THE RATIONAL VOTER IN AN AGE OF RED AND BLUE STATES: THE EFFECT OF PERCEIVED CLOSENESS ON TURNOUT IN THE 2004 PRESIDENTIAL ELECTION

A Thesis
submitted to the
Graduate School of Arts and Sciences
at Georgetown University
in partial fulfillment of the requirements for the degree of Master of Public Policy in the Georgetown Public Policy Institute

By

Scott Evan Epstein, B.A.

Washington, D.C.
April 10, 2007
THE RATIONAL VOTER IN AN AGE OF RED AND BLUE STATES: THE EFFECT OF PERCEIVED CLOSENESS ON TURNOUT IN THE 2004 PRESIDENTIAL ELECTION

Scott Evan Epstein, B.A.

Thesis Advisor: Michael J. Puma, M.S.

ABSTRACT

This paper studies the effect of perceived closeness on voter turnout in the 2004 presidential election. Previous research has not examined the closeness-turnout relationship in recent presidential elections, and has suffered from a lack of statistical power. In this paper, turnout is examined at the county level to increase statistical power, and campaign efforts by candidates, parties, and interest groups are controlled for in the model. A measure of perceived closeness is constructed using polls published in the media. The results provide support for the hypothesis that perceived closeness affected voter turnout in the 2004 presidential election, independent of the effect of campaign activities. This paper suggests that citizens were more likely to vote in competitive states for two reasons: (1) Citizens perceived that their votes were more likely to be decisive in these states, and (2) Citizens in competitive states were exposed to a greater number of campaign advertisements.
# TABLE OF CONTENTS

Introduction................................................................................................................1

Literature Review.........................................................................................................4

The Decline in Voter Turnout.......................................................................................4

Individual versus Systemic Determinants of Voter Turnout.....................................5

Individual Determinants............................................................................................5

The Closeness-Turnout Relationship: Rational Voting or Elite Mobilization?.........6

The Closeness-Turnout Relationship in U.S. Presidential Elections.........................9

Measuring Closeness..................................................................................................11

Summary of the Literature/Research Gaps.................................................................12

Research Methods......................................................................................................13

Data Sources..............................................................................................................14

Analysis Methods......................................................................................................19

Results.......................................................................................................................22

Policy Implications....................................................................................................29

References..................................................................................................................31
LIST OF TABLES

Table 1: Descriptive Statistics for County Level Data……………………………………..22
Table 2: Descriptive Statistics for State Level Data………………………………………22
Table 3: Turnout Models 1-4 Including All Variables and Significance Levels…………24
Table 4: Models 1-4 with Variables of Interest, Standard Errors, and P Values………..25
INTRODUCTION

In his classic book, *An Economic Theory of Democracy*, Anthony Downs (1957) claims that citizens’ choices about whether to vote are based on an informal calculation of the cost and benefits of voting. According to Downs, the decision on whether to vote will take the following basic form:

\[ V = P(D) \times B - C \]

Where \( V \) = the individual’s decision whether to vote, \( P(D) \) = the probability that an individual’s vote will decide the election, \( B \) = the benefit the individual derives from the desired election outcome, and \( C \) = the costs associated with voting. That is, an individual’s decision to vote is a function of the probability of casting a decisive vote multiplied by the net benefit of the desired election outcome. According to Downs’ theory, in close elections voters will have a higher probability of casting a decisive vote and therefore will be more likely to vote. Matsusaka and Palda (1993) call this the Downsian Closeness Hypothesis (DCH).

The Electoral College provides an interesting test of the DCH. Under the Electoral College system, the fifty states and the District of Columbia each hold separate presidential elections, and with the exception of Maine and Nebraska, each state’s Electoral College votes are awarded to the candidate who wins the plurality of the votes statewide (i.e., has the highest percentage of the vote). Because the states are
different in countless ways, there is considerable variation across states in the closeness of any particular election. The DCH predicts that, holding other factors constant, in states with closer contests, turnout will be higher.

A large body of research has examined the effect of perceived closeness on voter turnout, and the general conclusion is that turnout is moderately higher in close elections (Matsusaka and Palda 1993). It is possible, however, that the closeness effect may have been stronger in recent U.S. presidential elections as voters have become more aware of how closeness affects their probability of casting a decisive vote. First, the number of swing states has greatly decreased in the past 45 years. According to the “The Shrinking Battleground” by the Center for Voting and Democracy, there were 23 swing states in 1960, compared to only 13 states in 2004 (FairVote). In 1960, when almost half of the states were very competitive, voters may not have been particularly aware of the mechanics of the Electoral College system, and thus how closeness varies across states. More recently, as presidential races were fought in a smaller number of competitive states, the relationship between closeness and the value of their votes may have been clearer to voters. In particular, the controversial 2000 election may have made many Americans aware that it matters where they vote, i.e., many voters must have observed that votes in Florida were particularly meaningful. Finally, the media’s recent parsing of the country into “red” and “blue” states may have further focused voters’ attention on the Electoral College and the relative closeness of contests in different states.
There are a number of reasons why policymakers should be concerned with variation in voter turnout. First, high turnout increases the legitimacy of government. The higher the percentage of eligible voters that turn out, the greater is the likelihood that citizens will view the government’s power as legitimate. On the other hand, if turnout is low in certain states because people believe their votes are not important, the President’s legitimacy may be diminished.

Differences in turnout may also have policy implications. Robert Fleck (1999) and Paul Martin (2003) both find that turnout affects the distribution of federal resources. According to Fleck, counties with higher turnout received higher levels of funding from the Federal Emergency Relief Administration during the New Deal, and Martin has demonstrated that congressmen steer federal resources toward areas of their districts that have higher turnout.

Finally, differences in voter turnout may have harmful effects on the voters themselves. Finkel (1985) shows that voting leads to higher levels of political efficacy, i.e., citizens who vote have more faith in government, and more confidence that they have the ability to understand and participate in political affairs. Not surprisingly, Finkel also finds evidence that the causal relationship works in the opposite direction – higher political efficacy leads to higher participation. This evidence suggests that nonvoting may cause a downward spiral in political participation, as nonvoting leads to low political efficacy and low political efficacy leads in turn to even less participation.
LITERATURE REVIEW

The Decline in Voter Turnout

A focus of much of the research on voter turnout has been an examination of the dramatic decline since 1960. In 1960, national turnout (defined as the ratio of total votes to Voting Age Population) was 62.77 percent (Election Assistance Commission). Twenty years later, in 1980, national turnout had decreased by more than 10 percentage points to 52.56 percent. Brody (1978) notes that turnout decline is something of a puzzle because a number of demographic trends post-1960, including higher average levels of education and fewer barriers to voter registration, would have predicted an increase in turnout. Using pooled data from the American National Election Studies survey, Teixeira (1987) finds that some of the decline in turnout can be attributed to demographic changes during the time period, particularly the larger number of youths in the electorate, and the greater proportion of unmarried and highly mobile voters. Teixeira calculates that together these three demographic shifts explain 38 percent of the decline in turnout from 1960 to 1980. He attributes the remaining 62 percent of the turnout decline to decreases in newspaper readership, party identification, and political efficacy (the degree to which an individual has faith in government and believes she can understand and influence political affairs). Teixeira notes that Abramson and Aldrich (1982) found remarkably similar results; in their study, partisanship and political efficacy explained roughly two thirds of the decline in turnout from 1960 to 1980.
Individual versus Systemic Determinants of Voter Turnout

According to Leighley and Nagler (1992), the determinants of voter turnout can be divided into two major categories: (1) individual determinants and (2) systemic determinants. Individual determinants are qualities of the individual citizen that may affect her decision about whether to vote, such as age, income, and educational attainment. Systemic determinants are characteristics of the political environment that may affect turnout, such as the strictness of registration requirements or the presence of get out the vote operations. This paper is primarily concerned with two systemic determinants of turnout: voter’s perceptions of the closeness of elections and campaign activities.

Individual Determinants

Leighley and Nagler (1992) find that the effects of individual determinants, such as education, age, and income, have been fairly consistent over time. Using the Current Population Survey, they examine the determinants of voter turnout in the 1984 election, utilizing the same multivariate probit model employed by Wolfinger and Rosenstone (1980) for their classic book, *Who Votes?*, which examined the determinants of voting in the 1972 presidential election. In general, Leighley and Nagler find that their estimates are quite similar to Wolfinger and Rosenstone’s, showing that education and income are the most powerful predictors of voter turnout, i.e., individuals with higher levels of education and income are considerably more likely to vote, as are older
citizens. Although blacks were more likely to vote in both years than non-blacks, the effect was much greater in 1984 than in 1972. Leighley and Nagler also find that being married and being a union member both increased the probability of voting.

In addition to these demographic factors, other researchers have found that attitudes toward politics and party identification can affect voter turnout. For example, Teixeira (1987) finds that holding demographic factors constant, citizens with higher levels of political efficacy, newspaper readership, and attachment to a political party all vote in higher numbers.

**The Closeness-Turnout Relationship: Rational Voting or Elite Mobilization?**

This paper focuses primarily on systemic determinants of voter turnout, particularly the effect of election closeness. There is a rich literature on the closeness-turnout relationship. Inspired by the theoretical work of Downs (1957) and Riker and Ordeshook (1968), social scientists have produced a steady stream of empirical work on the topic since the early 1970s. In a meta-analysis of the existing research in 1992, Matsusaka and Palda (1993) identified 25 different studies that were performed starting in 1973, including studies in different countries (e.g., France, the United Kingdom, Canada, Israel, and the United States), and studies of different types of elections (federal, state, and local elections, and ballot referendums). Research on this topic continues to be produced regularly (Berch 1993; Hanks and Grofman 1998; Grofman,
As Cox and Munger (1989) explain, closeness may theoretically affect turnout in two different ways. One possibility is that closeness may encourage citizens to vote because they believe their votes have a high probability of deciding the election (the DCH theory). An alternative possibility is that closeness may inspire elites (e.g., candidates, parties, political action committees, etc.) to buy more television advertisements and knock on more doors, and in the process convince more voters to go to the polls. If some variable measuring campaign activity is not included in the model, it is impossible to tell whether the DCH or elite mobilization is responsible for the closeness-turnout relationship. The best studies on the closeness-turnout relationship therefore control for campaign activities in some way. For example, in their study of the 1982 U.S. House elections, Cox and Munger address this problem by including campaign expenditures in their model. They find that even when controlling for campaign expenditures, the coefficient on their measure of closeness still has the expected sign and is statistically significant, although it is small in magnitude. They conclude that closeness does inspire elite mobilization but that voters’ decisions are also affected by closeness directly.

Matsusaka and Palda (1993) argue that the moderate effect cited by Cox and Munger (1989) and others may not actually be due to the DCH. They contend that the closeness effect that remains after controlling for campaign expenditures may be
attributable to unobservable elite mobilization such as volunteer labor, or a spurious correlation due to aggregation bias. To investigate these possibilities, Matsusaka and Palda study closeness and turnout in the 1979 and 1980 Canadian legislative elections using two different data sources: aggregate election results (the type of data typically used in the literature) and survey data. Using aggregate election results data (macro regressions) and controlling for campaign expenditures, they find statistically significant support for the DCH. However, when they test the same elections using survey data for individual voters (micro regressions), the coefficient on closeness is in the predicted direction but is no longer statistically significant. Matsusaka and Palda favor the micro regressions because they provide a direct test of the voting decision, and are not subject to aggregation bias. They therefore conclude that previous support for the DCH may have been a result of spurious correlation or omitted variable bias.

Matsusaka and Palda make an important observation that closeness effects observed in the literature may be due to unobserved elite mobilization rather than the DCH. But their argument for the superiority of their micro regression results is unconvincing. As they readily admit, survey data on elections is subject to over-reporting. If the false voters are different on average from the actual voters, as research has suggested they are (Silver, Anderson, and Abramson 1986), bias may be introduced into the estimates of the closeness effect. It seems just as reasonable then that Matsusaka and Palda’s micro regressions may have been biased rather than their macro regressions.
On net, the literature supports the hypothesis that citizens are more likely to vote if they believe the election is close. A majority of the literature has found that turnout is moderately higher in closer elections, providing support for the DCH. In their meta-analysis, Matsusaka and Palda identify 43 independent estimates of the relationship between closeness and turnout. Out of these tests, 35 (81.4%) supported the DCH and 30 (69.8%) of these estimates were statistically significant. The most convincing studies, such as Cox and Munger (1989), control for mobilization by elites.

The Closeness-Turnout Relationship in U.S. Presidential Elections

The evidence on the effect of closeness on turnout in presidential elections has been mixed; some studies have provided support for the DCH while others have not. For example, Kau and Rubin (1976), in their study of state-level turnout in the 1964 presidential election, do not find evidence that supports the DCH (i.e., the coefficient on their measure of closeness is of the wrong sign and not statistically significant). There is reason to believe, however, that Kau and Rubin’s estimate is neither efficient nor reliable given the small sample size (50 state observations in a single year), and the failure to control for demographic variables which may have biased the coefficient on closeness.

Alternatively, Settle and Abrams (1976) compare national turnout across years examining the presidential elections from 1868 to 1972, controlling for total national campaign expenditures by major party candidates (adjusted to 1957 dollars). Their
estimate for the effect of closeness on turnout is negative and statistically significant, supporting the DCH. This study suggests that nationwide turnout is higher in presidential elections that are competitive on the national scale. However, Settle and Abrams’ findings do not answer the question of how variation in the closeness of statewide elections affects turnout across states.

Crain and Deaton (1977) provide more convincing evidence in support of the DCH, focusing on the effect of income on turnout in the 1972 presidential election, but including a measure of closeness as a control variable. Their results support the DCH, finding that closer elections are associated with higher turnout. The Crain-Deaton model is superior to the Kau-Rubin because it includes a number of important demographic controls such as race, sex, income, and age and therefore probably provides a more reliable estimate of the relationship between closeness and voter turnout. Foster (1984), however, sheds doubt on Crain and Deaton’s (1977) findings. Foster uses the Crain-Deaton model to test whether there was a closeness effect in the 1968, 1972, 1976, and 1980 presidential elections. Using a pooled sample of 200 observations (50 states in 4 elections), he finds that the coefficient on the closeness effect is in the predicted direction but is not statistically significant. Given that other research has found that the closeness effect is small, it is possible that the sample size is not large enough to pick up the effect.

In sum, there is not overwhelming evidence in favor of the DCH for U.S. presidential elections. Models that control for demographics seem to perform better,
but are still somewhat inconsistent. One problem in the literature is low statistical power because of the use of largely state-level data. With a higher number of observations, it may be possible to detect smaller closeness-turnout effects. Another problem is the lack of research on recent presidential elections. The existing research has focused on presidential elections in 1980 and earlier, and the effect of closeness on turnout may have changed over time.

**Measuring Closeness**

One major issue in the literature is how to measure perceived closeness. The classic approach was to construct closeness variables by using *ex post* election results data, typically either the percentage of the vote that the winner received (Barzel and Silberberg 1973; Kau and Rubin 1976) or the percentage margin (Matsusaka 1993). The assumption was that these *ex post* measures would be reasonable proxies for voters’ *ex ante* perceptions of closeness. Cox (1988) argues that regressing turnout on the percentage margin may produce a spurious correlation because total votes appears in both the numerator of turnout and the denominator of the percentage margin. Because of this concern, Cox (1989) uses the raw margin (the difference in votes between the winner and the loser) for his study of the 1982 U.S. House elections. This measure, however, can only be employed when comparing jurisdictions with roughly similar populations.
Although widely used in the literature, *ex post* measures of closeness are problematic because they may not always accurately reflect voters’ perceptions of closeness. Some researchers have attempted to solve this problem by constructing *ex ante* measures of perceived closeness using data that would be available to voters before an election. For example, Matsusaka and Palda (1993) use an *ex ante* model in their study of Canadian legislative elections, arguing that one source of information that voters could use to estimate closeness is the margin from the previous election. Matsusaka and Palda therefore use the margin of the 1979 election as a measure of perceived closeness for the 1980 election. Kunce (1999) uses an alternative *ex ante* measure of perceived closeness based on the results of polls released by newspapers within the last week before elections. Intuitively, the Kunce (1999) *ex ante* measure is more appealing than the Matsusaka-Palda (1993) measure. It does not seem likely that the average voter would be aware of the closeness of the previous election; in contrast, it is easy to imagine that voters would be exposed to polling released by the media, and that this polling might shape their perceptions of closeness.

**Summary of the Literature/Research Gaps**

In general, the literature has shown that in closer elections, turnout is moderately higher, although the evidence has not been universal. Controlling for elite mobilization seems to decrease the magnitude of the closeness effect but not eliminate it (Cox and Munger 1989). Despite general support for the DCH, there are significant gaps in the
literature. In particular, no published research on the effect of closeness on turnout in presidential elections post-1980 could be found. This is a striking omission, given that the downward trend in the number of swing states may have focused voters’ attention on how closeness varies across states. In addition, the literature on presidential elections has focused almost exclusively on *ex post* rather than *ex ante* measures of closeness, and has suffered from low statistical power because of the use of state-level data.

**RESEARCH METHODS**

Focusing on the 2004 presidential election, this paper investigates whether voters’ perceptions of how close presidential elections are in their states affect their decisions to vote controlling for other factors which are expected to affect voter turnout. Specifically, the hypothesis is that voters in states where the presidential contest is close (“swing states”) will turn out in higher numbers than voters in states where the presidential contest is not close (“spectator states”) because they believe their votes have a higher probability of deciding the election.

Previous research has examined the effect of perceived closeness on voter turnout, but has not focused on recent presidential elections. It is possible that the relationship between perceived closeness and voter turnout has changed over time, particularly as the number of swing states has greatly diminished, and the concept of
“red” and “blue” states has saturated our culture. In addition, Election 2000 has probably made more voters aware of the differing power of their votes. As a consequence, this paper adds to the research by examining the most recent presidential election.

This analysis also uses a novel approach to measuring closeness that appears to have been used in one previous study (Kunce 2001), and has never been used for a study on presidential elections. The perceived closeness variable is constructed using polling data from before the election rather the final margin itself. As discussed above, using the final result of an election is problematic given that some election results are surprises, i.e., sometimes the margin is much closer than is perceived before the election, and sometimes candidates win by a much larger margin than was expected.

Finally, the problem of low statistical power found in the literature is mitigated by examining turnout at the county level, and controlling for county-specific demographic factors. To my knowledge, this is the first study to use such an approach.

**Data Sources**

This paper employs data measured at two levels of aggregation: (1) data for 49 states (Alaska is excluded) and the District of Columbia; and, (2) data for 3,110 counties clustered within these states. The primary explanatory variable of interest, perceived closeness, is calculated at the state level based on state level polling data. However, to increase the sample size for higher statistical power, voter turnout is examined at the county level. Demographic controls are also included at the county
level.

The dependent variable, voter turnout, was constructed using two data sources. The denominator is an estimate of the Voting Age Population (VAP) by county, and comes from the U.S. Census Bureau’s Population Estimates Program (U.S. Census Bureau Population Division). VAP includes all residents 18 or older, including those ineligible to vote such as non-citizens and convicted felons. Voting Eligible Population (VEP) excludes these populations and is therefore a superior measure. Unfortunately, data for VEP was not available. However, VAP is used extensively in the literature and it is expected to be a reasonable proxy for the VEP. The numerator for the turnout variable is the total presidential vote by county compiled from local election results data by Dave Leip (Leip 2005).

The primary explanatory variable of interest, perceived closeness, is constructed from state polling data obtained from two websites: Dave Leip’s Atlas of U.S. Presidential Elections and RealClearPolitics (Dave Leip’s Atlas; RealClearPolitics). Seventy percent of the polls were listed on both websites, 25 percent were listed just on RealClearPolitics, and 5 percent were listed just on Dave Leip’s Atlas of U.S. Presidential Elections. Given that a large majority of the polls were listed by both websites, it seems reasonable to assume that together the two websites capture most of the polling that could have influenced voters. Two different specifications of the perceived closeness variable are used in the analysis. The first is an average of all polls from the beginning of October to Election Day; if there were no polls from October or
November, the most recent poll was used (all states had at least one poll in September or October). The second specification of the perceived closeness variable was created by averaging all polls after October 15\textsuperscript{th}; again, if no polls were taken post-October 15\textsuperscript{th}, the most recent poll was used.

The state polls were conducted by many different firms, and employed different sample sizes and methodologies, e.g., some were of registered voters and some were restricted to self-reported “likely voters.” If the goal was to compare the merits of polls, these differences would be problematic. For this analysis, however, the polls are used as a proxy for citizens’ perception about the closeness of the presidential race in their states. The assumption here is that the average citizen does not pay close attention to the nuances of sample sizes and polling methodology. Rather, they simply absorb the numbers through the media, and make somewhat unsophisticated judgments about closeness based on the reported results.

The primary control variable, campaign advertisements, was compiled from data available on CNN.com’s America Votes website (Campaign Media Analysis Group). The Center for Voting and Democracy used these data to provide descriptive analysis for its report, “Who Picks the President?” (FairVote). The specifications of campaign advertisement were created matching the time frame of the polling variables described above. The number of ads was used instead of total ad expenditures as a measure of variation in campaigning across states. This specification has the benefit of avoiding variation in media costs across states; however, the variable is limited in that it does not
distinguish ads by media market within the state or by ratings. This measure is, therefore, an approximation of the average voter’s exposure to campaign ads.

Other average county demographic variables included in the model (percent VAP 18 to 24 years of age, percent VAP 45 or older, percent VAP male, percent black, and percent Hispanic) come from the Census Bureau’s Populations Estimates Program (U.S. Census Bureau Population Division). Additional county-level covariates including percent of the population with a college degree or higher, percent of the population below the poverty line, percent of the population currently married, and percent of the population living in an urban area are derived from the Decennial Census (U.S. Census Bureau American FactFinder). Although 2004 demographic data are preferable, such information is not available. It is assumed that the 2000 data are a reasonable proxy for county-level characteristics in 2004.

Finally, the analysis includes a number of state and county level political variables, including the registration closing date (measured in the model as the number of days between the last day to register in the state and election day), and dummies for senatorial and gubernatorial races, marriage amendments, and competitive house elections. The dummy variables for state and local races are included because it is possible that these races could be more important to certain voters than the presidential election. Therefore, it is expected that turnout will be higher in states with senate and gubernatorial races, and marriage amendments on the ballot, and in counties where there is a competitive house election. The state closing dates were obtained from The
Book of the States 2005 (Council of State Governments 2005). The data on senatorial, gubernatorial, and marriage amendments were readily available in the public sphere (CNN.com; National Conference of State Legislatures). The competitive house races are the 58 races that the Cook Political Report was tracking immediately before the election (Cook Political Report). Each county that was at least partially within one of these 58 congressional districts received a 1 on the competitive house race dummy variable; otherwise it was coded as 0.

The data set consists of 3,110 counties with no missing data. Alaska was excluded from this analysis because it has no counties, as was the tiny Kalawao County in Hawaii, which had a population of 111 people in 2004, and for which no election results could be found. Three additional counties were also excluded because of their change in status from 2000 to 2004: Alleghany County, VA, Broomfield County, CO and Boulder County, CO. Each of these three counties changed their borders from 2000 to 2004. Alleghany County subsumed the formerly independent city of Clifton Forge in 2001. The city of Broomfield in Colorado split from Boulder County in 2001 and became its own county. Because both 2000 Decennial Census data and 2004 population estimates data are used in the analysis, including any of these three counties would have introduced bias into the results. The District of Columbia is included as a state with 1 county.
Analysis Methods

Data for this analysis are at both the state and county level. I use an OLS regression model with robust standard errors to correct for the clustering of counties within states. The model takes the following basic form:

$$Turnout_c = \beta_0 + \beta_1 Closeness_s + \beta_2 ElecVotes_s + \beta_3 Closeness_s \times ElecVotes_s + \beta_4 Campaign_s + \beta_5 Polit_{s,c} + \beta_6 Demo_c$$

The dependent variable is turnout at the county level. Turnout is defined as the ratio of total votes to Voting Age Population and is measured in percentage points. The primary variables of interest are Closeness, ElecVotes, and the interaction term Closeness\times ElecVotes. All three variables of interest are at the state level of aggregation. Closeness is the measure of perceived closeness. Three specifications of the variable are used in different models. The first specification, Polls Margin 1, is the average of the margins of all statewide polls from October 1st to Election Day; if no polls were taken in October or November, then the most recent poll was used. The second specification, Polls Margin 2, is the average of the margins of all statewide polls taken after October 15th; if no polls were taken after October 15th, the most recent poll was used. The third specification, Actual Margin, is the ex post actual margin of the election. It is anticipated that the sign on $\beta_1$ will be negative, showing that as voters perceive the race to be closer in their state, they will be more likely to vote. ElecVotes represents the number of electoral votes in the state. Theoretically, ElecVotes could be related to turnout in either of two ways. If citizens are basing their decision about
whether to vote on the probability that their state’s electoral votes will be decisive in the Electoral College tally, they will be more likely to vote in states with more electoral votes (larger states). If instead citizens are basing their decision about whether to vote on how decisive their individual votes will be in the statewide tally, they will be more likely to vote in smaller states (with fewer electoral votes). In the first case, the sign of $\_2$ would be positive. In the second case, the sign of $\_2$ would be negative. It is also possible that the effect of closeness will depend on the size of the state. For this reason, the interaction term $Closeness \times ElecVotes$ is included in the model.

$Campaign$ represents the primary control variable: total campaign ads, which consists of ads by the Bush and Kerry campaigns, the two parties, and interest groups. The variable is measured in thousands of ads for ease of interpretation. $Campaign$ is calculated at the state level, and has two specifications. The first specification, $Campaign Ads 1$, is the total number of campaign advertisements in the state from September 25th to Election Day. The second specification, $Campaign Ads 2$, is the total number of campaign advertisements in the state from October 17th to Election Day.

$Polit$ represents a vector of political and systemic variables, including 5 variables at the state level and 1 variable at the county level. The 5 state level variables are $Closing Date$, $South$, $Senate Race$, $Gubernatorial Race$, and $Marriage Measure$. $Closing Date$ is a measure of the ease of registering to vote in the state. It is defined as the number of days between the last day to register and Election Day. $South$ is a dummy variable for southern states; the eleven states of the confederacy are each
assigned a value of 1. *Senate Race, Gubernatorial Race,* and *Marriage Measure* are dummy variables for different types of statewide elections. A state is assigned a value of 1 for *Senate Race* if there was a race for the U.S. Senate, a value of 1 for *Gubernatorial Race* if there was a race for governor in the state, and a value of 1 for *Marriage Measure* if there was a ballot measure to ban gay marriage in the state. The county level variable is a dummy variable for competitive house races; a county is assigned a value of 1 if at least part of the county is located in a congressional district with a competitive race as defined by the Cook Political Report.

Finally, *Demo* represents a vector of 9 demographic variables measured at the county level. All of the variables are measured in percentage points. Three variables use Voting Age Population as the denominator: *Voting Age Population Male, Voting Age Population 18-24,* and *Voting Age Population 45 or Older.* Four variables use the total population of the county as the denominator: *Black, Hispanic, Urban* (the proportion of county residents living in urban areas), and *Poverty* (the proportion of county residents living under the federal poverty line). *Currently Married* is defined as the proportion of county residents age 15 or older that are currently married. Finally, *College* is defined as the proportion of county residents age 25 or older with at least a college degree.

Below are descriptive statistics for the variables employed in the model. County level and state level variables are listed separately.
Table 1: Descriptive Statistics for County Level Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting Age Population Male</td>
<td>49.14</td>
<td>40.53</td>
<td>69.88</td>
<td>3110</td>
</tr>
<tr>
<td>Voting Age Population 18-24</td>
<td>12.94</td>
<td>2.44</td>
<td>53.67</td>
<td>3110</td>
</tr>
<tr>
<td>Voting Age Population 45 or Older</td>
<td>52.80</td>
<td>12.72</td>
<td>78.05</td>
<td>3110</td>
</tr>
<tr>
<td>Turnout</td>
<td>57.97</td>
<td>15.29</td>
<td>100.00</td>
<td>3110</td>
</tr>
<tr>
<td>Black</td>
<td>9.30</td>
<td>0.00</td>
<td>86.18</td>
<td>3110</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6.95</td>
<td>.13</td>
<td>97.44</td>
<td>3110</td>
</tr>
<tr>
<td>Urban</td>
<td>40.18</td>
<td>0.00</td>
<td>100.00</td>
<td>3110</td>
</tr>
<tr>
<td>Currently Married</td>
<td>60.38</td>
<td>30.92</td>
<td>87.59</td>
<td>3110</td>
</tr>
<tr>
<td>College</td>
<td>16.51</td>
<td>4.92</td>
<td>63.75</td>
<td>3110</td>
</tr>
<tr>
<td>Poverty</td>
<td>14.18</td>
<td>0.00</td>
<td>56.92</td>
<td>3110</td>
</tr>
<tr>
<td>Competitive House Race</td>
<td>.22</td>
<td>0.00</td>
<td>1.00</td>
<td>3110</td>
</tr>
</tbody>
</table>

Table 2: Descriptive Statistics for State Level Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>Minimum</th>
<th>Maximum</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senate Race</td>
<td>0.66</td>
<td>0.00</td>
<td>1.00</td>
<td>50</td>
</tr>
<tr>
<td>Gubernatorial Race</td>
<td>0.22</td>
<td>0.00</td>
<td>1.00</td>
<td>50</td>
</tr>
<tr>
<td>Marriage Measure</td>
<td>0.22</td>
<td>0.00</td>
<td>1.00</td>
<td>50</td>
</tr>
<tr>
<td>South</td>
<td>0.22</td>
<td>0.00</td>
<td>1.00</td>
<td>50</td>
</tr>
<tr>
<td>Closing Date</td>
<td>21.46</td>
<td>0.00</td>
<td>31.00</td>
<td>50</td>
</tr>
<tr>
<td>Campaign Ads 1 (Post September 25th)</td>
<td>4956.06</td>
<td>0.00</td>
<td>55487.00</td>
<td>50</td>
</tr>
<tr>
<td>Campaign Ads 2 (Post October 16th)</td>
<td>2453.86</td>
<td>0.00</td>
<td>28817.00</td>
<td>50</td>
</tr>
<tr>
<td>Polls Margin 1 (Polls Post September 30th)</td>
<td>13.54</td>
<td>0.85</td>
<td>67.00</td>
<td>50</td>
</tr>
<tr>
<td>Polls Margin 2 (Polls Post October 15th)</td>
<td>13.57</td>
<td>0.85</td>
<td>67.00</td>
<td>50</td>
</tr>
<tr>
<td>Actual Margin</td>
<td>15.86</td>
<td>0.38</td>
<td>79.84</td>
<td>50</td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>10.70</td>
<td>3.00</td>
<td>55.00</td>
<td>50</td>
</tr>
</tbody>
</table>

RESULTS

The four models presented in Tables 3 and 4 vary in two ways: (1) How closeness is specified, and (2) Whether the interaction effect between closeness and electoral votes is included in the model. Model 1 uses Polls Margin 1 (the average of
all polls post September 30th) and includes the interaction effect. Model 2 uses Polls Margin 1, but does not include the interaction effect. Model 3 uses Actual Margin and includes the interaction effect. Finally, Model 4 uses Actual Margin, but does not include the interaction effect. Table 3 presents the coefficients on all variables included in the models, as well as significance levels. Table 4 only presents data on the variables of interest, but includes robust standard errors and p values as well as coefficients.

Models employing the shorter time frame specifications of the campaign ads variable (Campaign Ads 2) and the polls margin variable (Polls Margin 2) were estimated but are not presented here. Campaign Ads 1 and Campaign Ads 2 were very highly correlated, as were Polls Margin 1 and Polls Margin 2; consequently, the results for the models not presented were remarkably similar to the results in Models 1-4.

The results provide suggestive evidence that perceived closeness affected voter turnout in the 2004 presidential election. The closeness effect is statistically significant at the 10 percent level or higher in 3 of the 4 models. In Model 4 the effect is statistically significant at the 10 percent level, in Model 1 it is on the cusp of reaching significance at the 5 percent level, and in Model 3 the effect is significant at the 5 percent level. The closeness coefficient does not reach statistical significance at any conventional level in Model 2 (p = .139). Given the small sample size (50 observations for the closeness variable), it is reasonable to conclude that perceived closeness affected voter turnout in the 2004 presidential election.
Table 3: Turnout Models 1-4 Including All Variables and Significance Levels

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
</tr>
<tr>
<td>Statewide Variables of Interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.140 *</td>
<td>-0.123 **</td>
<td>-0.085 *</td>
<td></td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>-0.130 *</td>
<td>-0.104</td>
<td>-0.030</td>
<td></td>
</tr>
<tr>
<td>Closeness * Electoral Votes</td>
<td>0.007 *</td>
<td>--</td>
<td>0.005</td>
<td>--</td>
</tr>
<tr>
<td>Campaign Ads (in thousands)</td>
<td>0.123 ***</td>
<td>0.100 ***</td>
<td>0.111 ***</td>
<td>0.095 **</td>
</tr>
<tr>
<td>State and County Level Political and Systemic Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern State</td>
<td>-4.391 ***</td>
<td>-4.079 ***</td>
<td>-4.266 ***</td>
<td>-3.972 ***</td>
</tr>
<tr>
<td>Closing Date</td>
<td>-0.192 ***</td>
<td>-0.189 ***</td>
<td>-0.194 ***</td>
<td>-0.192 ***</td>
</tr>
<tr>
<td>Senate Race</td>
<td>-1.175</td>
<td>-1.321</td>
<td>-1.254</td>
<td>-1.346</td>
</tr>
<tr>
<td>Gubernatorial Race</td>
<td>-1.817</td>
<td>-1.903</td>
<td>-1.756</td>
<td>-1.760</td>
</tr>
<tr>
<td>Marriage Measure</td>
<td>2.209 *</td>
<td>2.127 *</td>
<td>2.218 *</td>
<td>2.151 *</td>
</tr>
<tr>
<td>Competitive House Race</td>
<td>1.498</td>
<td>1.568 *</td>
<td>1.529 *</td>
<td>1.580 *</td>
</tr>
<tr>
<td>County Level Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAP Male</td>
<td>-0.269 ***</td>
<td>-0.279 ***</td>
<td>-0.271 ***</td>
<td>-0.278 ***</td>
</tr>
<tr>
<td>VAP 18 to 24</td>
<td>0.050</td>
<td>0.062</td>
<td>0.056</td>
<td>0.065</td>
</tr>
<tr>
<td>VAP 45 and up</td>
<td>0.479 ***</td>
<td>0.484 ***</td>
<td>0.484 ***</td>
<td>0.487 ***</td>
</tr>
<tr>
<td>Black</td>
<td>0.062 *</td>
<td>0.057</td>
<td>0.063 *</td>
<td>0.060</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.091 *</td>
<td>-0.080 *</td>
<td>-0.089 *</td>
<td>-0.081 *</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.080 ***</td>
<td>-0.081 ***</td>
<td>-0.080 ***</td>
<td>-0.080 ***</td>
</tr>
<tr>
<td>Currently Married</td>
<td>0.133 *</td>
<td>0.140 *</td>
<td>0.138 *</td>
<td>0.145 *</td>
</tr>
<tr>
<td>College</td>
<td>0.529 ***</td>
<td>0.525 ***</td>
<td>0.530 ***</td>
<td>0.529 ***</td>
</tr>
<tr>
<td>Poverty</td>
<td>-0.194 *</td>
<td>-0.192 *</td>
<td>-0.185 *</td>
<td>-0.184 *</td>
</tr>
<tr>
<td>Constant</td>
<td>42.217 ***</td>
<td>41.072 ***</td>
<td>41.594 ***</td>
<td>40.622 ***</td>
</tr>
<tr>
<td>R Squared</td>
<td>0.657</td>
<td>0.654</td>
<td>0.657</td>
<td>0.656</td>
</tr>
</tbody>
</table>

Notes: *** p<.01; ** p<.05; * p<.10
Table 4: Models 1-4 with Variables of Interest, Standard Errors, and P Values

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>Robust Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Polls Margin with Interaction Effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.140</td>
<td>0.070</td>
<td>0.051</td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>-0.130</td>
<td>0.066</td>
<td>0.055</td>
</tr>
<tr>
<td>Closeness * Electoral Votes</td>
<td>0.007</td>
<td>0.004</td>
<td>0.067</td>
</tr>
<tr>
<td>Campaign Ads (in thousands)</td>
<td>0.123</td>
<td>0.037</td>
<td>0.002</td>
</tr>
<tr>
<td>Model 2: Polls Margin without Interaction Effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.079</td>
<td>0.053</td>
<td>0.139</td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>-0.023</td>
<td>0.038</td>
<td>0.557</td>
</tr>
<tr>
<td>Campaign Ads (in thousands)</td>
<td>0.100</td>
<td>0.036</td>
<td>0.008</td>
</tr>
<tr>
<td>Model 3: Actual Margin with Interaction Effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.123</td>
<td>0.060</td>
<td>0.047</td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>-0.104</td>
<td>0.069</td>
<td>0.139</td>
</tr>
<tr>
<td>Closeness * Electoral Votes</td>
<td>0.005</td>
<td>0.004</td>
<td>0.211</td>
</tr>
<tr>
<td>Campaign Ads (in thousands)</td>
<td>0.111</td>
<td>0.039</td>
<td>0.006</td>
</tr>
<tr>
<td>Model 4: Actual Margin without Interaction Effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>-0.085</td>
<td>0.047</td>
<td>0.077</td>
</tr>
<tr>
<td>Electoral Votes</td>
<td>-0.030</td>
<td>0.037</td>
<td>0.433</td>
</tr>
<tr>
<td>Campaign Ads (in thousands)</td>
<td>0.095</td>
<td>0.036</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Interestingly, Actual Margin seems to be a more efficient measure of closeness than Polls Margin. It turns out that the polling data did an excellent job of predicting the actual margin of the state elections; the correlation between Actual Margin and Polls Margin 1 is .9517. Actual Margin, however, has a smaller standard error in both the model with the interaction effect and the model without the interaction effect.
Apparently the theoretical objection to using an *ex post* measure of closeness is unwarranted. The traditional approach of using the actual margin of the election seems to be just as good if not better than constructing an *ex ante* measure.

The evidence on the effect of *ElecVotes* and the interaction effect is much less persuasive than the evidence on closeness. *ElecVotes* is significant at the 10 percent level in Model 1, but not in the other models, and the interaction effect is significant at the 10 percent level in Model 1, but not in Model 3. Focusing on Model 1, the results suggest that when a citizen makes her decision to vote or not, she takes into account whether her vote will be decisive in the statewide tally rather than whether her state’s electoral votes will be decisive in the Electoral College. As the state population increases (and thus its electoral votes), turnout decreases. The interaction effect suggests that the effect of closeness on turnout depends on the size of the state. According to Model 1, the effect of closeness is more powerful in smaller states than in larger states. That is, closeness will have a greater effect on turnout when comparing a small swing state like New Hampshire to a small spectator state like Rhode Island, than when comparing a large swing state like Florida to a large spectator state like New York. This makes intuitive sense; a citizen is more likely to believe that closeness can affect the decisiveness of his vote in a state with 1 million people than in a state with 30 million people. However, given that neither *ElecVotes* nor the interaction effect reaches statistical significance in Model 3, these results should be considered preliminary at best. It may be that 50 state observations is simply too small a sample size to pick up
these effects. Future research that increases sample size by examining more than one
election cycle may be able to better answer whether state size and the interaction
between state size and closeness affect turnout in presidential elections.

To evaluate the magnitude of the closeness effect, I examine Model 4. Because
Model 4 does not include the interaction effect, the coefficient on closeness in this
Model is nearer to the “pure” effect of closeness on turnout than the coefficient in
Model 3. According to Model 4, a 1 percentage point increase in Actual Margin is
accompanied by a .085 percentage point decrease in turnout. This is a small but
nontrivial effect. Increasing Actual Margin by 10 percentage points (the approximate
difference in margins between the quintessential swing state of Wisconsin and the
marginally competitive state of Arizona) is associated with a decrease in turnout of .85
percentage points. The average county in the sample has a Voting Age Population of
70,633 individuals; consequently, a .85 percentage point decrease in turnout in an
average-sized county would lead to 600 fewer people voting. Alternatively, if Actual
Margin were to increase by 30 percentage points (the difference in margins between
Wisconsin and the highly uncompetitive state of Oklahoma), turnout would decrease by
2.55 percentage points, and 1,800 fewer people would vote in an average-sized county.

It is important to note that the closeness effect is independent of the effect of
campaign advertisements on turnout. This result is consistent with previous literature
that has controlled for campaign activity (Cox and Munger 1989). As can be seen in
Table 4, Campaign Ads is highly statistically significant in each of the 4 models. The
models suggest that each additional 1,000 campaign advertisements boost voter turnout by .095 to .123 percentage points. To understand this finding, one needs to have a sense of the variation in campaign advertisements across states. The number of advertisements over this time period ranged from 0 for 22 of the states to a high of 55,487 advertisements in Florida. Among the states that received no campaign advertisements was California, the highest population state in the country. Now let us suppose that California residents were exposed to 30,228 campaign advertisements (the number of ads in the swing state of Pennsylvania). Using the most conservative estimate of Campaign Ads (Model 4), an additional 30,228 ads in California would lead to an increase in turnout of 2.872 percentage points. In 2004, California had a Voting Age Population of 26,266,383 people. If California was exposed to an additional 30,228 ads (from a base of 0) it is estimated that 754,371 more people would have voted statewide. Alternatively, if Campaign Ads was increased by 55,487 (the number of ads in Florida), turnout would be expected to increase by 5.271 percentage points in California, with 1,384,501 more people voting in the state. If we make the assumption that Campaign Ads is a reasonable proxy for campaign activity in general, the model provides impressive evidence that campaign activity can substantially boost voter turnout.

Finally, it should be noted that the models do an excellent job of predicting voter turnout. Each of the models explains over 65 percent of the variance in turnout across counties clustered within states. Of the 9 county level demographic covariates, 4 are
statistically significant at the 1 percent level, and another 4 are statistically significant at the 10 percent level. Only 1 of the demographic variables, Voting Age Population 18-24, fails to be significant in any of the models. Consistent with previous literature, education (College) and age (Voting Age Population 45 or Older) are the most powerful individual level determinants of turnout. Of the political and systemic factors, South and Closing Date are highly statistically significant and powerful predictors of turnout. The results also suggest that gay marriage amendments and competitive house races boosted turnout.

**Policy Implications**

This paper provides evidence that in the 2004 election, individuals were more likely to vote in swing states than in spectator states. This was a result of two phenomena. First, citizens perceived that their votes were more likely to be decisive in swing states, and second, citizens were encouraged to vote in swing states by a barrage of campaign advertisements that were nowhere to be seen in many other states.

These findings are important in terms of both policy and politics. Because only a small handful of states were highly competitive in 2004, and because on a national scale the election was quite competitive, it is likely that the Electoral College system depressed turnout in the 2004 election. In a system where the President was elected by a national popular vote, citizens would have been encouraged to vote because of how close the election was nationwide. Under the current Electoral College system, many voters were discouraged from voting because of where they lived.
It is likely that at least in the immediate future, presidential contests will continue to be close on the national scale but only competitive in a small group of swing states. As a result, the Electoral College will probably continue to depress voter turnout. Policymakers concerned with civic participation may be inclined to favor changes to the system that move away from the winner-take-all Electoral College System.
REFERENCES


