Models with Utility Maximization." *Journal of Econometrics* 43 (March): 373–88.

Bowden, Roger J. 1987. "Repeated Sampling in the Presence of Publication Effects." *Journal of the American Statistical Association* 82 (June): 476–84.

Brady, Henry, and Paul M. Sniderman. 1985. "Attitude Attribution: A Group Basis for Political Reasoning." *American Political Science Review* 79 (December): 1061–78.

Burden, Barry C., and David C. Kimball. 1998. "A New Approach to the Study of Ticket Splitting." *American Political Science Review* 92 (September): 533–44.

Calvert, Randall L., and John A. Ferejohn. 1983. "Coattail Voting in Recent Presidential Elections." *American Political Science Review* 77 (June): 407–19.

Campbell, Angus, and Warren E. Miller. 1957. "The Motivational Basis of Straight and Split Ticket Voting." *American Political Science Review* 51 (June): 293–312.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems.* New York: Cambridge University Press.

Dastoor, Naorayex K. 1985. "A Classical Approach to Cox's Test for Non-nested Hypotheses." *Journal of Econometrics* 27 (March): 363–70.

Enelow, James C., and Melvin J. Hinich. 1982. "Ideology, Issues, and the Spatial Theory of Elections." *American Political Science Review* 76 (September): 493–501.

Enelow, James C., and Melvin J. Hinich. 1983a. "Voting One Issue at a Time: The Question of Voter Forecasts." *American Political Science Review* 77 (June): 435–45.

Enelow, James C., and Melvin J. Hinich. 1983b. "Voter Expectations in Multi-Stage Voting Systems: An Equilibrium Result." *American Journal of Political Science* 27 (November): 820–7.

Ferejohn, John A., and Randall L. Calvert. 1983. "Presidential Coattails in Historical Perspective." *American Journal of Political Science* 28 (February): 127–46.

Fey, Mark. 1997. "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting." *American Political Science Review* 91 (March): 135–47.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections.* New Haven, CT: Yale University Press.

Fiorina, Morris P. 1988. "The Reagan Years: Turning to the Right or Groping Toward the Middle." In *The Resurgence of Conservatism in Anglo-American Democracies,* ed. Barry Cooper, Allan Kornberg, and William Mishler. Durham, NC: Duke University Press. Pp. 430–59.

Fiorina, Morris P. 1992. *Divided Government.* New York: Macmillan.

Fiorina, Morris P. 1994. "Response to Born." *Legislative Studies Quarterly* 19 (1): 117–25.

Franklin, Charles H. 1984. "Issue Preferences, Socialization, and the Evolution of Party Identification." *American Journal of Political Science* 28 (August): 459–78.

Franklin, Charles H., and John E. Jackson. 1983. "The Dynamics of Party Identification." *American Political Science Review* 77 (December): 957–73.

Frymer, Paul. 1994. "Ideological Consensus within Divided Government." *Political Science Quarterly* 109 (2): 287–311.

Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory.* Cambridge, MA: MIT Press.

Galambos, Janos. 1987. *The Asymptotic Theory of Extreme Order Statistics.* 2d ed. Malabar, FL: Robert E. Krieger.

Grossman, Sanford J. 1989. *The Informational Role of Prices.* Cambridge, MA: MIT Press.

Ingberman, Daniel, and John Villani. 1993. "An Institutional Theory of Divided Government and Party Polarization." *American Journal of Political Science* 37 (May): 429–71.

Isaki, Cary T., and Wayne A. Fuller. 1982. "Survey Design Under the Regression Superpopulation Model." *Journal of the American Statistical Association* 77 (March): 89–96.

Jackson, John E. 1975. "Issues, Party Choices and Presidential Votes." *American Journal of Political Science* 19 (May): 161–85.

Johnson, Norman L., Samuel Kotz and N. Balakrishnan. 1995. *Continuous Univariate Distributions, Volume 2.* 2d ed. New York: John Wiley & Sons.

Kreps, David M. 1977. "A Note on 'Fulfilled Expectations' Equilibria." *Journal of Economic Theory* 14 (February): 32–43.

Maddala, G. S. 1983. *Limited-dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press.

Markus, Gregory B. 1988. "The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis." *American Journal of Political Science* 32 (February): 137–54.

Markus, Gregory B., and Philip E. Converse. 1979. "A Dynamic Simultaneous Equation Model of Electoral Choice." *American Political Science Review* 73 (December): 1055–70.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models,* ed. Anders Karlqvist, Lars Lundqvist, Folke Snickars and Jörgen W. Weibull. New York: North-Holland. Pp. 75–96.

McKelvey, Richard D., and Peter C. Ordeshook. 1984. "Rational Expectations in Elections: Some Experimental Results Based on a Multidimensional Model." *Public Choice* 44 (1): 61–102.

McKelvey, Richard D., and Peter C. Ordeshook. 1985a. "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources." *Journal of Economic Theory* 36 (June): 55–85.

McKelvey, Richard D., and Peter C. Ordeshook. 1985b. "Sequential Elections with Limited Information." *American Journal of Political Science* 29 (August): 480–512.

McKelvey, Richard D., and Talbot Page. 1990. "Public and Private Information: An Experimental Study of Information Pooling." *Econometrica* 58 (November): 1321–39.

Mebane, Walter R. Jr. 1998. "Rational Expectations Coordinating Voting in American Presidential and House Elections." Paper presented at the 1998 Summer Methods Conference, University of California San Diego, July 23–26.

Miller, Alan William. 1997. *Atlas of United States Presidential Elections: 1932–1996.* Richmond, VA: Klipsan.

Miller, Warren E., and Arthur H. Miller. 1977. *The CPS 1976 American National Election Study* [computer file] (Study # 7381). Ann Arbor: Center for Political Studies, University of Michigan [original producer]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [producer and distributor].

Miller, Warren E., and the National Election Studies. 1982. *American National Election Study, 1980: Pre- and Post-Election Survey* [computer file] (Study # 7763). Ann Arbor: Center for Political Studies, University of Michigan [original producer]. 2d ICPSR ed. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [producer and distributor].

Miller, Warren E., and the National Election Studies. 1985. *American National Election Study, 1984: Continuous Monitoring Survey File* [computer file] (Study # 8298). Ann Arbor: Center for Political Studies, University of Michigan [original producer]. 1st ICPSR ed. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [producer and distributor].

Miller, Warren E., and the National Election Studies. 1986. *American National Election Study, 1984: Pre- and Post-Election Survey* [computer file] (Study # 8298). Ann Arbor: Center for Political Studies, University of Michigan [original producer]. 2d ICPSR ed. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [producer and distributor].

Miller, Warren E., and the National Election Studies. 1989. *American National Election Study, 1988: Pre- and Post-Election Survey* [computer file] (Study # 9196). Ann Arbor: Center for Political Studies, University of Michigan [original producer]. 2nd ICPSR ed. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [producer and distributor].

Miller, Warren E., Donald R. Kinder, Steven J. Rosenstone, and the National Election Studies. 1993. *American National Election Study, 1992: Pre- and Post-Election Survey* [enhanced with 1990 and 1991 data] [computer file] (Study # 6067). Conducted by University of Michigan, Center for Political Studies. Ann Arbor: University of Michigan, Center for Political Studies, and Inter-University Consortium for Political and Social Research [producers]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor].

Moran, P. A. P. 1971a. "The Uniform Consistency of Maximum-likelihood Estimators." *Proceed-*

ings of the Cambridge Philosophical Society 70 (3): 435–9.

Moran, P. A. P. 1971b. "Maximum-likelihood Estimation in Non-standard Conditions." *Proceedings of the Cambridge Philosophical Society* 70 (3): 441–50.

Nielsen, Lars Tyge, Adam Brandenburger, John Geanakoplos, Richard McKelvey, and Talbot Page. 1990. "Common Knowledge of an Aggregate of Expectations." *Econometrica* 58 (September): 1235–9.

Page, Benjamin I., and Calvin C. Jones. 1979. "Reciprocal Effects of Policy Preferences, Party Loyalties and the Vote." *American Political Science Review* 73 (December): 1071–89.

Palfrey, Thomas R. 1989. "A Mathematical Proof of Duverger's Law." In *Models of Strategic Choice in Politics*, ed. Peter C. Ordeshook. Ann Arbor: University of Michigan Press.

Palmer, John L., and Isabel V. Sawhill, eds. 1984. *The Reagan Record: An Assessment of America's Changing Domestic Priorities.* Cambridge, MA: Ballinger.

Resnick, Sidney I. 1987. *Extreme Values, Regular Variation, and Point Processes.* New York: Springer-Verlag.

Resnick, Sidney, and Rishin Roy. 1990. "Multivariate Extremal Processes, Leader Processes and Dynamic Choice Models." *Advances in Applied Probability* 22 (2): 309–31.

Rosenstone, Steven J., Donald R. Kinder, Warren E. Miller, and the National Election Studies. 1997. *American National Election Study, 1996: Pre- and Post-Election Survey* [computer file] (Study # 6896). 2nd release. Ann Arbor: University of Michigan, Center for Political Studies [producer]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor].

Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman. 1992. *Model Assisted Survey Sampling.* New York: Springer-Verlag.

SAS Institute. 1989–95. "SAS (r) PROC NLIN." Proprietary Software Release 6.11. Cary, NC: SAS Institute.

SAS Institute. 1990. *SAS/STAT User's Guide, Version 6, 4th ed.* Cary, NC: SAS Institute.

Self, Steven G., and Kung-Yee Liang. 1987. "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions." *Journal of the American Statistical Association* 82 (June): 605–10.

Table 1: Correspondences between Theoretical and Empirical Coordinating Vote Choice Models

Description	Theoretical Model Notation	Formal Definition	Empirical Model Notation	Formal Definition
Choice set	$K = \{RR, RD, DR, DD\}$	—	$K = \{RR, RD, DR, DD\}$	—
Voter i's choice	Y_i	equation 5	Y_i	data
Observable effects of choices on voter i's loss	x_{RRi}, x_{RDi}, x_{DRi}, x_{DDi}	equations 6a–6d	x_{RRi}, x_{RDi}, x_{DRi}, x_{DDi}	equations 6a–6d, 17a–17d
Probability that voter i makes each choice	μ_{RRi}, μ_{RDi}, μ_{DRi}, μ_{DDi}	equation 7	μ_{RRi}, μ_{RDi}, μ_{DRi}, μ_{DDi}	equation 7
Common knowledge expected House election outcome	$\bar{H}$	equation 9	$\hat{\bar{H}}$	equation 12
Common knowledge expected presidential election outcome	$\bar{P}$	equation 10	$\hat{\bar{P}}$	four methods
Fixed-point condition	$\bar{P}_{\ell\jmath} = \ell$, $\bar{H}_{\ell\jmath} = \jmath$	Appendix A	$(\hat{\bar{H}}^{(\imath)}, \hat{\bar{P}}^{(\imath)}) = (\hat{\bar{H}}^{(\imath+1)}, \hat{\bar{P}}^{(\imath+1)})$	Appendix A

Table 2: Parameter Estimates for the Coordinating and Noncoordinating Vote Choice Models

	Coordinating		Noncoordinating			Coordinating		Noncoordinating	
parameter	MLE	SE	MLE	SE	parameter	MLE	SE	MLE	SE
q	1.027	.068	1.065	.098	τ	.392	.041	.371	.047
b_P	3.916	.433	—	—	b_{E0}	.306	.111	—	—
b_{HP}	1.880	.492	—	—	b_{E1}	.478	.176	—	—
b_H	5.160	1.472	—	—	b_{RRE0}	—	—	−.294	.137
b_{PH}	.020	.080	—	—	b_{RRE1}	—	—	−.717	.225
b_{NC}	—	—	21.580	2.006	b_{RDE0}	—	—	−.096	.097
$\alpha_{D,76}$	1*	.150	.803	.082	b_{RDE1}	—	—	−.366	.171
$\alpha_{D,80}$	.405	.207	.693	.074	b_{DRE0}	—	—	.230	.084
$\alpha_{D,84}$	.774	.133	.883	.056	b_{DRE1}	—	—	−.193	.139
$\alpha_{D,88}$	.672	.158	.733	.066	$c_{P0,76}$	−1.054	.080	−1.128	.079
$\alpha_{D,92}$	.978	.091	.849	.045	$c_{P0,80}$	−.849	.101	−.883	.100
$\alpha_{D,96}$	.830	.083	1*	.047	$c_{P0,84}$	−.647	.077	−.756	.077
$\alpha_{R,76}$	.757	.195	.979	.088	$c_{P0,88}$	−.895	.086	−.906	.087
$\alpha_{R,80}$	.949	.095	1*	.063	$c_{P0,92}$	−.891	.102	−.975	.102
$\alpha_{R,84}$	.544	.111	.792	.056	$c_{P0,96}$	−1.140	.088	−1.270	.087
$\alpha_{R,88}$	.739	.009	.915	.071	$c_{H0,76}$	−.944	.084	−.993	.083
$\alpha_{R,92}$	.557	.151	.873	.055	$c_{H0,80}$	−.909	.105	−.930	.105
$\alpha_{R,96}$	.097	.257	.062	.120	$c_{H0,84}$	−.870	.081	−.953	.079
$\rho_{D,76}$	.991	.200	1*	.174	$c_{H0,88}$	−1.064	.089	−1.090	.084
$\rho_{D,80}$	.861	.201	.905	.199	$c_{H0,92}$	−.759	.092	−.841	.088
$\rho_{D,84}$	.968	.226	1*	.218	$c_{H0,96}$	−.854	.089	−.966	.092
$\rho_{D,88}$	.783	.208	.931	.233	$c_{P1,76}$	.220	.130	.209	.132
$\rho_{D,92}$	1*	.191	.890	.138	$c_{P1,80}$	.424	.129	.475	.129
$\rho_{D,96}$	.742	.159	.678	.196	$c_{P1,84}$	.479	.128	.518	.125
$\rho_{R,76}$	.648	.295	.749	.239	$c_{P1,88}$	.294	.157	.317	.160
$\rho_{R,80}$	1*	.193	1*	.168	$c_{P1,92}$	.346	.129	.355	.130
$\rho_{R,84}$	.779	.211	1*	.292	$c_{P1,96}$	.677	.150	.677	.151
$\rho_{R,88}$	.633	.258	.687	.259	$c_{H1,76}$	.221	.111	.114	.118
$\rho_{R,92}$	.558	.199	.861	.160	$c_{H1,80}$	.035	.111	.160	.113
$\rho_{R,96}$	1*	.217	.826	.314	$c_{H1,84}$	.008	.094	.066	.093
c_D	.902	.107	.980	.107	$c_{H1,88}$	.275	.123	.282	.125
c_{ID}	.774	.124	.836	.122	$c_{H1,92}$	.220	.095	.190	.098
c_I	1.922	.138	2.056	.137	$c_{H1,96}$	.325	.127	.404	.129
c_{IR}	2.820	.128	3.022	.126	c_{DEM}	1.015	.112	1.027	.113
c_R	2.884	.123	3.080	.121	c_{REP}	1.325	.121	1.328	.124
c_{SR}	3.542	.142	3.803	.138					

Note: Maximum likelihood estimates. Asterisk indicates a boundary-constrained parameter. Pooled NES survey data, 1976–96, $n = 4,859$. Log-likelihood values: coordinating model, -3186.1; noncoordinating model, -3225.1.

Table 3: Likelihood-ratio Test Statistics for the Constraint $\alpha_D = \alpha_R = 1$, by Year

Year	$-2(L_{\text{constrained}} - L)$
1976	1.8
1980	17.0
1984	25.4
1988	16.0
1992	12.8
1996	39.4

Note: The constraint is imposed separately for each year's α_D and α_R parameters. The critical value for a .05-level test using the χ^2_2 distribution is 5.99 (using χ^2_1, for 1976, the critical value is 3.84).

Table 4: 95% Confidence Intervals for α_D and α_R

Parameter	Lower Bound	Upper Bound	Parameter	Lower Bound	Upper Bound
$\alpha_{D,76}$	.79	1^*	$\alpha_{R,76}$	.38	1^*
$\alpha_{D,80}$	.03	.83	$\alpha_{R,80}$	.76	1^*
$\alpha_{D,84}$	.52	1^*	$\alpha_{R,84}$	.34	.79
$\alpha_{D,88}$	.36	.98	$\alpha_{R,88}$	.54	.99
$\alpha_{D,92}$	.80	1^*	$\alpha_{R,92}$	.27	.89
$\alpha_{D,96}$	.67	1^*	$\alpha_{R,96}$	0^*	.66

Note: Estimates are based on tabulation of an asymptotic mixture distribution of the kind derived in Moran (1971) and Self and Liang (1987), under the hypothesis $\alpha_{D,76} = \rho_{D,92} = \rho_{R,80} = \rho_{R,96} = 1$. Asterisk indicates a boundary-constrained value.

Table 5: Expected Probability that Republican President Wins ($\bar{P}$) and Proportion Republican in National House Vote ($\bar{H}$), for Alternative Estimates of $\bar{P}$, by Year

	Beta Approximation[a]		Normal Approximation[b]		Electoral Vote Proportion[c]		Simple Vote Proportion[d]	
Year	$\hat{\bar{P}}$	$\hat{\bar{H}}$	$\hat{\bar{P}}$	$\hat{\bar{H}}$	$\hat{\bar{P}}$	$\hat{\bar{H}}$	$\hat{\bar{P}}$	$\hat{\bar{H}}$
1976	.497	.465	.557	.465	.525	.466	.515	.466
1980	.618	.505	.821	.502	.643	.503	.613	.505
1984	.620	.473	.933	.470	.695	.468	.608	.473
1988	.540	.453	.740	.451	.628	.447	.553	.452
1992	.416	.431	.183	.431	.425	.431	.429	.431
1996	.430	.537	.317	.541	.431	.537	.448	.536
$\hat{L}_{\mathrm{C}}$	−3186.1		−3201.1		−3186.8		−3186.4	

Note: Computed using the parameter MLEs for each method of computing $\hat{\bar{P}}$ and 1976–96 NES data. MLEs for the beta approximation method appear in Table 2.

[a] Text method two.

[b] Text method three.

[c] Text method one.

[d] Text method four.

Table 6: Orderings of Ideal Points and Expected Party Policy Positions, by Year

	Ordering				
Year	$\theta_{Di} > \theta_i > \theta_{Ri}$	$\theta_{Ri} > \theta_i > \theta_{Di}$	$\theta_{Di}, \theta_{Ri} > \theta_i$	$\theta_i > \theta_{Di}, \theta_{Ri}$	$\theta_i = \theta_{Di} = \theta_{Ri}$
1976	2.9	25.9	17.1	53.1	1.0
1980	2.7	42.4	21.0	33.7	0.2
1984	1.8	50.5	22.8	24.9	0.0
1988	1.7	46.1	26.6	25.2	0.3
1992	2.9	39.7	29.2	27.5	0.8
1996	3.9	44.0	29.1	22.5	0.6

Note: Entries show the percentage of voters in each year who have the indicated ordering of ideal point and expected party policy positions. Computed using the parameter MLEs in Table 2 and 1976–96 NES data. Each observation is weighted by the sampling weight $1/\omega_i$.

Table 7: Vote Choices That Minimize the Effect on Expected Policy-related Loss, by Year

	$\text{argmin}_{ph\in K}(x_{phi} - z_{phi})$			
Year	RR	RD	DR	DD
1976	50.0	17.2	0.0	32.9
1980	60.3	0.8	6.0	32.9
1984	40.7	14.7	0.0	44.6
1988	57.6	1.3	8.7	32.4
1992	35.8	10.2	0.7	53.2
1996	45.5	5.6	0.1	48.8

Note: Entries show the percentage of voters in each year for whom the indicated vote choice minimizes $x_{phi} - z_{phi}$, $ph \in K$, which is the policy-related component of the effect the voter's choice has on the voter's expected loss. Computed using the parameter MLEs in Table 2 and 1976–96 NES data. Each observation is weighted by the sampling weight $1/\omega_i$. Voters with $\theta_i = \theta_{Di} = \theta_{Ri}$ are omitted.