The Informational Content of Campaign Advertising

Gregory J. Martin*

October 24, 2012

Abstract

Understanding the mechanisms by which political advertising affects voters is crucial for evaluating the welfare effects of campaign finance and election regulation. This paper develops a method to distinguish between two alternative mechanisms for advertising influence: an "informative" channel in which voters learn about candidate attributes through advertising, and a "persuasive" channel in which voters can be directly influenced by advertising even if it makes no contribution to the quality of information they possess about the sponsoring candidate. I separately identify the impact of each mechanism on voter choice by taking advantage of variation in voters' prior information. I first construct a dataset of all television advertisements aired in the 2002 and 2004 US Senate and Gubernatorial elections, matched with voting data at the media-market level. I then construct and estimate a structural model of vote choice that allows for both informative and persuasive effects of advertising. The results are largely consistent with the persuasive hypothesis. Using the estimated parameters, I conduct counterfactual analyses of several alternative campaign finance regimes and discuss implications for voter welfare.

*PhD Candidate in Political Economy, Stanford Graduate School of Business. Send correspondence to: gjmartin@stanford.edu

1 Introduction

With the rise of so-called "super PACs" following the Supreme Court's decision in *Citizens United*, questions about the effects of campaign advertising on election outcomes have risen to the forefront of public debate. Opponents of the ruling, which allows unlimited spending by outside groups on advertising intended to influence elections, argue that it gives undue influence to wealthy individuals and organizations, subverting democratic principles of equal representation. Proponents - including the *Citizens United* plaintiffs themselves - argue first that the constitution prohibits Congress from restricting individuals' political speech, and second that outside PACs are not manipulating voters but rather informing them about the consequences of the important choice they must make.

Though the level of media attention directed at the topic has risen recently, these arguments are not new. The *Citizens United* controversy is merely the latest instance of a long-running debate about the merits and consequences of the American system of private campaign finance. Campaign contributions as manipulation versus campaign contributions as free speech is a recurring theme in American politics since the dawn of the television age expanded candidates' opportunities for mass communication and made large-scale fundraising a necessary prerequisite of serious campaigns for national office. Because a large majority of campaign funds raised are spent on advertising, adjudicating this debate hinges in large part on understanding how advertising affects voters.[1]

One possibility is that campaign advertisements provide valuable information to voters. Political insiders and large donors might have better information on the relative strengths of candidates than do average voters; insiders might have interacted with the candidates personally and better understand their likely effectiveness in office, or they might have deeper understanding of certain policy issues and hence a better idea of the candidates' competence for addressing those issues. By observing these informed donors' contributions to candidates - or advertising paid for by those contributions - voters could infer the insiders' information without having to incur the costs of acquiring it themselves. Under this "informative" hypothesis, restricting campaign advertising, either directly by regulating advertisements or

[1] "Adjudicating" here refers exclusively to the aggregate welfare consequences of campaign finance policy, e.g., putting aside legal questions of the right to speech, the distinction between money and speech, and so forth.

indirectly by limiting campaign contributions, would tend to degrade the quality of elected candidates. If fundraising effectiveness is a positive correlate of performance in office, then limiting voters' ability to observe the signal limits the ability of high-quality candidates to distinguish themselves from mediocre ones.

A second possibility is that advertising alters voters' preferences directly. Viewing a campaign ad might instill a positive image of the sponsoring candidate in the voter's mind, independently of the candidate's underlying qualities. Or, ads might make a voter feel he is part of a broader movement of like-minded people, increasing his subjective utility from turning out and voting. Under this "persuasive" hypothesis, voters can be manipulated into voting against their pre-advertising preferences. Private campaign finance thus tilts elections towards candidates who can attract the support of rich individuals or organizations, biasing policy outcomes in favor of these groups at the expense of the mass public.

This paper develops a method to empirically distinguish these two mechanisms of advertising influence, building on an insight due to Ackerberg (2001) and Ansolabehere and Iyengar (1996). The strategy exploits the fact that the effectiveness of a candidate's informative advertising is decreasing in the precision of the voter's prior information about the candidate. As Ansolabehere and Iyengar put it, "the voters who are most likely to learn from advertising are those who lack other sources of information." The influence of persuasive advertising, on the other hand, is unrelated to this precision of prior belief. By observing variation in the effectiveness of advertising as a function of variation in levels of prior information, I can measure the relative importance of the two effects.

Electoral politics provides two useful measures of voters' prior information in elections. First, incumbency status is associated with more precise information, as voters have had a chance to observe the incumbent's - but not the challenger's - past performance in office. Second, Snyder and Strömberg (2010) show that the degree of overlap between political districts and newspaper circulation areas - which they call "congruence" - is an excellent predictor of voters' level of political knowledge and engagement. I exploit the marginal effects of incumbency status and newspaper congruence on advertising effectiveness to separately identify persuasive and informative effects of advertising.

While understanding the relative importance of the two channels is valuable on its own, the main questions of interest motivating this paper involve the effects of advertising on

the functioning of the democratic process. For this reason I develop a structural model of vote choice, analogous to modern methods of demand estimation in product markets, and estimate it in a sample of state-wide (Senate and Governor) races. With estimates of the structural parameters in hand, I can explore the consequences of various alternative regulatory scenarios, including advertising bans and public financing of elections.

The results show that political advertising appears to have measurable informational content. However, the influence of this information provision on behavior is small. The variance of voters' posterior beliefs - a measure of the uncertainty voters face when choosing their vote - falls by less than 10% on average as a result of exposure to advertising. The impact of the informative component of advertising on vote choice is typically an order of magnitude less than that due to the persuasive effect.

Both types of advertising effect, though statistically significant, are small enough in magnitude that even dramatic policy changes are unlikely to lead to significant changes in election outcomes, at least in elections involving an incumbent candidate.[2] In large part, this is a result of how few non-open-seat Senate and Governor elections are close races. In only a few marginal elections is it possible that symmetric restrictions on advertising - such as outright bans or expenditure caps - could swing the election. The estimated effects are large enough, however, that large asymmetries in spending could be decisive, and hence individual candidates have no incentive to unilaterally "disarm" and stop advertising. Conversely, there is still reason for concern about potential asymmetries in advertising levels that may arise in an unlimited, post-*Citizens United* world.

The remainder of the paper proceeds as follows. Section 2 describes related work and connects the present paper to past theoretical and empirical results in the literature. Section 3 develops the empirical model and outlines the approach to identification of its parameters. Section 4 describes the data, which consists primarily of advertising levels and election outcomes in 83 Senate and Gubernatorial races in the 2002 and 2004 general elections, at the level of the television market.[3] Section 5 presents the parameter estimates, and section

[2]The counterfactual analysis in this paper presumes a fixed pool of candidates. It is possible, of course, that such large policy changes could change the pool of potential candidates, for instance by altering the relative value of fundraising ability as compared to other political skills. Analysis of this type of compositional effect is well beyond the scope of the present paper but nonetheless an intriguing theoretical question.

[3]Television markets are typically defined to include a central city and its metropolitan area. Hence, a

6 describes the results of several counterfactual experiments. Section 7 concludes.

2 Related Literature

Models of advertising in political economics tend to fall into one of three camps, analogously to the classes of models of advertising in product markets described by Bagwell (2007). The primary distinction between political and product advertising that these models account for is that political advertising is typically not paid for directly by candidates, but rather is financed by contributions from a third party with distinct and perhaps competing interests. Aside from this complication, the modeling technology and assumptions are quite similar to standard microeconomic treatments of advertising, and can be similarly grouped according to their assumed information structure.

In the first group, advertising reveals verifiable information to voters. Advertising messages cannot be falsified; candidates are able only to make voters aware of facts that they may not have been previously. An example of this approach is Ashworth (2006). In Ashworth's setup, advertising is costly, and hence even a candidate who truly is better than voters believe her to be may choose not to advertise. Advertising expenditures must be financed by outside groups, who demand favors in exchange for their financial support. When voters observe advertising, they trade off the positive signal it reveals against the knowledge that the message must have been paid for by promises of favors to special interests.

In the second group, political advertising is non-verifiable, but nonetheless reveals information to voters in equilibrium. The actual content of advertising messages is cheap talk and hence non-credible, but voters can still learn something from the quantity of messages they observe. An example of this type of model is Prat (2002), in which an insider receives an early signal about the relative quality of the candidates, and then chooses to make a campaign contribution to either candidate. Candidates in turn use the insider's contribution to finance mass-media advertising that voters observe. Because the insider extracts promises of favors from candidates on which he can collect only if the candidate wins, he has an incentive to direct his contributions to the higher-quality candidate. The equilibrium level of

state-level race will typically encompass multiple media markets. I use within-race variation in advertising levels to identify the parameters.

advertising thus reveals the insider's signal to voters.

In the last group, political advertising contains no information and is purely manipulative. In Baron (1994), some fraction of voters are "uninformed" and respond to advertising in purely mechanical fashion: higher levels of advertising by one candidate increase these voters' propensity to vote for that candidate. The remaining "informed" voters are already partisans of one candidate or the other, and are unaffected by advertising. The effectiveness of political advertising in this setup thus depends primarily on the preponderance of the uninformed voter type in the electorate.

In all of these models, the welfare gain (if any) to the electorate from advertising comes from improved access to information about candidate quality and the resulting improvement in the average quality of elected candidates. The downside is the potential that this information may be paid for by the diversion of public resources to the donor groups to whom candidates are indebted. Ansolabehere, Figueiredo, and Snyder (2003) argue that this assumption that donors demand some return on their investment is flawed. They instead argue for a view of political contributions as a (normal) consumption good, analogous to charitable contributions; the benefits that donors expect to receive are strictly of the personal, "warm glow" variety, and do not come at the expense of the broader public.

Even if contributions are strictly non-instrumental, however, the fact that wealthier people tend to contribute more can still bias policy outcomes in favor of the wealthy. Campante (2011) shows that a sufficiently unequal distribution of income can make politicians dependent on a small group of very wealthy people for the financial support necessary to remain competitive in elections. To maintain the support of these large donors, politicians must reduce the degree of redistribution in their platforms relative to the majority-preferred level.

Hence, if advertising is purely persuasive, in the sense described above, private campaign finance is likely to be welfare-decreasing for the median voter - even in the absence of a direct money-for-influence relationship between politicians and donors. Conversely, the existence of some informational value of campaign advertising is a necessary, though not sufficient, condition for private campaign finance to be welfare-improving.

A few papers attempt to answer this question empirically, though most existing empirical work focuses on identifying overall advertising effects rather than distinguishing alternative mechanisms of influence. Huber and Arceneaux (2007) analyze voter response to advertising

in presidential elections using survey responses. They exploit the discontinuities at state lines created by the electoral college to separate causal effects of advertising from selection bias induced by the targeting strategies of presidential candidates. They find little evidence that advertising influences voters' assessments of candidates' policy positions, but significant evidence of effects on voter perceptions of candidates' personal characteristics. Ashworth and Clinton (2007), using a similar empirical strategy, find little evidence of an effect of advertising on turnout. Gerber, Gimpel, Green, and Shaw (2011) conduct a randomized field experiment within a single gubernatorial campaign, and measure voter response using a panel of polling data. They find that advertising effects are measurable but very short-lived.[4] This paper differs from these earlier works in that it estimates a structural model that allows separate identification of multiple channels of advertising influence. I focus on actual voting behavior, rather than survey or poll responses, and exploit cross-candidate variation that is unavailable in a single-election study. Finally, I employ new instrumental variables for advertising levels that allow me to broaden the scope of the analysis to cover the entire country, rather than a limited number of presidential "battleground" states.

Another empirical work closely connected to the present paper is Prat, Puglisi, and Snyder (2010), who look for indirect evidence of informational content of campaign advertising. Using a measure of legislator effectiveness in the North Carolina state legislature, they ask whether the amount of campaign funds raised is correlated with candidate quality. They find little overall correlation, suggesting that the amount that voters could learn by observing candidates' levels of advertising in a campaign is limited. I look for direct evidence of information provision in voters' responses to advertising funded by campaign contributions; though the data sets and empirical methods are quite different, my results are generally consistent with their conclusions.

3 Identification of Informative and Persuasive Effects

This section develops an empirical model and strategy to distinguish the first two types of advertising described in section 2, which I will refer to as "informative," from the third, which

[4]Interestingly, candidates in my sample seem either to not believe this result or to be acting very suboptimally, as they begin advertising at significant levels several months before the election: see Figure 2.

I refer to as "persuasive." The approach is a structural one, which builds on the modeling approach of Ackerberg (2003), who examines advertising in the market for single-serving yogurt. Ackerberg's insight is that the influence on behavior of informative advertising is decreasing in the precision of the consumer's prior beliefs about a product. At some point, the consumer learns all there is to know about a product and further information no longer influences his decision. Persuasive advertising affects consumers' utility directly and hence is unaffected by the precision of prior beliefs.

This insight applies not only to advertising in product markets, but to political advertising as well. Most elections in the US feature a contest between two candidates, one an incumbent, and the other a challenger. Voters, particularly the large majority who do not pay close attention to politics, can be expected to have better information about the incumbent than the challenger, due to having observed her past performance in office. The purely retrospective voter (as in, e.g. Ferejohn 1986) can learn something, however noisy and imperfect, about the competence of the incumbent by assessing his current situation. Such self-assessment tells him nothing about the quality of the challenger. Because voters' prior information about the incumbent is more precise, observing the relative effectiveness of challenger versus incumbent advertising tells us something about the importance of information provision in campaign advertising.

For intuition, consider the following idealized example. Imagine a voter whose only political knowledge is the fact that, in the time since the current incumbent took office, the factory where the voter previously worked has shut down and relocated overseas, leaving him out of work. Subsequently, the voter observes two advertisements that reveal the identical (and for the sake of argument, verifiable) piece of information about both the challenger and the incumbent - say, that both were chief executives of successful companies before beginning their political careers. When forming his expectations about the incumbent's future job performance, the voter weighs this new piece of information against the fact that her tenure in office thus far has not been good for his job prospects. In the challenger's case, the information about career history is all he has to go on. Hence, we can expect that the information provided in the ad will have a larger relative impact on the voter's posterior beliefs about the challenger than about the incumbent.

In contrast, consider the situation when the advertisements contain no information. For

example, each ad is just an image of the sponsoring candidate shaking hands with a smiling constituent; furthermore, the voter knows that both ads were paid for by an independent nonpartisan group whose only goal is to increase civic-mindedness and trust in government.[5] The effectiveness of this kind of advertising in persuading voters to vote for one candidate or the other is unaffected by the prior informational environment - that is, the fact that the voter has observed the incumbent's performance in office but not the challenger's.

In addition to this across-candidate variation, I also take advantage of variation in the quality of voters' prior information across different media markets within the same election. Snyder and Strömberg (2010) construct a measure of how well newspaper circulation areas fit into political districts, which they call "congruence." They show that this measure strongly affects the level of press coverage that representatives receive and, consequently, is an excellent predictor of how much voters in a district know about their representatives. Again, we can expect that higher-congruence areas in a statewide election, having better prior information, will update less and hence respond less strongly to informational advertising.

Hence, observing how advertising effectiveness varies with variation in voters' prior information reveals the relative importance of information provision in campaign advertising. The following subsections develop an empirical model that formalizes this notion. The model is a signaling game between voters and candidates in which voters attempt to learn about the candidates' relative quality by observing advertising signals. I first describe the model primitives and game sequence, and then discuss the identification of its parameters.

3.1 Model Primitives

I model an election between two candidates who compete for the support of a set of voters, indexed by i. Voters are partitioned into a set of media markets indexed by j; each voter resides in one and only one market. I refer to a generic candidate by the subscript k.

Candidates have three salient characteristics that voters care about. First, voters have

[5] I invoke this somewhat implausible scenario in order to emphasize that the notion of information in this paper is not restricted to verifiable information contained in the literal message of an advertisement, but also includes indirect learning of the type in Prat (2002). To be information-free in this sense it is necessary both that the ad's message be free of verifiable information, and that the voter not be able to infer anything about the candidate from the fact of the ad's existence.

preferences over the candidate's policy position p_k. For now I will leave the policy space in which p_k resides abstract. Second, voters care about the candidate's "quality" or "valence," ξ_k. Quality is a real number, higher levels of which are strictly preferred by all voters in the electorate. Legislator quality may be interpreted as the ability to provide constituency services or represent the district's particular interests in legislative bargaining, as in the theoretical treatments of Londregan and Romer (1993) or Groseclose (2001). Lastly, voters may care directly about the level of advertising that they observe supporting or opposing a candidate.[6] This term captures direct persuasive effects of advertising, unrelated to learning or information transmission.

Advertising is market-specific: voters in one market may observe different levels than do voters in another. I refer to the levels of advertising in a market j sponsored by a given candidate k by a_{jk}.[7] If all three of these parts were perfectly known to the voter, then the voter's utility function would be:

$$u_{ijk} = \beta_{0,i} + \beta_{1,i} p_k + \beta_{2,i} a_{jk} + \xi_k + \epsilon_{ijk} \tag{1}$$

Where ϵ_{ijk} is an iid logit error observed by the voter but unobserved by the econometrician. With the addition of an outside option - not voting - whose utility is normalized to zero, this utility specification defines a random coefficients logit model of demand in the style of Nevo (2000) or Berry, Levinsohn, and Pakes (1995).

This formulation implicitly assumes that voters vote "expressively," i.e. they choose the option (either vote for one of the candidates or don't vote) that gives the highest utility, without regard to strategic considerations. In the large elections that I study here, an individual voter has effectively zero probability of affecting the outcome, and hence models

[6] Another important distinguishing feature of political versus product advertising, which is usually neglected in theoretical treatments, is that candidates often run ads which only discuss the opposing candidate, in a negative frame ("attack ads"). In some specifications I will distinguish between promotional and attack advertising and allow each type to have a different informational content or persuasive effect.

[7] Again, in some specifications I will separate promotional and attack advertising and refer to these by a^P_{jk} and a^A_{jk}, respectively. The attack ads related to candidate k, a^A_{jk}, will be paid for by the opposing candidate $-k$. The derivations that follow aggregate both types of advertising together; all extend directly to the case where the effects of promotional and attack advertising may differ. Additionally, In some specifications I will include higher-order terms in the advertising levels to allow for nonlinear effects of advertising. These are omitted here for clarity of exposition.

of vote choice based on pivot probabilities will tend to predict negligible turnout. The expressive model has the advantage of both analytical simplicity and empirical plausibility. The constant term $\beta_{0,i}$ in this interpretation represents a utility cost associated with the act of turning out to vote. To make estimation of equation (1) feasible, I will restrict the random coefficients to vary with observable demographic characteristics in a manner described in section 3.3.

On the supply side of the market, candidates' preferences are simple: they want to win office. I will assume that there is no positional competition among candidates; a candidate's policy position is exogenous. This is not as strong a restriction as it first seems. In the empirical application I will use only within-race variation to identify the parameters. Hence, the mean differences between candidates - including the mean policy position - will be absorbed by the candidate fixed effects. All that I require is that the candidates' perceived policy platforms do not vary across markets within a race.

Candidates similarly cannot choose their quality, ξ_k; this is an exogenous attribute. The candidate's problem thus reduces to choosing how to deploy his advertising budget in order to maximize his likelihood of victory. This endogeneity of advertising allocations across markets introduces an identification problem in the estimation of advertising effects, as candidates may choose to allocate their advertising budgets preferentially to those markets which for unobserved reasons are more likely to favor (or oppose) the candidate. For instance, in the characterization of equilibrium in a simplified version of this problem given by Snyder (1989), candidates concentrate their resources on those markets where voters are closest to being evenly split. Thus a naive estimate of advertising effects would tend to overstate the effects for underdog candidates, and understate them for favorites. I return to this issue, and how to address it, in section 3.3.

3.2 Sequence

Equation (1) describes the voter's problem under full information. In order for advertising to be potentially informative, however, it must be the case that the voter is initially uncertain about one or more of the candidate's attributes, and therefore might learn something from observing advertising levels. I will model the voter as having full information about the

candidate's policy position, which I take to be fully revealed by the candidate's party label. However, he has incomplete information about each candidate's quality ξ_k.[8] The voter has a prior belief about the joint distribution of advertising levels and candidate quality, and hence may update his posteriors about the latter upon observing the former.

The sequence of events in the voter's learning process is as follows. Initially, the voter's prior belief about the joint distribution of ξ_k and a_k is that they follow the joint normal distribution:

$$\begin{pmatrix} \xi_k \\ a_k \end{pmatrix} \sim N\left(m_0, \Sigma_0\right) \tag{2}$$

Where $m_0 \equiv (m_0^{\xi}, m_0^{a})'$. The prior mean m_0 represents voters' expectations about the average candidate's quality, as well as how much such a candidate would advertise.[9] In the estimation section, I will allow the voter's prior on quality to vary by market and candidate: m_0^{ξ} will be replaced with $m_{0,jk}^{\xi}$. Variation in $m_{0,jk}^{\xi}$ captures the idea that certain markets may be, for unobserved reasons, predisposed to support one candidate over another; candidates may potentially use knowledge of these market-level variations in prior preferences in setting their advertising strategies. In some specifications I will also allow the prior expectations on advertising levels to vary by market, though in more restrictive parametric fashion.[10]

The prior covariance matrix Σ_0 describes the voter's belief about the relationship between advertising levels and candidate quality. Nonzero values of the off-diagonal elements of Σ_0 indicate that the voter believes that advertising reveals something about the candidate's quality. Such a relationship could hold due to either of the theoretical mechanisms discussed in section 2: either because advertising directly reveals verifiable information about the

[8]This formulation of uncertainty is consistent with Huber and Arceneaux (2007)'s empirical finding that advertising has no effect on voter perceptions of candidates' policy positions, but significant effects on perceptions of personal qualities.

[9]One way to think about this is to suppose that the voter has a model in her mind that describes how advertising is related to quality, e.g.: $a = f(\xi)$. The prior mean on ad levels m_0^a then represents the average candidate's advertising levels, e.g. $f(m_0^{\xi})$.

[10]Again, here we can think of the voter as having a model of advertising as a function of quality, though now with an additional level of sophistication: she may recognize that some markets are smaller or more costly than others and hence expect to see less ads in those markets than in bigger or cheaper markets. In this case, we have $a_j = f(\xi) + \epsilon_j$ and $m_{0,j}^a = f(m_0^{\xi}) + \epsilon_j$.

candidates, or because the underlying campaign finance game between knowledgeable donors and candidates indirectly reveals quality in equilibrium.

In general the model described here will not be able to distinguish between these alternative mechanisms of information transmission. However, rationality on the part of candidates would tend to rule out negative off-diagonals of Σ_0 in the case of "hard" information, as candidates would be unlikely to voluntarily reveal negative information about themselves. Such a relationship is possible in the "soft" information case if, for instance, voters believe high-spending candidates to be corrupted by their dependence on large donors. If the off-diagonals of Σ_0 are zero, then advertising has no informational content, and voters do not update their beliefs about candidate quality upon observing campaign ads.

Next, voters observe a noisy signal of the quality of each candidate. This signal can be interpreted as the individual voter's assessment of the candidate's likely performance in office. Signals are unbiased and on average reflect the candidate's true quality,[11] but are subject to an additive normal shock specific to each individual. The quality signal observed by an arbitrary voter i is thus:

$$\xi_{i,k} = \xi_k + \eta_k^{\xi}, \quad \eta_k^{\xi} \sim N(0, \sigma^2_{\xi,jk}) \tag{3}$$

After voters observe quality signals, candidates choose advertising levels in each of the media markets within the district. Voters observe the level of advertising in their home market, again with additive normal error:

$$a_{ijk} = a_{jk} + \eta_{ijk}^{a}, \quad \eta_{ijk}^{a} \sim N(0, \sigma^2_{a,jk}) \tag{4}$$

Following their observation of the advertising signals and the performance signal, voters update beliefs in Bayesian fashion. The resulting posterior beliefs are also distributed normally.[12] Posterior beliefs are normal with mean and variance given by:

[11] Unbiasedness implies that candidates cannot manipulate their performance signal.

[12] This section adapts the derivation of posteriors in Ackerberg (2003).

$$m_{ijk} = m_{0,jk} + \left(\Sigma_0^{-1} + \Phi_{jk}^{-1}\right)^{-1} \Phi_{jk}^{-1} \begin{pmatrix} \xi_{ijk} - m_{0,jk}^{\xi} \\ a_{ijk} - m_{0,jk}^{a} \end{pmatrix} \quad (5)$$

$$\Sigma_{jk} = \left(\Sigma_0^{-1} + \Phi_{jk}^{-1}\right)^{-1} \quad (6)$$

Where Φ_{jk} is the diagonal matrix with diagonal entries $(\sigma^2_{\xi,jk}, \sigma^2_{a,jk})$. The performance signals induce both a mean shift and a decrease in the variance in the voter's beliefs about each candidate's type. The variance reduction is greater for candidate-market combinations which have smaller values of $\sigma^2_{\xi,jk}$, e.g. smaller error variances in the quality signal. This reduction leaves proportionately less room for candidates' advertising to influence beliefs. Variation in $\sigma^2_{\xi,jk}$ across candidates and across markets thus produces variation in the marginal effectiveness of advertising; it is this variation that identifies the relative strength of the informative component.

Finally, voters choose the option (of voting for either candidate or not voting) that gives maximum expected utility, given their beliefs about each candidate. This is equivalent to making the choice that maximizes a version of equation 1 where the full-information quantity ξ_k is replaced by the mean of the posterior distribution defined above, and a_{jk} is replaced by the voter-specific value a_{ijk}. The expected utility becomes:

$$u_{ijk} = \beta_{0,i} + \beta_{1,i} p_k + \beta_{2,i} a_{ijk} + m_{ijk}^{\xi} + \epsilon_{ijk} \quad (7)$$

3.3 Estimation Procedure

The voting model defined thus far yields predicted likelihoods of voting for each candidate (plus the outside option of not voting), for every voter in a given district. Given knowledge of the random coefficients β_i, these can be computed by solving each individual's utility maximization problem conditional on the draw of the individual error terms $\epsilon_{ik}, \eta_{ik}^{\xi}, \eta_{ik}^{a}$. Of course, voter privacy laws imply that we cannot hope to obtain individual-level choice data. Fortunately, the model's parameters can be identified from aggregate vote data alone.

To make estimation from aggregate data feasible, it is first necessary to impose some restrictions on the random coefficients β_i. Following Nevo (2000), I take the random coeffi-

cients to be decomposable into a mean component β and an individual-specific shift that is a function of the individual's demographic or political characteristics, D_i:

$$\beta_i = \beta + \Pi D_i \tag{8}$$

Suppose β and D_i are b- and d-dimensional vectors, respectively. Then Π is a $(b \times d)$ matrix of coefficients, whose elements represent the marginal effect of changes in each demographic variable on an individual's preference for each candidate characteristic.

Each voter's utility from voting for a candidate k is thus the sum of two terms: a mean utility δ_{jk} common to all voters in market j, and an individual-specific term μ_{ijk}. We have:

$$\begin{aligned}
u_{ijk} &\equiv \delta_{jk} + \mu_{ijk} && (9)\\
\delta_{jk} &= \beta_0 + \beta_1 p_k + \beta_2 a_{jk} + \bar{m}^{\xi}_{jk} && (10)\\
\mu_{ijk} &= \beta' \begin{pmatrix} 0 \\ 0 \\ \eta^a_{ijk} \end{pmatrix} + (\Pi D_i)' \begin{pmatrix} 1 \\ p_k \\ a_{jk} + \eta^a_{ijk} \end{pmatrix} + \alpha_{jk} \begin{pmatrix} \eta^{\xi}_{ijk} \\ \eta^a_{ijk} \end{pmatrix} + \epsilon_{ijk} && (11)
\end{aligned}$$

Where $\bar{m}^{\xi}_{jk}$ is the mean perception, after exposure to advertising, of the quality of candidate k in market j. This quantity is a function of the learning parameters Σ_0, m_0, Φ as well as the advertising levels a_{jk}:

$$\bar{m}^{\xi}_{jk} = m^{\xi}_{0,jk} + \alpha_{jk} \begin{pmatrix} \xi_k - m^{\xi}_{0,jk} \\ a_{jk} - m^{a}_{0,jk} \end{pmatrix} \tag{12}$$

The unobserved differences in priors on quality $m^{\xi}_{0,jk}$ in (12) will ultimately enter the structural error term. The priors on ad levels $m^{a}_{0,jk}$ will in different specifications be taken to either be constant across markets, in which case they are absorbed by the candidate fixed effects, or a function of the share of the state population residing in the market.[13] The row vector α_{jk} in (12) and (11) is the first row of the matrix A_{jk}, where:

[13]The latter specifications presume that voters expect candidates to allocate ads to markets in proportion to the market's share of the state population.

$$A_{jk} = \left(\Sigma_0^{-1} + \Phi_{jk}^{-1}\right)^{-1} \Phi_{jk}^{-1} \tag{13}$$

Let $P_j(D)$ be the population distribution of demographics in market j. Similarly, let $P(\eta)$ and $P(\epsilon)$ be the population distributions of the individual specific error terms $\eta_{ijk}^{\xi}, \eta_{ijk}^{P}, \eta_{ijk}^{A}$, and ϵ_{ijk}. Then utility maximization implies predicted market shares for each candidate in each market. Denoting by V_k the set of realizations of $(D_i, \epsilon_{ijk}, \eta_{ijk}^{\xi}, \eta_{ijk}^{a})$ for which k is the preferred choice, the predicted share of candidate k in market j is given by:

$$s_{jk} = \int_{V_k} dP_j(D) dP(\eta) dP(\epsilon) \tag{14}$$

Under the assumption that the ϵ's follow an iid type 1 extreme value distribution, the outermost part of the integral can be evaluated analytically. The remaining integrals over $P_j(D)$ and $P(\eta)$ will be approximated by simulation. From the previous section, I take $P(\eta)$ to be a normal distribution, and hence easy to simulate from. I simulate draws from $P_j(D)$ without imposing parametric assumptions by means of a bootstrap procedure: given a sample of individuals from market j, I resample with replacement an arbitrary number of times. Denote the simulated draws of the first three components of μ_{ijk} (excluding the logit error ϵ_{ijk}) by μ'_{ijk}. Substituting in the definition of utility, using the analytical form to evaluate the outer integral, and replacing the inner integrals with the simulated approximation, yields:

$$s_{jk} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{e^{\delta_{jk} + \mu'_{ijk}}}{1 + \sum_l e^{\delta_{jl} + \mu'_{ijl}}} \tag{15}$$

I use the contraction-mapping procedure of Berry, Levinsohn, and Pakes (1995) to estimate the δ_{jk}'s by matching estimated vote shares to observed vote shares for each candidate in each market. Denoting the fitted values from this procedure by $\hat{\delta}_{jk}$, I define the structural error:

$$\omega_{jk} \equiv \hat{\delta}_{jk} - \left(\beta' \begin{pmatrix} 1 \\ p_k \\ a_{jk} \end{pmatrix} + \alpha_{jk} \begin{pmatrix} \xi_k \\ a_{jk} - m_{0,jk}^{a} \end{pmatrix} \right) \tag{16}$$

Note that the equation for ω_{jk} involves both the (unobserved) true quality ξ_k and the

candidate's policy position p_k. Ignoring these omitted variables could bias my estimates of advertising effects if either variable is correlated with advertising levels; for instance, if higher-quality candidates are able to raise more money to spend on advertising. One way to deal with this, following Nevo (2000), would be to estimate a dummy variable for each candidate - the equivalent of a "brand dummy" in the more typical consumer-products application. But given the nature of the dataset, which has many candidates and only a few markets per candidate, this approach is undesirable. Estimates of the candidate dummies are only consistent as the number of markets in a race tends to infinity, a condition which in this setting is clearly violated. Instead, I apply a fixed-effects transformation to the structural errors to eliminate the quality term, which is common across markets within a race:

$$\tilde{\omega}_{jk} = \omega_{jk} - \frac{1}{|L_k|} \sum_{l \in L_k} \omega_{lk} \tag{17}$$

Equation (17) also eliminates the candidate characteristics that do not vary across markets within a race. A benefit of this restriction is that measurement error resulting from using proxies for the difficult-to-quantify ideology measure p_k will no longer enter the error term. On the downside, we can no longer identify the policy coefficient β_1 or the turnout cost β_0. The demographic interactions on these variables in Π, however, remain identified. What is left behind by the transformation are cross-market variations in the prior predisposition to support one candidate or another - in the terms of the learning model, deviations from the state-wide average in $m_{0,jk}$:

$$\tilde{\omega}_{jk} = (1 - \alpha_{jk}^{(1)}) m_{0,jk} \tag{18}$$

$$\hat{m}_{0,jk} = \tilde{\omega}_{jk} / (1 - \alpha_{jk}^{(1)}) \tag{19}$$

Strictly speaking, equations (18) and (19) hold exactly only when $\alpha_{jk}^{(1)}$ is constant within a candidate. This condition will hold when the information-shifting covariates include only candidate-level attributes, such as incumbency status or previous offices held. It will not in specifications that include geographic measures of prior information quality, such as newspaper congruence. However, even in these specifications, in practice the variation in $\alpha_{jk}^{(1)}$ is

very small relative to the estimated scale of the fixed effects. Hence, (19) is likely to be a very close approximation. I discuss the construction and identification of α_{jk} further in the next section.

Given the endogenous allocation of advertising levels by the candidates, one worry is that candidates may choose to allocate advertising to those markets which for unobserved reasons are more (or less) likely to support the candidate. This kind of endogeneity would make estimates derived from directly fitting observed to predicted market shares inconsistent: the estimated persuasive effects of advertising would be contaminated by selection. To get around this difficulty, I require instruments for advertising levels, z_j, which are uncorrelated with the demeaned structural errors $\tilde{\omega}_{jk}$. Given (19), this implies that the instruments must be uncorrelated with variations in prior predispositions towards one or another candidate across markets within a race. The $(2N \times r)$ matrix of instruments for all observations will be denoted by Z.[14] I discuss the choice of instruments in section 4.

Putting all of this together, we can now define the criterion function to be minimized. The sample moments that will be used to form the criterion function are the interactions between these instruments and the demeaned structural errors. The resulting function has the usual GMM quadratic form:

$$Q_N(\beta, \Sigma_0, \Pi, \Phi) = \hat{m}'_{0,jk} Z W_N Z' \hat{m}_{0,jk} \tag{20}$$

Where W_N is an $(r \times r)$, positive definite weighting matrix. I initially estimate W_N using the two-stage least squares form, $(Z'Z)^{-1}$. After computing an initial estimate of the parameter values by minimizing Q_N with this choice of W_N, I estimate the optimal weighting matrix W_N^* as $\frac{1}{2N} \sum_j \sum_{k=1}^{2} \hat{m}_{0,jk}^2 z_j z_j'$, where $\hat{m}_{0,jk}$ are the estimated structural errors at the first-stage parameter estimate.

Finally, to compute confidence intervals, I implement the Laplace-Type Estimator (LTE) approach of Chernozhukov and Hong (2003). The LTE approach uses a Markov chain monte carlo (MCMC) method which replaces the standard likelihood function with a transformed version of the objective function in (20). The resulting MCMC draws can be used to construct

[14] I use N to denote the total number of market-race observations in the sample. As there are two candidates per race - one a Democrat, one Republican - the total number of vote share observations available for estimation is $2N$.

estimates and confidence intervals for the parameters of the model. Using the optimal weighting matrix (the estimate of W_N^* described above) in this process allows construction of asymptotically valid confidence intervals by simply taking quantiles of the MCMC draws.

3.4 Identification

At this point, a discussion of identification of the model parameters is in order. I discuss each group of parameters - advertising coefficients, variance parameters, and the remaining learning parameters - in turn.

3.4.1 Advertising and demographic coefficients

First, the elements of β corresponding to candidate attributes that vary by market are identified by exogenous variation - generated by the instrumental variables - in within-race levels of advertising. This category includes the coefficients representing the persuasive effects of advertising. As noted previously, coefficients on candidate attributes that do not vary with market, including all policy-related characteristics, are differenced out by the estimator and cannot be identified.

Second, informational effects of advertising - the off-diagonal elements of the leftmost column of Σ_0 - are identified by the marginal effects of incumbency and newspaper congruence on advertising effectiveness. Changes in advertising levels move the mean posterior on quality $\bar{m}_{jk}^{\xi}$ differently depending on the values of these information-shifter variables. Identification of these marginal effects does not rely on the functional form chosen for the main advertising effects; it is possible to include higher-order terms on advertising levels in β to allow for diminishing marginal returns on the persuasive effect. I include such higher-order terms in several specifications.

Third, the demographic terms in Π are identified by cross-market and cross-race variation in the distribution of demographic characteristics. This includes coefficients on characteristics which do not vary within a race. For instance, if some characteristic in D_i is reliably associated with a preference for Democratic candidates, then the demeaned structural errors in (17) will be reduced on average by a positive coefficient in the element of Π corresponding to the interaction of that characteristic with a dummy for Democratic partisanship.

3.4.2 Quality signal variance coefficients

As described above, the informational effects of advertising on voters' beliefs - parameterized by the covariance matrix Σ_0 - are identified by variation in advertising's influence on vote share as a function of variation in levels of newspaper congruence and incumbency status. In the learning model, these variables enter via the quality signal variance, $\sigma^2_{\xi,jk}$. In the absence of a normalization, it is not possible to separately identify the influence of congruence and incumbency on this variance along with the other learning parameters in Σ_0. I bring in some additional data to tie down the relationship between the information-shifting variables and the learning parameter $\sigma^2_{\xi,jk}$.

I use data from the Cooperative Congressional Elections Survey (CCES), a large-scale survey of voters in all 50 states. Among other things, the CCES asked voters to state their approval of the incumbent Senators and Governor in their state, on a four-point scale ranging from "strongly disapprove" to "strongly approve." I constructed a measure of variance in quality signals by estimating the variance in this measure of approval, conditional on voters' self-reported ideology, partisan affiliation, and demographic characteristics.

The levels of the CCES' approval variable are arbitrary, but they do have a natural ordering. I therefore treated the discrete approval ratings, which I label r_i, as thresholds of an unobserved latent continuous approval measure, $\tilde{r}_i$. I used the ordered probit framework of Hausman, Lo, and MacKinlay (1992), and estimated both the thresholds and coefficients using their maximum likelihood estimator. Specifically, the model is $\tilde{r}_i = \alpha X_i + \eta_i^{\xi}$, $\eta_i^{\xi} \sim N(0, \sigma^2_{\xi,i})$, where:

$$r_i = \begin{cases} 0 & \text{if } \tilde{r}_i < \lambda_1 \\ 1 & \text{if } \lambda_1 \leq \tilde{r}_i \leq \lambda_2 \\ 2 & \text{if } \lambda_2 \leq \tilde{r}_i \leq \lambda_3 \\ 3 & \text{if } \tilde{r}_i > \lambda_3 \end{cases} \tag{21}$$

Here, λ is a vector of estimated cutoff values, and α is a vector of coefficients on observables. X_i consists of variables measuring the ideological and partisan similarity of the respondent to the incumbent being evaluated: four dummy variables for the combination of respondent and incumbent partisanship, and one continuous measure of ideology. The ideol-

ogy measure is the absolute difference between the respondent's self-reported conservative-liberal rating and the conservative-liberal rating the respondent gives to the incumbent.[15] I also included a few demographic characteristics of the individual respondent in X_i, of which only one, unemployment status, appears to have a significant (and, as might be expected, negative) effect on respondents' average approval ratings. Estimates of (α, λ) are in table 1; the first two columns are OLS approximations where the integer approval ratings r_i are simply taken to be the dependent variable, and the third column is the full MLE latent approval model where both coefficients and cutoff values are jointly estimated.

As in Hausman, Lo, and MacKinlay (1992), I allow the error variances $\sigma^2_{\xi,i}$ to be heteroskedastic. I estimate the variance as a function of newspaper congruence in the respondent's media market, the office held by the politician being evaluated (Senate or Governor) and the respondent's demographics:[16]

$$\sigma^2_{\xi,i} = 1 - \gamma^2 W_i \tag{22}$$

Here, W_i is a vector of demographics plus newspaper congruence in the respondent's media market. The parameter of interest is the coefficient on newspaper congruence, γ^2_C. I report estimates of this parameter in the main results tables in section 5. I estimate it to be significantly different from zero, indicating that the residual variance in voters' assessments of the quality of their incumbent Senators and Governors is reduced in markets with higher congruence.

In the main estimation routine described above, I construct the DMA-state-candidate level signal variance using the congruence parameter γ^2_C just described along with an additional shifter for incumbency status, γ^2_I. The effect of incumbency status on the precision of voter beliefs can now be identified in the main estimation routine, relative to the impact of newspaper congruence. Putting these together, we have:

$$\sigma^2_{\xi,jk} = 1 - \gamma^2_C C_j - \gamma^2_I I_k \tag{23}$$

[15] These are both measured on 100-point scales, with 0 being extremely liberal and 100 being extremely conservative. Respondents were told to make these ideological assessments on a scale where "the average American" scores a 50.

[16] This method requires a normalization; I normalize the constant in the variance equation to 1.

Coefficient	(1)	(2)	(3)
Democratic Voter - Democratic Incumbent	0.518	0.498	0.709
	(0.008)	(0.025)	(0.036)
Democratic Voter - Republican Incumbent	-0.217	-0.187	-0.285
	(0.008)	(0.024)	(0.026)
Republican Voter - Republican Incumbent	0.345	0.380	0.442
	(0.008)	(0.032)	(0.028)
Republican Voter - Democratic Incumbent	-0.042	-0.085	-0.061
	(0.009)	(0.043)	(0.026)
Ideological Distance	-0.023	-0.023	-0.033
	(0.000)	(0.000)	(0.001)
Assessing a Senator	0.065	0.070	
	(0.006)	(0.057)	
Nonwhite	0.022	0.025	
	(0.007)	(0.015)	
College-Educated	0.008	0.005	
	(0.007)	(0.013)	
Employed	-0.010	-0.009	
	(0.006)	(0.009)	
Unemployed	-0.065	-0.058	-0.120
	(0.004)	(0.005)	(0.050)
Income $> 75K$	0.010	0.015	
	(0.006)	(0.011)	
Intercept	3.17		
	(0.008)		
λ_1			-1.99
			(0.077)
λ_2			-1.11
			(0.046)
λ_3			0.19
			(0.019)
Politician Fixed Effects?	No	Yes	No
Model	OLS	OLS	MLE

Standard Errors in Parentheses. N = 83,072

Table 1: Regression of approval ratings on ideological assessments and partisan affiliation.

3.4.3 Advertising signal variance coefficients

The remaining variance term is that of the advertising signals $\sigma^2_{a,jk}$. This can be found by the following simple procedure: noting that the unit of advertising quantity is impressions, assume that each individual in the population of each market is equally likely to be the "consumer" of a given impression. Then the quantity of advertising "consumed" by each member of the population follows a binomial distribution, with variance $a_{jk}(P_j - 1)/(P_j)^2$.

I use this variance estimate to scale the individual advertising errors η^a_{ijk}. By design, the distribution of these errors varies across markets j and candidates k within a race.

However, an estimate of the ad signal variance is also necessary for the construction of voters' posterior beliefs, as it enters the inverse variance matrix Φ^{-1}. For this purpose, within-race variation in $\sigma^2_{a,jk}$ is undesirable as, for instance, voters in markets that saw no ads would have undefined posteriors. I instead use a weighted average value of σ^2_a, computed at the election level, in the construction of posterior beliefs for all voters in a given election. Substantively, this restriction implies that voters are unaware that some markets may have higher *variance* across individuals in observations of advertising levels; all individuals in a given election believe that the ad signal they observe has the same variance. Voters are, of course, aware of cross-market differences in *means*.

3.4.4 Additional learning parameters

The remaining parameters to be estimated are the other elements of Σ_0: the top left element and the lower-right block. The top-left element represents voters' beliefs about the variance in quality ξ_k among the population of politicians. This parameter acts like the variance of the random constant term in a more standard random coefficients logit demand system; it scales up or down the effect of the random quality draw η^ξ_{ijk} on the voter's choice. As such, it is identified in the contraction mapping step. Each value of $\Sigma_{0,(1,1)}$ implies a different relationship between the observed vote shares and the estimated mean utilities $\hat{\delta}_{jk}$.

Finally, I estimate the lower-right block of Σ_0 directly from the data, under the assumption that voters have correct beliefs about the population variances of advertising levels. These parameters represent voters' belief about the population covariance matrix of levels

of promotional and attack advertisements.[17] I directly estimate this matrix by the sample covariance in levels of both types of advertising.

4 Data

In this section, I describe the dataset used to implement the estimation procedure described thus far. The major components required are data on advertising levels, election results, demographics, congruence, and instruments, all at the media-market level in elections containing multiple media markets. State-level offices (Governor and Senator) are geographically large enough that they typically include several media markets, and there a sufficient number of elections for these offices in each election cycle that it is feasible to use cross-sectional variation. I thus focus on these elections. My sample consists of the 105 Senate and gubernatorial races in the 2002 and 2004 elections which featured two major-party candidates. I next describe each of the components of the dataset in turn.

4.1 Advertising data

The most important component of the dataset is the levels of advertising chosen by each candidate in each of the media markets in the sample. My source for this data is the Wisconsin Advertising Project (WAP), which collected meticulous records of all network television advertisements[18] purchased by candidates in federal elections (House, Senate and Presidential races) as well as gubernatorial elections in the 2002 and 2004 election cycles.

An observation in the WAP data is an individual advertisement. The data includes numerous characteristics of each ad, including the time of day, station, and media market[19] in which the ad aired. The WAP also collects much useful information about the content of the advertising: whether it promotes the sponsoring candidate, attacks the opposing candidate, or "contrasts" the two; whether the focus of the message is policy or personal

[17] This lower right block will be a scalar in specifications where both types of advertising are combined together; in this case it is simply the population variance of advertising levels.

[18] In the top 100 media markets.

[19] I follow the WAP in using the "Designated Market Areas" (DMAs) defined by the Nielsen Media Research company as my media market definition throughout the paper. This is the industry standard market definition, widely used by buyers of television advertising. Figure 1 maps the geography of DMAs in the US.

characteristics; whether it cites endorsements from outside groups; and the estimated cost of the ad spot.

4.1.1 Summary statistics

There are roughly 940,000 observations of ads aired by candidates in my sample of races that are recorded in the dataset. Table 2 gives summary statistics for these advertisements, broken down between challenger and incumbent candidates. Inspection of the table reveals some significant differences between challengers and incumbents in their advertising strategies. Incumbents, for instance, are more likely to run pure promotional advertisements that do not mention the opposing candidate, whereas challengers are more likely to contrast themselves with the incumbent. Incumbents are more likely to cite policy specifics, whereas challengers focus on personal characteristics - likely because incumbents have a record of accomplishments that they can point to whereas challengers need to offer less specific promises of future achievement. Challengers are more likely to cite third-party endorsements.

Figure 2 plots the timing of ad purchases in elections, separately for challengers in open-seat races, challengers facing incumbent candidates, and incumbents. The figure shows the number of ads aired in the average campaign in each week leading up to the election. Open-seat races, particularly in Senate campaigns, feature the highest overall levels of ads. Not coincidentally, these races tend to be the most competitive. In races featuring and incumbent and a challenger, an interesting pattern emerges. Challengers tend to keep keep pace with or even out-advertise incumbents in the early weeks of the campaign, but towards the end incumbents rapidly increase their spending and overwhelm challengers in the last weeks of the campaign.

These summary measures are consistent with a world where voters have better prior information about incumbents than they do about challengers. Incumbents can afford to sit back and wait until the last minute to unleash their generally superior advertising resources, whereas challengers need to get their names out early if they want to be competitive. Challengers need to contrast themselves with the (better-known) incumbent, or associate themselves with other political figures whom voters know. Incumbents can expect that voters already know who they are, and can focus on extolling their own political accomplishments.

Variable	Value	Challenger	Incumbent
Ad Type	**Promote**	49.03	58.81
	Attack	28.39	29.58
	Contrast	22.14	11.41
Outside Support?	No	58.73	57.78
	Yes, a newspaper article	22.14	28.20
	Yes, other	12.46	8.34
Information Type	Personal characteristics	14.27	16.05
	Policy matters	47.77	53.44
	Both personal and policy	35.14	28.64
Endorsements?	No	90.52	92.27
	Yes	8.70	7.03
Program Type	News	42.73	40.34
	Non-news	57.27	59.66
	Total Cost	$398M	$190M
	N	637611	308237

Table 2: Summary of Advertising Characteristics by Candidate Status

4.1.2 Measurement issues

The three crucial features that the WAP data provides for my purposes are that it identifies the media market in which an ad aired, whether the ad promotes the sponsoring candidate or attacks the opponent, and the cost of the ad. A few manipulations are necessary to get this data into the form required for estimation.

Contrast ads The WAP classifies ads according to the main subject of the advertisement: ads may promote the sponsoring candidate, or attack the opposing candidate. The WAP also includes an intermediate third category, "contrast" ads, which fall somewhere in between. To limit the number of estimated parameters, I eliminate the intermediate category by assigning these ads to one or the other of the primary two categories. I assign them to the "promote" category if the WAP defines them as "more promote than attack," and similarly assign them to the "attack " category if they are described as "more attack than promote." Those contrast ads labeled "about equal promote and attack" are assigned to the "promote"

category.

Quantities Estimating the signaling model presented above requires a measure of the quantity of advertising that a candidate allocated to a given media market. The obvious measure of quantity - the total number of ad spots purchased - is inappropriate because, in general, spots are not a homogeneous commodity. For instance, a thirty second ad purchased on a highly-rated prime-time sitcom may be seen by an order of magnitude more viewers than the same thirty second ad run on a daytime soap opera.

The quantity that we want to measure is not seconds of advertising but *impressions*: the total number of viewers who can be expected to have seen the ad. While the WAP does not record impressions directly, they can be computed from the WAP data using the method described by Hartmann and Gordon (2011), which I follow. Because television advertising is sold by the ratings point,[20] dividing the total cost of the spot (recorded by the WAP) by the per-point price (available from sources described in section 4.5) and then scaling by the adult population of the media market yields an estimate of the number of viewers the spot reached. Posted prices on a per-1000-impressions basis are available by media market, program type,[21] and day part,[22] yielding a reasonably fine-grained match from advertising observations to the price used to compute quantity in units of impressions. The price data is described in detail in section 4.5.

4.2 Vote data

I collected voting data for the 105 Senate and gubernatorial races in 2002 and 2004 that featured two major-party candidates. Sources for voting data were the websites of state election boards and secretaries of state, as well as Congressional Quarterly's *Voting and Elections Collection*,[23] which gathers county-level election results for all federal races. I

[20] One ratings point represents 1% of the TV-watching population of the media market.

[21] Different prices prevail for news and non-news programs at the standard evening (5PM) and late (11PM) news slots.

[22] There are seven day parts, ranging from the relatively cheap daytime (9AM-4PM) to the most expensive prime-time (8PM-11PM) slots.

[23] `http://library.cqpress.com/elections/`

gathered voting data at the smallest geographic level available: in some cases precinct, in others state legislative districts, towns, or in the worst case counties.

I then aggregated voting results to the state-DMA level using geographic boundary files for the DMAs and the political divisions mentioned above. DMAs are geographically large and usually drawn to coincide with county boundaries, so in most cases there was little difficulty in matching each political division to a DMA. When, on occasion, DMA boundaries split a county and there was no finer-grained voting data available, I allocated the county's voting results to each DMA in proportion to the fraction of the county's population residing within each DMA.[24]

The number of DMAs in a state varies with the geographic and population size of the state. The largest and most populous states have the most DMAs within their borders: the state with the most markets is Texas, with 18. Some of the smallest states, like Rhode Island, have only one. Others have multiple DMAs but one or none that are in the top 100 media markets (by population) in the country. I exclude races in these states from the final estimation dataset, because I use only within-race variation and thus need at least two markets with available advertising data. The final sample, after these states are excluded, contains 75 races and 308 markets.

There is significant variation in vote shares across markets within a typical statewide race. Figure 3 shows the density of demeaned vote shares - i.e., each candidate's vote share in each market within a state, minus that candidate's average share in the entire state. The standard deviation is approximately 7.5 percentage points. This is the fundamental variation that the estimator will use to fit the model's parameters.

4.3 Demographic data

To fit the random coefficients logit model (7), I need to be able to simulate from the distribution of observable demographic factors in each market. The primary purpose of allowing the coefficients β to vary with demographics in my context is to allow for differences in political tastes across different demographic segments of the population. Secondarily, the random

[24]These population shares were computed using the block-level population datasets and associated geographic boundary files produced by the US Census.

coefficients model allows for the response to advertising to vary with demographics.

I collected demographic data from the Census bureau's Current Population Survey (CPS) in October and November of 2002 and 2004. The CPS provides individual-level data on race, age, sex, income, education, and labor force participation from a nationwide sample of households. Each survey respondent was matched to a state and to one of the DMAs based on their place of residence as recorded by the CPS. This method resulted in a sample of individuals and their associated demographic characteristics for every state-DMA.

I augmented the CPS data with information on rates of DMA-level partisanship from the National Annenberg Election Surveys (NAES) conducted in 2000 and 2004.[25] Using the individual-level data, I estimated a within-DMA model of partisan affiliation as a function of CPS demographics. For every CPS individual, I then drew a partisan affiliation from the multinomial distribution[26] with probabilities given by the mean probabilities in the individual's DMA of residence plus the demographic marginal effects. In the estimation procedure, I resample with replacement from this market-specific pool when computing the simulated approximation to (14).

Demographic	Democratic	Republican
Nonwhite	0.17	-0.17
	(0.01)	(0.01)
College-Educated	0.00	0.02
	(0.01)	(0.01)
Employed	-0.04	-0.02
	(0.01)	(0.01)
Unemployed	-0.04	-0.08
	(0.02)	(0.02)
Income $> 75K$	-0.03	0.06
	(0.01)	(0.01)

Includes DMA fixed effects. N=25,806.

Table 3: Linear Probability Model of NAES Party ID on CPS demographics

[25] I used the 2004 values for 2004 elections and interpolated between the 2000 and 2004 values for the 2002 elections.

[26] With possible outcomes given by the set $\{Democrat, Republican, Independent\}$.

To keep the specification as parsimonious as possible, I limited the demographics used in estimation to partisan affiliation alone.[27] Thus, the vector D_i used in the estimation is simply a vector of two dummy variables: whether the individual identifies as a Democrat or as a Republican. The omitted category is voters who identify as independents.

Table 4 shows that NAES party ID is a reliable predictor of vote shares. The dependent variable is vote share (with one observation per candidate per DMA-state) and the independent variables are the probability of voters in the DMA-state stating a preference for the Democratic and Republican parties respectively. The estimates imply that shifting a market from 100% Democratic to 100% Republican would produce roughly a 40-point swing in vote shares.

Demographic	Estimate
DMA share of own partisans	0.21
	(0.04)
DMA share of opposing partisans	-0.22
	(0.04)

Dependent variable is vote share (ranges from zero to one). All regressions include candidate fixed effects.

Table 4: Regression of vote shares on NAES party ID.

4.4 Congruence Data

Snyder and Strömberg (2010) construct a measure of how well newspaper circulation areas fit into political districts, which they call "congruence." Suppose a district j has a set of papers P_j with positive circulation. Then, congruence is defined as:

$$Cong_j = \sum_{p \in P_j} MarketShare_{pj} ReaderShare_{pj} \tag{24}$$

They show that this measure affects the level of coverage that representatives receive

[27] However, given the way the partisanship variables are simulated conditionally on other demographic characteristics, this specification implicitly accounts for other characteristics as well.

and, consequently, is an excellent predictor of how much voters in a district know about their representatives. Their paper, however, focuses on House elections, and hence they construct their measure at the level of the congressional district. As my sample consists of media markets in statewide races, I need to construct the measure somewhat differently. I adapt their measure and use their circulation data from 2002 and 2004 to compute congruence at the DMA level.

I construct this DMA-level congruence for statewide races as follows. I identify the set of papers P_j with positive market share in DMA-state j. I then compute a weighted average, by market share in j, of the share of each paper's readership that resides in the same state. This is exactly like equation (24) with the exception that the district used in $MarketShare_{pj}$ is the DMA-state whereas the district used in $ReaderShare_{pj}$ is the state. This measure represents an average value of the in-state readership of papers circulating in DMA-state j.

As a check that this measure does, in fact, predict voters' levels of political information, I included the DMA-level congruence measure as an independent variable in several regressions on survey response data from the CCES. As described in section 3.4.2, newspaper congruence appears to reduce the residual variance in voters' assessments of the quality of incumbent politicians, consistent with voters in higher-congruence markets having more precise beliefs about their representatives' types. Table 5 shows an influence of congruence on a more direct measure of information: the probability that a respondent correctly identifies the party of the incumbent politician representing her state. The results show that higher congruence is associated with higher likelihood of correctly answering this question. The magnitude of the effect of congruence - which, by definition, ranges from zero to one - is comparable to the effect of having a college education. The difference between the pooled OLS and fixed effects estimators indicates that this difference is primarily a cross-state, rather than within-state, phenomenon.

4.5 Instruments

Finally, consistently estimating the parameters of (7) requires some source of exogenous variation in levels of advertising. Instruments are necessary to avoid the selection problem that candidates may allocate advertising to markets on the basis of knowledge about the

Coefficient	(1)	(2)
Intercept	0.844	
	(0.007)	
News Congruence of DMA	0.095	0.016
	(0.007)	(0.009)
Assessing a Senator	-0.028	-0.028
	(0.002)	(0.002)
Nonwhite	-0.024	-0.028
	(0.002)	(0.002)
College-Educated	0.052	0.051
	(0.002)	(0.002)
Employed	-0.012	-0.010
	(0.001)	(0.002)
Unemployed	-0.064	-0.061
	(0.004)	(0.005)
Income $> 75K$	0.039	0.039
	(0.002)	(0.002)
State Fixed Effects?	No	Yes

Standard Errors in Parentheses. N = 105,240

Table 5: Linear probability model of ability to correctly identify the party of the incumbent.

markets' prior likelihood of supporting the candidate.

Fortunately, instruments which are likely to affect the candidates' choices of where to advertise but are unrelated to unobserved political factors are available. I exploit variation in the (per-impression) prices of television advertising across markets within a state, which I collected from the *Media Market Guides* produced by Spot Quotations and Data (SQAD). SQAD publishes quarterly tables of prices of television advertising on a per-impression basis for all 210 DMAs in the United States.[28] Because candidates have limited resources available to purchase advertising, they are likely to shift advertising into relatively low-cost markets at the expense of high-cost markets, relative to the allocation they would choose in a hypothetical world of costless advertising.

[28]Since the relevant demographic that candidates want to reach is voting-age adults, I use SQAD's listed prices on a per-1000 impression basis within the "Adults 18+" demographic.

I used the 3rd quarter prices from the year *before* the relevant election as my instruments, for two reasons. One, to avoid mechanical correlation induced by the use of the same-year prices to construct advertising quantities, described in section 4.1.[29] Two, to eliminate the potential problem that the same-year prices may be driven in part by the desirability of the market for political advertisers. Local station affiliates can derive significant revenues in the pre-election months from political advertisers, and hence it is at least conceivable that their price-setting process may reflect some market-level political unobservables. In the prior year, no election was happening and hence prices should reflect only the factors that make particular markets more or less attractive to standard, consumer-products advertisers. Figures 4 and 5 plot different summaries of the price data, showing that significant variation exists in the normalized (per-1000-impressions) prices of advertising across DMAs.

A second source of exogenous variation arises from the fact that DMAs are defined in terms of metropolitan areas, which often cross state boundaries. Hence, many states contain "stub" parts of DMAs whose main population centers are outside the state, but include some smaller urban or suburban communities in the "stub" state.

For example, the city of Vancouver, Washington, home to about 150,000 residents, lies directly across the Columbia river from the much larger city of Portland, Oregon. Vancouver, Portland, and Portland's Oregon suburbs share a single television market, and Vancouver residents watch television broadcast by local Portland affiliates. Reaching these viewers requires buying advertising impressions on Portland television stations. From the perspective of a candidate running for a Washington office, many of these impressions will be wasted on Oregon residents who are ineligible to vote in Washington elections.

Discontinuities at state borders, as in the Vancouver example, increase the effective price of advertising for state-level candidates in markets that cross state boundaries. I therefore will use both the prices themselves and the prices divided by the fraction of the total DMA population that resides in the state of the election[30] as instruments for advertising levels.

The instruments discussed thus far - effective prices of advertising in markets within a state - will be identical for both candidates within a given race. This feature may introduce

[29]For discussion of this procedure, see Hartmann and Gordon (2011).

[30]Population shares were computed from the Census bureau's intercensal population estimates for 2002 and 2004.

an identification issue if the price instruments have the same, or nearly the same, effect on a candidate's own advertising levels as they do on her opponent's. To avoid this potential issue, I collected data on candidates' fundraising activity,[31] and computed each candidate's total funds raised from individual and Political Action Committee (PAC) sources as of the August 1st before the election. The idea here is that candidates with bigger fundraising budgets are likely to be less sensitive to price differences across media markets. Thus, interacting the price instruments with candidates' budgets generates instrumental variables which influence the allocation decisions of the two candidates in a race differently.[32] The final set of instruments consists of each of the two fundraising totals (individual and PAC) divided by each of the eight raw prices and eight effective (scaled by in-state share of DMA population) prices, for a total of 32 instruments. With this construction, each represents the number of ad impressions that could be purchased if a candidate spent her entire budget in a given market and time slot.

Table 6 shows fit statistics for regressions of the set of instruments on levels of promotional, attack, and total (the sum of promotional and attack) advertising. The instruments appear to strongly predict both types of advertising, though the fit is somewhat better for promotional advertising - likely because there are more observations with zero attack ads than there are for promote ads.

Fit statistic	Promote Ad Q	Attack Ad Q	Total Ad Q
R^2	0.48	0.40	0.50
F-statistic	12.2	8.73	13.2

All regressions include candidate fixed effects.

Table 6: First stage regressions of ad levels on instruments.

[31] Data for federal races comes from the Center for Responsive Politics (`http://www.opensecrets.org`). Data for state races is from Follow the Money (`http://www.followthemoney.org`).

[32] It is of course possible that fundraising totals are correlated with the unobserved quality term ξ_k. However, because I use within-race variation to identify the parameters, the fixed effects remove ξ_k from the structural error term. Because a candidate's budget is a constant, it will by definition be uncorrelated with the varitions in priors ($m^{\xi}_{0,jk}$) that remain in the structural error term after applying the fixed effects transformation.

5 Results

Regression approximations Before moving to the estimates of the structural parameters, I will first present the results of some more straightforward regressions that approximate the model described in section 3. These regressions use the identical data set and the same basic instruments described in the previous section, and hence will be comparable to the structural estimates. The downside is that they are not based on a model of utility maximization and therefore will not be useful for the counterfactuals of section 6.

Tables 7 and 8 show two different specifications of the relationship between advertising levels and vote shate. The first is a fully homogeneous logit model, as in Berry (1994), where the dependent variable is the difference in log vote shares between a candidate and the outside option (not voting). For each market within a race, there are therefore two observations of the dependent variable: one for the Democratic candidate and one for the Republican. Each has the form $\log(s_{jk}) - \log(s_{j0})$. I include candidate fixed effects in all specifications, and report two stage least squares estimates using the identical set of instruments described in section 4.5.

Table 8 shows a linear specification, where the dependent variable is the Democratic share of the two-party vote. In this version, there is only one observation per market within a race; the two party share here is explicitly a function of both candidates' advertising. In this version the fixed effects are at the race (rather than candidate) level, and the instruments are simply the market-level, previous year's advertising prices.

In both specifications, including the informational interaction terms appears to be important. The congruence interaction terms are of similar magnitude but opposite signs from the main effects. Since the congruence measure ranges from zero to one, this implies that the effects of advertising on vote share would be nearly eliminated in a maximally-congruent market. The direction of the incumbency interactions also opposes the main effects, though the magnitudes are smaller.

To get an idea of the magnitude of the estimated coefficients, the units in Table 7 are logged differences in vote shares and advertising quantities in 1000 impressions per capita. The median levels in the sample are around 50 impressions per capita, with standard deviation around 10. The estimate in column 3 of Table 7 implies that for two initially evenly

Variable	(1)	(2)	(3)
Total Ad Q	0.22	-0.04	2.21
	(0.26)	(0.26)	(1.51)
News Congruence * Ad Q			-1.90
			(1.53)
Incumbency * Ad Q			-1.55
			(0.73)
DMA share of own partisans		1.11	1.11
		(0.34)	(0.34)
DMA share of opposing partisans		-0.96	-0.98
		(0.34)	(0.34)

Dependent variable is the log difference in market share between candidate and outside option (not voting). Ad quantities in 1000 impressions per capita. All regressions include candidate fixed effects and instruments for ad levels.

Table 7: Reduced form regression of vote shares on ad levels (total) and interactions with information variables.

matched candidates advertising at the sample median level, if one raised her advertising levels by one standard deviation, she would gain about half a percentage point of the two-party vote. While not an overwhelming effect, this would be enough to be decisive in a close election.

Structural estimates Table 9 shows parameter estimates from the structural model, with specifications increasing in complexity from left to right. Column (1) allows for only a persuasive effect of advertising, fixing informational effects to zero; the utility function here is equation (7) without the expectation term $E_{ij}[\xi_k]$. The specification includes demographic interaction terms on the advertising effects, using NAES party ID's as demographic variables. Column (2) adds the possibility of informational effects: I report the first column of the prior covariance matrix Σ_0, along with the error-scaling parameter on congruence.[33] Column (3) adds an effect of incumbency on the variance of quality signals. Negative values of these

[33]Note that this parameter, as described in section 3.4.2, is estimated by maximum likelihood using a separate survey dataset.

Variable	(1)	(2)	(3)
Democratic Ad Q	2.50	0.62	2.42
	(0.59)	(0.24)	(0.94)
Republican Ad Q	-3.44	-0.22	-0.06
	(0.75)	(0.25)	(1.14)
News Congruence * Democratic Ad Q			-2.32
			(1.02)
News Congruence * Republican Ad Q			0.26
			(1.25)
Democratic Incumbent * Democratic Ad Q			-0.54
			(0.21)
Republican Incumbent * Republican Ad Q			0.20
			(0.41)
DMA share of Democratic partisans		0.54	0.55
		(0.11)	(0.11)
DMA share of Republican partisans		-0.38	-0.39
		(0.14)	(0.14)

Dependent variable is Democratic share of two-party vote (ranges from zero to one). All regressions include race fixed effects and price instruments for ad levels.

Table 8: Reduced form regression of Democratic vote shares on ad levels and interactions with information variables.

coefficients indicate that news congruence and incumbency status have the effect of reducing the variance of the quality signals. Column (4) allows for the possibility of nonlinear effects of persuasive advertising by including squared terms in the ad level.

Columns (5) and (6) repeat specifications (3) and (4), but with the addition of moments generated by the necessary conditions for equilibrium in the advertising allocation game between the two candidates. (See section 6 for an analysis of the supply side equilibrium). In these specifications, I add an additional moment which is constructed by computing the average difference from zero of the supply-side first order conditions, described in equation (30). These conditions imply that for two markets to which a candidate allocates positive quantities of advertising, the impact on the election result per dollar of spending must be equalized between the two markets. Because the full set of such conditions is clearly not

independent - in a race with three markets A,B, and C, the difference in the first order condition between markets A and B is related to the difference between A and C - I follow the suggestion of Bajari, Benkard, and Levin (2007) and select a random sample from the conditions that includes each market at most once.

Parameter	(1)	(2)	(3)	(4)	(5)	(6)
Total Ad Q	0.63	0.58	0.64	0.59	0.61	0.61
(Total Ad Q)^2				0.31		0.28
Σ_ξ		1.08×10^{-3}	1.23×10^{-4}	2.77×10^{-4}	1.39×10^{-3}	3.82×10^{-4}
$\Sigma_{\xi,a}$		-1.74×10^{-3}	-5.89×10^{-4}	-1.39×10^{-4}	-1.83×10^{-3}	-3.10×10^{-4}
News Congruence		-0.12	-0.12	-0.12	-0.12	-0.12
Incumbent			-0.10	-0.14	-0.11	-0.09
Supply Side Moments?	No	No	No	No	Yes	Yes

Table 9: Structural estimates of advertising and information parameters - total ad quantity specifications.

A few things stand out from the table. First, the main (persuasive) coefficients on advertising levels are positive and, consistent with the reduced form results reported above, fairly small in magnitude. Columns (4) and (6) indicate that the persuasive effect appears to have increasing marginal returns, as the squared terms are positive. However, given the units of the advertising quantity measure, which is measured in 1000 impressions per capita and is typically on the order of 0.05, the squared effect is negligible and the estimated function is very nearly linear.

Second, informational effects appear to counteract the sponsoring candidate's intentions. The off-diagonal element of the prior covariance matrix Σ_0 is negative, indicating that voters update their beliefs *downwards* upon observing ads. Though this result appears counterintuitive, it is entirely consistent with a world where voters are concerned about a donations-for-influence exchange between politicians and major donors. Higher levels of advertising may indicate to voters that a candidate is more indebted to the donors who contributed to finance the ads. The overall effect of advertising - the persuasive plus the informative effect - remains positive, implying that it remains in candidates' interest to advertise despite the negative influence on voter beliefs.

Third, adding the supply side moment induces very little change in the persuasive coefficient. It does appear to increase the value of the informative coefficient relative to the corresponding specification without the supply side moment, indicating that candidates may be targeting ads relatively heavily to lower-congruence markets.

Finally, higher newspaper congruence reduces the variance of quality signals, as does being an incumbent. The magnitude of the incumbency effect on voter information levels is very similar in all specifications to that of newspaper congruence.

Parameter	(1)	(2)	(3)	(4)	(5)	(6)
Promote Ad Q	-0.05	0.20	-0.29	1.17	0.71	0.38
(Promote Ad Q)^2				-6.23		-1.58
Opponent's Attack Ad Q	-1.23	-0.85	-0.90	0.29	-0.43	-1.25
(Opponent's Attack Ad Q)^2				-1.09		0.03
Σ_ξ		2.64×10^{-3}	4.40×10^{-4}	7.65×10^{-3}	1.48×10^{-2}	3.42×10^{-2}
$\Sigma_{\xi,P}$		1.27×10^{-3}	5.97×10^{-4}	2.30×10^{-3}	-2.15×10^{-3}	2.60×10^{-3}
$\Sigma_{\xi,A}$		7.05×10^{-3}	3.90×10^{-4}	2.90×10^{-4}	-1.57×10^{-3}	2.73×10^{-3}
News Congruence		-0.12	-0.12	-0.12	-0.12	-0.12
Incumbent			-0.20	-0.07	-0.24	-0.02
Supply Side Moments?	No	No	No	No	Yes	Yes

Table 10: Structural estimates of advertising and information parameters - promote / attack specifications.

Promote / attack variant Table 10 shows estimates of a variant of the model which estimates separate coefficients for promotional and attack advertising. I extend the voter's utility from a given candidate to include the candidate's *own* promotional advertising and her *opponent's* attack advertising. The prior matrix Σ_0 in this case expands to be 3×3, allowing each type of advertising to have different informational effects as well.

This form is interesting because so much of the existing literature on political advertising has focused specifically on the effects of negative ads. However, it is important to note that the instruments I use affect candidates' choices of where to run ads and how many to buy - e.g., the total ad quantities purchased in each market - but have no influence on candidates' choices of ad *content*. Given that a candidate has chosen to make an ad purchase in some

market, the instruments have no further influence on the candidate's decision to use that ad time to promote herself or attack her opponent. Hence, coefficient estimates in the split promote / attack specifications may partially reflect candidates' targeting strategies along with the causal effect of ad exposure on vote choice.

In these specifications, imposing the supply-side conditions at the estimation stage may be especially useful. Adding the supply-side moments uses the information contained in candidates' targeting strategies to infer voters' response to each type of advertising.

The table shows that the persuasive effects of candidates' own promotional advertising are typically positive, and the effects of their opponents' attack advertising typically negative. However, in a few of the demand-side-only specifications, the persuasive coefficients have the "wrong" sign, implying that candidates would be better off not advertising. Adding the supply side moments restores the expected directionality, and results in coefficients on the squared terms that imply decreasing marginal effectiveness of advertising (though, again, the squared coefficient for negative advertising is negligible in practice and the function is effectively linear).

Comparing these estimates with the total quantity coefficients from table 9 suggests that candidates are likely targeting promotional ads to markets which for unobserved reasons are more likely to support the candidate, and attack ads to markets which are less likely.This result is consistent with the contention in Ansolabehere and Iyengar (1996) that candidates use negative ads to depress the turnout of their opponent's supporters.

The directionality of informational effects in the promote/attack variant are less consistent than in the total quantity version, though typically both coefficients are positive. Because the attack ad measure here is the quantity of attack ads run by the candidate's opponent, a positive coefficient indicates that voters' inference from this type of advertising counteracts the sponsor's intention. Again, as in the total ads specification, this result points to a "soft" information interpretation, where voters learn by understanding the relationship between donors and candidates rather than by discovering previously unknown facts. This contrasts with the contention of Geer (2006) that negative ads contain more "hard," or factual, information.

Tables 11 and 12 show sample estimates of the demographic marginal terms, for the specification reported in column (6) of tables 9 and 10, respectively. As expected, Democratic

partisans are much more likely to vote for Democratic candidates than are Republican partisans. Markets with higher Democratic partisanship also tend to have lower turnout overall. Interestingly, the marginal terms on advertising levels generally go in the opposite direction of the main effects for both kinds of partisans, suggesting that voters with established political affiliations are less influenced by persuasive advertising than are independents. In the promote / attack specification, the marginal terms on attack ads for both kinds of partisans are large enough that the net effect is significantly positive, suggesting that partisans may be "fired up" by seeing their favored candidate attacked.

Demographic	Total Ad Q	Democrat	Constant
Democratic	-1.31	0.47	-0.21
Republican	-0.02	-0.36	0.11

Table 11: Structural estimates of demographic interaction terms - total ad specification.

Demographic	Promote Ad Q	Attack Ad Q	Democrat	Constant
Democratic	-0.61	2.29	3.07	-2.71
Republican	2.65	2.70	-9.97	0.32

Table 12: Structural estimates of demographic interaction terms - promote / attack specification.

Interpretation Because the structural parameters are difficult to interpret directly, Tables 13 to 15 show two measures of the effects of advertising on voter behavior at these estimated parameter values.

Table 13 reports the percentage reduction in voters' posterior variance in beliefs over quality. These reductions are computed by comparing the top-left element of the posterior variance matrix Σ - defined by equation 6 - under the estimated parameter vector to the same element in a hypothetical world where the variance of advertising signals $\eta^P_{ijk}, \eta^A_{ijk}$ is set to infinity. When the variance of advertising signals goes to infinity, voters must fully discount the advertising signal and do not update their posteriors at all. The effect on the

precision of voters' posterior beliefs is non-negligible, but still relatively small - typically less than ten percent.

Specification	(2)	(3)	(4)	(5)	(6)
Totals	10.0%	10.0%	0.26%	8.9%	9.3%
Promote / Attack	4.2%	6.2%	3.56%	2.5%	2.3%

Table 13: Posterior variance reduction due to informational component of advertising.

Tables 14 and 15 report the average change in the log odds ratio of voting for the sponsoring candidate over her opponent following a 10-impression-per-capita increase in ad levels, decomposed into persuasive and informative components. The results show that in all specifications, the persuasive component is the dominant influence on behavior, typically larger by an order of magnitude than the informational component. The magnitude of the changes in the log odds ratio indicate that an increase of this size in one candidate's total advertising quantity would make an initially indifferent voter a little more than 3% more likely to vote for the candidate who sponsored the additional advertising.[34]

Cand. Type	Effect	(2)	(3)	(4)	(5)	(6)
Incumbent	Persuasive	0.031	0.034	0.032	0.033	0.033
	Informative	-0.003	-0.001	0.000	-0.003	-0.001
Challenger	Persuasive	0.031	0.034	0.032	0.033	0.033
	Informative	-0.003	-0.001	0.000	-0.004	-0.001

Table 14: Changes in log odds ratio due to persuasive and informative components - total ad quantity specifications.

This result may initially appear inconsistent with the reduced-form specification presented earlier, as that specification indicated that the inclusion of the informational interaction terms was important. What reconciles the two is that in the sample, the levels of newspaper congruence are quite high on average - the mean is over 0.8. While informational

[34]Note that the odds ratio measures the probability of voting for one candidate over that of voting for the other. Since there is a third option - not voting - the effect on a candidate's vote share will be smaller than 3%.

Cand. Type	Ad Type	Effect	(2)	(3)	(4)	(5)	(6)
Incumbent	Promote	Persuasive	0.006	-0.009	0.038	0.023	0.012
		Informative	0.001	0.001	0.002	-0.002	0.003
Challenger	Promote	Persuasive	0.006	-0.009	0.038	0.023	0.012
		Informative	0.001	0.001	0.005	-0.004	0.005
Incumbent	Attack	Persuasive	-0.024	-0.025	-0.172	-0.012	-0.044
		Informative	0.001	0.001	0.000	-0.002	0.004
Challenger	Attack	Persuasive	-0.024	-0.025	-0.172	-0.012	-0.044
		Informative	0.001	0.001	0.001	-0.003	0.005

Table 15: Changes in log odds ratio due to persuasive and informative components - promote / attack specifications.

effects are important in low-congruence markets, the typical market has high congruence and hence little influence of advertising on vote share due to informational effects.

The same interpretation is evident in the results of the structural model. The maximum potential informational effect of advertising on voters' posterior beliefs occurs in the limit where $\sigma_\xi^2 \to \infty,\ \sigma_a^2 \to 0$. In this limit, voters completely discount the initial quality signal observation and place all weight on the ad signal observation. The limiting informational coefficients on advertising in this case (the elements of the vector α_{jk} associated with promotional and attack advertising) are given by:

$$\lim_{\sigma^2_{\xi,jk}\to\infty,\, \sigma^2_{a,jk}\to 0} \alpha_{jk}^{(2)} = \frac{\Sigma_{21}}{\det \bar{\Sigma}} \tag{25}$$

$$\lim_{\sigma^2_{\xi,jk}\to\infty,\, \sigma^2_{a,jk}\to 0} \alpha_{jk}^{(3)} = \frac{\Sigma_{31}}{\det \bar{\Sigma}} \tag{26}$$

Where $\bar{\Sigma}$ is the bottom right sub-matrix of Σ_0, which is set to the sample covariance of advertising levels. As the determinant of this covariance matrix in the sample is on the order of 10^{-7}, the estimates in table 10 imply the potential for informational effects of advertising much larger than the estimated persuasive effects in markets with very little prior information. However, because the typical level of newspaper congruence is high,

voters' prior beliefs are already quite precise prior to the introduction of advertising and hence it is difficult for advertising to shift beliefs in the average market.

6 Counterfactuals and Welfare

Given the parameter estimates reported in 5, it is possible to explore the consequences of alternative regulatory regimes on vote shares and election outcomes. I consider a few plausible alternative regimes that have been suggested by advocates of election reform: restricting advertising expenditures across the board, banning negative advertising, or implementing a public-finance system in which advertising budgets are equalized across candidates and paid for by the federal government. The main questions of interest are how these changes would affect vote shares, turnout, and most importantly, overall voter welfare.

6.1 Defining welfare

The appropriate definition of welfare in the model presented here is not obvious, as the utility specification allows advertising to influence voter utility not only by potentially changing who wins or loses the election, but also directly, through the presence or absence of ads themselves. Taken literally, this feature implies that regulatory changes that alter equilibrium advertising levels could affect voter welfare even if they have no effect whatsoever on election outcomes.

I instead choose to interpret the utility function in (7) as a *decision utility*, as distinct from *experience utility*.[35] In my welfare measures, I focus only on the latter. I define the experience utility that a voter receives as the combination of policy and quality characteristics of the elected candidate. This definition excludes direct advertising utility, and hence allows welfare to vary as a result of variation in advertising levels *only* if those changes induce change in the distribution of candidates who are elected. Persuasive effects of advertising and the turnout cost in this interpretation are transient; they may alter a voter's choice but are irrelevant to his utility after the election is over.

The structural model defined previously allows me to compute expected differences of this experience utility across competing candidates in a race in a simple manner. Recall

[35]For a discussion of alternative welfare criteria, see e.g. Bernheim and Rangel (2009)

from equation (10) that the mean utility of candidate k in market j is defined as:

$$\delta_{jk} = \beta_0 + \beta_1 p_k + \beta_2 a_{jk} + \bar{m}^{\xi}_{jk} \tag{27}$$

These mean utilities are computed by the contraction-mapping procedure used in the estimation step. With estimates of the parameters in hand, we can subtract from δ_{jk} the advertising component $\beta_2 a_{jk} + \alpha_{jk}\,[0, a_{jk}]'$. If we take the difference across the two candidates in the same race, we are left with the following difference in structural errors:

$$\Delta\omega_j = \beta_1(p_1 - p_2) + \alpha^{(1)}_{j1}(\xi_1 - m^{\xi}_{0,j1}) - \alpha^{(1)}_{j2}(\xi_2 - m^{\xi}_{0,j2}) + m^{\xi}_{0,j1} - m^{\xi}_{0,j2} \tag{28}$$

This expected mean difference in utility includes three components. The first, $\beta_1(p_1 - p_2)$ is a policy term representing the mean difference in utility from the policy positions of the two candidates. The second, $\alpha^{(1)}_{j1}(\xi_1 - m^{\xi}_{0,j1}) - \alpha^{(1)}_{j2}(\xi_2 - m^{\xi}_{0,j2})$, is an expected quality difference determined by the voter's observation of the candidates' job performance signals. The last term, $m^{\xi}_{0,j1} - m^{\xi}_{0,j2}$ is a difference in priors, which may vary across markets.

Recall that we have an estimate for $(1 - \alpha^{(1)}_{jk})m^{\xi}_{0,jk}$, given by equation 18. Subtracting these prior belief differences from (28) leaves behind only the quality and policy difference, which by construction is constant across markets within a race. To this value I add the population-weighted average value of the demographic interaction term on party in ΠD_i. This cross-candidate difference will be the measure of welfare used in the counterfactual simulations. The difference enters welfare calculations with positive sign if the Democratic candidate wins, and with negative sign if she loses.

6.2 Candidates' objectives

Predicting the effects of changes in campaign advertising regulation requires understanding how candidates would change their advertising strategies under alternative regimes. I will assume that candidates have full knowledge of voters' utility parameters and that they act so as to maximize their overall vote share in the election, conditional on their opponent's

strategy.[36] In other words, I take advertising allocations to be the Nash equilibrium of a game where each candidate solves:

$$\max_{\vec{a}_k} \frac{\sum_j s_{j,k}(a_{j,k}, a_{j,-k})P_j}{\sum_j (s_{j,k}(a_{j,k}, a_{j,-k}) + s_{j,-k}(a_{j,k}, a_{j,-k}))P_j} \tag{29}$$

Subject to a budget constraint of the form $\sum_j p_{j,k}a_{j,k} \leq B_k$. Here $s_{j,k}$ is the vote share of candidate k in market j, which is defined by equation 15; P_j is the total voting-age population of market j who reside in the state of the election; p_j is the price of advertising in market j; and B_k is the candidate's advertising budget. Note that since the unit of $a_{j,k}$ is impressions per person in the media market, the price $p_{j,k}$ will have units of dollars times media market population. The prices thus account for the fact, noted in section 4.5, that certain markets may have higher effective prices for state-level candidates if much of their population resides outside of the state. I will take the candidates' budgets to be fixed at the level observed in the data, and use the average observed prices paid for $p_{j,k}$.[37]

Under some conditions on the utility parameters that are satisfied in the estimates reported above,[38] the resulting game is concave and an equilibrium exists. Given new constraints, new equilibria can be computed using the gradient descent method of Rosen (1965). Since the share functions s_{jk} are averages of logistic functions s_{ijk} over the set of voters in a market, the gradients of the share functions are well defined and can be computed analytically.

First Order Conditions The constrained optimization problem described by equation 29 generates a set of first order conditions for each candidate. The conditions take the following form for a given candidate k and any two markets j, l in which k advertises at a positive level:

[36] An alternative specification would be to assume that candidates maximize their probability of victory. I choose to use the maximize-vote-share objective for computational reasons: it is far simpler to compute and has properties that ensure the existence of equilibrium in the two-candidate game.

[37] As described previously, there are multiple prices for each market for different day-parts. I compute an average price by dividing the candidates' total dollar expenditures in each market by the quantity of ads (in impressions) purchased.

[38] $s'_{j,k}(a_{j,k}, \cdot) \geq 0, s''_{j,k}(a_{j,k}, \cdot) \leq 0$.

$$\frac{P_j}{p_{j,k}}\left[\frac{\partial s_{j,k}}{\partial a_{j,k}}V_{-k} - \frac{\partial s_{j,-k}}{\partial a_{j,k}}V_k\right] = \frac{P_l}{p_{l,k}}\left[\frac{\partial s_{l,k}}{\partial a_{l,k}}V_{-k} - \frac{\partial s_{l,-k}}{\partial a_{l,k}}V_k\right] \tag{30}$$

Where V_k here is candidate k's total votes receieved in the election, e.g. the sum of $s_{j,k}P_j$ over all markets in the election. In specifications that separate promotional and attack advertising, similar conditions will hold on the partial derivatives of the shares with respect to a candidate's promotional and attack advertising, both within the same market and across markets.

6.3 Alternative regulatory scenarios

I simulate the consequences of six possible alternative regulatory scenarios. Each involves a hypothetical change to the regulations governing campaign advertising in the elections in the sample. In each scenario, I impose additional constraints on the candidates' spending allocations and/or changes to the parameters, as described below. I then compute new, counterfactual equilibrium advertising levels and recompute vote shares and ultimate election outcomes.

The six scenarios I investigate are defined as follows:

Baseline This scenario imposes only the constraint that candidates must spend no more than their observed spending, and allows candidates to reallocate ads across markets (or between promotional and attack advertising) as necessary. The purpose of this scenario is as a measure of the fit of the model to the data: it tests how close candidates' observed advertising strategies are to the equilibrium strategies generated by the model.

Ad Ban In this scenario, I impose a ban on all television advertising. I restrict ad levels to zero for all candidates and all markets in the sample.

Negative Ban In this scenario, I prohibit negative advertising only. Candidates can substitute by reallocating their advertising budgets to purchase additional promotional advertising.

Expenditure Cap This scenario imposes a maximum (but not a minimum) on candidates' ad spending. The cap is set at \$0.11 per capita, the 20^{th} percentile level in the dataset.

Democrats Don't Advertise To simulate the effects of asymmetric changes in ad levels, in this scenario I prevent Democratic candidates from advertising at all. I leave Republican candidates unconstrained.

Public Finance In this scenario, I equalize ad spending across each candidate in a race by assigning each candidate the same budget. I also eliminate the informational content of ads, as under a publicly-funded system, voters could not learn anything about candidate quality from the quantity of ads deployed.

6.4 Results of counterfactual simulations

Figure 6 shows a measure of fit of the predictions generated by the baseline scenario to the actual data.[39] The figure compares the observed vote share of the Democratic candidate in each race on the horizontal axis, plotted against the simulated vote share of the Democratic candidate in that race on the vertical axis. The fit is quite close, with most races lying near the 45 degree line.

Figure 7 shows the distribution of three outcome measures - the vote share of the Democratic candidate, the vote share of the incumbent candidate, and overall turnout - under the baseline scenario and the five alternatives described previously. Perhaps not surprisingly, given the small magnitudes of the parameter estimates, the overall effects on the distribution of election outcomes are all quite small.

The symmetric scenarios, where both candidates are equally affected by new constraints on spending, produce close to negligible effects on aggregate outcomes. There is a small increase in overall turnout rates under the negative ban scenario, due to the negative persuasive coefficient associated with this type of advertising. Figure 9 shows this movement in turnout rates in more detail.

[39] All results in this section use parameter estimates from Specification 5 in table 10.

As advertising levels are estimated to be close to orthogonal to candidate quality, changes in advertising also do not systematically change the electoral performance of the candidates associated with higher quality: even extreme restrictions on ad levels do not degrade the quality of voters' choices. Figure 10 plots the distribution of races by the change in welfare between the baseline and total-ban scenarios. In the vast majority of races, the outcome of the election does not change and hence there is no change in welfare. In the few races that do flip, there is no apparent bias in the direction of change; the races are evenly split between increases and decreases in welfare. This rather striking result is a consequence of the very limited degree of belief-updating due to advertising exposure in the typical race.

These results point to a strong "arms race" element of advertising expenditures in elections. If either candidate were to unilaterally decide not to advertise, the election results could flip in a number of relatively close races in the sample. However, bilateral reductions like symmetric bans or spending caps have little net effect on the outcome, because each candidate's expenditures cancel out those of the other. The "Democrats Don't Advertise" scenario investigates such an asymmetric change in advertising levels (albeit an extreme one), in which Democratic candidates' ability to advertise is shut off but Republican candidates are unaffected.

Here, we see a fall in Democratic vote share of around three percentage points in the typical race. Figure 8 shows that these losses are concentrated in the most competitive (and most heavily advertised) races - exactly the cases where advertising is most likely to swing the outcome. Though three points is not an overwhelming handicap, it is enough to discourage any candidate who expects she might face a close race from unilaterally choosing not to advertise. The individual candidate's incentive is still to advertise as much as possible, despite the futility of advertising from the aggregate perspective.

Because fundraising decisions must be made well in advance, and candidates are likely to be risk averse, no candidate who foreseeably might face a close race is likely to forego fundraising for this contingency. And once the candidate has raised the money, there is little reason not to spend it, even if the race turns out not to be so close.[40] As most of this

[40]Of course, candidates who expect to win comfortably can save some of their campaign funds as a "war chest" for use in future, more competitive, contests. Candidates in losing battles, on the other hand, have no incentive to save at all.

fundraising effort will end up being wasted on non-competitive elections, the parties have developed some mechanisms to more effectively coordinate the party's fundrasing activity and reduce unnecessary effort. The Democratic Senatorial Campaign Committee (DSCC) and National Republican Senatorial Committee (NRSC) are both examples of institutions that centralize fundraising and then distribute funds to the races that end up being most competitive, thereby increasing the aggregate efficiency of the party's fundraising and advertising activity.

7 Conclusion

This paper has provided a step towards understanding the effect of private campaign finance on the performance of the electoral mechanism. The model developed here allows advertising to influence voters both by signaling candidate attributes and by direct persuasion. Candidates in the model act strategically: they seek out the most cost-effective advertising purchases and may change their advertising strategies in response to changes in conditions. This strategic behavior yields instruments for advertising levels in the form of variation in the effective prices of advertising across media markets within a statewide race. Separate identification of the two causal channels derives from the interaction of advertising effectiveness with variables that shift the precision of voters' prior beliefs, before exposure to advertising.

The results show that campaign advertising contains some measurable informational content. Voters appear to be capable of drawing inferences about candidate qualities by observing levels of advertising: they respond more strongly to ads from candidates and in markets for which less non-advertising information is available. Nonetheless, the estimates demonstrate that voters' pre-advertising uncertainty is already quite low in the typical race, and hence additional information provided by advertising has little room to influence beliefs. Accordingly, on average the direct persuasion response is larger, by a wide margin, than the learning response to advertising.

Overall, both effects of advertising are real but relatively small in magnitude. Furthermore, the levels of advertising of the opposing candidates in most races are fairly balanced, and in large measure each candidate's spending simply cancels out the effects of her oppo-

nent's, yielding little net change in vote shares for either candidate. As a result, in simulations I find that bilateral reductions in ad levels that affect both candidates symmetrically, such as expenditure limits or outright advertising bans, tend to induce little change in election outcomes. Because the primary mechanism for advertising influence is orthogonal to candidate attributes, such reductions do not decrease the quality of the pool of elected candidates and hence are welfare-neutral. If there is a social cost associated with the fundraising activities that pay for campaign ads - for instance, that fundraising requires policy concessions to the interest groups that comprise the donor pool - then the overall welfare effects will be positive.

A few important issues are left unaddressed by this analysis and will remain for future work. First, I do not explicitly model the relationship between candidates and donors. As a result, the counterfactuals presented here treat candidates' advertising budgets as exogenous. An interesting research question is how changes in contribution limits or other regulatory mechanisms would influence the incentives of donors and hence alter the levels of funding available to candidates representing different constituencies. Second, I treat the candidates' strategic problem as a static choice, when of course real campaigns are months-long affairs during which voter opinion and outside events can change unpredictably. Expanding the candidate-side analysis to a fully dynamic model of advertising decisions could yield estimates of a number of quantities of interest. For instance, estimates of the marginal value of additional advertising dollars to different types of candidates, at different points in a campaign, would be useful for understanding in what situations candidates are likely to have more or less bargaining power vis-a-vis donors. I leave these questions for future research.

References

Ackerberg, D. (2001). Empirically distinguishing informative and prestige effects of advertising. *The RAND Journal of Economics 32*(2), 316–333.

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review 44*(3), 1007–1040.

Ansolabehere, S., J. M. d. Figueiredo, and J. M. Snyder (2003). Why is there so little money in U.S. politics? *The Journal of Economic Perspectives 17*(1), 105–130.

Ansolabehere, S. and S. Iyengar (1996). *Going negative: how political advertisements shrink and polarize the electorate*. New York: Free Press.

Ashworth, S. (2006). Campaign finance and voter welfare with entrenched incumbents. *American Political Science Review 100*(01), 55–68.

Ashworth, S. and J. Clinton (2007). Does advertising exposure affect turnout? *Quarterly Journal of Political Science 2*(1), 27–41.

Bagwell, K. (2007). The economic analysis of advertising. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, Volume 3, pp. 1701 – 1844. Elsevier.

Bajari, P., C. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331–1370.

Baron, D. (1994). Electoral competition with informed and uniformed voters. *American Political Science Review*, 33–47.

Bernheim, B. D. and A. Rangel (2009, February). Beyond revealed preference: Choice-Theoretic foundations for behavioral welfare economics. *The Quarterly Journal of Economics 124*(1), 51–104.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994, July). Estimating Discrete-Choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Campante, F. R. (2011). Redistribution in a model of voting and campaign contributions. *Journal of Public Economics 95*(7-8), 646 – 656.

Chernozhukov, V. and H. Hong (2003). An MCMC approach to classical estimation. *Journal of Econometrics 115*(2), 293–346.

Ferejohn, J. (1986, July). Incumbent performance and electoral control. *Public Choice 50*(1-3), 5.

Geer, J. G. (2006). *In defense of negativity : attack ads in presidential campaigns*. Chicago: University of Chicago Press.

Gerber, A. S., J. G. Gimpel, D. P. Green, and D. R. Shaw (2011). How large and long-lasting are the persuasive effects of televised campaign ads? results from a randomized field experiment. *American Political Science Review 105*(01), 135–150.

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science 45*(4), pp. 862–886.

Hartmann, W. and B. Gordon (2011). Advertising effects in presidential elections. Stanford GSB Working Paper.

Hausman, J. A., A. W. Lo, and A. MacKinlay (1992, June). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics 31*(3), 319–379.

Huber, G. and K. Arceneaux (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science 51*(4), 957–977.

Londregan, J. and T. Romer (1993). Polarization, incumbency, and the personal vote. In W. A. Barnett (Ed.), *Political economy: institutions, competition, and representation*, Proceedings of the Seventh International Symposium in Economic Theory and Econometrics, pp. 355–378. Cambridge University Press.

Nevo, A. (2000). A practitioner's guide to estimation of Random-Coefficients logit models of demand. *Journal of Economics & Management Strategy 9*(4), 513–548.

Prat, A. (2002). Campaign advertising and voter welfare. *Review of Economic Studies 69*(4), 999–1017.

Prat, A., R. Puglisi, and J. M. Snyder (2010). Is private campaign finance a good thing? estimates of the potential informational benefits. *Quarterly journal of political science 5*(3), 291–318.

Rosen, J. B. (1965, July). Existence and uniqueness of equilibrium points for concave N-Person games. *Econometrica 33*(3), 520–534.

Snyder, J. M. (1989, May). Election goals and the allocation of campaign resources. *Econometrica 57*(3), 637–660.

Snyder, J. M. and D. Strömberg (2010, April). Press coverage and political accountability. *Journal of Political Economy 118*(2), 355–408.

Figures

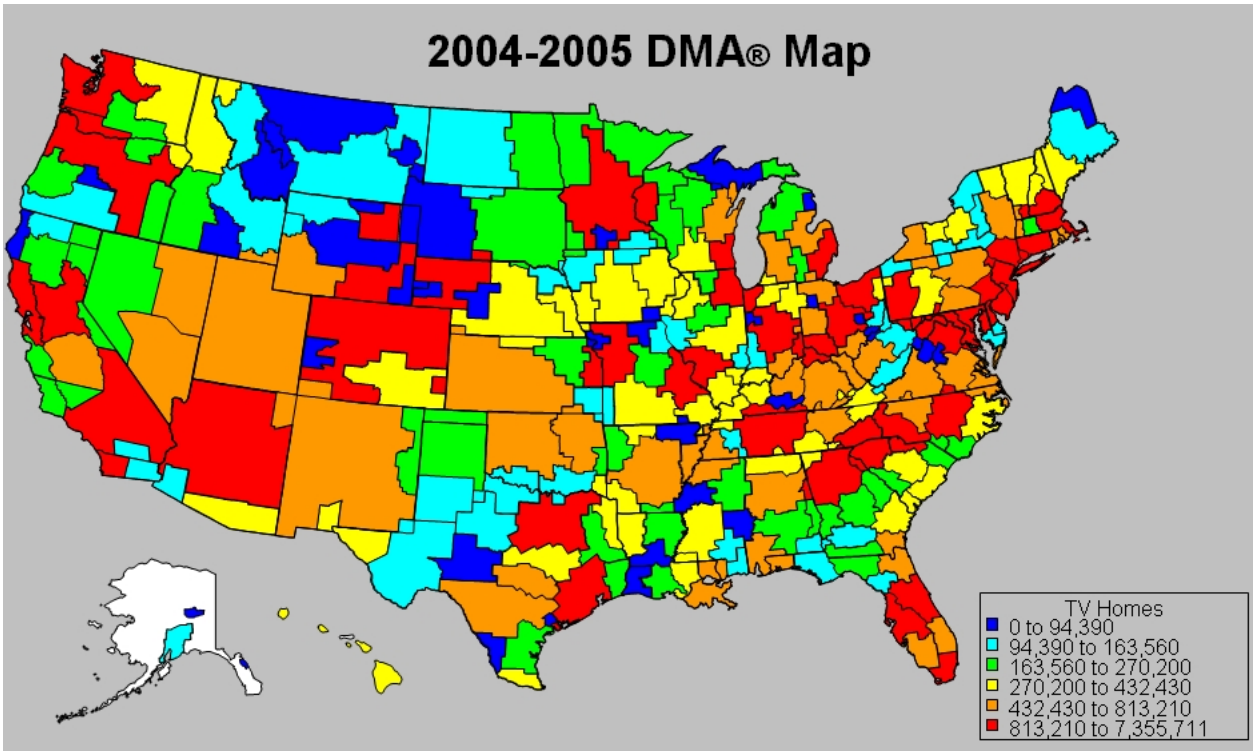

Figure 1: Map of Nielsen Designated Market Areas (DMAs)

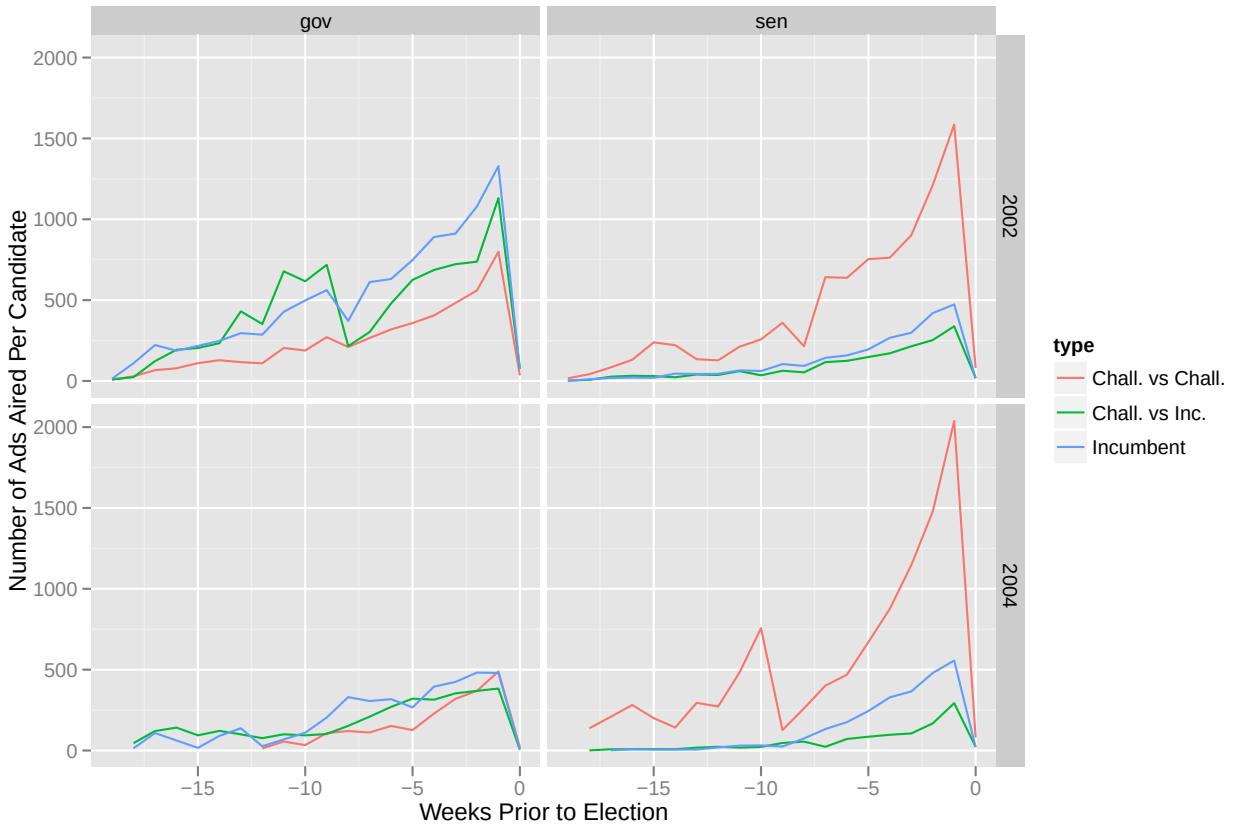

Figure 2: Timing of Challenger versus Incumbent Advertising

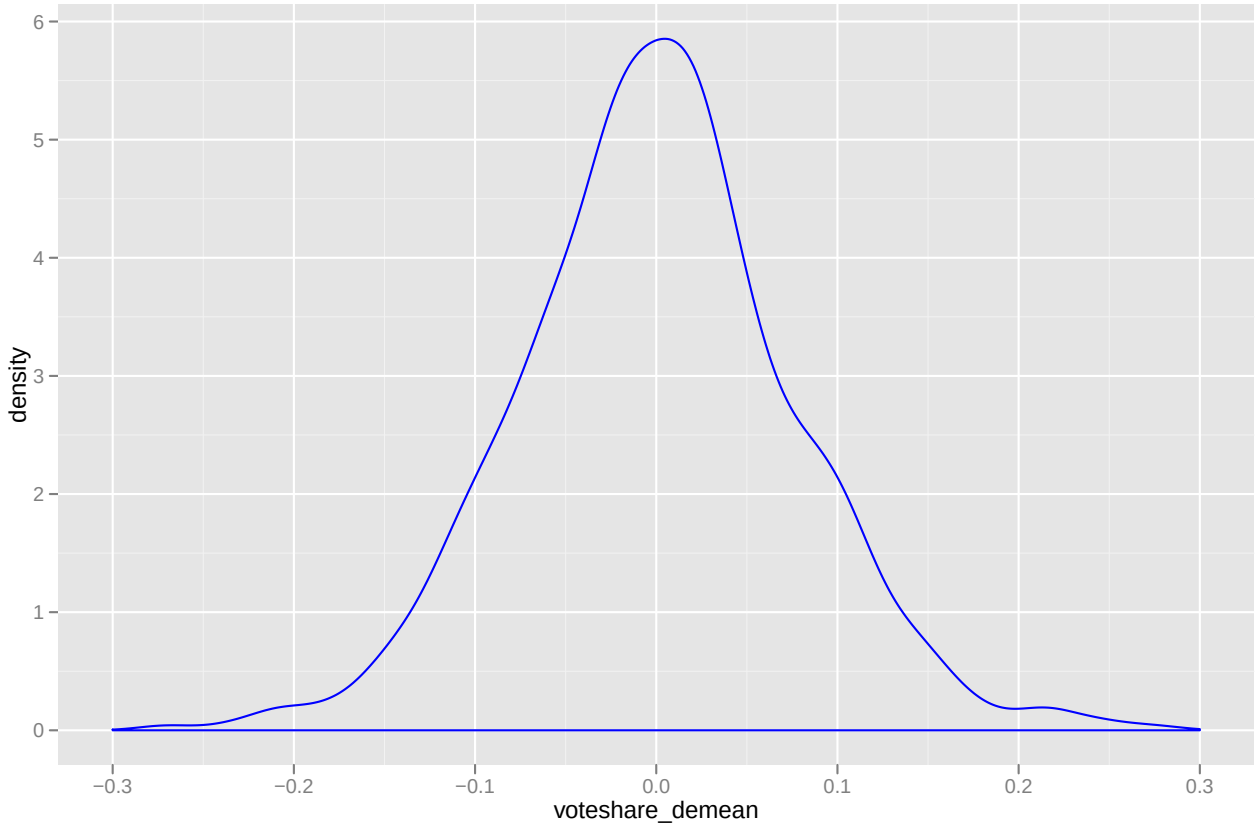

Figure 3: Density of demeaned vote shares at the DMA level.

Figure 4: Normalized advertising prices, by day part and DMA

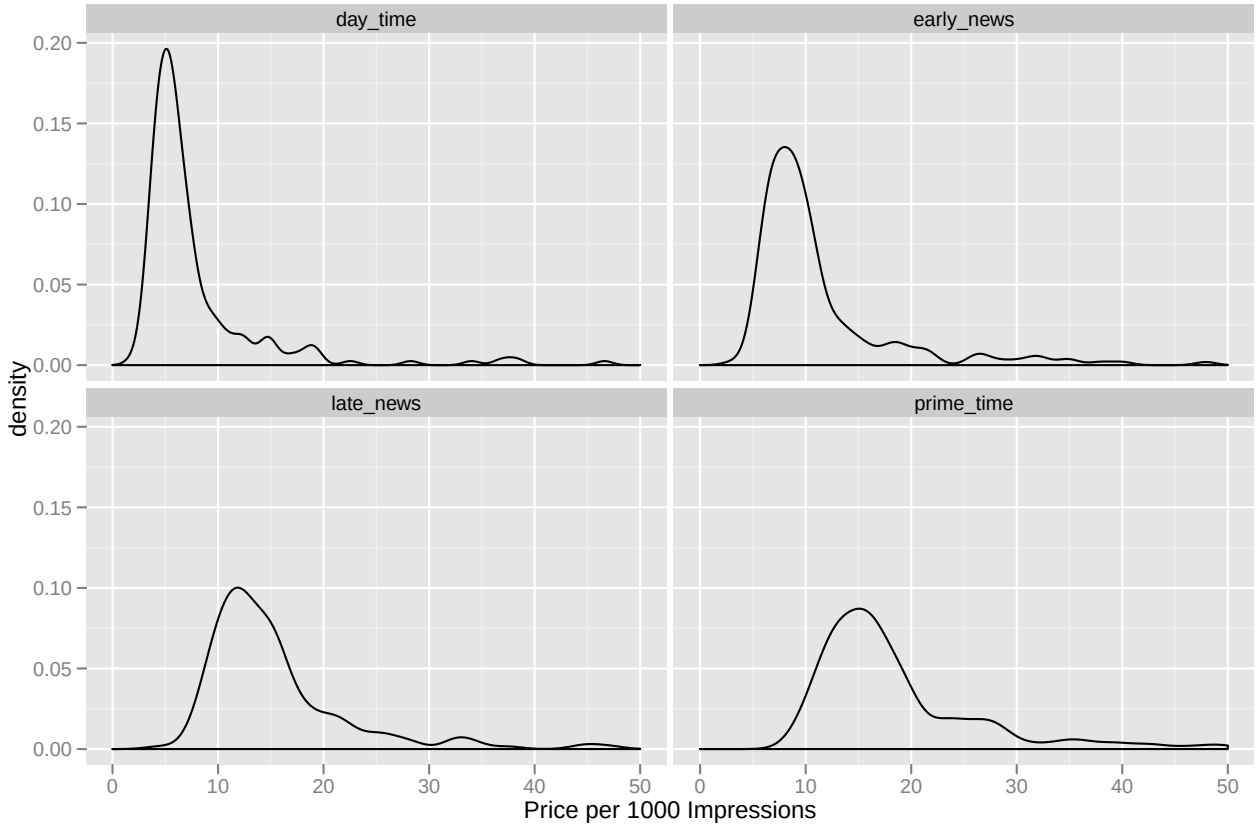

Figure 5: Density of normalized advertising prices, by day part.

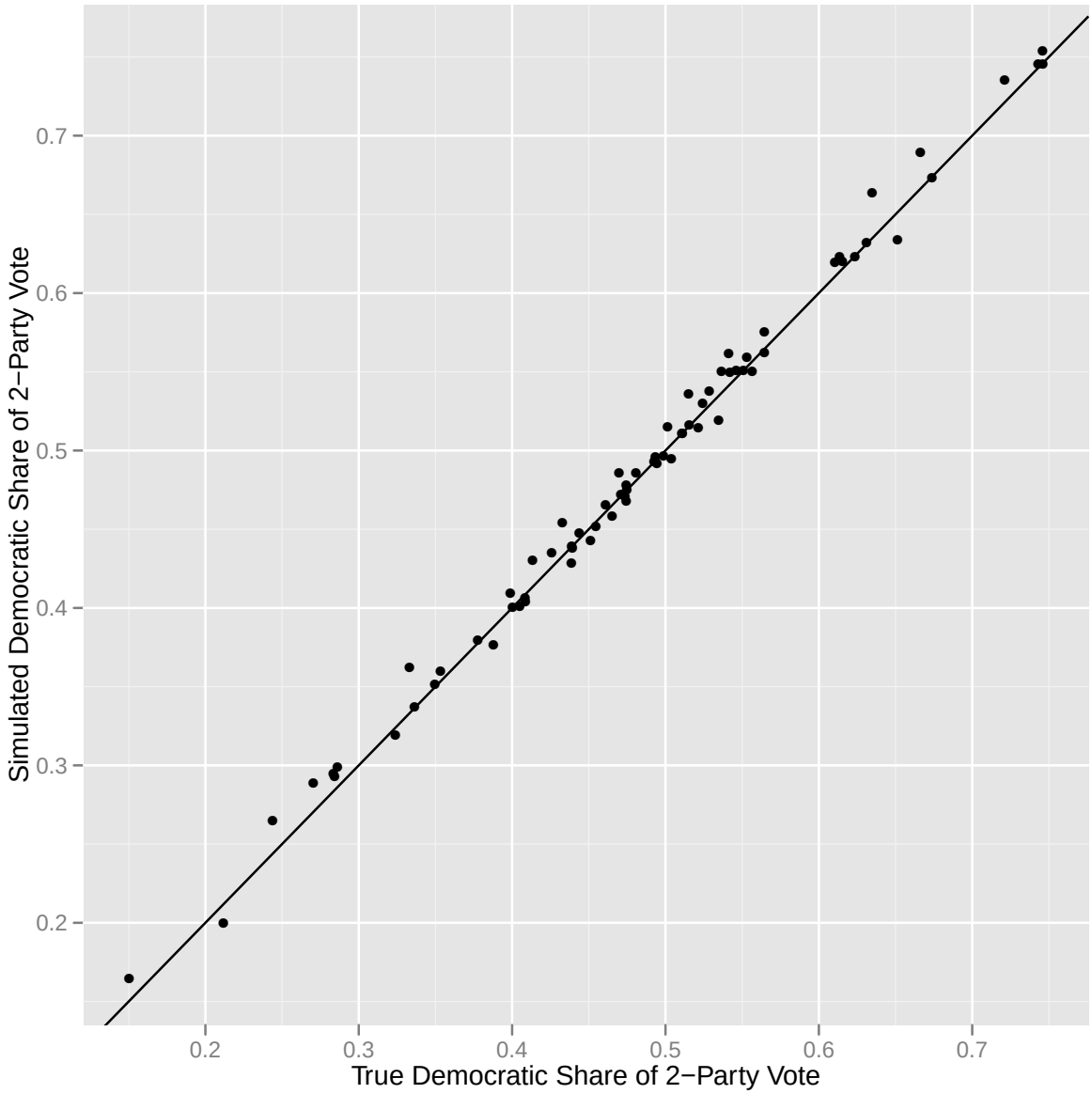

Figure 6: Predicted vs. actual Democratic vote shares.

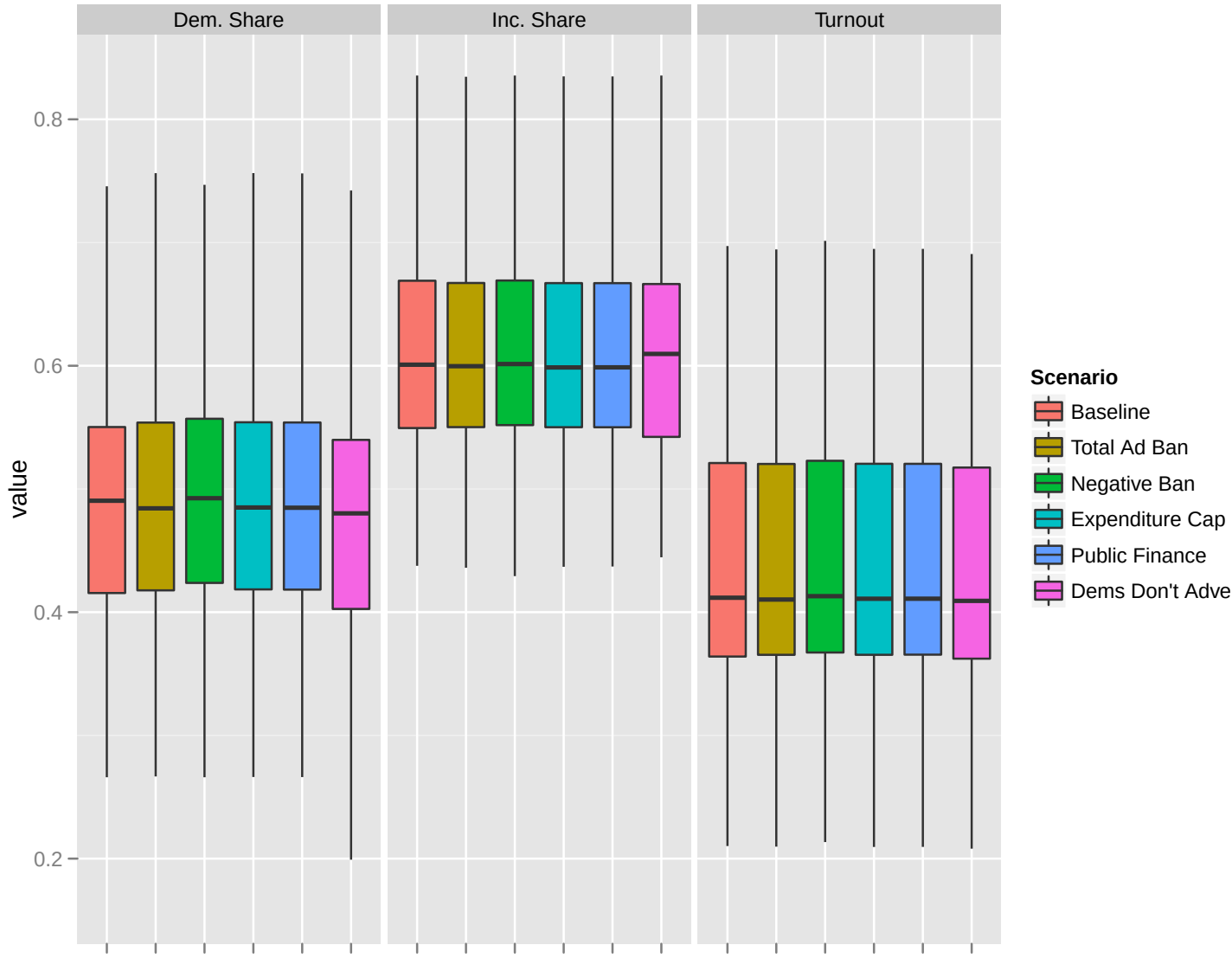

Figure 7: Outcomes under various counterfactual scenarios.

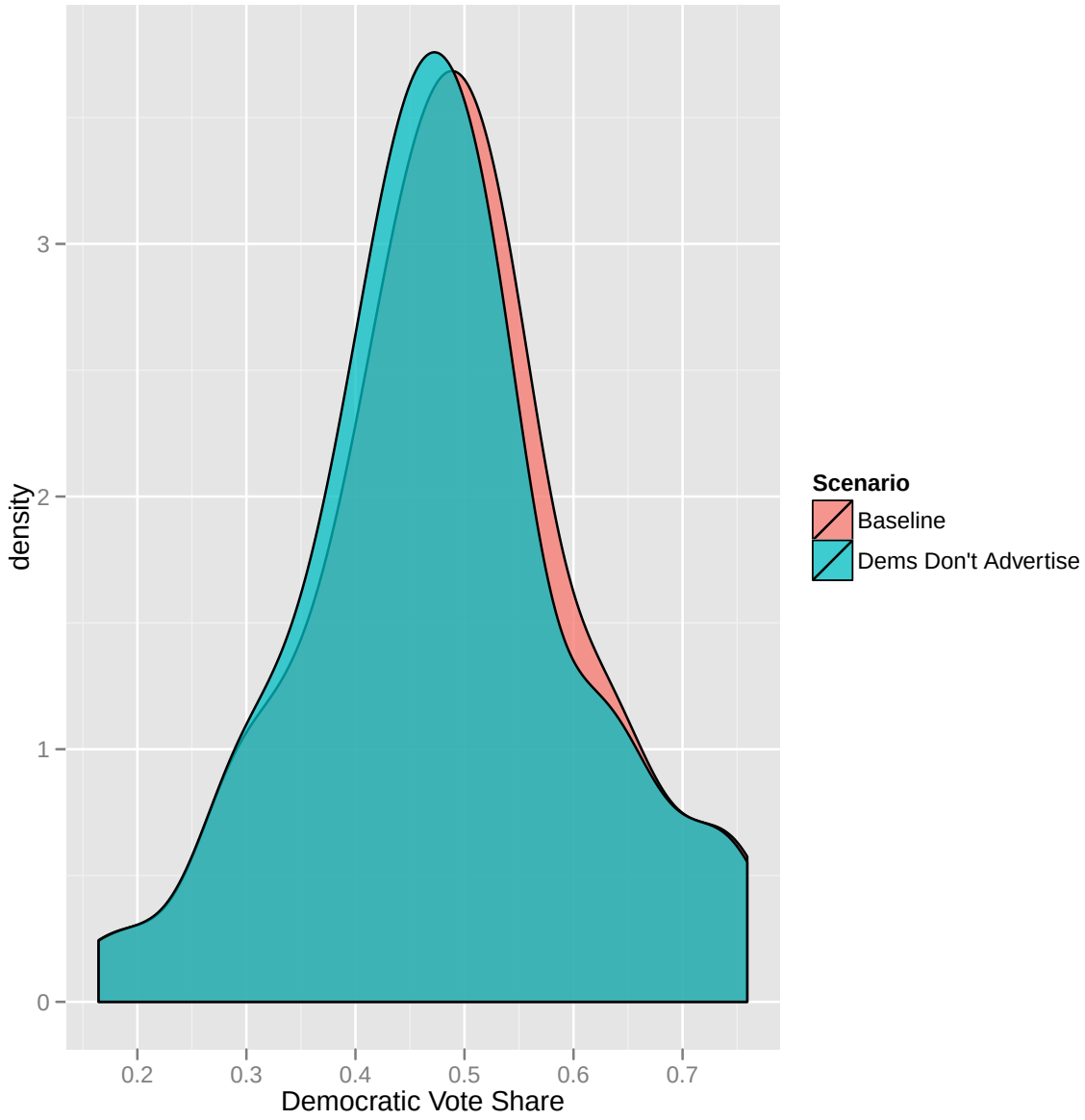

Figure 8: Density of Democratic vote share under baseline and no-Democratic-advertising scenarios.

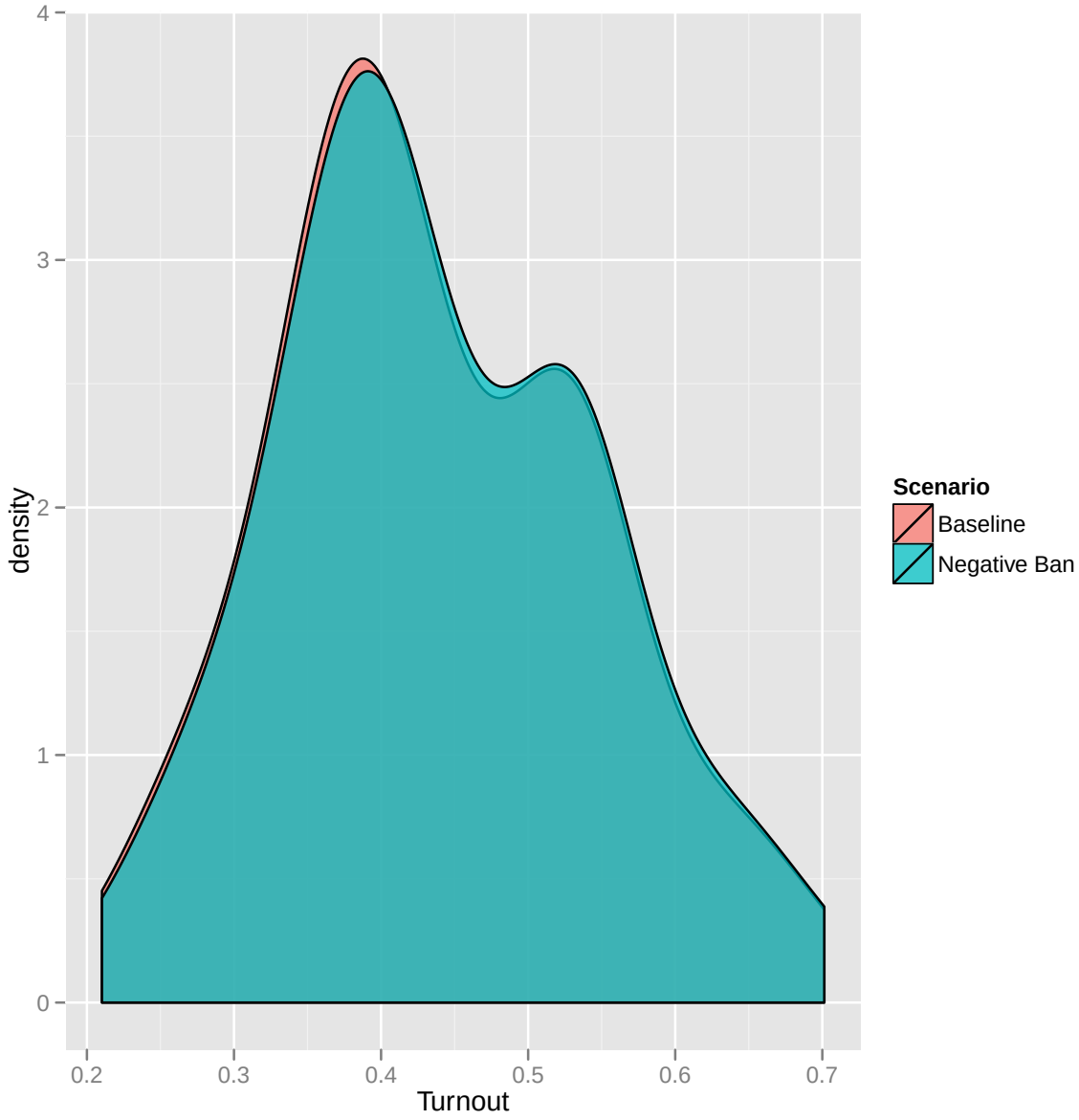

Figure 9: Density of voter turnout under baseline and negative-ban scenarios.

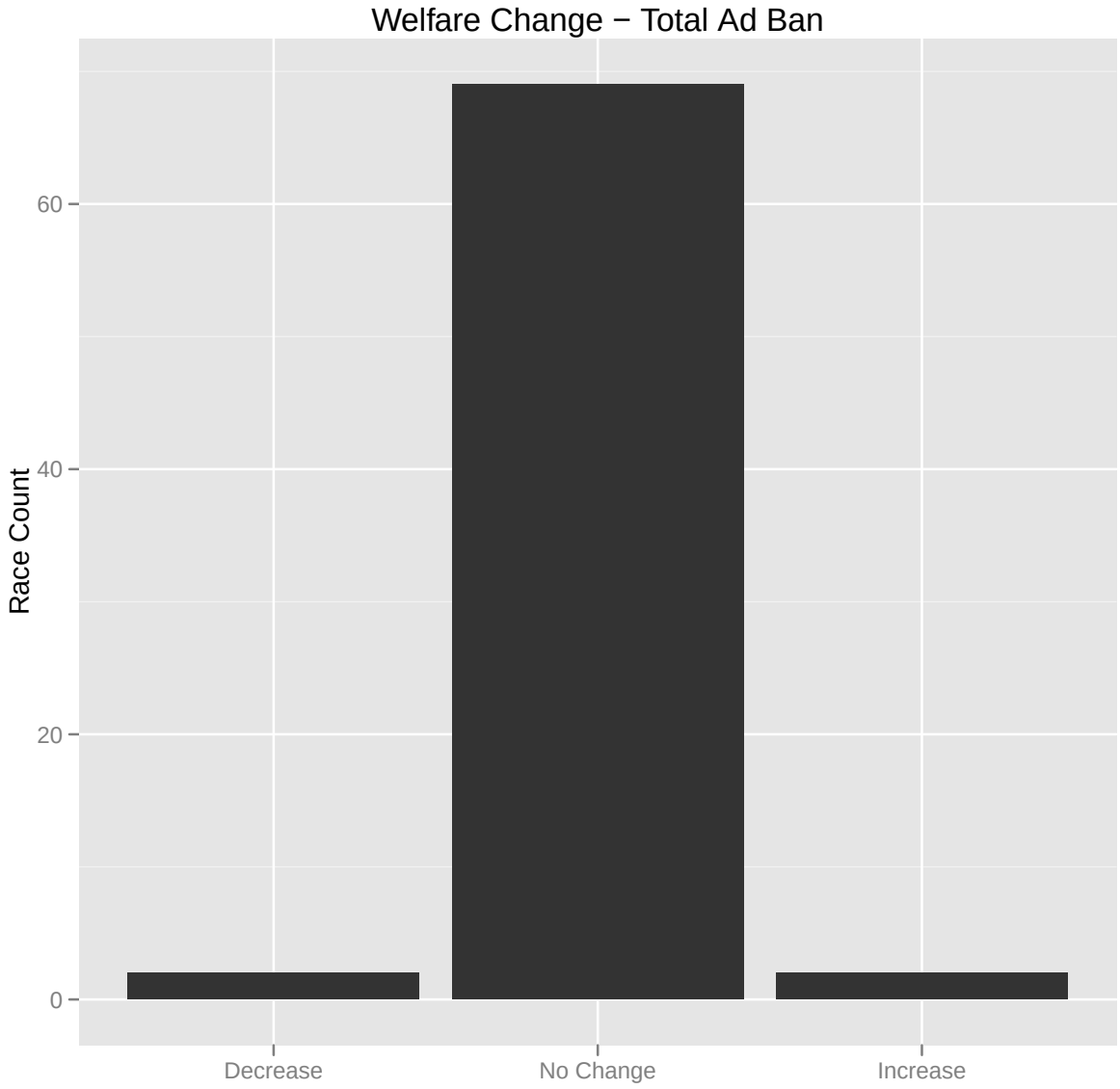

Figure 10: Distribution of change in welfare, by race, between baseline and total-ban scenarios.