

DISCUSSION PAPER SERIES

No. 3372

OPTIMAL CAMPAIGNING IN PRESIDENTIAL ELECTIONS: THE PROBABILITY OF BEING FLORIDA

David Strömberg

PUBLIC POLICY



Centre for **E**conomic **P**olicy **R**esearch

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP3372.asp

OPTIMAL CAMPAIGNING IN PRESIDENTIAL ELECTIONS: THE PROBABILITY OF BEING FLORIDA

David Strömberg, Institute for International Economic Studies (IIES),
Stockholm University and CEPR

Discussion Paper No. 3372
May 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: David Strömberg

May 2002

ABSTRACT

Optimal Campaigning in Presidential Elections: The Probability of Being Florida*

This Paper delivers a precise recommendation for how presidential candidates should allocate their resources to maximize the probability of gaining a majority in the Electoral College. A two-candidate, probabilistic-voting model reveals that more resources should be devoted to states which are likely to be decisive in the Electoral College and, at the same time, have very close state elections. The optimal strategies are empirically estimated using state-level opinion polls available in September of the election year. The model's recommended campaign strategies closely resemble those used in actual campaigns. The Paper also analyses how the allocation of resources would change under the alternative electoral rule of a direct national vote for president.

JEL Classification: C50, C72, D72, H50 and M37

Keywords: elections, political campaigns and public expenditure

David Strömberg
Institute for International
Economic Studies
Stockholm University
S-106 91 Stockholm
SWEDEN
Tel: (46 8) 16 43 76
Fax: (46 8) 16 14 43
Email: david.stromberg@iies.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=131482

* I thank Steven Brams, Steve Coate, Antonio Merlo, Torsten Persson, Gerard Roland, Tom Romer, Howard Rosenthal, Jim Snyder, Jörgen Weibull, and seminar participants at UC Berkeley, Cornell University, the CEPR/IMOP Conference in Hydra, New York University, Stanford University, University of Pennsylvania, Princeton University, and the Wallis Conference in Rochester. Previous versions have been circulated under the titles: 'The Lindbeck-Weibull model in the Federal US Structure,' and 'The Electoral College and Presidential Resource Allocations.'

Submitted 07 March 2002

1. Introduction

This paper explores how the Electoral College shapes incentives for presidential candidates to allocate resources across states. It does so by developing a probabilistic voting model of electoral competition under the US Electoral College system.¹ The model delivers a precise recommendation for how presidential candidates, trying to maximize the probability of gaining a majority in the Electoral College, should allocate their resources. This recommendation is fully characterized, both theoretically and empirically. The recommendations of the model are then compared to the actual presidential campaign visits across states during the 1988-2000 presidential elections, and to presidential campaign advertisements across media markets in the 2000 election. The actual allocation of these resources closely resembles the equilibrium allocation in the model. The paper finally analyses how the allocation of advertisements across media markets would change under an institutional reform, namely the transition to a direct national vote for president. The principles guiding the allocation are quite different under the two systems, and the incentives to favor certain markets are much stronger under the Electoral College than under the Direct Vote, causing a more unequal distribution of resources.

In an early game-theoretic analysis of the effects of the Electoral College system, Brams and Davis (1974) find that presidential candidates should allocate resources disproportionately in favor of large states. They use a model where votes are cast with equal probability for each candidate, and where the candidates maximize their expected number of electoral votes. Their result is disputed by Colantoni, Levesque and Ordeshook (1975) who instead argue that a proportional rule, modified to take into account the closeness of the state election, predicts actual campaign allocations better. In his model of two-party competition for legislative seats, Snyder (1989) allows parties to have advantages in certain districts and finds that equilibrium campaign allocations are higher in close districts. Further, if the goal of the parties is to maximize the probability of winning a majority of seats, then allocations are also higher in districts which are more likely to be pivotal. Finally, more resources will be spent in safe districts of the advantaged party than in the safe districts of the other party.

¹In this system, a direct vote election is held in each state and the winner of the vote is supposed to get all of that states electoral votes. Then all the electoral votes are counted, and the candidate who receives most votes wins the election. (The fact that Maine and Nebraska organize their presidential elections by congressional district is disregarded in this paper.)

Inspired by these results, Nagler and Leighley (1992) empirically investigate state-by-state campaign expenditures on non-network advertising in 1972 and find these to be higher in states with closer elections and more electoral votes. A related empirical literature studies the political influence on allocations of federal funds across states. Wright (1974) finds that federal spending between 1933 and 1940 was higher in states with higher “political productivity”, a measure depending on the electoral votes per capita, the variability in the vote share of the incumbent government in past elections, and the predicted closeness of the presidential elections. For a more recent contribution to this literature, see Wallis (1996).

A separate theoretical literature has analyzed the policy effects of plurality versus proportional representation election systems. For example, Persson and Tabellini (1999), and Lizzeri and Persico (2001), find that under plurality rule governments tend to overprovide redistributive spending because its benefits can be more easily targeted to voters than public goods.

The main contribution of this paper is that it develops a model that is empirically estimable and allows for explicit solutions. It therefore ties together, in a precise way, theoretical insights similar to those of Brams and Davis (1974), Snyder (1989), Persson and Tabellini (1999), and Lizzeri and Persico (2001), with empirical results on actual campaigns or distribution of federal funds, similar to those of Nagler and Leighley (1992), and Wright (1974). The model also extends the theory of the Electoral College. It allows for differences in preferences across states, it allows for vote outcome across states to be correlated, and it allows for explicit solutions. This yields new theoretical insights. The model also reveals a link between all of the above literature and the literature concerning “voting power”, that is, the probability that a vote is decisive in an election (Banzhaf (1968), Chamberlain and Rothschild (1988), Gelman and Katz (2001), Gelman, King and Boscardin (1998)), and Merrill (1978)). The equilibrium allocation of resources is found to be proportional to the “voting power” under the Electoral College system, but not under Direct Vote.

This is the first in a series of three papers. Here, I develop a theory of political redistribution under the Electoral College and test it on instruments where the presidential candidates have clear control and clear objectives. In Strömberg (2002a), I study an area where presidential control is less clear and objectives are more multi-faceted, but where the welfare effects are larger. That paper finds similar patterns of Electoral College effects on the allocation of federal civilian employment across states 1948-1996. In Strömberg (2002b), I study voter participation. That paper adds political competition under the Electoral College to the

model of Shachar and Nalebuff (1999). As these applications, and the discussion of the change to a Direct Vote system, show, the estimable probabilistic-voting model developed in this paper is very general and can be applied to a wide variety of electoral settings and questions.

Section 2 develops the theoretical model. It also estimates the probability distribution for election outcomes suggested by the model empirically, and uses these estimates to interpret the equilibrium. Section 3 confronts the models predictions with actual campaign efforts. Section 4 addresses the allocational effects of a change to a Direct Vote system. Finally, Section 5 discusses the results and concludes.

2. Model

Two presidential candidates, indexed by superscript R and D , try to maximize their expected probability of winning the election by selecting the number of days, d_s , to campaign in each state s , subject to the constraint

$$\sum_{s=1}^S d_s^J \leq I,$$

$J = R, D$. In each state s , there is an election. The candidate who receives a majority of the votes in that state gets all the e_s electoral votes of that state. After elections have been held in all states, the electoral votes are counted, and the candidate who gets more than half those votes wins the election.

There is a continuum of voters, each indexed by subscript i , a mass v_s of which live in state s . Campaigning in a state increases the popularity of the campaigning candidate among voters in that state, as captured by the increasing and concave function $u(d_s^J)$.² The voters also care about some fixed characteristics of the candidates, captured by parameters R_i, η_s , and η . The parameter R_i represents an individual-specific ideological preference in favor of candidate R , and η_s and η represent the general popularity of candidate R . The voters may vote for candidate R or candidate D , and voter i in state s will vote for D if

$$\Delta u_s = u(d_s^D) - u(d_s^R) \geq R_i + \eta_s + \eta. \quad (2.1)$$

²This paper does not address the question of why campaigning matters. This is an interesting question in its own right, with many similarities to the question of why advertisements affect consumer choice.

At the time when the campaign strategies are chosen, there is uncertainty about the popularity of the candidates on election day. This uncertainty is captured by the random variables η_s and η . The candidates know that the S state level popularity parameters, η_s , and the national popularity parameter, η , are independently drawn from cumulative distribution functions $G_s = N(0, \sigma_s^2)$, and $H = N(0, \sigma^2)$ respectively, but they do not know the realized values.

The distribution of voters' ideological preferences, R_i , within each state is $F_s = N(\mu_s, \sigma_{fs}^2)$, a normal distribution with mean μ_s and variance σ_{fs}^2 . The means of the states' ideological distributions may shift over time, but the variance is assumed to remain constant. The share of votes that candidate D receives in state s is

$$F_s(\Delta u_s - \eta_s - \eta).$$

This candidate wins the state if

$$F_s(\Delta u_s - \eta_s - \eta) \geq \frac{1}{2},$$

or, equivalently, if

$$\eta_s \leq \Delta u_s - \mu_s - \eta.$$

The probability of this event, conditional on the aggregate popularity η , and the campaign visits, d_s^D , and d_s^R , is

$$G_s(\Delta u_s - \mu_s - \eta). \tag{2.2}$$

Let e_s be the number of votes of state s in the Electoral College. Define stochastic variables, D_s , indicating whether D wins state s

$$\begin{aligned} D_s &= 1, \text{ with probability } G_s(\cdot), \\ D_s &= 0, \text{ with probability } 1 - G_s(\cdot). \end{aligned}$$

The probability that D wins the election is then

$$P^D(d^D, d^R, \eta) = \Pr \left[\sum_s D_s e_s > \frac{1}{2} \sum_s e_s \right]. \tag{2.3}$$

However, it is difficult to find strategies which maximizes the expectation of the above probability of winning. The reason is that it is a sum of the probabilities of all possible combinations of state election outcomes which would result in D

winning. The number of such combinations is of the order of 2^{50} , for each of the infinitely many realizations of η .

A way to cut this Gordian knot, and to get a simple analytical solution to this problem, is to assume that the candidates are considering their approximate probabilities of winning. Since the η_s are independent, so are the D_s . Therefore by the Central Limit Theorem of Liapounov,

$$\frac{\sum_s D_s e_s - \mu}{\sigma_E}$$

where

$$\mu = \mu(d^D, d^R, \eta) = \sum_s e_s G_s(\Delta u_s - \mu_s - \eta), \quad (2.4)$$

and

$$\sigma_E^2 = \sigma_E^2(d^D, d^R, \eta) = \sum_s e_s^2 G_s(\cdot)(1 - G_s(\cdot)), \quad (2.5)$$

is asymptotically distributed as a standard normal. The mean, μ , is the expected number of electoral votes. That is, the sum of the electoral votes of each state, multiplied by the probability of winning that state. The variance, σ_E^2 , is the sum of the variances of the state outcomes, which is the e_s^2 multiplied by the usual expression for the variance of a Bernoulli variable. Using the asymptotic distribution, the approximate probability of D winning the election is

$$\tilde{P}^D(d^D, d^R, \eta) = 1 - \Phi\left(\frac{\frac{1}{2}\sum_s e_s - \mu}{\sigma_E}\right).$$

The error made from using the approximate probability of winning is discussed in Section 2.2, and Appendix 6.8.

Candidate D maximizes the approximate probability of winning the election

$$\max_{d^D} P^D(d^D, d^R) = \max_{d^D} \int \tilde{P}^D(d^D, d^R, \eta) h(\eta) d\eta$$

subject to the constraint

$$\sum d_s^D = I.$$

Candidate R also maximizes his approximate probability of winning. This game has a unique, interior, pure-strategy equilibrium characterized by the proposition

below.³ In Section 2.1, the functions $u(d_s)$ are chosen to ensure interior equilibria. Non-interior equilibria are characterized in Appendix 6.2.

Proposition 1. *The unique pair of strategies for the candidates (d^D, d^R) that constitute a NE in the game of maximizing the expected probability of winning the election must satisfy $d^D = d^R = d^*$, and for all s and for some $\lambda > 0$*

$$Q_s u'(d_s^*) = \lambda, \quad (2.6)$$

where

$$Q_s = - \int \frac{\partial}{\partial \Delta u_s} \Phi \left(\frac{\frac{1}{2} \sum_s e_s - \mu}{\sigma_E} \right) h(\eta) d\eta.$$

Proof: See Appendix 6.1.

Proposition 1 says that presidential candidates who are trying to maximize their probability of winning the election should spend more time in states with high values of Q_s . This follows since $u'(d_s^*)$ is decreasing in d_s^* . The following pages will be devoted to exploring what Q_s represents and how to measure it.

First, note that Q_s consists of two additively separable parts:

$$Q_s = - \int \left(\frac{\partial \Phi(\cdot)}{\partial \mu} \frac{\partial \mu}{\partial \Delta u_s} + \frac{\partial \Phi(\cdot)}{\partial \sigma_E} \frac{\partial \sigma_E}{\partial \Delta u_s} \right) h(\eta) d\eta = Q_{s\mu} + Q_{s\sigma}. \quad (2.7)$$

One arises because the candidates have an incentive to influence the expected number of electoral votes won by D , that is the mean of the normal distribution. The other arises because the candidates have an incentive to influence the variance in the number of electoral votes. The empirical discussion will be organized to discuss each term separately.

A qualified guess is that Q_s is approximately the joint "likelihood" that a state is actually decisive in the Electoral College and, at the same time, has a tied election. I will call states who are ex post decisive in the Electoral College and have tied elections *decisive swing states*. In the 2000 election, Florida was a

³Note that the candidates have diametrically opposed preferences, in other words, this is a zerosum game. This implies that the equilibrium strategies are as if each candidate tried to minimize the maximum probability that the other candidate could get by allocating resources across states. The same equilibrium would result in a game where one candidate moved first, and then the other, taking the first candidate's strategy as given. The first candidate would minimize the maximum that the second player could attain. And the second player would maximize this probability of winning given the first candidate's strategy.

decisive swing state. In contrast, neither New Mexico nor Wyoming were *decisive swing states*. While New Mexico was a swing state with a very close election outcome, it was not decisive in the Electoral College since Bush would have won with or without the votes of New Mexico. While Wyoming was decisive in the Electoral College, since Gore would have won the election, had he won Wyoming, it was not a swing state.

The above guess is based on the fact that the probability of being a *decisive swing state* replaces Q_s in the equilibrium condition of the model without the Central Limit approximation.⁴ Further, Appendix 6.3 shows that $Q_{s\mu}$ and $Q_{s\sigma}$ approximately equals the first and second-order parts of a second-order Taylor-expansion of the approximate probability of being a *decisive swing state*. In the empirical section, the values of the analytical expression for Q_s will be compared with the probability that a state is decisive in the Electoral College and, at the same time has a state margin of victory of less than two percent. To measure this probability, I now estimates the probability distribution for election outcomes.

2.1. Estimation

In equilibrium, both candidates choose the same allocation, so that $\Delta u_s = 0$ in all states. The Democratic vote-share in state s at time t equals

$$y_{st} = F_{st}(-\eta_{st} - \eta_t) = \Phi\left(\frac{-\mu_{st} - \eta_{st} - \eta_t}{\sigma_{fs}}\right),$$

where $\Phi(\cdot)$ is the standard normal distribution, or equivalently,

$$\Phi^{-1}(y_{st}) = \gamma_{st} = -\frac{1}{\sigma_{fs}}(\mu_{st} + \eta_{st} + \eta_t). \quad (2.8)$$

For now, assume that all states have the same variance of preferences, $\sigma_{fs}^2 = 1$, and the same variance in state-specific shocks, σ_s^2 .⁵ Further assume that the mean of the preference distribution, μ_{st} , depends on a set of variables X_{st} , so that the estimated equation is

$$\gamma_{st} = -(\beta X_{st} + \eta_{st} + \eta_t). \quad (2.9)$$

⁴Unfortunately, I can not compute the analytical solution of that model.

⁵These assumptions will be removed in Section 4. However, the estimates become imprecise if separate values of μ_{st} , σ_{fs} , and σ_s are estimated for each state using only 14 observations per state. Therefore, the more restrictive specification will be used for most of the paper.

The parameters β , σ_s and σ are estimated using a standard maximum-likelihood estimation of the above random-effects model.⁶

The variables in X_{st} are basically those used in Campbell (1992). The national variables are: the Democratic vote share of the two-party vote share in trial-heat polls from mid September (all vote-share variables x are transformed by $\Phi^{-1}(x)$); second quarter economic growth; incumbency; and incumbent president running for re-election. The state variables for 1948-1984 are: lagged and twice lagged difference from the national mean of the Democratic two-party vote share; the first quarter state economic growth; the average ADA-scores of each state's Congress members the year before the election; the Democratic vote-share of the two-party vote in the midterm state legislative election; the home state of the president; the home state of the vice president; and dummy variables described in Campbell (1992). After 1984, state-level opinion-polls were available. For this period, the state-level variables are: lagged difference from the national mean of the Democratic vote share of the two-party vote share; the average ADA-scores of each state's Congress members the year before the election; and the difference between the state and national polls. The other state-level variables were insignificant when state polls were included. The coefficients β and the variance of the state level popularity shocks, σ_s^2 , are allowed to differ for when opinion polls were available and when they were not. The equation yields forecasts by mid September of the election year. The data-set contains state elections for the 50 states 1948-2000, except Hawaii and Alaska which began voting in the 1960 election. During this period there were a total of 694 state-level presidential election results. Of this total, 13 state elections were excluded, leaving a total of 681 observations. Four elections in Alaska and Hawaii were excluded because there were no lagged vote returns. Nine elections are omitted because of idiosyncrasies in Presidential voting in Alabama in 1948, and 1964, and in Mississippi in 1960; see Campbell (1992).

⁶The model was also extended to include regional swings. In this specification, the election result in one state equals

$$y_{st} = F_{st} (\eta_{st} + \eta_{rt} + \eta_t),$$

where η_{rt} denotes independent popularity shocks in the Northeast, Midwest, West, and South. The estimated variances of the state and national level shocks are similar to those estimated without allowing for regional shocks, $\sigma_{s,post1984} = 0.084$, and $\sigma = 0.038$, see Appendix 6.4. The standard deviation of the regional shock is $\sigma_r = 0.054$ before state level forecasts were available in 1988. However, after 1988, the standard deviation of the regional shocks is zero. Taking into account the information of september state-level opinion polls, there are no significant regional swings. Therefore, the simpler specification without regional swings is used below.

The estimation results are shown in Table 1. The standard deviation of the state level shocks after 1984, σ_s , equals 0.077, or about 3% in vote shares. This is more than twice as large as that of the national shocks, $\sigma = 0.033$. The average error in state election vote forecasts is 3.0 percent and the wrong winner is predicted in 14 percent of the state elections. This is comparable to the best state-level election-forecast models (Campbell, 1992; Gelman and King, 1993; Holbrook and DeSart, 1999; Rosenstone, 1983).

2.2. Characterization of equilibrium

Next, I test whether Q_s approximately equals the probability of being a *decisive swing state*. To this end, one million electoral vote outcomes were simulated for each election 1988-2000 by using the estimated state-means, and drawing state and national popularity-shocks from their estimated distributions. Then, the share of elections where a state was decisive in the Electoral College and at the same time had a state election outcome between 49 and 51 percent was recorded. This provides an estimate which should be roughly equal to Q_s . Figure 2.1, contains these shares on the y -axis and values computed from the analytic expression of Q_s , on the x -axis. Large states are trivially more likely to be decisive. To check that the correlation between Q_s and the simulated values is not just a matter of size, the graph on the right contains the same series divided by the state's number of electoral votes. The simple correlation in the diagram to the right is 0.997. So the two variables are interchangeable, for practical purposes. The 0.003 difference could result on the Q_s -side from using the approximate probability of winning the election, and on the simulation-side from using a finite number of simulations and recording state election results between 49 and 51 percent, whereas theoretically it should be exactly 50 percent.

To illustrate the discussion of how Q_s varies across states, I will use the year 2000 election, see Figure 2.2. Based on polls available in mid September, 2000, Florida, Michigan, Pennsylvania, California, and Ohio were the states most likely to be decisive in the Electoral College and at the same time have a state election margin of less than 2 percent. This happened in 2.2 to 3.4 percent of the simulations.

The analytic expression for Q_s explains exactly why some states are more likely to be *decisive swing states*. First, Q_s is roughly proportional to the number of electoral votes. The reason is that the change in the expected number of

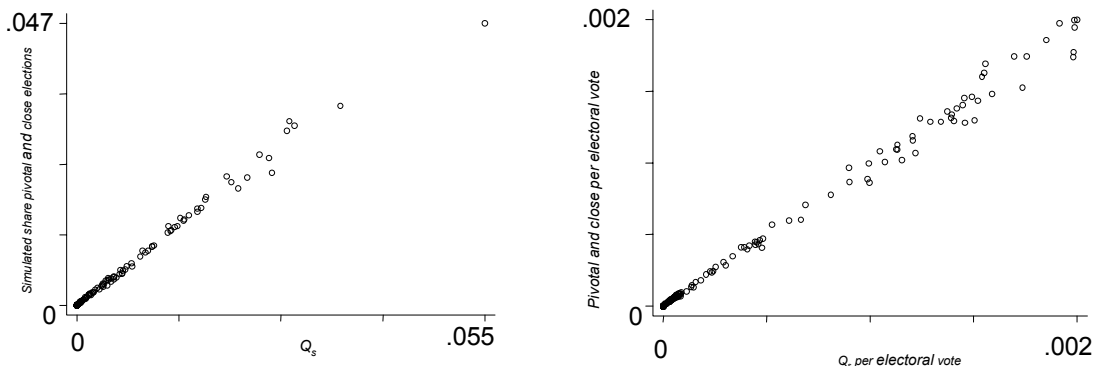


Figure 2.1: Q_s and simulated probability of being a *decisive swing state*

electoral votes in response to an extra candidate visit to a state is proportional to the number of electoral votes of that state. Therefore, so is $Q_{s\mu}$. The change in the variance, and therefore $Q_{s\sigma}$, is proportional to the state's electoral votes squared. As $Q_{s\sigma}$ is generally considerably smaller than $Q_{s\mu}$, Q_s is roughly proportional to the number of electoral votes. This implies that candidates should, on average, spend more time in large states. However, for states of equal size there is considerable variation.⁷

This can be seen in Figure (2.3). The x -axis shows the forecasted Democratic vote share. The circular dots show the share of the simulated elections where a state was decisive in the Electoral College and at the same time had a state-election outcome between 49 and 51 percent, per electoral vote. The solid normal-form line shows $Q_{s\mu}/e_s$ which arises because the candidates try to affect the expected number of electoral votes, see equation (2.7). This part of Q_s/e_s accounts for most of the variation in the simulated values. It explains why states like New York and Texas are never in a million simulated elections decisive in the Electoral College and at the same time have close state elections, while in Florida, Michigan, Pennsylvania, and Ohio this happens quite frequently. The solid line is in fact a normal distribution, multiplied by a constant. It is characterized by three features: its amplitude, its mean, and its variance.

⁷This can be contrasted to the finding that voters in larger states should receive more than proportional attention (Banzaf 1967, Brams and Davis 1974, Gelman and Katz 2001). Their results depend on all voters being equally likely to vote for one candidate or the other. My results differ since voters are not equally likely to vote for each candidate, and since there are aggregate popularity shocks, see Chamberlain and Rothschild (1981).

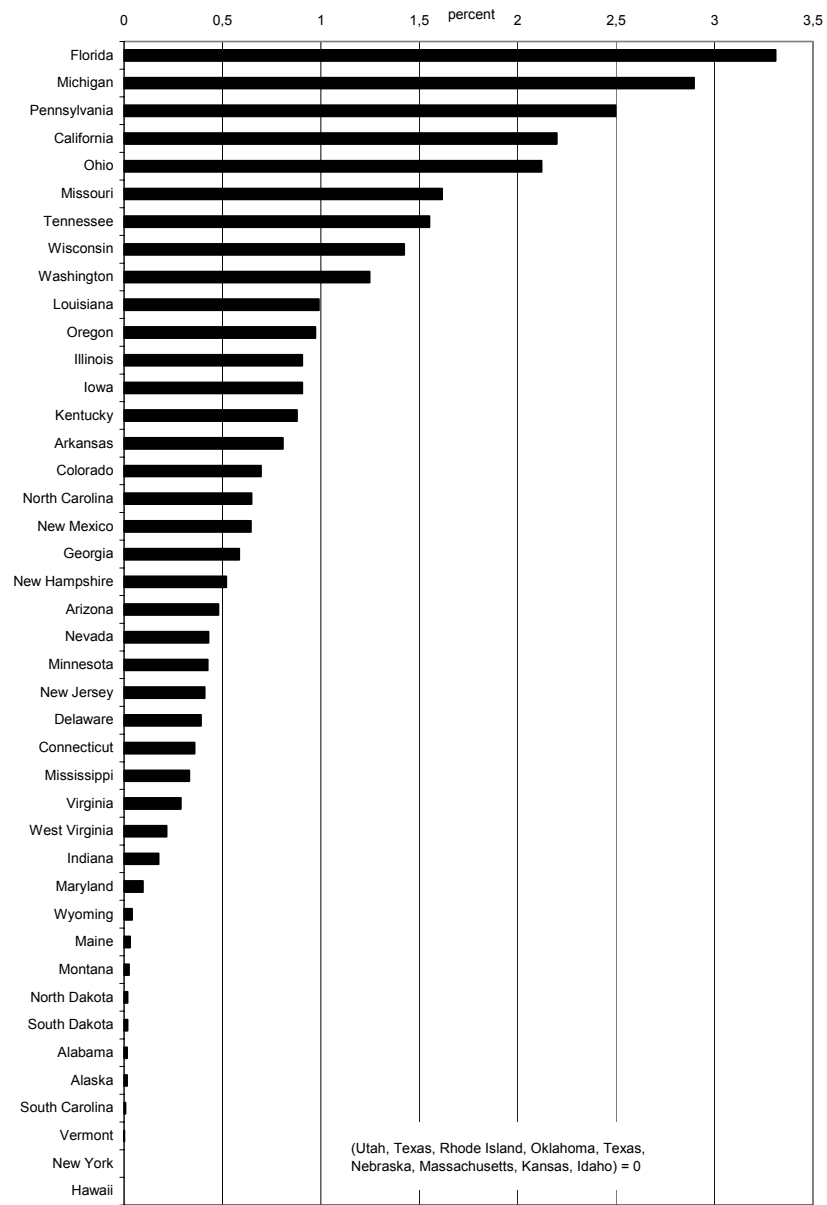


Figure 2.2: Joint probability of being pivotal and having a state margin of victory less than two percent, based on September 2000 opinion polls.

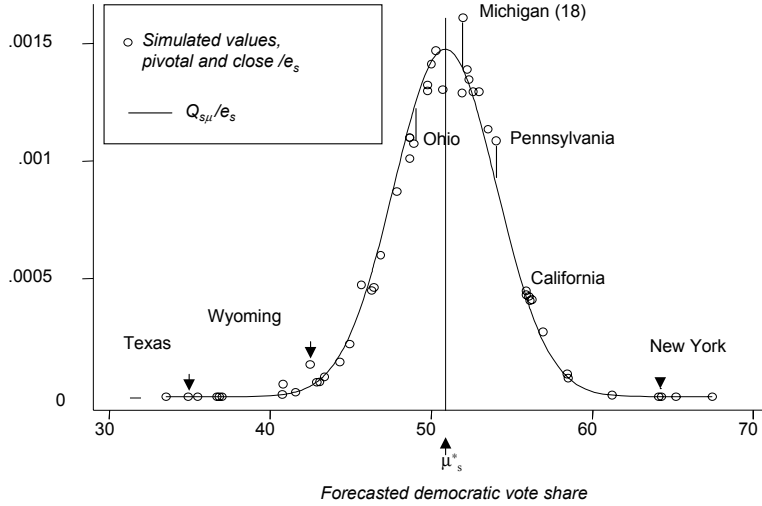


Figure 2.3: Probability of being a *decisive swing state* per electoral vote.

The amplitude of all $Q_{s\mu}$ is trivially higher when the national election is expected to be close. Define $\tilde{\eta}_t$ to be the national popularity-swing which would give equal expected Electoral Vote shares, $\mu(\tilde{\eta}) = \frac{1}{2} \sum_s e_s$. Then $Q_{s\mu}$ is larger when $\tilde{\eta}_t$ is close to zero, that is, when the national election is expected to be close.⁸ The value of $\tilde{\eta}_t$ affects all states in a single election in the same way. It explains why the average $Q_{s\mu}$ varies between elections.

Notice that the mean is located slightly above 50%. Since we have an analytic expression for Q_s , it is possible to say exactly why this is the case. The mean equals

$$\mu_{st}^* = -\frac{\sigma^2}{\sigma^2 + (\sigma_E/a)^2} \tilde{\eta}_t, \quad (2.10)$$

where

$$a_t = \sum_s e_s g_s(-\mu_{st} - \tilde{\eta}_t).$$

The mean always lies between a pro-Republican state bias of $\mu_{st} = 0$, which corresponds to a 50% forecasted Democratic vote share, and $\mu_{st} = -\tilde{\eta}_t$, which approximately corresponds to the forecasted national Democratic vote share. (If the Democrats are ahead by 60-40 nationally, then a pro-Republican swing $\tilde{\eta}_t$,

⁸See equation (6.4) in the Appendix.

corresponding to about 10%, is needed to draw the election. Therefore $\mu_{st} = -\tilde{\eta}_t$ corresponds to 10% pro-democrat bias in a state, that is, a vote share of 60-40.)

The intuition is the following. Suppose that the Democrats are ahead 60-40 in the national polls, 50-50 in Texas, and 60-40 in Pennsylvania. A candidate visit may only affect a state outcome in swing states, where the state election is close. Candidate visits are therefore more likely to influence the outcome in forecasted swing states like Texas, than in states like Pennsylvania. For this reason, candidates should target states like Texas.

However, the candidates must condition their visit strategies on what must be true for a state to be a swing state on election day. If Texas is still a swing state on election day, then the Democrats are probably winning by a landslide and Texas will not be decisive. If Pennsylvania is a swing state, then it is likely that the election at the national level close and Pennsylvania decisive. For this reason candidates should target states like Pennsylvania. The logic resembles that of the *winners curse* in auction theory. There the size of the bid only matter when the bid is highest, and the bidders must condition their bid on the circumstance in which it matters. Here the visit only matters if the state is a swing state, and the candidates must condition their visits on this circumstance.

The model shows how to strike a balance between high average influence (Texas) and influence when it matters (Pennsylvania). Basically, the less correlated the state election outcomes are, the more time should be spent in 50-50 states like Texas. This is evident from equation (2.10). The smaller the variance of the national popularity-swings, σ^2 , the more important it is to target states with expected outcomes close to 50-50. In the extreme case where this variance equals zero, then $\mu_s^* = 0$ and most time should be spent in states like Texas. The reason is that without national swings, the state outcomes are not correlated, and Texas being a swing state on election day carries no information about the outcomes in the other states. (The winner's curse does not arise in auctions with independent private values.) In the extreme case that σ approaches infinity, μ_s^* approaches $-\tilde{\eta}$. Therefore most time should be spent in states like Pennsylvania with a 60-40 expected outcome. In my estimates maximum attention should typically be given to states in the middle, 55-45 in this example. In September of 2000, Gore was ahead by 1.3 percentage points. The maximum $Q_{s\mu}/e_s$ was obtained for states where the expected outcome was a Democratic vote share of 50.8 percent, as illustrated in Figure 2.3.

People who are familiar with the market CAPM model may prefer the following analogy. Assets trivially attract more investment if they yield higher returns on

average (like Texas), but also if they yield higher return in recessions when returns are more valuable (like Pennsylvania). The larger the aggregate shocks (national popularity-swings), causing deep recessions and high booms, the more important it is for assets to yield high return in recessions.

Although the normal-shaped curve in Figure 2.3 explains most of the variation in Q_s/e_s , there are some noteworthy discrepancies. First, Wyoming and two other states to the left of the center are noticeably above the normal-shaped curve. The reason is that I could not find state-level opinion poll data for these states, and the forecasts for these states are more uncertain. These states actually lie on a normal-shaped curve with a higher variance than that drawn in Figure 2.3.⁹ These observations illustrate one effect of improved forecasting on the allocation of resources. Better state-level forecasts lead to a more unequal allocation of campaign resources as the variance of the normal-shaped distribution of Figure 2.3 decreases. States with forecasted vote shares close to the center of that distribution would gain while states far from the center would lose. Better national-level forecasts has a similar effect.

In Figure 2.3, note also that around its peak, the normal-shaped curve is far from the simulated probabilities of being a *decisive swing state* per electoral vote. States to the right of μ_s^* , like Michigan and Pennsylvania, generally lie above the curve, while states on the left, like Ohio, generally lie below. The difference between the simulated values and $Q_{s\mu}/e_s$ arises because the candidates also have incentives to influence the variance of the electoral vote distribution, even if this means decreasing the expected number of electoral votes, see $Q_{s\sigma}$ in equation (2.7).

To get the intuition of why this is rational, consider the following example from the world of ice-hockey. One team is trailing by one goal and there is only one minute left of the game. To increase the probability of scoring an equalizer, the trailing team pulls out the goalie and puts in an extra offensive player. Most frequently, the result is that the leading team scores. But the trailing team does not care about this, since they are losing the game anyway. They only care about increasing the probability that they score an equalizing goal, which is higher with

⁹The variance of the normal-form distribution,

$$\tilde{\sigma}^2 = \sigma_s^2 + \frac{1}{\left(\frac{1}{\sigma^2} + \frac{1}{(\sigma_E/a)^2}\right)},$$

depends on the variance in the state, and national, level popularity shocks.

an extra offensive player. Therefore, it is better to increase the variance in goals, even though this decreases net expected goals.

Similarly, presidential candidates who are behind should try to increase variance in electoral votes. This is done by spending more time in large states where this candidate is behind (putting in an extra offensive player). This is compensated by fewer visits to states where this candidate is ahead (pulling the goalie). Candidates who are ahead should try to decrease variance in electoral votes, thus securing their lead. This is done by spending more time in large states where this candidate is ahead, and reducing the number of visits where this candidate is behind. This leads both candidates to spend more time in large states where the expected winner is leading. This resounds the result by Snyder (1989) that parties will spend more in safe districts of the advantaged party than in safe districts of disadvantaged party.

To formally see why a trailing candidate increases the variance by spending more time in states with many electoral votes where he is behind, consider equation (2.5) showing the variance, conditional on a national shock. The variance in the number of electoral votes from a state is proportional to these votes squared. Therefore, the effect on the total variance is larger in large states. Further, the variance in a state outcome is higher the closer the expected result is to a tie. By visiting a state where the leading candidate is ahead, the trailing candidate moves the expected result closer to a tie, and increases the variance in election outcome. Similarly, decreasing the number of visits to a state where the lagging candidate is leading increases the variance

Figure 2.4 illustrates this effect in the year 2000 election. It plots the values of the analytical expression for $Q_{s\sigma}/e_s$. The lagging candidate (Bush) should put in extra offensive visits in states like Michigan and Pennsylvania, at the cost of weakening the defense of states like Ohio. The leading candidate (Gore) should increase his defense of states like Michigan and Pennsylvania, at the cost of offensive visits to Ohio.

3. Relation between Q_s and actual campaigns

This section will compare the equilibrium campaign strategies to actual campaign strategies. The first sub-section will investigate presidential candidate visits to states in the 2000 election, and also more loosely discuss visits during the 1988-1996 elections. The second sub-section will study the allocation of campaign advertisements across media markets in the 2000 campaign. Finally, the last sub-

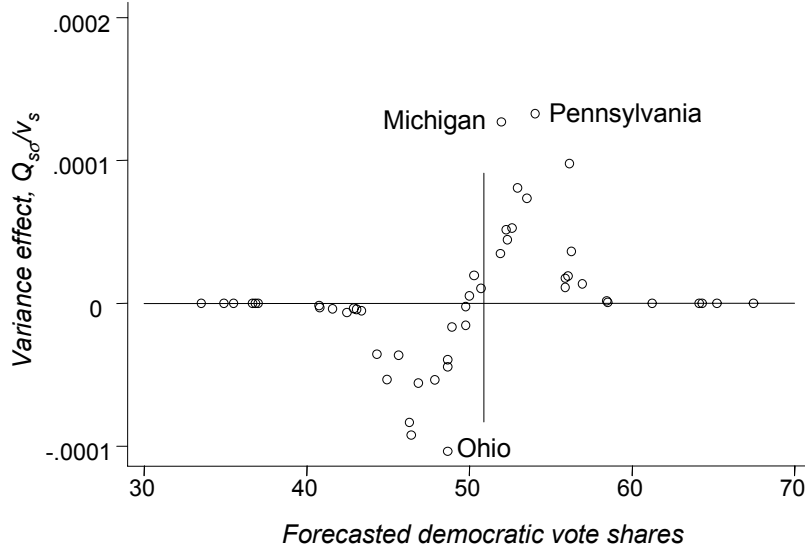


Figure 2.4: Incentive to influence variance

section estimates the impact of the actual campaigns on the election results.

3.1. Campaign visits

If one assumes log utility, then the optimal allocation, based on equation (2.6) is,

$$\frac{d_s^*}{\sum d_s^*} = \frac{Q_s}{\sum Q_s}, \quad (3.1)$$

and the number of days spent in each state should be proportional to Q_s .

The Bush and Gore campaigns were very similar to the equilibrium campaign based on September opinion polls. The actual number of year 2000 campaign visits, after the party conventions, and Q_s , are shown in Figure 3.1.¹⁰ Campaign visits by vice presidential candidates are coded as 0.5 visits. The model and the candidates' actual campaigns agree on 8 of the 10 states which should receive most attention. Notable differences between theory and practice are found in Iowa, Illinois and Maine, which received more campaign visits than predicted, and Colorado, which received less. Perhaps extra attention was devoted to Maine since

¹⁰I am grateful to Daron Shaw for providing me with the campaign data.

its (and Nebraska's) electoral votes are split according to district vote outcomes. Other differences could be because the campaigns had access to information of later date than mid September, and because aspects not dealt with in this paper matter for the allocation. The raw correlation between campaign visits and Q_s is 0.91. For Republican visits the correlation is 0.90 and for Democratic visits, 0.88. A tougher comparison is that of campaign visits per electoral vote, d_s/e_s , with Q_s/e_s . The correlation between d_s/e_s and Q_s/e_s was 0.81 in 2000.

Next, I look at the 1996, 1992, and 1988 campaigns. For these campaigns, only presidential visits are available. The correlation between visits and Q_s during those years are: 0.85, 0.64, and 0.76 respectively. But this is mainly a result of presidential candidates spending more time in large states. For the 1996, 1992, and 1998 elections, the correlation between d_s/e_s and Q_s/e_s was 0.12, 0.58, and 0.25 respectively. An explanation for the poor fit in 1996 and 1988 may be that these elections were, *ex ante*, very uneven. The expected Democratic vote shares in September of 1996, 1992, and 1988 were 56, 50, and 46 percent. In uneven races, perhaps the candidates have other concerns than maximizing the probability of winning the election.

A possible explanation for the difference between the actual and optimal campaigns is that presidential candidate visits target media markets instead of states. In Appendix 6.6, this situation is modelled. The main new feature is that there are spillovers across states as media markets cross state boundaries. This increases the number of visits seen in New York and Massachusetts. The presidential candidates choose to visit the media markets in New York and Boston because they cross into states which are important for re-election concerns. However, this does not explain why Iowa, Illinois and Maine received more visits, or why Colorado received less than expected. Instead, as the candidates did not visit New York and Massachusetts, this decreases the correlation between the actual and optimal visits.

A complication is that candidates should consider in what media markets their visit will be reported, rather than what markets they visit. A presidential candidate visit to L.A. may be covered also in surrounding Californian media markets. A more direct way to study targeting of media markets is to examine in which media markets the campaigns choose to air their advertisements.

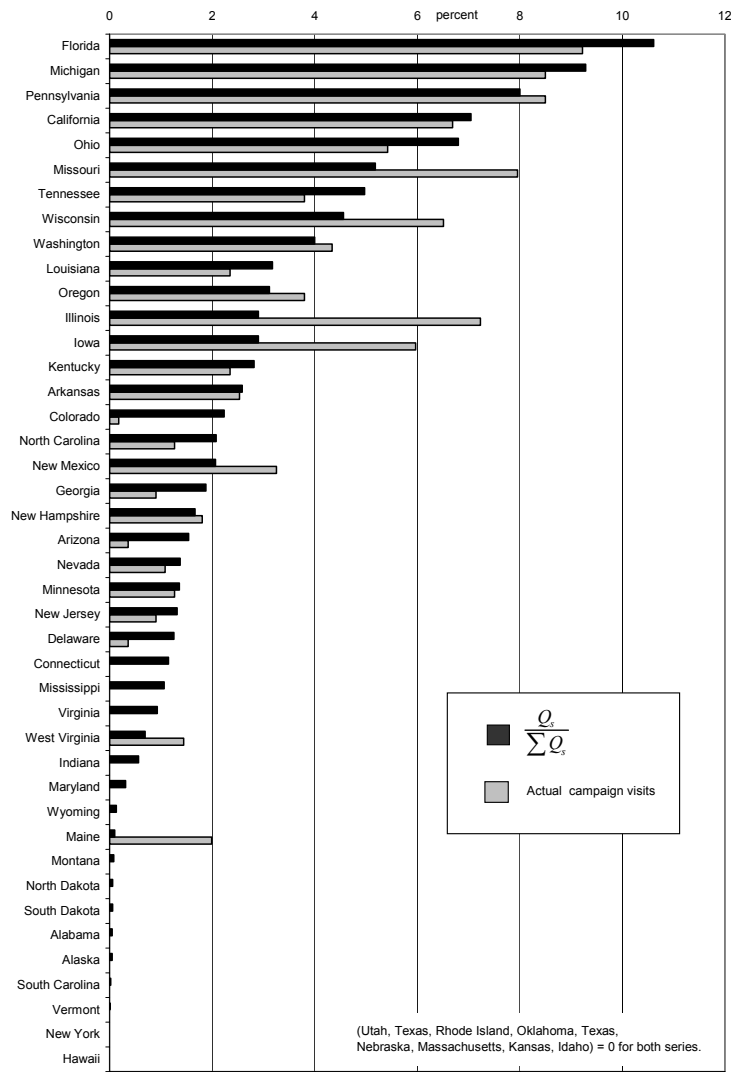


Figure 3.1: Actual and equilibrium campaign visits 2000

3.2. Campaign advertisements

Appendix 6.5 models the decision of presidential candidates to allocate advertisements across Designated Market Areas (DMAs).¹¹ In that model, two presidential candidates have a fixed advertising budget I to spend on a_m ads in each media market m subject to

$$\sum_{m=1}^M p_m a_m^J \leq I,$$

$J = R, D$, where p_m is the price of an advertisement. Media market m contains a mass v_{ms} voters in state s . Voters are affected by campaign advertisements as captured by the increasing and concave function $w(a_m)$. A voter i in media market m in state s will vote for D if

$$\Delta w_m = w(a_m^D) - w(a_m^R) \geq R_i + \eta_s + \eta.$$

In equilibrium both candidates choose the same advertising strategy. Advertising in media market m is increasing in $\frac{Q_m}{p_m}$, where

$$Q_m = \sum_{s=1}^S Q_s \frac{n_{ms}}{n_s}.$$

Q_m is the sum of the Q_s of the states in the media market weighted by the share of the population of state s that lives in media market m .

The advertisement data is from the 2000 election and was provided by the Brennan Center.¹² It contains the number and cost of all advertisements relating to the presidential election, aired in the 75 major media markets between September 1 and Election Day. The data is disaggregated by whether it supported the Republican, Democrat, or independent candidate, and by whether it was paid for by the candidate, the party or an independent group. The cost estimates, p_m , are average prices per unit charged in each particular media market. The estimates are done by the Campaign Media Analysis Group. Advertisements were only aired in 71 markets. Therefore there are only cost estimates in these 71 markets. The

¹¹A DMA is defined by Nielsen Media Research as all counties whose largest viewing share is given to stations of that same market area. Non-overlapping DMAs cover the entire continental United States, Hawaii and parts of Alaska.

¹²The Brennan Center began compiling this type of data for the 1998 elections. According to them, no such data exists elsewhere for any other election. This is a new and unique database.

data set recorded a total of 174 851 advertisements, for a total cost of \$118 million, making an average price of \$680. The Democrats spent \$51 million, while Republicans spent \$67 million. To measure total campaign efforts, I sum together the advertisements by the candidates, the parties and independent groups supporting the Democratic or Republican candidate.

The model and the data agree on the two media markets where most ads should be aired (Albuquerque - Santa Fe, and Portland, Oregon); see Figure 3.2. These two markets has the highest effect on the win probability per advertising dollar. In third place the model puts, Orlando - Daytona Beach - Melbourne, while the data has Detroit (number four in the model). The correlation between actual campaign advertisement and equilibrium advertisement is 0.75. That few advertisements were aired in Denver is consistent with the few candidate visits to Colorado, see Figure 3.1. The few advertisements in Lexington are more surprising, since candidate visits to Kentucky were close to the equilibrium number.

To see why Albuquerque - Santa Fe gives a large effect per advertising dollar, note that

$$\frac{Q_m}{p_m} = \sum_s \underbrace{\frac{Q_s}{e_s}}_{(i)} \underbrace{\frac{e_s}{n_s}}_{(ii)} \underbrace{\frac{1}{p_m/n_m}}_{(iii)} \frac{n_{ms}}{n_m}. \quad (3.2)$$

Albuquerque - Santa Fe covers a population of 1.4 million in New Mexico ($\frac{n_{ms}}{n_m} = 0.95$), and 70 000 in Colorado ($\frac{n_{ms}}{n_m} = 0.05$). (i) Since New Mexico has a forecasted Democratic vote-share of 51.8%, it has a very high value of Q_s per electoral vote, see Figure 2.3. (ii) Since New Mexico is a small state with only 1.8 million inhabitants, it has a high number (2.7) of electoral votes per capita. (iii) At the same time, the average cost of an ad per million inhabitants in the media market is only \$209, compared to the average media-market cost, which is \$270. In comparison, the Detroit media market lies entirely in Michigan which has the highest value of Q_s per electoral vote. However, being a fairly large state, Michigan only has 1.8 electoral votes per million inhabitants. Further, the average cost of an ad in Detroit is \$239. Therefore the, the marginal impact on the probability of winning per dollar is lower than in Albuquerque - Santa Fe.

To see whether the actual advertisements responded independently to changes in price and Q_m , assume that $u(a_m) = \ln(a_m)$. Then

$$\ln(a_m^*) = c + \ln(Q_m) - \ln(p_m). \quad (3.3)$$

For the 53 media markets where advertisements supporting the Democratic or

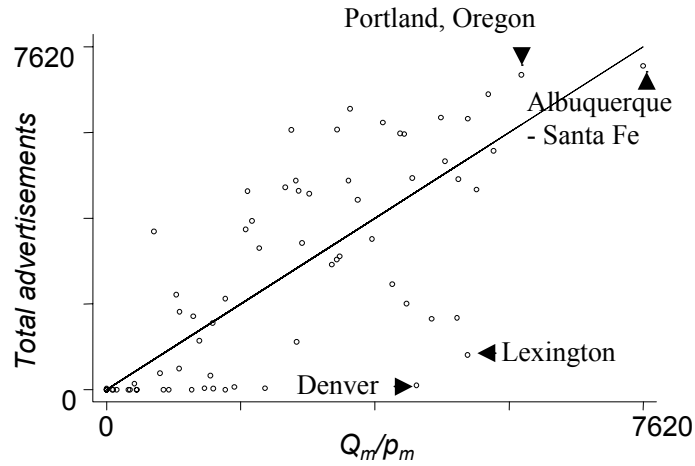


Figure 3.2: Total number of advertisements Sept. 1 to election day and Q_m/p_m , for the 75 largest media markets

Republican candidates were aired:

$$\ln(a_m) = 13.37 + 1.28 \ln(Q_m) - .94 \ln(p_m).$$

(2.21) (.32) (.35)

The candidates were responsive, both to changes in Q_m and p_m , and the elasticities are both close to one. Finally, one can note that since the correlation between price and market size is close to one (0.92), there is no clear relationship between market size and the number of ads ($\text{corr}(d_m, n_m) = -0.09$).

Via the price, the size is instead captured in the costs. Assuming log utility, equilibrium expenditures, $p_m a_m^*$, are proportional to Q_m . Empirically, the simple correlation between advertisement costs, $p_m a_m$, and Q_m is 0.88. Figure 3.3 plots equilibrium and actual advertising costs by market.

3.3. Estimating the effect of campaign visits on election outcomes

To complete the description of optimal strategies, the decreasing marginal impact of campaign visits and advertisements should be estimated. This has been relegated to this last section since the estimation is not fully consistent with theory, and because this estimation is rather imprecise. If Democrats and Republicans allocate campaign visits according to this theory, and have the same information,

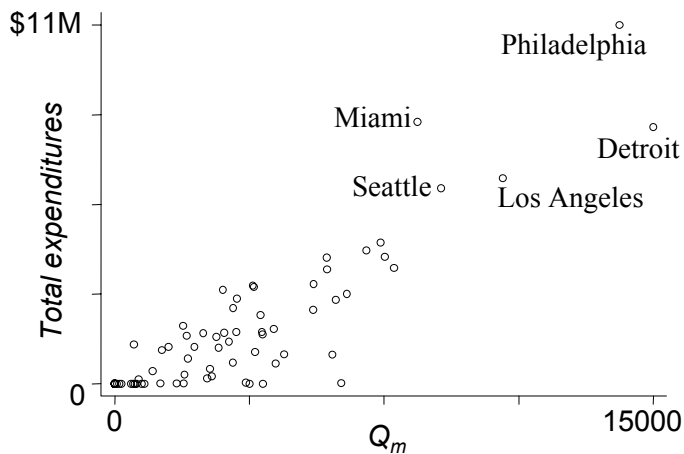


Figure 3.3: Total advertisement expenditures Sept. 1 to election day and Q_m , for the 75 largest media markets

then $\Delta u_s = 0$, and no effects can be estimated. In reality they do not. Under the assumption that Δu_s is not correlated with the popularity shocks, the effect of campaign visits may be estimated by including Δu_s in equation (2.9) and re-arranging

$$\underbrace{\gamma_{st} + \hat{\beta}X_{st}}_{\hat{\varepsilon}_{st}} = \Delta u_s - \eta_{st} - \eta_t.$$

If one assumes the functional form

$$u_s(d_s) = \gamma d_s^\alpha,$$

then the parameters γ and α determine the strength and decreasing marginal impact of campaign visits. Estimating the equation

$$\hat{\varepsilon}_{st} = \gamma \left((d_{st}^D)^\alpha - (d_{st}^R)^\alpha \right) - \eta_{st} - \eta_t$$

yields the estimated parameter values, $\hat{\alpha} = 0.34$, and $\hat{\gamma} = 0.018$. The estimate implies that if the Gore spent one and Bush no days in a state, then Gore would gain 0.7 percentage points; if Gore spent two and Bush one, then Gore would gain 0.2 percentage points; if Gore spent ten and Bush seven days in a state (as was the

case in Florida), then Gore would gain 0.2 percentage points.¹³ These effects are similar to those of Shaw (1999) who estimated the effect of one extra campaign visit to 0.8 extra points in the opinion polls, which, according to the estimates in this paper, corresponds to an increase in of 0.4 percentage points in the election.

In this specification, the equilibrium allocation is

$$\frac{d_s^*}{\sum d_s^*} = \frac{Q_s^{\frac{1}{1-\alpha}}}{\sum Q_s^{\frac{1}{1-\alpha}}}. \quad (3.4)$$

The estimated $\hat{\alpha}$ implies that the marginal impact of an additional campaign visit declines slower than the earlier logarithmic utility specification. Therefore, equilibrium campaign visits increase more than proportionally to Q_s .

It is not meaningful to do the same analysis for the TV-advertising, since there are too few observations. The advertising data is only available for the 2000 election. Also, while the advertising data is by media market, the vote data is only by state. Still a simple look at some correlations may be informative. The correlation between the forecasting error, $\hat{\varepsilon}_{st}$, and $a_s = \sum_{m \in s} \frac{n_{ms}}{n_s} (a_{mst}^D - a_{mst}^R)$ ¹⁴ is 0.24. Figure 3.4 plots the difference in this weighted number of advertisements in a state against the forecasting error in percent.

¹³A complication is that if $\Delta u_{st} \neq 0$, then the estimated equation (2.9) is incorrectly specified. However, including Δu_{st} and re-estimating this equation makes little difference (the correlation between Q_s/v_s estimated with and without Δu_{st} is 0.996).

¹⁴This specification makes sense since the share D votes in state s equals

$$\begin{aligned} \sum_{m \in s} \frac{n_{ms}}{n_s} \Phi \left(\frac{\Delta u_m - \eta_s - \eta}{\sigma_s} \right) &\approx \sum_{m \in s} \frac{n_{ms}}{n_s} \left(\Phi \left(\frac{-\eta_s - \eta}{\sigma_s} \right) + \varphi \left(\frac{-\eta_s - \eta}{\sigma_s} \right) \frac{\Delta u_m}{\sigma_s} \right) \\ &= \Phi \left(\frac{-\eta_s - \eta}{\sigma_s} \right) + \varphi \left(\frac{-\eta_s - \eta}{\sigma_s} \right) \frac{1}{\sigma_s} \sum_{m \in s} \frac{n_{ms}}{n_s} \Delta u_m \\ &\approx \Phi \left(\frac{\sum_{m \in s} \frac{n_{ms}}{n_s} \Delta u_m - \eta_s - \eta}{\sigma_s} \right). \end{aligned}$$

Therefore

$$\hat{\varepsilon}_{st} \approx \gamma \sum_m \frac{n_{ms}}{n_s} \left((a_{mst}^D)^\alpha - (a_{mst}^R)^\alpha \right) + \eta_{st} + \eta_t.$$

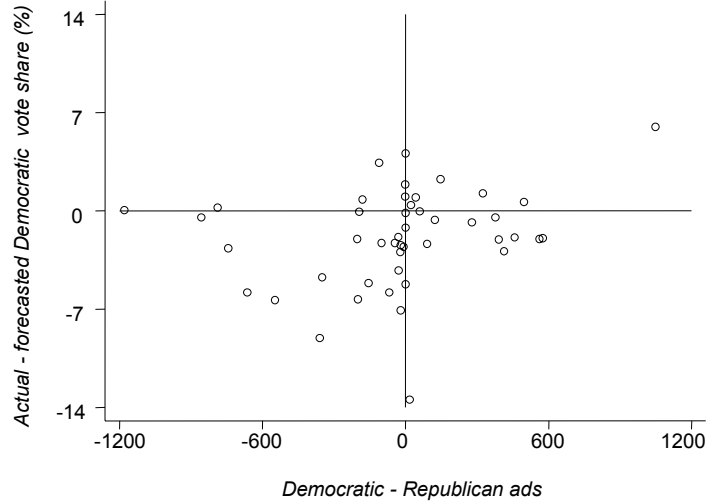


Figure 3.4:

4. Direct national presidential vote

This section will explore the distributional effects of an institutional reform, namely, the change to a direct vote for president. First the equilibrium under Direct Vote will be calculated using the same methodology that was used for the Electoral College. Next, the differences between allocation under the Electoral College and Direct Vote will be discussed. The section ends with a discussion of which electoral system is likely to generate a more unequal distribution of resources.

Suppose the president is elected by a direct national vote. The number of Democratic votes in state s is then equal to

$$v_s F_s(\Delta u_s - \eta - \eta_s).$$

The Democratic candidate wins the election if

$$\sum_s v_s F_s(\Delta u_s - \eta - \eta_s) \geq \frac{1}{2} \sum_s v_s.$$

The number of votes won by candidate D is asymptotically normally distributed

with mean and variance

$$\begin{aligned}\mu_v &= \sum_s v_s \Phi \left(\frac{\Delta u_s - \mu_s - \eta}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}} \right), \\ \sigma_v^2 &= \sigma_v^2(\Delta u_s, \eta).\end{aligned}\tag{4.1}$$

See Appendix 6.7 for the explicit expression for σ_v^2 . The probability of a Democratic victory is

$$P^D = 1 - \int \Phi \left(\frac{\frac{1}{2} \sum_s v_s - \mu_v}{\sigma_v} \right) d\eta.$$

Both candidates again choose election platform subject to the budget constraint. Given that the concavity conditions are satisfied, the following proposition characterizes the equilibrium allocation.

Proposition 2. *A pair of strategies for the parties (d^D, d^R) that constitute a NE in the game of maximizing the expected probability of winning the election must satisfy $d^D = d^R = d^*$, and for all s and for some $\lambda > 0$*

$$Q_{sv} u'(d_s) = \lambda.\tag{4.2}$$

The variable Q_{sv} measures the expected number of marginal voters in state s , evaluated at combinations of national shock and state level shocks which would cause a draw, weighted by the likelihood of these shocks.¹⁵

The main differences between allocation under Direct Vote and under the Electoral College are evident from the expressions for μ_v and μ . First, the number of electoral votes in μ has been replaced by the number of popular votes in μ_v . The incentives to visit states under the Electoral College was roughly proportional to the number of electoral votes. Under Direct Vote, these incentives are instead roughly proportional to the number of popular votes.

Second, the variance in state shocks σ_s^2 in μ has been replaced by the sum of variances in state shocks and preferences, $\sigma_s^2 + \sigma_{fs}^2$, in μ_v . A consequence of this is that allocation under Direct Vote is not very sensitive to the forecasted political bias (vote shares) in the states, μ_s . Since σ_{fs} is about thirteen times larger than σ_s , this is as if the state-level shocks in the Electoral College model were fourteen

¹⁵This is shown in Appendix 6.7. The correlation between Q_{sv} and the average marginal voter densities, evaluated all simulated national election outcomes with a margin of victory closer than 2%, is 0.999.

times their actual size. As discussed in Section 2.2, a larger variance in the state-level shocks implies that $Q_{s\mu}/e_s$ depends less on forecasted vote shares.

This second difference also implies that while μ_v depends on σ_{fs} , μ does not. Under the Direct Vote system, candidates compare how many extra votes they would get by visiting one state compared to the number of votes they would win by visiting another. Therefore the allocation is sensitive to the share of marginal voters in each state, which is captured by the variance in the state preference distribution, σ_{fs} . On the contrary, the share of marginal voters is not important under the Electoral College system. Under this system, the presidential candidates care about whether they win the support of the median voter in the state, they do not care about how many marginal voters there are to his left or right.

Although Q_{sv} is not very sensitive to forecasted Democratic vote shares in the states, it is quite sensitive to the forecasted national vote share. The amplitude of all Q_{sv} is higher when the national election is expected to be close. This is since the likelihood of a draw is then much higher.

Under Direct Vote, the candidates also have an incentive to influence the variance in the vote outcome. Again, candidates who are behind try to increase this variance and candidates who are ahead try to decrease it. The trailing candidate increases the variance by spending more time in large states where he is behind. This moves the expected result closer to 50-50, which increases the number of marginal voters, and thus increases the variance. However, since the share of marginal voters is not very sensitive to expected vote shares, this influence is small.

The allocation under Direct Vote depends crucially on the estimated variance in the preference distribution, σ_{fs}^2 . Therefore, the restriction $\sigma_{fs} = 1$ is removed in the maximum likelihood estimation of equation (2.8), as well as the assumption that σ_s is the same for all states. The identification of σ_{fs} and σ_s is not trivial. If the election outcome in a certain state varies a lot over time, is this because the state has many marginal voters or is it because the state has been hit by unusually large shocks which have shifted voter preferences? The model solves this problem by identifying σ_{fs} by the response in vote shares to shocks that are common to all states, and shocks which are measurable. Specifically, σ_{fs} are empirically identified by the covariation between vote shares and: economic growth at national and state level, incumbency variables, home state of the president and vice president, and dummy variables. States where the vote share outcome covary strongly with economic growth, etc., are thus estimated to have many marginal voters. Maine is estimated to have the largest share of marginal voters while California has the smallest. The variance in the state popularity-shocks, σ_s^2 , is largest in the southern

states, such as Mississippi and South Carolina, and lowest in Ohio, Indiana and Michigan.

Which political system creates a more unequal distribution of resources? I will look at the allocation of advertising expenditures per capita under the two systems. Assuming log utility, equilibrium advertisement expenditures, $p_m a_{mv}^*$, are proportional to

$$\frac{p_m a_m^*}{n_m} = \frac{Q_m}{n_m} = \sum_{s=1}^S \underbrace{\frac{e_s}{n_s}}_{(i)} \underbrace{\frac{Q_s}{e_s}}_{(ii)} \frac{n_{ms}}{n_m},$$

$$\frac{p_m a_{mv}^*}{n_m} = \frac{Q_{mv}}{n_m} = \sum_{s=1}^S \underbrace{\frac{v_s}{n_s}}_{(i)} \underbrace{\frac{Q_{sv}}{v_s}}_{(ii)} \frac{n_{ms}}{n_m}.$$

Since the variables determining the allocation under the two systems are different, it is not possible on theoretical grounds to determine which system will generate more unequal distribution of resources. This will depend on (i) whether electoral votes per capita varies more than voter turnout (votes per capita), and (ii) whether the probability of being a *decisive swing state* per electoral vote varies more than the share of marginal voters. The left hand plot in Figure 4.1 shows that Electoral votes per capita million varies more than voter turnout. While Wyoming has four times as many electoral votes per capita as Texas, Minnesota only has 1.7 times as high voter turnout as Hawaii. The right hand plot shows that the probability of being a *decisive swing state* per electoral vote varies more than the share of marginal voters. While Maine is estimated to have twice as large a share of marginal voters as California, Michigan has infinitely larger $\frac{Q_s}{e_s}$ than New York or Texas.

Given these results, it is not surprising that the equilibrium allocation of advertisement expenditures across advertisement markets is much more equal under Direct Vote. This is shown in Figure 4.2.

Finally Figure 4.3 shows the per capita weight (Q_s respectively Q_{sv}) given to states under the Electoral College system and the Direct Vote system. Under the present system, the allocation of campaign visits and advertisements has been approximately proportional to this weight, see Figures 3.1 and 3.3. The scale on the y-axis is normalized so that 1 represents an equal per capita weight to all states. For example, under the Electoral College, Delaware receives about five times the weight they would receive under equal per capita treatment. New

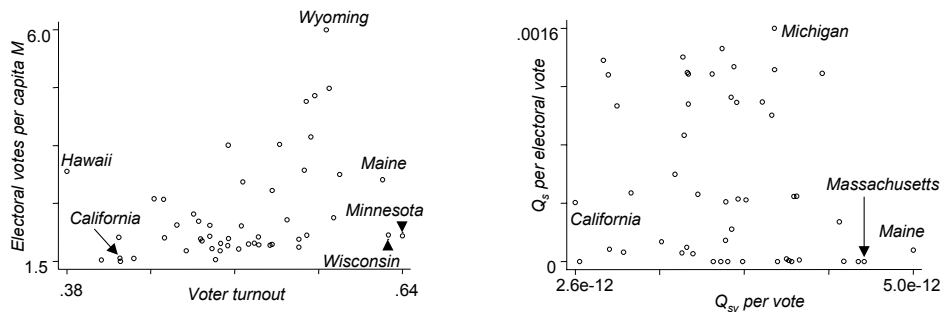


Figure 4.1: Variables creating unequal distribution under the Electoral College and Direct Vote

Hampshire and New Mexico are also winners under the Electoral College system while states like Texas, Massachusetts and Utah would gain from a reform.

Incentives to redistribute campaign resources unequally are much stronger under the Electoral College than under Direct Vote. As I show in Strömberg (2002a), the incentives to allocate federal civilian employment unequally for political purposes, are also much stronger under the Electoral College than under Direct Vote. However, these incentives are weaker for federal employment than for campaign expenditures. The reason is that the uncertainty about the election outcome is larger when decisions about federal employment are taken than during the election campaign. Therefore Q_s are more evenly distributed across states when relating to federal employment. As a result, while the principles for political campaigning (studied by Brams and Davis (1974), or Nagler and Leighley (1992)) are the same as for political distribution of funds (studied by, for example Wright (1974) and Wallis (1996)), the optimal allocations are different.

Note that Q_{sv} does *not* measure average "voting power", that is, the probability that an average voter in state s is decisive in the Direct Vote system. It measures the expected number of voters in state s who are decisive and at the same time swing voters (indifferent between voting for D and R). In contrast, under the Electoral College, Q_s measures "voting power". In this model with a continuum of voters, a vote is decisive when the state is decisive and the state election is exactly tied. This implies that while the allocation of resources under the Electoral College is determined by the "voting power", this is not true under Direct Vote.

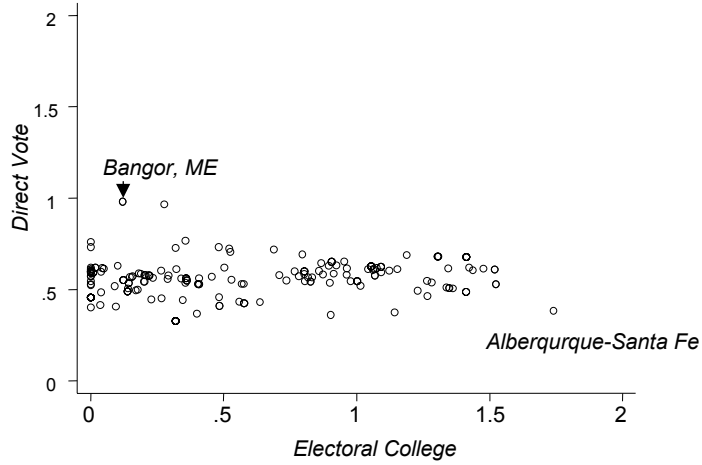


Figure 4.2: Optimal advertisement expenditures (dollars per capita) by advertisement market

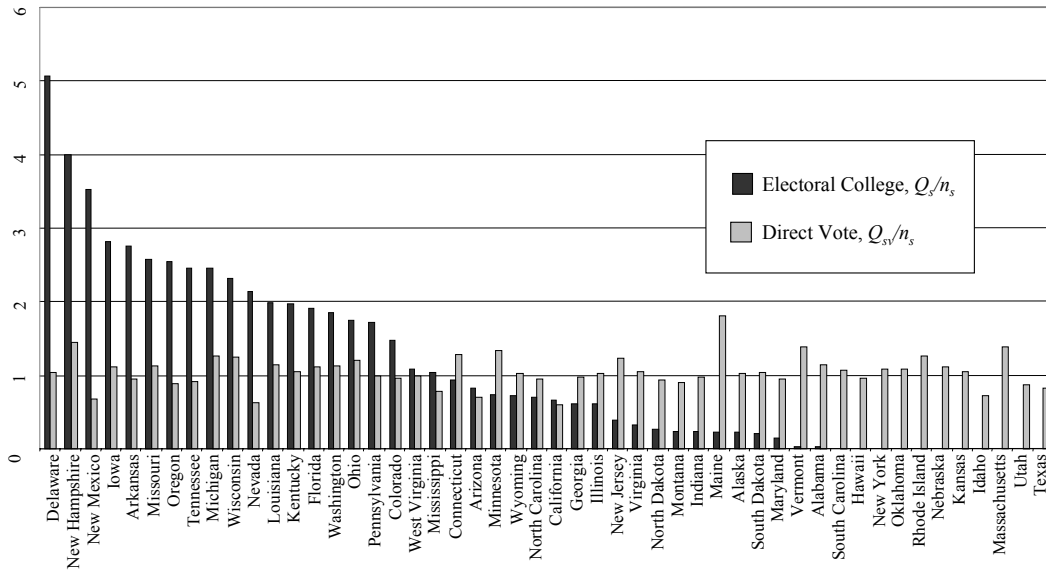


Figure 4.3: Treatment under the Electoral College and Direct Vote

5. Conclusion

This paper explores how the Electoral College shapes incentives for presidential candidates to allocate resources across states. It does so by developing a probabilistic voting model of electoral competition under the Electoral College system. The model delivers a precise recommendation for how presidential candidates, trying to maximize the probability of gaining a majority in the Electoral College, should allocate their resources. Basically, more resources should be devoted to states who are likely to be *decisive swing states*, that is, states who are decisive in the electoral college and, at the same time, have very close state elections. The probability of being a *decisive swing state* is fully characterized, both theoretically, and empirically.

The theoretical solutions show, first, that the probability of being a *decisive swing state* is roughly proportional to the number of electoral votes. Second, this probability *per electoral vote* is highest for states who have a forecasted state election outcome which lies between a draw and the forecasted national election outcome. For example, suppose that the Democrats are ahead 60-40 in the national polls and in Pennsylvania state polls, while the Texas state polls show a draw. On one hand, candidate visits are more likely to influence the state election in forecasted swing states like Texas than in states like Pennsylvania. On the other hand, if Texas is still a swing state on election day, then the Democrats are probably winning by a landslide anyway, while if Pennsylvania is a swing state on election day, then the election at the national level is likely to be close.

The model shows how to strike a balance between high average influence (Texas) and influence when it matters (Pennsylvania). The more correlated state outcomes are, the more attention should be given to states like Pennsylvania. The maximum attention should typically be given to states with polls halfway between a draw and the national polls, around 55-45 in this example. The model also shows that candidates who are trailing should try to go for large states where they are behind. Even if this decreases the expected number of electoral votes that the candidate gets, it increases the variance in the outcome and therefore the probability of winning.

The model is applied to presidential campaign visits across states during the 1988-2000 presidential elections, and to presidential campaign advertisements across media markets in the 2000 election. The actual allocation of these resources closely resembles the optimal allocation in the model. In the 2000 election, the correlation between optimal and actual visits by state is 0.91, and the correlation

between optimal and actual advertisement expenditures by advertising market is 0.88.

The paper finally analyses how the allocation of advertisements across media markets would change under an institutional reform, namely the transition to a direct national vote for president. The allocational principles are quite different under the two systems, and the incentives to favor certain markets are much stronger under the electoral college than under the direct vote, causing a more unequal distribution of resources.

References

- [1] Anderson, Gary M. and Robert D. Tollison, "Congressional Influence and Patterns of New Deal Spending, 1933-1939", *Journal of Law and Economics*, 1991, 34, 161-175.
- [2] Banzhaf, John R., "One man, 3.312 votes: a mathematical analysis of the Electoral College", *Villanova Law Review* 13, 1968, 304-332.
- [3] Brams, Steven J. and Morton D. Davis, "The 3/2 Rule in Presidential Campaigning." *American Political Science Review*, March 1974, 68(1), 113-34.
- [4] Campbell, James E. (1992). Forecasting the presidential vote in the states. *American Journal of Political Science*, 36, 386-407.
- [5] Chamberlain, Gary, and Michael Rothchild, "A note on the probability of casting a decisive vote", *Journal of Economic Theory* 25, 1981, 152-162.
- [6] Cohen, Jeffrey E., "State-level public opinion polls as predictors of Presidential Election Results", *American Politics Quarterly*, Apr98, Vol. 26 Issue 2, 139.
- [7] Colantoni, Claude S., Terrence J. Levesque, and Peter C. Ordeshook, "Campaign resource allocation under the electoral college (with discussion)", *American Political Science Review* 69, 1975, 141-161.
- [8] Gelman, Andrew, and Jonathan N. Katz, "How much does a vote count? Voting power, coalitions, and the Electoral College", *Social Science Working Paper* 1121, 2001, California Institute of Technology.
- [9] Gelman, Andrew, and Gary King, "Why are American presidential election campaign polls so variable when votes are so predictable?", *British Journal of Political Science*, 1993, 23, 409-451.
- [10] Gelman, Andrew, Gary King, and W. John Boscardin, "Estimating the probability of events that have never occurred: when is your vote decisive?", 1998, *Journal of the American Statistical Association*, 93(441), 1-9.
- [11] Hobrook, Thomas M. and Jay A. DeSart, "Using state polls to forecast presidential election outcomes in the American states", *International Journal of Forecasting*, 1999, 15, 137-142.

- [12] Lindbeck, Assar, Jörgen W. Weibull, "Balanced-budget redistribution as the outcome of political competition", *Public Choice*, 1987, 52, 273-97.
- [13] Lizzeri Alessandro and Nicola Persico, "The provision of public goods under alternative electoral incentives", *American Economic Review*, 91(1), 2001, 225-239.
- [14] Merrill, Samuel III, "Citizen Voting Power Under the Electoral College: A Stochastic Voting Model Based on State Voting Patterns", *SIAM Journal on Applied Mathematics*, 34(2), 376-390.
- [15] Nagler, Jonathan, and Jan Leighley, "Presidential campaign expenditures: Evidence on allocations and effects", *Public Choice* 72, 1992, 319-333.
- [16] Persson, Torsten, and Guido Tabellini, "Political Economics: explaining economic policy", MIT Press, 2000.
- [17] Persson, Torsten and Guido Tabellini, "The Size and Scope of Government: Comparative Politics with Rational Politicians," *European Economic Review* 43, 1999, 699-735.
- [18] Rosenstone, Steven J., "Forecasting presidential elections", 1983, New Haven: Yale University Press.
- [19] Shachar, Ron, and Barry Nalebuff, "Follow the leader: Theory and evidence on political participation", *American Economic Review* 89(3), 1999, 525-547.
- [20] Shaw, Daron R., "The Effect of TV Ads and Candidate Appearances on Statewide Presidential Votes, 1988-96", *American Political Science Review* 93(2), 1999, 345-361.
- [21] Strömberg, "The Electoral College and the Distribution of Federal Employment", Stockholm university, 2002a, in preparation.
- [22] Strömberg, "Follow the Leader 2.0: Theory and Evidence on Political Participation in Presidential Elections", Stockholm university, 2002b, in preparation.
- [23] Snyder, James M., "Election Goals and the Allocation of Campaign Resources", *Econometrica* 57(3), 1989, 637-660.

- [24] Wallis, John J., "What Determines the Allocation of National Government Grants to the States?", NBER Historical paper 90, 1996.
- [25] Wright, Gavin , "The Political Economy of New Deal Spending: An Econometric Analysis", Review of Economics and Statistics 56, 1974, 30-38.
- [26] Holbrook and DeSart, 1999; Rosenstone, 1983).

6. Appendix

6.1. Proof of Proposition 1

Symmetry. The best-reply functions of candidates D and R are characterized by the first order conditions

$$\begin{aligned} Q_s u' (d_s^D) &= \lambda^D, \\ Q_s u' (d_s^R) &= \lambda^R. \end{aligned}$$

Therefore,

$$\frac{u' (d_s^D)}{u' (d_s^R)} = \frac{\lambda^D}{\lambda^R}, \quad (6.1)$$

for all s . Suppose that $d^D \neq d^R$. This means that $d_s^D < d_s^R$ for some s , implying that $\lambda^D > \lambda^R$ by equation (6.1). Because of the budget constraint, it must be the case that $d_{s'}^D > d_{s'}^R$ for some s' , which implies $\lambda^D < \lambda^R$, a contradiction. Therefore, $\lambda^D = \lambda^R$ which implies $d_s^D = d_s^R$ for all s .

Uniqueness: Suppose there are two equilibria with equilibrium strategies d and d' corresponding to $\lambda > \lambda'$. The condition on the Lagrange multipliers implies $d_s > d'_s$ for all s which violates the budget constraint. Therefore, the only possibility is $\lambda = \lambda'$ which implies $d_s = d'_s$ for all s .

6.2. Non-interior equilibria

A NE (d^{D*}, d^{R*}) in the game of maximizing the expected probability of winning is characterized by

$$\frac{\partial P^D (d^{D*}, d^{R*})}{\partial d_s^D} = Q_s (u (d^{D*}) - u (d^{R*})) u' (d_s^{D*}) = Q_s u' (d_s^{D*}) = \lambda^D, \quad \text{if } d_s^{D*} \in (0, I],$$

$$\frac{\partial P^D (d^{D*}, d^{R*})}{\partial d_s^{D*}} = Q_s u' (d_s^{D*}) < \lambda^D, \quad \text{if } d_s^{D*} = 0.$$

Similarly for R :

$$\frac{\partial (1 - P^D (d^{D*}, d^{R*}))}{\partial d_s^{R*}} = Q_s u' (d_s^{R*}) = \lambda^R, \quad \text{if } d_s^{R*} \in (0, I],$$

$$\frac{\partial (1 - P^D (d^{D*}, d^{R*}))}{\partial d_s^{R*}} = Q_s u' (d_s^{R*}) < \lambda^R, \quad \text{if } d_s^{R*} = 0.$$

Suppose $\lambda^R > \lambda^D$. First, note that both d_s^{R*} and d_s^{D*} are weakly increasing in Q_s . Suppose R visits ($d_s^{R*} > 0$) the x states with the highest Q_s and D visits the y states with highest Q_s . In states which both candidates visit

$$\frac{u'(d_s^{D*})}{u'(d_s^{R*})} = \frac{\lambda^D}{\lambda^R} < 1,$$

so $d_s^{R*} < d_s^{D*}$. Since D spends more time in all states which both visit, D must visit fewer states, and $x > y$. Therefore, there must be some state s' which R visits but D does not. In this state

$$\lambda^D > Q_s u'(0) > Q_s u'(d_s^{R*}) = \lambda^R.$$

But this contradicts the assumption $\lambda^R > \lambda^D$. Therefore, non-interior equilibria are also symmetric, $d^D = d^R$. In these equilibria, both candidates make the same number of visits to the x states with the highest Q_s .

6.3. Derivation of Q_s

From Proposition 1,

$$\begin{aligned} Q_s &= Q_{s\mu} + Q_{s\sigma} \\ &= e_s \int \frac{1}{\sigma_E} \varphi(x(\eta)) g_s(-\mu_s - \eta) h(\eta) d\eta \\ &\quad + \frac{e_s^2}{\sigma_E^2} \int \varphi(x(\eta)) x(\eta) \left(\frac{1}{2} - G_s(\cdot) \right) g_s(-\mu_s - \eta) h(\eta) d\eta, \end{aligned}$$

where

$$x(\eta) = \frac{\frac{1}{2} \sum_s e_s - \mu}{\sigma_E}.$$

To see that Q_s is approximately the joint probability of a state being decisive in the Electoral College at the same time as having a close election, note that disregarding state s' , the electoral electoral vote outcome, $\sum_{s \neq s'} D_s e_s$, is approximately normally distributed with mean

$$\mu_{-s'} = \mu - e_s G_s(\cdot) \tag{6.2}$$

and variance

$$\sigma_{E-s'}^2 = \sigma_E^2 - e_s^2 G_s(\cdot) (1 - G_s(\cdot)). \tag{6.3}$$

The electoral votes of state s' are decisive if

$$\sum_{s \neq s'} D_s e_s \in \left(\frac{\sum_s D_s}{2}, \frac{\sum_{s'} D_s}{2} - e_s \right).$$

The probability of this event is approximately

$$P = \Phi \left(\frac{\frac{\sum_s D_s}{2} - \mu_{-s'}}{\sigma_{E-s'}} \right) - \Phi \left(\frac{\frac{\sum_s D_s}{2} - e_s - \mu_{-s'}}{\sigma_{E-s'}} \right).^{16}$$

First, make a second-order Taylor-expansion of P around the point

$$x_0(\eta) = \frac{\frac{\sum_s D_s}{2} - \mu}{\sigma_{E-s'}},$$

so that

$$P \approx \varphi(x_0) \frac{e_s}{\sigma_{E-s'}} + \frac{e_s^2}{\sigma_{E-s'}^2} \varphi(x_0) x_0 \left(G_s(\cdot) - \frac{1}{2} \right).$$

Next, given a national shock, η , the probability that the outcome in the state lies within two percent of a draw equals

$$G_s [(-\mu_s - \eta) - \sigma_{f_s} \Phi^{-1}(.49)] - G_s [(-\mu_s - \eta) - \sigma_{f_s} \Phi^{-1}(.51)].$$

To a first-order approximation, this equals:

$$g_s [-\mu_s - \eta] \sigma_{f_s} (\Phi^{-1}(.51) - \Phi^{-1}(.49)).$$

Conditional on the national shock, the events that the state has a close election and that the state is decisive are independent. Therefore the joint probability is the product of the two probabilities. The unconditional probability of being decisive and having a close election is approximately

$$\sigma_{f_s} (\Phi^{-1}(.51) - \Phi^{-1}(.49)) * \int \left(\varphi(x_0) \frac{e_s}{\sigma_{E-s'}} + \frac{e_s^2}{\sigma_{E-s'}^2} \varphi(x_0) x_0 \left(G_s(\cdot) - \frac{1}{2} \right) \right) g_s(-\mu_s - \eta) h(\eta) d\eta$$

The only difference between the above expression and Q_s , apart from the scale factor for two percent closeness, is that σ_E has been replaced by the smaller $\sigma_{E-s'}$. The difference between $\sigma_{E-s'}$ and σ_E is small (typically around one percent).

¹⁶This way of calculating the probability of being pivotal was developed by Merrill (1978).

The values of Q_s reported in the paper have been scaled by $\Phi^{-1}(.51) - \Phi^{-1}(.49)$ to be comparable to the simulated values of Section 2.1. In 3.3 percent of the 1 million simulated elections, Florida was decisive in the Electoral College and had a state margin of victory of less than 2 percent. The scaled Q_s for Florida was 3.5 percent. To get the probability that the state margin of victory is within, say 1000, votes, Q_s should be scaled by $\Phi^{-1}\left(\frac{1}{2} + \frac{500}{v_s}\right) - \Phi^{-1}\left(\frac{1}{2} - \frac{500}{v_s}\right)$. Using this formula, the probability of a state being decisive in the Electoral College, and at the same time having an election result with a state margin of victory less than 1000 was 0.00015 in Florida in the 2000 election. The probability that this would happen in any state was .0044. The probability of a victory margin of one vote in Florida is 0.15 per million, and the probability of this happening in any state is 4.4 per million. The state where one vote is most likely to be decisive is Delaware, where it is decisive .4 times in a million elections.

To arrive at a simpler form for $Q_{s\mu}$, do a first order Taylor expansion of the mean of the expected number of electoral votes $\mu(\eta)$ around $\eta = \tilde{\eta}$ for which $\mu(\tilde{\eta}) = \frac{1}{2} \sum_s e_s$, that is the value of the national shock which makes the expected outcome a draw. With this approximation

$$\mu(\eta) = \sum_s e_s G_s(-\mu_s - \eta) \approx \frac{1}{2} \sum_s e_s - a(\eta - \tilde{\eta}),$$

$$a = \sum_s e_s g_s(-\mu_s - \tilde{\eta}).$$

Since the mean of the electoral votes, $\mu(\eta)$, is much more sensitive to national shocks than is the variance $\sigma_E(\eta)$, the latter is assumed fixed

$$\sigma_E^2(\eta) = \bar{\sigma}_E^2 = \sum_s e_s^2 G_s(-\mu_s - \tilde{\eta})(1 - G_s(-\mu_s - \tilde{\eta})).$$

Then

$$\frac{1}{\sigma_E(\eta)} \varphi\left(\frac{\frac{1}{2} \sum_s e_s - \mu}{\sigma_E(\eta)}\right) \approx \frac{1}{\bar{\sigma}_E} \varphi\left(\frac{\eta - \tilde{\eta}}{\bar{\sigma}_E/a}\right)$$

This approximation is very good. Figure 6.1 shows the true, tfi, and the approximated functions, pfi. The values are calculated for an interval of four standard deviations centered around $\tilde{\eta}$ in the 2000 presidential election. We now have

$$Q_{s\mu} \approx e_s \frac{1}{\bar{\sigma}_E} \int \varphi\left(\frac{\eta - \tilde{\eta}}{\bar{\sigma}_E/a}\right) g_s(-\mu_s - \eta) h(\eta) d\eta.$$

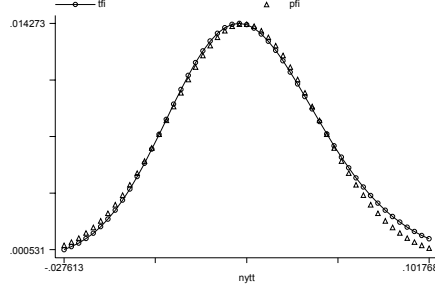


Figure 6.1:

Integrating over η ,

$$Q_{s\mu} \approx e_s \frac{\omega}{2\pi} \exp \left(-\frac{1}{2} \frac{\sigma_s^2 \tilde{\eta}^2 + (\bar{\sigma}_E/a)^2 \mu_s^2 + \sigma^2 (\tilde{\eta} + \mu_s)^2}{\sigma_s^2 \sigma^2 + (\bar{\sigma}_E/a)^2 \sigma^2 + (\bar{\sigma}_E/a)^2 \sigma_s^2} \right) \quad (6.4)$$

where

$$\omega^2 = \left(\frac{1}{(\bar{\sigma}_E/a)^2} + \frac{1}{\sigma_s^2} + \frac{1}{\sigma^2} \right)^{-1}.$$

$Q_{s\mu}$ may be written as

$$Q_{s\mu} \approx e_s \frac{\omega}{2\pi} \exp \left(-\frac{1}{2} \frac{c + (\mu_s - \mu_s^*)^2}{\tilde{\sigma}^2} \right),$$

where

$$\mu_s^* = -\frac{\sigma^2}{\sigma^2 + (\bar{\sigma}_E/a)^2} \tilde{\eta},$$

and

$$\tilde{\sigma}^2 = \sigma_s^2 + \frac{1}{\left(\frac{1}{\sigma^2} + \frac{1}{(\bar{\sigma}_E/a)^2} \right)}.$$

$Q_{s\sigma}$ is calculated using numerical integration.

6.4. Regional swings

The model is now extended to allow for regional popularity swings. These swings are captured by the parameters η_r , ($r = 1, 2, 3$, or 4 , depending on whether the

state is in the Northeast, the Midwest, the South or the West). A voter i in state s in region r will vote for D if

$$\Delta u_s = u_s(d_s^D) - u_s(d_s^R) \geq R_i + \eta_s + \eta_r + \eta.$$

All four regional swing parameters, η_r , are independently drawn from the same normal distribution with mean zero and variance σ_r^2 .

The equilibrium ($\Delta u_s = 0$) election result in one state equals

$$y_{st} = F_{st}(-\eta_{st} - \eta_{rt} - \eta_t).$$

Inverting

$$\Phi^{-1}(y_{st}) = \gamma_{st} = -\frac{1}{\sigma_f} (\mu_{fst} + \eta_{st} + \eta_{rt} + \eta_t).$$

This is a hierarchical random-effects model. Assuming that $\sigma_f = 1$, and given the national and regional shocks, γ_{st} is normally distributed with mean $-(\mu_{fs} + \eta_{rt} + \eta_t)$ and variance σ_s^2 . Let h be the probability density function of γ_{st} :

$$h(\gamma_{st}; \eta_{rt}, \eta_t) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left(-\frac{1}{2} \frac{\left(\gamma_{st} + \frac{1}{\sigma_f} (\mu_{fs} + \eta_{rt} + \eta_t)\right)^2}{\sigma_s^2}\right).$$

The joint density of state outcomes (there are S_r states in region r), equals

$$h(\gamma) = \int_{-\infty}^{\infty} \left(\prod_{r=1}^R \int_{-\infty}^{\infty} \prod_{s=1}^{S_r} h(\gamma_{st}; \eta_{rt}, \eta_t) g_r(\eta_{rt}) d\eta_{rt} \right) h(\eta_t) d\eta_t.$$

The parameters are estimated using maximum likelihood estimation. As reported in the main text, the estimated state and national level shock variances are similar to those estimated without allowing for regional shocks, $\sigma_{s,post1984} = 0.084$, and $\sigma = 0.038$. The standard deviation of the regional shock is $\sigma_r = 0.054$ before state level forecasts were available in 1988. However, after 1988, the standard deviation of the regional shocks is zero. Taking into account the information of September state-level opinion polls, there are no significant regional swings.

Even though there were no significant regional shocks in this case, it is interesting to know how the equilibrium would change with regional shocks. Conditional

on the national and regional shocks, the state outcomes within each region are independent. Using the Central Limit Theorem for all outcome states within region r ,

$$\Pr \left(\sum_{s=1}^{S_r} D_{sr} e_{sr} \leq x \mid \eta_r, \eta \right) \approx \Phi \left(\frac{x - \mu_r}{\sigma_{E_r}^2} \right),$$

where

$$\begin{aligned} \mu_r &= \mu_r(d^D, d^R, \eta_r, \eta) = \sum_{s=1}^{S_r} e_{sr} G_{sr}(\Delta u_{sr} - \mu_{sr} - \eta_r - \eta), \\ \sigma_{E_r}^2 &= \sigma_{E_r}^2(d^D, d^R, \eta_r, \eta) = \sum_{s=1}^{S_r} e_{sr}^2 G_{sr}(\cdot) (1 - G_{sr}(\cdot)). \end{aligned}$$

To estimate the distribution of electoral votes, independent of the regional shocks, do a Taylor-expansion of μ_r around $\mu(d^D, d^R, \eta_r = 0, \eta)$, and assume $\sigma_{E_r}^2 = \sigma_{E_r}^2(d^D, d^R, \eta_r = 0, \eta)$:

$$\begin{aligned} &\Pr \left(\sum_{s=1}^{S_r} D_{sr} e_{sr} \leq x \mid \eta \right) \\ &\approx \int_{\eta_r} \Phi \left(\frac{x - \sum_{s=1}^{S_r} e_{sr} G_{sr}(\Delta u_{sr} - \mu_{sr} - \eta) + a_r \eta_r}{\sigma_{E_r}} \right) g(\eta_r) d\eta_r \\ &= \Phi \left(\frac{x - \mu_{sr}}{\sqrt{\sigma_{E_r}^2 + a_r^2 \sigma_r^2}} \right), \end{aligned}$$

where

$$a_r = \sum_{s=1}^{S_r} e_{sr} g_{sr}(\Delta u_{sr} - \mu_{sr} - \eta).$$

So the regional election outcome is approximately normally distributed with mean

$$\mu_{sr} = \sum_{s=1}^{S_r} e_{sr} G_{sr}(\Delta u_{sr} - \mu_{sr} - \eta),$$

and variance

$$\sigma_{E_r}^2 + a_r^2 \sigma_r^2.$$

The total election outcome, the sum of the regional outcomes, is normally distributed with mean and variance

$$\begin{aligned}\mu &= \mu(d^D, d^R, \eta) = \sum_{r=1}^R \sum_{s=1}^{S_r} e_s G_s(\Delta u_s - \mu_s - \eta) = \sum_{s=1}^S e_s G_s(\Delta u_s - \mu_s - \eta), \\ \sigma_E^2 &= \sigma^2(d^D, d^R, \eta) = \sum_{r=1}^R (\sigma_{E_r}^2 + a_r^2 \sigma_r^2) \\ &= \sum_{s=1}^S e_s^2 G_s(\cdot) (1 - G_s(\cdot)) + \sum_{r=1}^R \left(\sum_{s=1}^{S_r} e_s g_s(\Delta u_s - \mu_s - \eta) \right)^2 \sigma_r^2.\end{aligned}$$

The mean is the same as without regional swings. However, allowing for regional correlation increases the variance in the electoral vote outcome, σ_E^2 . This will have two effects. Most importantly, Q_s becomes less sensitive to vote shares, as the variance of the normal form curve of Figure 2.3 increases. Therefore it becomes more important to visit large states, such as California, rather than states where the outcome is close to 50.8 percent. Secondly, it becomes more important to be close to 50 percent relative to the national shock. This is apparent from the mathematical expressions for the mean and variance of that distribution, see equation (2.10), and footnote (4).

6.5. Campaign advertisements

This Appendix analyses the decision of presidential candidates to allocate advertisements across media markets. Two presidential candidates: R and D , select the number of ads, a_m , in each media market m subject to

$$\sum_{m=1}^M p_m d_m^J \leq I,$$

$J = R, D$. Media market m contains a mass v_{ms} of voters in state s . Voters are affected by campaign advertisements as captured by the function $w(a_m)$. A voter i in media market m in state s will vote for D if

$$\Delta w_m = w(a_m^D) - w(a_m^R) \geq R_i + \eta_s + \eta.$$

In each media market m in state s , the individual specific preferences for candidates, R_i , are distributed with cumulative density function F_s . The state and national-level popularity-swings are drawn from the same distributions as before.

The share D votes in media market m in state s equals

$$F_s(\Delta w_m - \eta_s - \eta).$$

D wins the state if

$$\sum_m v_{ms} F_s(\Delta w_m - \eta_s - \eta) \geq \frac{v_s}{2}.$$

Define the total (state and national) swing which causes a draw in the state:

$$\bar{\eta}_s(a^D, a^R) : \sum_m v_{ms} F_s(\Delta w_m - \bar{\eta}_s(a^D, a^R)) = \frac{v_s}{2}. \quad (6.5)$$

D wins state s if

$$\eta_s + \eta \leq \bar{\eta}_s(a^D, a^R).$$

Conditional on the aggregate shock η , and the platforms, a^D , and a^R , this happens with probability

$$G_s(\bar{\eta}_s(a^D, a^R) - \eta).$$

The function $\bar{\eta}_s(a^D, a^R)$ now plays the same role as Δu_s in Section 2, see equation (2.2). The rest of the analysis is in that section, only exchanging $\bar{\eta}_s(a^D, a^R)$ for Δu_s . The best reply functions of D is characterized by

$$\sum_{s=1}^S Q_s \sum_{m=1}^M \frac{\partial \bar{\eta}_s(a^D, a^R)}{\partial a_m^D} = \lambda^D p_m. \quad (6.6)$$

Similarly, the best reply function of R is characterized by equation (6.6), replacing superscripts D by R . Because of the fixed budget constraint, the allocations must be symmetric, $\lambda^D = \lambda^R = \lambda$. Differentiating equation (6.5), and evaluating it in equilibrium yields

$$\frac{\partial \bar{\eta}_s(a^D, a^R)}{\partial a_m^D} = \frac{v_{ms}}{v_s} w'(a_m^D).$$

Proposition 3. *A pair of strategies for the parties (a^D, a^R) that constitute a NE in the game of maximizing the expected probability of winning the election must satisfy $a^D = a^R = a^*$, and for all s and for some $\lambda > 0$*

$$Q_m w'(a_m^*) = \lambda p_m, \quad (6.7)$$

where

$$Q_m = \sum_{s=1}^S Q_s \frac{v_{ms}}{v_s}.$$

Q_m is the sum of the Q_s of the states in the media market weighted by the share of the voting population of state s that lives in media market m .

Last, assuming that votes per capita

$$\frac{v_{ms}}{n_{ms}} = t_s$$

is the same for all media markets within each state,

$$Q_m = \sum_{s=1}^S Q_s \frac{n_{ms}}{n_s}.$$

6.6. Visits to media markets

One possible objection is that candidates do not choose which states to visit, but rather which media markets. Suppose that instead of buying ads, the candidates can decide which media markets to visit. A voter i in media market m in state s will vote for D if

$$\Delta u_m = u(d_m^D) - u(d_m^R) \geq R_i - D_i + \eta_s + \eta.$$

This model is the same as the advertising model with the same price for a visit, $p_m = 1$, in all media markets. The equilibrium condition is

$$u'(d_m^*) Q_m = \lambda. \tag{6.8}$$

To compare this allocation to the optimal allocation when the state was the unit of analysis, the visits to a media market was distributed across states according to the share of the media market population that lives in that state:

$$\tilde{d}_s = \sum_m \frac{n_{ms}}{n_m} d_m^*.$$

Q_m and \tilde{d}_s were calculated using data on the 75 largest Designated Market Areas (DMA's) used by Nielsen Media Research. A DMA consists of all counties whose largest viewing share is given to stations of that same market area. Non-overlapping DMA's cover the entire continental United States, Hawaii and parts of Alaska. The resulting allocation of visits to media markets, distributed across states, is shown in Figure 6.2, and compared to the equilibrium visits to states.

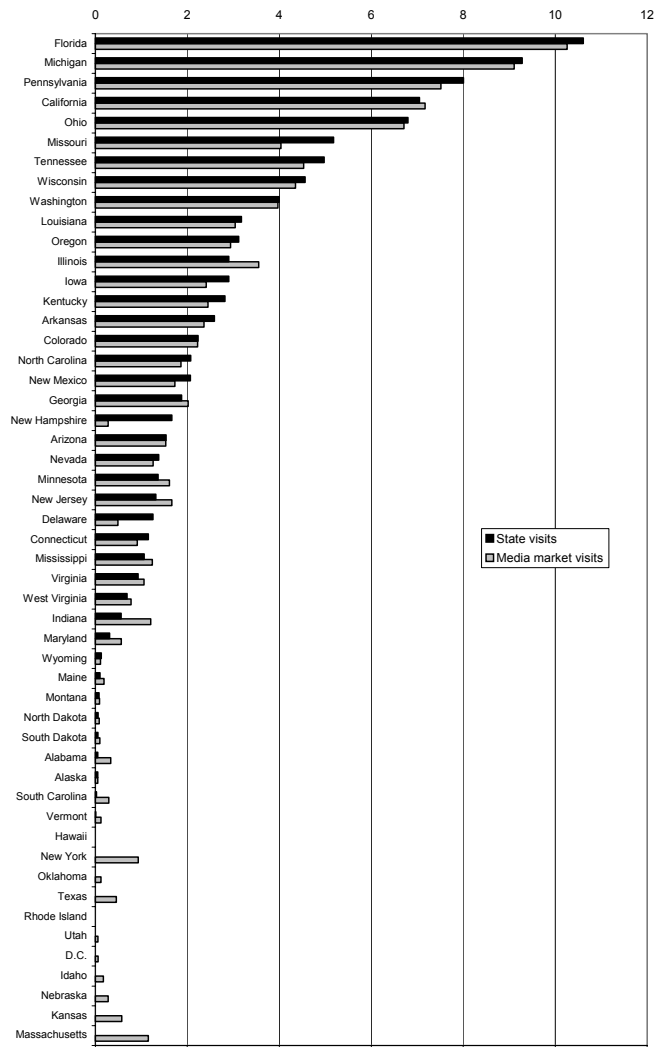


Figure 6.2: Visits to media markets, allocated across states

The resulting allocations are quite similar. Since media markets cross state borders, there are spillovers. This is why New York and Massachusetts receive more visits, and New Hampshire less, if candidates target media markets. Second, states which are more difficult to cover receive more visits. To get the intuition, consider two identical states, except that one is covered by a single media market and the other by two of equal size. Suppose both candidates make four visits to each state: in the state with two media markets they visit both markets twice. In this case, the marginal impact of a visit on the probability of winning must be higher in the state with the two media markets. In this state each voter has only seen the candidates twice while each voter in the other state has seen the candidate four times.

It may also be interesting to know the average number of times people living in a state get the chance to see the candidates. The number of equilibrium visits to an average media market in state s is

$$\bar{d}_s = \sum_m \frac{n_{ms}}{n_s} d_m^*.$$

This is highest in Delaware, Michigan, and Pennsylvania. Because of its many media markets, Florida only comes in eighth place.

6.7. Direct presidential vote

The mean outcome in state s is

$$\begin{aligned} \mu_{vs}(z^L, z^R, \eta) &= \int F_s(\Delta u_s - \eta - \eta_s) g_s(\eta_s) d\eta_s \\ &= \Phi\left(\frac{\Delta u_s - \eta - \mu_s}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right).^{17} \end{aligned}$$

The expression for the variance in one state is, by definition,

$$\sigma_{vs}^2 = v_s^2 \int \left(\Phi\left(\frac{\Delta u_s - \eta - \eta_s - \mu_s}{\sigma_{fs}}\right) - \Phi\left(\frac{\Delta u_s - \eta - \mu_s}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right) \right)^2 g_s(\eta_s) d\eta_s,$$

¹⁷This is not completely trivial. Contact the author for details.

and the total variance is

$$\sigma_v^2 = \sum_s \sigma_{vs}^2.$$

The probability of D winning the election is

$$P^D(z^D, z^R) = 1 - \int \Phi\left(\frac{\frac{1}{2}\sum_s v_s - \mu_v}{\sigma_v}\right) h(\eta) d\eta.$$

where:

$$\begin{aligned} \mu_v &= \sum_s v_s \Phi\left(\frac{\Delta u_s - \mu_s - \eta}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right), \\ \sigma_v^2 &= \sigma_v^2(\Delta u_s, \eta). \end{aligned} \tag{6.9}$$

The derivative of $P^D(z^D, z^R)$ with respect to Δu_s equals

$$Q_{sv} = Q_{sv\mu} + Q_{sv\sigma},$$

where

$$Q_{sv\mu} = \frac{v_s}{\sigma_v \sqrt{\sigma_s^2 + \sigma_{fs}^2}} \int \varphi\left(\frac{\frac{1}{2}\sum_s v_s - \mu_v}{\sigma_v}\right) \varphi\left(\frac{\Delta u_s - \mu_s - \eta}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right) h(\eta) d\eta,$$

$$Q_{sv\sigma} = \frac{1}{\sigma_v} \int \varphi\left(\frac{\frac{1}{2}\sum_s v_s - \mu_v}{\sigma_v}\right) \left(\frac{\frac{1}{2}\sum_s v_s - \mu_v}{\sigma_v}\right) \frac{\partial \sigma_v}{\partial \Delta u_s} h(\eta) d\eta,$$

and

$$\begin{aligned} \frac{\partial \sigma_v^2}{\partial \Delta u_s} &= 2v_s^2 \int \left(\Phi\left(\frac{\Delta u_s - \eta - \eta_s - \mu_s}{\sigma_{fs}}\right) - \Phi\left(\frac{\Delta u_s - \eta - \mu_s}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right) \right) \\ &\quad \left(\frac{1}{\sigma_{fs}} \varphi\left(\frac{\Delta u_s - \eta - \eta_s - \mu_s}{\sigma_{fs}}\right) - \frac{1}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}} \varphi\left(\frac{\Delta u_s - \eta - \mu_s}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}}\right) \right) g_s(\eta_s) d\eta_s \end{aligned}$$

Using this information, $Q_{sv\mu}$ may be calculated either using the same type of approximation as used for $Q_{s\mu}$, or by numerical integration. The second term,

$Q_{sv\sigma}$, is calculated by numerical integration. However, it turns out that this effect is negligible compared to $Q_{sv\mu}$.

The interpretation of $Q_{sv\mu}$ is the following. Conditional on η , the expected number of swing voters is

$$v_s \int_{-\infty}^{\infty} f_s(\Delta u_s - \eta - \eta_s) g_s(\eta_s) d\eta_s = \frac{v_s}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}} \varphi \left(\frac{\Delta u_s - \mu_s - \eta}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}} \right).$$

and the likelihood of a tie is

$$\frac{1}{\sigma_v} \varphi \left(\frac{\frac{1}{2} \sum_s v_s - \mu_v}{\sigma_v} \right).$$

Therefore, the unconditional expected number of swing voters in case of a tie is

$$\frac{v_s}{\sigma_v \sqrt{\sigma_s^2 + \sigma_{fs}^2}} \int \varphi \left(\frac{\frac{1}{2} \sum_s v_s - \mu_v}{\sigma_v} \right) \varphi \left(\frac{\Delta u_s - \mu_s - \eta}{\sqrt{\sigma_s^2 + \sigma_{fs}^2}} \right) h(\eta) d\eta,$$

which is the same as $Q_{sv\mu}$.

6.8. Is the electoral vote outcome normally distributed?

This appendix plots the distributions of election outcomes for all elections 1948-2000, with and without using the asymptotic normal distribution. To avoid numerical integration over η , the following extra approximation is exploited. Using the asymptotic distribution, the approximate probability of D vote share higher than V_s is

$$\begin{aligned} & \int 1 - \Phi \left(\frac{V_s - \mu}{\sigma_E} \right) h(\eta) d\eta \\ & \mu \approx \mu(\eta = 0) - \sum_s e_s g_s(-\mu_s) \eta \\ & \approx \int 1 - \Phi \left(\frac{V_s - \mu(0) + a\eta}{\sigma_E} \right) h(\eta) d\eta = 1 - \int \Phi \left(\frac{\eta + \frac{1}{a}(V_s - \mu(0))}{\sigma_E/a} \right) h(\eta) d\eta \\ & = 1 - \Phi \left(\frac{\frac{1}{a}(V_s - \mu(0))}{\sqrt{(\sigma_E/a)^2 + \sigma^2}} \right) = \Phi \left(\frac{V_s - \mu(0)}{\sqrt{\sigma_E^2 + a^2\sigma^2}} \right). \end{aligned}$$

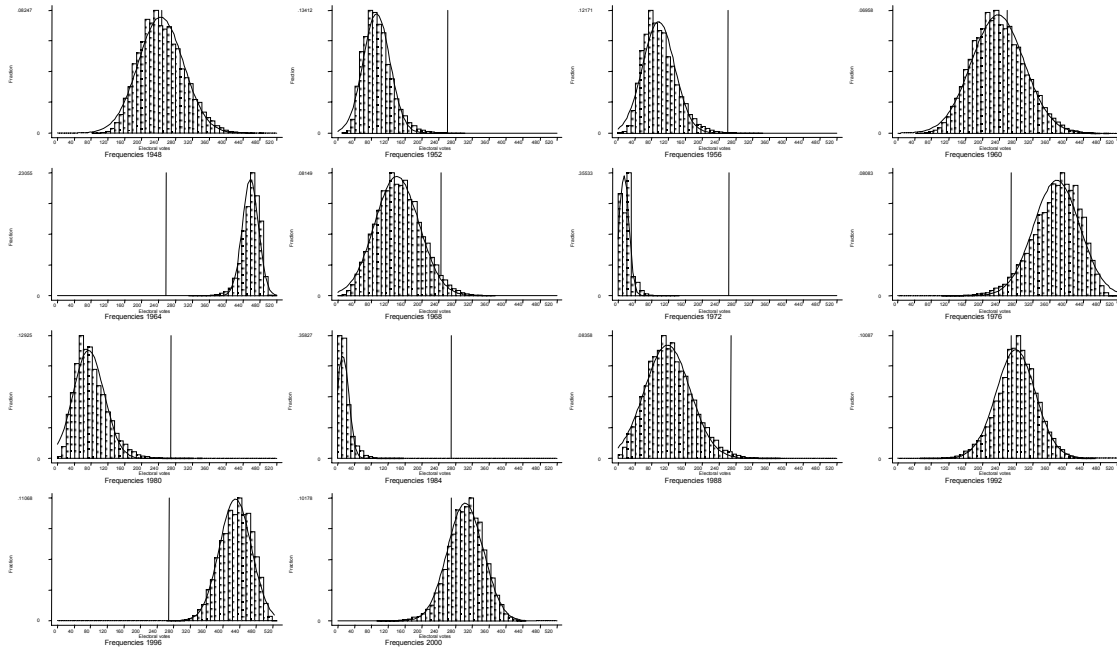


Figure 6.3: Frequency histograms for 100 000 simulated elections, and their asymptotic distribution, for each election 1948-2000

In the Figure 6.3, this distribution is plotted together with the frequencies of the simulated electoral vote outcomes. As is evident from the graphs, simulated distributions are not very different from the approximations. The solid line denotes half the number of electoral votes. In Figure 6.4, the simulated and approximate cumulative frequencies are plotted. The non-approximated distributions are nice and uni-modal, and the approximate distributions seem to follow them closely. In terms of the decision problem of the candidates, the relevant statistic is the correlation between Q_s and the joint likelihood of a state being decisive and having a close election, as shown in Figure 2.1.

6.9. Data definitions and sources

The data-set contains state elections for the 50 states 1948-2000, except Hawaii and Alaska which began voting in the 1960 election. During this period there

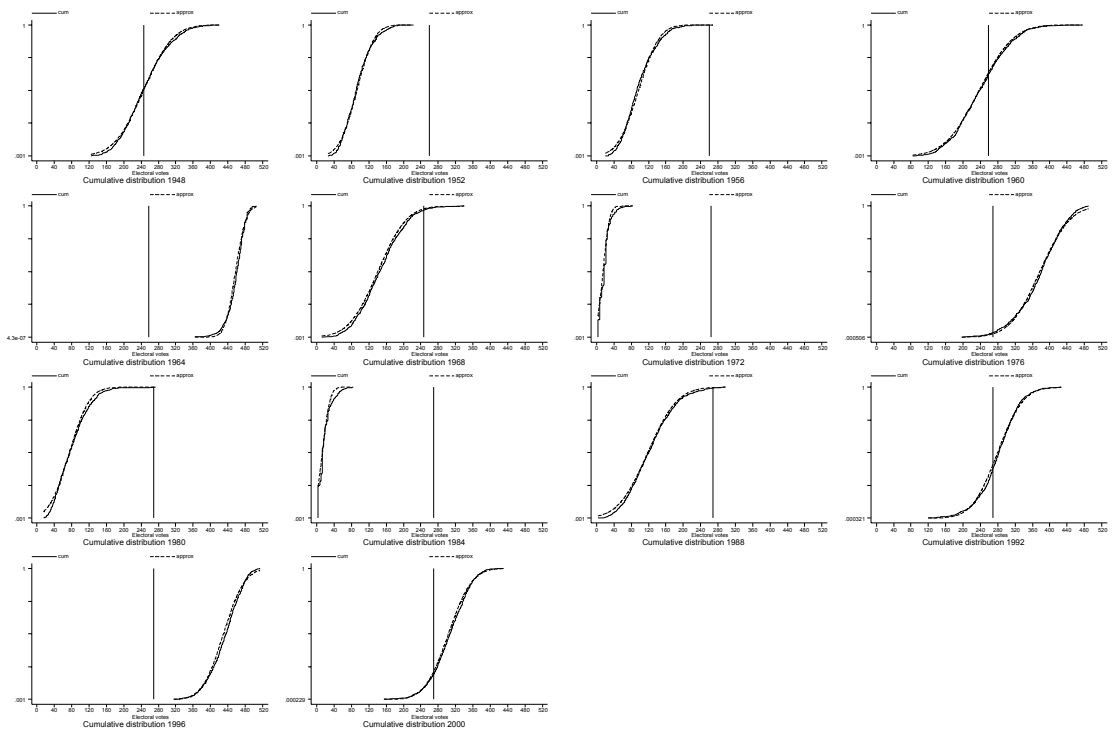


Figure 6.4: Simulated and asymptotic cdf, 100 000 simulated elections, for each election 1948-2000.

were a total of 694 state-level presidential election results. Of this total, 13 state elections were excluded, leaving a total of 681 observations. Four elections in Alaska and Hawaii were excluded because there were no lagged vote returns. Nine elections are omitted because of idiosyncracies in Presidential voting in Alabama in 1948, and 1964, and in Mississippi in 1960; see Campbell 1992.

- Dmvote: state Democratic percentage of the two-party presidential vote. Source: 1940-1944, ICPSR Study 0019; 1948-1988, Campbell; 1992, 1996, Statistical Abstract of the United States 2000; 2000, Federal Election Commission, 2000 OFFICIAL PRESIDENTIAL GENERAL ELECTION RESULTS.
- Electoral votes won (by state). Source: National Archives and Records Administration.
- National trial-heat poll results. Source: 1948-1996, Campbell (2000); 2000, Gallup.
- Second quarter national economic growth, multiplied by 1 if Democratic incumbent president and -1 if Republican incumbent president. Source: August or September election year issue of the Survey of Current Business, U.S. Department of Commerce, Bureau of Economic Analysis.
- Growth in personal state's total personal income between the prior year's fourth quarter and the first quarter of the election year, standardized across states in each year, multiplied by 1 if Democratic incumbent president and -1 if Republican incumbent president. Source: Survey of Current Business, U.S. Department of Commerce, Bureau of Economic Analysis.
- Incumbent: 1 if incumbent president Democrat, -1 if incumbent president Republican.
- Presinc: 1 if incumbent Democratic president seeking re-election, -1 if incumbent Republican president seeking re-election.
- President's home state: 1 if Democratic president home state, -1 if Republican (0.5 and -0.5 for large states (New York, Illinois, California). Source: Campbell 1948-1988; 1992-2000: Dave Leip's Atlas of U.S. Presidential Elections.

- Vice president's home state: 1 if Democratic president home state, -1 if Republican (0.5 and -0.5 for large states (New York, Illinois, California). Source: Campbell 1948-1988; 1992-2000: Dave Leip's Atlas of U.S. Presidential Elections.
- Average ADA-scores: Average ADA-scores of state's members in Congress year before election. Source: Tim Groseclose (<http://faculty-gsb.stanford.edu/groseclose/homepage.htm>).
- Legis: Partisan division of the lower chamber of the state legislature after the previous midterm election. Index is Democratic share of state legislative seats above the 50% mark. Two states, Nebraska and Minnesota, held nonpartisan state legislative elections for all (Nebraska) or part (Minnesota) of the period under study. In the case of Nebraska, the state legislative division was estimated based on the ranking of states of Wright, Erikson, and McIver's state partisan rankings based on public opinion data. Using this index, Nebraska was assigned the mean partisan division of the state most similar to it on the public opinion index, the nearly equally Republican state of North Dakota. The partisan division of the Minnesota legislature in its nonpartisan years (before 1972) is coded as the mean of its partisan division once it reformed to partisan elections (62% Democratic). Washington D.C. was as having the same partisan division as Maryland. Source: 1948-1988, Campbell; 1992-2000, Statistical Abstract of the United States.
- State-level opinion polls. Democratic share of two party vote. Source: Pre-election issues of the Hotline (www.nationaljournal.com).
- Regional dummy variables, see Campbell.
- Participation rate: 1948-1988, Shachar and Nalebuff (1999), voting age population 1992-2000, and votes cast 2000, Federal Election Commission, votes cast 1992, 1996, Statistical Abstract of the United States.

Table 1. Dependent variable: γ_{st} , Φ^1 (democratic share of two-party vote)

National variables:	
Opinion poll	0.453 (0.053)
Second quarter economic growth	0.064 (0.010)
Incumbent president running for re-election	0.053 (0.023)
Incumbency	-0.036 (0.018)
Constant	-0.049 (0.010)
State variables, 1948-1984	
Lagged democratic share of two-party vote, difference from national mean	0.266 (0.032)
Twice lagged democratic share of two-party vote, difference from national mean	0.216 (0.025)
Home state of presidential candidate	0.175 (0.028)
Home state of vice presidential candidate	0.056 (0.023)
First quarter state economic growth.	0.017 (0.005)
Average ADA-scores	0.0021 (0.0003)
Democratic vote-share in midterm state legislative election	0.020 (0.007)
State variables, 1988-2000	
Lagged democratic share of two-party vote, difference from national mean	0.515 (0.081)
Twice lagged democratic share of two-party vote, difference from national mean	0.074 (0.070)
Average ADA-scores	0.0009 (0.0004)
State-level opinion poll	0.389 (0.050)
σ	0.033 (0.007)
$\sigma_{s1948-1984}$	0.102 (0.003)
$\sigma_{s1948-1984} - \sigma_{s1988-2000}$	0.025 (0.005)
Average prediction error (percentage points)	3.0
Number of observations	681