NBER WORKING PAPER SERIES

VOTER TURNOUT: THEORY AND EVIDENCE FROM TEXAS LIQUOR REFERENDA

Stephen Coate
Michael Conlin

Working Paper 8720
http://www.nber.org/papers/w8720

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2002

We thank Tim Besley, Mark Bils, Stacy Dickert-Conlin, Tim Feddersen, Ted O'Donoghue, Jan Ondrich, Alvaro Sandroni, and Tim Vogelsang for helpful comments and discussions. We are also indebted to Randy Yarborough at the Texas Alcohol Beverage Commission for providing the local liquor law election data. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Voter Turnout: Theory and Evidence from Texas Liquor Referenda
Stephen Coate and Michael Conlin
NBER Working Paper No. 8720
January 2002
JEL No. D72

ABSTRACT

This paper uses data from Texas liquor referenda to explore a new approach to understanding voter turnout, inspired by the theoretical work of Harsanyi (1980) and Feddersen and Sandroni (2001). It presents a model based on this approach and structurally estimates it using the referendum data. It then compares the performance of the model with two alternative models of turnout. The results are encouraging: the structural estimation yields sensible parameter estimates and the model performs better than the two alternatives considered.

Stephen Coate
Department of Economics
Cornell University
Ithaca, NY 14853
and NBER
sc163@cornell.edu

Michael Conlin
Department of Economics
Syracuse University
Syracuse, NY 13244

1 Introduction

In Texas, as in most of the United States, the sale of liquor is heavily regulated. Moreover, regulations vary across different localities within the state. Some are "dry", completing prohibiting the retail sale of alcoholic beverages. Others permit the sale of only certain types of liquor or require that liquor be consumed "off premise". What is particularly interesting about Texas is that alcohol regulations at the local level are determined directly by the citizens.[1] A citizen wishing to change regulations in his community can get the change voted on in a referendum. Such referenda are commonplace, with over 500 elections between 1976 and 1996.

These liquor referenda appear a promising vehicle for understanding voter turnout. First, turnout varies widely. In some communities over 75% of the voting age population show up to vote, while in others turnout is less than 10%. Second, since there is a limited set of regulations that are actually proposed, the issues decided by the referenda are basically the same across jurisdictions. Third, the referenda are typically held separately from other elections, so that the only reason to go to the polls is to vote on the proposed change in liquor law.

This paper uses data from these referenda to explore a new approach to understanding voter turnout, inspired by the theoretical work of Harsanyi (1980) and Feddersen and Sandroni (2001). It presents a model based on this approach and structurally estimates it using the referendum data. It then compares the performance of the model with two alternative models of turnout. The results are encouraging: the structural estimation yields sensible parameter estimates and the model performs better than the two alternatives considered.

The approach begins with the observation that a referendum creates a contest between two groups: those who support it and those who oppose it. The winner of the contest is the group the most of whose members vote. Individual group members are presumed to want to "do their part" to help their side win. This is not because they receive a transfer from other group members for doing so - they simply adhere to the belief that this is how a citizen should behave in a democracy. "Doing their part" is understood to mean showing up to vote if their costs of going to the polls are below some "reasonable level".

Explaining turnout amounts to understanding what determines these "reasonable levels". Here, in the spirit of Harsanyi (1980) and Feddersen and Sandroni (2001), individuals are assumed to determine their "reasonable levels" by asking themselves the following question: What would be the level of voting costs below which those on my side of the issue should vote that would maximize the payoff of a representative member of my group? Thus, individuals behave according to the rule that would be best for the group as a whole. Rais-

[1] Local jurisdictions do not have complete discretion over liquor regulation. For example, the State of Texas prohibits the sale of alcoholic beverages in certain residential areas and within 300 feet of a church, school or hospital. There are also statewide restrictions on the times alcohol may be sold.

ing the reasonable level creates more turnout and hence raises the probability of the group's preferred outcome. On the other hand, it increases the expected voting costs incurred by group members. Balancing these two considerations determines the reasonable level.

The reasonable level so determined will depend on the turnout expected from the opposing group. The greater the number of votes from the opposition, the higher the reasonable level must be in order to ensure any given chance of success. The reasonable level for individuals in one group will therefore depend on that expected to be chosen by individuals in the opposition. Accordingly, it is natural to think of the reasonable levels as being determined in a game in which individuals from the two groups simultaneously choose their levels. In equilibrium, individuals in each group must be satisfied with their reasonable level, given the level they expect the opposing group to choose.

The equilibrium reasonable levels depend on election specific characteristics like the relative sizes of the two competing groups, the importance of the issue decided by the referendum to the groups, and expected voting costs. Understanding these relationships yields predictions for how turnout should depend on election specific characteristics. In this way, the approach yields a theory of turnout.

The specific model developed in the paper assumes that all supporters enjoy the same benefit and all opposers incur the same cost if the referendum passes. This sidesteps the interesting question of how the burden of voting should be shared among group members with differing intensities of preference. In addition, the fraction of supporters is assumed to be the realization of a random variable with a Beta distribution. Individuals do not observe the realization but do know the parameters of the Beta distribution. This captures the idea that individuals will be aware of general characteristics of their fellow citizens - such as age and religious affiliation - that will influence the likely distribution of supporters. Finally, the cost of voting for each supporter and opposer is assumed to be the realization of an independent random variable uniformly distributed on an identical support. The model is therefore described by five parameters: the benefit of the proposed change to supporters; the cost to opposers; the two parameters of the Beta distribution; and the upper bound of the support of voting costs. The equilibrium reasonable levels for supporters and opposers depend on these parameters and these, together with the realization of the fraction of supporters, determine the turnout of supporters and opposers.

Our data include information on the type of referendum, the votes for and against, and when the referendum was voted on. We also know the size of the voting population and many characteristics of the jurisdiction in question at the time of the election. This includes the religious affiliations of the county population and the liquor regulations in neighboring communities. Finally, we know weather conditions on the day of voting. We specify functional forms which relate our parameters to these characteristics and then estimate the coefficients. In this procedure, we draw on the work of Shachar and Nalebuff (1999).

The two alternative models we consider maintain the same underlying assumptions concerning the environment but postulate different voting behavior.

They are *ad hoc*, in the sense that they do not provide an account of why people behave in the postulated way. Nonetheless, they do capture ideas that have been expressed in the literature. The *intensity hypothesis* says that people are more likely to vote the more intensely they feel about an issue. The *popularity hypothesis* says that individuals are more likely to vote if they believe that many of their fellow citizens share their position on the issue. Using the non-nested hypothesis test of Vuong (1989), we reject the hypotheses that these alternative models and our basic model are equally close to the true data generating process in favor of the alternative hypothesis that our basic model is closer.

The organization of the remainder of the paper is as follows. The next section explains how the paper relates to previous work on voter turnout. Section 3 describes the institutional details concerning the referenda that we study and presents the raw data. Section 4 presents the model. Section 5 describes how we estimate it and section 6 discusses the results. Section 7 outlines the two alternative models and compares their performance with the basic model. Section 8 concludes with suggestions for further research.

2 Relationship to the turnout literature

Understanding voter turnout is a central problem in political economy. Turnout is sensitive to the specific characteristics of elections. Political parties understand this and policy stances are fashioned to "bring out the base" or discourage the opposition's base. Accordingly, turnout not only determines which option wins but also shapes the policy options from which voters select. Reflecting this importance, there has been a considerable amount of work on the subject. Here, we briefly point out where our paper fits into the literature. The reader is referred to Aldrich (1993), (1997), Fiorina (1997), Grossman and Helpman (2001), and Matsusaka and Palda (1993), (1999) for broader overviews and discussion.

The well-known *calculus of voting* model of turnout (Downs (1957), Riker and Ordeshook (1968)) defines the benefits of voting as $pB + d$ where p is the probability of swinging the election, B is the gain from having one's preferred candidate win, and d is the benefit a citizen feels from doing his civic duty or expressing his preference. A voter votes if these benefits exceed the direct cost of voting, denoted c, which includes the time taken to get to the polls and so on. To get a useful theory of turnout, it is necessary to understand how these variables depend on election specific characteristics.

Since the benefits from doing one's duty seem rather nebulous, it is tempting to look at the pB term to understand turnout. The *pivotal-voter* model of Ledyard (1984) and Palfrey and Rosenthal (1985) provides a natural way of endogenizing the probability that a voter will swing the election. However, the obvious problem with this approach would seem to be that p is sufficiently small in any large election that changes in pB are likely to be minuscule across elections. Thus, many have questioned the fruitfulness of a theory of turnout based on minuscule changes in a minuscule number (see, for example, Green and Shapiro (1994)). Formal support for this concern is provided by Palfrey and Rosenthal's well known result that in a sufficiently large electorate the only citizens who vote in equilibrium are those for whom d is no smaller than c.[2] Accordingly, significant variations in turnout in large elections must arise from variations in the fraction of the population for whom d is no smaller than c.

More recently, researchers have turned to the d term. An interesting line of work has assumed that this term can be influenced by leaders (see, for example, Shachar and Nalebuff (1999)). The idea is that in close elections or in elections where there is much at stake, community and political leaders put in more effort exhorting their fellows to vote and this leads to higher turnout. The effort decisions of political leaders are rational because their efforts can sway large groups of voters. Exactly why such exhortions are successful is not clear, which seems a difficulty with the approach.[3]

An alternative strategy is to think more deeply about where individuals' notions of duty in the voting context may come from. Harsanyi (1980) argues that voting may usefully be understood as individuals acting according to the dictates of rule-utilitarianism. A *rule-utilitarian* takes the action that, if taken by all rule-utilitarians, would maximize aggregate utility.[4] Harsanyi illustrates his argument by considering an environment in which a fixed number of votes are needed to pass a policy that would raise aggregate utility. Each citizen faces the same cost of voting and chooses a probability of voting that, if adopted by all, would maximize aggregate utility. The key insight is that the optimal probability is between zero and one. Not everybody should stay home, because that would mean the policy would not pass. But not everybody should vote because that would result in a surfeit of votes, imposing unnecessary costs on society. In this way, the logic of rule utilitarianism yields an elegant theory of turnout. In terms of the calculus of voting model, Harsanyi is effectively assuming that d is large enough so that everyone does their duty but rejects the implicit assumption that doing one's duty always involves voting.

Harsanyi's insight is developed much further by Feddersen and Sandroni (2001). They consider the more relevant environment of a two candidate plurality rule election in which citizens have heterogeneous voting costs. Feddersen and Sandroni first point out a problem with Harsanyi's argument in this context. With two candidates to choose from, a rule utilitarian has to choose not only whether to vote but also for whom to vote. All rule utilitarians would vote for the candidate that maximizes aggregate utility and, accordingly, if only rule utilitarians voted, the optimal voting rule would be such that turnout is minimal. Since all voters would be voting for the same candidate, it is best for society as a whole to minimize the number of individuals incurring voting costs.

To deal with this problem, Feddersen and Sandroni introduce disagreement

[2] Palfrey and Rosenthal's result is for symmetric equilibria in a model where voters are imperfectly informed about each others' voting costs and preferences.

[3] In one of the first papers stressing the importance of group leaders, Uhlaner (1989) assumed that leaders offered transfers to group members in exchange for their votes. However, even when transfers are interpreted most broadly, this practice does not seem particularly widespread.

[4] This is distinct from an *act-utilitarian* who, in any given situation, takes the action that maximizes aggregate utility.

on which candidate maximizes aggregate utility. There are two groups of rule utilitarians with opposing views. Individuals in each group follow a voting rule that, if followed by all in their group, would maximize aggregate utility given the behavior of individuals in the opposing group. Feddersen and Sandroni show that the groups' voting rules prescribe individuals to vote if and only if their voting costs are below some critical level. The two groups' critical cost levels can be derived from a game in which each group member chooses a critical level to maximize the utility of a representative member of his group.

While there are differences in the details, Feddersen and Sandroni's game has the same basic structure as the one studied in this paper.[5] The key difference between the two approaches is thus one of interpretation. In the approach studied here, two candidate elections divide the population into two groups (the supporters of each candidate) and all individuals follow the voting rule that maximizes the payoff of a representative member of their group. They are therefore motivated only to help those on their side of the issue. In Feddersen and Sandroni's model, all individuals who vote follow the voting rule that they believe maximizes aggregate utility. The merit of Feddersen and Sandroni's approach is that all behavior follows from the single postulate that citizens are rule utilitarians. This has significant theoretical appeal. However, in our view, it does not seem particularly plausible to suppose that voters only care about the aggregate benefits associated with different candidates and that all turnout is driven by differences in beliefs. Moreover, it is not clear why citizens should have different beliefs, nor what should determine the relative sizes of the two groups.

The approach developed here is also related to the work of Morton (1987), (1990). She studies a two candidate election and assumes that the population is exogenously divided into groups with different policy preferences. Each group collectively and simultaneously decides how many of its members should vote in order to maximize the group's aggregate benefit. The choice trades off the policy benefit associated with changing the outcome of the election with the cost to members of voting. In Coasian fashion, Morton is not specific on why the groups behave in this way: "The model assumes that groups invest resources (financial or otherwise) which provide group members with the individualized incentives necessary to vote. These resources are then transformed into votes by the groups." (Morton (1987) page 120). This paper's approach may be considered as a special case of Morton's in which there are only two groups - supporters and opposers. While we prefer our interpretation, there is nothing in

[5] There are four main differences in the details. First, in Feddersen and Sandroni's model, the payoff of each group member depends on the expected voting costs of the average citizen in the community as a whole as opposed to the average citizen in his group. This is because group members care about aggregate utility. Second, in the model of this paper, the two groups may differ in the intensity of their preference for their preferred candidates. Third, in Feddersen and Sandroni's model, the fraction of each group who behave "ethically" (i.e., as rule utilitarians) is random. Non-ethical voters abstain. Fourth, in Feddersen and Sandroni's model, ethical voters will only follow the optimal rule if their payoff from ethical behavior (the d term) exceeds their voting cost. As in Harsanyi (1980), the model of this paper implicitly assumes that d is sufficiently large that individuals always do their part.

the empirical work to distinguish it from a story where supporters and opposers collectively determine which of their members should vote.[6]

While there is a large empirical literature on turnout, there are very few papers that try to structurally estimate models of turnout. Hansen, Palfrey and Rosenthal (1987) use data on school budget referenda to try to structurally estimate the pivotal-voter model. Given its complexity, they must make strong assumptions to undertake the estimation. In particular, they assume that the population is equally divided between supporters and opposers and that supporters and opposers have identical benefits from their preferred outcomes. They then estimate the parameters of the distribution of voting costs. Our simpler model permits estimation of the distribution of supporters and opposers, the benefits of supporters and opposers, and the distribution of voting costs.

Shachar and Nalebuff (1999) use state-by-state voting in U.S. presidential elections to structurally estimate a model based on the "follow the leader" approach. Their model assumes that Democratic and Republican leaders in each state expend effort to impact the outcome of the presidential election.[7] Leaders' ability to have an impact depends on how followers respond and on the expected closeness of the race (at both state and national levels). The former is a parameter of the model to be estimated and the latter depends on the distribution of Democrats and Republicans in the population, which is estimated from past election outcomes. The authors conclude that voters do respond to effort and that effort is higher in races that are predicted to be closer. Shachar and Nalebuff's model is an equilibrium model in that the leaders from the two parties in each state choose their effort levels simultaneously. This gives it a similar flavor to our model.

3 Preliminaries

3.1 Institutional background

Chapter 251 of the Texas Alcoholic Beverage Code states that "On proper petition by the required number of voters of a county, or of a justice precinct or incorporated city or town in the county, the Commissioners' Court shall order a local election in the political subdivision to determine whether or not the sale of alcoholic beverages of one or more of the various types and alcoholic contents shall be prohibited or legalized in the county, justice precinct, or incorporated city or town". Thus, citizens can propose changes in the liquor laws of their communities and have their proposals directly voted on in referenda. Such

[6] Morton (1987) endogenizes the policy choices office-seeking candidates would make given this group voting behavior. Building on this work, Filer, Kenny and Morton (1993) present a group voting model where candidates propose tax schemes that differ in the degree of progressiveness. The paper empirically tests the model's qualitiative predictions using county-level turnouts in the 1948, 1960, 1968 and 1980 presidential elections.

[7] One drawback with the study is that, in reality, voters are voting on many other issues at the same time as they are casting their presidential ballot.

direct democracy has a long history in Texas liquor regulation, with local liquor elections dating back to the mid 1800s.[8]

The process by which citizens may propose a change for their jurisdiction is relatively straightforward. The first step involves applying to the Registrar of Voters for a petition. This only requires the signatures of ten or more registered voters in the jurisdiction. The hard work comes after receipt of the petition. The applicants must get it signed by at least 35% of the registered voters in the jurisdiction and must do this within thirty days.[9] If this hurdle is successfully completed, the Commissioners' Court of the county to which the jurisdiction belongs must order a referendum be held. This order must be issued at its first regular session following the completion of the petition and the referendum must be held between twenty and thirty days from the time of the order. All registered voters can vote and if the proposed change receives at least as many affirmative as negative votes, it is approved.

Citizens may propose changes for their entire county, their justice precinct, or the city or town in which they reside. The state is divided into 254 counties and each county is divided into justice precincts.[10] Accordingly, a justice precinct lies within the county to which it belongs. By contrast, a city may spillover into two or more justice precincts. If only part of a city belongs to a particular justice precinct that has approved a change, then that part must abide by the new regulations. However, if the city then subsequently approved a different set of regulations, then they would also be binding on the part contained in the justice precinct in question. Effectively, current regulations are determined by the referendum most recently approved. Over our data period, citizens almost always choose to propose changes at the city or justice precinct level rather than at the county level.

Importantly for the purposes of our study, liquor referenda are typically held separate from other elections. Section 41.01 of the Texas Election Laws sets aside four dates each year as uniform election dates.[11] These are the dates when presidential, gubernatorial, and congressional elections are held. In addition, other issues are often decided on these days such as the election of aldermen, and the approval of the sale of public land and bond issuances. Elections pertaining to these other issues may occur, but rarely do, on dates other than uniform election days. Liquor referenda, in contrast, do not typically occur on uniform election dates. This reflects the tight restrictions placed by Chapter 251 on the timing of elections.[12]

[8] From 1919 to 1935, these elections were abolished as a result of prohibition. Since 1935, the process of citizen-democracy has been governed by the procedures described in Chapter 251 of the Texas Alcoholic Beverage Code.

[9] Prior to 1993, the number of signatures needed was 35% of the total number of votes cast in the last preceding gubernatorial election.

[10] The number of justice precincts in a county range from 1 to 8.

[11] These are the third Saturday in January, the first Saturday in May, the second Saturday in August and the first Tuesday after the first Monday in November.

[12] Interestingly, the Texas state government voted in 2001 to require liquor law referendum votes to occur on one of the four uniform election dates. This was to avoid the costs of holding referenda separately.

3.2 Data

We assembled data on 363 local liquor elections in Texas between 1976 and 1996 where prior to the election the voting jurisdictions prohibited the retail sale of all alcohol.[13] Information on these elections was obtained from the annual reports of the Texas Alcoholic Beverage Commission (TABC). These reports contain the county, justice precinct, city or town voting on the referendum, the date of the election, the proposed change, and the number of votes cast for and against. As indicated in Table 1, the elections differed in the degree to which restrictions were relaxed: 147 proposed permitting the selling of beer only or beer and wine; 144 proposed permitting the sale of all alcoholic beverages for off-premise consumption only (i.e., liquor stores); and 72 proposed not only that all beverages be sold but they may be consumed off- and on-premise (i.e., bars as well as liquor stores). Of these 363 elections, 2 were at the county level, 133 were at the justice precinct level and 228 were at the city or town level. While not indicated in Table 1, at least one election occurred in 125 different counties. While certain counties had multiple local liquor elections during this 20 year period, approximately two-thirds of the 363 elections involved jurisdictions which account for a single election. For those jurisdictions that had multiple elections, these elections often occur a number of years apart.

We supplemented our election data with information on county-, city- and town-level populations, by age, obtained from the United States Census. Using this information, we estimate the voting age population and the population over the age of 50 at the time of an election. Table 1 indicates that the mean voting age population in the 363 jurisdictions is 4,415 at the time of the elections.

We also attempted to find information that might tell us about the attitudes of citizens towards the selling of alcohol. Using the county-level population and county-level information on the number of adherents to Baptist denominations from *Churches & Church Membership in the United States*, we constructed estimates of the fraction of the county population that is Baptist at the time of an election. We use this fraction as a proxy for the fraction of Baptists in each of the 363 jurisdictions; thereby, implicitly assuming that Baptists are uniformly distributed throughout each county. As indicated in Table 1, the average fraction of the county population that is Baptist is 0.48.

Information about what type of alcohol could be sold and where it could be consumed elsewhere in the county was obtained from the annual reports of the TABC. The alcohol policy being voted on in a third of the elections is more liberal than the alcohol policy in the rest of the county. Monthly information on the number of alcohol related road accidents in each county was obtained from the Texas Department of Public Safety. The average number of alcohol related accidents per capita in a county for the twelve months prior to an election is 0.00204. Finally, whether the jurisdiction is located in a Metropolitan Statistical Area (MSA) was determined using classifications obtained from the 1996 United States Census. Table 1 shows that 44% of the elections involved jurisdictions

[13] See the data appendix for a description of this data collection process. The appendix also contains a detailed explanation of how certain variables are created.

located in an MSA.

In an attempt to get information about the costs of voting, daily weather conditions at 44 weather stations in Texas was obtained from the United States Carbon Dioxide Information Analysis Center (CDIAC). The weather conditions on the day of each election are taken to be the same as those measured at a weather station in close proximity to the voting jurisdiction. While many of these elections occurred on rainy days, as expected, few occurred on days when snow fell (see Table 1). Besides the weather, whether the election occurred on a weekend and whether the election occurred in the summer are also likely to affect the costs of voting. The majority of the 363 elections occurred on a weekend while slightly more than a quarter occurred in June, July or August.

3.3 Some basic facts

Of our 363 referenda, 150 were approved and 213 were rejected by the voters. The percent of the voting population that voted for the referendum averaged 17% across the 363 elections while the percent voting against averaged 19%. The average turnout in these elections (calculated by dividing total votes by voting age population) is 0.36 but there is substantial variation across elections. Figure 1 presents the turnout information in a histogram where the vertical axis measures the number of elections in each turnout category. While a number of elections had turnout rates over 0.75, the majority had less than a third of the voting age population vote. Interestingly, average turnout is significantly higher in city-level elections (0.44 compared to 0.21). In addition, average turnout was significantly higher when the referendum involved off-premise consumption of all alcohol as opposed to just beer and wine or off- and on-premise consumption.

The elections tended to be close. When closeness is defined as the difference between votes for and against divided by total votes, the average closeness is 0.25. The histogram of this measure is depicted in Figure 2 and demonstrates that while the majority of the elections are relatively close, there are outliers. Unlike turnout, the average value of this measure of closeness does not differ significantly between city-level elections and justice precinct- or county-level elections. It is also similar across the different types of regulatory changes.

It is natural to ask whether the data support the familiar idea that turnout is higher in close elections. This all depends on how we measure closeness. Proceeding as in Figure 2, there is a slight positive relationship (correlation coefficient of 0.12). This positive relationship is stronger if closeness is defined as the difference between votes for and against divided by the voting age population (correlation coefficient of 0.58). However, there is a negative relationship between turnout and closeness when closeness is defined as the difference between votes for and against (correlation coefficient of -0.11).

4 The Model

Consider a community that is holding a referendum on relaxing liquor laws. For analytical tractability, we adopt the fiction that the community has a continuum of citizens. These citizens are divided into supporters and opposers. Each supporter is willing to pay b for the relaxation, while each opposer is willing to pay x to avoid it.

Each citizen knows whether he is a supporter or an opposer, but not the fraction of citizens in each category. However, all citizens know that the fraction of supporters in the population, denoted μ, is the realization of a random variable with range $[0, 1]$ distributed according to the *Beta Distribution*.[14] Thus, the probability density function of the random variable is

$$h(\mu; \nu, \omega) = \mu^{\nu-1}(1-\mu)^{\omega-1}/B(\nu, \omega)$$

where ν and ω are parameters known by the citizens and $B(\nu, \omega)$ is the Beta function

$$B(\nu, \omega) = \int_0^1 \mu^{\nu-1}(1-\mu)^{\omega-1}d\mu.$$

The expected fraction of supporters under this distributional assumption is $\nu/(\nu+\omega)$ and the variance is $\nu\omega/[(\nu+\omega)^2(\nu+\omega+1)]$. We will assume that both ν and ω exceed 1 which implies that the density is hump shaped.

Citizens must decide whether or not to vote in the referendum. If they do, supporters vote in favor and opposers vote against. Voting is costly, with each citizen i facing a cost of voting c_i where c_i is the realization of a random variable uniformly distributed on $[0, c]$. Citizens know c but do not observe the voting costs of their fellows. We assume that individuals want to "do their part" to help their side win, where doing one's part means showing up to vote if one's costs of voting are below some "reasonable level". Individuals determine these reasonable levels by asking themselves "what would be the critical level of voting costs below which those on my side of the issue should vote that would maximize the payoff of a representative member of my group?".

Letting the critical voting costs for the two groups be denoted by γ_s and γ_o, if citizen i is a supporter he votes if $c_i \leq \gamma_s$ and if he is an opposer he votes if $c_i \leq \gamma_o$. The probability that a supporter votes is the probability that γ_s exceeds c_i, which is γ_s/c. Similarly, the probability that an opposer votes is γ_o/c. Thus, the referendum passes when $\mu\gamma_s/c > (1-\mu)\gamma_o/c$ or, equivalently, when $\mu > \gamma_o/(\gamma_s+\gamma_o)$. The probability that the referendum passes is therefore:

$$\pi(\gamma_s, \gamma_o; \nu, \omega) = \int_{\frac{\gamma_o}{\gamma_s+\gamma_o}}^1 h(\mu; \nu, \omega)d\mu.$$

[14] Shachar and Nalebuff (1999) model uncertainty in the fraction of the population who are Democrats in a similar way. However, they assume that the fraction of Democrats is the realization of a random variable with a normal distribution. This has the obvious drawback that it can take on values outside the interval $[0, 1]$.

Prior to the realization of his voting cost, a supporter's expected payoff from the critical level γ_s given that the reasonable level of opponents is γ_o is given by:

$$\pi(\gamma_s, \gamma_o; \nu, \omega)b - \frac{\gamma_s^2}{2c}.$$

The first term represents the expected policy benefits stemming from the referendum passing and the second term represent expected voting costs, given that a supporter will vote only if $c_i \leq \gamma_s$.[15] Similarly, an opposer's expected payoff from the critical level γ_o given that supporters have reasonable level γ_s is

$$-\pi(\gamma_s, \gamma_o; \nu, \omega)x - \frac{\gamma_o^2}{2c}.$$

Accordingly, we define a pair of critical levels (γ_s^*, γ_o^*) to be an *equilibrium* if

$$\gamma_s^* \in \arg\max_{\gamma_s \in [0,c]} \{\pi(\gamma_s, \gamma_o^*; \nu, \omega)b - \frac{\gamma_s^2}{2c}\}$$

and

$$\gamma_o^* \in \arg\max_{\gamma_o \in [0,c]} \{-\pi(\gamma_s^*, \gamma_o; \nu, \omega)x - \frac{\gamma_o^2}{2c}\}.$$

We say that (γ_s^*, γ_o^*) is an *interior equilibrium* if both γ_s^* and γ_o^* are between 0 and c.

We are now able to establish the following result.[16]

Proposition 1 *If (γ_s^*, γ_o^*) is an interior equilibrium, then*

$$\gamma_s^* = \left(\frac{c(\sqrt{x})^{\nu}(\sqrt{b})^{\omega+2}}{(\sqrt{x}+\sqrt{b})^{\nu+\omega}B(\nu,\omega)}\right)^{\frac{1}{2}},$$

and

$$\gamma_o^* = \left(\frac{c(\sqrt{x})^{\nu+2}(\sqrt{b})^{\omega}}{(\sqrt{x}+\sqrt{b})^{\nu+\omega}B(\nu,\omega)}\right)^{\frac{1}{2}}.$$

This proposition shows that, if there exists an interior equilibrium, it is unique and, moreover, the equilibrium reasonable levels are related to the parameters in a relatively simple way. As intuition would suggest, both reasonable levels are increasing in the maximal voting cost c, although γ_s^*/c and γ_o^*/c are decreasing in c. The reasonable level for each group is increasing in the gains or losses to that group caused by the referendum passing.[17]

[15] Given that a supporter will vote if and only if his voting cost is less than γ_s, his expected voting costs are $\int_0^{\gamma_s} c_i \frac{dc_i}{c} + \int_{\gamma_s}^{c} 0 \frac{dc_i}{c}$ which simplifies down to $\frac{\gamma_s^2}{2c}$.

[16] The proofs of this and the next proposition are in the Appendix.

[17] If the model were interpreted in the same way as in Feddersen and Sandroni (2001), μ would be the fraction of the population who believed that the change would be good for society, b would be their estimate of the average net benefit of the change, and x would be the estimate of the average net cost of the change of those who thought it was a bad idea. The objective

While providing a nice characterization of the equilibrium, the proposition leaves open the question of existence. There is no general guarantee that an equilibrium will exist - the payoff functions of supporters and opposers are not quasi-concave functions of their own reasonable cost levels. Indeed, it is not difficult to find parameter values for which no equilibrium exists.[18] In such circumstances, one of the critical levels described in the proposition is not a best response for the group in question. This is typically because it would be better for that group not to vote at all than to vote for cost levels below the reasonable level identified in the proposition. This arises, for example, when one group (say, supporters) is expected to be much smaller than the other (i.e., $\nu/(\nu+\omega)$ is small). While the cost level in the proposition is always positive and implies a positive level of turnout, if supporters are very unlikely to win they may be better off just giving up and staying home. But then if supporters are staying home, the optimal reasonable cost for opposers becomes very small, which then provides supporters an incentive to vote.

When can we be sure that a pair of critical cost levels satisfying the conditions of Proposition 1 is actually an equilibrium? Our next proposition provides some useful sufficient conditions.

Proposition 2 *Suppose that $(\gamma_s^*, \gamma_o^*) \in (0, c]^2$ satisfies (i) the conditions of Proposition 1, (ii) the "second order" conditions*

$$(\nu+2)\gamma_s^* > (\omega-2)\gamma_o^* \ \ \& \ (\omega+2)\gamma_o^* > (\nu-2)\gamma_s^*,$$

and (iii) the "better than staying home" conditions

$$\pi(\gamma_s^*, \gamma_o^*; \nu, \omega)b \geq (\gamma_s^*)^2/2c \ \ \& \ (1-\pi(\gamma_s^*, \gamma_o^*; \nu, \omega))x \geq (\gamma_o^*)^2/2c.$$

Then (γ_s^*, γ_o^*) is an equilibrium.

The second order conditions in (ii), together with the conditions of Proposition 1, imply that the payoff functions of supporters and opposers are locally strictly concave at (γ_s^*, γ_o^*). The better than staying home conditions in (iii) ensure that at (γ_s^*, γ_o^*) the payoffs of supporters and opposers are at least as high as if they simply choose not to vote. The proof of the proposition amounts to showing

function of those favoring the change would be $\pi(\gamma_s, \gamma_o; \nu, \omega)b - (\frac{\nu}{\nu+\omega}\frac{\gamma_s^2}{2c} + \frac{\omega}{\nu+\omega}\frac{\gamma_o^2}{2c})$ and that of those opposed to the change would be $-\pi(\gamma_s, \gamma_o; \nu, \omega)x - (\frac{\nu}{\nu+\omega}\frac{\gamma_s^2}{2c} + \frac{\omega}{\nu+\omega}\frac{\gamma_o^2}{2c})$. Equilibrium would be defined in the same way and conditions analagous (but different) to those presented in Proposition 1 can be obtained. Since the two interpretations have different implications for the equilibrium reasonable costs, it is in priniciple possible to ask which best fits the data. However, empirically, it is not clear what the parameters ν and ω should depend on in the Feddersen and Sandroni interpretation.

[18] Feddersen and Sandroni (2001) deal with this existence problem in their model by assuming that the fraction of individuals in each group who behave "ethically" (i.e., according to the the dictates of rule utilitarianism) is uncertain. Under the assumption that the two groups care equally intensely about the election, Feddersen and Sandroni show that equilibrium exists and is unique if the fraction of "ethicals" in each group is uncertain, independent, and uniformly distributed.

that the payoff functions can have at most one interior local maximum in which case these three conditions are sufficient to imply that γ_s^* is a best response to γ_o^* and vice versa.

In our empirical work, we estimate the determinants of the exogenous variables (b, x, c, ν, ω) assuming that supporters and opposers use the reasonable levels described in Proposition 1. Of course, this is only legitimate if these are indeed equilibrium reasonable levels. We can check this using Proposition 2. Our estimates imply values of the exogenous variables (b, x, c, ν, ω) for each jurisdiction which, in turn, imply values of the reasonable costs (γ_s^*, γ_o^*) via the equations of Proposition 1. If these implied values satisfy the second order conditions and the better than staying home conditions of Proposition 2, then we know that (γ_s^*, γ_o^*) really are equilibrium reasonable costs given (b, x, c, ν, ω).

5 Estimation

We assume that for each jurisdiction j, $\nu_j = 1 + \exp(\beta_v \cdot z_{vj})$ and $\omega_j = 1 + \exp(\beta_\omega)$ where β_ν is a vector of parameters to be estimated, z_{vj} is a vector of jurisdiction specific characteristics that may influence the mix of supporters and opposers, and β_ω is a parameter to be estimated. We further assume that $x_j = \exp(\beta_x \cdot z_{xj} + \varepsilon_j)$ and $b_j = \exp(\beta_b \cdot z_{bj} + \varepsilon_j)$ where β_x and β_b are vectors of parameters to be estimated, z_{xj} and z_{bj} are vectors of jurisdiction and referendum specific characteristics that may affect supporters' benefits and opposers' costs and ε_j is the realization of some random variable distributed according to the standard normal distribution. The assumption that ε_j is a common shock to both supporters' benefits and opposers' costs allows us to derive the likelihood function. Finally, we assume that $c_j = \exp(\beta_c \cdot z_{cj})$ where β_c is a vector of parameters to be estimated and z_{cj} is a vector of jurisdiction and referendum specific variables that may impact voting costs. The functional forms are selected to ensure that the Beta distribution parameters ν_j and ω_j are greater than one and that x_j, b_j, and c_j are non-negative.

Our task is to estimate the parameters $\Omega = \{\ \beta_\nu,\ \beta_\omega,\ \beta_x,\ \beta_b,\ \beta_c\}$. To construct the likelihood function, fix Ω and consider a particular jurisdiction j. Suppose that the fraction of the population voting in favor of the referendum is v_{sj} and against is v_{oj}. Then, according to the model, $v_{sj} = \mu_j \gamma_{sj}^*/c_j$ and $v_{oj} = (1 - \mu_j)\gamma_{oj}^*/c_j$ where μ_j is the fraction of the voting population who are supporters and γ_{sj}^* and γ_{oj}^* are the equilibrium reasonable cost levels for supporters and opposers.

Using the formulas presented in Proposition 1, it follows that

$$v_{sj} = \frac{\mu_j}{c_j}\left(\frac{c_j(\sqrt{x_j})^{\nu_j}(\sqrt{b_j})^{\omega_j+2}}{(\sqrt{x_j}+\sqrt{b_j})^{\nu_j+\omega_j}B(\nu_j,\omega_j)}\right)^{\frac{1}{2}}$$

and that

$$v_{oj} = \frac{(1-\mu_j)}{c_j}\left(\frac{c_j(\sqrt{x_j})^{\nu_j+2}(\sqrt{b_j})^{\omega_j}}{(\sqrt{x_j}+\sqrt{b_j})^{\nu_j+\omega_j}B(\nu_j,\omega_j)}\right)^{\frac{1}{2}}$$

Substituting in our functional forms and rearranging, we can write these as:

$$v_{sj} = \mu_j \sqrt{\exp \varepsilon_j} K_j$$

and

$$v_{oj} = (1 - \mu_j) \sqrt{\exp \varepsilon_j} K_j (\frac{\sqrt{\widehat{x}_j}}{\sqrt{\widehat{b}_j}}),$$

where $\widehat{b}_j = \exp(\beta_b \cdot z_{bj})$, $\widehat{x}_j = \exp(\beta_x \cdot z_{xj})$ and

$$K_j = (\frac{(\sqrt{\widehat{x}_j})^{\nu_j} (\sqrt{\widehat{b}_j})^{\omega_j + 2}}{(\sqrt{\widehat{x}_j} + \sqrt{\widehat{b}_j})^{\nu_j + \omega_j} B(\nu_j, \omega_j) c_j})^{\frac{1}{2}}.$$

We can now solve these two equations for the realizations of μ_j and ε_j implied by any given choice of parameters Ω. In this way, we obtain:

$$\mu_j = \frac{v_{sj} \sqrt{\widehat{x}_j}}{v_{oj} \sqrt{\widehat{b}_j} + v_{sj} \sqrt{\widehat{x}_j}}$$

and

$$\varepsilon_j = 2 \ln(v_{oj} \sqrt{\widehat{b}_j} + v_{sj} \sqrt{\widehat{x}_j}) - 2 \ln(K_j \sqrt{\widehat{x}_j}).$$

These equations define μ_j and ε_j as functions of the turnouts (v_{sj}, v_{oj}). Using the distributions of μ_j and ε_j, we can now compute the probability of observing any pair of turnouts (see the Appendix for the derivation). Letting $Z_j = (z_{xj}, z_{bj}, z_{vj}, z_{cj})$, the probability density function for (v_{sj}, v_{oj}) is

$$g(v_{sj}, v_{oj} \mid \Omega, Z_j) = \frac{2(\sqrt{\widehat{x}_j})^{\nu_j} (v_{sj})^{\nu_j - 1} (\sqrt{\widehat{b}_j})^{\omega_j} (v_{oj})^{\omega_j - 1}}{(v_{oj} \sqrt{\widehat{b}_j} + v_{sj} \sqrt{\widehat{x}_j}))^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}),$$

where

$$\begin{aligned} \varsigma_j \quad = \quad & (\nu_j + \omega_j) \ln(\sqrt{\widehat{x}_j} + \sqrt{\widehat{b}_j}) + \ln B(\nu_j, \omega_j) + \ln c_j \\ & + 2 \ln(v_{oj} \sqrt{\widehat{b}_j} + v_{sj} \sqrt{\widehat{x}_j}) - (\nu_j + 2) \ln \sqrt{\widehat{x}_j} - (\omega_j + 2) \ln \sqrt{\widehat{b}_j}. \end{aligned}$$

Accordingly, our *likelihood function* is

$$L(\Omega) = \prod_{j=1}^{J} g(v_{sj}, v_{oj} \mid \Omega, Z_j). \qquad (1)$$

Any given estimate of the parameters $\Omega = \{ \beta_\nu, \beta_\omega, \beta_x, \beta_b, \beta_c \}$ implies values of the exogenous variables for each jurisdiction j.[19] These, in turn, imply

[19] Thus, $\nu_j = 1 + \exp(\beta_v \cdot z_{vj})$, $\omega_j = 1 + \exp(\beta_\omega)$, and $c_j = \exp(\beta_c \cdot z_{cj})$. Moreover, $x_j = \exp(\beta_x \cdot z_{xj} + \varepsilon_j)$ and $b_j = \exp(\beta_b \cdot z_{bj} + \varepsilon_j)$ where $\varepsilon_j = 2 \ln(v_{oj} \sqrt{\widehat{b}_j} + v_{sj} \sqrt{\widehat{x}_j}) - 2 \ln(K_j \sqrt{\widehat{x}_j})$.

values of the reasonable costs via the equations of Proposition 1. Unconstrained maximization of the likelihood function generates parameter estimates which, for some jurisdictions, imply values of the reasonable costs that exceed the maximum possible cost. Since this is clearly inconsistent with the model, we must maximize the likelihood function subject to the feasibility constraints that $\gamma^*_{sj} \leq c_j$ and $\gamma^*_{oj} \leq c_j$ for each jurisdiction j.[20]

To see how to impose these constraints, observe that for each jurisdiction j

$$\frac{\gamma^*_{sj}}{c_j} = \frac{v_{sj}}{\mu_j} = \frac{v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}}{\sqrt{\widehat{x}_j}},$$

and that

$$\frac{\gamma^*_{oj}}{c_j} = \frac{v_{oj}}{1-\mu_j} = \frac{v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}}{\sqrt{\widehat{b}_j}}.$$

Using these, the feasibility constraint for jurisdiction j can be written as

$$(\frac{v_{sj}}{1-v_{oj}})^2 \leq \frac{\widehat{b}_j}{\widehat{x}_j} \leq (\frac{1-v_{sj}}{v_{oj}})^2.$$

Substituting in for $\widehat{b}_j$ and $\widehat{x}_j$, yields

$$\ln(\frac{v_{sj}}{1-v_{oj}})^2 \leq \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj} \leq \ln(\frac{1-v_{sj}}{v_{oj}})^2. \qquad (2)$$

We can now solve for the parameters that maximize the likelihood function subject to these constraints.[21] As noted in the previous section, we may then use Proposition 2 to check whether the estimated reasonable cost levels are actually an equilibrium given the values of the exogenous variables. Remarkably, this is the case for every jurisdiction.

6 Results

The empirical results of the basic model are shown in Tables 2 and 3. Table 2 presents the parameter estimates that maximize the likelihood function in Equation (1) subject to the constraints specified in Equation (2). Using these

[20] By imposing these feasibility constraints, we are requiring that the choice of parameters must satisfy the conditions of Proposition 1 when either γ^*_{sj} and γ^*_{oj} equals c_j. This restricts the choice of parameters in a marginally tighter way than is implied by the model. This is because the conditions of Proposition 1 need not be satisfied if either group's reasonable cost level is at the boundary. In the boundary case, the first order conditions are in the form of weak inequalities rather than equalities. Since this dampens the ability of the model to fit the data, it will in no way compromise our conclusions about the relative performance of the model.

[21] The Appendix contains a detailed description of how we solved the constrained maximization problem.

parameter estimates, Table 3 presents some aggregate information about the implied values of the model's exogenous variables for the 363 jurisdictions.

We expect the fraction of supporters in a jurisdiction to be a function of the fraction of baptists, the fraction of citizens over the age of 50, the number of alcohol related accidents in the prior year, and the jurisdiction's perception of the effects of a less restrictive alcohol policy. This perception is likely to depend on whether the jurisdiction is a city (compared to a justice precinct or county) and whether the jurisdiction is located in an MSA. Therefore, we allow the model's Beta distribution to vary across jurisdictions by specifying ν as a function of these five variables.

The coefficient estimates in Table 2 imply that the average expected percentage of supporters across the 363 jurisdictions is 5.07/(5.07+4.14)=55% (see Table 3). The estimates also indicate that increasing by ten percent the fraction of baptists or the fraction of voting age population over the age of 50, decreases supporters by approximately one percent. This suggests that both baptists and people over 50 years old are ten percent more likely to oppose the referendum than non-baptists and people under the age of 50, respectively.[22] The number of prior year's alcohol related accidents in the county increases the fraction of supporters. This is consonant with Baughman, Conlin, Dickert-Conlin and Pepper (2000) who find that the number of alcohol related accidents in Texas counties may actually decline with a less restrictive alcohol policy.[23] In addition, the fraction of supporters is slightly less in cities and significantly less in jurisdictions located in an MSA. The large negative coefficient associated with the MSA variable implies that the fraction of supporters is eight percent less, on average, if the jurisdiction is located in an MSA. The reason for this might be that a jurisdiction that allows the sale of alcohol attracts more outsiders if it is in an urban compared to a rural area. While some residents may perceive this as a benefit, others may be concerned about the type of people the alcohol would attract. Alternatively, it may be that black market liquor is more readily available in urban areas.

We allow supporters' benefit and opposers' cost of a passed referendum to depend on the type of referendum, whether the jurisdiction voting is a city and whether passing the referendum would result in the jurisdiction having a more liberal alcohol policy than any other jurisdiction in the county. The coefficient estimates indicate that all of these factors have large and statistically significant effects. As for the type of referendum, Table 2 shows that the supporters' benefit and opposers' costs are greater when the vote pertains to off-premise consumption of all alcohol than when it involves beer and wine. The average marginal effects are to increase the supporter's benefit by 0.17 and the opposers'

[22] The fact that these variables are measured at the county-level and religious affiliations are available at the county level only in 1970, 1980 and 1990, makes these noisy measures of the fraction of baptists and the fraction of people over the age of 50 in the jurisdiction at the time of the election. This may explain why their coefficients are not statistically significant.

[23] While these law changes decrease the implicit price of alcohol for the jurisdiction, they also reduce the travel distance required to obtain the alcohol. Baughman, Conlin, Dickert-Conlin and Pepper (2000) find that for certain alcohol policy liberalizations, this second effect dominates in regards to alcohol related accidents.

costs by 0.10. These are relatively large given the average benefits and costs to supporters and opposers (see Table 3). While the positive coefficients associated with off-premise consumption were expected, the negative coefficients associated with off- and on-premise consumption of all alcohol were not.

The positive coefficients associated with a referendum being voted on by a city suggests that the benefits and costs to voters of a less restrictive alcohol policy are greater in city-level elections. A possible explanation is that residents in more densely populated areas (such as cities compared to justice precincts) are more likely to feel the impact of an alcohol policy liberalization. Referenda involving a more liberal alcohol policy than exists in the rest of the county also increases supporters' benefits and opposers' costs. We expect these positive effects since if surrounding jurisdictions are tightly regulated, this should increase the impact of a relaxation. As for the average marginal effects of a city referendum and a referendum involving the most liberal policy in the county, they are 0.49 and 0.24 for the supporters' benefit and 0.66 and 0.40 for the opposers' costs, respectively.

The coefficients in Table 2 indicate that while the cost of voting does depend on whether the election is held on the weekend, the weather conditions on the day of the election and summer-time elections do not significantly effect the cost of voting.[24] The positive coefficient of 0.282 associated with the weekend indicator variable implies an average marginal effect of 0.65 on the upper support of voting costs. We felt that voting on a weekend might be more costly because individuals would have to give up their leisure time in order to vote. The marginal effect is significant given that the average upper support of voting costs across the 363 jurisdictions is 2.45 (see Table 3).

The average values of ν_j, w_j, b_j and x_j presented in Table 3 indicate that opposers feel much more intensely about the issue than do supporters. This greater intensity translates into opposers having significantly higher reasonable cost levels on average than supporters. The average reasonable cost level is 0.76 for supporters and 0.96 for opposers. These yield average turnout rates of 31% for supporters and 42% for opposers.

The predictions of the model's exogenous variables and reasonable cost levels imply a value of μ_j for each district. This can be combined with b_j and x_j to provide a measure of the average net benefit of the proposed change $\mu_j b_j - (1 - \mu_j)x_j$. The change passes a standard cost-benefit test if and only if this average net benefit is positive. Of the 363 jurisdictions, 110 had a positive net benefit. While all of these 110 referenda did pass, 49 of the referenda with a negative net benefit also passed. This is a nice demonstration of the familiar idea that majority voting does not produce surplus maximizing outcomes.

A proposed change with a positive net benefit does not imply that holding

[24] Note that there is no constant term in the cost of voting function. By not including a constant term we are setting the upper support of voting costs equal to one for elections held on a non-summer weekday where there is no rain nor snow and the temperature is zero degrees Fahrenheit. The normalization is required because we cannot infer benefits and costs from the number of people who vote for and against the referendum. Instead, we can only infer relative benefits and relative costs.

a referendum is desirable because of the transactions costs associated with voting. Holding the referendum passes a cost-benefit test if and only if $\mu_j(b_j - (\gamma^*_{sj})^2/2c_j) - (1-\mu_j)(x_j + (\gamma^*_{oj})^2/2c_j)$ is positive. Only 39 referenda had a positive net benefit when voting costs are included.[25] This suggests that the case for this form of direct democracy is weak when evaluated on conventional cost-benefit grounds.

The following section presents two alternative models by which to test the appropriateness of the current model. Because the comparison of the three models will be primarily based on how well the models fit the data, it is interesting to note that the model just estimated explains 54 percent of the variation in turnout for the referenda and 49 percent of the variation in turnout against the referenda.

7 Alternative Models

The previous section showed how to estimate the parameters of our turnout model assuming that it was the correct model. While the fact that the parameter estimates seem reasonable is comforting, the results give us no reason to believe that our basic model is a good approximation of voting behavior. To provide evidence on this, we compare it with two simple alternatives. They are ad hoc, in the sense that they do not provide an account of why people behave in the postulated way, but they do capture basic ideas that have been expressed in the turnout literature.

7.1 The intensity hypothesis

The *intensity hypothesis* asserts that people are more likely to vote the more intensely they feel about an issue. This is consistent with an *expressive* view of voting (see, for example, Brennan and Lomasky (1993)). Voting is like cheering at a football game and you are more likely to cheer the more you care about the outcome. Formally, we assume that supporters vote if their voting cost is less than

$$\gamma_s = \alpha b$$

and opposers vote if their voting cost is less than

$$\gamma_o = \alpha x,$$

where $\alpha > 0$. Here, the parameter α measures the strength of citizens' desire to express themselves through voting which may depend upon community characteristics. The key restriction is that both supporters and opponents share the same α. Under this specification, the probability that a supporter votes is the

[25] The Texas state government's recent move to require that liquor referenda be held on uniform election dates should help in this respect by spreading the transactions costs over a number of ballot issues. However, since it will also impact the likely turnout pattern, it may also increase the set of referenda with negative net benefits that pass.

probability that γ_s exceeds his voting cost, which is $\gamma_s/c = \alpha b/c$. Similarly, the probability that an opposer votes is $\gamma_o/c = \alpha x/c$.

To estimate the model, we assume as in the basic model that for each jurisdiction j, $\nu_j = 1 + \exp(\beta_v \cdot z_{vj})$, $\omega_j = 1 + \exp(\beta_\omega)$, $x_j = \exp(\beta_x \cdot z_{xj} + \varepsilon_j)$, $b_j = \exp(\beta_b \cdot z_{bj} + \varepsilon_j)$ and $c_j = \exp(\beta_c \cdot z_{cj})$. We further assume that $\alpha_j = \exp(\beta_\alpha \cdot z_{\alpha j})$ where β_α is a vector of parameters to be estimated and $z_{\alpha j}$ is a vector of jurisdiction specific characteristics that may affect citizens' desires to express themselves. The parameters to be estimated are $\Omega = \{\ \beta_v, \beta_\omega, \beta_b,\ \beta_x, \beta_\alpha, \beta_c\}$.

To construct the likelihood function, fix Ω and consider a particular jurisdiction j. Then, according to the intensity hypothesis,

$$v_{sj} = \mu_j \exp \varepsilon_j \frac{\alpha_j \widehat{b}_j}{c_j},$$

and

$$v_{oj} = (1-\mu_j) \exp \varepsilon_j \frac{\alpha_j \widehat{x}_j}{c_j},$$

where $\widehat{b}_j = \exp(\beta_b \cdot z_{bj})$ and $\widehat{x}_j = \exp(\beta_x \cdot z_{xj})$. Letting $Z_j = (z_{vj}, z_{bj}, z_{xj}, z_{\alpha j}, z_{cj})$, the probability density function for (v_{sj}, v_{oj}) is then[26]

$$g(v_{sj}, v_{oj}; \Omega, Z_j) = \frac{(\widehat{x}_j)^{\nu_j}(v_{sj})^{\nu_j - 1}(\widehat{b}_j)^{\omega_j}(v_{oj})^{\omega_j - 1}}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}),$$

where

$$\varsigma_j = \ln c_j + \ln(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj}) - \ln \alpha_j - \ln \widehat{x}_j - \ln \widehat{b}_j.$$

As in the basic model, we must ensure that the parameter estimates are such that γ_{sj} and γ_{oj} are less than the upper support of the distribution of voting costs. Therefore, we maximize the likelihood function subject to the constraints that $\gamma_{sj} \leq c_j$ and $\gamma_{oj} \leq c_j$ for each jurisdiction j. These constraints impose the following restrictions on the parameter estimates: for all jurisdictions j,

$$\ln(\frac{v_{sj}}{1 - v_{oj}}) \leq \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj} \leq \ln(\frac{1 - v_{sj}}{v_{oj}}).$$

Table 4 shows the parameter values that maximize the likelihood function subject to the above constraints and Table 5 contains average implied values of the exogenous variables $(\nu_j, \omega_j, \alpha_j, b_j, x_j, c_j)$ based on the parameter estimates. We use the same variables to explain variation in ν, b, x, and c as in the basic model. The coefficients in Table 4 suggest that the effects of these variables are similar to those in the basic model. As indicated by the implied values in Table 5, the average expected percentage of supporters predicted by the intensity model is almost identical to that in the basic model (54% compared to 55%). While it is impossible to directly compare the implied values of b, x and c across

[26] Again, the Appendix provides the derivation.

models, note that while in both models the implied value of the opposers' cost is greater than the supporters' benefit, on average, this difference is much larger in the basic model.

The strength of citizens' desire to express themselves through voting is likely to vary across jurisdictions depending on the jurisdiction's religious composition, age distribution, and size. The coefficients in Table 4 associated with the fraction of county population that is baptist, the fraction of county voting age population over 50, and voting age population suggest that individuals who are baptist, over the age of 50, and reside in smaller jurisdictions have a stronger desire to express themselves through voting.

7.2 The popularity hypothesis

The *popularity hypothesis* asserts that people are more willing to vote if they expect that many of their fellow citizens share their position on the issue. The idea is that the returns from voting are in the form of social approval (or avoidance of disapproval) from those who share one's position (as discussed, for example, by Coleman (1990)). Accordingly, a supporter failing to vote for the referendum when most other citizens are supporters, will experience more disapproval than if most citizens are opponents.

Recalling that the mean value of μ is $\frac{\nu}{\nu+\omega}$, we can capture this idea by assuming that supporters vote if their voting cost is below

$$\gamma_s = \alpha \frac{\nu}{\nu + \omega},$$

while opposers vote if their voting cost is below

$$\gamma_o = \alpha \frac{\omega}{\nu + \omega},$$

where $\alpha > 0$. This says that individuals from each group are more likely to vote the larger is the expected proportion of their group in the population. The parameter α measures the value of obtaining approval in the election. This will depend on community characteristics (such as size) and also on the salience of the election to citizens. Again, the key restriction is that it is the same for both supporters and opponents. Under this specification, the probability that a supporter votes is the probability that γ_s exceeds his voting cost, which is $\gamma_s/c = \alpha\nu/c(\nu + \omega)$. Similarly, the probability that an opposer votes is $\gamma_o/c = \alpha\omega/c(\nu + \omega)$.

To estimate the model, we assume as in the basic model that for each jurisdiction j, $\nu_j = 1+\exp(\beta_v \cdot z_{vj})$, $\omega_j = 1+\exp(\beta_\omega)$ and $c_j = \exp(\beta_c \cdot z_{cj})$. We also assume that $\alpha_j = \exp(\beta_\alpha \cdot z_{\alpha j} + \varepsilon_j)$ where β_α is a vector of parameters to be estimated, $z_{\alpha j}$ is a vector of jurisdiction and referendum specific characteristics that may affect the value of obtaining approval and ε_j is the realization of some random variable distributed according to the standard normal distribution. The vector of parameters to be estimated is therefore $\Omega = \{ \beta_\nu, \beta_\omega, \beta_\alpha, \beta_c\}$.

To construct the likelihood function, fix Ω and consider a particular jurisdiction j. Then, according to the popularity hypothesis,

$$v_{sj} = \mu_j \exp \varepsilon_j \frac{\widehat{\alpha}_j}{c_j} \cdot \frac{\nu_j}{\nu_j + \omega_j},$$

and

$$v_{oj} = (1 - \mu_j) \exp \varepsilon_j \frac{\widehat{\alpha}_j}{c_j} \cdot \frac{\omega_j}{\nu_j + \omega_j},$$

where $\widehat{\alpha}_j = \exp(\beta_\alpha \cdot z_{\alpha j})$. Letting $Z_j = (z_{\alpha j}, z_{vj}, , z_{cj})$, the probability density function for (v_{sj}, v_{oj}) is then

$$g(v_{sj}, v_{oj}; \Omega, Z_j) = \frac{(\omega_j)^{\nu_j}(v_{sj})^{\nu_j - 1}(\nu_j)^{\omega_j}(v_{oj})^{\omega_j - 1}}{(\nu_j v_{oj} + \omega_j v_{sj})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}),$$

where

$$\varsigma_j = \ln(v_j v_{oj} + \omega_j v_{sj}) + \ln(v_j + \omega_j) - \ln \widehat{\alpha}_j + \ln c_j - \ln v_j - \ln \omega_j.$$

The constraints that $\gamma_{sj} \leq c_j$ and $\gamma_{oj} \leq c_j$ for each jurisdiction j impose the following restrictions on the parameter estimates: for all j

$$\frac{v_{sj}}{1 - v_{oj}} \leq \frac{1 + \exp(\beta_v \cdot z_{vj})}{1 + \exp(\beta_\omega)} \leq \frac{1 - v_{sj}}{v_{oj}}.$$

Tables 6 and 7 show the optimal parameter values and the implied average values of the exogenous variables $(\nu_j, \omega_j, \alpha_j, c_j)$.[27] We use the same variables to explain variation in ν and c as in the basic model. As indicated by the implied values in Table 7, the average expected percentage of supporters predicted by the popularity model is 49% which is less than the 55% predicted by the basic model. Besides that associated with the city indicator variable, the coefficients in Table 7 suggest that the effects of the variables on the fraction of supporters are similar to those in the basic model. The positive and statistically significant coefficient of 0.085 in the popularity model indicates that the fraction of supporters increases by two percentage points if the election involves a city. As for the cost of voting, the only coefficient estimate in Table 6 that is appreciably different than the analogous estimate in Table 2 is the one associated with the weekend indicator variable. While both models predict that having an election on a weekend increases the cost of voting, the predicted effect is larger for the basic model. In addition, the popularity model predicts a much lower average cost of voting than the basic model.

The social approval supporters and opposers obtain from voting is likely to depend not only on the referendum but also the characteristics of the jurisdiction. We expect the social approval from voting to be greater the more important the referendum and the more individuals interact with others who have similar views regarding the liquor referendum. The coefficients in Table 6

[27] The parameter estimates in Table 6 are such that none of the constraints bind.

indicate that social approval increases when the referendum involves off-premise consumption of alcohol, when the referendum involves a city, and when the surrounding jurisdictions have more restrictive alcohol policies. While these effects are as expected, the negative coefficient associated with the off- and on-premise consumption is surprising. Interestly, the direction and magnitude of these effects on social approval are similar to those on supporters' benefit and opposers' costs in the basic model. As for the size of the jurisdiction, the negative and statistically significant coefficient of -0.024 associated with voting age population suggests that increasing a jurisdiction's voting age population by 5,000 increases the social approval from voting by, on average, 11.5 percent.

7.3 Comparing the models

To test the validity of the basic model relative to the intensity and popularity models, we use the directional test for non-nested models proposed by Vuong (1989). Vuong proposes a likelihood-ratio based statistic to test the null hypothesis that two competing models are equally close to the true data generating process against the alternative hypothesis that one model is closer. Vuong proves that the difference between the maximum log-likelihood values of Model A and Model B divided by the product of the standard deviation of the difference in the log likelihood value for each observation and the square root of the number of observations has a standard normal distribution if the two models are equivalent. Vuong also demonstrates that the null hypothesis that Models A and B are equivalent can be rejected when the alternative hypothesis is that Model A (B) is better than Model B (A) if the above test statistic is greater (less) than critical value c (-c) obtained from the standard normal distribution for some significance level.

The maximum log-likelihood values for the basic, intensity, and popularity models are 749.20, 704.05, and 692.26 respectively. Table 8 presents the value of Vuong's test statistic for the three possible null hypotheses. These values indicate the null hypothesis that the basic model is equivalent to the popularity model and that the basic model is equivalent to the intensity model can be rejected at the five percent significance level when the alternative hypothesis is that the basic model is better. However, these null hypotheses cannot be rejected when the alternative hypotheses are that the popularity and intensity models are better, respectively. Table 8 also indicates that the null hypothesis that the popularity and intensity models are equivalent can be rejected at the five percent significance level when the alternative hypothesis is that the intensity model is better. Thus, the basic model appears the best, with the intensity model the runner up.

8 Conclusion

This paper has made use of a unique data set to structurally estimate a model of voter turnout and to statistically compare it with two simple alternatives. The results are encouraging: the structural estimation yields reasonable coefficient estimates and the model performs better than the alternatives. This suggests that the approach to thinking about turnout that underlies the model warrants serious consideration.

There are many different directions for future research on this approach (see also Feddersen and Sandroni (2001)). From an empirical perspective, it would be worth comparing the performance of the model presented here with that of the pivotal-voter model. While there are good reasons to be sceptical about its abilities to explain turnout, the pivotal voter model represents in many respects the simplest way of thinking about voting behavior. Thus, it should only be rejected if it can be shown to be outperformed by some coherent alternative. This has yet to be demonstrated. The data set used in this paper is appropriate for studying the pivotal-voter model and the model we have developed is a coherent alternative. It remains to structurally estimate the pivotal-voter model and compare its performance. Given its great complexity, this will be a challenging task.

From a theoretical perspective, it would be interesting to think about the implications of heterogeneity in supporters' and opposers' preferences. It seems likely that, within groups, those voters who care less intensely about an issue will have lower reasonable cost levels. This may reflect considerations of equity in the allocation of the costs of voting. Another interesting topic is how to think about elections with three or more candidates. While such elections naturally divide the population into groups of supporters, it is no longer obvious how supporters of an underdog candidate should vote. This is particularly the case when there are differences among group members in their second choice candidate. Finally, more thought should be given to the justification of the behavior postulated here. Why should citizens believe that this is how they should behave in elections? Moreover, even if they do believe it, what forces compel them to do their part?

References

[1] Aldrich, John, [1993], "Rational Choice and Turnout," *American Journal of Political Science*, 37, 246-78.

[2] Aldrich, John, [1997], "When is it Rational to Vote," in Mueller, Dennis (ed), *Perspectives on Public Choice,* Cambridge: Cambridge University Press.

[3] Baughman, Reagan; Conlin, Michael; Dickert-Conlin, Stacy and John Pepper, [2000], "Slippery when Wet: The Effects of Local Alcohol Access Laws on Highway Safety," *Journal of Health Economics*, forthcoming.

[4] Brennan, Geoffrey and Loren Lomasky, [1993], *Democracy and Decision: The Pure Theory of Electoral Preference*, Cambridge: Cambridge University Press.

[5] Coleman, James S., [1990], *Foundations of Social Theory*, Cambridge MA: Harvard University Press.

[6] Downs, Anthony, [1957], *An Economic Theory of Democracy*, New York: Harper and Row.

[7] Feddersen, Timothy J. and Alvaro Sandroni, [2001], "A Theory of Participation in Elections," mimeo, University of Rochester.

[8] Filer, John; Kenny, Lawrence and Rebecca Morton, [1993], "Redistribution, Income and Voting," *American Journal of Political Science*, 37, 63-87.

[9] Fiorina, Morris, [1997], "Voting Behavior," in Mueller, Dennis (ed), *Perspectives on Public Choice,* Cambridge: Cambridge University Press.

[10] Green, Donald and Ian Shapiro, [1994], *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*, New Haven: Yale University Press.

[11] Grossman, Gene and Elhanan Helpman, [2001], *Special Interest Politics*, Cambridge MA: MIT Press.

[12] Hansen, Stephen; Palfrey, Thomas and Howard Rosenthal, [1987], "The Downsian Model of Electoral Participation: Formal Theory and Empirical Analysis of the Constituency Size Effect," *Public Choice*, 52, 15-33.

[13] Harsanyi, John C., [1980], "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior," *Theory and Decision*, 12, 115-33.

[14] Ledyard, John O., [1984], "The Pure Theory of Large Two-Candidate Elections," *Public Choice*, 44, 7-41.

[15] Matsusaka, John and Filip Palda, [1993], "The Downsian Voter Meets the Ecological Fallacy," *Public Choice*, 77, 855-78.

[16] Matsusaka, John and Filip Palda, [1999], "Voter Turnout: How Much Can we Explain?" *Public Choice*, 77, 855-78.

[17] Morton, Rebecca B., [1987], "A Group Majority Voting Model of Public Good Provision," *Social Choice and Welfare*, 4, 117-31.

[18] Morton, Rebecca B., [1991], "Groups in Rational Turnout Models," *American Journal of Political Science*, 35, 758-76.

[19] Palfrey, Thomas and Howard Rosenthal, [1985], "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79, 62-78.

[20] Riker, William H. and Peter Ordeshook, [1968], "A Theory of the Calculus of Voting," *American Political Science Review*, 62, 25-42.

[21] Shachar, Ron and Barry Nalebuff, [1999], "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 89, 525-47.

[22] Uhlaner, Carole J., [1989], "Rational Turnout: The Neglected Role of Groups," *American Journal of Political Science*, 33, 390-422.

[23] Vuong, Quang H., [1989], "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses," *Econometrica*, 57, 307-333.

9 Appendix

9.1 Information on the data

A total of 526 local liquor elections are identified in the annual reports of the Texas Alcoholic Beverage Commission between 1976 and 1996. We use 363 of these in our estimation. Of the 163 elections we do not use, 64 were missing critical information[28] and 43 involved elections where other items seem likely to have been voted on at the same time.[29] To keep the basic issue constant across elections, we focus on proposals to move from a completely "dry" status where the selling of any alcohol is prohibited at the retail level.[30] Therefore, we eliminate the 53 elections where the jurisdiction was not "dry" prior to the election. Finally, in order to structurally estimate our model, we drop three elections where zero votes were cast against the referendum.[31]

The United States Census Bureau provides annual county-level populations, by age. This information allows us to determine the population and the fraction of the population over 50 at the time of the election when the jurisdiction voting is either an entire county or a justice precinct. The population of a justice precinct at the time of an election is estimated by dividing the county population by the number of justice precincts in the county. We expect this to be a relatively good approximation based on information provided by the Texas Legislative Council indicating that justice precincts are selected so that each in a particular county has roughly the same number of residents. The fraction of the justice precinct over the age of 50 is assumed to be the same as in the county.

In addition to the county-level information, the Census provides the total population of many cities and towns in 1970, 1980 and 1990. For cities and towns, the voting age population is estimated at the time of an election by linearly interpolating and extrapolating the information provided by the Census Bureau. Consider the city of Novice in Coleman county which had an election on January 6, 1987. Novice had a voting age population of 129 in 1980 and 140

[28] Specifically, 12 observations did not identify the precise nature of the changes proposed by the referendum, 15 elections occurred in cities not identified in the United States Census and 37 occurred in justice precincts where the precise number of justice precincts in the county could not be identified with confidence.

[29] We sent letters to the clerks of the 180 counties which had liquor elections over the period requesting information on whether other issues were being voted on at the same time. Almost half sent copies of the notes from the Commissioners Court's meeting or a copy of the official document containing the results of the election. Both of these identified all items that were voted on at the same time as the local liquor referendum. Most of the other county clerks sent letters indicating whether the liquor law referendum was the only item on the ballot. A few county clerks either did not respond or could not determine all items on the ballot. Of the 43 elections we suspect might have been held with other issues, 24 were ones for which we could not get a response from the relevant county clerk and which were held on uniform election days. The remaining 19, were ones that we knew for certain were held with other issues. Of these, 16 were held on uniform election dates.

[30] A jurisdiction can prohibit the retail sale of all alcohol while still allowing private clubs (including the VFW, American Legion and other fraternal organizations) to serve alcohol.

[31] All of these elections had at least one vote for the referendum.

in 1990. By linearly interpolating this information, we estimate Novice's voting age population to be 136.7 at the time of the election. If the election occurred in 1993, we would estimate the population by using Novice's voting age population in 1990 and assuming that this population grew at the same rate as the county's voting age population between 1990 and 1993. Because Coleman's voting age population grew -1.56 percent from 1990 to 1993, we would estimate the voting age population in Novice in 1993 to be 137.8. A similar extrapolation is used for elections prior to 1980 in cities and towns whose populations were not reported by the 1970 Census (but were in 1980 and 1990). As with the justice precincts, the fraction of a city's or town's population over the age of 50 is assumed to be the same as in the county.

Churches & Church Membership in the United States provides county-level information on the number of adherents to Baptist denominations. It is published every ten years (in 1970, 1980, and 1990). The total number of Baptists in a county at the time of an election is estimated by linearly interpolating and extrapolating this information. By dividing this number by the county population, we obtain an estimate of the fraction of the county population that is Baptist. We use this fraction as a proxy for the fraction of Baptists in each of the 363 jurisdictions; thereby, implicitly assuming that Baptists are uniformly distributed throughout each county.

The United States Carbon Dioxide Information Analysis Center (CDIAC) collects daily observations of maximum temperature, minimum temperature, precipitation and snowfall from 1,062 weather stations (of which 44 are located in Texas) comprising the United States Historical Climatology Network. We calculate the midpoint of the maximum and minimum temperature at each weather station on the day of an election and use this measure of temperature in our specification.

9.2 Proofs

Proof of Proposition 1: Let (γ_s^*, γ_o^*) be an interior equilibrium. Then, (γ_s^*, γ_o^*) must satisfy the pair of first order conditions:

$$\frac{\partial \pi(\gamma_s^*, \gamma_o^*)}{\partial \gamma_s} b = \frac{\gamma_s^*}{c}$$

and

$$-\frac{\partial \pi(\gamma_s^*, \gamma_o^*)}{\partial \gamma_o} x = \frac{\gamma_o^*}{c}.$$

We know that

$$\frac{\partial \pi(\gamma_s, \gamma_o)}{\partial \gamma_s} = h(\frac{\gamma_o}{\gamma_o + \gamma_s}; \nu, \omega) \frac{\gamma_o}{(\gamma_o + \gamma_s)^2}$$

and

$$-\frac{\partial \pi(\gamma_s, \gamma_o)}{\partial \gamma_o} = h(\frac{\gamma_o}{\gamma_o + \gamma_s}; \nu, \omega) \frac{\gamma_s}{(\gamma_o + \gamma_s)^2}.$$

It follows that the two first order conditions imply that

$$\gamma_s^* = \frac{\sqrt{b}}{\sqrt{x}}\gamma_o^*.$$

Substituting this into the first of the two first order conditions, we find that

$$(\gamma_o^*)^2 = h(\frac{\sqrt{x}}{\sqrt{b}+\sqrt{x}};\nu,\omega)\frac{c(\sqrt{b})(\sqrt{x})^3}{(\sqrt{b}+\sqrt{x})^2}$$

which implies that

$$(\gamma_s^*)^2 = h(\frac{\sqrt{x}}{\sqrt{b}+\sqrt{x}};\nu,\omega)\frac{c(\sqrt{b})^3(\sqrt{x})}{(\sqrt{b}+\sqrt{x})^2}.$$

Thus,

$$\gamma_o^* = [h(\frac{\sqrt{x}}{\sqrt{b}+\sqrt{x}};\nu,\omega)\frac{c(\sqrt{b})(\sqrt{x})^3}{(\sqrt{b}+\sqrt{x})^2}]^{\frac{1}{2}},$$

and

$$\gamma_s^* = [h(\frac{\sqrt{x}}{\sqrt{b}+\sqrt{x}};\nu,\omega)\frac{c(\sqrt{b})^3(\sqrt{x})}{(\sqrt{b}+\sqrt{x})^2}]^{\frac{1}{2}}.$$

For the Beta distribution, we have that

$$h(\frac{\sqrt{x}}{\sqrt{b}+\sqrt{x}};\nu,\omega) = \frac{\sqrt{x}^{\nu-1}\sqrt{b}^{\omega-1}}{(\sqrt{x}+\sqrt{b})^{\nu+\omega-2}B(\nu,\omega)},$$

and substituting this into the above formulas yields the characterization stated in the proposition. QED

Proof of Proposition 2: We need to show that γ_s^* maximizes the supporters' payoff $\pi(\gamma_s,\gamma_o^*;\nu,\omega)b - \frac{\gamma_s^2}{2c}$ subject to the constraint that $\gamma_s \in [0,c]$ and that γ_o^* maximizes the opposers' payoff $-\pi(\gamma_s^*,\gamma_o;\nu,\omega)x - \frac{\gamma_o^2}{2c}$ subject to the constraint that $\gamma_o \in [0,c]$. We prove only the former claim, since the argument for the latter is analogous.

If γ_s^* did not maximize the supporters' payoff, there must exist some $\widehat{\gamma}_s$ that would yield a higher payoff. By condition (iii) of the Proposition, we know that $\widehat{\gamma}_s \neq 0$. Define the function $\varphi : [0,c] \to \Re$ as follows:

$$\varphi(\gamma_s) = \pi(\gamma_s,\gamma_o^*;\nu,\omega)b - \frac{\gamma_s^2}{2c}.$$

Note first the following important claim.

Claim: Suppose that $\varphi'(\widetilde{\gamma}_s) = 0$ for some $\widetilde{\gamma}_s \in (0,c]$. Then, $\varphi''(\widetilde{\gamma}_s)$ has the opposite sign from $(\nu+2)\widetilde{\gamma}_s - (\omega-2)\gamma_o^*$.

Proof: We have that

$$\varphi'(\gamma_s) = \frac{\partial \pi(\gamma_s,\gamma_o^*)}{\partial \gamma_s}b - \frac{\gamma_s}{c},$$

and that

$$\varphi''(\gamma_s) = \frac{\partial^2 \pi(\gamma_s, \gamma_o^*)}{\partial \gamma_s^2} b - \frac{1}{c}.$$

Observe that

$$\frac{\partial^2 \pi(\gamma_s, \gamma_o)}{\partial \gamma_s^2} = -\frac{\gamma_o}{(\gamma_o + \gamma_s)^3}[h_\mu \frac{\gamma_o}{\gamma_o + \gamma_s} + 2h],$$

so that

$$\varphi''(\gamma_s) = -\frac{\gamma_o^*}{(\gamma_s + \gamma_o^*)^3}[h_\mu \frac{\gamma_o^*}{\gamma_s + \gamma_o^*} + 2h]b - \frac{1}{c}.$$

For the Beta distribution, for all (γ_s, γ_o) we have that

$$h = \frac{\gamma_o^{\nu-1}\gamma_s^{\omega-1}}{(\gamma_o + \gamma_s)^{\nu+\omega-2}B(\nu, \omega)},$$

and

$$h_\mu = \frac{\gamma_o^{\nu-2}\gamma_s^{\omega-2}}{(\gamma_o + \gamma_s)^{\nu+\omega-3}B(\nu, \omega)}[(\nu - 1)\gamma_s - (\omega - 1)\gamma_o],$$

so we may write

$$h_\mu = \frac{[(\nu - 1)\gamma_s - (\omega - 1)\gamma_o](\gamma_o + \gamma_s)h}{\gamma_o \gamma_s}.$$

Moreover, the fact that $\varphi'(\widetilde{\gamma}_s) = 0$ implies that

$$\frac{1}{c} = hb \frac{\gamma_o^*}{\widetilde{\gamma}_s(\widetilde{\gamma}_s + \gamma_o^*)^2}.$$

It follows that

$$\varphi''(\widetilde{\gamma}_s) = -\frac{\gamma_o^* hb}{(\widetilde{\gamma}_s + \gamma_o^*)^2}\{\frac{[(\nu - 1)\widetilde{\gamma}_s - (\omega - 1)\gamma_o^*]}{\widetilde{\gamma}_s(\widetilde{\gamma}_s + \gamma_o^*)} + \frac{2}{(\widetilde{\gamma}_s + \gamma_o^*)} + \frac{1}{\widetilde{\gamma}_s}\}.$$

Since $\gamma_o^* > 0$, the sign of $\varphi''(\widetilde{\gamma}_s)$ is the opposite of the sign of

$$\frac{[(\nu - 1)\widetilde{\gamma}_s - (\omega - 1)\gamma_o^*]}{\widetilde{\gamma}_s(\widetilde{\gamma}_s + \gamma_o^*)} + \frac{2}{(\widetilde{\gamma}_s + \gamma_o^*)} + \frac{1}{\widetilde{\gamma}_s}.$$

This is positive if $(\nu + 2)\widetilde{\gamma}_s > (\omega - 2)\gamma_o^*$ and negative if $(\nu + 2)\widetilde{\gamma}_s < (\omega - 2)\gamma_o^*$. The Claim now follows. QED

Suppose first that $\widehat{\gamma}_s > \gamma_s^*$. Consider the problem

$$\min\{\varphi(\gamma_s) : \gamma_s \in [\gamma_s^*, \widehat{\gamma}_s]\}$$

Since φ is continuous and the constraint set is compact, the problem has a solution which we denote by $\widetilde{\gamma}_s$. Note that the solution must lie in the interior of $[\gamma_s^*, \widehat{\gamma}_s]$. To see this note that $\widetilde{\gamma}_s$ must be less than $\widehat{\gamma}_s$ since $\varphi(\widehat{\gamma}_s) > \varphi(\gamma_s^*)$. In addition, we know that by condition (i) $\varphi'(\gamma_s^*) = 0$, and by condition (ii) and the Claim, $\varphi''(\gamma_s^*) < 0$. This means that for γ_s slightly larger than γ_s^* that $\varphi(\gamma_s) < \varphi(\gamma_s^*)$. Since φ is smooth, it follows that $\varphi'(\widetilde{\gamma}_s) = 0$ and $\varphi''(\widetilde{\gamma}_s) \geq 0$. By the Claim, we have that $\varphi''(\widetilde{\gamma}_s) \geq 0$ if and only if $(\nu+2)\widetilde{\gamma}_s \leq (\omega-2)\gamma_o^*$. But we know from condition (ii) and the fact that $\widehat{\gamma}_s > \gamma_s^*$ that $(\nu+2)\widetilde{\gamma}_s > (\nu+2)\gamma_s^* > (\omega-2)\gamma_o^*$ so this is impossible. Thus, $\widehat{\gamma}_s$ cannot be greater than γ_s^*.

Now suppose that $\widehat{\gamma}_s < \gamma_s^*$. Without loss of generality, we may assume that $\widehat{\gamma}_s$ solves the problem:

$$\max\{\varphi(\gamma_s) : \gamma_s \in [0, \gamma_s^*]\}.$$

Since $\widehat{\gamma}_s \in (0, \gamma_s^*)$, we know that $\varphi'(\widehat{\gamma}_s) = 0$ and $\varphi''(\widehat{\gamma}_s) \leq 0$. By the Claim, we know that $(\nu+2)\widehat{\gamma}_s \geq (\omega-2)\gamma_o^*$. Now consider the problem

$$\min\{\varphi(\gamma_s) : \gamma_s \in [\widehat{\gamma}_s, \gamma_s^*]\}$$

The problem has a solution which we denote by $\widetilde{\gamma}_s$. Note that the solution must lie in the interior of $[\widehat{\gamma}_s, \gamma_s^*]$. To see this note that $\widetilde{\gamma}_s$ must be greater than $\widehat{\gamma}_s$ since $\varphi(\widehat{\gamma}_s) > \varphi(\gamma_s^*)$. In addition, we know that by conditions (i) and (ii) and the Claim, $\varphi'(\gamma_s^*) = 0$ and $\varphi''(\gamma_s^*) < 0$. This implies that for γ_s slightly smaller than γ_s^* that $\varphi(\gamma_s) < \varphi(\gamma_s^*)$. Since φ is smooth, it follows that $\varphi'(\widetilde{\gamma}_s) = 0$ and $\varphi''(\widetilde{\gamma}_s) \geq 0$. By the Claim, we have that $\varphi''(\widetilde{\gamma}_s) \geq 0$ if and only if $(\nu+2)\widetilde{\gamma}_s \leq (\omega-2)\gamma_o^*$. But we know that $(\nu+2)\widetilde{\gamma}_s > (\nu+2)\widehat{\gamma}_s > (\omega-2)\gamma_o^*$ so this is impossible. Thus, $\widehat{\gamma}_s$ cannot be smaller than γ_s^*. QED

9.3 Deriving the probability density function for (v_{sj}, v_{oj})

9.3.1 The basic model

Define the functions $u_{sj} : [0,1] \times \Re \to \Re$ and $u_{oj} : [0,1] \times \Re \to \Re$ as follows:

$$u_{sj}(\mu, \varepsilon) = \mu K_j \sqrt{\exp \varepsilon}$$

and

$$u_{oj}(\mu, \varepsilon) = (1-\mu) K_j \frac{\sqrt{\widehat{x}_j}}{\sqrt{\widehat{b}_j}} \sqrt{\exp \varepsilon}.$$

Then, as shown in the text, we have that

$$v_{sj} = u_{sj}(\mu_j, \varepsilon_j),$$

and

$$v_{oj} = u_{oj}(\mu_j, \varepsilon_j).$$

In addition, we know that μ_j and ε_j are continuous random variables having joint probability density function

$$h_j(\mu_j, \varepsilon_j) = \frac{\mu_j^{\nu_j - 1}(1-\mu_j)^{\omega_j - 1}}{B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_j^2}{2}).$$

We know that $v_{sj} = u_{sj}(\mu_j, \varepsilon_j)$ and $v_{oj} = u_{oj}(\mu_j, \varepsilon_j)$ define a one-to-one mapping from (μ_j, ε_j) space to (v_{sj}, v_{oj}) space. Thus, the joint probability density function of (v_{sj}, v_{oj}) is given by:

$$g_j(v_{sj}, v_{oj}) = h_j(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))\,|\Psi|$$

where $(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ are defined implicitly by the relations $v_{sj} = u_{sj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and $v_{oj} = u_{oj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and

$$\begin{aligned} |\Psi| &= \begin{vmatrix} \frac{\partial \mu_j}{\partial v_{sj}} & \frac{\partial \mu_j}{\partial v_{oj}} \\ \frac{\partial \varepsilon_j}{\partial v_{sj}} & \frac{\partial \varepsilon_j}{\partial v_{oj}} \end{vmatrix} \\ &= \frac{\partial \mu_j}{\partial v_{sj}} \frac{\partial \varepsilon_j}{\partial v_{oj}} - \frac{\partial \mu_j}{\partial v_{oj}} \frac{\partial \varepsilon_j}{\partial v_{sj}}. \end{aligned}$$

As shown in the text,

$$\mu_j(v_{sj}, v_{oj}) = \frac{v_{sj}\sqrt{\widehat{x}_j}}{v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}},$$

and

$$\varepsilon_j(v_{sj}, v_{oj}) = 2\ln(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}) - 2\ln(K_j\sqrt{\widehat{x}_j})$$

Note that

$$\frac{\partial \mu_j}{\partial v_{sj}} = \frac{v_{oj}\sqrt{\widehat{b}_j}\sqrt{\widehat{x}_j}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^2},$$

and

$$\frac{\partial \mu_j}{\partial v_{oj}} = \frac{-v_{sj}\sqrt{\widehat{b}_j}\sqrt{\widehat{x}_j}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^2}.$$

Moreover,

$$\frac{\partial \varepsilon_j}{\partial v_{oj}} = \frac{2\sqrt{\widehat{b}_j}}{v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}}$$

and

$$\frac{\partial \varepsilon_j}{\partial v_{sj}} = \frac{2\sqrt{\widehat{x}_j}}{v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}}$$

Thus, we have:

$$|\Psi| = \frac{2\sqrt{\widehat{b}_j}\sqrt{\widehat{x}_j}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^2}.$$

This implies that

$$\begin{aligned} g_j(v_{sj}, v_{oj}) &= \frac{\mu_j(v_{sj}, v_{oj})^{\nu_j - 1}(1 - \mu_j(v_{sj}, v_{oj}))^{\omega_j - 1}}{B(\nu_j, \omega_j)} \\ &\quad \times \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_j(v_{sj}, v_{oj})^2}{2}) \cdot \frac{2\sqrt{\widehat{b}_j}\sqrt{\widehat{x}_j}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^2} \\ &= \frac{2(\sqrt{\widehat{x}_j})^{\nu_j}(v_{sj})^{\nu_j - 1}(\sqrt{\widehat{b}_j})^{\omega_j}(v_{oj})^{\omega_j - 1}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \\ &\quad \times \frac{1}{\sqrt{2\pi}} \exp(-\frac{(\ln[\frac{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})}{K_j\sqrt{\widehat{x}_j}}]^2)^2}{2}). \end{aligned}$$

Next observe that, after substituting in for K_j, we may write

$$[\frac{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})}{K_j\sqrt{\widehat{x}_j}}]^2 = \frac{(\sqrt{\widehat{b}_j} + \sqrt{\widehat{x}_j})^{\nu_j + \omega_j} B(\nu_j, \omega_j) c_j (v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^2}{(\sqrt{\widehat{x}_j})^{\nu_j + 2}(\sqrt{\widehat{b}_j})^{\omega_j + 2}},$$

so that

$$g_j(v_{sj}, v_{oj}) = \frac{2(\sqrt{\widehat{x}_j})^{\nu_j}(v_{sj})^{\nu_j - 1}(\sqrt{\widehat{b}_j})^{\omega_j}(v_{oj})^{\omega_j - 1}}{(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}),$$

where

$$\begin{aligned} \varsigma_j &= (\nu_j + \omega_j) \ln(\sqrt{\widehat{b}_j} + \sqrt{\widehat{x}_j}) + \ln B(\nu_j, \omega_j) + \ln c_j \\ &\quad + 2\ln(v_{oj}\sqrt{\widehat{b}_j} + v_{sj}\sqrt{\widehat{x}_j}) - (\nu_j + 2)\ln(\sqrt{\widehat{x}_j}) - (\omega_j + 2)\ln(\sqrt{\widehat{b}_j}). \end{aligned}$$

9.3.2 The intensity hypothesis

Define the functions $u_{sj} : [0,1] \times \Re \to \Re$ and $u_{oj} : [0,1] \times \Re \to \Re$ as follows:

$$u_{sj}(\mu, \varepsilon) = \mu T_j \exp \varepsilon$$

and

$$u_{oj}(\mu, \varepsilon) = (1 - \mu) T_j \frac{\widehat{x}_j}{\widehat{b}_j} \exp \varepsilon.$$

where $T_j = \frac{\alpha_j \widehat{b}_j}{c_j}$. Then, as shown in the text, we have that

$$v_{sj} = u_{sj}(\mu_j, \varepsilon_j),$$

and

$$v_{oj} = u_{oj}(\mu_j, \varepsilon_j).$$

In addition, we know that μ_j and ε_j are continuous random variables having joint probability density function

$$h_j(\mu_j, \varepsilon_j) = \frac{\mu^{\nu_j - 1}(1-\mu)^{\omega_j - 1}}{B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_j^2}{2}).$$

By the same logic as for the basic model, the joint probability density function of (v_{sj}, v_{oj}) is given by:

$$g_j(v_{sj}, v_{oj}) = h_j(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))\,|\Psi|$$

where $(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ are defined implicitly by the relations $v_{sj} = u_{sj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and $v_{oj} = u_{oj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and

$$|\Psi| = \frac{\partial \mu_j}{\partial v_{sj}} \frac{\partial \varepsilon_j}{\partial v_{oj}} - \frac{\partial \mu_j}{\partial v_{oj}} \frac{\partial \varepsilon_j}{\partial v_{sj}}.$$

Solving for $\mu_j(v_{sj}, v_{oj})$ and $\varepsilon_j(v_{sj}, v_{oj})$, we obtain:

$$\mu_j(v_{sj}, v_{oj}) = \frac{\widehat{x}_j v_{sj}}{\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj}},$$

and

$$\varepsilon_j(v_{sj}, v_{oj}) = \ln(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj}) - \ln T_j \widehat{x}_j.$$

Note that

$$\frac{\partial \mu_j}{\partial v_{sj}} = \frac{\widehat{b}_j \widehat{x}_j v_{oj}}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})^2}$$

and

$$\frac{\partial \mu_j}{\partial v_{oj}} = \frac{-\widehat{b}_j \widehat{x}_j v_{sj}}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})^2}.$$

Moreover,

$$\frac{\partial \varepsilon_j}{\partial v_{oj}} = \frac{\widehat{b}_j}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})}$$

and

$$\frac{\partial \varepsilon_j}{\partial v_{sj}} = \frac{\widehat{x}_j}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})}$$

Thus, we have:

$$|\Psi| = \frac{\widehat{b}_j \widehat{x}_j}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})^2}.$$

This implies that

$$g_j(v_{sj}, v_{oj}) = \frac{(\widehat{x}_j)^{\nu_j}(v_{sj})^{\nu_j - 1}(\widehat{b}_j)^{\omega_j}(v_{oj})^{\omega_j - 1}}{(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}),$$

where

$$\varsigma_j = \ln c_j + \ln(\widehat{b}_j v_{oj} + \widehat{x}_j v_{sj}) - \ln \alpha_j - \ln \widehat{x}_j - \ln \widehat{b}_j.$$

9.3.3 The popularity hypothesis

Define the functions $u_{sj} : [0,1] \times \Re \to \Re$ and $u_{oj} : [0,1] \times \Re \to \Re$ as follows:

$$u_{sj}(\mu, \varepsilon) = \mu S_j \exp \varepsilon$$

and

$$u_{oj}(\mu, \varepsilon) = (1-\mu) S_j \frac{\omega_j}{\nu_j} \exp \varepsilon.$$

where $S_j = \frac{\widehat{\alpha}_j}{c_j} \cdot \frac{\nu_j}{\nu_j + \omega_j}$. Then, as shown in the text, we have that

$$v_{sj} = u_{sj}(\mu_j, \varepsilon_j),$$

and

$$v_{oj} = u_{oj}(\mu_j, \varepsilon_j).$$

In addition, we know that μ_j and ε_j are continuous random variables having joint probability density function

$$h_j(\mu_j, \varepsilon_j) = \frac{\mu^{\nu_j - 1}(1-\mu)^{\omega_j - 1}}{B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varepsilon_j^2}{2}).$$

By the same logic as for the basic model, the joint probability density function of (v_{sj}, v_{oj}) is given by:

$$g_j(v_{sj}, v_{oj}) = h_j(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))\, |\Psi|$$

where $(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ are defined implicitly by the relations $v_{sj} = u_{sj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and $v_{oj} = u_{oj}(\mu_j(v_{sj}, v_{oj}), \varepsilon_j(v_{sj}, v_{oj}))$ and

$$|\Psi| = \frac{\partial \mu_j}{\partial v_{sj}} \frac{\partial \varepsilon_j}{\partial v_{oj}} - \frac{\partial \mu_j}{\partial v_{oj}} \frac{\partial \varepsilon_j}{\partial v_{sj}}.$$

Solving for $\mu_j(v_{sj}, v_{oj})$ and $\varepsilon_j(v_{sj}, v_{oj})$, we obtain:

$$\mu_j(v_{sj}, v_{oj}) = \frac{\omega_j v_{sj}}{\nu_j v_{oj} + \omega_j v_{sj}}$$

and

$$\varepsilon_j(v_{sj}, v_{oj}) = \ln(\nu_j v_{oj} + \omega_j v_{sj}) - \ln \omega_j S_j.$$

Note that

$$\frac{\partial \mu_j}{\partial v_{sj}} = \frac{\nu_j \omega_j v_{oj}}{(\nu_j v_{oj} + \omega_j v_{sj})^2}$$

and

$$\frac{\partial \mu_j}{\partial v_{oj}} = \frac{-\nu_j \omega_j v_{sj}}{(\nu_j v_{oj} + \omega_j v_{sj})^2}.$$

Moreover,

$$\frac{\partial \varepsilon_j}{\partial v_{oj}} = \frac{\nu_j}{(\nu_j v_{oj} + \omega_j v_{sj})}$$

and

$$\frac{\partial \varepsilon_j}{\partial v_{sj}} = \frac{\omega_j}{(\nu_j v_{oj} + \omega_j v_{sj})}.$$

Thus, we have:

$$|\Psi| = \frac{\nu_j \omega_j}{(\nu_j v_{oj} + \omega_j v_{sj})^2}.$$

This implies that

$$g_j(v_{sj}, v_{oj}) = \frac{(\omega_j)^{\nu_j}(v_{sj})^{\nu_j - 1}(\nu_j)^{\omega_j}(v_{oj})^{\omega_j - 1}}{(\nu_j v_{oj} + \omega_j v_{sj})^{\nu_j + \omega_j} B(\nu_j, \omega_j)} \cdot \frac{1}{\sqrt{2\pi}} \exp(-\frac{\varsigma_j^2}{2}).$$

where

$$\varsigma_j = \ln(\nu_j v_{oj} + \omega_j v_{sj}) + \ln c_j + \ln(\nu_j + \omega_j) - \ln \omega_j - \ln \widehat{\alpha}_j - \ln \nu_j.$$

9.4 Solving the constrained maximization

For the basic model, the problem is to maximize (1) subject to the constraint (2) holding for each jurisdiction j. Observe that the constraints described in (2) have a nice linear structure. Moreover, we use the same four district specific variables to explain variation in b and x and they are all dummy variables. Thus, $z_{bj} = z_{xj} = (1, \delta_{1j}, \delta_{2j}, \delta_{3j}, \delta_{4j})$ where $\delta_{ij} \in \{0, 1\}$ for $i = 1, 2, 3, 4$. Since there are only sixteen possible values for the vector $(\delta_{1j}, \delta_{2j}, \delta_{3j}, \delta_{4j})$, we can divide the districts into sixteen different categories according to their $(\delta_{1j}, \delta_{2j}, \delta_{3j}, \delta_{4j})$ vectors. We then need only impose the constraint for the two outliers in terms of $\ln(\frac{v_{sj}}{1-v_{oj}})^2$ and $\ln(\frac{1-v_{sj}}{v_{oj}})^2$ values in each group. This reduces the number of feasibility constraints to thirty two.

We use the following iterative procedure to solve the problem. First, we group the districts into the sixteen categories $t \in \{1, ..., 16\}$ according to the values of their $(\delta_{1j}, \delta_{2j}, \delta_{3j}, \delta_{4j})$ vectors. We then select from each category the district for which $\ln(\frac{v_{sj}}{1-v_{oj}})^2$ is maximized, denoted $j(t)$. Next, for each t, we maximize the likelihood function imposing only the single constraint that $\ln(\frac{v_{sj}}{1-v_{oj}})^2 = \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj}$ for district $j(t)$. We then let $T(1)$ denote the set of all t for which the solution satisfies the constraints that for all t', $\ln(\frac{v_{sj(t')}}{1-v_{oj(t')}})^2 \leq \beta_b \cdot z_{bj(t')} - \beta_x \cdot z_{xj(t')}$ and let $t(1)$ be the $t \in T(1)$ which yields the highest value of the likelihood function.

The second step in the iteration involves considering all possible pairs (t, t') such that $t, t' \notin T(1)$ and for each of these maximize the likelihood function imposing only the constraints that $\ln(\frac{v_{sj}}{1-v_{oj}})^2 = \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj}$ for districts $j(t)$ and $j(t')$. We then let $T(2)$ denote the set of all such (t, t') pairs for which the solution satisfies all the constraints and let $(t, t')(2)$ be the (t, t') pair in $T(2)$ yielding the highest value of the likelihood function.

The third step is to consider all possible triples (t, t', t'') such that $t, t', t'' \notin T(1)$ and $(t, t'), (t, t''), (t', t'') \notin T(2)$ and for each of these maximize the likelihood function imposing only the constraints that $\ln(\frac{v_{sj}}{1-v_{oj}})^2 = \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj}$ for districts $j(t), j(t')$ and $j(t'')$. We then let $T(3)$ denote the set of all such (t, t', t'') triples for which the solution satisfies all the constraints and let $(t, t', t'')(3)$ be the (t, t', t'') triple in $T(3)$ which yields the highest value of the likelihood function.

We continue doing this until the mth step in which it is not possible to find any admissible m-tuples. We then compare the value of the likelihood function associated with $t(1)$, $(t, t')(2)$, $(t, t', t'')(3)$, etc. and pick the constraint combination that results in the highest value. This solves the problem of maximizing the likelihood function subject to the feasibility constraints $\ln(\frac{v_{sj}}{1-v_{oj}})^2 \leq \beta_b \cdot z_{bj} - \beta_x \cdot z_{xj}$ for all districts j. It also turns out to satisfy the constraints that $\beta_b \cdot z_{bj} - \beta_x \cdot z_{xj} \leq \ln(\frac{1-v_{sj}}{v_{oj}})^2$ for all j and hence is the solution to the full problem.

A similar iterative procedure is use for the intensity model. This procedure was not required for the popularity model because the coefficient estimates of the unconstrained maximization problem satisfied all constraints.

TABLE 1: Summary Statistics for the 363 Elections

Fraction of Referenda involving only Beer/Wine	0.40 (147 Elections)
Fraction of Referenda involving Off-Premise Consumption	0.40 (144 Elections)
Fraction of Referenda involving Off- and On-Premise Consumption	0.20 (72 Elections)
Fraction of Referenda involving an Entire County	0.01 (2 Elections)
Fraction of Referenda involving a Justice Precinct	0.37 (133 Elections)
Fraction of Referenda involving an Incorporated Town or City	0.63 (228 Elections)
Voting Age Population in Jurisdiction	Mean = 4145 Standard Deviation=7464
Fraction of County Voting Age Population Over the Age of 50	0.40
Fraction of County Population that is Baptist	0.48
Fraction of Referenda involving more Liberal Policy than County	0.33 (121 Elections)
Number of Alcohol Related Accidents in County Divided by County Population (1,000) in past 12 months	Mean = 2.04 Standard Deviation=0.71
Fraction of Jurisdiction Located in an MSA	0.44 (158 Elections)
Average Temperature on Day of Election (Fahrenheit)	Mean = 65.2 Standard Deviation=16.0
Rainfall on Day of Election (tenths of inches)	Mean = 0.96 Standard Deviation=3.10
Snowfall on Day of Election (tenths of inches)	Mean = 0.06 Standard Deviation=1.05
Fraction of Referenda that occurred on Saturday or Sunday	0.70 (255 Elections)
Fraction of Referenda that occurred in Summer	0.27 (97 jurisdictions)
Fraction of Referenda that Pass	0.41 (150 Elections)
Fraction of Voting Age Population Voting For Referendum	Mean = 0.17 Standard Deviation=0.12
Fraction of Voting Age Population Voting Against Referendum	Mean = 0.19 Standard Deviation=0.13
Turnout (# of votes/ Voting Age Population)	Mean = 0.36 Standard Deviation=0.22
Closeness (Difference between Votes For and Against Divided by Total Votes Cast)	Mean = 0.25 Standard Deviation=0.19

Figure 1: Turnout Histogram

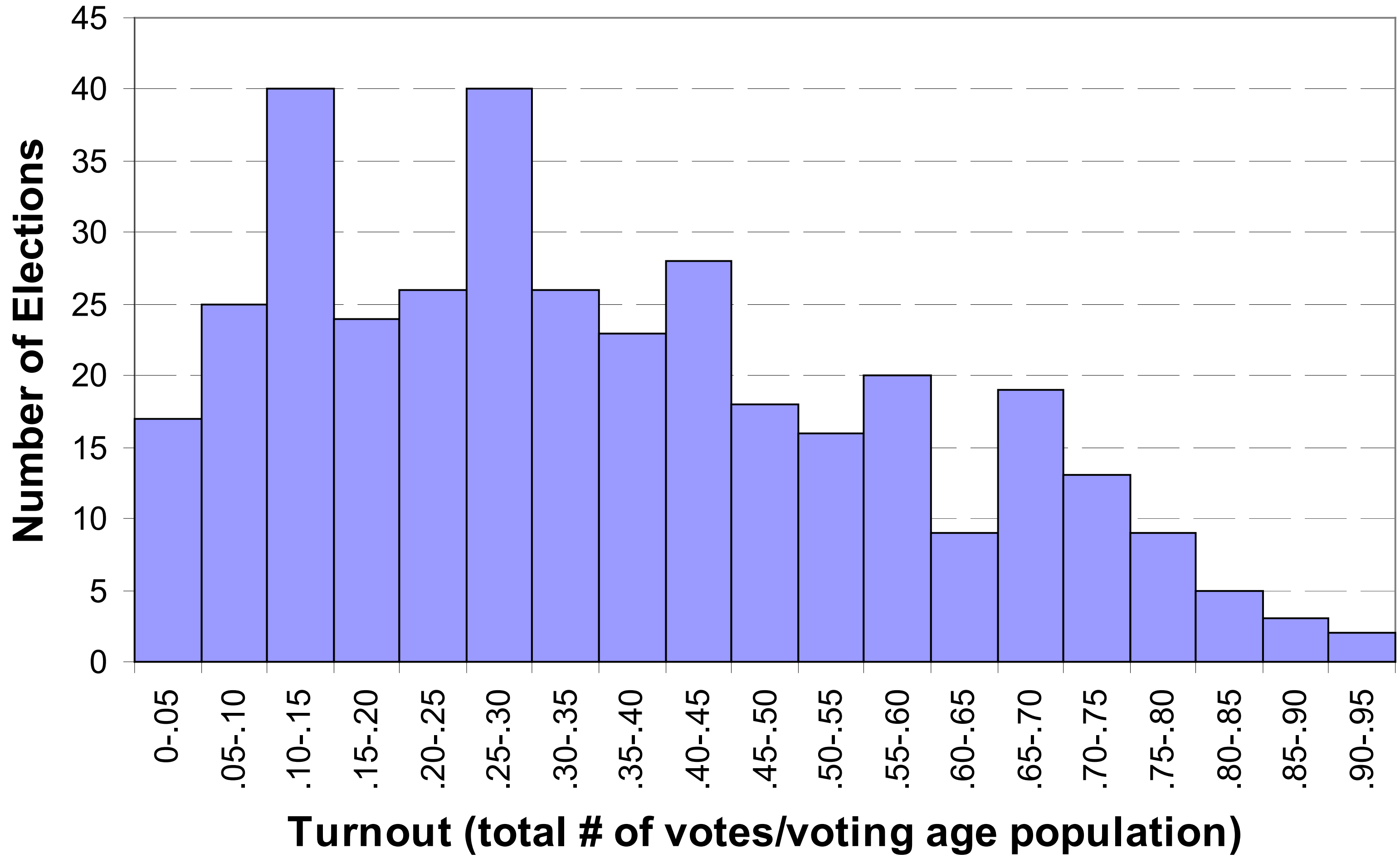

Figure 2: Closeness Histogram

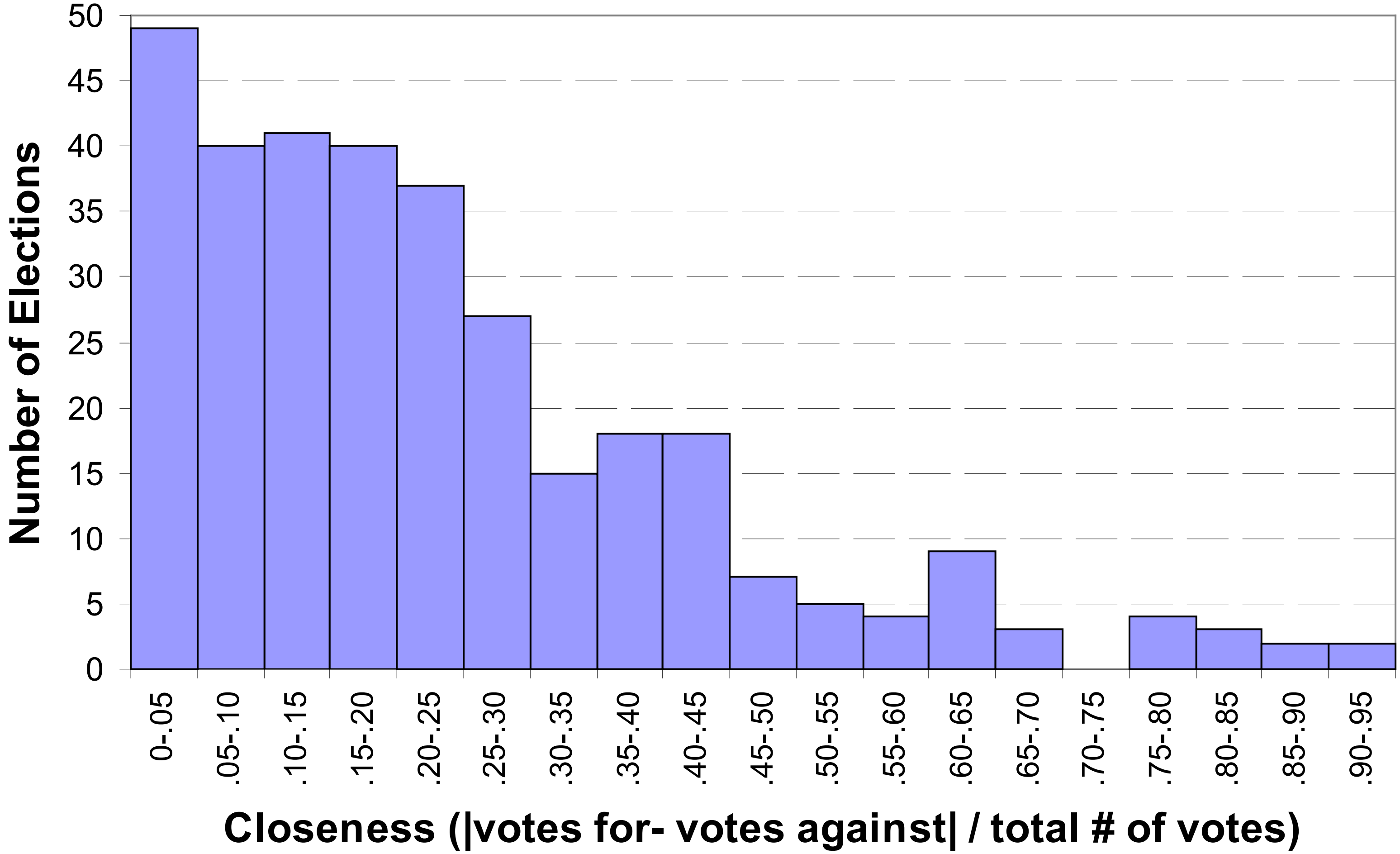

TABLE 2: BASIC MODEL
N=363 , Log Likelihood=749.20

VARIABLES	Coefficients
v:	
Fraction of County Population that is Baptist	-0.531 (0.477)
Fraction of County Voting Age Population Over the Age of 50	-0.428 (0.572)
Alcohol Related Accidents in County in Prior Year	0.044 (0.062)
Indicator Variable for Town or City Referendum	-0.083 (0.084)
Indicator Variable for Jurisdiction being Located in an MSA	-0.413** (0.118)
Constant	1.948** (0.369)
w:	
Constant	1.144** (0.099)
b:	
Indicator Variable for Off-Premise Consumption of Alcohol	0.379** (0.134)
Indicator Variable for Off- and On-Premise Consumption of Alcohol	-0.610** (0.173)
Indicator Variable for Town or City Referendum	1.888** (0.112)
Indicator Variable for Most Liberal Policy in County	0.568** (0.115)
Constant	-2.758** (0.70)
x:	
Indicator Variable for Off-Premise Consumption of Alcohol	0.142 (0.158)
Indicator Variable for Off- and On-Premise Consumption of Alcohol	-1.045** (0.184)
Indicator Variable for Town or City Referendum	1.497** (0.113)
Indicator Variable for Most Liberal Policy in County	0.614** (0.115)
Constant	-1.848** (0.709)
c:	
Indicator Variable for Election on Weekend	0.282** (0.117)
Rainfall on Day of Election (tenths of inches)	0.002 (0.002)
Snowfall on Day of Election (tenths of inches)	0.067 (0.052)
Average Temperature on Day of Election (Fahrenheit)	0.009 (0.024)
Average Temperature Squared	0.00002 (0.0002)
Indicator Variable for Election in Summer	-0.103 (0.161)

Standard errors are in parentheses. * Statistically significant at the .10 level; ** Statistically significant at the .05 level.

TABLE 3: ESTIMATED VALUES FROM BASIC MODEL

Model Parameters	Mean Estimates
v	5.07 (0.91)
w	4.14 (0)
x	0.92 (0.97)
b	0.59 (0.65)
c	2.45 (0.56)

Standard errors are in parentheses.

TABLE 4: INTENSITY MODEL
N=363 , Log Likelihood=704.05

VARIABLES	Coefficients
v:	
Fraction of County Population that is Baptist	-0.519
	(0.401)
Fraction of County Voting Age Population Over the Age of 50	-0.497
	(0.577)
Alcohol Related Accidents in County in Prior Year	0.071
	(0.057)
Indicator Variable for Town or City Referendum	-0.080
	(0.087)
Indicator Variable for Jurisdiction being Located in an MSA	-0.333**
	(0.116)
Constant	1.850**
	(0.290)
w:	
Constant	1.165**
	(0.103)
α:	
Fraction of County Population that is Baptist	0.971
	(0.496)
Fraction of County Voting Age Population Over the Age of 50	1.022
	(0.621)
Voting Age Population (1,000)	-0.021**
	(0.008)
b:	
Indicator Variable for Off-Premise Consumption of Alcohol	0.063
	(0.131)
Indicator Variable for Off- and On-Premise Consumption of Alcohol	-0.154
	(0.156)
Indicator Variable for Town or City Referendum	0.896**
	(0.115)
Indicator Variable for Most Liberal Policy in County	0.165
	(0.123)
Constant	-2.440**
	(0.739)
x:	
Indicator Variable for Off-Premise Consumption of Alcohol	-0.036
	(0.133)
Indicator Variable for Off- and On-Premise Consumption of Alcohol	-0.342*
	(0.159)
Indicator Variable for Town or City Referendum	0.700**
	(0.115)
Indicator Variable for Most Liberal Policy in County	0.237*
	(0.126)
Constant	-2.053**
	(0.740)
c:	
Indicator Variable for Election on Weekend	0.133
	(0.117)
Rainfall on Day of Election (tenths of inches)	0.001
	(0.002)
Snowfall on Day of Election (tenths of inches)	0.017
	(0.052)
Average Temperature on Day of Election (Fahrenheit)	0.004
	(0.023)
Average Temperature Squared	0.0000
	(0.0002)
Indicator Variable for Election in Summer	-0.069
	(0.162)

Standard errors are in parentheses. * Statistically significant at the .10 level; ** Statistically significant at the .05 level.

TABLE 5: ESTIMATED VALUES FROM INTENSITY MODEL

Model Parameters	Mean Estimates
v	4.93 (0.78)
w	4.21 (0)
α	2.28 (0.55)
x	0.24 (0.15)
b	0.20 (0.13)
c	1.42 (0.13)

Standard errors are in parentheses.

TABLE 6: POPULARITY MODEL
N=363 , Log Likelihood=692.26

VARIABLES	Coefficients
v:	
Fraction of County Population that is Baptist	-0.374* (0.214)
Fraction of County Voting Age Population Over the Age of 50	-0.334 (0.326)
Alcohol Related Accidents in County in Prior Year	0.053* (0.032)
Indicator Variable for Town or City Referendum	0.085* (0.049)
Indicator Variable for Jurisdiction being Located in an MSA	-0.190** (0.062)
Constant	1.437** (0.173)
w:	
Constant	1.240** (0.091)
α:	
Indicator Variable for Off-Premise Consumption of Alcohol	0.130 (0.120)
Indicator Variable for Off- and On-Premise Consumption of Alcohol	-0.336** (0.146)
Indicator Variable for Town or City Referendum	0.818** (0.110)
Indicator Variable for Most Liberal Policy in County	0.270** (0.115)
Voting Age Population (1,000)	-0.024** (0.007)
Constant	-0.678 (0.698)
c:	
Indicator Variable for Election on Weekend	0.135 (0.117)
Rainfall on Day of Election (tenths of inches)	0.001 (0.002)
Snowfall on Day of Election (tenths of inches)	0.022 (0.052)
Average Temperature on Day of Election (Fahrenheit)	0.005 (0.024)
Average Temperature Squared	0.0000 (0.0002)
Indicator Variable for Election in Summer	-0.060 (0.161)

Standard errors are in parentheses. * Statistically significant at the .10 level; ** Statistically significant at the .05 level.

TABLE 7: ESTIMATED VALUES FROM POPULARITY MODEL

Model Parameters	Mean Estimates
v	4.34 (0.36)
w	4.46 (0)
α	1.06 (0.65)
c	1.50 (0.15)

Standard errors are in parentheses.

TABLE 8: VUONG'S NON-NESTED TEST STATISTIC

		Vuong's Test Statistic
Null Hypothesis	Basic Model = Intensity Model	2.18
	Basic Model = Popularity Model	2.81
	Intensity Model = Popularity Model	2.08