# Simulation of Increased Youth Turnout on the Presidential Election of 2004

Drew Kurlowski
University of Missouri – Columbia
dak6w7@mizzou.edu

*Abstract*

Youth voting has become a major issue in campaign politics. The 18 to 24 year old population, or 'Generation Y', represents the largest group of nonvoting but eligible Americans, with overall turnout for this group declining more than13% between 1972 and 2000. This research explores the potential impact of high levels of participation among this group on presidential politics.

Traditional studies suggest that increased voter turnout should favor Democratic candidates, because nonvoters share characteristics with those who vote for Democrats. Empirical research has failed to demonstrate a clear relationship, however, between higher turnout and Democratic success. I engage this literature by exploring the impact of full turnout among one of the most heavily mobilized groups in America, young voters, on the presidential election of 2004.

Using a novel approach to voter studies, I marry census data with voter exit polls to estimate a partisan preference in nonvoting youth. I then use this partisan preference to estimate the impact of higher youth voter turnout on the national vote totals for both John Kerry and George Bush. Because of the Electoral College and the institutional characteristics of Presidential elections in this country, however, this national impact really has no meaning. Thus, I simulate the impact of full youth voter turnout state by state and assess the impact on electoral vote totals.

The findings from these analyses suggest that youth nonvoters in 2004 were more likely to vote Democratic and full turnout among this group would have significantly increased John Kerry's share of the popular vote. Even if every person in this heavily mobilized group had turned out, however, it would have had essentially no impact on electoral vote totals and, thus, no impact on the outcome of 2004 election. Based on these findings, I conclude that previous studies on the potential and actual impact of turnout on presidential politics may have arrived at disparate conclusions because they did not take into account the institutional realities of presidential selection in this country.

Traditional research into voter turnout has espoused the view that increasing voter turnout helps produce more Democratic votes. "Those elements with the lowest voting rates consist in larger measure of persons with Democratic rather than Republican predispositions…Democratic candidates would have been strengthened…by bringing out substantially the entire electorate" (Key 1964). Even with this simple presupposition, there are few cases of direct tests of the theory. The logic behind these assumptions stems from the idea that the majority of non-voters are from lower socio-economic levels. These minority and working class Americans have tended to lean Democratic in past elections.

However, not all youth voters can be grouped into lower socio-economic levels. Youth span the continuum of race, gender, and income. This fact leads to obstacles in making assumptions about their behavior. Even with these obstacles, the idea that increasing their turnout would benefit Democrats has still prevailed in recent literature. Many scholars, such as Piven and Cloward (1988) and Lijphart (1997) have supported this presumption. Lijphart, for example, advocates mandatory voting in order to assure greater economic and social equality. In addition, Radcliff's study of Presidential elections up to 1980 has supported this traditional view (1994).

The notion that increased turnout does not favor Democrats has also been examined in recent literature. For example, DeNardo has tried to show that the relationship between turnout and party success is complex. He goes on to argue that there are many instances when Republicans gain strength as he illustrates with his "defection effect" (1980). Other research has backed up DeNardo's original finings. Nagel and McNulty have found similar results in Senatorial and Gubernatorial elections in their

research (1996). These disparate conclusions in literature suggest there is more to be studied here.

Interestingly, with gradually increasing turnout, we have not seen overwhelming Democratic victories. In fact, the exact opposite has been happening, as we see from the 2000 and 2004 elections of relatively staunch conservative, George W. Bush. These seemingly contradictory facts leave many questions to be answered. Why have elections over the past 30 years shown no correlation between increased turnout and Democratic victories? It would seem that even if increased turnout leads to more Democratic votes, there is another factor in play keeping Democrats from winning elections.

Another issue with the recent targeting of youth voters is the assumption that youth are a unified voting bloc. The basis of the theory that increased turnout aids Democrats is that the mobilized group being studied always has similar values and needs. This theory is reinforced because the groups used as examples in this research are all very similar. Examples are made out of minority voters, which have proven to be a unified voting bloc in many elections. In addition, lower and working class Americans have also shown they are a unified voting bloc through their common affiliation with labor unions and similar organizations. These groups all share common issues and difficulties. However it is hard to find a unifying thread among American youth. They come from all socio-economic backgrounds, are of all races and all genders. There is no common demographic or economic characteristic to unify them.

Most voter studies have focused on popular vote totals. Sometimes this route is taken because the elections in questions are decided by those totals, such as elections for Congress (Nagel and McNulty, 1996; Citrin, Schickler, and Sides 2003). These popular

vote totals are important, and do offer insights into the behavior of the group being studied, but they miss out on a complete analysis of the impact. Looking at popular vote totals may show that an individual has a certain partisan preference, but we do not elect our Presidents by popular vote totals. Thus, I take into account the institutional realities of the Electoral College in order to better answer questions about the effect of increased youth turnout on presidential selection.

In order to completely analyze the effect of this issue we must ask two questions. First, are youth voters a demographic that favors Democrats? Also, will increasing turnout amongst youth voters yield Democratic victories? It is clear that for the past 30 years of Presidential elections, there is no clear correlation between increased turnout overall and Democratic victories. Even with this fact, youth voters do seem to lean towards a more Democratic vote. The full effect of this and the possibility for Democratic victories still can not be answered with the existing research. If youth could act as a unified voting bloc, a simulation of the 2004 election with full youth turnout would show if it is possible for this group to have an impact on the election of a Democratic president.

*Why Youth Voters?*

In this study, I decided to target youth voters for a few different reasons. First, and foremost, I chose this group because it is assumed to lean Democratic and has been targeted for heavy mobilization for the past few Presidential elections. In the 2004 election it is estimated that various interest groups and voter organizations spent between 50 and 70 million dollars to reach out to young voters. The Pew Charitable Trust alone spent over 16 million dollars in grants funding youth voter outreach programs. Youth

between the ages of 18 and 24 comprise about 13% of the voting eligible population (CPS 2004) and with current trends youth could become more important in the future. As of 2000 the 0-17 year old population was an estimated 72.4 million people, this figure rivals the population of the 'baby boomer' generation (CPS 2004). With a potentially large group of youth filtering into the voting age population, it is important to be able to better understand how and why youth vote.

These basic assumptions and mobilization efforts make voters a particularly interesting group to study. Other voting blocs such as minorities and women are more understood and have more predictable behavior since they have come forward at times to be influential groups. The power and influence of youth have yet to be seen.

Even though we know youth are being mobilized and they most likely lean Democrat, there is still one last unanswered question. What is the effect of mobilization, and what potential do youth have to change the landscape of Presidential elections. Although it is difficult to tell if the efforts of youth mobilization groups are effective, it is possible to determine the effect of increased voter turnout on presidential elections. This simulation of the 2004 election should answer some important questions about who they are and their capabilities.

*Hypothesis*

Although empirical research suggests higher voter turnout would benefit Democrats, the effect on the actual outcome of an election has not been studied thoroughly. For many reasons, I believe that even though youth may tend to lean slightly Democratic, there would be no effect on the overall outcome of the election.

First, the very nature of youth voters suggests the election outcome would not change. Literature stating that youth favor Democrats still suggests that some youth would vote Republican. For this reason, the partisan vote split would give votes to each party candidate. This minimizes the chance of any large scale vote shift.

Secondly, the smaller nature of youth as a voting bloc suggests that even if they were all to vote for one candidate, that they could not overcome the rest of the voting public. Because of institutional realities such as the two party system and the Electoral College, it does not matter if one party would gain a larger vote share unless they can gain majority with it. For reasons of size, it seems improbable that youth would be able to tip the scales.

*Data and Methods*

By using exit polls from the 2004 Presidential election and census I determined a partisan preference of nonvoting youth. This was done using a logistic regression with vote choice as the dependant variable and age, income, gender, and race as my multiple independent variables (Table 1). These specific variables were necessary because they were common between the exit poll and census data. Although there are other variables that can also be very strong indicators of partisan preference, they were not common between the two data sources, and therefore could not be used. After the dissolution of the VNS (Voter News Service), who conducted polls on Election Day, we have to look to Edison Media Research and Mitofsky International, and its new polls to get the data. Their methods are extremely similar to the extinct VNS and they are generally accepted as the new standard in exit polls.

It has been suggested in the current literature that these demographic variables can accurately predict vote choice. The logistic coefficients for my control variables match well with current research regarding these demographics. The next step in the process is to marry the logistic coefficients with nonvoter demographic data. For this data, the CPS (Current Population Survey) is used. The annual voter supplement that is published contains the same variables as the earlier exit poll data and although it does not include vote choice, it does have a variable that shows whether the respondent voted or not.

Some of the variables, including age and income had to be recoded to match the values for the exit poll data. After recoding, it was possible to create a subset consisting of only non-voters using the variable that showed whether or not the respondent had voted. This is important, because we can not simply apply our coefficients to all people. The assumption is that non-voters as a group will look different than voters. Therefore with the new subset it is assured that we are predicting the vote choice of the correct people.

By multiplying the coefficients by each variable we come up with a new variable that needs to be fine tuned in order to become a probability. By taking the exponential function of the new productions in the equation, a new zero to one probability is formed. In this probability, one represents a Democratic vote and zero represents a Republican vote. I call these predicted probabilities the partisan preference.

With the new predicted probabilities and the CPS data in hand, the next step is to calculate new vote totals. First, the mean of the predicted probabilities for youth from each state are calculated (Table 2). Next, it is necessary to find out how many nonvoting youth were actually in each state. This figure was found by calculating an estimated

turnout percentage from the CPS voter supplement. This percentage is then applied to population estimates from the census to estimate the state by state non-voting population for 18 to 24 year olds. Now that we have a nonvoting population figure and a predicted probability for voting, multiplying the two together yields the number of new votes for the Democratic candidate, since a predicted probability of one (1) is associated with a Democrat vote. The other portion of the population would be a Republican Vote. These new votes are added to state by state vote totals for each candidate. The change in state totals is more important than an overall national change. With this data can tell if there would be a change in the distribution of electoral votes, and therefore an overall change in the outcome of the election. This simulation is laid out in Table 3, part a. and b.

One weakness with this marriage of data is the need to use common variables between Edison and the CPS. The CPS does not ask questions about vote preference or party preference, so I am not able to use these predictors in our model. I am limited to age, race, gender, income, and education as predictors for our model. Citrin, Schickler, and Sides (2003) used this model and found that their predictions in senate races correlated highly with actual election outcomes. Use of political party identification as one of the predictors made no significant difference in the model's accuracy.

*Findings and discussion*

Overall, the simulation showed no change in the outcome of the election except for the state of New Mexico. In this case, the electoral votes swapped from Bush to Kerry. However, the small number of votes (5) could not sway the overall outcome of the election. This proves the hypothesis that the youth could not change the overall outcome

of the election. The actual results of the simulated election are shown in Table 3. This is not the only finding that can be drawn from the research. Looking at the mean of the coefficients for all states, it is true that youth tend to lean more Democratic. This held in 49 out of the 50 states. This strengthens the argument that increased youth turnout would benefit Democrats to a certain extent, but the degree to which it would help seems diluted with the findings of this experiment.

The process of analyzing the regression was fairly simple using the statistical package Stata9. The 18 to 24 year old group was more likely to produce a Kerry vote than the other age groups. This is in line with the view that young voters tend to be more Democratic. Also, as expected, when looking at race, the minority groups identified in the logistic regression have coefficients significantly higher than whites. These coefficients do not show any radical or extraordinary results. For all of our control variables, the coefficients produced are in line with the view in current literature. The proportionate reduction in error for the logistic regression was .269 showing a good improvement over the modal choice. This is encouraging, as the model's accuracy directly impacts the accuracy of the simulation.

In addition, Table 1 shows the logistic coefficients and their corresponding z and *p*-values. In some cases, such as race, the *z*-scores were extremely high, indicating a good predictor of vote choice. Calculating a 95% confidence interval also assisted in the simulation. When both the upper and lower bounds were substituted into the simulation, there were no changes in the overall outcome of the election. This is probably due to the fact that with such minute changes, the large scale implication is virtually negated. Because of this evidence I feel confident that the model was accurate for the simulation.

*Conclusion*

The findings from my simulation show that as recent research has suggested, youth nonvoters in 2004 were more likely to vote Democratic. The slight advantage Democrats receive allows them to pick up extra popular votes in most states. The overall popular vote totals did not fully shift between Kerry and Bush, but Kerry was able to get a 1% increase in votes. This simulation shows that youth voters are not capable of completely changing the landscape of this Presidential election, but perhaps their turnout, combined with other factors, could increase the chance of a Democratic victory.

While this shift in popular vote totals is interesting, there are other factors that need to be considered in order to have a complete understanding of how the election would have turned out. Because we elect the President through the Electoral College, we can not simply stop with an analysis of popular votes. When these votes are simulated state by state, the Kerry advantage is significantly weakened. Even with his 1% overall gain, Kerry was only able to take one state away from Bush. New Mexico's shift from Kerry to Bush may be more complex than what appears on the surface. With the actual outcome being very close, and with the large population of minorities, there may be other forces working. Whether or not Kerry's New Mexico victory can be directly related to youth or not, it was still not enough to shift the balance. Kerry would have still come 14 electoral votes shy of the 270 needed for election. This shows that even if every person in the group had turned out, it would have had essentially no impact on electoral vote totals and, thus, no impact on the outcome of 2004 election.

Based on these findings, two conclusions arise. First, previous studies on the potential and actual impact of turnout on Presidential politics, specifically those focusing on youth voting, may be flawed. The assumption that youth as a voting bloc have the power to sway an election may be one of the first assumptions that needs to be reexamined. Further research into this issue could lend insight into whether or not this is true. Secondly, the focus on popular vote totals in past simulations does not truly represent the impact on elections that they are trying to examine. True impact on a Presidential election comes in the form of electoral votes. It is possible that the impact of youth voters may have been exaggerated due to neglecting the institutional realities of Presidential selection. This research asks many new questions that need to be studied further in order to come to a definite conclusion on the matter. However, this research shows that past works on the subject may have arrived at disparate conclusions due to assumptions and neglect of the Electoral College.

Table 1. Logistic Coefficients

| Variable | Variable Label | Coefficient | Std. Err. | z |
|---|---|---|---|---|
| A18to24 | Aged 18 to 24 | 0.3376353 | 0.0346193 | 9.75 |
| A25to29 | Aged 25 to 29 | 0.2668036 | 0.03628 | 7.35 |
| A30to39 | Aged 30 to 39 | 0.0653621 | 0.0297372 | 2.2 |
| A45to49 | Aged 45 to 49 | 0.1614431 | 0.0329485 | 4.90 |
| A50to59 | Aged 50 to 59 | 0.290668 | 0.0296165 | 9.81 |
| A60to64 | Aged 60 to 64 | 0.1688345 | 0.0397972 | 4.24 |
| a65to74 | Aged 65 to 75 | 0.2169413 | 0.0384842 | 5.64 |
| over75 | Over Age 75 | 0.4320188 | 0.0491131 | 8.8 |
| under15 | Income under $15,000 | 0.530701 | 0.0432599 | 12.27 |
| i15to30 | $15,000 to $30,000 | 0.3049729 | 0.0375002 | 8.13 |
| i30to50 | $30,000 to $50,000 | 0.1120395 | 0.0350751 | 3.19 |
| i50to75 | $50,000 to $75,000 | 0.0080654 | 0.0347961 | 0.23 |
| i75to100 | $75,000 to $100,000 | 0.071846 | 0.0372616 | 1.93 |
| i100to150 | $100,000 to $150,000 | 0.1208545 | 0.0392272 | 3.08 |
| black | Black | 2.248633 | 0.0400078 | 56.2 |
| latino | Latino | 0.7000971 | 0.0344896 | 20.3 |
| asian | Asian | 0.7942104 | 0.069555 | 11.42 |
| other | Other Race | 0.6948505 | 0.0534795 | 12.99 |
| man | Male | -0.1950661 | 0.0163155 | -11.96 |

N=66879 Psuedo r^2=0.0691 PCC = 60.20% PRE = 0.219

Table 2. Predicted Probabilities By State

| | Obs | Mean P.P. | Std. Dev. | | Obs | Mean P.P. | Std. Dev. |
|---|---|---|---|---|---|---|---|
| AL | 52 | 0.6067988 | 0.191001 | MT | 53 | 0.5380361 | 0.0796102 |
| AK | 52 | 0.5395406 | 0.1021014 | NE | 99 | 0.5304578 | 0.0909167 |
| AZ | 45 | 0.5788301 | 0.1082361 | NV | 57 | 0.5676573 | 0.136121 |
| AR | 69 | 0.5994464 | 0.1598381 | NH | 53 | 0.4921084 | 0.0715171 |
| CA | 253 | 0.6260642 | 0.1141231 | NJ | 67 | 0.5723216 | 0.1695238 |
| CO | 88 | 0.5909856 | 0.1367932 | NM | 69 | 0.6186059 | 0.1080821 |
| CT | 53 | 0.5316772 | 0.1293303 | NY | 154 | 0.6039545 | 0.1647208 |
| DE | 50 | 0.6058033 | 0.1776518 | NC | 90 | 0.6382793 | 0.1829087 |
| DC | 25 | 0.8441256 | 0.1480855 | ND | 65 | 0.5381453 | 0.0911505 |
| FL | 124 | 0.5876717 | 0.1633033 | OH | 116 | 0.5291793 | 0.1128756 |
| GA | 69 | 0.6353714 | 0.1920694 | OK | 53 | 0.6111006 | 0.1664722 |
| HI | 67 | 0.6421713 | 0.081067 | OR | 49 | 0.5560265 | 0.0954341 |
| ID | 55 | 0.5682419 | 0.1030054 | PA | 123 | 0.5150339 | 0.1034084 |
| IL | 122 | 0.5385818 | 0.1226646 | RI | 52 | 0.5583615 | 0.1493673 |
| IN | 81 | 0.5512878 | 0.1469369 | SC | 61 | 0.6582582 | 0.1985424 |
| IA | 69 | 0.5245325 | 0.0831945 | SD | 76 | 0.5107808 | 0.0576912 |
| KS | 74 | 0.560325 | 0.1363848 | TN | 83 | 0.5513407 | 0.1127816 |
| KY | 49 | 0.5298959 | 0.0799111 | TX | 208 | 0.6395349 | 0.1465368 |
| LA | 38 | 0.6686929 | 0.2023174 | UT | 71 | 0.5114546 | 0.0648861 |
| ME | 63 | 0.5273279 | 0.0773052 | VT | 56 | 0.5068966 | 0.0734859 |
| MD | 56 | 0.5864491 | 0.1626205 | VA | 55 | 0.5783407 | 0.1692946 |
| MA | 61 | 0.5520637 | 0.1403571 | WA | 52 | 0.5526731 | 0.1148691 |
| MI | 88 | 0.5407873 | 0.1475594 | WV | 32 | 0.5093277 | 0.0571931 |
| MN | 60 | 0.5277871 | 0.1032464 | WI | 55 | 0.5497347 | 0.1255904 |
| MS | 34 | 0.6197853 | 0.1895069 | WY | 70 | 0.5285127 | 0.0799374 |
| MO | 70 | 0.5391217 | 0.1030325 | | | | |

Table 3. Summary of Simulation

| | Actual 2004 Election | | | Simulated 2004 Election | | |
|---|---|---|---|---|---|---|
| | Kerry | Bush | Total | Kerry | Bush | Total |
| Popular Vote Totals | 59,028,111 | 62,040,610 | 121,068,721 | 68,072,918 | 68,441,803 | 136,514,721 |
| % Popular Vote | 48.76% | 51.24% | 100.00% | 49.86% | 50.14% | 100.00% |
| Electoral Vote Totals | 251 | 286 | 537 | 256 | 281 | 537 |
| Outcome | | | Bush Wins | | | Bush Wins |

# References

Citrin, Jack, Eric Schickler, and John Sides. 2003. “What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections.” *American Journal of Political Science*. 74(1):75-90.

DeNardo, James. 1980. “Turnout and the Vote: The Joke’s on the Democrats.” *American Political Science Review* 74(2):406–20.

Key, V. O. Jr. 1964. *Politics, Parties and Pressure Groups.* New York: Thomas Y. Crowell Company.

Lijphart, Arend. 1997. “Unequal Participation: Democracy’s Unresolved Dilemma.” *American Political Science Review* 91(1):1–14.

Nagel, J. H., McNulty, J. E. 1996. “Partisan effects of voter turn out in senatorial and gubernatorial elections.” *American Political Science Review*, *90*(4):780-793.

Piven, Frances Fox, and Richard A. Cloward. 1988. *Why Americans Don’t Vote*. New York: Pantheon Books.

Radcliff, B. 1994. “Turn out and the Democratic vote.” *American Politics Quarterl.y* *22*(3):259-276.