What Led to the Ban on Same-Sex Marriage in California?: Structural Estimation of Voting Data on Proposition 8

Vardges Levonyan*

April 2015

Abstract

The voting literature has largely analyzed voter turnout and voter behavior separately, with a focus on individual election outcomes. This is in spite of the fact that multiple elections are on one ballot, and turnout is determined by participation in all elections. I present a model of voter turnout and behavior in multiple elections. The assumptions are consistent with individual election preferences and decision is derived from utility maximization. I also provide necessary moment conditions for identification. The framework is applied to analyze turnout and voting choices in the 2008 California elections for the US presidential election and Proposition 8 ballot initiative. The exit polls and initial election results made national headlines by linking the historic turnout of African-Americans for Presidential candidate Obama in helping pass Proposition 8. I structurally estimate the demographic preferences for each of the election choices. I find that the African-American turnout and voting share for Proposition 8 was lower than predicted by the exit polls. As a counterfactual, I use the estimated model to look at the turnout and outcome of Proposition 8, without the presidential race on the ballot. As predicted by the model and estimates, I find lower voter turnout that are on par with midterm elections. I also find a lower share of Yes votes on Proposition 8 – enough that the referendum would not have passed.

*vlevonyan@ethz.ch. Discussions with and helpful comments by Alberto Alesina, Stephen Ansolabehere, Greg Crawford, Daniel Chen, Ed Glaeser, Claudia Goldin, Horacio Larreguy, Greg Lewis, Ariel Pakes, and Martin Schonger, and seminar participants at Harvard University, ETH, University of Zurich, and ISNIE, EARIE and ALEA conferences are gratefully acknowledged. All errors are my own.

1 Introduction

It is the norm in the US, and most places in the world, that multiple elections take place at the same time. A single general election ballot typically includes national, state, and local races, along with referendums and initiatives. The vast literature on voting outcomes has largely taken each election individually in a vacuum, ignoring the presence and possible effect of other elections happening simultaneously. The literature on multiple elections has focused on explaining political phenomena between different elections such as split-ticket voting, either as an equilibrium of strategic considerations (Alesina and Rosenthal 1996) or due to information variation (Degan and Merlo 2011) between races.

Even if voter preferences have no strategic components and voters are perfectly informed of each candidate, multiple simultaneous elections might still affect each outcome simply due to turnout. A voter considers her utility for participating in all elections, and would go to vote only if the utility of participating is larger than the utility of staying home.[1] Thus, even if voting choices between candidates in a given election are due to individual preferences for each candidate, the overall turnout in the election will depend on all the races taking place at that time.

It is well-documented that turnout varies greatly between elections (Blais 2000). The most evident example is that US presidential elections have had historically higher voter turnouts than midterm elections. Voters respond more strongly to presidential races than to Congressional or local elections. (This could be due to various factors, such as the prestige level of the office or simply more campaign advertising). For heterogeneous voters, adding or excluding certain races like a presidential ticket will affect the composition of voters who go to the polls. This will, in turn, impact shares in other races. It may even impact election outcome.

The prior literature has predominantly looked at turnout separately from voting behavior. And the ecological inference on heterogeneous voters has looked at individual election outcomes. I present a model of turnout and voting decisions in multiple elections. When analyzing each election outcome separately, the presence of additional races will not impact other election results. The framework I propose considers the voter's utility of participation as a direct function of choices in all elections and derives turnout endogenously from utility maximization.

I borrow from discrete choice demand estimation theory in a single market (Berry et al. 1995) and apply the setup to voting decisions in a single election. I then extend the model

[1] Or net-utility, if explicitly accounting for costs.

to allow for selection in multiple elections and derive utility of turnout that is consistent with the model's assumptions on voting preferences. To my knowledge, this framework is the first to extend discrete choice literature to multiple elections, and the setup can also be used to analyze simultaneous consumption decisions in multiple markets. I also provide identification condition for estimation, as the usual moment conditions rely on precinct-level variation of candidate characteristics. I propose new moment conditions when the same candidate choices are on each ballot across precincts, as is typically the case during elections. The new moment conditions are once again consistent with the assumptions of the model.

I apply the setup to estimate turnout and outcomes in the 2008 California general elections. The election results in California made national headlines when exit polls singled out the historic turnout of African-American voters for the first-ever Black presidential candidate, Obama, along with their strong preference for passing Proposition 8, which banned same-sex marriage in the state. While the share of the African-American voting bloc was not sufficient to make up the difference of the close election results of Proposition 8, the story nevertheless brought to the forefront the issue of turnout for one election possibly affecting the outcome of other elections on the same ballot.

I estimate the preferences for presidential candidates Obama and McCain and voting Yes or No on Proposition 8 (Prop 8) for each demographic. When compared to exit poll results, African-Americans had a lower turnout and lower share of Yes votes on Prop 8. (Although they came out in large numbers and the majority voted Yes on Prop 8). As a counterfactual, I look at the turnout and election outcome of Prop 8 without the presidential race on the ballot. The structural setup of the model allows for such analysis, and the counterfactual results are consistent with the observed preferences of voting in both elections. I find lower turnout without the presidential election as predicted by the model and on a par with midterm election results. I also find that Proposition 8 would have most likely been defeated.

The intuition of the model predicts that part of the voting population who chose to vote for president and Prop 8, may abstain from voting when the choice is for Proposition 8 only.[2] The intuition for the outcome of Proposition 8 is similar. The estimates indicate the demographic groups with a strong preference for one of the presidential candidates: Hispanics and Blacks overwhelmingly voted in favor of Obama. Incidentally, they are also the demographic groups with the highest share of Yes votes on Proposition 8. Eliminating the

[2] It is clear that voting in two elections will give at least as much utility as voting in one election, since the voter may simply abstain from the second election. Given that a very small percentage actually abstain from the election, the voter gains positive utility from at least one of the choices in each election. This implies that almost all voters will gain a strictly higher utility from voting in two elections versus one.

presidential election from the ballot reduces turnout, but not uniformly across demographics. Voters with the strongest preference for one of the presidential candidate are less likely to vote in this case. In other words, the drop in turnout for Blacks and Hispanics will be higher than for other demographics. Thus, the share of remaining Yes votes will decrease more than the share of No votes. The drop is enough to overturn the election outcome of Prop 8.

The remainder of the paper is organized as follows. Section 2 provides a review of relevant literature on discrete choice models, voting, and the 2008 election. Section 3 presents the data sources and compares them to polling data. Section 4 presents the model, and derives identification. Section 5 presents the main results, and Section 6 concludes.

2 Literature Review

2.1 Related Literature

The extensive literature on voting also includes the choices of voters over multiple elections (Alesina and Rosenthal 1995, and Alesina and Rosenthal 1996). Prior research largely looks at strategic choices across multiple elections. Alesina and Rosenthal (1996), for instance, have shown that interesting dynamics can arise when the ballot choice set goes from one to two elections. Such political normalities as a divided government and a midterm reversion can, in fact, be borne out as stable equilibrium outcomes. I abstract away from strategic considerations and treat a voting choice as a pure consumption choice. This is done in part because voters' choices have an infinitesimal impact on election outcomes, and partly because it is harder to devise a strategic storyline for election choices consisting of a nationwide presidential campaign and a state referendum. A lack of strategic choices still does not preclude considering the elections simultaneously. Voter action in one election may still affect the results in another due to the participation constraint of going to the voting booth. I treat the voting choice as a consumption choice, which naturally leads to analyzing demand literature.

Recent advances in demand estimation have found applications in many markets, or even fields of economics. Pioneered by the seminal work of Berry et al. (1995) (henceforth, BLP), a structural model of discrete choice random coefficients setup is used to study industries as varied as automobiles (BLP), cereals (Nevo 2001), movies (Einav 2007), and TV broadcast (Goolsbee and Petrin 2004). The models are based on individual optimizing decisions, aggregated to obtain market shares. The model's main advantage is the ability to match aggregate market-level (macro-level) data that has total product sales and characteristics with consumer demographics. One does not need to have individual-level (micro-level) data that matches consumer characteristics with their purchases. The models are general enough to produce reasonable markups and substitution patterns.

This paper extends the BLP framework and develops a model that analyzes consumer participation and consumption in multiple markets. Many purchases take place simultaneously in multiple markets, e.g., an individual buying milk and bread in the same shopping trip, making this is a relevant setup for demand analysis. Estimating bread and milk markets separately might produce incorrect results, especially in terms of participation. The consumer considers her choices in both markets and weights the total utility of bread and milk purchases in the decision to go to the store.

I develop a framework that allows for a joint decision to participate in several markets. The decision will be a function of the utilities of all the products, and I show the assumptions to be consistent with the BLP setup. My methodology is unique for two reasons. First, I derive demand estimates using only aggregate, macro-level, data. Prior literature that looks at purchases in multiple markets uses individual-level data, tying consumers to their bundle of products.[3] To my knowledge, my research is the first to rely on only aggregate data. Complicating the matter of using only aggregate data is that shares are usually reported for individual products, but not for bundles of products, which are the relevant choices for the consumer. In other words, suppose there are two possible choices for bread – $bread1, bread2$, and two possible choices for milk – $milk1, milk2$. If the consumer purchases one of each, her choice set will be from $(bread1, milk1)$, $(bread1, milk2)$, $(bread2, milk1)$, and $(bread2, milk2)$. The aggregate shares are usually reported not for the combination of each of these four bundles, but for $bread1, bread2, milk1$, and $milk2$ separately. I am able to derive the shares of each possible bundle from the product market shares.

Second, I obtain identification despite no product-level variation. The main identification from logit-type models like BLP comes from product space and product characteristic variations across markets, most notably, price. I show that consumer variation across markets, such as demographic differences, could be used instead to carry out estimation. My methodology once again conforms to the assumptions of the BLP setup.

The BLP approach has been used widely to analyze different markets and forms of competition, using the key methodology – heterogeneous preferences and endogenous prices. These include Nevo (2001), Petrin (2002), and Berry et al. (2004). Applications to other forms of competition extend to advertising (Ackerberg 2001, Anand and Shachar 2011), the expansion of satellite broadcast (Goolsbee and Petrin 2004), geographic distribution (Davis

[3] See, e.g., the class of AIDS models.

2006), and the real estate market (Wong 2013), among others.

In the voting literature, ecological inference on election outcomes has a long tradition (King 1997). Analyzing voting data in a discrete choice framework dates back to as early as Poole and Rosenthal (1985). Applying the BLP framework to voter preferences has been used by Rekkas (2007), Gordon and Hartmann (2013), and Martin (2013). In all these cases, campaign spending is used as a substitute for price in the individual utility specification, which provides key moment conditions and identification of the model. In my setup I am able to derive identification despite no campaign spending or any other product characteristic that varies at the market-level.

Hendel (1999) presents a model of multiple markets. However, his model allows for multiple purchases in a given market, e.g., a firm buying multiple PCs or even various brands of PCs. In my setup, the choice is clearly for up to one candidate in each election, and different election races are related primarily by appearing on the same ballot. Perhaps the paper closest in spirit to this one is by Degan and Merlo (2011). The authors develop a structural model of participating in multiple elections, and apply the setup to presidential and Congressional House races on the same ballot. My approach is different from theirs in several key aspects. First, they use individual, micro-level data, whereas my setup can be used with macro-level data. Second, they model the decision to take part in an election as a function of civic duty, which is the same for both elections. Whether to participate in voting, and the difference in voters' preferences for candidates, is due to information (and misinformation) the voter has for each race. By eliminating one of the elections, the outcome in the other election would not change. In my case, the utility from each race depends on observable voter characteristics, such as demographics. The decision to vote is directly determined by all the elections taking place. The utility from turnout is derived as the total utility from choosing a candidate in each election.

2.2 Proposition 8 Background

State amendments and propositions to allow or ban gay marriage were among the newsworthy issues in the 2008 general election. The issue had come to the forefront since the Massachusetts State High Court decision in 2003, and many states subsequently moved to add constitutional amendments through public referendums. In 2008, three additional states – Arizona, California, and Florida – had ballot measures prohibiting same-sex marriage. Probably none received as much media attention as did Proposition 8 in California. As stated on the ballot, voting Yes to Prop 8 would add a provision to the state constitution that "only marriage between a man and a woman is valid or recognized in California".[4] Earlier in 2008, the California State Supreme Court granted permission for same-sex marriages, and by election time, many same-sex marriages had already taken place. Passing Prop 8 would invalidate the State Supreme Court ruling, and the status of already-issued same-sex marriage certificates would be in jeopardy.

Proposition 8 passed with 7,001,084 (52.24%) Yes to 6,401,482 (47.76%) No votes. The passage sparked many protests and demonstrations. Its affect on existing same-sex marriages, and even the validity of the entire proposition, was challenged in court. Through successive appeals to higher courts, the case reached the United States Supreme Court. In June 2013, the US Supreme Court declined to take up the case, effectively handing down the lower court ruling overturning Proposition 8. (The court also ruled that all same-sex marriages be federally recognized.) On a nationwide scale, there is also political discussion to pass legislature for broader gay rights (for instance, President Obama's second inaugural speech).[5]

Perhaps more interesting, in the aftermath of the election, was the analysis of the passage itself. Less than 5% of the total vote separated the Yes and No choices. Any one determining factor could have been the deciding factor between the passage of the proposition and its defeat. The only available data immediately after the election were exit polls. Most produced similar results, and as the CNN polls[6] or *The New York Times* polls[7] indicate, for instance, the demographic breakdown of California voters showed that Whites, Hispanics, Asians, and Other races were almost evenly split in favor of, or against, Proposition 8. Across a gender divide, males and females were also almost evenly split in their sentiment for Prop 8. The only glaring exceptions were African-Americans/Blacks, who were an overwhelming 70% in favor of Prop 8. The runaway story from the exit polls was that the Black population turned out in great numbers to vote for the first-ever Black presidential candidate, Barack Obama, and in the process, helped tip the scales in favor of the passage of Prop 8. This storyline was reported by all the major newspapers, such as *The New York Times*,[8] the *Los Angeles Times*,[9] *The Washington Post*,[10] and other media outlets.[11]

I apply my framework to analyze voting in the 2008 elections in California by looking at the presidential race and the Proposition 8 ballot measure. I look at the relative weight

[4] http://vig.cdn.sos.ca.gov/2008/general/text-proposed-laws/text-of-proposed-laws.pdf#prop8
[5] http://www.whitehouse.gov/the-press-office/2013/01/21/inaugural-address-president-barack-obama
[6] http://www.cnn.com/ELECTION/2008/results/polls/#CAI01p1
[7] http://elections.nytimes.com/2008/results/states/exitpolls/california.html
[8] http://www.nytimes.com/2008/11/06/us/politics/06marriage.html
[9] http://articles.latimes.com/2008/nov/08/local/me-gayblack8
[10] http://www.washingtonpost.com/wp-dyn/content/article/2008/11/06/AR2008110603880.html
[11] E.g., http://www.slate.com/articles/health_and_science/human_nature/2008/11/original_skin.html

of each demographic on the passage of Proposition 8 by using the official election results for the 2008 general election. This method has several advantages over poll results.[12] Primarily, the polling results have weights applied to match the population average. Such weights may produce distorted results when looking at only a particular demographic. Second, the official election data uses the entire population of the voting count, rather than the poll sample, which may not be representative or could suffer from other types of sample bias. Also, the election results are actual choices made, whereas poll responses are self-reported. Third, the overall official turnout, broken down by precincts, provides a more direct comparison to that of previous years. Finally, the election results also include absentee and overseas ballots, which are not fully represented in exit polls but are, nevertheless, becoming relevant portions of overall electorate counts. According to the official vote count by the Secretary of State of California, 41.64% of all votes were cast by mail.[13]

The main drawback of using election results data is that they are precinct-level (macro-level), whereas poll data is individual-level (micro-level). That is, at the poll level, one knows the demographics of an individual, matched with her voting choices. To compute voting preferences across all races then becomes a matter of straightforward computation. In election results, however, one only knows the overall votes for each election, and the breakdown of possible combinations of election choices is not given. To overcome the lack of detailed data, I introduce a structural model of voter turnout and voting in multiple elections. Thus, with my model I can potentially explain the behavior of going to vote on a presidential election, and in the process, also voting on Prop 8. At the heart of my structural model is discrete choice estimation.

3 Data

3.1 Census Data

The two main sources of my data are the US Census,[14] and the voting data from the California Statewide Database.[15] The US Census, conducted decennially, provides a detailed description of the population and businesses at a local geographic level. The smallest level of aggregation depends on the choice of variables and is selected such that individual entries cannot be identified from the aggregate data. Race and gender characteristics for the US population are

[12] Greiner and Quinn (2013) provides a statistical model to combine election results with poll results.

[13] http://www.sos.ca.gov/elections/sov/2008-general/sov_complete.pdf

[14] http://www.census.gov

[15] http://swdb.berkeley.edu

broken down to block level, the finest geographic level possible. Other variables, like income, are provided over a larger geographic area. I am primarily interested in racial characteristics of the population and their impact on elections, and so I collect census data at the block level. Moreover, I restrict the population age to 18 and over, which is the relevant fraction of Americans who can vote.[16] There is no census for the year 2008, so I use the census figures from 2000 and 2010 to extrapolate the population demographics for 2008.

The California Statewide Database is an online redistricting database commissioned by the state itself. It has detailed voting and registration data for all elections since 1992, broken down at the county, district, and precinct levels. Moreover, it provides mapping between census geography units – blocks, and election geography units – precincts. I use the mapping to match race characteristics to voting outcomes across precincts.[17]

Regarding race categories, the 2000 Census differed from the previous ones in that it introduced two or more options as possible choices for the category of race.[18] In addition to mixed race, ethnicity was now primarily divided between Hispanic/Latino and non-Hispanic/Latino. The 2010 Census followed in the same manner, making the distinction between ethnicity and race more explicit. In the 2010 Census, the possible choices for single race are: White, Black or African American, American Indian and Alaska Native, Asian, Native Hawaiian and Other Pacific Islander, and finally, Other. Additionally, it includes all possible combinations of mixed races, starting with any two different races to a maximum of six. The population numbers are reported in two ways. One set of tables is organized along race categories. The other is first broken down by two ethnicities – Hispanic/Latino, and non-Hispanic/Latino – and the non-Hispanic/Latino population is further tabulated by race. Without the ethnicity option and the difference between the first and second methods of reporting, the Hispanic/Latino population identifies themselves primarily as either White, or Other race.[19]

I make use of the second way of reporting census numbers to have a separate category for Hispanics, which is one of the important demographic characteristics in my estimation. The population numbers vary greatly among the six races, as shown in Table 1. I combine the individual races into five major, mutually exclusive categories – Hispanic (henceforth,

[16]The portion of the population who is ineligible to vote may still be an issue. They may not be US citizens, or haven't registered in time for the elections. I explore alternative baseline populations.

[17]The mapping file does not match with perfect accuracy either with the census blocks or with election precincts. This mismatch is the main source of discrepancy when trying to merge the two datasets; whenever such a discrepancy was too large, I looked for alternate places for datasets, where I could obtain information on missing precincts and blocks.

[18]http://www.census.gov/dmd/www/pdf/d02p.pdf

[19]Only about 2% identify as some other race or a mixed race.

Hispanic); White (henceforth, White); Black or African American (henceforth, Black); Asian or Native Hawaiian and Other Pacific Islander (henceforth, Asian); American Indian and Alaska Native or Other (henceforth, Other). [20] This method is dictated by the OMB Directive 15[21] and seems to be the most widely used method of categorization in practice and in the literature (Greiner and Quinn 2013). I also follow the same directive and combine two or more races into a single race in the following fashion.[22] All the people identified as Hispanic ethnicity, are classified as Hispanic. Within the non-Hispanic population, if one of the races with which a person identifies is Black, then that person is classified as Black. Otherwise, if one of the races a person identifies with is Asian, then that person is classified in the Asian category. Otherwise, if it is a mix of White and Other, I classify them as White. Finally, if a person is a mix of Other races, I put that person in the Other category. This is one possible way of combining multiple races into one, and there are certainly other ways of aggregating mixed races. Since the vast majority of the Census, or about 98%, are either Hispanic or single race, the variations in the breakdown of multiple races into single ones would not have a material impact on the results. In summary, this enables me to place the Census population into one of possible five possible categories: White, Black, Asian, Hispanic, and Other.

3.2 Voting Data

Table 1 shows data from the California Statewide database website. The 2008 general election had 8,274,473 votes for the Democratic Presidential candidate (Obama), and 5,011,781 votes for the Republican Presidential candidate (McCain); 7,001,084 Yes votes for Proposition 8, with 6,401,482 No votes. Given the close result for Prop 8, it is not surprising that possible explanations and theories have been topics of intense analysis and scrutiny. These numbers are an exact match with the official count of the California election results as reported by the Secretary of State of California.[23] It is interesting to note that these numbers do not coincide precisely with the initial press release figures printed in *The New York Times*[24]or *USA Today*, for example.[25][26] The press release numbers underreport the outcomes for both candidates

[20]Alternatively, I combine the Native Hawaiian and Other Pacific Islander, American Indian and Alaska Native, and Other race into one, Other race. Another possible alternative is to put American Indian and Alaska Native into the Hispanic ethnicity, and keep Other race as a stand alone category.

[21]http://www.whitehouse.gov/omb/fedreg_directive_15

[22]Remember that two or more races can come from the combination of the following races: White, Black, Asian, and Other.

[23]http://www.sos.ca.gov/elections/sov/2008-general/sov_complete.pdf

[24]http://elections.nytimes.com/2008/results/states/california.html

[25]http://www.usatoday.com/news/politics/election2008/ca.htm

[26]http://content.usatoday.com/news/politics/election2008/

and both choices on Proposition 8. The initially reported results were even closer for Prop 8, making discussions about demographics particularly significant. This is primarily due to the fact that numbers reflecting the initial counts are not updated to include all possible remaining ballots (absentee, questionable, and otherwise) that are added to the final tally. The eventual voting count is finalized and certified many weeks after the election, and it is usually a formality as the results are known by this time. (Unless, of course, the close results trigger a recount). Nevertheless, I work with complete election results, which includes all accepted votes.

3.3 Merging

The block-by-block addition of the 2000 Census figures, reported in Table 1, matches the official Census count[27] of 24,621,819 people in California who are 18 years of age and over. The 2010 Census[28] summation also matches the official figure of 27,958,916 Californians who are 18 years old and over.

I first merge 2000 and 2010 Census datasets to estimate the population size and demographic breakdown for 2008. I do so by assuming a linear trend of population growth for each block-level demographic group. For each 2000 Census block, I match it with the corresponding 2010 block (or blocks, if there has been any redrawing between censuses), and compute the average of each demographic category for the 2008 population.[29] In this way, I can ensure that each block-demographic combination has its own growth rate rather than averaging several growth rates together – for instance, faster growing blocks or demographics with slower growing blocks. In addition, I do not have to worry about computing an unreasonably large population for very fast growing block-demographics since I calculate weighted means of two endpoints as opposed to forecasting into the future.

With my approach, I get a total population of 27,446,193 of California residents for 2008 who are 18 years old and over. Table 1 reports the demographic breakdown of the total

StateDetailResultsByState.aspx?oi=I&rti=G&cn=1&sp=CA

[27]http://www.census.gov/census2000/states/ca.html,
http://www.census.gov/census2000/xls/ca_tab_1.xls

[28]http://www.census.gov/2010census/popmap/ipmtext.php?fl=06

[29]If the mapping from a 2000 Census block does not lead to a unique 2010 Census block, I match it to all the 2010 census blocks that have an area in common with the 2000 block. More specifically, take a 2000 block, with area A. Then take all the 2010 blocks that contain some part of A. If the sum of the areas of all those 2010 blocks is B, and their total population is P, then the relevant 2010 population for the 2000 block will be P*A/B. Thus I assume equal demographic density between neighboring blocks, which provides proportional weight, relative to their areas. With this approach, I ensure correct mapping under various scenarios from 2000 blocks to 2010 blocks, such as 1-to-1, many-to-1, or 1-to-many. The approach then extends naturally to the many-to-many block mappings.

population: 12,538,434 are White, 1,746,243 are Black, 8,824,512 are Hispanic, 3,886,755 are Asian, and the remaining 450,249 are Other race. Since there is no official census count for 2008 as a comparison, I can instead look at the American Community Survey figures,[30] also conducted by the census. There, the population estimate for 2008 is 27,420,473, which is very close to the figure I obtain. Given that their estimate error is $+/-0.1\%$, it is heartening to see that my estimate falls within their margin of error. I do not use American Community Survey results for the 2008 population as those figures are not broken down to block-level observations, even though they have a very detailed racial breakdown of the population.[31]

I then merge the 2008 population figures with 2008 voting data. As with any merger, I do not get a perfectly accurate matching. I also eliminate small precincts – those with total votes or a total population of less than 100. This ensures that the precinct has a large enough sample size to be treated as a market for discrete choice estimation.[32] I also eliminate precincts where the total vote is larger than the overall population. Finally, I remove the precincts with less than a 5% or over a 95% participation rate. These constitute only a handful of precincts, and such outlier rates come from precinct grouping and not necessarily from population voting choices. It also ensures that the outside alternative of not going to vote is well-defined in my estimation.[33] After merging the three datasets – 2000 Census, 2010 Census and 2008 voting data, I finish with 26,532,519 people for whom I have block-level demographic and precinct-level voting information. I also account for 18,807 of the 25,423 precincts that have a total of 12,445,659 votes, from which 7,536,220 voted for a Democratic president, 4,499,673 for a Republican; 6,361,321 voted Yes on Prop 8, while 5,797,454 voted No. Table 5 shows that there is roughly an equal proportional decrease in the number of votes and races as in my matched sample and I get the same aggregate racial composition and voting shares as in the census numbers and the overall voting tally.[34] It is worthwhile to note that I account for the vast majority of the population and the votes, and since I capture the aggregate demographic and voting variation, my results can be extended to the entire population. Plus, given that I have the detailed breakdown of all block-level data, I can map the results for all Californians and estimate voting outcomes not just on the merged data, but for the entire population.

[30] http://www2.census.gov/acs2008_1yr/prod/SelectPopulationProfile/State/California.csv

[31] Including nationalities and ethnicities.

[32] In theory, the market shares are the total population shares, and if a sample size is used, one needs to account for sample size variance as well.

[33] I do not eliminate the tails for the voting choices, as I believe it will omit valuable information about demographic choices.

[34] I am currently investigating the cause of imperfect mapping from blocks to precincts, and will possibly raise the matching percentage between the datasets. However, a higher matching rate will not affect the results.

The summary statistics for the 2008 elections in Table 1 show the variation of demographics and voting outcomes. Figure 1 plots the relationship between the share of each demographic group in the precinct and the percentage of the vote that the Democratic candidate, Obama, received. Figure 2 draws the same picture of voter demographics on the share of Yes votes on Prop 8. A few patterns immediately emerge from the figures. For the White population in both votes, the dispersion grows as the share of the White population increases. The 100% (or nearly 100%) all-White precincts have the most variance on voting choices. This indicates no clear preference for either candidate nor choice for Prop 8 by White voters. Such a dispersion can be accounted for by the unobserved consumer characteristic. The model setup will still enable me to have different preferences for minority demographics, which can differ from each other as evident from the figures.

The graphs for the minority demographics show a completely different picture. Figures 1 and 2 depict a generally favorable preference both for Obama in the presidential election and for a Yes vote on Prop 8. Some other interesting points merit discussion. In the Prop 8 figure, the shares of Yes in largely Black precincts have roughly the same mean as the shares in largely Asian and Hispanic precincts. In the presidential election, there is a clear favorable bias toward Obama by Black, Asian, and Hispanic demographics.[35] The Hispanic and Black populations both approach close to 100% as their precinct shares increase, with the Black population having a smaller variance for moderate to high shares. It is encouraging to have this confirmed in the raw data, as conventional wisdom dictates that Blacks were strongly in favor of Obama.[36] Later in the estimation, however, this causes problems for the discrete choice estimation with no unobserved voter characteristic, as the coefficient for Blacks is undefined (or rather approaches positive infinity).[37] For the Other demographic, there is not enough variation at high levels to obtain reasonable preference estimates. In the estimations, I also find very low participation by Other voters. Therefore, I combine the Other demographic with the White demographics, and treat this as the baseline demographic.

Figures 3 and 4 provide a mechanical intuition of why exit poll results might have provided incorrect results for the subpopulations of different demographics. At the heart of the issue is the question of whether there is reason to believe that the exit polls could provide incorrect predictions for demographic preferences over ballot choices. When the exit poll raw numbers are aggregated, they are weighted to give, on average, correct predictions for the

[35] To get the overall impact of demographics over preferences, the entire distribution is needed.

[36] Though probably not close to 100%.

[37] I have also analyzed the results of the 2004 election and the optimization does not suffer from this problem.

entire population over the elections.[38] As such, by restricting the sample to a subpopulation – such as certain demographics – the outcomes might be strongly weighted in one direction or another, and the results for that subset could be swayed in either direction while being correct for the entire population.[39]

The figures show the plot of the imputed exit poll versus actual voting numbers for each precinct. Figure 3 does this for the Going to Vote share, and Figure 4 for the share of voting Yes on Proposition 8. For each precinct, I calculate the expected number of total voters and Prop 8 Yes votes based on the exit poll ratios. I then subtract the actual number of votes and Prop 8 Yes votes. The data points are then plotted by first being grouped into 20 equal-sized bins, based on the share of the demographic percentage in a precinct. I also fit a linear model. Looking at the Going to Vote share, it is evident that exit polls overstate the number of Blacks and Hispanics who went to the polls and understate the number of Whites and Asians than actual numbers indicate. The exit polls suggest a higher participation in precincts with large Black or Hispanic populations than the actual turnout. Similarly, they show lower turnout in overwhelmingly White or Asian precincts compared to actual election results. As for Prop 8 votes, there is again overstatement of the Black percentage of votes than is borne out in the election results. As the raw data indicates, the exit poll figures seem to overstate both the Black participation and preference for voting Yes on Prop 8; this is what I find in my own estimation, as well. They also seem to suggest understatement of Hispanic voting preferences on Yes for Prop 8, which is also borne out in the results.

4 Model

4.1 The Benchmark Model

The main obstacle to using discrete choice setup for the voting data is that there is one market in the discrete choice model. On the other hand, if the voter decides to participate, she chooses one candidate or ballot choice in all the elections taking place at the time. Further complicating matters is the fact that voting results are reported by each race (or "market") instead of combinations of races. For instance, suppose there are only two elections on the ballot – for president and governor, and only two candidates in each race – Democrat and Republican. The reported election results tell only how many people have voted for a Democrat in the

[38] http://www.ncpp.org/drupal57/files/Weighting.pdf

[39] That is partly why the margin of error for the subpopulation is usually larger than for the entire sample, and it is possible for the estimates not to conform to the actual results, even after accounting for the error.

presidential election, and how many have voted for a Democrat in the gubernatorial election, but not how many have voted all Democrat – (Democrat, Democrat) – in both races.[40]

One approach might be to treat each election race count, along with the outside option numbers, as its own individual market. This, however, does not conform to the individual utility maximization, as presumably the voter makes the choice to go to an election by considering her utility for all the races taking place, and not just from a specific election. Moreover, there is no clear, intuitive way to combine the results to obtain voting counts for say, the (Democrat, Democrat) option that is derived from preferences. I present a nested logit specification of voter demand. The framework allows me to model the decision of simultaneously choosing products from multiple markets. Moreover, my analysis remains consistent with individual utility maximization, and does not require additional assumptions or conditions beyond what is imposed by the discrete choice literature.

For the setup of the model, suppose there are $m_1, m_2, ..., m_L$ markets, each offering $J_1, J_2, ..., J_L$ number of products. Without loss of generalization, I assume that the outside option is common to all markets. That is, if the consumer chooses to participate in one of the markets, then she will participate in all markets. Alternatively, I can introduce an outside option to the product space for each individual market and eliminate the aggregate outside option, as then the aggregate outside option will be the joint union of all individual outside options. Figure 5 presents the graphical tree representation of the consumer decision.

The consumer i chooses the products $c_{j1}, c_{j2}, ..., c_{jL}$ in each of the markets $m_1, m_2, ..., m_L$, respectively. The number of possible combinations for the bundle of products is $J_1 \times J_2 \times ... \times J_L + 1$ and in the ideal setup, I would have the actual market shares for each combination matched with the model's predicted market shares.[41] However, in the voting data (or in a general macro-level demand data), I only have the shares of the outside good and the goods in each of the markets m_t individually, and not the combination. This gives me a total of $J_1 + J_2 + ... + J_L + 1$ observable market shares. Since in general, $J_1 + J_2 + ... + J_L + 1 << J_1 \times J_2 \times ... \times J_L + 1$, I will not have enough identification to back out $J_1 \times J_2 \times ... \times J_L + 1$ market shares without additional assumptions. I show that the discrete choice framework is sufficient to obtain all the shares of the possible combinations.

My discrete choice logit consists of two loops. On the outer loop, the voter decides whether to go to vote, or not. Conditional on going to the election, I then have the inner loop, where the voter i makes the choice for her preferred candidate for each election. For a given

[40]Such cross-tabulation of results is usually found in polls and surveys only.

[41]Even with that ideal setup, there can also be an issue of how to specify the utility as a function of the bundle, if the product characteristics enter non-linearly into the mean utility.

election or market m_t, the inner loop reduces to a standard discrete choice. In particular, let the possible products have mean utilities $\delta_1, ..., \delta_{J_t}$, the individual characteristics of the consumer are $\nu_{i1}, ..., \nu_{ip}$ and the product characteristics of j are $x_{j1}, ..., x_{jq}$. Let k enumerate the relevant consumer and producer characteristic interactions in market m_t from the maximum possible pq pairs. Then, the utility for choosing product j will be:

$$u_{ijt} = \delta_{jt} + \sum_k \alpha_k \nu_{ik} x_{kj} + \varepsilon_{ijt}$$

where ε_{ijt} is i.i.d. Type-I extreme value error term. Note that the mean utility term δ_{jt} includes all product characteristics, including the unobserved product characteristic.

In each market m_t then the consumer chooses the product with the highest utility. Since ε_{ijt} is i.i.d. Type-I, the optimization can be integrated out with a closed form solution for the probability. The probability of choosing good s is:

$$\Pr(jt = s) = \frac{\exp\left(\delta_s + \sum \alpha_k \nu_{ik} x_{ks}\right)}{\sum\limits_{r=1}^{J_t} \exp\left(\delta_r + \sum \alpha_k \nu_{ik} x_{kr}\right)} \tag{1}$$

These probabilities are then computed for all the products in all the markets. For the integration to get market shares, the standard assumption in the discrete choice literature is used, that ε_{ijt} are independent across products and consumers. I can go one step further and assume that ε_{ijt} are also independent across markets. This assumption is rather innocuous as the error terms for products within a market are more likely to be correlated than the error terms for products across markets.[42] Then, since each market product choice is independent from other markets, the probability of selecting a particular bundle $jt_1 = s_1, ..., jt_L = s_L$ will be a product of individual probabilities:

$$\Pr\left(jt_1 = s_1, ..., jt_L = s_L\right) = \frac{\exp\left(\delta_{s_1} + \sum \alpha_k \nu_{ik} x_{ks_1}\right)}{\sum\limits_{r_1} \exp\left(\delta_{r_1} + \sum \alpha_k \nu_{ik} x_{kr_1}\right)} \cdot ... \cdot \frac{\exp\left(\delta_{s_L} + \sum \alpha_z \nu_{iz} x_{zs_L}\right)}{\sum\limits_{r_L} \exp\left(\delta_{r_L} + \sum \alpha_z \nu_{iz} x_{zr_L}\right)} \tag{2}$$

The difference in the notation between k and z allows the possibility that in different markets different consumer-producer characteristic pairs might be relevant. They would also enter the individual utility specification with different magnitudes, specified by α. This prob-

[42] If needed, there is an option to add correlation for the error terms across the products, similar to Berry (1994), and Cardell (1997). Another way to interpret the standard discrete choice assumption is that only the product characteristics, along with the interaction of product and consumer characteristics determine the correlation between product choices.

ability is computed for each individual and each market combination. The aggregate market share is computed by integrating over all the individuals in the market, which is equivalent to integrating over the distributions of all consumer characteristics. This gives the market shares for each of the $J_1 \times J_2 \times ... \times J_L$ product combinations that represent all possible shares conditional on participating in the market.

The final stage is to bring back the decision to participate or abstain from Figure 5, and compute the model predicted probabilities of both decisions. To do this, note that the probability shares for the combination of the goods can be rewritten as:

$$\Pr\left(jt_1 = s_1, ..., jt_L = s_L\right) = \frac{\exp\left(\delta_{s_1} + ... + \delta_{s_L} + \sum \alpha_k \nu_{ik} x_{ks_1} + ... + \sum \alpha_z \nu_{iz} x_{zs_L}\right)}{\sum\limits_{r_1 ... r_L} \exp\left(\delta_{r_1} + ... + \delta_{r_L} + \sum \alpha_k \nu_{ik} x_{kr_1} + ... + \sum \alpha_z \nu_{iz} x_{zr_L}\right)} \quad (3)$$

This expression can be treated as if being a solution to a discrete choice random coefficients model, with mean utilities of $\delta_{s_1} + ... + \delta_{s_L}$, for the product combination of $\{s_1, ..., s_L\}$, and the bundle's product characteristics being the union of individual product characteristics: $\bigcup\{x_{ks_1}, \ldots, x_{ks_L}\}$. The consumer characteristics set is also the union of all consumer characteristics, if not all consumer attributes enter in all product utilities. Looking at the probabilities this way, the solution in equation (3) can be represented as the solution in equation (1) to a standard discrete choice optimization problem. The joint selection of individual products in different markets is therefore equivalent to a single discrete choice decision for the entire bundle, with the mean utility of $\delta_{s_1} + ... + \delta_{s_L}$, and product attribute set of $\bigcup\{x_{ks_1}, \ldots, x_{ks_L}\}$.

To finalize the model, I also need to add the leftmost branch of the decision tree: whether or not to participate in the market. Having established the selection of product bundles as a familiar discrete choice optimization, adding another "product" in the choice set can still be shown as an expression like in equation (1). In the discrete choice setup, the maximum of all error terms, net of mean utilities and interaction terms (i.e. the error term that gives highest utility), will also have Type-I extreme value distribution (Cardell 1997). If $\{s_1, ..., s_L\}$ is the maximizing combination, then the decision between choosing that, or not participating at all entails a choice between two terms with the same Type-I distributive error terms. Once again, the solution to this discrete choice setup will have the familiar expression for the probabilities of the shares. A general combination of product probabilities, including the outside option to the choice set amounts to adding the expression $\exp\left(\delta_0\right)$ to the denominator of equation (3). Alternatively, I scale all the inner choices of participating by δ_g, called the utility of "going", which will again normalize the outside option's mean utility to 0, as is customary in the literature.

The entire decision tree of the consumer can be computed as a two-stage estimation. First, one would estimate the decision to go or to abstain. This will provide the mean utility of going, δ_g. Then, conditional on going, one would estimate the individual market shares: $\{\delta_{s_1}, ..., \delta_{s_L}\}$. The mean utility of the bundle, $\{s_1, ..., s_L\}$, will then be: $\delta_{s_1} + ... + \delta_{s_L} + \delta_g$, with the product attribute set being the union of all the markets' product attributes. Such a setup provides correct within and participation shares. Moreover, one can do counterfactual analysis if the underlying conditions in the decision tree, or in any of the markets, change.

4.2 Moment Conditions

The typical identification for the logit specification is again problematic in my case. At the macro level with only market shares available, variation in product choices and product characteristics (price, in particular) across markets is used to set up the appropriate moment conditions. This is not applicable in my case. In the California elections I analyze, the same choices are on the ballot in each precinct. In the voting literature the candidate/product characteristic is typically constant across markets, unless it includes region-specific campaign spending (Rekkas 2007, Gordon and Hartmann 2013, and Martin 2013). Even with detailed candidate characteristics, in my voting scenario the product dummies will encompass all other individual characteristics. Therefore, I do not even "open up" the mean utility term δ_{jt}, and instead estimate it in its entirety.[43]

I use the variation in aggregate consumer characteristics across markets to get identification. Taken at face value, the mean utility term in the logit specification is composed of only product characteristics and should be orthogonal to individual consumer characteristics. By aggregating the consumer characteristic over the market, I obtain the market demographic distribution for that characteristic. The orthogonality of the distribution term with the mean utility term still holds from the independence of individual characteristics. This provides the basis for my estimation. That is, if $\delta_{jt} - E_j(\delta_{jt})$ is the mean utility of product j in precinct t, normalized to have mean 0 across all precincts, and w_t is the distribution parameter of consumer characteristics, (e.g., the share of whites in precinct t), then:

$$E\left(\left(\delta_{jt} - E_j\left(\delta_{jt}\right)\right) w_t\right) = 0 \tag{4}$$

The orthogonality applies to all products and all consumer demographic components. Since the model assumptions call for complete independence, one may also use higher order

[43] micro-BLP (2004) also takes a similar approach.

terms for distribution components.

The above procedure appears as if the mean utility is treated like an error term. Viewed this way, my moment condition takes the shape of a standard econometric assumption of identification around the error distribution. In fact, such an approach is not far fetched. In traditional logit specification, the mean utility term is split into two components: observed and unobserved product characteristics. The unobserved product characteristic then serves as the error term, around which the moment conditions are built. I treat the entire mean utility, net of mean, as an unobserved product characteristic and form the moment conditions accordingly. One can still add additional information in the δ_{jt} term, such as fixed effects. I do that by including county dummies, which restricts the variation to within-county deviations.

5 Estimation

5.1 Going to Vote

I compare my benchmark model with a parsimonious model that incorporates some of the same qualities as the logit model, and computation is only OLS. The drawback of the parsimonious model is the inability to incorporate unobserved consumer characteristics. Also, it is not micro-founded, and therefore the estimation can be treated as ad hoc. When there is only one racial demographic in the precinct, and no other consumer characteristic including unobserved consumer characteristic, the shares can be expressed in terms of mean-utility as:

$$\ln s_{jt} - \ln(1 - s_{jt}) = \delta_{jt} + \alpha_r$$

where α_r is the utility contribution of race r. For all possible demographies, the more general specification can be written as:

$$\ln s_{jt} - \ln(1 - s_{jt}) = \delta_{jt} + \sum I_r * \alpha_r$$

where I_r is indicator for race r. I can then extend it for the fractions of races. That is:

$$\ln s_{jt} - \ln(1 - s_{jt}) = \delta_{jt} + \sum s_{rt} * I_{rt} * \alpha_{rt} \tag{5}$$

It is important to note that this is not a theoretically correct specification, as the individual shares need to be aggregated first before applying the logit transformation rather than doing aggregation over the logit transformation of the shares.

Written this way, the mean utility δ_{jt} once again becomes the error plus the constant term. Table 2 presents the specification, OLS, along with the Discrete Choice model. The Discrete Choice model estimation is done two ways: one without unobserved consumer characteristic, and the other – the Full Model – with the unobserved characteristic present. In the OLS, the dependent variable is not (log of) share, but the logistic transformation. I have included county fixed effects in all specifications to restrict the mean utility deviations to be among precincts within a given county. All the coefficients are precisely estimated. This is due in part to having a large number of markets – the precincts. The negative coefficient of all the regressed demographics indicates that Whites have the highest probability of voting. Other minority races are less likely to participate. The positive constant coefficient shows that Whites are more likely to vote than not to vote. This is not surprising, since they constitute the majority of the population and roughly 50% of Californians aged 18 and over (75% of the registered voters) voted.[44]

For the remaining demographics, the sum of the constant and the beta coefficient determines their likelihood to vote. The average likelihood of voting doesn't necessarily imply the average ratio of voting share, as the latter would depend on precinct populations and the mean utility terms. In my case, all other demographics are less likely to vote than not to vote. This is perhaps a surprising finding for the Black population, as the media consensus was an historic turnout of Black voters. Even though their turnout is higher than that of Asians, Hispanics, and Others, I find that it is much less than that of White voters. More than half of the Black population living in California, though a lesser percentage of registered Black voters, stayed home during the election. For a robustness check, I also compare the Black turnout in 2008 to the turnout in the previous presidential election of 2004.

The Other population has a large negative coefficient, suggesting that this group is not made up of active voting participants. Their low participation translates to a statistically very small contribution to the subsequent analysis of the presidential election and Prop 8. I therefore combine the Other population with the White demographics. The estimates remain roughly identical from this inclusion. I also tried including the Other population with other demographics (e.g., the Black or Hispanic population) and the results again do not change. I include the Other race with the White population as their total population will form the baseline demographics to which the other demographic preferences will be compared. The analysis is still robust despite the fact that, out of all demographics, the Whites are the most likely to vote, and the Other population is the least likely to vote.

[44] http://www.sos.ca.gov/elections/sov/2008-general/sov_complete.pdf

The Hispanic population also has a large negative coefficient of participation. This translates to Hispanics having a smaller share in total vote participation than the Asians, even though there are three times more Hispanics than Asians in the population count. This is consistent with many Hispanic residents not being eligible to vote, and it is heartening to see the estimation picking up this effect.[45] It is interesting to note that the R-squared in the OLS regression is almost 54%. Just the demographic component explains over half of the variation in participation to vote. Adding the unobserved consumer characteristic will improve the fit of the model, as is also evident from the figures.

When comparing the imputed probabilities of going to vote across demographics to the poll numbers, I find that the polls overestimate the participation of Black and Hispanic voters, and underestimate the participation of White and Asian voters. This is consistent with the intuitive results from Figure 3, and also lessens the magnitude of Black voters' impact on Prop 8. To calculate the overall impact, I also need to calculate the preferences over elections ballots, which I turn to next.

5.2 Presidential Election and Prop 8

I next calculate the shares for the two elections – Presidential race and Prop 8. Tables 3a and 3b report the results. The nested logit specifices the relevant demographic shares as the population of voters, and not the entire population of residents. In the estimation for presidential election with no unobserved consumer characteristic, the preference for Black voters for Democratic candidate, Obama, was problematic. As evident from Figure 1, the share of votes for Obama quickly approaches 100% as the share of Black voters increases. The precinct-specific mean utilities, δ_{jt}, are free-form to pick up any precinct-wide excess preferences. However, even such general specification is not enough to account for the strong preference of Black voters for Obama. In the optimization, the coefficient would approach infinity. This is especially problematic for calculation, as in the discrete choice setup, it would involve taking the exponent of a very large number. To partly offset the computations, I instead compute the share of the Republican candidate, McCain. This still produces large negative numbers for the Black population, but allows for the optimization to converge.[46] The large numbers are still problematic in calculations for standard errors, as the multinomial standard errors are divided by the square of the shares, thus making the variance for the Black coefficient especially large. The Full Model specification does not suffer from such estimation

[45]The large negative coefficient of Asian voters conforms with the anecdotal evidence that many remain permanent residents and do not become citizens.

[46]The computer handles $e^{-\infty}$ better than e^{∞}.

issues, in part because even with strong preference, the presence of another random variable reduces the choice probability to strictly between 0 and 1.

In the presidential election, the preferences for Black, Asian and Hispanic demographics are strongly for Obama. It is interesting to note that, with the correct specification of demographic shares, I show all races being pro-Obama. This is consistent with the CNN poll findings. The R-squared in the OLS is almost 40%, indicating that demographics alone is a strong enough predictor for preference for the presidential candidate.

The Prop 8 estimate predictions on the voting preference match almost exactly with the CNN poll estimates for Asian and White demographics. They are both slightly in favor of voting No on Prop 8. For the Black and Hispanic demographics the results, however, vary from the CNN poll. Similar to the poll results, I find strong preference of both demographic groups to vote Yes on Prop 8. The magnitudes however, are different from the polls. I find that the Black voters are less inclined to vote Yes on Prop 8, and the Hispanics are, in turn, more inclined to vote Yes, than is suggested by the poll results. This finding, combined with the fact that more Hispanic voters participated in the election than did Black voters, suggests that they were more pivotal in the passing of Prop 8 than the Black voters.

5.3 Robustness and Counterfactual Analysis

As a robustness check, I also analyze the 2004 voter turnout and presidential elections in California. Table 4 reports the results. The coefficients for Black voter turnout are very similar to the 2008 estimates. This shows that Black voter turnout, relative to Whites, was not higher in 2008 compared to the prior presidential election. Of course, that can still be consistent with the fact that more Black voters took part in the 2008 election, if more people in general, including Black voters, participated in the elections in 2008. This can be seen by a larger constant coefficient, implying that more baseline White voters went to vote in 2008. Convergence for discrete choice setup with no unobserved characteristic is no longer an issue for the 2004 results, owing to the fact that 2004 Democratic nominee Kerry was not as popular among Blacks as was Obama. Another interesting finding is the larger support for the Republican nominee by the Hispanic demographics in 2004 than in 2008. This is consistent with the anecdotal evidence that 2004 Republican nominee Bush was more popular among Hispanics than subsequent Republican candidates, McCain and Romney.

Would Prop 8 have passed without a presidential election on the same ballot? The setup's main advantages include counterfactual analysis: having a micro-founded framework allows one to estimate differences when the underlying components change. By analyzing going to multiple elections, one can look at alternative consumer behavior when some of the possibilities are eliminated. I look at the voter turnout and the total votes for Proposition 8, without the presidential election on the ticket. The setup would be similar to voting on Prop 8 during a midterm election year. I find that voter turnout is lower – at 40.8%. This is much less than the reported turnout of close to 50%, and is more in line with midterm election turnout. This further signifies the important role that presidential elections play on turnout, and it is encouraging to see reasonable estimates from the counterfactual. It is important to note that such analysis would not be possible with the single election framework, as each election is independent from others, meaning that results will not change with elimination (or addition) of different elections.

More interestingly, I also find that Proposition 8 would gather only 49.3% of the vote. It would, therefore, most likely fail to pass. The results are driven by participation in elections, which is explicitly modeled and accounted for in my setup. Certain minorities, such as Blacks and Hispanics, were in favor of both Obama and Proposition 8. Eliminating the presidential ticket on the ballot would drop the mean aggregate utility from elections below the participation threshold for some of them. If previously, they had a strong preference to vote for Obama and Proposition 8 versus not going to vote, their decision is now reduced to voting only on Proposition 8 versus not going to vote. A share of the population who have participated before would therefore stay home on election night. I find that enough people in favor of Prop 8 abstain from voting, compared to those voting No on Prop 8, which tips the scale against passing Proposition 8.

6 Conclusion

The paper develops a methodology to impute bundles of product choices from only the aggregate market shares. I also provide identification when there is no product variation across markets. I show my estimation is consistent with BLP assumptions of discrete choice analysis. I use the methodology to analyze the voting outcome for 2008 California elections. More specifically, I look at the voting preferences on presidential election and Prop 8 ballots, for each demographic group. I find that the participation of the Black population is largely exaggerated in the media, based on the exit polls, and they were more likely not to go to vote than to vote. Their results are largely consistent with the estimates from the 2004 General Election. There is a very strong preference of the Black voters for presidential candidate Obama, consistent with the poll results. However, I find that the preference of the Black population for choosing Yes on Prop 8 is not as high as suggested by the polls. The Hispanic population is more likely to vote Yes, than the Black population. This, coupled with the fact that more Hispanics voted in the 2008 election that did Blacks, suggests that the role of the Black population in helping pass Prop 8 is largely exaggerated.

Further extensions to this paper would be to expand the model to include possible correlations between different markets. In my setup, I can include a parameter σ that shows the correlation between each of the choices of the nest. The implementation is straightforward and allows an arbitrary level of substitution between choices (Berry 1994, Cardell 1997, Einav 2007). One reason is to see how substitutable election choices are relative to each other and across elections. In American politics, there is evidence that such strong preferences do not constitute the majority of the electorate. A large number of voters are in fact independent, and election campaigns are mainly targeted to attract those votes. Incorporating correlation can provide an empirical estimation of voter polarization, as my model can be set up to account for the correlation through the σ coefficient.

Perhaps an even more far-reaching analysis would be to extend the voting outcome on gerrymandering. The unique nature of gerrymandering is that voting districts are constructed by the state legislature to conform to certain requirements, such as the Voting Rights Act (1965). The general permission to pattern in any specific shape has resulted in some uniquely shaped districts. Gerrymandering is typically done at the state level, by state legislatures. Voters, however, mostly go to elections to vote for national candidates. My analysis can help answer the question of the extent that gerrymandering is affected by such a setup. Also, how do local candidates optimize their electability through gerrymandering, knowing that voter turnout is largely driven by more national candidates and issues? This is left for future work.

References

Ackerberg, Daniel, "Empirically Distinguishing Informative and Prestige Effects of Advertising," *The RAND Journal of Economics*, 2001, *32* (2), pp. 316–333.

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

_ **and** _ , "A Theory of Divided Government," *Econometrica*, 1996, *64* (6), pp. 1311–1341.

Anand, Bharat and Ron Shachar, "Advertising, the Matchmaker," *The RAND Journal of Economics*, 2011, *42* (2), pp. 205–245.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), pp. 242–262.

_ **, James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), pp. 841–890.

_ **,** _ **, and** _ , "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), pp. 68–105.

Blais, Andre, *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*, Pittsburgh: University of Pittsburgh Press, 2000.

Cardell, N. Scott, "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 1997, *13* (2), pp. 185–213.

Davis, Peter, "Spatial Competition in Retail Markets: Movie Theaters," *The RAND Journal of Economics*, 2006, *37* (4), pp. 964–982.

Degan, Arianna and Antonio Merlo, "A Structural Model of Turnout and Voting in Multiple Elections," *Journal of the European Economic Association*, 2011, *9* (2), pp. 209–245.

Einav, Liran, "Seasonality in the U.S. Motion Picture Industry," *The RAND Journal of Economics*, 2007, *38* (1), pp. 127–145.

Goolsbee, Austan and Amil Petrin, "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 2004, *72* (2), pp. 351–381.

Gordon, Brett and Wesley Hartmann, "Advertising Effects in Presidential Elections," *Marketing Science*, 2013, *32* (1), 19–35.

Greiner, D. James and Kevin Quinn, "Exit Polling and Racial Bloc Voting: Combining Individual-Level and R x C Ecological Data," 2013. Working Paper.

Hendel, Igal, "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *The Review of Economic Studies*, 1999, *66* (2), pp. 423–446.

King, Gary, *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton: Princeton University Press, 1997.

Martin, Gregory, "The Informational Content of Campaign Advertising," 2013. Working Paper.

Nevo, Aviv, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 2001, *69* (2), pp. 307–342.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, *110* (4), pp. 705–729.

Poole, Keith and Howard Rosenthal, "A Spatial Model for Legislative Roll Call Analysis," *American Journal of Political Science*, 1985, *29* (2), pp. 357–384.

Rekkas, Marie, "The Impact of Campaign Spending on Votes in Multiparty Elections," *The Review of Economics and Statistics*, 2007, *89* (3), pp. 573–585.

Voting Rights Act, Public Law 89-110, 79 Stat. 437, 1965.

Wong, Maisy, "Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore," 2013. Working Paper.

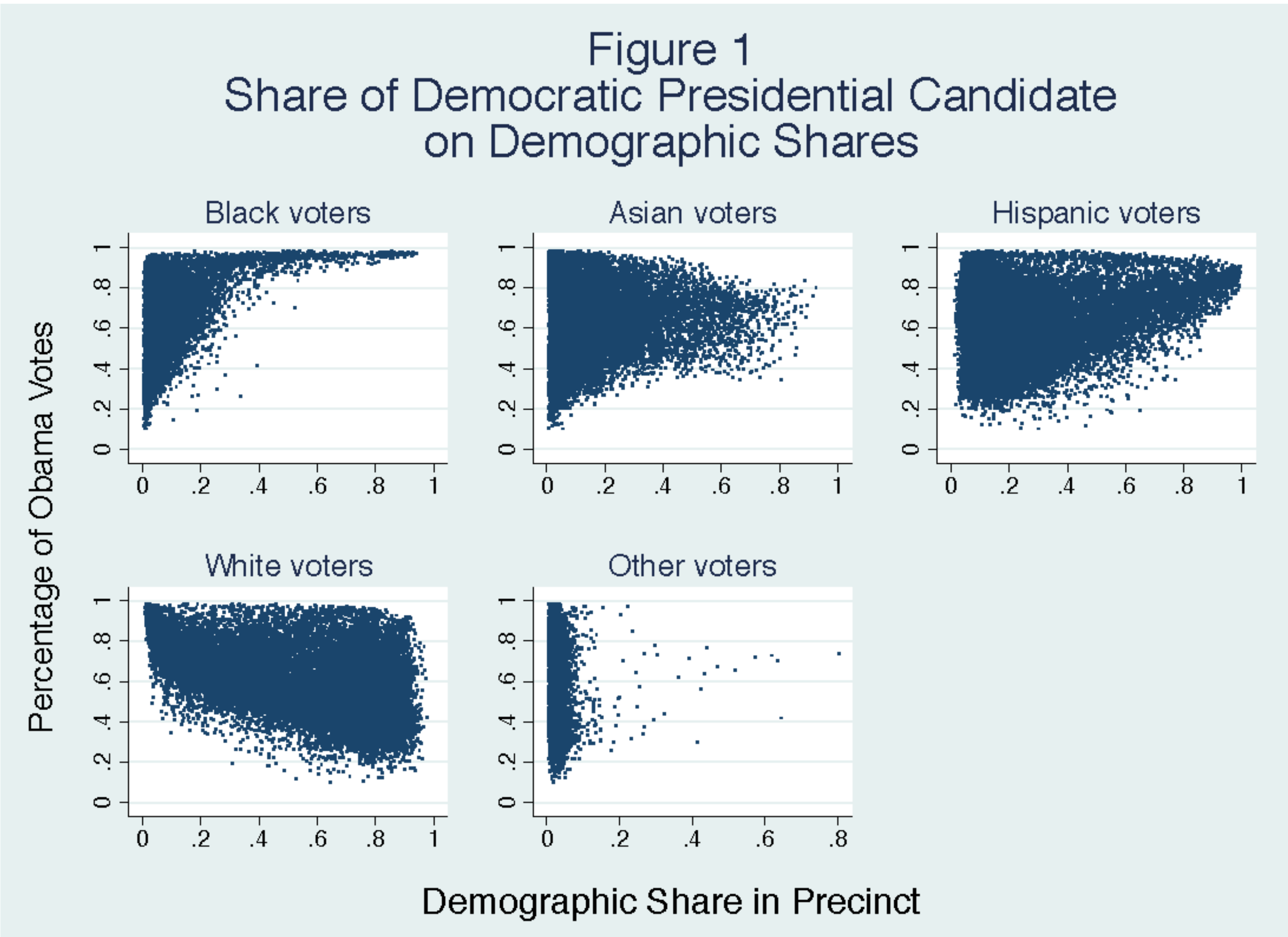


Figure 1
Share of Democratic Presidential Candidate
on Demographic Shares
Black voters
Asian voters
Hispanic voters
White voters
Other voters
Percentage of Obama Votes
0
.2
.4
.6
.8
1
Demographic Share in Precinct

Figure 2
Share of Proposition 8 Yes Votes
on Demographic Shares

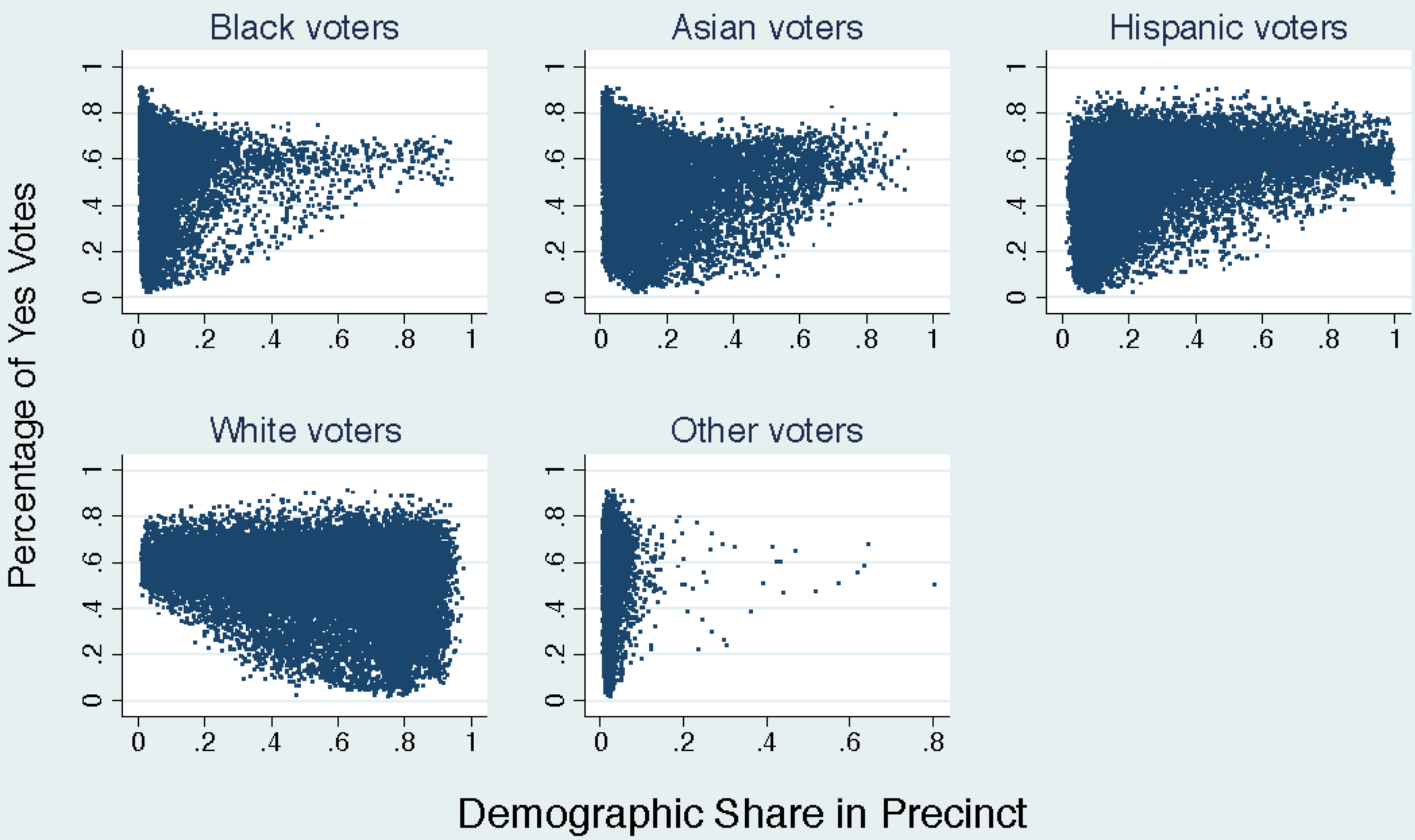
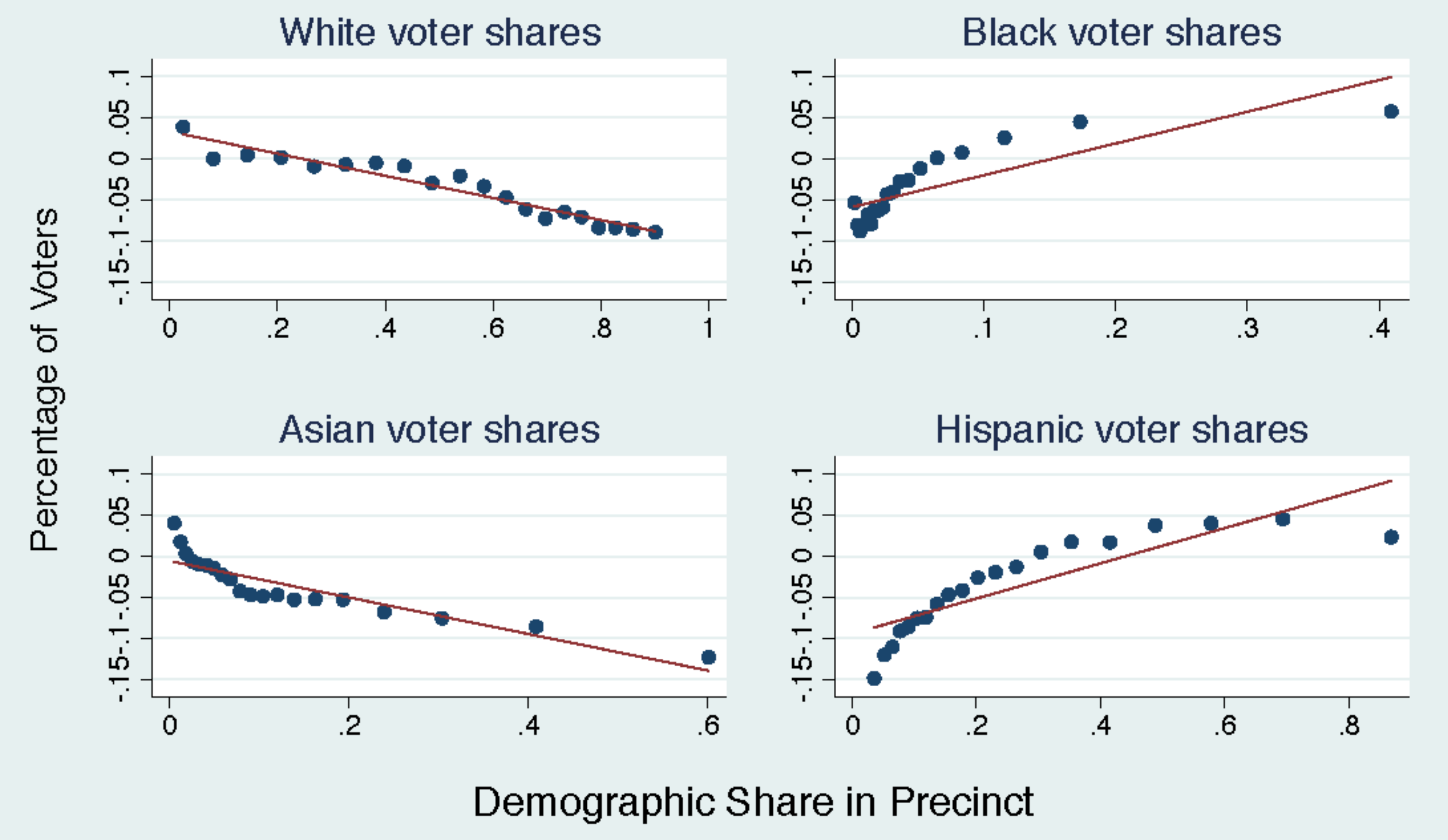


Figure 3
Difference in Share of Participation
on Demographic Shares



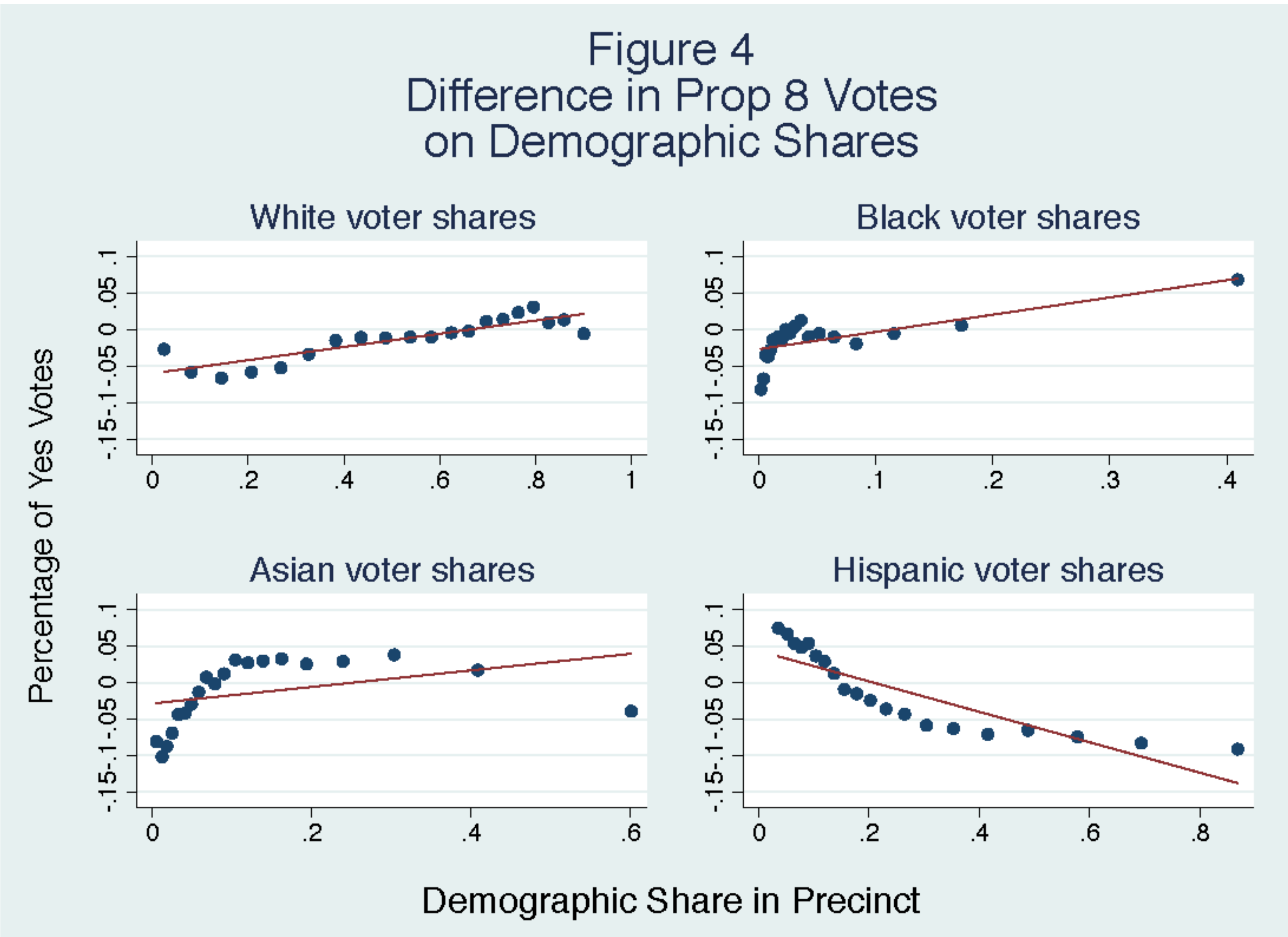


Figure 4
Difference in Prop 8 Votes
on Demographic Shares
White voter shares
Black voter shares
Asian voter shares
Hispanic voter shares
Percentage of Yes Votes
-.15 -.1 -.05 0 .05 .1
0 .2 .4 .6 .8 1
0 .1 .2 .3 .4
0 .2 .4 .6
0 .2 .4 .6 .8
Demographic Share in Precinct

Figure 5. The

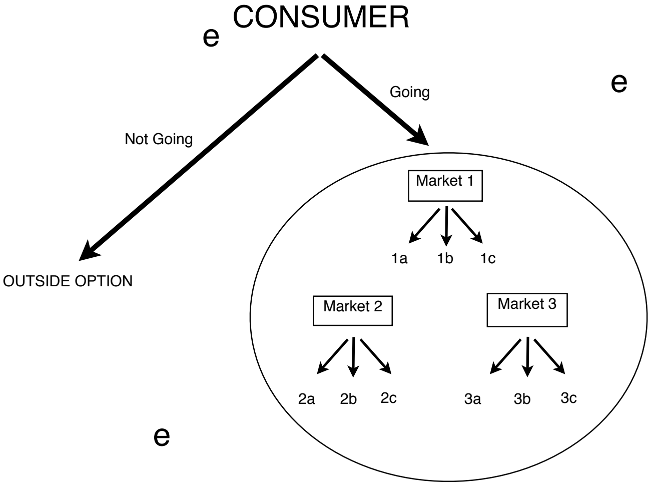

CONSUMER
Not Going
Going
OUTSIDE OPTION
Market 1
1a
1b
1c
Market 2
2a
2b
2c
Market 3
3a
3b
3c

Table 1. Summary Statistics

CNN Poll results for voter demographics and shares of election choice

	Pop share	Proposition 8 Yes	Proposition 8 No	Presidential Election Dem	Presidential Election Rep	Presidential Election Other
White	63%	49%	51%	52%	46%	2%
African-American	10%	70%	30%	94%	5%	1%
Latino	18%	53%	47%	74%	23%	3%
Asian	6%	49%	51%	64%	35%	1%
Other	3%	51%	49%	55%	41%	4%

Total votes cast

Proposition 8		
Yes	7,001,084	52.24%
No	6,401,482	47.76%
Presidential Race		
Democrat	8,274,473	61.1%
Republican	5,011,781	37.0%
Other	275,646	1.9%

Population Demographics

	Census 2000	Census 2010	2008	Mean Share	St Dev	Min	Max
White	12,593,932	16,101,075	11,991,899	52%	27%	0%	97%
Black	1,528,106	1,725,443	1,686,836	6%	10%	0%	93%
Asian	2,793,113	3,648,163	3,804,575	14%	15%	0%	92%
Hispanic	6,915,731	10,517,229	8,618,187	27%	23%	1%	99%
Other	542,033	799,820	431,456	2%	2%	0%	80%
Total	24,621,819	27,958,916	26,532,954				

Notes: The Mean, Standard Deviation, Minimum and Maximum are calculated per precinct.

Table 2. The Effect of Demographics on Going to Vote

	OLS	$\nu = 0$	OLS	$\nu = 0$	Full Model
Black	-0.729**	-0.955**	-0.730**	-0.891**	-0.961**
	(0.0426)	(0.122)	(0.0437)	(0.115)	(0.074)
Asian	-1.637**	-1.694**	-1.470**	-1.538**	-1.6951**
	(0.0293)	(0.051)	(0.0296)	(0.047)	(0.046)
Hispanic	-2.759**	-3.082**	-2.664**	-2.973**	-2.986**
	(0.0193)	(0.049)	(0.0196)	(0.047)	(0.03)
Other	-7.141**	-14.635**			
	(0.222)	(5.42)			
ν					0.067
					(0.0434)
Const	1.355**		1.184**		
	(0.0101)		(0.00879)		
N	18807	18807	18807	18807	18807
R-sq	0.536		0.511		

Notes: Standard errors in parentheses. + Significant at 10%; * Significant at 5%; ** Significant at 1%.

Estimated Share of Voters by Demographics

	Discrete Choice	CNN Poll
White	70%	66%
Black	7%	10%
Asian	12%	6%
Hispanic	11%	18%

Table 3a. The Effect of Demographics on Presidential Candidate Choices

	OLS	$\nu = 0$	Full Model
Black	-4.008**	-19.2112	-7.1176**
	(0.0423)	(598835.8)	(0.2949)
Asian	-0.950**	-1.09778**	-0.76015**
	(0.0384)	(0.0294)	(0.0635)
Hispanic	-1.245**	-1.77063**	-.92331**
	(0.0320)	(0.0238)	(0.2695)
ν			-0.3527**
			(0.0956)
Const	-0.0713**		
	(0.00852)		
N	18807	18807	18807
R-sq	0.391		

Notes: The Dependent variable is the share of votes for the Republican Candidate (McCain). Standard errors in parentheses. + Significant at 10%; * Significant at 5%; ** Significant at 1%.

Table 3b. The Effect of Demographics on Proposition 8 Choices

	OLS	$\nu = 0$	Full Model
Black	0.281**	0.32022**	0.8764**
	(0.0409)	(0.022691)	(0.0981)
Asian	-0.0482	-0.03363+	-0.43205**
	(0.0371)	(0.022176)	(0.1472)
Hispanic	1.176**	1.1973**	1.3534**
	(0.0309)	(0.025678)	(0.04438)
ν			0.7085**
			(0.01068)
Const	-0.0567**		
	(0.00823)		
N	18807	18807	18807
R-sq	0.079		

Notes: The Dependent variable is the share of Yes votes on Proposition 8. Standard errors in parentheses. + Significant at 10%; * Significant at 5%; ** Significant at 1%.

Imputed and Poll Probabilities of Voting Yes on Prop8

	Discrete Choice	CNN Poll
Black	57%	70%
Asian	49.10%	49%
Hispanic	74.40%	53%
White	48.50%	49%

Table 4. Demographics on Going to Vote and Presidential Choice (2004)

	Going to Vote			Presidential Election		
	OLS	ν = 0	Full Model	OLS	ν = 0	Full Model
Black	-0.708**	-0.864**	-0.9185**	-1.789**	-4.742**	-6.26**
	(0.0425)	(0.0866)	(0.0606)	(0.0317)	(0.16)	(0.398)
Asian	-1.302**	-1.35**	-1.562**	-1.056**	-1.38**	-1.322**
	(0.033)	(0.045)	(0.0433)	(0.033)	(0.036)	(0.069)
Hispanic	-2.809**	-3.25**	-3.175**	-0.629**	-1.614**	-1.202**
	(0.0213)	(0.055)	(0.041)	(0.0127)	(0.0232)	(0.039)
ν			-0.045*			-0.422**
			(0.018)			(0.0875)
Const	1.022**			1.406**		
	(0.0088)			(0.0055)		
N	18307	18307	18307	18307	18307	18307
R-sq	0.624			0.469		

Notes: Standard errors in parentheses. + Significant at 10%; * Significant at 5%; ** Significant at 1%.

Table 5. Sample and Population Shares

Overall Population Racial Breakdown

White	Black	Asian	Hispanic	Other
45.7%	6.4%	14.2%	32.2%	1.6%

My Sample Racial Breakdown

White	Black	Asian	Hispanic	Other
45.2%	6.4%	14.3%	32.5%	1.6%

Overall Population Voting Breakdown

Participation	Obama	McCain	Prop 8 Yes	Prop 8 No
50.1%	61.0%	37.0%	52.2%	47.8%

Overall Voting Breakdown between Obama/McCain, Yes/No

Pres. Partic.	Obama	McCain	Prop 8 Partic.	Yes	No
48.4%	62.3%	37.7%	48.8%	52.2%	47.8%

My Sample Breakdown between Obama/McCain, Yes/No

Pres. Partic.	Obama	McCain	Prop 8 Partic.	Yes	No
48.1%	62.6%	37.4%	48.5%	52.3%	47.7%