CALTECH/MIT
VOTING TECHNOLOGY PROJECT

A multi-disciplinary, collaborative project of
the California Institute of Technology – Pasadena, California 91125 and
the Massachusetts Institute of Technology – Cambridge, Massachusetts 02139

WHERE THE GOOD SIGNATURES ARE:
The Number and Validity Rates of Initiative Petition Signatures Gathered in California Counties

Frederick J. Boehmke
University of Iowa

R. Michael Alvarez
Caltech

Where the Good Signatures Are:
The Number and Validity Rates of Initiative Petition Signatures Gathered In California Counties

Frederick J. Boehmke
University of Iowa

R. Michael Alvarez
California Institute of Technology

January 23, 2004

Where the Good Signatures Are:

The Number and Validity Rates of Initiative Petition Signatures Gathered In California Counties

Abstract

One of the biggest hurdles to qualifying an initiative for the ballot is gathering the required number of signatures. Yet little is known about these signatures' representativeness or demographic origin. Using data from eight recent California initiatives, we study the distribution of signatures across counties. We then conduct regression analysis to study how characteristics of counties relate to the number of signatures gathered there. Besides the source of petition signatures, we also study the results of checks performed by county officials on these signatures to determine whether there are sufficient valid signatures (i.e., from registered voters) in order to qualify the petition for the ballot. We then conduct regression analysis of validity rates by county. Our findings indicate that a variety of social and political factors influence both the number and validity rate of signatures across counties, though the results are more consistent for the former. Finally, we demonstrate that signature gathering campaigns have consequences for outcomes by relating the intensity of the signature campaign within a county to ballot roll-off.

1   Introduction

The process of direct legislation provides citizens and organized interests the opportunity to submit legislation directly to a state's voters rather than attempt to persuade the legislature to enact it for them. The process has seen a resurgence in use and attention over the last quarter century, particularly in the last decade or so. As the frequency of use has increased, so has the amount of scholarly and popular attention. The majority of this attention has been in three areas: the initiative process' consequences for state policy outcomes, whether it infringes on minority rights and whether voters are sufficiently informed to make direct policy decisions. Less attention has been paid to the process whereby initiatives qualify for the ballot. Further, almost all of the studies of the qualification process are legal studies, case-study analyses or are descriptive studies based on interviews. In this paper we study one of the more important aspects of the initiative qualification process: signature gathering.

To see the importance of qualification hurdles, consider the fact that since adopting the initiative process in 1911, California has had 1,187 initiatives titled for circulation. Of those, only 290 (25%) have reached the ballot (Shelley 2002). This discrepancy is certainly caused by many factors, including legal issues, legislative action and lack of resources, but gathering the necessary signatures is often considered the greatest hurdle in successfully qualifying an initiative. Across the twenty-four states with some form of direct legislation, signature requirements vary from four to fifteen percent of votes cast in the previous election. This produces great disparities in the number of signatures required per state; while California has one of the lower percentage requirements at five percent for statutory and eight percent for constitutional initiatives, this translates into almost half a million and two-thirds of a million signatures, respectively.

This vast number of signatures required has led critics to proclaim that qualifying an initiative is not a test of public support, but a test of the depth of its supporters' pockets. Ballot access, it follows, has effectively been restricted to large and wealthy groups rather than the grassroots movements that Populist and Progressive reformers envisioned availing

themselves of the process. While there is still debate about whether broad-based interests have been sidelined from the process, it is clear that it takes a significant amount of money to place measures on the ballot.[1] It is often said that an individual or group can place any measure on the California ballot if they are willing to spend a million dollars, though costs can increase when many petitions are circulating and time is running out.

Another way in which reformers have proposed placing new controls on the initiative process is by imposing geographic distribution requirements on the signatures required to get a proposal onto an election ballot. Forcing proposal proponents to get signatures from every county in a state, or to obtain some minimal threshold of signatures from every county, could increase the costs of proposing a measure for the ballot.[2] Also, it could help insure that proposals that make it to the ballot have widespread geographic support in a state.

Despite a number of studies and government inquiries into the initiative process that have generated a general understanding of how signature gathering is organized and how the validation process works, there is no work that we know of that analyzes the source and validity of actual signatures.[3] What type of people sign petitions? Do signature gatherers tend to focus their efforts on specific counties? Are certain types of voters or counties more likely to produce invalid signatures? How do petition campaigns influence individual behavior and ballot outcomes? Would a geographic distribution requirement have an impact on the number and type of ballot measures that qualify for an election ballot?

These questions are important for a variety of reasons. First and foremost among them is the need to develop an ever-better understanding of one of the most crucial, costly parts of an increasingly important component of political activity in dozens of states. Vast sums of money are spent in many states (like California) in the attempt to qualify measures for the ballot, and there is now an industry that works on signature gathering for political groups and candidates; Ellis (2003) compares the signature gathering process for initiatives to primaries for candidates. Studying how the signature gathering process works is therefore critical for producing a better understanding of the politics of the initiative process. Beyond a need to understand process, however, it is also important to understand how specific steps in the

process affect its accessibility to different interests in society. Since the initiative process is used by citizen groups that are often under-represented among those that typically lobby the legislature (Gerber 1999), whether and how the signature gathering process restricts ballot access vis-a-vis traditional groups is an important component for understanding the effect of the initiative process on representation. If it is easier to gather signatures from certain types of voters then it becomes possible that issues associated with certain interests may qualify for the ballot more easily. Similarly, the issue of signature validation is also important since if signatures from certain types of voters or from specific counties are less likely to be valid, then there will again be consequences for which issues reach the ballot. In addition, systematic invalidation of signatures from specific voters may undermine their equality of representation.

We address these questions in this paper by analyzing signatures gathered for eight California initiative petitions that made it to the ballot between 2000 and 2003. We obtained county-level reports from the Secretary of State's Office that include the total number of signatures gathered, the number of signatures verified and the number that were valid. We use these data to study variation in the number of signatures gathered per county as well as variation in the validity of signatures per county. We also analyze the distribution of signatures relative to population. We then conduct regression analysis using models for event count data that allow us to relate county characteristics to the number of signatures gathered as well as to the validity rate. Our results indicate that a variety of social and political variables influence both the number of validity of signatures gathered. Lastly, we provide evidence that signature gathering campaigns matter for outcomes by regressing roll-off rates on the intensity of signature gathering in a county.

2   The Signature Gathering Process

To satisfy the demand for placing measures on the ballot, an entire industry arose in California to manage all stages of initiative campaigns. Estimates put the total number of firms

involved in the initiative process at nearly one hundred, almost all of which are located in California (McCuan, Bowler, Donovan and Fernandez 1998). The number of firms involved in signature gathering is much smaller, with about a half dozen full-time firms in California (McCuan, Bowler, Donovan and Fernandez 1998). The largest firms are Kimball Petition Management and National Petition Management; smaller firms include American Petition Consultants, National Voter Outreach, Masterton and Wright, and Progressive Campaigns (Broder 2000).

The existence of an initiative consulting industry is not a recent phenomenon, however. Almost immediately after the adoption of direct legislation provisions in California in 1911, temporary consultants began to spring up to assist campaigns. While attempting to qualify a referendum in 1912, dairymen reportedly paid ten cents a signature to gather 23,000 signatures (Goebel 2002); a more typical rate at the time was about five cents per signature, corresponding to around 80 cents today (Ellis 2003). The first full-time permanent firm, Whitaker and Baxter's Campaigns Inc., was established in 1930 and handled five or six initiatives per election (McCuan, Bowler, Donovan and Fernandez 1998). Very early on, many states responded to the role of professional circulators by banning their use.[4]

Today, these petition management firms typically hire independent contractors, often on a regional basis, to gather signatures. Contractors typically make five to twenty cents per signature. These "crew chiefs" then hire individuals to actually circulate the petitions, paying anywhere from twenty-five to fifty cents per signature (Broder 2000; McCuan, Bowler, Donovan and Fernandez 1998). On a good day, a petition circulator can make fifty to eighty dollars an hour; even more when time is running out and groups are desperate to qualify their measure.

While the intent of Progressive reformers may have been to generate public discussion of issues during the signature gathering process, it rarely works that way. Petition circulators usually approach individuals in high-traffic areas, such as shopping malls, and ask them if they are registered to vote and, if so, to sign their petitions. People who ask for more information about the issue are generally offered a one or two sentence summary of the

proposal or are asked to step aside as they read the petition. If someone is not sure how they feel about the measure, circulators often appeal to democratic instincts by asking to them to sign so the proposal can get on the ballot for the people to decide. There is little incentive for circulators to spend time explaining the measure when most people who are approached will sign with little prompting — between half and three-quarters of individuals sign without reading the petition (Cronin 1989).[5].

In a potentially extreme, but oft-cited example of the signature gathering process, Ed Koupal describes what he refers to as the "hoopla process":

> First, you set up a table with six petitions taped to it, and a sign in front that says, SIGN HERE. One person sits at the table. Another person stands in front. That's all you need — two people.
>
> While one person sits at the table, the other walks up to people and asks two questions. We operate on the old selling maxim that two yeses make a sale. First, we ask them if they are registered in that county. If they say yes to that, we immediately push them up to the table where the person sitting points to a petition and says, "Sign this." By this time, the person feels, "Oh goodie, I get to play," and signs it. If the table doesn't get 80 signatures an hour using this method, it's moved the next day.[6]

At least in part because of similar portrayals of the signature gathering process, states have attempted to regulate it more strictly. In hope of reducing the relationship between money and ballot access, the most widely adopted approach has been to ban the use of paid signature gatherers. Yet the Supreme Court has ruled in *Meyer v. Grant* [486 U.S. 414 (1988)] that such bans violate the First Amendment and that there is not a sufficient state interest in regulating the process to reduce fraud to warrant such an infringement (Ellis 2003; Lowenstein and Stern 1989; Tolbert, Lowenstein and Donovan 1998).[7] Colorado responded to this decision by requiring petition circulators to wear name badges, but the Court also ruled that a violation. The closest states have come so far is to require them to wear badges

or that say "volunteer" or "paid" signature gatherer or that they are registered to vote in that state.

In California – the state that is the subject of our study – there are two types of initiative measures that are required to be signed by a certain number of registered voters: initiative statutes and initiative constitutional amendments.[8] Initiative statutes are measures that would amend California statute, and require signatures equivalent to 5% or more of the votes cast for all candidates in the previous gubernatorial election.[9] Initiative constitutional amendments seek to amend the state constitution and require signatures equivalent to at least 8% of the votes cast for all candidates in the past gubernatorial election. Unlike some other initiative states, California has no distribution requirement that requires a certain share of signatures to originate in a fixed number of counties.[10] California does, however, have one of the shortest signature-gathering periods, allowing only one hundred and fifty days from the date the petition is titled to gather the required number of signatures.

After the signed petitions are submitted to the state, there are a variety of procedures for checking if there are sufficient valid signatures from registered voters to qualify the measure. Most states check either all of the submitted signatures or a random sample. In California, county officials check three percent of signatures submitted in their county, with a five hundred signature minimum.[11] Each official then projects the expected number of valid signatures in their county and reports the result to the Secretary of State's office.[12] Once all counties have reported, the projected statewide total number of valid signatures is calculated. If the number of estimated valid signatures is greater than 110% of the total number required, the initiative qualifies for the ballot; if the number is less than 95% of that required, it is rejected. When the total falls between these two points, counties conduct a full verification of all signatures submitted.

States vary somewhat in their definition of valid signatures as well. Except for North Dakota (which has no voter registration), signatures must come from registered voters. In California, voters are required to list their address as well. If the address does not match the one listed with their voter registration, the signature may be deemed invalid. In Colorado,

signatures have been invalidated if voters names do not match exactly what is on the registration rolls (Tolbert, Lowenstein and Donovan 1998). Other states invalidate petitions that have stray marks on them, including coffee stains. Because a certain number of signatures will fail to meet the potentially stringent validity requirements, firms tend to "overqualify" their petitions, gathering up to 175% of the number actually required.

3   The Distribution of Total and Invalid Signatures

To study the types of characteristics that make counties more attractive for gathering signatures and whether there is systematic variation in the validity rate of those signatures, we obtained data on eight petitions from the California Secretary and State's office. These data include the following information for each petition and each county: the number of signatures gathered, the number of signatures sampled, and the number of valid signatures in the sample.[13] In addition, two of these potential initiatives were projected to have total signatures within the 95-110% range of the number required for qualification, triggering a full check of all signatures. We analyze the number of invalid signatures from the full checks separately from the initial random sample checks.

The potential ballot measures that we have data on all had sufficient signatures to reach the ballot; some were overqualified by wide margins while one just barely made it. Further, four of our measures are constitutional amendments and needed to gather signatures equal to eight percent of gubernatorial turnout rather than the five percent required for statutory initiatives. We provide basic information about each of the eight petition signature efforts in Table 1.

[**Insert Table 1 here.**]

The measures involved cover a wide array of issues, over a range of election cycles. Three of the measures in our database were on the general election ballot in the 2000 presidential election. Petition number 830, which was Proposition 35 on the ballot, covered state public works projects and issues associated with the use of private contracts for those projects: this

7

passed with 55.2% of the votes cast in this election. Also in the 2000 general election were petition 656 (on the ballot as Proposition 36) and petition 874 (Proposition 37). The former dealt with drug treatment programs, and passed with 60.9% of votes cast; the latter would have lowered the vote threshold for passage of new taxes, and this measure failed to pass, with only 47.9% of votes cast supporting passage.

One of the petition measures we have data on was on the 2002 primary election ballot: petition number 918, which became Proposition 45 on the spring primary ballot. This measure would have altered the term limits law for legislators in California, and it was defeated after receiving 42.3% of votes cast. From the 2002 general election we have data on three measures: 935, 936, and 952. Measure 935 was renumbered to be Proposition 51 on the ballot, regarding the distribution of transportation taxes, and it received only 42.2% of yes votes and failed to pass. Measure 936 became Proposition 52, an initiative that would have changed voter registration laws to usher in election day registration in California; this measure also failed to pass, gathering only 40.9% of yes votes. Last on the ballot in the 2002 general election was Proposition 49, which had been circulated for signatures as measure 952. This measure provided new funding for before and after school programs, and it passed after receiving yes votes from 56.7% of voters.

The last measure in our database was 933, which was placed on the October 2003 statewide recall ballot as Proposition 54. This measure would have barred the state from collecting racial and ethnic data, but it failed to pass with only 36.1% of voters casting ballots in support of passage.

Next, we present descriptive statistics for each initiative in Table 2. The total number of signatures for the four statutory measures ranges from 713,849 to 784,984. This represents significant overqualification since only 419,260 signatures were needed to meet the five percent threshold. The constitutional initiatives needed about 670,816 signatures; the four in our sample had between 957,370 and 1,089,042 signatures. While this represents an average overqualification of seventy-five percent for the statutory and fifty percent for the constitutional proposals, it is still the case that one of each type had a sufficiently low validity rate

in the random sample check that full counts were triggered.

[**Insert Table 2 here.**]

More specifically, we find that for all of the ballot measures we have data on, signatures collected were 162% more than required. There is not an enormous amount of variability in the overqualification percentages across each measure, with a low of 143% (measure 187) and a high of 187% (measure 952). This suggests that signature gathering campaigns aim for a target of over 150% of necessary signatures.

Turning now to the county level, we find that each petition had signatures gathered from every county, with the exception of one measure (952) for which no signatures were submitted in Alpine county. This is interesting in light of the fact that California has no distribution requirements that force circulators to gather a minimum number from a fixed number of counties.[14] The number of signatures gathered does vary widely, however, with twenty-six instances of counties with ten or fewer signatures in our sample. At the other extreme is Los Angeles County with over a third of a million signatures gathered for one petition and an average of 210,000 for the statutory initiatives and 276,000 for the constitutional initiatives. The only other counties that break six digits are San Diego (three times) and Orange (once). The average number is around 17,600 for constitutional initiatives and 12,800 for statutory proposals.

Yet despite this wide disparity in the number of signatures gathered per county, the overall distribution is surprisingly equitable. To determine this we constructed two variables — one that measures the proportion of a petition's signatures gathered in a county and a second that measures the proportion of Californians living in the same county. These two variables have a correlation of 0.97 and a regression of the former on the latter produces a constant of zero and a coefficient of 0.999.[15] This indicates an almost perfectly equitable distribution of signatures across counties relative to their size. It also suggests that the imposition of a distribution requirement would only distort the signature gathering process and that large areas such as Los Angeles and San Diego are not overrepresented in terms of determining what makes it to the ballot.

Using our data, we can investigate whether a distribution requirement would have any effect on whether the eight initiatives in our study would have reached the ballot. Since California does not have a distribution requirement we can not say exactly what would happen, but we can see how stringent a requirement would have to be to keep our initiatives from qualifying. Since we have county data and since the majority of distribution requirements are county-based, we focus on this type of requirement. Using data for the vote in the 1998 gubernatorial race, we can calculate the number of counties in which the statewide signature threshhold was met or exceeded. For statutory initiatives this means counting the number of counties where the number of signatures obtained exceed five percent of the county's turnout in the 1998 governor's race; for constitutional initiatives we use an eight percent threshhold. These calculations indicate that it would take a strict distribution requirement to disqualify these initiatives, particularly statutory ones. The four statutory proposals met the threshhold in twenty-eight to thirty-four of the fifty-eight counties; constitutional proposals did so in twenty-two to thirty counties. Except for Nevada and Wyoming, county-based requirements are generally for half the counties or fewer (and often involve percentage less than the statewide requirement). Given that our petitions met the requirement in about half the counties without even trying, a low or moderate requirement would probably not inhibit qualification.

An alternate way to interpret how equitably signatures are distributed across counties is to calculate a Gini index for the number of signatures per capita in each county. Gini indices are commonly used as a measure of inequality and vary from zero (perfect equality) to one (perfect inequality); they provide us a way to make comparisons between initiatives. The Gini indices for our eight initiatives are included in our table and vary from 0.38 for petitions 830 and 952 to 0.50 for petition 865. To put these values in perspective, they are about the same as those for the distribution of income across U.S. households. In combination with the correlation and coefficients discussed in the previous paragraph, these results suggest that while, on average, signatures are gathered equitably from across counties, there is significant variation across counties of similar size.

In terms of validity, there is wide variation in the proportion of submitted signatures that are determined to be valid during the random sample validation.[16] Overall, our seven random sample checks indicate an average validity rate of 78%.[17] Ignoring cases with ten or fewer signatures, the validity rate varies from fifty-nine percent to ninety-six percent, with an average of seventy-eight percent.[18] In addition, there is little difference in the validity rate of constitutional and statutory initiatives. This suggests that petition managers' strategy of gathering extra signatures is well justified (albeit based on our sample of successfully qualified petitions) and that almost one third extra signatures is needed just to average the minimum required.[19]

Examining the two cases where validation was performed on all the signatures, we find a 75.5% validity rate, slightly below the validity rate for the random sample tests. This is not due to differences between the proposals in the two samples, either: for the one measure that we have both random and full sample validation, the validity rates are 78.3% and 75.6%, respectively. The difference is highly statistically significant.[20]

These results provide an interesting view of the signature gathering process in California, suggesting that signatures are gathered rather equitably from among the fifty-eight counties and that the overall validity rate of these signatures does not vary much at the aggregate level. Yet there is considerable variation across initiatives in terms of the absolute and per capita distributions across counties; there is also considerable variation across counties in the validity rate. Our next task is to explore whether these variations are systematically related to political and social characteristics of the counties.

4   Signature and Validity Rates By County

In this section we study the signature gathering process and the strategies employed by petition managers by determining whether they focus on certain counties. It is also important to study how social and demographic factors relate to signature gathering as this is the most important and expensive hurdle to getting issues on the ballot. Following the analysis of

the number of signatures gathered per county, we study how the same factors relate to the validity of signatures gathered. In particular, we are interested in whether there exists systematic variation in validity rates across counties. Lastly, we link petition campaigns to individual behavior by relating the intensity of signature gathering campaigns within a county to ballot roll-off rates on the same measure.

The objective of petition managers is generally straightforward: gather signatures sufficient to qualify specific petitions for the ballot. As a business, we should expect them to do so as cheaply as possible, subject to the necessity of gathering enough valid signatures to surpass the qualification threshold. Petition managers therefore face a potential trade-off between paying more to gather fewer high-quality signatures and paying less for more low-quality signatures. High-quality signatures have a greater chance of being valid and should therefore be worth more from a business perspective. There is evidence of this trade-off in our data: the number of signatures gathered in a county is negatively, though weakly, correlated with the validity rate of those signatures and the validity rate for constitutional proposals, which require 60% more signatures, is one percent lower than for statutory proposals.

This logic is also consistent with statements from signature gatherers, who emphasize their strategy of going to locations with lots of people: "The most effective version of this technique is for a single circulator to work long, slow-moving lines of people waiting to get into a movie, play, concert or other event."[21] Or, in the words of Ellis: "The key variable is not the attitude of those who sign the petition .... the key variable is the number of people who can be solicited" (Ellis 2003). While opportunities of this sort exist in any county, we expect that petition managers will focus their efforts on high density counties and test this by including a variable *Population Density* that measures the number of people per square mile in each county. We expect that denser counties should produce more signatures, but that they should have greater invalidity rates. Additionally, we control for a county's *Total Population* and expect it to increase the number of signatures, but also to increase the invalidity rate.

To control for social and economic characteristics, we include measures of race, age,

education, unemployment and income. For some of these variables, we have no specific expectation regarding their effect on the number of signatures, though they may have a specific effect on validity rates. In general, we expect validity rates to be lower in areas where there is greater mobility since one would expect that to increase the chance that a voter's current address is out of date with that listed on the official registration rolls. In addition, areas with younger, poorer citizens may produce signatures from people who are less likely to be registered at all. Thus we would expect greater *Median County Age*, greater *Per Capita Income*, greater percent of the population with a *High School Education* and lower *Unemployment Rates* to be associated with lower shares of invalid signatures.

Lastly, we also include measures of political characteristics, including partisanship and competitiveness. Partisanship, measured with the *Democratic Vote Share* in the 2002 gubernatorial election, may have a different effect for specific initiatives since the issue content of a proposal may help determine which areas petition managers target. Of course given the wisdom that the vast majority of signers do not read the petition, partisanship may not have a large effect. There may also be an indirect effect of issue content since some initiative consultant and petition management firms tend to specialize in issues associated with a specific ideological perspective (McCuan, Bowler, Donovan and Fernandez 1998). If firms tend to focus on certain geographic areas and certain issues are more likely to be taken to certain firms, then this could produce a partisan effect as well. Of course, signature gatherers generally carry multiple petitions at once, so this effect is probably muted.

Competition may matter because of its relationship to overall political activity. Further, since competition is generally thought to be associated with greater levels of political fraud, we might expect more competitive counties to produce more invalid signatures.[22] We measure competition with the *Margin in the Governor's Race* — the absolute value of the difference between the Republican and Democratic vote shares in the 2002 gubernatorial election.

Taking these factors into account, we estimate models for the number of signatures gathered in each county and then for the proportion of invalid signatures in each county. Validity rates are analyzed separately for the seven random samples and the two full samples. We

also estimate separate models for each initiative.

Because our dependent variable is the number of signatures gathered in a county, we employ models for event count data. Event count models such as Poisson and negative binomial regression are superior to linear regression models for these type of data. The difference is particularly important for our analysis since we have many counties that produce very small numbers of signatures; linear regression models have been shown to produce biased coefficients (King 1989a). While the negative binomial and Poisson models are quite similar, the Poisson suffers from the limitation that it assumes that the variance is equal to the mean ($E[Y] = Var[Y]$). The negative binomial relaxes this assumption by allowing for overdispersion, which permits the variance to be greater than the mean. When this is the case, the Poisson parameter estimates are correct, but the standard errors are incorrect (King 1989b).[23]

Overdispersion is caused by two phenomena: contagion and heterogeneity. Contagion would occur if the gathering of one signature in a county increases the chances of another signature being gathered in the same county during the same time period. Since signatures are gathered by individuals attempting to gather as many signatures as possible, it seems quite likely that contagion exists. Heterogeneity would arise if different voters within a county have different probabilities of signing a petition or if different signatures gathered within the same county have different probabilities of being deemed invalid. Both of these seem likely as well. Thus we expect our data to exhibit strong overdispersion and estimate negative binomial regression models.

4.1 Total Signatures

The analysis of the number of signatures gathered per county is presented in Table 3. Overall the results indicate that a variety of factors influence how many signatures are gathered per county. In addition, the models greatly increase our ability to explain the data, as evidenced by the small p values associated with the χ^2 statistics for explanatory power. The results

14

are very consistent across the eight initiatives in terms of the direction of the effect as well as the significance of the different independent variables. Because of this consistency, we discuss only the overall patterns rather than focus on the effects for individual initiatives.

[**Insert Table 3 here.**]

Our expectations about the strategies of signature gatherers are borne out by the data. More signatures are gathered from larger counties and more signatures are gathered from denser counties. Both of these variables are significant at the five percent level or better in every model. To obtain a better understanding of the substantive impact of these variables on the number of signatures gathered per county, we calculate first differences for each independent variable. These marginal effects represent the change in the dependent variable resulting from an increase in each independent variable from one standard deviation below its mean to one standard deviation above its mean, holding all other independent variable constant at their mean values. These results are presented in the top panel of Table 4. The calculations indicate that changes in population have a slightly greater effect than changes in density: population has a marginal effect of 6,222 whereas density has a marginal effect of 4,260.[24] Recall that the average number of signatures per county is about 15,000.

[**Insert Table 4 here.**]

There are also important effects for our social and demographic variables. We find that counties with a greater median age produce significantly fewer signatures, both in the pooled regression and for five of the eight separate petitions. The marginal effect is around six thousand for statutory measures and eight thousand for constitutional initiatives. At the same time, counties with greater per capita income produce significantly more signatures, with a first difference of almost 4,700. Some of the other social and demographic variables produced more mixed or non-existent results: the percent of the county's voting age population that has a high school degree does not affect the number of signatures gathered, while the unemployment rate has a negative effect that is significant for only two of the eight measures.

In terms of racial characteristics, we find that counties with a greater proportion of blacks yield more signatures, though we find no relationship between the proportion of Hispanic

15

voters and the number of signatures. The first differences indicate that the former has an effect of 4,532 signatures in the pooled model.

Lastly, turning to our political factors, we find effects of both partisanship and competitiveness. Counties that exhibited greater support for the Democratic candidate in the 2002 gubernatorial election produce significantly fewer signatures; the coefficients are all significant except for the last three statutory proposals. Interestingly, the direction of the effect is consistent across initiatives, suggesting that circulators do not seek out ideologically sympathetic voters. Competitiveness produces more signatures — when the margin of victory is greater, significantly fewer signatures are gathered from that county. Overall, partisanship has the third largest largest impact on the number of signatures gathered (almost 6,000) whereas margin's 4,600 signature effect puts it right in the middle. For the four constitutional proposals, the marginal effect of partisanship lies between two-thirds and three-fourths of the average number of signatures per county.

4.2   Signature Validity

We now turn to our analysis of the validity rate of the signatures gathered in each county. We perform joint and separate analyses for each initiative for which we have random sample results as well as for the two initiatives for which we have full validation results. Our dependent variable in these analyses is the number of invalid signatures in each county. Of course, this number is heavily dependent on the number of signatures checked in each county. Fortunately, event count models can accommodate variation in the maximum number of possible events, commonly referred to as exposure, across units. This is accomplished by including the natural logarithm of the number of signatures checked as an independent variable. If the coefficient on this variable is constrained to one, the model is equivalent to estimation of the percent valid in a county. Rather than impose this constraint, however, we follow the recommendations of Maddala (1983) and King (1989b) to estimate a coefficient for this variable.[25]

The results for the seven random samples are presented in Table 5.[26] The results are not as striking as for the number of signatures gathered, but there are some interesting findings.[27] The most consistent variable is unemployment — counties with greater unemployment rates have a significantly higher rate of invalid signatures.[28] This finding obtains in the overall results as well as in six of the seven individual initiatives. The marginal effects reported in the second panel of Table 4 indicate that the first difference effect for unemployment is about twenty signatures. Given that the average number of invalid signatures in the random samples is almost 144, this represents about a 14% change. The percent of a county that is black has a positive and significant effect in two of the initiatives and in the combined model. The marginal effect is about four signatures.

[Insert Table 5 here.]

Less consistent results obtain for population and density. The effect of population on the number of invalid signatures is positive for five of the seven initiatives, but only one of these five coefficients is significant. This is consistent with our expectation based on the incentives for signature gatherers. Yet the effect is negative in both and significant in one of the other cases, which is contrary to our expectation. Population density also produces positive and negative effects, though the two significant cases are positive. This is more consistent with our expectation about the trade-off between signature volume and signature quality. The marginal effects for these two variables are relatively small, however.

Most of the other variables produce no systematic results. Partisanship and is general positive, but never significant whereas margin is mostly negative but never significant. Age is negative and significant in two of the seven initiatives as well as in the pooled results; the marginal effects in these two cases are -16 and -24, respectively. Finally, education is significant and positive in two cases and per capita income is never significant.

The results for the two full checks are presented in Table 6. For the initiative that is included in the previous table, the results are similar in terms of patterns of significance, though age and population swap significance. For the one initiative that we do not have random sample results for, there is mixed evidence for our expectations regarding population

size and density effects — both are significant at the 0.01 level, though population is again found to have a negative effect. Unemployment also has significant effects, though only the 0.10 level. Overall, the full sample tests appear to produce results similar to the random sample tests, though they provide additional evidence for some of the variables that had mixed results.

[Insert Table 6 here.]

4.3   Consequences of Signature Gathering for Outcomes

What is the political relevance of the signature gathering process for ballot outcomes? Besides the obvious necessity of gathering signatures to meet qualification hurdles, the signature gathering process may have consequences for individuals' voting and participation decisions. For example, the Progressive ideal of citizen debate on specific proposals would be for signature gathering to generate broader discussion of the merits of ballot measures. As voters are confronted with petitions to sign they become interested in the topic and seek to become more informed about the proposal and potential alternatives. This may lead to a more informed and politically active electorate, which might be more likely to have well-formed opinions on ballot measures and may be more likely to express those opinions by voting. In fact, there is already evidence that direct legislation states have greater political participation (Boehmke 2002; Tolbert, Grummel and Smith 2001; Tolbert, McNeal and Smith 2003 Smith 2001) and possibly more informed voters (Tolbert, McNeal and Smith 2003).

Given these findings, the existence of direct legislation campaigns may increase voter participation, but also the variation in the intensity of those campaigns may produce variable responses in terms of behavior. Some specific hypotheses can be tested using our data on county-level signature gathering. For example: counties with more intense signature campaigns may have higher turnout; counties with more intense campaigns may have lower ballot roll-off rates (the difference between the number of ballots cast in a election and the number of votes tabulated for each question on the ballot) for initiatives since voters in that area may be more likely to have formed an opinion; more intense campaigns may produce

different aggregate preferences on the measure. In this paper we focus on the second of these by determining whether the intensity of signature gathering in a county influences ballot roll-off.

Our dependent variable for this analysis is the proportion of voters in a county who turned out to vote, but did not cast a vote on each of our eight ballot measures.[29] On average, nine percent of voters who cast a ballot abstain from voting on our each of our eight ballot measures; Proposition 49 has the lowest rate at seven percent whereas Proposition 51 has the highest at eleven percent. The standard deviation in abstention rates across counties is about 3.2%. To test whether the intensity of the signature gathering campaign for that Proposition influences the roll-off rate in a county, we include a variable that measures the intensity of the signature gathering effort in that county, *Signatures Per Capita*. More intense campaigns correspond to a greater number of signatures per capita, leading us to predict a negative relationship between intensity and roll-off.

In addition to our intensity measure, we include a variety of control variables for characteristics including education, median age, income, race and ethnicity, density, partisanship and competitiveness. Because our proposals were on four different ballots and because there may have been features unique to each not captured by our control variables, we also include a series of indicator variables for each proposal. Because our dependent variable is the proportion of voters who abstained on specific measures, we estimate a grouped logit model.[30] The results are presented in Table 7.

[Insert Table 7 here.]

The grouped logit results are strong overall, with an R^2 of 0.58. In addition, they provide evidence in favor of our hypothesis: counties with more signatures per capita have a lower roll-off rate. This provides evidence consistent with direct legislation as a means to increase voter's awareness of issues on upcoming ballots and demonstrates that there is an important relationship between the signature gathering process and political behavior. In addition, it indicates that direct legislation not only increases political activity at the state level, but that variation in direct legislation campaigns within a state generate local variation in individual's

responses.

5   Discussion and Conclusion

This paper provides an empirical investigation into the process of signature gathering for initiatives. For scholars and critics of the direct legislation process, these results provide a first look at the consequences of the strategies of signature gatherers for the representativeness of petition signers. At the aggregate level, our results show a remarkable equity in the geographic distribution of signatures relative to population. Yet we also find pervasive evidence that a host of social and political factors are related to the signature gathering process, suggesting that the answer is more nuanced than suggested by the aggregate data.

It is also important to observe that the pattern of signatures is extremely consistent across all eight petitions that we analyze. Not only do the same factors matter for determining the number of signatures gathered, but they operate in the same direction in every case. This suggests that petition gatherers engage in little targeting across measures, either ideologically or socially. Of course, this type of activity may happen at a more local level than used in our analysis. This suggests that the ability to qualify an initiative may not depend strongly on its issue content, but is almost entirely determined by the ability to pay for sufficient signatures as the common wisdom suggests.

These findings may have implications for concerns about the importance of distribution requirements. In some states these requirements have been ruled unconstitutional since they effectively provide greater influence for voters in smaller counties. Since California has no such requirement and our results indicate no undue reliance on large population centers such as Los Angeles, the effect of distribution requirements can only be distortionary in the sense that they would produce an over-reliance on smaller counties. In addition, our examination of a hypothetical distribution requirement indicates that only the most onerous of requirements would have inhibited our initiatives from qualifying, even ignoring the fact that they were oblivious to our counterfactual experiment.

While our analysis does reveal an equitable distribution of signatures on average, we also find a fair amount of variation in the number of signatures gathered in similarly sized counties, as evidenced by the Gini indices and the results of our regression analyses. Petition managers or signature gatherers appear to place a greater emphasis on counties comprised of potentially disaffected and disinterested voters, rather than find counties with ideological and partisan profiles that are likely to fit well with the specific issue content of each measure.

This conclusion is based on the pattern of relationships found in our regression analyses: counties that produce more signatures are younger, bigger, denser and with a larger black population. These factors correspond to disaffected citizens who are typically less likely to be politically active. These counties may be targeted either because voters may be less trustful of elected officials or, more cynically, they may be more likely to sign petitions without much reflection or objection. In addition, we also find more signatures from counties that vote Republican and are more competitive. As California politics has been dominated by the Democratic party in recent years, signature gatherers may feel that Republican voters are more likely to sign petitions because they fell that the current constellation of control does not sufficiently represent their interests.

Our results for signature validity are less consistent, though there is evidence of higher invalidity rates in denser counties and counties with greater proportions of blacks and higher unemployment rates. The strong and consistent effect of unemployment may be capturing the effect of other, correlated variables (e.g., it may be related to mobility which increases the chances of a signature not matching the address on the official registration rolls) or it may be influencing the quality of signature gatherers that may be contracted to work locally. This may be a case where further analysis and field interviews may help elucidate the exact relationship.

Last, we examined one way in which signature gathering campaigns could influence the eventual election outcome on each of the different ballot measures in our study. We specifically focused on the impact of signature gathering activities on the relative roll-off rates for each ballot measures, and we found that the more signatures gathered in a county, the

lower the eventual roll-off rate. This is just one way in which signature gathering activities might influence ballot measure campaigns and voting behavior, and this should be a subject of future research.

Notes

[1]Arguments on one side suggest that the initiative process is dominated by wealthy business interests (Broder 2000; Ellis 2002; Schrag 1998; Smith 1998); arguments on the other side observe that money is rarely enough to successfully pass an initiative and that the benefits of the process still accrue mainly to broader-based citizen groups (Boehmke 2002, 2003; Donovan, Bowler, McCuan and Fernandez 1998; Ernst 2001; Gerber 1999).

[2]Some distribution requirements are based on other political regions, including state electoral districts or Congressional districts.

[3]An excellent extended discussion of signature gathering and the issues surrounding it is contained in Ellis (2003). Briefer discussions are contained in Broder (2000); Cronin (1989); Ellis (2002); Magelby (1984); and Tolbert, Lowenstein and Donovan (1998).

[4]Oregon first attempted to ban them in 1909; WA, SD and OH banned them about five years later. These bans were stricken down by the Supreme Court ruling in *Meyer v. Grant* in 1988 (Ellis 2003).

[5]In fact, one circulator admitted that people were more likely to sign if he told them it was his birthday (Ellis 2003)

[6]Ed Koupal, originally quoted by *California Journal* and cited in Cronin (1989).

[7]See Lowenstein and Stern (1989) for an extended discussion of this decision.

[8]Other forms of direct legislation in California include the referendum and recall, both of which require a certain number of signatures be gathered. While the process is almost certainly the same, in this paper we focus on the far more common statutory and constitutional measures.

[9]Originally, this requirement was 8%, but was dropped in 1966 due to flagging usage.

[10]A typical distribution requirement exists in Nebraska, where a five percent signature requirement must be met in at least two-fifths of its counties.

[11]In cases where there are less than five hundred total signatures submitted, all are checked.

[12]As a potentially interesting aside, the files we received from the Secretary of State's office indicate that (at least for some petitions) as many as twenty-five counties did not generate their own random numbers for the sampling procedure, rather these numbers were supplied by the Secretary of State's office.

[13]Between 2000 and 2003, there were six initiative constitutional amendments that made it to the ballot; we have data for four of these six measures (we lack data for Propositions 38 and 39 in the 2000 general election). In this same period, there were five initiative statutes that made it to the ballot, of which we have four; we did not receive data on Proposition 50 in the 2002 general election. Proposition 38 in the 2000 general election focused on school vouchers, while Proposition 39 regarded lowering the voting threshold for school bonds to 55%. Proposition 38 failed to pass, receiving only 29.4% of yes votes, while Proposition 39 passed with 53.4% voting yes. Proposition 50 in the 2002 general election concerned water quality, water projects, and wetland protection; this measure passed with 55.3% voting yes. We received this data as part of research we were undertaking for another project, as in the course of communications with officials in the Elections Division we learned that they retained some data on signature checks for some recent ballot measures. We asked them to provide all of the data they had retained. We received a spreadsheet with data

on the eight ballot measures, and only these eight. Expansion of this database, including more data from earlier proposed ballot measures, is the subject of future research.

[14]And even if it did, no states have distribution requirements that require at least one signature from every county (or district in some cases).

[15]If we exclude Los Angeles the correlation is 0.92 and the slope coefficient is 1.06. Obviously a linear regression is not the best model to run; we use it for illustrative purposes only.

[16]We do not have random sample validation results for one initiative, though we do have the full sample results for that case.

[17]This estimate is comparable to the rates obtained by the Ohio Secretary of State's Office in a 1983 study. The petition circulated by professionals had a validity rate of 68% and the two petitions circulated by volunteers had a validity rate of 84% (Lowenstein and Stern 1989).

[18]This average is the same even if we weight by county size.

[19]Firms also do their own internal validity checks to make sure they can meet their obligations.

[20]We performed a t-test for difference in means between the average percent valid for the two samples for this one initiative, producing a t statistic of -6 (n=58).

[21]California Commission on Campaign Financing, quoted in Broder (2000).

[22]Of course, the absence of a distribution requirement in California means that signatures from one county are just as valuable as signatures from any other county. There be little incentive to intentionally falsify signatures because of this, however, there are documented instances where fraudulent signatures have been submitted.

[23]Another possibility is that the variance is less than the mean, known as underdispersion. This case, along with the other two (overdispersion and equidispersion) can simultaneously be estimated using King's generalized event count model (King 1989b). Our data exhibit overdispersion, so we use negative binomial regression with constant overdispersion.

[24]The magnitude of the first differences varies mostly by whether the proposal is a statutory or constitutional proposal. This is not because the variables matter more, it is because more signatures must be gathered for the latter.

[25]The reason for including the natural logarithm follows from the model's properties: if the expected proportion of events is $E[Y]/Y_{max}$, we can write $E[Y]/Y_{max} = exp(X\beta)$, so $E[Y] = exp(X\beta) * Y_{max} = exp(X\beta) * exp(ln(Y_{max})) = exp(X\beta + ln(Y_{max}))$.

[26]The number of observations decreases in a couple of cases because some counties did not have entries for the sample validity results.

[27]In addition to the negative binomial model reported, we also ran a grouped logit analysis, which produced very similar results in term of sign and significance.

[28]An additional variable that we were unable to include would be the proportion of signatures gathered by volunteers versus professionals. Unlike California, Oregon requires circulators to report this figure. The Ohio Secretary of State's study and statements by professional circulators suggest that volunteers have a lower invalidity rate (Ellis 2003; Lowenstein and Stern 1989).

[29]These data are available on the California Secretary of State's website: http://www.ss.ca.gov.

[30]See Greene (1993) or Maddala (1983) for more information on grouped logit (also referred to minimum logit chi-square method).

References

Boehmke, Frederick J. 2002. "The Effect of Direct Democracy on the Size and Diversity of State Interest Group Populations." *The Journal of Politics*: (3) 827-844.

Boehmke, Frederick J. 2003. "Using Auxiliary Data to Estimate Selection Bias Models, With an Application to Interest Groups' Use of the Direct Initiative Process." *Political Analysis* 11(3): 234-254.

Bowler, Shaun, Todd Donovan, and Caroline J. Tolbert, eds. 1998. *Citizens As Legislators: Direct Democracy in the United States.* Columbus: Ohio State University.

Broder, David S. 2000. *Democracy Derailed.* United States: Harcourt Brace.

Cronin, Thomas E. 1989. *Direct Democracy: the Politics of Initiative, Referendum and Recall.* Cambridge: Harvard University Press.

Donovan, Todd, Shaun Bowler, David McCuan and Ken Fernandez. 1998. "Contending Players and Strategies: Opposition Advantages in Initiative Campaigns." In *Citizens As Legislators: Direct Democracy in the United States*, edited by Shaun Bowler, Todd Donovan, and Caroline J. Tolbert. Columbus: Ohio State University Press.

Ellis, Richard. 2002. *Democratic Delusions: The Initiative Process in America.* Lawrence: University of Kansas Press.

Ellis, Richard. 2003. "Signature Gathering in the Initiative Process: How Democratic is it?" *Montana Law Review* 64: 35-97.

Ernst, Howard R. 2001. "The Historical Role of Narrow-Minded Interests." In *Dangerous Democracy: The Battle Over Ballot Initiative in America*, edited by Larry J. Sabato, Howard R. Ernst and Bruce A. Larson. United States of America: Rowman and Littlefield Publishers, Inc.

Gerber, Elisabeth R. 1999. *The Populist Paradox: Interest Group Influence and the Promise of Direct Legislation.* Princeton: Princeton University Press.

Gerber, Elisabeth R. 2001. "The Logic of Reform: Assessing Initiative Reform Strategies." In *Dangerous Democracy: The Battle Over Ballot Initiative in America*, edited by Larry J. Sabato, Howard R. Ernst and Bruce A. Larson. United States of America: Rowman and Littlefield Publishers, Inc.

Goebel, Thomas. 2002. *A Government by the People: Direct Democracy in America, 1890-1940.* Chapel Hill: The University of North Carolina Press.

Greene, William H. 1993. *Econometric Analysis.* Englewood Cliffs: Prentice Hall.

King, Gary. 1989a. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for The Exponential Poisson Regression Model." *American Journal of Political Science* 32: 838-863.

King, Gary. 1989b. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33: 762-784.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(April):341-355.

Lascher, Edward L. Jr; Michael G.Hagen and Steven A. Rochlin. 1996. "Gun Behind the Door? Ballot Initiatives, State Policies and Public Opinion." *Journal of Politics* 58: 760-775.

Lowenstein, Daniel Hays and Robert M. Stern. 1989. "The First Amendment and Paid Initiative Petition Circulators: A Dissenting View and a Proposal." *Hastings Constitutional Law Quarterly* 17: 175-224.

Maddala, G.G. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Magleby, David B. 1984. *Direct Legislation: Voting On Ballot Propositions in the United States.* Baltimore: Johns Hopkins University Press.

Matsusaka, John G. 1995. "Fiscal Effects of the Voter Initiative." *Journal of Political Economy* 103: 587-623.

McCuan, David, Shaun Bowler, Todd Donovan and Ken Fernandez. 1998. "California's Political Warriors: Campaign Professionals and the Initiative Process." In *Citizens as Legislators: Direct Democracy in the United States.* Edited by Shaun Bowler, Todd Donovan and Caroline Tolbert. Columbus: Ohio State University Press.

National Conference of State Legislatures. 2002. "Initiative and Referendum in the 21st Century: Final Report and Recommendations of the NCSL I&R Task Force." Available from http://www.ncsl.org/programs/press/2002/pr020725I&R.htm (downloaded on 08/30/02).

Sabato, Larry J., Howard R. Ernst and Bruce A. Larson, eds. 2001 *Dangerous Democracy: The Battle Over Ballot Initiative in America.* United States of America: Rowman and Littlefield Publishers, Inc.

Schrag, Peter. 1998. *Paradise Lost: California's Experience, America's Future.* New York: The New Press.

Shelley, Kevin. 2002. *A History of California Initiatives.* Report prepared by the California Secretary of State's Office.

Smith, Daniel. 1998. *Tax Crusaders.* NY: Routledge.

Smith, Mark A. 2001. "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout." *American Journal of Political Science* 45(3): 700-706.

Tolbert, Caroline, Daniel H. Lowenstein and Todd Donovan. 1998. "Election Laws and Rules for Initiatives." In *Citizens as Legislators: Direct Democracy in the United States.* Edited by Shaun Bowler, Todd Donovan and Caroline Tolbert. Columbus: Ohio State University Press.

Tolbert, Caroline, John Grummel, and Daniel A. Smith. 2001. "The Effects of Ballot Initiatives on Voter Turnout in the United States." *American Politics Research* 29: 625-48.

Tolbert, Caroline, Ramona McNeal, and Daniel A. Smith. 2003. "Enhancing Civic Engagement: The Effect of Direct Democracy on Political Participation and Knowledge." *State Politics and Policy Quarterly* 3: 23-41.

Tomz, Michael, Jason Wittenberg, and Gary King (2001). CLARIFY: Software for Interpreting and Presenting Statistical Results. Version 2.0 Cambridge, MA: Harvard University, June 1. http://gking.harvard.edu

U.S. Bureau of the Census. Various years. *Statistical Abstract of the United States.* Washington: US Government Printing Office.

Table 1: Descriptive Information on Ballot Measure Petitions

Petition Number	Measure Number	Description	Election	Yes Vote
830	35	Public works projects: Use of private contracts for engineering and architectural services. Initiative constitutional amendment and statute	2000 General	55.2
865	36	Drugs. Probation and treatment program. Initiative statute	2000 General	60.9
874	37	Fees: Vote requirements. Taxes. Initiative constitutional amendment	2000 General	47.9
918	45	Legislative term limits. Local voter petitions. Initiative constitutional amendment	2002 Primary	42.3
935	51	Transportation: Distribution of existing motor vehicle sales and use taxes. Initiative statute	2002 General	42.2
936	52	Election day registration. Voter fraud penalties. Initiative statute	2002 General	40.9
952	49	Before and after school programs. State Grants. Initiative statute	2002 General	56.7
933	54	Classification by race, ethnicity or color, or national origin. Initiative constitutional amendment	2003 Recall	36.1

Table 2: Descriptive Statistics for Number of Signatures and Validity Rates by Petition Number

	All	830	865	874	918	933	935	936	952
Total Signatures Submitted	7,054,936	957,370	713,849	1,089,042	1,062,509	983,761	729,296	734,125	784,984
Total Signatures Required	4,360,304	670,816	419,260	670,816	670,816	670,816	419,260	419,260	419,260
Overqualification	162%	143%	170%	162%	158%	147%	174%	175%	187%
Average Signatures per County	15,237	16,506	12,308	18,777	18,319	16,961	12,574	12,657	13,772
Standard Deviation	36,839	32,904	30,763	47,778	48,072	39,639	29,448	27,211	34,159
Maximum	339,023	193,357	215,179	319,092	339,023	256,372	203,702	182,452	242,187
Minimum	1	10	1	2	1	6	1	1	5
Gini Coefficient (Per Capita Signatures)	0.44	0.38	0.50	0.45	0.49	0.42	0.39	0.43	0.38
Counties Meeting % Threshhold	23	23	28	26	22	30	32	33	34
Average Sample Validity per County	78%		81%	81%	74%	78%	78%	77%	77%
Standard Deviation	10%		7%	6%	12%	7%	14%	13%	5%
Maximum	96%		90%	90%	85%	93%	96%	89%	85%
Minimum	59%		62%	67%	59%	63%	59%	67%	63%
Average Full Sample Validity per County	75%	75%				76%			
Standard Deviation	8%	8%				8%			

N=58. Minima and Maxima for validity rates calculated using counties that had more than ten signatures for that initiative.

Table 3: Negative Binomial Regression Estimates for Total Number of Signatures Gathered per County

	All	830	865	874	918	933	935	936	952
Population Density	2.31**	2.72**	3.22**	3.72**	2.89**	1.59**	1.87**	1.80**	1.81**
	(0.62)	(0.60)	(1.11)	(0.79)	(0.71)	(0.58)	(0.71)	(0.60)	(0.65)
Total population	3.09**	2.52**	3.26**	3.48**	3.33**	3.36**	2.89**	2.82**	3.19**
	(0.32)	(0.35)	(0.34)	(0.37)	(0.32)	(0.39)	(0.33)	(0.32)	(0.33)
Median Age	−105.52*	−82.67	−101.01*	−96.78	−94.04*	−134.03*	−110.75	−137.57**	−139.40**
	(58.40)	(60.46)	(58.94)	(65.32)	(56.32)	(74.78)	(69.94)	(67.79)	(67.76)
Percent Black	0.86**	1.20**	0.87**	1.04**	1.08**	0.84**	0.80**	0.68*	0.64
	(0.34)	(0.33)	(0.40)	(0.38)	(0.35)	(0.36)	(0.37)	(0.37)	(0.39)
Percent Hispanic	0.05	0.12	−0.03	0.08	−0.05	0.14	0.12	0.05	0
	(0.12)	(0.14)	(0.13)	(0.14)	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)
Percent Completed HS	0.09	−0.44	−1.19	−0.41	−2.92	2.08	1.19	1.19	0.38
	(4.14)	(4.24)	(4.19)	(4.36)	(4.15)	(5.05)	(4.77)	(4.64)	(4.63)
Percent Unemployed	−7.83	−11.24*	−8.92	−8.56	−11.56*	−6.92	−9.22	−4.9	−7.78
	(5.81)	(6.49)	(6.08)	(6.34)	(6.93)	(7.19)	(6.96)	(6.84)	(7.27)
Per capita income	4.81**	5.57**	4.69**	5.23**	6.04**	4.67**	4.17**	5.02**	4.61**
	(1.16)	(1.14)	(1.50)	(1.34)	(1.24)	(1.01)	(1.19)	(1.15)	(1.23)
Democrat Vote for Governor	−3.94**	−4.82**	−4.27**	−7.28**	−4.92**	−5.38**	−2.25	−2.03	−2
	(1.34)	(1.26)	(2.07)	(1.62)	(1.38)	(1.26)	(1.63)	(1.46)	(1.46)
Margin in Gov. Race	−2.84**	−2.82**	−3.93**	−5.01**	−3.86**	−3.38**	−1.57	−1.98	−1.61
	(1.22)	(1.18)	(1.79)	(1.42)	(1.49)	(1.02)	(1.20)	(1.25)	(1.40)
Constant	13.73**	13.57**	14.71**	15.30**	16.14**	13.94**	12.14**	12.83**	13.77**
	(1.94)	(2.25)	(2.12)	(2.06)	(2.35)	(2.25)	(2.16)	(2.15)	(2.14)
Dispersion	10427.6	9674.26	7510.97	10039	9384.29	10157.95	7700.19	8353.33	7542
SE	2070.79	2484.8	2185.19	2621.35	2322.76	2687.14	1912.73	1997.78	1855.14
χ^2_{10}	824.3	541.3	1,053.6	895.2	1,288.2	783.1	709.3	691.5	1,000.5
p	0	0	0	0	0	0	0	0	0
Observations	824	541	1054	895	1288	783	709	691	1000

Standard errors in parentheses. * Significance at 10% level, ** at 5% level.

Table 4: Marginal Effects of Independent Variables on Number of Signatures and Number of Invalid Signatures

	All	830	865	874	918	933	935	936	952
Number of Signatures Submitted									
Population Density	4260	5564	4403	7116	5429	3168	2784	2847	2784
Total population	6222	5549	4829	7170	6805	7425	4711	4900	5385
Median Age	-7364	-6276	-5148	-6802	-6571	-10460	-6290	-8533	-8386
Percent Black	4532	7091	3353	5609	5829	4854	3409	3106	2799
Percent Hispanic	1118	2877	-444	1759	-1061	3344	2196	1042	81
High School Education	109	-604	-1091	-515	-3738	2860	1213	1297	402
Percent Unemployed	-3493	-5590	-2921	-3876	-5258	-3359	-3345	-1877	-2905
Per capita income	4684	6014	3341	5171	5987	4950	3276	4227	3751
Dem Vote for Gov.	-5814	-7937	-4628	-11546	-7421	-8849	-2640	-2543	-2425
Margin in Gov. Race	-4638	-5074	-4764	-8570	-6439	-6072	-2059	-2769	-2178
Number of Invalid Signatures									
ln(Sample Size)	284.78		229.62	270.69	276.72	308.20	357.23	290.03	280.25
Population Density	2.14		-2.23	-3.05	-0.02	8.54	5.63	3.63	2.41
Total population	1.53		7.34	6.13	4.02	-4.50	-6.18	1.75	1.14
Median Age	-12.03		-16.24	-8.28	-7.13	-23.58	-15.61	-12.17	-0.02
Percent Black	4.05		3.83	-0.81	16.89	-0.02	-1.82	1.94	11.78
Percent Hispanic	5.40		-5.29	12.31	-12.47	-3.24	11.74	3.28	10.18
High School Education	16.97		14.57	24.53	-10.79	21.33	20.56	16.73	5.11
Percent Unemployed	20.15		22.22	17.42	5.57	28.35	25.21	19.80	19.30
Per capita income	-0.21		3.30	1.82	3.86	-3.91	-1.58	-1.25	10.58
Dem Vote for Gov.	1.69		2.07	4.69	-7.84	4.72	3.25	0.80	-6.38
Margin in Gov. Race	-0.12		-0.38	-4.08	3.99	-1.44	1.81	0.41	-1.98

Marginal effects represent the change in the predicted value of the dependent variables when each independent variable is increased from one standard deviation below its mean to one standard deviation above its mean, holding all other variables fixed at their respective means.

Table 5: Negative Binomial Regression Estimates for Number of Invalid Signatures in Sample Check by County

	All	865	874	918	933	935	936	952
ln(Sample Size)	0.95**	0.94**	0.97**	0.88**	1.02**	1.06**	0.92**	0.96**
	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
Population Density	0.12	−0.16	−0.18	0	0.43**	0.40**	0.22	0.12
	(0.11)	(0.12)	(0.19)	(0.12)	(0.18)	(0.20)	(0.18)	(0.22)
Total population	0.08	0.49**	0.33	0.17	−0.21	−0.41**	0.1	0.05
	(0.09)	(0.12)	(0.22)	(0.17)	(0.15)	(0.15)	(0.14)	(0.15)
Median Age	−18.77*	−31.88**	−12.89	−8.99	−32.33*	−31.47	−20.42	−0.02
	(11.15)	(12.84)	(17.35)	(18.13)	(18.51)	(20.54)	(20.60)	(15.63)
Percent Black	0.08*	0.1	−0.02	0.27**	0	−0.05	0.04	0.21**
	(0.05)	(0.07)	(0.08)	(0.08)	(0.09)	(0.10)	(0.09)	(0.09)
Percent Hispanic	0.03	−0.03	0.06	−0.05	−0.01	0.07	0.02	0.04
	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)
Percent Completed HS	1.43*	1.53*	2.03*	−0.73	1.57	2.28	1.55	0.37
	(0.84)	(0.91)	(1.17)	(1.36)	(1.61)	(1.77)	(1.52)	(1.34)
Percent Unemployed	4.78**	6.66**	4.11**	1.07	5.90**	7.67**	5.09**	3.96**
	(1.07)	(1.54)	(1.78)	(1.92)	(1.76)	(1.90)	(1.64)	(1.20)
Per capita income	−0.02	0.46	0.2	0.34	−0.4	−0.23	−0.16	1.07
	(0.41)	(0.38)	(0.49)	(0.44)	(1.07)	(0.98)	(1.18)	(1.10)
Democrat Vote Share Gov, 2002	0.12	0.19	0.34	−0.46	0.3	0.3	0.06	−0.4
	(0.19)	(0.26)	(0.34)	(0.34)	(0.39)	(0.42)	(0.37)	(0.35)
Absolute % Margin in Gov race	−0.01	−0.03	−0.27	0.21	−0.08	0.15	0.03	−0.11
	(0.19)	(0.23)	(0.27)	(0.22)	(0.31)	(0.32)	(0.23)	(0.28)
Constant	−2.07**	−1.88**	−3.00**	0.11	−2.05**	−3.21**	−1.81**	−1.99*
	(0.75)	(0.75)	(1.02)	(1.07)	(0.98)	(1.13)	(0.88)	(1.04)
Dispersion	3.12	0.94	1.87	1.87	1.89	2.1	1.51	1.7
SE	0.49	0.37	0.61	0.69	0.62	0.69	0.54	0.68
χ^2_{11}	37480	22757	30727	19547	25490	11328	13063	21231
p	0	0	0	0	0	0	0	0
Observations	394	57	58	58	57	54	55	55

Standard errors in parentheses. * Significance at 10% level, ** at 5% level.

Table 6: Negative Binomial Regression Estimates for Number of Invalid Signatures in Full Checks by County

	All	830	933
ln(Number of Signatures)	$1.07**$	$1.06**$	$1.09**$
	(0.02)	(0.03)	(0.02)
Population Density	$0.52**$	$0.66**$	$0.33*$
	(0.09)	(0.14)	(0.19)
Total population	$-0.38**$	$-0.42**$	$-0.32*$
	(0.09)	(0.09)	(0.17)
Median Age	-7.02	-2.21	-21.89
	(17.32)	(23.60)	(19.33)
Percent Black	0.08	0.12	0.01
	(0.06)	(0.08)	(0.06)
Percent Hispanic	0	0	0
	(0.03)	(0.05)	(0.03)
Percent Completed HS	-0.23	-0.43	0.4
	(1.16)	(1.54)	(1.42)
Percent Unemployed	$4.16**$	$5.01**$	$3.84**$
	(1.41)	(1.97)	(1.69)
Per capita income	0.12	0	0.24
	(0.26)	(0.28)	(0.31)
Democrat Vote for Governor	0.29	0.4	0.24
	(0.21)	(0.29)	(0.25)
Margin in Gov. Race	0.18	-0.02	0.35
	(0.21)	(0.24)	(0.30)
Constant	$-2.13**$	$-2.11*$	$-2.17**$
	(0.78)	(1.18)	(0.97)
Dispersion	36.65	32.12	29.27
SE	7.43	7.77	9.41
χ^2_{11}	62821	33765	38970
p	0	0	0
Observations	115	58	57

Standard errors in parentheses. * Significance at 10% level, ** at 5% level.

Table 7: Grouped Logit Regression Estimates for Ballot Roll-Off by County

Signatures per Capita	$-0.26**$
	(0.11)
Percent Unemployed	-2.12
	(1.62)
Percent Completed HS	0.80
	(1.04)
Percent Hispanic	0.01
	(0.03)
Percent Black	0.02
	(0.05)
Per capita income	-0.28
	(0.36)
Population Density	$0.28**$
	(0.10)
Median Age	$-40.38**$
	(11.46)
Democrat Vote for Governor	$1.44**$
	(0.21)
Margin in Governor's Race	$-0.49**$
	(0.18)
Petition 865	$0.25**$
	(0.04)
Petition 874	$0.18**$
	(0.04)
Petition 918	$-0.15**$
	(0.05)
Petition 933	$-0.26**$
	(0.04)
Petition 935	$0.21**$
	(0.04)
Petition 936	$-0.14**$
	(0.04)
Petition 952	$-0.23**$
	(0.04)
Constant	$-1.68**$
	(0.71)
Observations	464
R^2	0.58

Standard errors in parentheses. * Significance at 10% level, ** at 5% level.