The B.E. Journal of Economic Analysis & Policy

Advances

Volume 10, *Issue* 1 | 2010 | *Article* 98

Protecting Minorities in Large Binary Elections: A Test of Storable Votes Using Field Data

Alessandra Casella* Shuky Ehrenberg†
Andrew Gelman‡ Jie Shen**

*Columbia University, ac186@columbia.edu
†JGB Capital, shuky.ehrenberg@aya.yale.edu
‡Columbia University, gelman@stat.columbia.edu
**UC Irvine, jshen3@uci.edu

Recommended Citation
Alessandra Casella, Shuky Ehrenberg, Andrew Gelman, and Jie Shen (2010) "Protecting Minorities in Large Binary Elections: A Test of Storable Votes Using Field Data," *The B.E. Journal of Economic Analysis & Policy*: Vol. 10: Iss. 1 (Advances), Article 98.
Available at: http://www.bepress.com/bejeap/vol10/iss1/art98

Protecting Minorities in Large Binary Elections: A Test of Storable Votes Using Field Data*

Alessandra Casella, Shuky Ehrenberg, Andrew Gelman, and Jie Shen

Abstract

The legitimacy of democratic systems requires the protection of minority preferences while ideally treating every voter equally. During the 2006 student elections at Columbia University, we asked voters to rank the importance of different contests and to choose where to cast a single extra "bonus vote," had one been available — a simple version of Storable Votes. We then constructed distributions of intensities and electoral outcomes and estimated the probable impact of the bonus vote through bootstrapping techniques. The bonus vote performs well: when minority preferences are particularly intense, the minority wins at least one contest with 15-30 percent probability; when the minority wins, aggregate welfare increases with 85-95 percent probability. The paper makes two contributions: it tests the performance of storable votes in a setting where preferences were not controlled, and it suggests the use of bootstrapping techniques when appropriate replications of the data cannot be obtained.

KEYWORDS: voting, minorities, storable votes, elections, bootstrap

*We thank the editor of this journal, two anonymous referees and S. Turban for their comments; A. Sacarny for help with programming; and R. Davidson for discussions on the bootstrap. The research was supported by NSF grants SES-00214013 and SES-0084368 and by the Guggenheim Foundation. A. Casella thanks NBER for its hospitality, and numerous seminar audiences for comments.

1 Introduction

When voters must choose between two alternatives, majority voting works well, with a single important drawback: the winning alternative commands the wider support, but not always the more intense. At least since Madison, Tocqueville, and Mill, political thinkers have argued that a necessary condition for the legitimacy of a democratic system is to set limits to the *tyranny of the majority* and provide some expression to intense minority preferences.[1] With increased recourse to direct democracy, the goal becomes particularly important because direct democracy deprives minorities of the protections afforded by a diverse legislature. Devising systems to protect minorities when binary choices are at stake is thus central to the establishment of referendums as legitimate tools of policy-making.[2] The challenge is to do so while treating every voter equally, and avoiding the inertia and inefficiencies of super-majorities or veto powers.

A possible answer comes from *Storable Votes*, a simple voting mechanism in which individuals voting over multiple binary proposals are granted, in addition to a regular vote for each proposal, one or more bonus votes to cast as desired. Decisions are then taken according to the majority of votes cast. Voters are allowed to single out the issues that they each consider most important and cast on them their bonus votes, making it possible for the minority to win occasionally. But because every voter is treated equally, the majority loses only if, on average, its members consider a given issue a low priority, not deserving of bonus votes, while members of the minority do not. The majority loses ”when it should”, from an efficiency point of view. Although counterexamples can be found, typically the increase in minority representation comes together with an increase in aggregate expected welfare, relative to simple majority voting.[3]

Storable votes resemble *Cumulative Voting*, a voting system occasionally used in corporations and local jurisdictions in the United States and recommended by the courts exactly to redress violations of fair voting rights. As with storable votes, cumulative voting assigns to each voter a budget of votes to spend freely over multiple choices, but cumulative voting applies to a single election of multiple representatives, while storable votes apply to multiple decisions, each between

---

[1]For a recent discussion, see for example Guinier, 1994. Dahl,1956 and 1989, provide a critical analysis of the arguments; Issacharoff, Karlan and Pildes, 2002 discuss the legal implementation and dilemmas.

[2]Gerber, 1999, and Matsusaka, 2004.

[3]Casella, 2005; Casella, Palfrey and Riezman, 2008. A very similar voting scheme is proposed in Hortala-Vallve, 2007. As shown by Jackson and Sonnenschein, 2007, the idea of linking different decisions to elicit individuals' intensity of preferences can be applied generally. The mechanism proposed by Jackson and Sonnenschein achieves the first best asymptotically, as the number of linked decisions becomes large, but is significantly more complex than storable votes.

two alternatives. The strategic games induced by the two voting schemes are very different–think for example of the very different criteria for victory in the election–and to our knowledge existing analyses of cumulative voting do not discuss its efficiency properties.[4]

Storable votes are simple, and in countries where voters are routinely presented several referendums at the same time–in the United States, for example–could be implemented as a minor modification of existing voting practices. So far, tests have been very encouraging, but they have been limited to small groups in laboratory settings.[5] The tight controls of the laboratory should be complemented by field studies where the number of voters is large, and voters' preferences are observed, as opposed to being induced. This paper reports on a test based on field data collected during an actual election. The goal is to evaluate the impact of storable votes on the probability of minority victories, aggregate welfare, and voters' ex-post inequality.

In the spring of 2006, we attached a short survey to students' election ballots in two different schools at Columbia University, asking students to rank the importance assigned to all binary contests on the ballot, and to indicate where they would have cast an additional bonus vote, had one been available. An identifier connected responses and actual voting choices, allowing us to construct distributions of intensities and to propose welfare measures of the electoral outcomes, both without and with the bonus vote. Bootstrapping techniques provide estimates of the bonus vote's probable impact. The bonus vote choices indicated by the survey respondents are hypothetical only - elections are delicate matters and not surprisingly we were not allowed to change the actual voting system. Thus, as a robustness check, in addition to the use of the bonus vote reported in the survey, we study three alternative plausible rules for casting the bonus vote. For each of the four cases, we estimate three measures: (1) the frequency with which the bonus vote allows the minority to win at least one election in each set; (2) the difference in aggregate welfare, comparing the hypothetical outcome using bonus votes to simple majority voting; and (3) the impact of the bonus vote on ex post inequality. We find that the bonus vote works well: when minority preferences are particularly intense, the minority wins at least one of the contests with 15–30 percent probability, ex post inequality falls, and yet when the minority wins aggregate welfare increases with 85–95 percent probability. When majority and minority preferences are equally intense, the effect of the bonus vote is smaller and more variable, but on balance still positive.

[4]Experiences with cumulative voting in local elections are discussed in Bowler et al., 2003, and Pildes and Donoghue, 1995. Cox, 1990, studies the scheme theoretically, Gerber et al., 1998, test it experimentally. Cumulative voting was advocated particularly by Guinier, 1994.

[5]Casella, Palfrey and Gelman, 2006; Casella, Palfrey and Riezman, 2008, Casella, 2010, Hortala-Vallve and Llorente-Saguer, 2010.

Student elections are rarely considered worthy of study because they are low-stake contests, measuring students' popularity more than the attractiveness of their electoral platforms, in a population that is hardly representative of a typical electorate. But our work is not about extrapolating political tendencies from campus elections, nor, because it focuses on binary contests only, is it really about the Columbia campus elections themselves. Our objective is to illustrate the voting mechanism and its comparison to simple majority voting in realistic large scale elections where the distribution of preferences *is not controlled by the researchers* — this is the central requirement here. For our purposes, the true determinants of the voters' preferences are secondary.

The paper proceeds as follows. In section 2 we discuss briefly the theory behind the idea of granting a bonus vote in large elections; section 3 describes the design of the survey and the data; section 4 studies the bonus vote choice; section 5 describes the bootstrapping exercise and its results, and section 6 concludes. A copy of the questionnaire (Figure A.2) and some additional data and results (Figure A.1) are reported in the Appendix.

2 The Theory

We begin by summarizing briefly the theory of storable votes in large elections. The results described in this section are drawn from Casella and Gelman (2008), where the model is phrased in terms of contemporaneous referendums over several proposals. Here, in line with the student elections, the two outcomes of each election are labelled as two different candidates.

A large number N of voters are asked to vote, contemporaneously, on a set of K unrelated elections (with $K > 1$). Each election E_k is between two candidates, a_k and b_k, with $k = 1, \ldots, K$. Voters are asked to cast one vote in each election, but in addition are given a single bonus vote that can be spent on any of the elections. Each election is won by the candidate with most votes, including bonus votes.

The preferences of voter i in election k are summarized by a valuation v_{ik}. By convention, a negative valuation indicates that i favors candidate a, and a positive valuation that i favors b. The valuation's absolute value, denoted by v_{ik}, is the voter's differential utility from having his or her preferred candidate win the election — the *intensity* of the voter's preferences. Voter i's utility function is then $U_i = \sum_{k=1}^{K} u_{ik}(E_k)$ where $u_{ik}(E_k) = v_{ik}$ if i's favorite candidate wins E_k, and 0 otherwise. Note that what matters is only the differential utility from one's preferred candidate winning as opposed to losing. The formulation presented here normalizes to zero the utility from losing, but is identical to assigning some utility v_{ik}^{w} to winning and some disutility $-v_{ik}^{l}$ to losing, where $v_{ik}^{w} + v_{ik}^{l} = v_{ik}$.

Individual valuations are drawn independently across individuals from a joint distribution $\mathscr{F}(\mathrm{v}_1,\ldots,\mathrm{v}_K)$. The assumption of a common distribution is in practice equivalent to assuming that we have no additional knowledge about individual voters (which, for example, would allow different distributions for men and for women, or, in our specific data, for English and for history majors). In this section, although not in the empirical analysis, we also assume that individual preferences are independent across elections, and thus voters' valuations over election k are drawn from some distribution $F_k(\mathrm{v})$. Each individual knows his or her own valuation over each proposal, and the probability distribution $F_k(\mathrm{v})$ of the others' valuations. There is no cost of voting.

Simple majority voting is designed to give weight to the *extent* of support for a candidate — the mass of voters who in election k prefer a to b. Storable votes allow voters to express not only the direction of their preferences but also to some extent, through the bonus vote, their *intensity*. How important accounting for intensity is depends crucially on the shape of the distributions $F_k(\mathrm{v})$.

Suppose first that the environment is fully symmetric, and nothing systematic distinguishes either the elections or the two sides in each election: $F_k(\mathrm{v}) = F(\mathrm{v})$ for all k and $F(\mathrm{v})$ symmetric around 0. Then: (a) in equilibrium each voter casts the bonus vote in the election to which the voter attaches highest intensity; (b) with positive probability, one or more of the elections, although not all, are won by a candidate supported by a minority of the electorate; and (c) ex ante expected utility with storable votes is higher than ex ante expected utility with simple majority voting.[6]

The results are clean, but the assumption of full symmetry is strong and in fact minimizes the importance of the voting rule. If the number of voters is large, the empirical frequencies of the preference draws must approximate more and more precisely the theoretical distributions. In the limit, in each election the two candidates come to be supported by an equal mass of voters with equal distribution of intensity, and thus, from a welfare point of view, become interchangeable. Although storable votes dominate majority voting for all finite population of voters, asymptotically all selection rules must be equivalent.

Allowing for asymmetries is important, but not only are asymmetries difficult to handle analytically, it is also not clear how best to model them. Suppose first that all asymmetries came from the *extent* of support: in each election, one

[6] To be precise, in the scenario described in the text the welfare superiority of storable votes holds only if the value of the bonus vote is below a threshold that depends on the shape of F. In a more general model, however, where the probability of supporting either candidate is not $1/2$ but is stochastic and distributed according to some distribution symmetric around $1/2$, the result holds for all values of the bonus vote. The assumption of symmetry across elections ($F_k = F$) simplifies the analysis but can be dropped fairly easily.

candidate is more popular than the other, and the difference in popularity — the difference in the expected mass of supporters — in general varies across elections. Conditional on supporting either candidate, however, the distribution of preferences intensity is equal across the two groups of supporters.[7] This is the environment to which majority voting is ideally suited. Then: (a) in equilibrium all voters cast their bonus vote in the election expected to be closest; (b) storable votes and simple majority yield identical outcomes: the minority never wins; (c) both voting rules are ex ante efficient.

Now suppose instead that the source of asymmetry is the *intensity* of support. In all elections, the two candidates are equally popular ex ante, but intensities are not equally distributed. With no loss of generality, suppose that in each election k, the mean intensity of supporters of candidate a_k is higher than the mean intensity of supporters of candidate b_k.[8] Then: (a) in equilibrium each voter casts the bonus vote in the election with highest intensity; (b) a minority candidate is expected to win occasionally with positive probability; (c) ex ante expected utility with storable votes is higher than expected utility with simple majority if a voter's highest valuation is expected to be on a candidate of type a; and (d) the difference in ex ante utility does not disappear asymptotically. Condition (c) is satisfied automatically if, for example, the distribution of intensities among a_k supporters first-order stochastically dominates the distribution of intensities among b_k supporters. More generally, numerical simulations show that violating condition (c) is possible but, as long as supporters of candidate a_k have higher expected intensity, not easy: the number of elections K must be large enough, and some counter-intuitive constraints on the shapes of the distributions of intensities must be satisfied.[9]

Thus, with the qualification just described, the theoretical analysis leads us to expect that storable votes should be superior, in ex ante utility sense, to simple majority. In realistic situations, however, both extent and intensity of support are likely to differ across candidates and across elections. It is difficult to say anything general in such cases because equilibrium strategies will depend on the exact shape of the distributions of valuations. The limitations of a purely theoretical analysis are one of the main motivations of this study: What shapes do the distributions of valuations take in practice? How well do storable votes behave when information about such distributions is imprecise?

---

[7] Formally, call $G_{ak}(v)$ the distribution of intensity for supporters of candidate a in election k, and p_k the ex ante probability of favoring candidate a in election k. Then suppose $G_{ak} = G_{bk} = G$ for all k, but $p_k \neq p_s$ for $k \neq s$. The previous assumption: $F_k = F$ for all k and F symmetrical around 0, corresponds to $G_{ak} = G_{bk} = G$ for all k, *and* $p_k = p = 1/2$ for all k.

[8] Suppose $p_k = p = 1/2$, but $G_{ak} = G_a \neq G_{bk} = G_b$ for all k. In particular suppose that G_a has higher mean than G_b: $Ev_a > Ev_b$.

[9] See the discussion in Casella and Gelman (2008).

3 The Test

3.1 The design

Several of Columbia's schools hold elections in the spring, and the students' organizations and the deans at the School of General Studies (GS) and Columbia College (CC) agreed to collaborate with us. In each school, voters elected representatives for multiple positions: GS students voted on a total of twenty different elections and CC students on twelve. We selected the subset of elections with two candidates or two mutually exclusive party lists only — three elections in GS (Board President, Alumni Representative and International Representative), and four elections in CC (Executive Board, Senator-Two Year Term, Senator-One Year Term, and Academic Affairs Representative)[10].

All voting was electronic, and at the end of the ballot students were invited to participate in our survey. In the GS case, interested students clicked on a link and were redirected to a web page containing the survey. Students' votes and their responses to the survey were matched and saved under anonymous identifiers and were later forwarded to us by the student bodies supervising the elections. In the CC case, the survey was on paper, because of logistical difficulties with the voting stations. At the end of their electronic ballot, CC students interested in answering the survey were given a number to be copied at the top of the paper questionnaire. The number allowed us to link their responses to their actual votes, again forwarded to us anonymously after the voting was concluded.

Understandably, the organizers of the elections were concerned with keeping our interference minimal, and the survey had to be short. We asked two sets of questions: first, how much the voter cared about the outcome of each of those elections, on a scale from 1 (not at all) to 10 (a lot); second, in which of these elections the voter would cast a single additional bonus vote in support of his or her favorite candidate, had one been available. The paper questionnaire for CC is reproduced in Figure A.2 in the Appendix; the electronic GS questionnaire was identical, with one exception — being electronic, and thus faster, we added a question about expected election outcomes.

Prior to the elections, students in both schools were informed about the survey through a school-wide email message, and through posters and fliers distributed widely throughout campus. Possible prizes from answering the survey were adver-

[10]We excluded the elections for class presidents or representatives because they concern different subsets of the electorates. We also excluded two elections where one of the candidates was accused of irregularities and later disqualified.

tised: respondents from each school would take part in a lottery with iPods and $20 gift certificates at Barnes & Noble awarded to the winners.[11]

3.2 The data

Out of a total of 1161 GS students, 476 voted in the GS elections, and of these 297 answered our survey; in the College, 2057 voted out of a potential electorate of 4073, and 644 answered the survey. After eliminating CC questionnaires that were either unreadable or unmatchable to actual voters, we cleaned the data according to the following criteria: (1) we assigned a score of 0 to any election in which a respondent abstained, a plausible option logically, and the only one available since we could not match the score to a voting choice; (2) we assigned a score of 1 to any election in which a respondent voted but left the ranking blank, and (3) we eliminated from the sample respondents who stated that they would cast the bonus vote in an election in which they in fact abstained, implying that the survey answers were patently untruthful or confused. Because the CC questionnaires were on paper, they were missing automatic completeness checks that the electronic program forced instead on the GS students. In particular, nine CC respondents did not indicate where they would have cast the bonus vote. But choosing not to use the bonus vote is a legitimate choice, and we left these respondents in the sample. After cleaning the data, we were left with 276 responses in the GS sample and 502 in the CC sample, or a valid response rate among voters of 58 percent in GS and 24 percent in CC.

Participation in the survey was voluntary, and individuals in our samples appear more engaged in the elections than the rest of the electorate. Table A.1 in the Appendix tests the hypothesis that in each school both the rest of the electorate and our sample are random draws from a common population. When we look at abstention rates, the hypothesis is rejected at the 5 percent confidence level for all GS elections and for two of the four CC elections — predictably individuals who answered the questionnaire have significantly lower abstentions rates. More surprisingly, we also find differences in preferences for candidates. When we look at support for the winning candidate among voters, the hypothesis of random draws from a common population is rejected in the President election in the GS sample and in the Academic Affairs election in the CC data — although the majority candidate does not change, both elections were much closer in our sample than in the full electorate, a surprising finding for which we have no explanation. Statistical techniques would allow us to reweigh the samples to achieve representativeness,

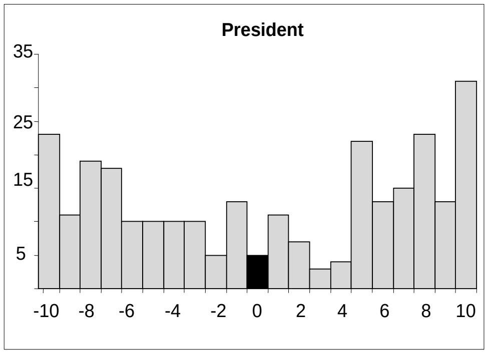
[11] The prizes were 2 iPods and 8 gift certificates for the CC lottery (with a potential electorate of 4,073); and 1 iPod and 5 gift certificates for the GS lottery (with a potential electorate of 1,161 students).

but we decided not to do so for two reasons. First, for consistency with the theory, we selected only elections with binary choices. Thus our set of elections is also limited, and even with representative samples the test could not apply to the actual Columbia student elections. Second, the biases we see in the data are not predictive of any clear bias in the performance of storable votes. Recall that we are interested in the comparison between storable votes and simple majority. It is difficult to see which effect if any should result from lower abstention rates. As for the smaller margins of victory, the addition of a bonus vote may result into higher probability of minority victories, but there is no ground on which to expect either higher or lower welfare. For these reasons, in the remainder of the paper we will ask not what the impact of the bonus vote would have been in the Columbia elections, but what it would have been in the subset of elections we consider, in a population for which our samples are representatives. As discussed in the Introduction, the interest of the study is not in the elections per Se, but in testing the impact of the bonus vote when we do not control the distributions of valuations and the voters' information and beliefs, two conditions indeed satisfied by our samples.

3.3 The distributions of the scores

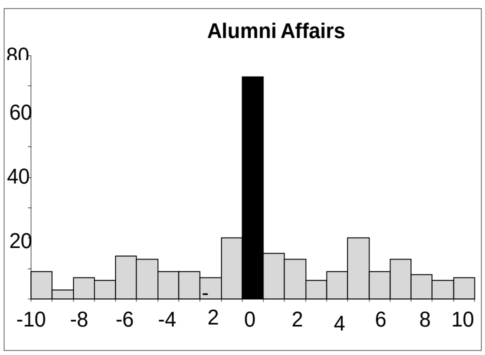
Figure 1 reports the distributions of the scores assigned by respondents to each election, for both GS and CC. Scores are labelled positive or negative, according to which of the two candidates the respondent voted for; for ease of reading, in each election we assign positive scores to the candidate commanding a majority in our sample. Consider for example the election for Board President in the GS sample, at the top of the figure. The scores assigned by Susannah's supporters are plotted on the positive axis; those of Liz's supporters on the negative axis, while the cell at zero, in black, reports abstentions. The histogram tells us that 23 of Liz's supporters and 31 of Susannah's assigned to the election a score of 10; 11 of Liz's supporters and 13 of Susannah's assigned it a score of 9, etc. The box next to the histogram summarizes the main data concerning this election. In our sample of 276 students, 142 voters supported Susannah, 129 supported Liz, and 5 abstained; 208 stated that they would have cast the bonus vote in this election, and, attributing the bonus votes according to each respondent's actual candidate choice, Susannah again wins a majority of the bonus votes (110 against 98). Susannah did in fact win the election among all voters, with 241 votes in favor, versus 188 for Liz and 47 abstentions. The box reports also the average score assigned to the election by the supporters of each candidate: the average score is slightly higher among Susannah's supporters (6.7 versus 6.3 for Liz's supporters).

GS



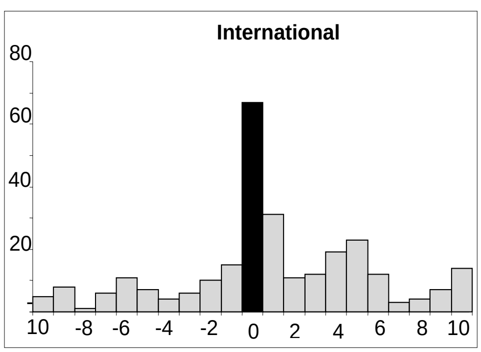
Susannah >0
Votes: Susannah 142; Liz 129; Abs 5
Bonus Votes: Susannah 110; Liz 98
Average Score: Susannah 6.7; Liz 6.3

ELECTION RESULTS
Susannah 241; Liz 188; Abs 47



Bob >0
Votes: Bob 106; Maria 97; Abs 73
Bonus Votes: Bob 17; Maria 19
Average Score: Bob 5; Maria 4.8

ELECTION RESULTS
Maria 151; Bob 144; Abs 181



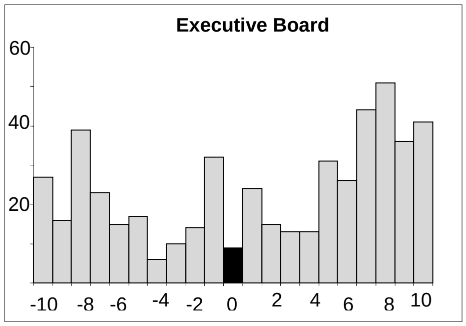
Makiko >0
Votes: Makiko 136; Liron 73; Abs 67
Bonus Votes: Makiko 21; Liron 11
Average Score: Makiko 4.5; Liron 4.7

ELECTION RESULTS
Makiko 203; Liron 113; Abs 160

(a) GS

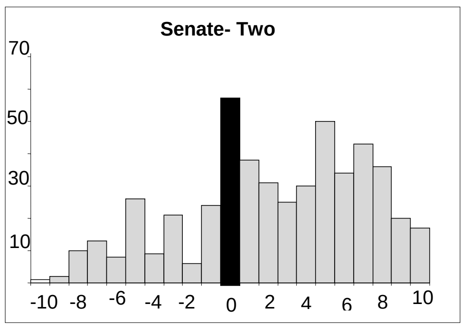
Figure 1: Histograms of recorded scores

CC



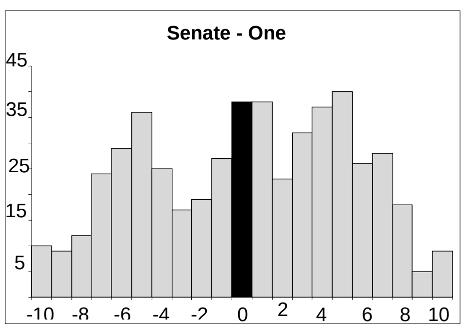
Open Columbia >0
Votes: O. C. 294; Evolution 199; Abs. 9
Bonus Votes: O. C. 205; Evolution 105
Average Score: O. C. 6.5; Evolution 5.9

ELECTION RESULTS
O. C. 1107; Evolution 819; Abs 131



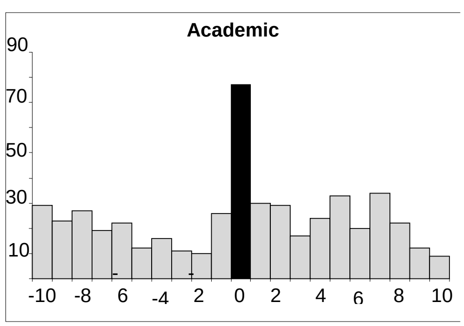
Tiffany >0
Votes: Tiffany 324; Gerry 120; Abs 58
Bonus Votes: Tiffany 48 ; Gerry 9
Average Score: Tiffany 5.2; Gerry 4.3

ELECTION RESULTS
Tiffany 1233; Gerry 533; Abs 291



David >0
Votes: David 256; Yihe 208; Abs 38
Bonus Votes: David 29; Yihe 29
Average Score: David 4.5;Yihe 4.9

ELECTION RESULTS
David 1015; Yihe 809; Abs 233



Alidad >0
Votes: Alidad 230; Ehizoje 195; Abs. 77
Bonus Votes: Alidad 30 ; Ehizoje 38
Average Score: Alidad 4.9; Ehizoje 6.1

ELECTION RESULTS
Alidad 1060; Ehizoje 634; Abs 363

(b) CC

Figure 1: Histograms of recorded scores (cont.)

In the GS sample, the election for President was overwhelmingly the most salient: abstentions are less than one tenth of the second most attended election (International), and average scores are about seventy percent higher than in the other elections. Two of the three elections (President and Alumni Affairs) were close elections, and in these elections not only is the extent of support similar between the two candidates, but the distributions of voters' scores are also approximately symmetrical across the two sides. The third election (International) was quite lopsided. Note that if the elections had been held with bonus votes, none of the outcomes would have changed: in the President and in the International election, the majority winner commands a majority of the bonus votes; it does not in Alumni Affairs, but the difference is small (2 votes), and cannot override the difference in regular votes.

In the CC sample, there is again one election that voters in our sample considered most salient (Executive Board), with both low abstentions and higher average scores, but the difference with respect to the others is less pronounced than in GS. The Senate-Two Year election was a landslide, but the others, and in particular Senate-One Year and Academic Affairs were close elections. The Academic Affairs election, with scores distributed according to the histogram at the bottom of Figure 1(b), is particularly interesting. Here Alidad won a majority of the votes (230 versus 195 for Ehizoje), but the average score is higher among Ehizoje's supporters (6.1 versus 4.9 for Alidad). The distribution is not symmetric, with the majority of Ehizoje's supporters attributing high importance to the outcome, while the scores given by Alidad's supporters are concentrated in the middle range. Storable votes are designed to address situations of this type, and not surprisingly Ehizoje receives about 25 percent more bonus votes than Alidad (38 to 30), although the difference of 8 votes would not have been enough to counter the majority advantage. In the CC sample too bonus votes would not have changed any of the outcomes.

The scores reflect the importance attached by respondents to the different elections, and we interpret them as measures of intensity of preferences. More precisely, and in line with the theoretical model, we read the scores as proxies for the differential utility that each respondent attaches to having the preferred candidate win that specific election, as opposed to the opponent. The important information is the *relative* score assigned by each respondent to different contests: it is preferable to have one's favorite candidate win an election rated as a 4 than an election rated as a 2. To give a measurable meaning to intensity of preferences, we must interpret the scores as not only ordinal but cardinal values: we use the simplest linear mapping, so that winning an election rated as a 4 is not only preferable but twice as valuable as winning an election rated as a 2. The distributions of the scores in Figure 1 are then the empirical counterpart of the distributions of preferences discussed in the theory.

If the scores measure intensity of preferences, we can construct measures of aggregate welfare in our samples. The most immediate utilitarian measure simply sums all scores over supporters of each candidate and calls *efficient* in any election the victory of the candidate whose supporters have higher aggregate scores. For example, in the Board President election in GS, Susannah's supporters are both more numerous and have higher average score, guaranteeing a higher aggregate score (954, compared to 810 for Liz's supporters), leading to the conclusion that efficiency favored Susannah's victory.[12]

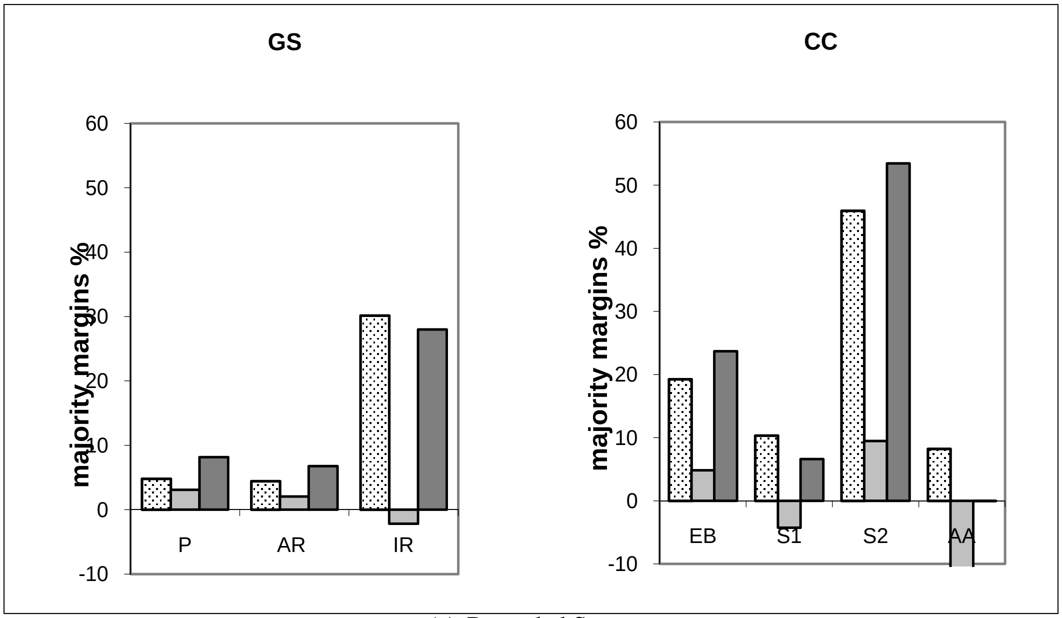
As welfare measure, the sum of the reported scores has the important advantage of being closest to the questionnaire. The difficulty is that in general, evaluated over all elections, it will give different weights to different voters, reflecting differences in the total scores that each of them has assigned. An alternative is to construct and sum *normalized* scores: scores obtained by constraining all individuals to the same total score. Specifically, if the number of elections is K and s^R_{ik} is the reported score assigned by individual i to election k, normalized score s_{ik} equals $s^R_{ik}/\sum_{r=1}^{K} s^R_{ir}$, with the property that the sum of normalized scores assigned by a single individual always equals 1. A normalized utilitarian measure of welfare is then the sum of the normalized scores over supporters of each winning candidate.

The histograms describing the distributions of normalized scores are reported in the Appendix. In Figure 2 we summarize the properties of our data, using both reported and normalized scores. The figure reports margins in favor of the majority, in each election, in terms of the number of votes, aggregate, and average scores.[13] Figure 2(a) is based on reported scores, as was Figure 1, while Figure 2(b) uses normalized scores.

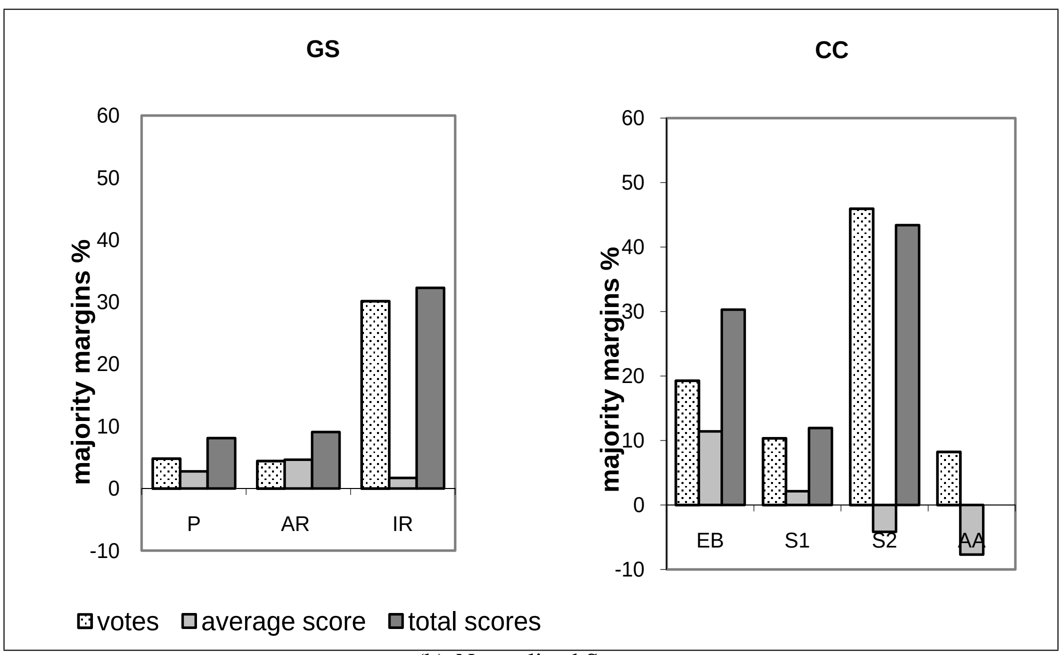
The two sets of figures are not identical, suggesting that the normalization does play some role: because it accounts for each individual's overall scoring pattern, the normalization in general rescales differently scores from different individuals, and thus changes both average and total scores. However, the figures make

[12]Note, for clarity, that normalizing the utility loss to zero in case of defeat has no effect on the welfare criterion. Suppose to the contrary that v^w_{ik} and $-v^l_{ik}$ were the utility gain and loss to voter i in election k from (i's favorite candidate) winning or losing. Then efficiency favors candidate a in election k if $\sum_{i\in A_k} v^w_{ik} - \sum_{i\in B_k} v^l_{ik} \geq \sum_{i\in B_k} v^w_{ik} - \sum_{i\in A_k} v^l_{ik}$ where A_k is the set of voters who favor a and B_k is the set of voters who favor b. Clearly this is equivalent to $\sum_{i\in A_k}(v^w_{ik}+v^l_{ik}) \geq \sum_{i\in B_k}(v^w_{ik}+v^l_{ik})$. Thus, for example, assigning utility equal to one's score in case of victory, and disutility equal to one's negative score in case of loss is equivalent to multiplying all scores by 2, with no effect on the efficiency criterion.

[13]If $\mathbf{M}$ is the size of the majority, and $\mathbf{m}$ the size of the minority, the margin of victory among voters is simply: $(\mathbf{M}-\mathbf{m})/N$. Using s to denote the reported score in Figure 2(a) and the normalized score in Figure 2(b), the aggregate score margin in favor of the majority is calculated as: $(\sum_{i\in M} s_i - \sum_{i\in m} s_i)/\sum_{i=1}^{N} s_i$, and the average score margin is: $(\hat{s}_M - \hat{s}_m)/(\hat{s}_M + \hat{s}_m)$ where $\hat{s}_M = (\sum_{i\in M} s_i)/\mathbf{M}$ and $\hat{s}_m = (\sum_{i\in m} s_i)/\mathbf{m}$.



(a) Recorded Scores



(b) Normalized Scores

Figure 2: Margins in favor of the majority, per election, in the samples of respondents. The first series (dotted) is the margin of victory, in votes, ignoring the bonus vote. The second series (light grey) is the average score margin. The third series (dark grey) is the aggregate score margin.

clear that the qualitative conclusions are robust: if efficiency is measured by higher aggregate scores, the efficient outcome in each election is unchanged whether we refer to reported or normalized scores — the dark grey columns always have equal sign in Figures 2(a) and 2(b). In the GS sample, efficiency always supports the majority choice. In the CC sample, efficiency supports the majority side on three of the four elections, but the majority's total score margin is barely negative, with a margin so small as to be undetectable in the figure, in the fourth (Academic Affairs) where the larger size of the majority is countered by the stronger intensity of preferences of the minority. Again, the observation holds for both sets of scores.

3.4 The bonus vote decision

In the GS sample, 89 percent of the voters stated they would cast their bonus vote in an election to which they assigned their highest score, and 81 percent did so in the CC sample. If voters knew little about all elections and were aware of their lack of information, this would be both the simplest and the optimal strategy. Which other criteria influenced their choice? Figure 3 shows the relevant data, election by election, for both GS and CC. Each diagram reports, for all respondents who said they would cast the bonus vote on that specific election, the reported score assigned to that election, on the vertical axis, and to the highest ranked of the other elections, on the horizontal axis. If the election selected for the bonus vote is the highest score election, then the respondent is indicated by a point above the 45 degree line; if not, by a point below the 45 degree line.[14]

The most salient election in each data set (President in GS and Executive Board in CC) was selected to receive the great majority of the bonus votes, and was also the election most respondents ranked as most important to them. In all elections, some bonus votes were cast by respondents who ranked a different contest higher. The number of such bonus votes is relatively small in both the GS President election (8 percent) and the CC Executive Board election (12 percent), but less so, in relative terms, in the other contests, reaching 50 percent in the Senate-One Year election in CC. Over the full electorate — and it is the full electorate that a voter would consider when deciding where to cast the bonus vote — Senate-One Year was the closest election in CC. If voters rationally weighed the probability of being pivotal, then we would expect the pattern seen in the CC sample. In the GS data set, on the other hand, the bonus vote choice does not seem to correlate in any transparent manner with the realized margin of victory: the closest election was

[14]The only relevant information in the figure is the ordinal ranking of the elections, and reported scores show the difference across elections more clearly. Normalized scores yield a rescaled but otherwise identical figure.

Alumni Affairs, with a margin of 2 percent, but the fraction of voters who cast the bonus vote in that election while ranking a different one higher is smaller than the fraction in the International election, the most lopsided, with a margin of victory of 28 percent.

In the theoretical model, the bonus vote choice depends on the *common knowledge* of the distributions of valuations in the electorate at large. But we do not have any evidence that preferences were common knowledge. Indeed we have some evidence to the contrary: as mentioned earlier, in the GS questionnaire, electronic and thus faster, we added a question about expected election outcomes. In the election for President, more than 80 percent of the students filling the questionnaire answered the question, but about half of them predicted the wrong winner, and a quarter predicted that she would win by a large margin; in the other two elections, more than half of those filling the questionnaires chose not to answer the question, in line with the large abstention rate, but among those who did respond, a majority predicted the wrong winner in the Alumni Affairs election, half of them predicting a victory by a large margin, and one third did so in the International election.[15] These answers may reflect something other that rational calculations of expected outcomes, but can hardly be read as support for common knowledge of the full distributions of valuations. The lack of information by the voters could explain the anomalies noticed in Figure 3.

If preferences were not common knowledge, we have no basis for testing rigorously the strategic behavior predicted by the theory described earlier. In fact, we have two additional reasons to be cautious. First, we do not know the distribution of valuations for the electorate as a whole: it is the choice of the entire electorate that any single voter is trying to influence, but the distributions of valuations we construct from the survey refer to a sample only — the respondents of the survey. Second, when asking where the respondent would cast the bonus vote, the questionnaire did not state explicitly that in the thought experiment all other voters would also be casting their bonus vote. We doubt that answers would have been different otherwise, but we do not know how the question was interpreted.

For all of these reasons, we limit ourselves to a descriptive exploration of the bonus vote choices in our samples. We will use the reported bonus vote choices later, when we estimate the probable impact of the bonus vote on outcomes, as one of four plausible behavioral rules that voters may follow. In addition to the actual survey responses, it will be useful to have a concise description of the patterns we see in the data. A simple statistical model provides such a description.

[15]The exact numbers are reported in Casella, Ehrenberg, Gelman and Shen (2008).

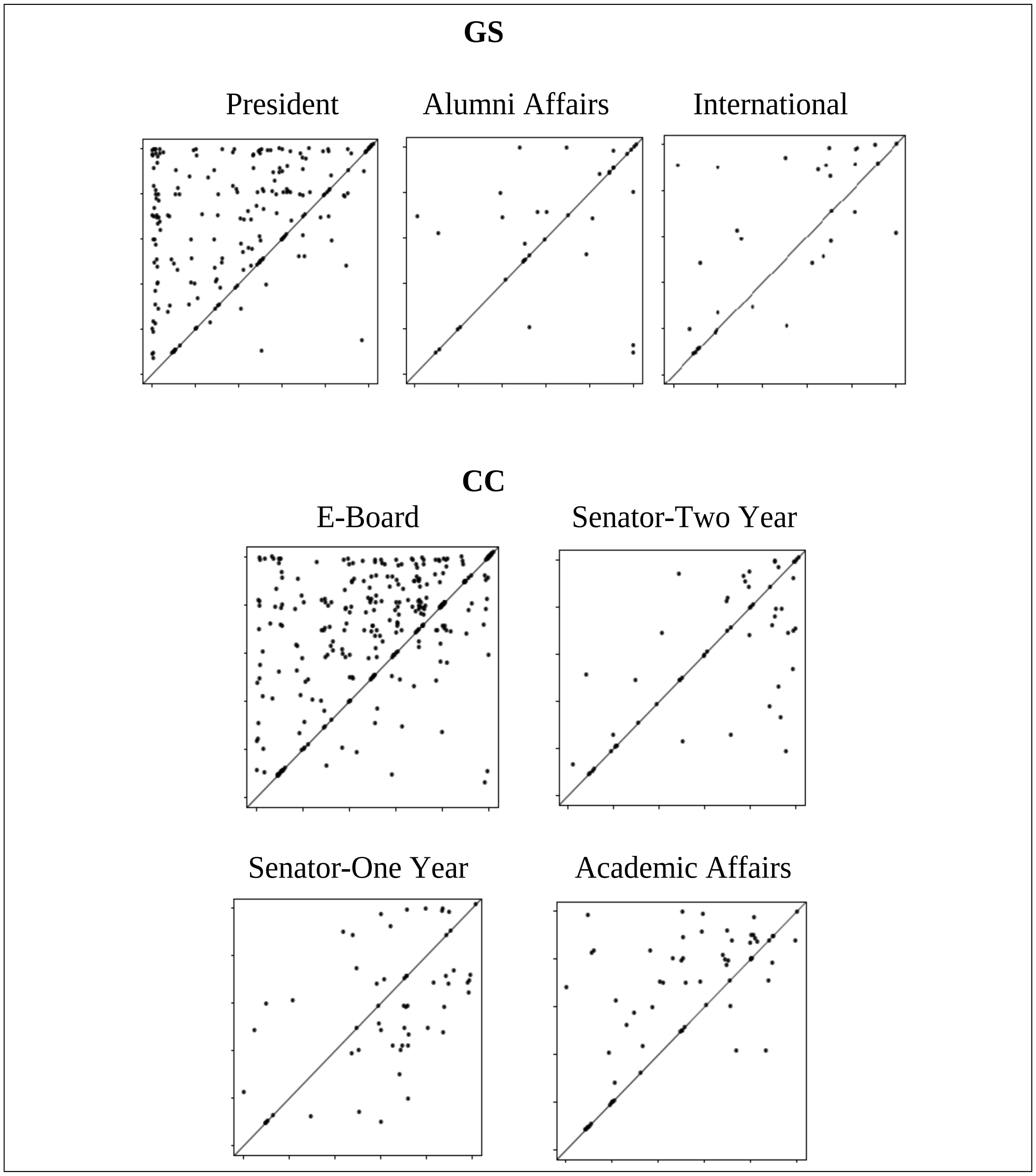


Figure 3: Relative raw score of the election receiving the bonus vote. Each point is an individual response. The vertical axis is the recorded score for the election selected for the bonus vote; the horizontal axis is the highest recorded score among all other elections.

3.5 A simple statistical model

The model describes respondents' bonus vote choices through a set of elementary criteria, each of which specifies how to cast the bonus vote. Each respondent's answer, given his or her scores, can then be written in terms of the probability of following the different criteria. We posit four mutually exclusive criteria: (1) cast the bonus vote in the election with highest score; (2) cast the bonus vote in the most salient election (President for GS, and Executive Board for CC); (3) cast the bonus vote in the closest election; (4) some other criterion we ignore, and such that the choice appears to us fully random. We suppose that each criterion is followed with some probability, which we call *pmax* for criterion 1, *psal* for criterion 2, *pclose* for criterion 3 and *prand* for criterion 4, and we describe a respondent's choice through these probabilities. For example, consider a GS voter whose highest score is on the President election, and who indicates that he would cast the bonus vote on that election. Under our model, this behavior occurs with probability $pmax + psal + (1/3)prand$. If the voter assigns the highest score also to a second election, then the probability of the observed behavior becomes $(1/2)pmax + psal + (1/3)prand$, and correspondingly for the other cases. Assuming that respondents' choices are independent, the likelihood of observing the data set is simply the product of the probabilities of all individual choices. The probabilities *pmax*, *psal*, *pclose*, and *prand* can then be estimated immediately through maximum likelihood or Bayesian methods.

We have estimated the probabilities on the two data sets separately, and in both cases maximum likelihood and Bayesian estimation yield identical results, summarized in Table 1, with standard errors are in parentheses.[16]

Table 1: Population frequencies of three behavioral criteria as estimated from a simple statistical model of observed bonus vote choices (with standard errors in parentheses)

	GS	CC
pmax	0.55 (0.06)	0.52 (0.04)
psal	0.34 (0.06)	0.30 (0.04)
pclose	0.01 (0.01)	0.04 (0.01)

[16] In conducting the Bayesian inference, we assigned uniform prior distributions to each of *pmax*, *psal* and *pclose* and used the software Bugs (Spiegelhalter et al., 1994, 2002). The Bayesian approach works well here because with moderate sample sizes, maximum likelihood estimates can be at the boundary of the parameter space (as discussed in general terms in chapter 4 of Gelman et al., 2004).

In both data sets, *pmax* is highest, reflecting the large frequency with which the bonus vote is targeted to one's highest ranked election. As for the remaining probabilities, *psal*, the probability of casting the bonus vote on the most salient election, is relatively large and significantly different from zero; *pclose*, the probability of casting the bonus vote on the election with smallest margin of victory, is always small and in the GS sample insignificant. These latter results are theoretically puzzling: the intensity of a voter's preferences should be fully captured by the score, and, as mentioned, the use of the bonus vote should be higher in close elections. However, the two most salient elections were also reasonably close elections, the second closest in both data sets, and attracted much more debate and attention prior to voting than any of the others. It is possible that voters chose them disproportionately not because they were salient, but because in fact they knew them to be close contests, and had little information about the other elections.

The statistical model we have posited is a synthetic description of the data, and should not be read as providing a test of the theory. In particular, the probabilities with which the different criteria are followed are treated as exogenous parameters constant across voters, as opposed to being the expression of each voter's optimal strategies, dependent on the voter's full set of scores, as they would in a strategic model.

4 Estimating the Probable Impact of the Bonus Vote

As shown in Figure 1 and stated earlier, in our samples no outcome would have changed with the addition of the bonus vote, given how respondents stated they would cast it. Per se, the observation is not an argument against storable votes: according to our welfare measure, the majority outcomes in the GS elections are all efficient, and the one inefficiency we find in the CC sample is barely detectable. In addition, the fact is not very informative: because the bonus vote links the outcomes of the bundle of elections over which the voters can choose to use it, we have only two independent data points, one for each school, too few to form an estimate of the bonus vote's probable impact. To obtain an estimate of the bonus vote's probable impact, ideally we would want to replicate the same elections many times, with many different electorates whose preferences are all drawn from the same underlying distribution. We cannot rerun the elections, but we can approximate such iterations by bootstrapping our data.[17] The bootstrapping simulations will provide us with an estimate of the impact of the bonus vote in a population for which our samples are representative. Note that the exercise is valid under the

[17] The classical references are Efron, 1979, and Efron and Tibshirani, 1993. For a recent treatment, see Davidson and MacKinnon, 2006.

assumption that the addition of the bonus vote does not alter the voters' preferences between the two candidates of each election. Given the binary choice, this would hold for rational voters, and seems in general quite plausible.

We implement the bootstrap assuming that preferences are independent across individuals, but not necessarily across elections for a single individual. We sample with replacement n individuals from our data sets, where $n = 276$ for GS and $n = 502$ for CC: for each individual, we sample the scores assigned to all elections and the choice of where to cast the bonus vote. A sample then corresponds, for each school, to a distribution of preferences (choice of candidate and score) over each election, and a bonus vote choice for each voter. We replicate this procedure 10,000 times, and for each replication we calculate outcomes and measures of welfare if the elections were held with simple majority voting, or with the bonus vote.

We consider four alternative rules governing the use of the bonus vote. Rule A is closest to the data: we use the bonus vote choice actually made by the individuals sampled in the bootstrapping. However, it is not clear that this is the appropriate rule: individuals' stated choices reflect individuals' beliefs about the electorate as a whole, whereas the elections simulated in the bootstrap concern the samples of respondents only. The most constructive reaction, in our opinion, is to simulate the elections using multiple possible rules governing the use of the bonus vote. Each of the rules we have chosen can be supported or disputed, but it is the robustness of the results across different rules that gives us confidence in their relevance. Rule B applies the statistical model described at the end of the previous section to each bootstrap sample; rule C states that every individual casts the bonus vote in the election with highest score (randomizing with equal probability if more than one election has the highest score); rule D replicates rule C but excludes, as targets of bonus votes, the most lopsided election in each sample, International in GS and Senate-Two year in CC.

The bonus vote can give weight to the intensity of preferences, and thus, given the distributions of preferences of the survey respondents, the hypothesis is that with high probability storable votes should be equivalent to majority voting in the GS elections, but have more impact and better welfare properties in the CC elections. In what follows we describe the analysis based on normalized scores because they seem theoretically somewhat superior. We have verified that the results are not sensitive to using either reported or normalized scores, as is to be expected given Figure 2.

4.1 The frequency of minority victories

The first question is the frequency with which, using the bonus vote, at least one of the elections in each set is won by the minority candidate. As shown in Figure 4, in the GS bootstrap samples the frequency is stable across the different rules at around 15 percent. In the CC samples, the frequency increases monotonically as we move from rule A to rule D: from 11 to 15 to 23 to 29 percent

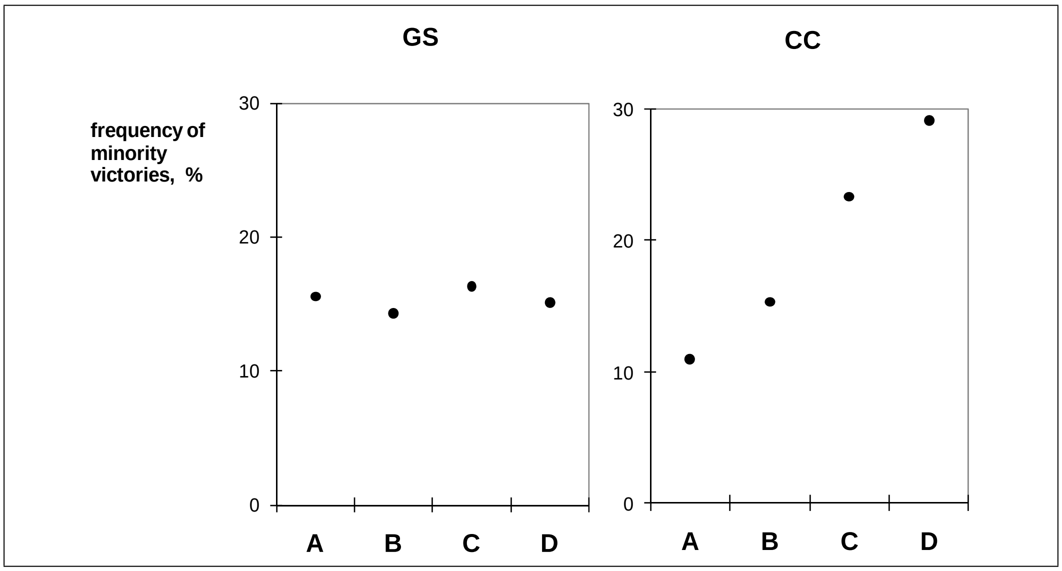


Figure 4: Percentage frequency with which at least one election in each set is won by the minority, out of 10,000 bootstrap samples, when voters cast their bonus vote according to each of four different rules.

The aggregate number, appropriate because of the linkage across elections imposed by the bonus vote, is the result of different impacts on the specific elections. Tables 2(a) for GS and 2(b) for CC report the frequency with which any individual election is won by the minority, for each bonus vote rule. The last column reports the aggregate frequency depicted in the figure. (In all cases, we assigned a weight of 1/2 to any sample where either simple majority or the bonus vote resulted in a tie).

Predictably, the numbers reflect the distributions of preferences over each election. The most lopsided elections (International Representative in GS, and Executive Board and Senate-Two Year in CC) are never won by the minority candidate. The election most susceptible to the impact of the bonus vote is Academic Affairs

Table 2: Frequency of minority victories in 10,000 bootstrap samples, based on four different assignment rules for each of the elections in our study

(a) GS

	President	Alumni	International	Aggregate
Rule A	5%	11%	0	16%
Rule B	6%	9%	0	14%
Rule C	8%	9%	0	16%
Rule D	7%	8%	0	15%

(b) CC

	Exec Board	Senate-Two	Senate-One	Academic	Aggregate
Rule A	0	0	2%	9%	11%
Rule B	0	0	1%	14%	15%
Rule C	0	0	1%	23%	23%
Rule D	0	0	1%	28%	29%

in CC; moving from rule A to rule D the number of bonus votes cast on Academic Affairs progressively increases; so does the frequency with which the minority candidate wins, and so does the aggregate frequency.[18] In the GS samples, the impact of the bonus vote reflects mostly how close the different contests are. In the CC samples, again the bonus vote affects exclusively the two closest elections, but the higher intensity of the minority supporters in Academic Affairs plays a clear additional role in tilting the results in the minority's favor when the bonus vote is available.

It is instructive to compare these numbers to the frequency with which a minority victory is efficient, in the different elections. In the GS samples, the frequency is 12 percent for President,14 percent for Alumni, and 0 for International Representative. In the CC samples, it is 0 for both Executive Board and Senate-Two Year,1 percent for Senate-One Year and 40 percent for Academic Affairs. The bonus vote brings the frequency of minority victories closer to efficiency, although still falling short. Whether minority victories occur in the same samples in which they are in fact efficient is the question addressed in the next subsection.

[18]In CC, rule B assigns some bonus votes to the closest election. In estimating the statistical model, we identified as closest the election with the lowest margin of victory in the full electorate. In the bootstrap exercise, we must limit ourselves to the sample. Taking into account the bonus votes, Senate-One Year is the closest election in 26 percent of the bootstrap samples, and Academic Affairs in 72 percent. The remaining 2 percent of bootstrap samples are ambiguous — either election can be closest, depending on how the bonus votes are cast. The ambiguity does not affect the frequencies reported in Table 2(a).

We used 10,000 bootstrap samples to guarantee that the experimental error (the binomial standard error of the observed frequencies due to the finite size of the sample) is negligible. As a test, we recalculated the frequencies in Tables 2(a) and 2(b) with 20,000 independent bootstrap samples, confirming that all frequencies changed by less than half of one percent.

Our first conclusion then is that in both sets of data and for all rules, although particularly in CC for rules C and D, the bonus vote allows the minority to win with substantial probability.

4.2 The impact of minority victories on aggregate welfare

How costly are minority victories in terms of aggregate welfare? Our second result is that minority victories not only are not costly, but in fact typically come with aggregate welfare gains.

In each election, realized welfare is defined as the sum of all (normalized) scores of all voters who supported the winner. With the bonus vote, the relevant unit is again the full set of elections in each school, for each bootstrap sample. More precisely, if M_k is the set of voters whose favorite candidate commands a majority of regular votes in election k, then welfare with majority voting, W_M is defined as $W_M = \sum_k \sum_{i \in M_k} s_{ik}$ where s_{ik} is the score assigned to election k by voter i. Similarly, welfare with storable votes W_{SV} is defined as $W_{SV} = \sum_k \sum_{i \in \mathscr{M}_k} s_{ik}$ where $\mathscr{M}_k$ is the set of voters whose favorite candidate commands a majority of all votes, including bonus votes, in elections k. We select all bootstrap samples where the minority wins at least one election: in each set of elections there exists at least one k for which $\mathscr{M}_k \neq M_k$. Among those samples only, the upper panel of Figure 5 shows the frequency with which $W_{SV} > W_M$, and the lower panel shows the mean percentage welfare change (mean $(W_{SV} - W_M)/W_M$).[19]

For all four rules in the CC data and for three of the four in the GS data the mean effect on welfare is positive. In the CC data, the result is strong: among all bootstrap samples where the minority wins at least one election, an increase in aggregate welfare occurs from a minimum of 86 percent of the time (rule A) to a maximum of 96 percent (rule C). The minority must be winning when its preferences are more intense than the majority preferences. In the GS data, the mean welfare change is always smaller than in CC, and in one case, when it is smallest, the sign is negative (rule A); with rules B, C and D the expected impact is positive, but the frequency with which aggregate welfare rises reaches a maximum at 72 percent (rule D).

[19] In the case of ties, we assign a probability of 50 percent to the victory of either side.

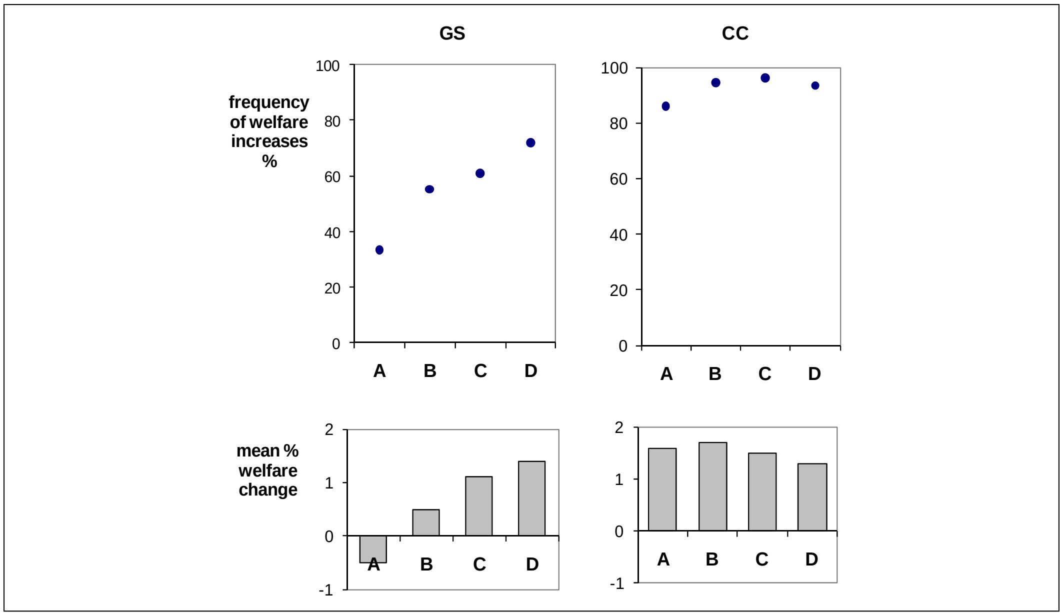


Figure 5: Percentage frequency of aggregate welfare increases and mean percentage welfare change, relative to majority voting, when the minority wins at least one election.

In addition to generating the summary results reproduced in Figure 5, the bootstrap exercise provides us with the full distribution of the impact of the bonus vote on aggregate welfare across all bootstrap samples. Figure 6 presents the histograms of the percentage difference in aggregate welfare between each of the four bonus vote rules and majority voting, in all samples where, using the bonus vote, the minority wins at least one election in each set. The frequency of positive welfare changes reported above, in the upper panel of Figure 5, corresponds to the positive mass in each of the histograms. The histograms show clearly the concentration of the probability mass around the positive mean welfare change in the CC samples, while the distribution is more dispersed in the case of GS. The higher variability of the GS results is consistent with approximately symmetric preferences across the two sides in elections affected by the bonus vote. Summary measures of dispersion — standard errors and bootstrap confidence intervals — are obtained easily; we omit them here, preferring to present the full distributions. Because our focus is on samples where the minority wins at least one election, the number of samples analyzed here is between 1,400 and 1,600 for GS and between 1,000 and 3,000 in CC, depending on the bonus vote rule (see Table 2). We have compared the results reported here to those obtained from a different set of 10,000 bootstrap samples, confirming that the experimental error remains negligible.

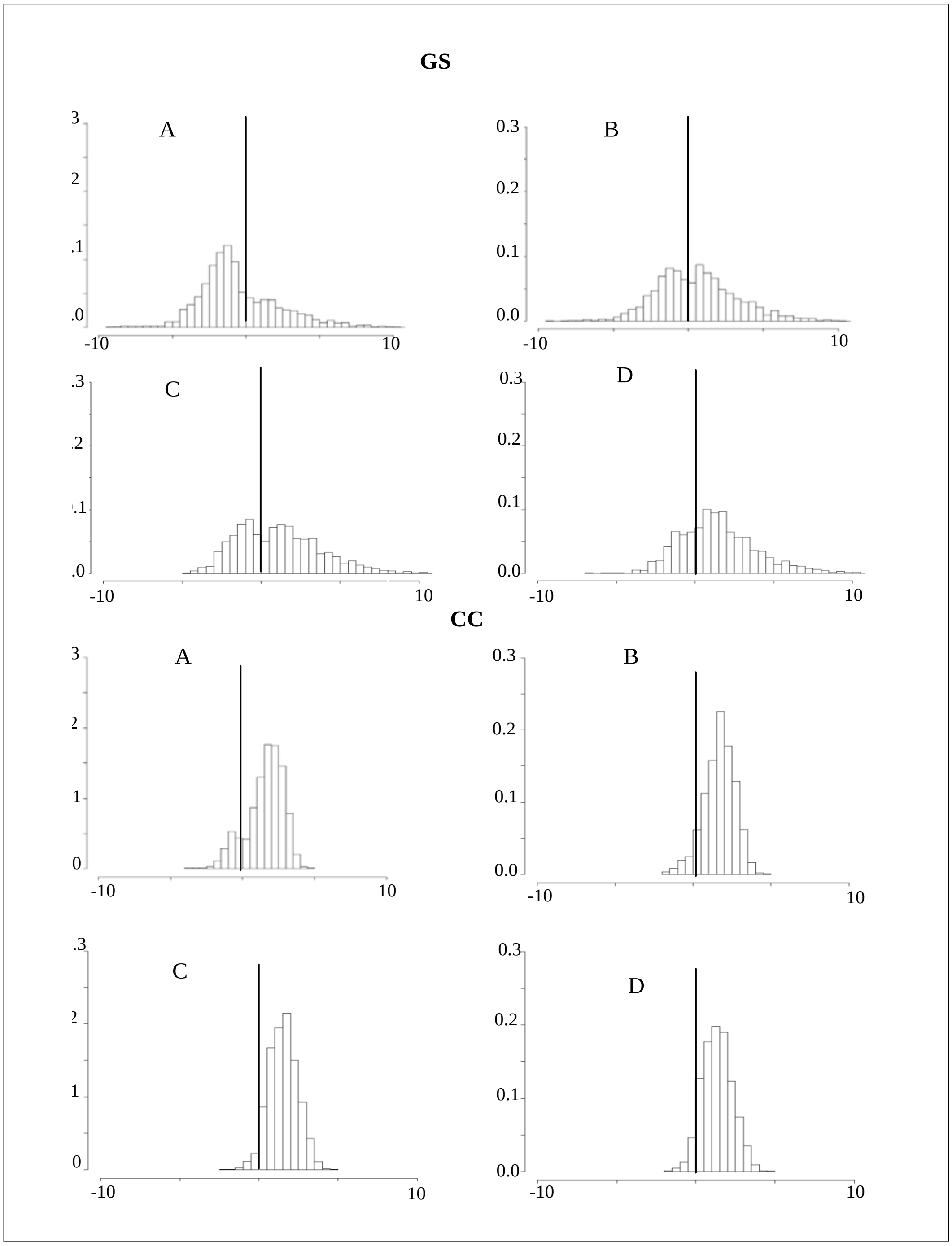


Figure 6: Histograms of percentage welfare differences between the bonus vote scheme and majority voting in all samples where, with the bonus vote, the minority wins at least one election.

4.3 Inequality

Minorities should win occasionally not only because aggregate efficiency is higher when intensity of preferences is recognized, but also because preferences can be correlated across elections, and individuals who find themselves repeatedly on the minority side will feel unrepresented by the political system. If it comes at no cost to aggregate welfare, a more equal distribution of decision-making influence seems desirable in itself. Our data allow us to construct the full distribution of ex post utility across voters in each of the bootstrap samples, where individual utility is defined as the sum of the voter's scores over all elections won by the voter's preferred candidate. Formally, recalling that $\mathscr{M}_k$ is the set of voters whose favorite candidate commands a majority of all votes, including bonus votes, in elections k, ex post utility of voter i, U_i, is defined as $U_i = \sum_{k|i \in \mathscr{M}_k} s_{ik}$. The distribution across voters thus captures the different frequencies with which each voter's preferred candidate wins, weighted by the relative importance the voter assigns to each election. If instead of summing over all elections for which i belongs to $\mathscr{M}_k$, we sum over all elections for which i belongs to M_k — the set of voters whose favorite candidate commands a majority of regular votes in election k — we obtain for comparison ex post utility with majority voting. We can then calculate the impact of the bonus vote on voters' ex post inequality.

Average realized utility distributions are shown in Figure 7. The bins on the horizontal axis are ordered in increasing levels of utilities, each bin corresponding to a 5 percent range, from lowest (0) to highest (1). The height of each bin is the fraction of voters with utility falling into the bin, averaged over all samples where, with the bonus vote, the minority wins at least one election. The black line corresponds to the bonus vote (rule A), and the grey line to simple majority. Here rule A is fully representative of the results obtained with the other bonus vote rules. The higher polarization of the GS distributions reflects the smaller number of contests (3, as opposed to 4 in CC) and the overwhelming priority assigned to a single election (President). The impact of the bonus vote on the distribution is hardly detectable in the GS samples but is clear in the CC sample, where the frequency of realizations at the two ends of the distribution is reduced in favor of larger mass in the center.

Figure 8 reports the impact of the bonus vote on voters' inequality as measured by the Gini coefficient. In the CC data, inequality declines unambiguously: over all bootstrap samples where the bonus vote allows the minority to win at least one election, inequality declines from a minimum of 90 percent of the samples (rule A) to a maximum of 98 percent (rule D). Across all bonus vote rules, the mean percentage decrease in the Gini coefficient is 12–13 percent. In the GS data, on the

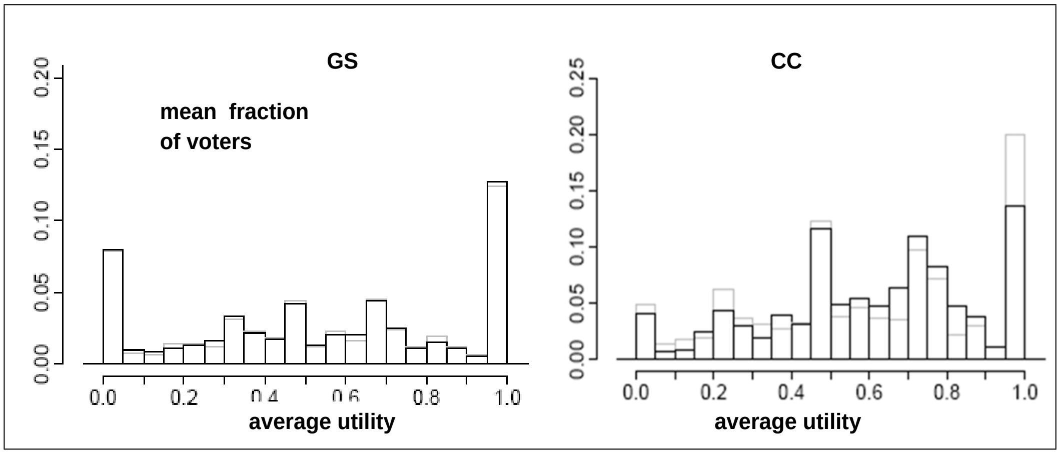


Figure 7: Realized utility distribution, simple majority(grey) and bonus vote, rule A, averaged over all samples in which the minority wins at least one election.

other hand, the bonus vote's effect on inequality is of inconsistent sign and small magnitude. The Gini coefficient declines barely more than 50 percent of the times with rules B, C and D, and barely less with rule A. The effects are quantitatively small: the mean percentage change in the coefficient is an increase of 2 percent.

The different results reflect the different roles of the bonus vote in the two data sets. In the GS data, the elections affected by the bonus vote (President and Alumni Affairs) are particularly close, with little difference in volume and intensity of support between the two candidates. The side benefiting from the bonus vote is not consistent across bootstrap samples, and thus the average impact of the bonus vote on equality of representation is small and similarly non consistent. In the CC data, on the other hand, the main role of the bonus vote is to overthrow the majority victory in the Academic Affairs election, on the strength of the more intense minority preferences. The distribution of preferences is sufficiently asymmetric to ensure that in the bootstrap samples the bonus vote consistently helps the same candidate. And because the minority in the Academic Affairs election is also disproportionately on the minority side in the other elections, the final result is a positive, sizeable improvement in equality.[20]

[20]The precise numbers reflected in Figure 8 are reported in the Appendix, together with 95 percent confidence intervals. An alternative measure of inequality — the ratio of the average utility of individuals at the bottom 20 percent of the utility distribution to average utility — yields identical conclusions, and is also reported in the Appendix, see Table A.2 .

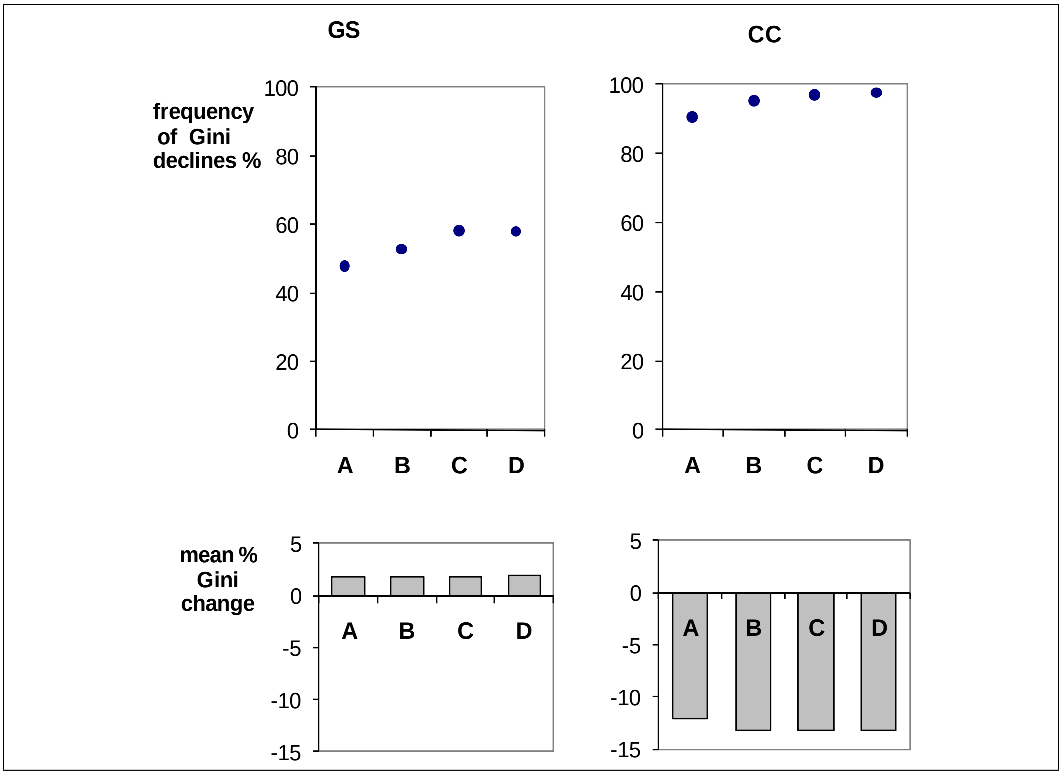


Figure 8: Percentage frequency of more egalitarian distributions of realized utility, and mean percentage decline in the Gini coefficient, relative to majority voting, when the minority wins at least one election.

5 Conclusions

Storable votes are a simple mechanism with the potential to improve fairness with little if any cost, in fact with some gains, to efficiency. In our data, they perform well: when the margins of victory are small, they allow the minority to win occasionally, with no substantive effects on welfare and inequality if the two sides have preferences of similar intensity, but with improvements in both if minority preferences are particularly intense. Storable votes should be tested in higher stake contests, particularly in referendums, for which they are expressly designed. Using a new voting system in real elections is, appropriately, difficult to do, and assembling an experimental data set rich enough to evaluate the system's performance statistically is probably impossible. Survey methods, coupled with bootstrap resampling, are a more practical route. One of the goals of this work is to suggest that the bootstrap methodology is ready to be exploited, easily and cheaply.

A Appendix

Table A.1: Testing the representativeness of the samples. Comparison of abstention rates and margins of victory in and outside the samples. The star indicates that the hypothesis that both the rest of the electorate and our samples are random draws from a common population is rejected at the 5 percent confidence level. Both in the electorate as a whole and in our samples: (i) abstention rates are calculated for each election for students who voted at least once, and thus do not reflect the fraction of students who did not take part in voting at all; (ii) the share of votes for the winner is calculated among voters in each election, and thus ignores abstentions.

(a) GS

	Abstentions			Share of votes for winner		
	non-sample	sample	diff.	non-sample	sample	diff.
President	0.21	0.02	0.19*	0.63	0.53	0.10*
Alumni Affairs	0.54	0.26	0.28*	0.59	0.48	0.11
Int'l Repr.	0.46	0.23	0.24*	0.63	0.65	-0.02

(b) CC

	Abstentions			Share of votes for winner		
	non-sample	sample	diff.	non-sample	sample	diff
Exec Board	0.08	0.02	0.06*	0.57	0.60	-0.03
Senate-1	0.15	0.12	0.03	0.69	0.73	-0.04
Senate-2	0.13	0.08	0.05*	0.56	0.55	0.01
Academic Affs.	0.18	0.15	0.03	0.65	0.54	0.11*

Table A.2: Measures of inequality when the minority wins at least one election. Because we are focussing on bootstrap samples where majority voting and the bonus vote lead to different outcomes, to each bonus vote rule corresponds a slightly different data set. The values reported in the table for majority rule correspond to bootstrap samples where majority and Rule A lead to different outcomes (the values in the other cases are almost identical)

(a) GS

Gini Coefficient

	Maj	Bv A	Bv B	Bv C	Bv D
Mean	0.347	0.353	0.353	0.353	0.354
95% CI	(0.20, 0.40)	(0.21, 0.40)	(0.21, 0.40)	(0.22, 0.40)	(0.22, 0.39)
Prob bv lower (%)		47.9	52.8	58.2	58

Ratio of average utility of bottom 20% to sample average

	Maj	Bv A	Bv B	Bv C	Bv D
Mean (%)	9.08	7.89	7.44	7.14	6.78
95% CI	(0, 47.1)	(0, 45.0)	(0, 44.9)	(0, 43.7)	(0, 43.1)
Prob bv higher (%)		52.5	51.2	53.5	49.4

(b) CC

Gini Coefficient

	Maj	Bv A	Bv B	Bv C	Bv D
Mean	0.280	0.246	0.243	0.243	0.243
95% CI	(0.24, 0.30)	(0.22, 0.29)	(0.22, 0.28)	(0.22, 0.27)	(0.22, 0.27)
Prob bv lower (%)		90.5	95.3	97.1	97.7

Ratio of average utility of bottom 20% to sample average

	Maj	Bv A	Bv B	Bv C	Bv D
Mean (%)	27.27	33.88	34.42	34.42	34.38
95% CI	(22.5, 37.7)	(24.9, 41.8)	(25.9, 41.5)	(27.2, 41.0)	(27.4, 41.2)
Prob bv higher (%)		88.6	94.0	96.6	97.0

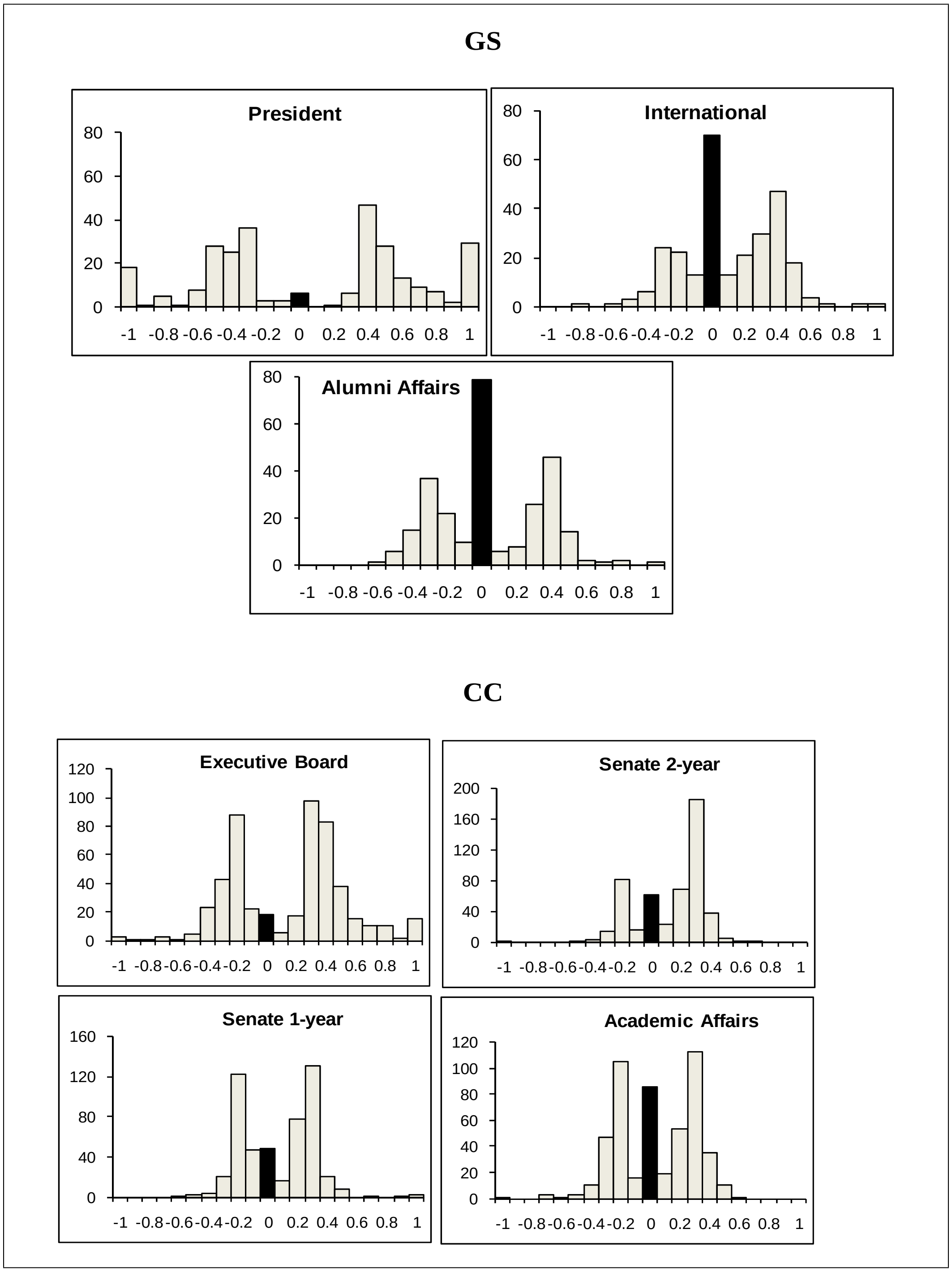


Figure A.1: Histogram of normalized scores.

Columbia University Economics Department **Voting Experiment**

Thank you for choosing to participate in this experiment. By entering your number below, and attaching your post-it to the questionnaire you acknowledge that you have read, and understand the consent form (attached to the ballot box). Please enter your number here:

A lottery will be conducted among all participants who have followed the instructions of the experiment and written the number above legibly. The winner will receive an iPod.

This experiment asks you how much you care about the outcome of some of the elections.

Each election can be scored on a scale of 1-10, with 10 meaning "care very much" and 1 meaning "care not at all". In order to indicate a box, please cross out the relevant number with an X like this:

9

E-Board

1	2	3	4	5	6	7	8	9	10

Senate (1 year)

1	2	3	4	5	6	7	8	9	10

Senate (2 year)

1	2	3	4	5	6	7	8	9	10

Academic Affairs

1	2	3	4	5	6	7	8	9	10

Suppose now that, in addition to the regular votes you cast earlier, you had 1 additional vote to cast in favor of your candidate in one of the elections. You can choose any of these four elections as you see fit. Which one would you choose? (NOTE: This is for experimental purposes only. Your answer will **not** change the outcome of the actual elections in any way).

Please check the box under the election you choose. Please check **only one** box.

E-Board	**Senate (1 year)**	**Senate (2 year)**	**Academic Affairs**
Evolution	David	Tiffany	Ehizoje
Open Columbia	Eric	Gerry	Alidad
☐	☐	☐	☐

Thank you for participating. And good luck with the lottery!

Figure A.2: Instructions

References

Bowler, S., T. Donovan, and D. Brockington (2003): *Electoral reform and minority representation: local experiments with alternative elections*, Ohio State University Press.

Casella, A. (2005): "Storable votes," *Games and Economic Behavior*, 51, 391–419.

Casella, A. (2010): "Agenda control as a cheap talk game: Theory and experiments with Storable Votes," *Games and Economic Behavior*.

Casella, A., S. Ehrenberg, A. Gelman, and J. Shen (2008a): "Protecting Minorities in Binary Elections: A Test of Storable Votes Using Field Data," .

Casella, A. and A. Gelman (2008): "A simple scheme to improve the efficiency of referenda," *Journal of Public Economics*, 92, 2240–2261.

Casella, A., A. Gelman, and T. R. Palfrey (2006): "An experimental study of storable votes," *Games and Economic Behavior*, 57, 123–154.

Casella, A., T. R. Palfrey, and R. Riezman (2008b): "Minorities and Storable Votes," *Quarterly Journal of Political Science*, 3, 165–200.

Cox, G. W. (1990): "Centripetal and Centrifugal Incentives in Electoral Systems," *American Journal of Political Science*, 34, 903.

Dahl, R. A. (1989): *Democracy and its critics*, Yale University Press.

Dahl, R. A. (2006): *A preface to democratic theory*, University of Chicago Press.

Davidson, R. and J. MacKinnon (2006): *Bootstrap Methods in Econometrics*, Palgrave Macmillan Ltd, Hound- mills, Basingstoke, Hampshire RG61 6XS, UK., chapter 25.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, 7, 1 – 26.

Efron, B., R. Tibshirani, and R. J. Tibshirani (1993): *An introduction to the bootstrap*, Chapman & Hall.

Gelman, A. (2004): *Bayesian data analysis*, Chapman & Hall/CRC.

Gerber, E. R. (1999): *The populist paradox: interest group influence and the promise of direct ...*, Princeton University Press.

Gerber, E. R., R. B. Morton, and T. A. Rietz (1998): "Minority Representation in Multimember Districts," *The American Political Science Review*, 92, 127.

Guinier, L. (1995): *The tyranny of the majority: fundamental fairness in representative democracy*, Free Press.

Hortala-Vallve, R. (2007): "Qualitative voting," *Department of Economics Discussion Papers Series (Oxford University)*, 320.

Hortala-Vallve, R. and A. Llorente-Saguer (2010): "A simple mechanism for resolving conflict," *Games and Economic Behavior*.

Issacharoff, S., P. S. Karlan, and R. H. Pildes (2008): *The Law of Democracy: Legal Structure of the Political Process, 2008 Supplement*, West Group.

Jackson, M. O. and H. F. Sonnenschein (2007): "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75, 241–257.

Matsusaka, J. G. (2004): *For the many or the few: the initiative, public policy, and American democracy*, University of Chicago Press.

Pildes, R. and K. Donoghue (1995): "Cumulative Voting in the United States," *U. Chi. Legal F.*, 241–314.

Spiegelhalter, D., A. Thomas, N. Best, and W. Gilks (1995): "BUGS - Bayesian inference Using Gibbs Sampling Version 0.50," .