Games and Economic Behavior 113 (2019) 461–481

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

The dark side of the vote: Biased voters, social information, and information aggregation through majority voting ☆

Rebecca B. Morton [a,*], Marco Piovesan [b], Jean-Robert Tyran [c,b]

[a] Department of Politics, NYU NYC and NYU Abu Dhabi, 19 West 4th Street, 2nd Floor, New York, NY 10012, United States of America
[b] Department of Economics and Center for Economic Behavior and Inequality (CEBI), University of Copenhagen, Øster Farimagsgade 5, 1353, Copenhagen, Denmark
[c] Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria

ARTICLE INFO

Article history:
Received 29 March 2015
Available online 3 November 2018

JEL classification:
C92
D7
D02
D03

Keywords:
Voting
Information aggregation
Cognitive biases
Wisdom of crowds
Social information

ABSTRACT

We experimentally investigate information aggregation through majority voting when voters are biased. In such situations, majority voting can have a "dark side," that is, result in groups making choices inferior to those made by individuals acting alone. In line with theoretical predictions, information on the popularity of policy choices is beneficial when a minority of voters is biased, but harmful when a majority is biased. In theory, information on the success of policy choices elsewhere de-biases voters and alleviates the inefficiency. However, in the experiment, providing social information on success is ineffective and does not sufficiently de-bias voters.

1. Introduction

One of the benefits of having democratic choice is the ability of voting to aggregate dispersed information in society. The argument, going back to de Condorcet (1785), is simple: if each voter's judgment is more likely to be right than wrong, the collective choice in a majority vote is going to be better (more likely to be right) than the average judgment of individuals acting alone. This is what we call the "bright side" of the vote. The argument applies in situations in which a "right" policy exists, voters have a common interest to implement the right policy, but all voters are uncertain about which policy is right. Examples of common interest situations are numerous. We all generally agree that we want to prevent financial crises or mass unemployment and that we would like to have world peace and an unpolluted environment. But we disagree over which policies are best to reach those goals.

☆ We gratefully acknowledge financial support by the Austrian Science Fund (FWF) under project no. S 12027-G16 and the Department of Economics, University of Copenhagen. We also thank Niels Bruhn Christensen, Jakob Muller Karstoft, Lars Reenberg, and Michael Schroder for their research assistance with this project and participants at presentations of earlier versions of this paper at the 2011 Piorat Political Economy Conference, Nuffield College Oxford Experimental Social Sciences Seminar, Queensland University of Technology, and the University of Queensland. All errors remain the responsibility of the authors.

* Corresponding author.
E-mail addresses: rebecca.morton@nyu.edu (R.B. Morton), marco.piovesan@econ.ku.dk (M. Piovesan), jean-robert.tyran@univie.ac.at (J.-R. Tyran).

https://doi.org/10.1016/j.geb.2018.10.008
0899-8256/

Nevertheless, the argument that in such situations group choices can be better than individual ones is based on various simplifying assumptions. We theoretically and experimentally address two key assumptions and what they imply for a "dark side" of the vote to exist. The first assumption is that all voters are more likely to be right than wrong when judging a particular issue. While the standard approach to information aggregation allows for some voter uncertainty about what is the right policy, it assumes that voters' judgments are not systematically mistaken. Yet, mounting evidence suggests that people may be biased in some instances (e.g. when making judgments about risky prospects), and in some cases a majority of voters may be biased. We provide a simple game-theoretic model in which voters vary in their competence in making inferences as a basis for our experiment. The model predicts both what we call a "bright side" and a "dark side" of the vote. Voting is beneficial when a majority is unbiased, but harmful when not. That is, the decision made by majority rule can be worse than that made by an average individual acting alone when a majority of voters make incorrect inferences.

The second simplifying assumption we address is that voters form their judgments independently. However, voting is often preceded by debate and flows of social information (as in opinion polls, news reports, and surveys) which may affect voters' judgments in similar ways. For example, voters might learn how popular some choices are in other countries, subnational regions, or localities, but not whether the choices are successful or not. Alternatively, voters might learn how happy individuals in other countries, subnational regions, or localities are with their overall collective choices, but not the specifics of the choices that these voters have made. The consequences of such social information are ambivalent in theory and practice. In general, social information may undermine the efficiency of information aggregation or strengthen it.

We study two types of social information: voters either learn about other voters' opinions (i.e. how popular a particular policy is, as in an opinion poll) or they learn about how successful other, very similar, electorates were in making decisions on a particular topic (but not what exact policy they implemented). Our model predicts that the effects of social information depend on whether a majority of voters is biased or not. If a majority makes correct inferences on average, social information tends to be beneficial. Specifically, social information about previous success does no harm, and social information about opinions improves the informational efficiency of voting. However, when a majority of voters makes biased judgments, providing social information may help or harm informational efficiency. In this case, our model predicts that social information on opinions makes matters worse (further reduces informational efficiency) but social information about success improves matters. The reason for this beneficial effect, i.e. for "brightening up the dark side", is that social information on success "de-biases" voters. Intuitively speaking, when a voter learns that other (similar) groups got it all wrong, the voter will (rationally) reconsider his views and vote against his earlier judgment (or prejudice in that case). The reason is that he knows he is most likely similar to these other voters and therefore his original judgment is likely to be wrong, too.

In the experiment, we find support for all of these predictions, with one important exception. We find support for the "bright side" of the vote (i.e. voting is productive when a majority of voters is more likely to be right than wrong), and for a beneficial effect of social information (information on opinions improves efficiency, information on success has no effect). We also find that the "dark side of the vote" is real. When voters are more likely to get it wrong than right, voting is counterproductive. And providing social information on the popularity of policies can make matters even worse. Our first major contribution is to provide an experimental measurement of the dark side and the effect of social information on opinions on that measurement.

Our second major contribution is that we find that, in contrast to theoretical predictions, social information on success has no clear de-biasing effect in our experiment. With reference to a measure of cognitive ability, we discuss to what extent this result is driven by cognitive limitations and the higher level of reasoning required for de-biasing to be successful. We find evidence that cognitive limitations explain the tendency to make incorrect choices and that those with higher cognitive abilities are slightly better able to interpret social information.

While the consequences of biases have been studied extensively for market outcomes (e.g. Ganguly et al., 2000; Gneezy et al., 2003, and Fehr and Tyran, 2005), we are, to the best of our knowledge, the first to experimentally investigate and measure the consequences of incorrect inferences for information aggregation in majority voting (see Kerr et al., 1996 for a general discussion). Our paper is related to a long stream, starting with Shaw (1932), of experimental studies investigating the ability of individuals vs. groups in making correct choices (e.g. Blinder and Morgan, 2005 and Slembeck and Tyran, 2004) but these studies do not focus on majority voting. For a survey, see Charness and Sutter (2012).

Our results are also related to the large literature on the extent that social information might lead to "herding" and inefficient choices. Bikhchandani et al. (1998) demonstrate how social information concerning previous decisions of others can lead to inefficient information aggregation as individuals ignore their own information and follow others' choices. On the other hand, Estlund (1994) demonstrates how social information can lead to information aggregation through voting. Hung and Plott (2001) provide experimental results showing how inefficient herding may occur in private decision-making but that social information through sequential voting can lead to more efficient group choices. Others have considered whether similar herding and information cascades can occur when voting is sequential within a given election (for experimental studies, see Morton and Williams, 1999; Hung and Plott, 2001; Battaglini et al., 2007). When voting in the same election, theory predicts that concerns about pivotality in the electorate can make herding less likely than in private decision-making. In contrast to these earlier voting studies, our voting groups are making independent choices and thus are more similar to the study of information cascades in private decision-making. Our treatment in which voters receive social information on the choices of previous groups (but not whether they were successful or not), can be seen as an extension of experiments testing such "herding" in private individual decision-making to sequential choices by independent groups.

Section 2 of the paper presents the model and section 3 explains how our experimental design tests the predictions of the model. Section 4 presents the experimental results and section 5 provides some concluding remarks.

2. A model of voting with incorrect inferences and social information

2.1. Basic setup and the voting equilibrium

Our model and experiment build on existing work on information aggregation through voting.[1] We consider a voting game with an odd number of participants, $n \geq 3$. The number of participants is common knowledge. Participants vote for one of two options, a or b (abstention is not allowed) in a majority rule election j.[2] The option that receives a majority of the votes in election j is declared the winner in that election (an odd number of votes with no abstention implies that there is always a clear winner). There are two states of the world A and B for each election, which occur with equal probability and are independent of the state of the world in other possible elections. In each election voters have homogeneous preferences. That is, all voters have the same utility function. We normalize voters' utility from election j to equal 1 if either option a is selected in state of the world A or b is chosen in state of the world B, and 0 otherwise.

Before election j occurs, voter i receives an imperfect signal of the world, $\sigma_{ij} \in \{a, b\}$. Define $p_i^j \in [0, 1]$ as the probability that voter i in election j receives an a signal when the state of the world is A and a b signal when the state of the world is B. We call p_i^j voter i's signal quality in election j. Voters do not know their true signal quality for election j when they vote or the true signal qualities of other voters in election j. Importantly, we assume that signals can be on average incorrect for a given election j; that is, we allow for $0 \leq p_i^j < 0.5$, such that an a signal implies that it is more likely that the state of the world is B than it is A. This assumption has not received much attention in the theoretical or experimental literature so far (see Bottom et al., 2002 for an exception). The reason might be that (in a context with 2 alternatives) voters need to be both biased and not aware of their bias for voters' biases to be consequential (otherwise they would just vote counter to their signal, which we discuss later in some depth). Interestingly, this possibility has been considered by Condorcet:

> "In effect, when the probability of the truth of a voter's opinion falls below $\frac{1}{2}$, there must be a reason why he decides less well than one would at random. The reason can only be found in the prejudices to which this voter is subject."[3]

Define p_i as the mean signal quality of voter i across elections, i.e. the expected value of p_i^j holding i constant, but varying j; p^j as the mean signal quality across voters in a single election j, i.e. the expected value of p_i^j, holding j constant and varying i; and p as the mean signal quality across voters and elections (varying both i and j). We assume that the p_i^j are drawn from voter-specific distributions with constant variances, $p_i \in [0, 1]$. Hence, voters may vary in the distribution of their signal qualities such that some may have greater mean signal qualities across elections than others. Some voters may expect that their information is on average positively informative ($p_i > 0.5$), others may expect that their information is on average negatively informative ($p_i < 0.5$), and some may expect that on average their information is not useful in either direction, that is uninformative ($p_i = 0.5$).

One way to visualize the game facing voters is to imagine that there are two jars, red and blue, which represent the two states of the world as in the typical Condorcet jury experiment. First one of the jars is chosen to be the "true jar." In the typical experiment the jars each contain a mixture of red and blue balls with the red jar holding more red balls than the blue jar and vice-versa. Voters receive their signals by drawing a ball from the true jar. Hence, if the true jar is the red jar, voters are more likely to draw a red ball and receive a red signal. In our game we are essentially positing that each voter has an individualized jar whose color matches the state of the world, so if the state of the world is red, then each voter has their own red jar and vice-versa. However, the mixture of balls across voters' jars varies such that some voters may even have a red jar with more blue balls than red balls and vice-versa. Thus, it is possible for some voters to be more likely to receive blue signals when the true jar is red and vice-versa.

The predictions for the voting game with fully rational voters without social information in a particular election j are straightforward. We provide a detailed derivation of this result in Auxillary Materials Appendix A. There we restrict our analysis to pure-strategy symmetric equilibria, in which all voters who receive the same signal and have the same expectations concerning their average signal quality use the same strategy. In solving for the voting equilibria, we assume that voters condition their vote choice on being pivotal. We demonstrate that, assuming voters do not use weakly dominated

[1] For game theoretic studies of the Condorcet Jury problem see Austen-Smith and Banks (1996), Wit (1998), McLennan (1998), Feddersen and Pesendorfer (1998), and Coughlan (2000). Experimental studies include Ladha et al. (1996), Guarnaschelli et al. (2000), Bottom et al. (2002), Ali et al. (2008), and Mechtenberg and Tyran (2016).

[2] Arguably allowing absention might mitigate some of the effects we observe if voters with low signal qualities are more likely to abstain from participation as found in previous research (Battaglini et al., 2008, 2010 and Morton and Tyran, 2011). However, in situations that are similar to our experiment where subjects are not a priori asymmetrically informed as in studies about voting with endogenous information acquisition and abstention, voters typically vote even when uninformed (Elbittar et al., 2018 and Großer and Seebauer, 2016).

[3] See de Condorcet (1785), cited after Baker (1976), p. 62.

strategies, a unique equilibrium exists in which voters sincerely vote their signals if their $p_i > 0.5$, vote opposite their signals if $p_i < 0.5$, and randomize if $p_i = 0.5$.

Arguably, the voters most relevant to our analysis are those for whom $p_i > 0.5$. These voters expect that their information is positively informative and thus are more likely to vote incorrectly when for a particular election, their information is negatively informative, i.e. when $0 \leq p_i^j < 0.5$, the type of biased voting discussed above. But the other types of voters are also interesting. Voters who expect their information is uninformative should ignore their signals and always randomize their choices. Voters who expect that their information is negatively informative, if rational, are sufficiently self-aware to understand that they are so biased such that they actually act to counter to those biases, voting the opposite of their inclinations. They are most likely to vote "incorrectly" when their information is actually positively informative, i.e. when $0.5 < p_i^j \leq 1$. But a self-aware voter i for whom $p_i < 0.5$ is theoretically indistinguishable in his or her behavior (i.e. likelihood of voting correctly) from a voter k with an expected value of $p_k = (1 - p_i) > 0.5$. Moreover, such self-aware voters are also empirically indistinguishable if we only observe the correctness of their voting behavior, but not their signals, as in our experiment.

Thus, in the analysis that follows we assume that for all voters $p_i \geq 0.5$. *That is, we assume that voters either expect their information to be on average positively informative or expect that on average their information is uninformative; we assume that no voter expects that their information on average is negatively informative. We furthermore assume that for at least one voter* $p_i > 0.5$. Our experiment is setup such that voters are on average "correct" and in some treatments we attempt to induce voters with this expectation.[4] Therefore, when in a given election for all voters $0 \leq p_i^j < 0.5$, we expect that the majority of votes will be incorrect and the group will choose the wrong option because they expect that on average they are correct.

2.2. Equilibrium behavior with social information

2.2.1. Social information about opinions

The information we study is "public" in the sense that everyone obtains it, it is free in the sense that voters do not have to pay or search for it. It is "social" in the sense that it refers to what other people think or have done (rather than to the physical environment etc.). Social information about opinions is often provided to voters when they observe other voters choosing in similar elections, public opinion polls, or surveys. We model a voting situation in which voters receive social information about opinions of other voters in a similar situation. That is, assume that there are now two groups of voters, group 1 and group 2, who independently vote over the exact same election j, a and b, with the same consequences for each group. To clearly pin down the effects of informational spillovers, we assume that the realized state of the world is the same; that is, if the state of the world is A in group 1, it is also A in group 2, and vice versa. The two groups are the same size, n. Voters' preferences are exactly the same in both groups and the realized signal qualities are the same. However, the choices of one group have no effect on the utility of members of the other group except through the information link.

Group 1 voters choose first and make their choices exactly as we assume in the previous subsection, with no social information. Then group 2 voters choose, but they are given information about the distribution of choices of group 1 voters (i.e., how popular the options are in group 1) before they choose to vote. Specifically, define n_k as the total number of votes for option k in group 1 and $q = n_k/n$, that is, the proportion of votes in group 1 for option k.[5] Voters in group 2 are told q and $(1 - q)$ before they choose. Note that group 2 voters do not learn whether group 1 voters' choices were "correct" in the sense that the voters' choices maximized their utility by choosing the option that matched the state of the world but the proportions that have chosen a and b. Hence, if for the majority of voters $p_i^j < 0.5$, then it is likely that group 1 members voted a majority for the option that did not match the state of the world (given that we have assumed that the number of voters who expect their information to be on average positively informative outweigh the number of voters who expect their information to be on average negatively informative).

As we show in Appendix A, group 1 members sincerely vote their signals if their signals are on average positively informative and vote randomly if their signals are on average uninformative. But what about group 2 voters? We also show in the Appendix A that voting decisions of group 2 voters should depend on their signals and the size of q. Specifically, we show that voter i who has received an a signal, has signals which are on average positively informative, and knows q, will prefer to vote as follows and vice-versa if they receive a b signal:

$$\begin{array}{ll} \text{If } 1 > n(1-2q) & \text{Vote for } a \\ \text{If } 1 < n(1-2q) & \text{Vote for } b \\ \text{If } 1 = n(1-2q) & \text{Indifferent} \end{array}$$

Voters who expect on average that their signals are uninformative should always vote for the option preferred by the previous group.

[4] Although we cannot empirically distinguish between self-aware voters who expect that their information is on average correct and those who expect the opposite, our experimental results concerning voter behavior in response to social information on successes (explained below) suggests that our assumption is reasonable.

[5] To simplify notation we drop the subscript j from our variables.

Hence, when the size of the majority voting in favor of option b in group 1 is more than one vote (in our experiment with $n = 5$ voters in a group implies $q < 40\%$), then voters in group 2 who have received an a signal should ignore their signals and vote for b. As discussed in the Introduction, our result is an extension of the literature on herding and information cascades in independent individual choices (see Bikhchandani et al., 1998) to sequential independent collective choices.

2.2.2. Social information about success

In contrast to receiving information about opinions and voting choices, a different type of social information is provided when voters learn about whether previous groups' collective choices are successful but not about the choices made by these groups. Voters might receive this information by observing the degree to which other voting groups are pleased or not with governmental decisions. For example, voters in one state in the U.S. may observe the economic well-being of voters in another state or their degree of satisfaction with their government officials. Such information may be provided by surveys or news reports. However, they may not know the specifics of the policies that led to these consequences. The idea here is that voters learn whether other groups made smart (successful) choices in deciding on a particular issue, but not details of what they chose.

In analyzing social information about success, we make the same simplifying assumptions as in the discussion of social information about opinions in the previous sections. But now, voters in group 2 are given information about the distribution of correct choices of group 1 voters before they choose to vote. Specifically, define n_c as the total number of *correct* votes in group 1 and $c = n_c/n$, the proportion of votes in group 1 voting for the option that matched the state of the world. Voters in group 2 are told c before they choose. Note that group 2 voters do not learn the proportions that have chosen a and b, i.e. how they voted, but simply whether the outcome of the voting was utility maximizing. Again, we expect that group 1 voters who expect that on average their signals are positively informative should sincerely vote their signals (see Appendix A) and those who expect that their signals are on average uniformative should randomize. We continue to assume that group 2 voters condition their vote choices on the event that they are pivotal and focus on pure-strategy symmetric equilibria in which voters who receive the same signal choose the same strategy.

The crucial effect of providing social information about success is that voters obtain new information on the realized value of the p^j in group 1, not available in the other cases. In the other cases, a voter's best guess as to the probability that his or her signal is correct is given by the parameter p_i, the expected value of his or her true signal quality. However, in the situation in which voters receive social information on the success of group 1, that is, c, they have additional information about the distribution of p_i^j that is unavailable to voters without social information or voters with social information on opinions only. If for all voters, $p_i > 0.5$ and thus they vote their signals, then c is a sample expected value of the mean of true signal qualities across voters in election j, p^j.

Assuming that group 2 voters are Bayesian updaters, voter i will use a weighted average of his or her prior (p_i) and the social information (c) received. In particular, we predict that the expectation of p^j of voter i in group 2, which we designate $\widehat{p}_j$, is a weighted average of p_i and c, as follows (where α is the weight placed on the new social information, $0 \leq \alpha \leq 1$):

$$\widehat{p}_j = \alpha c + (1 - \alpha) p_i \tag{1}$$

Suppose now that instead of there being just one group that votes prior to group 2, there are many such groups without social information choosing simultaneously and group 2 voters are told the average of the observed correct rates across these groups. It is well known that the mean of these sample proportions approaches the true value of p^j. In our experiment we provide subjects with the mean proportions across multiple groups and thus one might conjecture that the weight α placed on this average value of c, which we call $\bar{c}$ would approach 1. In the analysis that follows we make the strong assumption that $\alpha = 1$. We show in Appendix A that rational voters will vote as if there is only private information when $\bar{c} > 0.5$, vote contrary to how they would vote if there is only private information when $\bar{c} < 0.5$, and are indifferent between options when $\bar{c} = 0.5$.

Intuitively, voters learn the share of voters in other groups who made the correct choice (but not what it was). A rational voter who learns that a majority of voters in other groups were right ($\bar{c} > 0.5$), votes according to his or her own signal (if his or her signal is on average positively informative and randomizes if his or her signal is on average uninformative). That is, the social information has no value in this case. But when a majority of voters in other groups are wrong ($\bar{c} < 0.5$), the voter will vote contrary to his or her private signal if his or her signals are on average positively informative. The social information suggests to the voter that since he or she is similar to others who voted before, and their signals appeared to have led them astray, then it is probably likely that their signals are wrong and he or she should vote strategically instead.

Note that the same logic holds for voters who expect on average their signal to be uninformative. That is, they have received a signal and while on average across elections their signals are uninformative, in a given election these signals may be informative. Hence, the information from other groups provides them with information on the likely correctness of their signal as well. Therefore, they should follow the same logic as voters for whom $p_i > 0.5$. Social information on success should then de-bias voters.

Of course there are other means by which voters might be de-biased and avoid the dark side. For example, voters might learn from experience with choices that are similar over time when they are more likely, as a majority, to make biased inferences. In our experiment we do not allow for such learning to occur as subjects are not given feedback after each choice as discussed below. We are interested in the situation in which the choices are new to voters and thus not something over which they have sufficient information about their own capabilities. We also believe that the learning that can occur in the lab through participating in thirty elections is not likely to be observed naturally. We believe that it is more the case that voters can gain information not from their own prior experience in this case but for the experiences of other similar voters in other jurisdictions as when countries are voting in referenda sequentially over whether to join or leave the European Union. Hence, we focus on the possibility that experiences of others can de-bias voters.

2.3. Efficiency of voting choices

What do these theoretical results imply about the efficiency of information aggregation in the groups? First, consider the situation in which no social information exists. How efficient is voting one's signal in this case? We define Informational Efficiency of Majority Voting as the equilibrium probability with which a group makes the correct decision through majority voting. Label the expected probability of making the correct choice under majority voting absent social information as $E(P_U(p^j))$. Recall that p^j is the average signal quality of voters in an election j. For a group of five voters as in our experiment, $E(P_U)$ is given by

$$E(P_U(p^j)) = \left(p^j\right)^5 + 5\left(p^j\right)^4(1-p^j) + 10\left(p^j\right)^3(1-p^j)^2 \quad (2)$$

In a typical election, voting leads to more efficient outcomes than the average voter if voters on average make correct inferences. For example, if $p^j > 0.5$, $E(P_U(p^j))$ is greater than p^j. However, when $p^j < 0.5$, i.e., the average voter makes incorrect inferences on a particular issue, voting will result in less efficient information aggregation than the average voter as the probability of making the correct choice will be less than p^j.

Note that when $p^j > 0.5$, it is not necessarily the case that the group choice is more efficient than the choice of all of the voters choosing individually. That is, consider a group of three voters with signal qualities given by 0.9, 0.65, 0.65. In this case for all three voting their signals, the group choice is correct with probability 0.83 and $p^j = 0.73$, so the group choice is better than the average voter's choice. However, the group choice is not better than if the voter with the high quality signals made the choice.[6] Conversely, when $p^j < 0.5$, it is possible for the group choice to be better than the choice of some of the individual voters, but still worse than the choice made by the average voter. Hence, our results comparing group choices with individual choices are for the case of the comparison to the average individual voter, not necessarily in comparison to the voter with the highest quality of signal.

Now consider voting behavior when voters have social information on opinions as compared to group voting without social information. The expected probability of choosing the utility maximizing option in this case, which we label $E(P_O(p^j))$, is equal to the expected probability that the correct option won with more than a one-vote margin of victory in group 1 plus the expected probability of voting correctly when everyone votes their signals in group 2 times the expected probability that the margin of victory in the previous group was no more than one vote. This expected probability can be shown to be equal to the following in the case of five voters:

$$\begin{aligned} E(P_O(p^j)) = &\left(\left(p^j\right)^5 + 5\left(p^j\right)^4(1-p^j)\right) \\ &+10\left(\left(p^j\right)^3(1-p^j)^2 + (1-p^j)^3\left(p^j\right)^2\right)\left(\left(p^j\right)^5 + 5\left(p^j\right)^4(1-p^j) + 10\left(p^j\right)^3(1-p^j)^2\right) \end{aligned} \quad (3)$$

As in the case where no social information exists, when $p^j > 0.5$, then $P_O(p^j) > p^j$ and vice versa when $p^j < 0$. Therefore, information aggregation through voting with social information on opinions is on average more efficient than an individual voting alone. Furthermore, when $p^j > 0.5$, then $E(P_O(p^j)) > E(P_U(p^j))$, but when $p^j < 0.5$, then $E(P_O(p^j)) < E(P_U(p^j))$. Thus, voting with social information on opinions is more efficient than voting without social information when inferences are on average correct but more inefficient than voting without social information when inferences are on average incorrect. However, on average, voting with social information on opinions is more efficient than voting without social information, since it is more likely that inferences are on average correct.

When voters have social information on success the probability of choosing the utility maximizing option, which we label $P_C(p^j)$, depends on whether p^j is greater or less than 0.5. When $p^j > 0.5$, $E(P_C(p^j)) = E(P_U(p^j))$. But when $p^j < 0.5$, then $E(P_C(p^j)) = E(P_U(1-p^j))$. Thus, for the case of five voters we have:

$$\begin{aligned} E(P_C(p^j)) &= \left(p^j\right)^5 + 5\left(p^j\right)^4(1-p^j) + 10\left(p^j\right)^3(1-p^j)^2 && \text{If } p^j > 0.5 \\ E(P_C(p^j)) &= (1-p^j)^5 + 5(1-p^j)^4 p^j + 10(1-p^j)^3\left(p^j\right)^2 && \text{If } p^j < 0.5 \\ E(P_C(p^j)) &= 0.5 && \text{If } p^j = 0.5 \end{aligned} \quad (4)$$

[6] In such cases it is better for the two voters with signal qualities given by 0.65 to abstain and delegate the choice to the voter with signal quality equal to 0.90. However, if the difference in signal qualities are not large (say the three voters had signal qualities equal to 0.83, 0.79, 0.79), then the group choice is more efficient than the choice made by the most informed voter. See Morton and Tyran (2011) for experiments that explore these configurations and voters have the option to abstain.

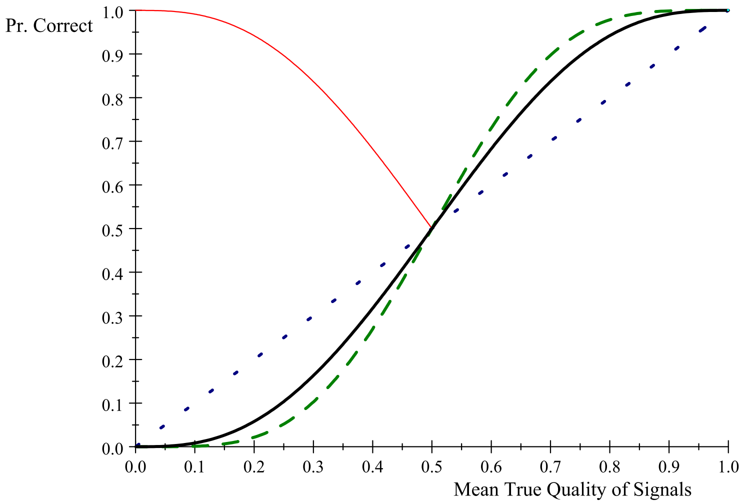

Fig. 1. Probability of optimal choice as a function of p^j. (Dotted line represents mean signal quality $= p^j$; solid black line $= E(P_U(p^j))$ & $E(P_C(p^j))$ when $p^j > 0.5$; dashed line $= E(P_O(p^j))$; solid red line $= E(P_C(p^j))$ when $p^j < 0.5$.) (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Hence, we find that social information about success is equivalent in efficiency to no social information when $p^j > 0.5$, but is more efficient than either the case of no social information and social information on opinions when $p^j < 0.5$. Social information on success is clearly superior in efficiency to voting without social information and individual choice. However, social information on success is not necessarily more efficient than social information on opinions. The greater the variance in p^j (that is the greater the variance in average signal quality across elections) and the more likely it is that inferences are on average incorrect, the more likely social information on success is superior to social information on opinions.

Fig. 1 summarizes these efficiency results for the case of five voters.[7] The vertical axis measures the probability of choosing the best option as a function of p^j, the average true quality of signals, measured along the horizontal axis. The dotted 45 degree line represents the case where this probability equals p^j as in individual choice where individuals follow their signals; $E(P_U(p^j))$ is given by the solid black line; $E(P_O(p^j))$ is given by the dashed line; and $E(P_C(p^j))$ when $p^j < 0.5$ is given the solid red line (and by the solid black line when $p^j > 0.5$). These theoretical results are also summarized below as Predictions 1, 2, and 3.

Prediction 1 *(Efficiency of Majority Voting without Social Information).* When signals are on average correct, then majority voting is more efficient at information aggregation than the average voter, but when signals are on average incorrect, majority voting is less efficient.

Prediction 2 *(Efficiency of Majority Voting with Social Information on Opinions).* When signals are on average correct, then majority voting with social information on opinions is more efficient at information aggregation than both majority voting without social information and the average voter, but when signals are on average incorrect, majority voting with social information on opinions is less efficient than both.

Prediction 3 *(Efficiency of Majority Voting with Social Information on Success).* When signals are on average correct, then majority voting with social information on success is more efficient at information aggregation than the average voter and equivalent in efficiency to majority voting without social information, but less efficient than majority voting with social information on opinions. When signals are on average incorrect majority voting with social information on success is more efficient than the other three cases.

3. Experimental design

3.1. General procedures

The experiment took place at the Laboratory for Experimental Economics (LEE) of the University of Copenhagen (Denmark). The experiment consisted of a total of 13 sessions: 4 sessions for each of 3 treatments, described below, plus one

[7] Obviously, as n increases the probability of making correct choices through majority voting when $p_j > 0.5$ converges to one both with and without social information. When $p_j < 0.5$, this probability converges to zero without social information and with social information on opinions, but converges to one with social information on successes.

additional individual decision-making session. In each session, 15 to 25 subjects participated. Subjects were recruited using the online system Orsee (Greiner, 2015) and all participants were undergraduate students of the University of Copenhagen. No subject had previous experience with similar experiments and each subject participated in only one session. The experiment was programmed using the software z-Tree (Fischbacher, 2007). At the beginning of each session, subjects received a copy of the instructions available in the Auxiliary Materials Appendix B. We followed the experimental procedures of anonymity, incentivized payments, and neutrally worded instructions that are typically used in such experiments. Overall, 291 subjects participated and earned, on average, 190 Danish Krone (DKK, approx. 25 Euro). Each session lasted approximately 1–2 hours.

3.2. Creating situations where inferences can be incorrect

Our theoretical formulation makes precise predictions about how subjects should vote and the efficiency of information aggregation through voting in situations in which the true quality of signals given to voters is uncertain and subjects may make incorrect inferences. We are most interested in the "dark side" of the vote, i.e., the effects incorrect inferences may have on the extent that majority voting can effectively aggregate information. We also wish to discover how social information may hinder or help the ability of voters to aggregate information through voting, particularly when inferences are on average incorrect.

Previous experiments on information aggregation through voting typically make the inference problem for voters exceedingly easy. In a typical such experiment, as we discussed above, subjects are told there are two jars, one red and one blue. Each jar has, say, 8 balls. In the red jar there are 6 red balls and 2 blue balls and in the blue jar there are 6 blue balls and 2 red balls. A jar is randomly chosen from a known probability distribution but subjects are not told the identity of the true jar. Each subject then randomly chooses a ball from the unknown jar (with replacement). In expectation, then, subjects should conclude that the true color of the jar has a higher probability of matching the ball each has drawn. Evidence suggests that almost all subjects are able to make the correct inference; that is, in these experiments subjects generally vote the color of the ball they receive as a signal in situations in which sincere voting is predicted such as under majority rule voting. Not surprisingly, typically experimentalists find that majority voting leads to more informed choices than the individuals would reach acting alone.[8]

In our experiment we wished to use decision problems which vary in difficulty, including situations in which it is possible that a majority is more likely to be wrong than right. Therefore, in our experiment subjects were presented with a series of quiz questions with two answers, labeled A or B. After extensive pre-testing, 30 questions were chosen. The majority of the questions, although they ranged in difficulty, were on average answered correctly in our pre-testing. But we also included a minority of questions in which most people display cognitive biases and make systematic incorrect inferences as shown in several previous studies (see, for instance, Hoorens, 1993) and in our pre-testing.[9]

The correct answer to a question, then, is the true "state of the world" in our theoretical setup. Subjects were told simply that the answer could be either A or B before reading a question. Hence, before reading a question, subjects should have on average expected either answer was equally likely (in fact they were equally likely). Subjects received their individual signals when they read the questions. Our experimental environment was therefore in some ways more parallel to the target environment of much of the theory of information aggregation in voting (like jury decision-making) than previous experiments as in actual juries individuals are all given common information either verbally or in a written transcript but each individual's understanding of that information is supposedly subject to independent random shocks and their own abilities or competence.

Subjects answered the questions sequentially, but were not told the answers to any questions until all had been completed. Between-subject communication was not allowed. We did not give feedback because we felt that subjects might update their prior beliefs on the difficulty of the questions by learning the consequences of previous answers. That is, suppose a subject approached most questions with the belief that he or she is on average correct. But after answering several "misleading questions" in which they are shown to be wrong because of their cognitive biases by learning the results (something that they would not learn when feedback is withheld), then they may approach later questions with more skepticism of their own abilities which we wished to avoid.[10]

Although our laboratory experiments arguably has more ecological validity, our manipulation has the same advantages over field studies of voting groups that exist in previous laboratory experiments in that we controlled the choices before the subjects and could randomize the type of social information received. Moreover, we knew the answers to the questions

[8] For example, in Guarnaschelli et al. (2000) the probability that an individual voter acting alone was correct was 70% when voting his or her signal (which voters did 94% of the time under majority rule) but groups deciding by majority rule were correct more than 70% of the time on average, depending on the size of the group and the true jar chosen.

[9] We also conducted a separate incentivized individual decision-making treatment as part of the experiment. In this session we found answers followed a similar distribution as discussed below.

[10] In one of our treatments, Success, described above, subjects did receive information about the difficulty of questions that might have influenced perceptions about future questions. Therefore, as described below, in one set of sessions and treatments we gave all subjects information on the overall difficulty of questions from a previous nonvoting incentivized session to anchor their perceptions.

Table 1
Treatment description.

All voters and groups answered 30 questions			
Treatment	Subjects	Information from Baseline	Information from NonVoting
NonVoting	25	None	None
Baseline 1	45	None	None
Baseline 2	50	None	p from NonVoting
Opinions 1	35	q in Baseline 1	None
Opinions 2	45	q in Baseline 2	p from NonVoting
Success 1	45	$\bar{c}$ in Baseline 1	None
Success 2	45	$\bar{c}$ in Baseline 2	p from NonVoting

and thus had an objective measure of the true state of the world. The questions we used can be found in the Auxiliary Materials.[11]

3.3. Treatments

We conducted three main treatments: Baseline, Opinions, and Success. In all the treatments, each question involved two stages: 1) Subjects indicated which answer they thought was correct ("Choice Stage") and 2) Subjects had the possibility to confirm (or switch) their answer ("Confirmation Stage"). Before each question, subjects were randomly re-matched in anonymous groups of 5. Therefore, if there were 25 subjects in a session, for each question there were 5 groups of 5, which were randomly drawn for each question. Simple majority voting was used to determine a group's decision. As the number of voters was odd and abstention was not allowed, we had no tie elections. Each subject received 10 DKK (approx. 1.4 Euro) for every correct group decision independently of how they individually voted. For each of these three treatments we conducted four separate sessions.

The treatments differed only in the information provided to the subjects between the "Choice Stage" and the "Confirmation Stage." In Baseline we do not provide any information between the two stages. In Opinions, subjects were told how popular the alternatives (A and B) were among voters in the two previous sessions of Baseline (q in section 2.2.1), while in Success, subjects were told the percentage of individuals who provided the correct answer in Baseline ($\bar{c}$ in section 2.2.2). In addition, in both stages in all treatments, subjects were asked to indicate how certain they were about their answer in a scale from 1 (not certain) to 5 (certain). The measure of certainty was not incentivized. We also conducted one session for a NonVoting Treatment with 25 subjects in which subjects answered the questions sequentially without feedback and were paid based on the individual accuracy of their responses.

We conducted the Baseline, Opinions, and Success sessions in two different sets with roughly half of the observations in each set. Henceforth, Baseline 1, Opinions 1, Success 1 refer to sessions from the first set and Baseline 2, Opinions 2, and Success 2 refer to the sessions from the second set. In the first set of sessions subjects had exactly the information described above. In the second set of sessions, we provided the subjects with one additional piece of information. Specifically, in the instructions we provided subjects with information on the overall difficulty of the questions by giving them information from the NonVoting sessions. Specifically, they were asked to read the following message which was shown on their computer monitors: "The questions you will be asked vary in difficulty. In a previous session in which subjects answered these questions (there was no voting), only 16% got the hardest question right while 100% got the easiest question right. On average, subjects answered 60% of the questions correctly." Subjects were given this information once at the beginning of the experiment.

We instituted the second set of sessions because we were concerned that subjects may change their beliefs during the experiment about the overall difficulty of the questions that might influence their answers and lessen the validity of the experiment as a test of the theory. That is, the theory posits that subjects expect that in an unknown question they will be "on average" correct. Hence, in the second set of sessions we attempted to avoid this difficulty. As we will see in the results section, there are only minor differences in behavior between the two sets of sessions.

Note that in Opinions and Success we used the results from the corresponding Baseline sessions for the social information provided. That is, subjects in Opinions 1 and Success 1 received social information on opinions and successes, respectively, of Baseline 1 voters for each question between giving their initial and final answers to each question while subjects in Opinions 2 and Success 2 received information on opinions and successes, respectively, of Baseline 2 voters.

Table 1 summarizes the relevant information and the main characteristics of each treatment.

[11] These questions have been chosen not for their practical relevance but for their quality of having clear-cut right and wrong answers, and we can credibly communicate to subjects that they do. The advantage of our design is that we have (by virtue of pretesting) quite precise knowledge about the accuracy with which subjects answer these questions. We are thus able to compose the questions with $p^j > 0.5$ and $p^j < 0.5$ such that we know $p > 0.5$ with high confidence. However, the technique does not allow us to know or control p_i^j or p_i for a particular subject.

Also note that we are not interested as such in how subjects vote on these particular questions. In fact, these are issues on which a group would ideally ask a trusted expert (or consult a lexikon).

As noted in the theoretical section, it is impossible both theoretically and within our empirical design to distinguish between the choices a rational voter i makes for whom $p_i < 0.5$ (he expects that on average across elections his signals are negatively informative) and the choices made by a rational voter k for which $p_k = 1 - p_i > 0.5$ (she expects on average across elections her signals are positively informative). Hence, we assume that for all voters $p_i \geq 0.5$. Under that assumption, we did attempt to measure the extent that our subjects perceived their overall signal quality across elections in NonVoting. That is, at the end of the session but before receiving feedback on their performance we asked subjects to estimate the number of questions that they answered correctly. We rewarded them with an additional 50 Danish Kroner if they estimated the right number of correct answers, deducting 5 Danish Kroner for every integer difference between their prediction and the true number. The mean response of the subjects was 67% of the questions, with only one subject estimating that he or she was wrong more than right.[12] Only five subjects appeared to suggest that they may have been answering randomly all the time by estimating that they got exactly 15 (half) of the questions correctly. Most appeared to believe that they were more right than wrong (19 out of the 25 subjects).

Subjects generally overestimated their performance as the average number of questions answered correctly across subjects was 17.96 or on average 59.9% correct. Nineteen of the 25 subjects were more right than wrong, two subjects were right exactly half of the time, and four were more wrong than right. The highest number of questions answered correctly was 24 and the lowest was 12. We did not ask this question of subjects in our voting games as individual choices in the voting game can depend on inferences voters make about the best choice based on the event in which they are pivotal (what he or she believes others are likely to know and choose), whereas individual choices in NonVoting are expected to not be so dependent on these perceptions.

3.4. Cognitive reflexion test

In addition to the questions in the experiment, at the end of the experiment in Baseline, Opinions, and Success, subjects completed a simple Cognitive Reflection Test (CRT), reported on in Frederick (2005). In the CRT subjects were asked three questions (which were not incentivized and subjects were given as long as they wished to answer the questions). These questions are also listed in Appendix C.[13] Each of these questions has an intuitive response that is wrong, yet the questions themselves are relatively easy once the answer is explained. As Frederick (2005) demonstrates the CRT test has high predictive validity in measuring cognitive abilities comparable to other measures used in the literature that involve much more extensive questions and longer completion times. As we expect our subjects to vary in their abilities to make correct inferences, we use the CRT test as a measure of these differences in our empirical analysis of individual behavior. We construct a measure of performance in the CRT as the percentage of correct responses, hereafter *CRT Score*. Finally, subjects in all the sessions except for NonVoting also answered a number of demographic questions measuring age, gender, major area of study, nationality, and income.

4. Experimental results

Table 2 summarizes our main experimental results. In Table 2 we present the percentages of individual and group correct choices by treatment. We present these percentages across all questions as well as for questions which we label "easy," "hard," and "misleading." We define an easy question as one for which 60% or more individuals in NonVoting made correct choices (15/30), a hard question is one for which less than 60% but more than 40% of individuals in NonVoting made correct choices (7/30), and a misleading question is one for which 40% or less individuals in NonVoting made correct choices (8/30). Easy questions, then, are questions where we expect that 1) voting can be more informationally efficient than individual choices and that 2) social information on opinions can improve informational efficiency. Misleading questions are questions where we expect the opposite. We expect our treatments to have little effect on choices when questions are hard. We use choices in NonVoting to define these three sets of questions since these choices were incentivized individual choices without any voting considerations and thus give an independent measure of question difficulty. Individual choices in the voting games were not incentivized separately to prevent hedging. Moreover, as noted above, individual choices in the voting games can be influenced by information inferred from being pivotal, unlike the individual choices in NonVoting. Note that we considered other classifications with no qualitative difference in results.

Overall we find no large differences in overall subject choices between the two versions in each of Baseline, Opinions, and Success, which yield remarkably similar proportions of correct responses. The largest difference occurs between Opinions 1 and Opinions 2 with subjects in Opinions 1 answering slightly more questions correctly, which is not significant at conventional levels.[14] When we disaggregate by easy, hard, and misleading questions, we find that the primary difference

[12] This proportion is significantly greater than 0.5 using a one-tailed test, z statistic = 1.72, Pr = 0.04.

[13] In Baseline 2 a different, but similar set of three CRT questions was asked from those asked in the rest of the treatments. There is no statistical difference between the mean sum of correct answers to the CRT questions for the two versions, t statistic = 1.27, Pr = 0.21, so we combine the two versions in our empirical analysis. The results are the same if we only use the version used in the majority of the sessions. We did not administer the CRT in NonVoting.

[14] Comparing the proportions correct in Baseline 1 with Baseline 2 yields a z statistic = 0.17, Pr = 0.86; for Opinions 1 with Opinions 2 = 1.87, Pr = 0.06; and for Success 1 with Success 2 = 0.36, Pr = 0.72.

Table 2
Individual and collective choices by treatment.

Treatment	Percent correct choices											
	Overall			Easy*			Hard*			Misleading*		
	Individ.	Groups	Obs.**	Individ.	Groups	Obs.**	Individ.	Groups	Obs.**	Individ.	Groups	Obs.**
Individual	60%	NA	750	79%	NA	375	53%	NA	175	30%	NA	200
Baseline All	59%	63%	2,850	76%	87%	1,425	50%	53%	665	36%	28%	760
Baseline 1	59%	63%	1,350	75%	84%	675	50%	52%	315	39%	33%	360
Baseline 2	59%	63%	1,500	77%	89%	750	51%	53%	350	33%	23%	400
Opinions All	60%	60%	2,400	89%	97%	1,200	45%	41%	560	20%	8%	640
Opinions 1	63%	62%	1,050	91%	96%	525	47%	43%	245	22%	14%	280
Opinions 2	59%	59%	1,350	88%	98%	675	44%	40%	315	18%	3%	360
Success All	61%	65%	2,700	78%	90%	1,350	49%	50%	630	40%	29%	720
Success 1	61%	63%	1,350	79%	90%	675	47%	44%	315	40%	29%	360
Success 2	62%	66%	1,350	78%	90%	675	51%	56%	315	40%	29%	360

*Easy = questions (15/30) for which ≥60% individuals in IT made correct choices, Hard = questions (7/30) for which <60% but >40% made correct choices, and Misleading = questions (8/30) for which <40% made correct choices.
***Number of individual choices. For number of groups, divide by 5.

between Opinions 1 and Opinions 2 occurs for the easy questions which is significant, but that no other differences are significant.[15] Hence, for the remaining analyses we present our primary results at the aggregate levels. In the statistical analyses we control for differences in sessions and the two sets of sessions.[16]

4.1. Is majority voting more informationally efficient than individual choice?

Prediction 1 concerns information aggregation in Baseline. Specifically, we expect that when individuals largely make correct inferences, group choices are better than the average individual's choice and when individuals largely make incorrect inferences, group choices are inferior to the average individual choice. When we compare individual choices in Baseline with group choices, we find that there are significantly more correct group choices when questions are easy, 76% of individuals make correct choices, but group choices are correct 87% of time (z statistic = 4.18, Pr = 0.00). For hard questions there is no statistical advantage of groups over individuals, 50% of individuals make correct choices and group choices are correct 53% of time (z statistic = 0.54, Pr = 0.59). And, in keeping with our prediction, when questions are misleading, using a one-tailed test groups are significantly more likely to be incorrect than individuals, 36% of individuals are correct, but only 28% of groups are correct (z statistic = 1.87, Pr =0.06).

Fig. 2 illustrates these comparisons in more detail by question. The questions are divided by type and ranked by difficulty for subjects in NonVoting. When we compare group choices with individual choices in NonVoting we find similar results to those reported above. For easy questions, we see that the voting groups in Baseline are correct a greater or equal to percentage of the time than individuals in NonVoting for 13 of the 15 questions, but for hard questions they are correct a greater percentage of the time for only 2 out of the 7 questions, and for misleading questions for only 2 out of 8.

Result 1 provides support for our first prediction:

Result 1 *(Group Choices with No Social Information).* Majority voting results in more informationally efficient choices than the average individual when questions are easy, but no more informationally efficient choices than the average individual when questions are hard. Majority voting results in less informationally efficient choices than the average individual when choices are misleading.

4.2. Does social information improve information efficiency?

Prediction 2 states that social information on opinions induces groups to make better decisions than absent such social information when signals are on average correct, but worse decisions when signals are on average incorrect. As expected we find that for easy questions groups in Opinions make many more correct choices (97%) compared to the Baseline (87%) and when questions are misleading make fewer correct choices (8%) compared to the Baseline (28%), both differences are

[15] For the comparison between Baseline 1 and Baseline 2 and easy questions the z statistic = 1.00, Pr = 0.32; hard questions = 0.27, Pr = 0.79; and misleading questions = 1.84, Pr = 0.07. The values for the comparisons between Opinions 1 and Opinions 2 are 2.00, Pr = 0.05; 0.74, Pr = 0.46; and 1.47, Pr = 0.14. The values for the comparisons between Success 1 and Success 2 are 0.33, Pr = 0.74; 1.12, Pr = 0.26; and 0.00, Pr = 1.00.

[16] If we compare the two sets of sessions at group level choices we only find one significant difference between Opinions 1 and Opinions 2 for misleading questions. However, group choices may vary simply due to randomization of assignment to groups and the relevant comparisons are individual voter choices.

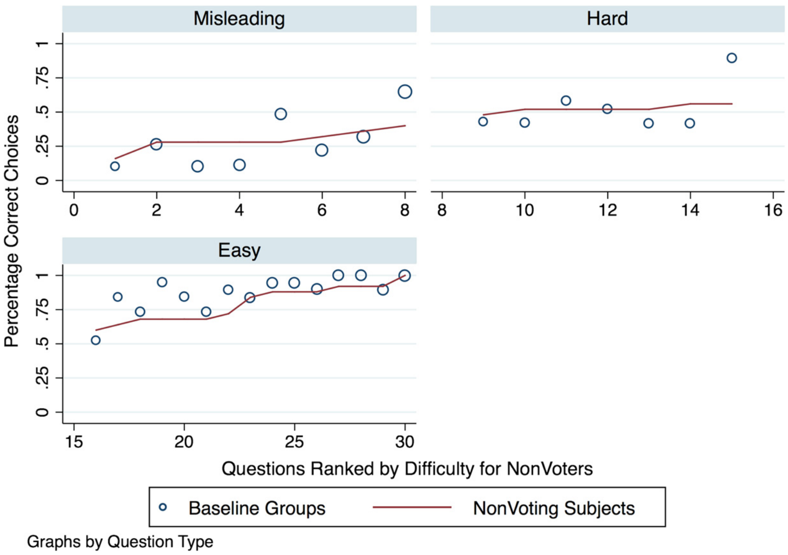

Fig. 2. Percentage correct baseline group choices versus percentage correct nonvoting choices.

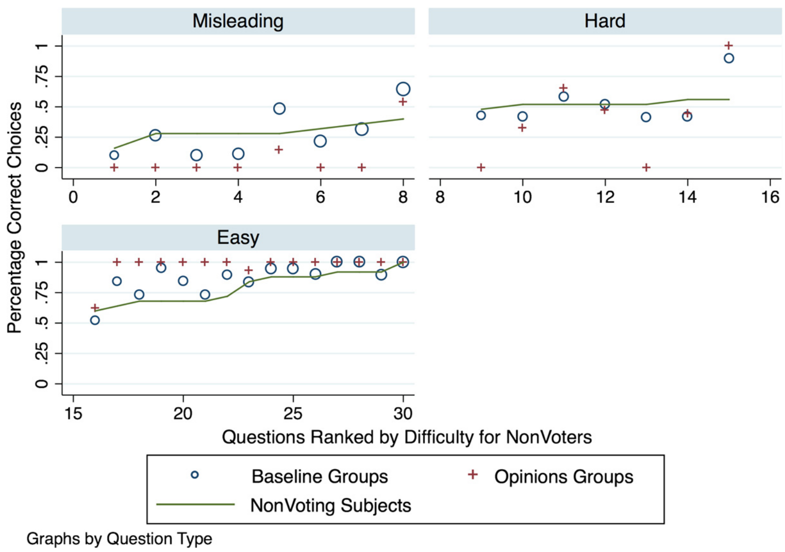

Fig. 3. Percentage correct opinions and baseline group choices versus percentage correct nonvoting choices.

highly significant. For hard questions, groups in Opinions make less correct choices (41%) than in the Baseline (53%) as well but the difference is not significant at conventional levels.[17]

Fig. 3 compares the percentage correct choices of groups in Opinions, Baseline, and subjects in NonVoting by question difficulty as in Fig. 2. Fig. 3 illustrates how for misleading questions the percentage of correct choices in Opinions is lower than in Baseline for *every* question and for easy questions the percentage of correct choices in Opinions is at least as large as Baseline for *every* question. For 5 of the 7 hard questions, the percentages correct in Opinions and Baseline are, as expected, closely aligned.

Our third prediction is that social information on success should not affect information efficiency of voting when voters are on average correct, but should improve information efficiency when voters are on average incorrect. In order to evaluate this prediction we compare group choices in Success with those in Baseline. As predicted, we find no significant difference between group choices in Success and Baseline when questions are classified as easy (90% compared to 87%, z statistic =

[17] The z statistic for the comparison of easy questions = 4.15, Pr = 0.00; for hard questions = 1.81, Pr = 0.07; and for misleading questions = 4.25, Pr = 0.00.

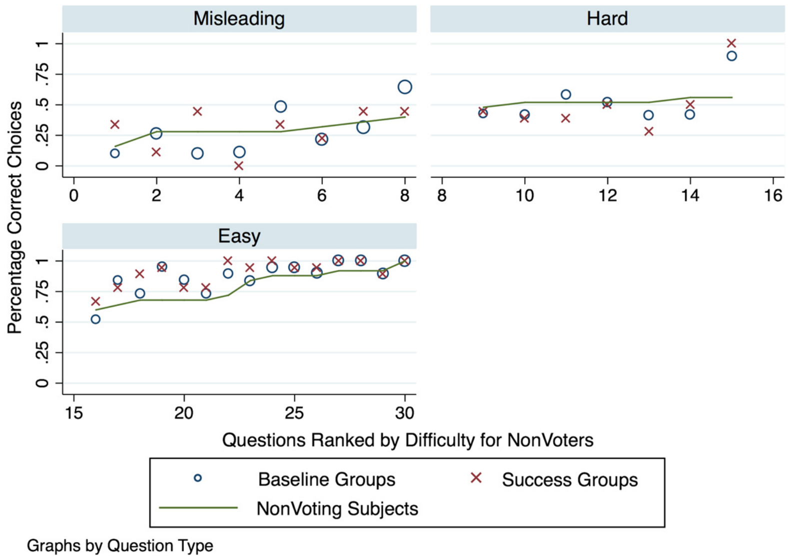

Fig. 4. Percentage correct success and baseline group choices versus percentage correct nonvoting choices.

1.24, Pr = 0.21) and hard (50% compared to 53%, z statistic = 0.42, Pr = 0.67). However, contrary to expectations, we also do not find a significant difference when questions are classified as misleading (29% compared to 28%, z statistic = 0.29, Pr = 0.77).

Fig. 4 illustrates these similarities between Success and Baseline by Question as in the previous figures. As the figure shows, there are significant overlaps between the percentage correct for all three question types. For the misleading questions, the percentage correct is greater for only 3 out of 8 questions, and worse for 4 out of 8, overlapping for one question. Social information on success does not lead to higher information efficiency.

One possible problem with the analysis above is that we do not control for differences in question difficulty within our categories of easy, hard, and misleading questions. There is also variation, as a consequence, in the social information received within our categories. In order to control for question difficulty, variation in social information received, and possible session effects, we estimated a probit regression where the dependent variable is the probability a group chooses correctly as a function of question difficulty and the precise information received. We measured *Question Difficulty* as 1 minus the percent correct answers on the question in NonVoting. We measured *Infor Opinions* as the victory size information from Baseline that subjects received in Opinions (the percentage of votes received by each option in the previous groups) and we measured *Infor Success* as the percent correct information from Baseline that subjects received in Success (the percentage of correct choices made in previous groups). We further interacted these information variables with question type generating six variables: *Infor Opinions Easy, Infor Opinions Hard, Infor Opinions Misleading, Infor Success Easy, Infor Success Hard, and Info Success Misleading*, as we expected the effects of the information to depend on the question type. The results from the probit analysis are presented in Table D1 in the Supplemental Online Appendix D.

Fig. 5 presents the marginal effects of the independent variables on the probability of a group choosing correctly. The results discussed above hold. We find that when questions are easy, information on opinions significantly increases the probability a group chooses correctly and significantly decreases that probability when questions are hard or misleading, even when controlling for variations in the information and question difficulty. And as above, we find no evidence that social information on success increases the probability a group chooses correctly except for a small effect when questions are easy.

Hence, we find strong support for our prediction that social information improves majority voting decisions when questions are easy (the bright side), but makes majority decisions worse when questions are misleading (the dark side) as hypothesized in Prediction 2. But our results provide no support for Prediction 3. These results are summarized below:

Result 2 *(Group Choices with Social Information on Opinions).* Social information on opinions leads to more informationally efficient group choices by majority voting than without such information when questions are easy, but less efficient group choices by majority voting than without such information when questions are misleading.

Result 3 *(Group Choices with Social Information on Success).* We find no evidence that social information on success mitigates the effects of incorrect inferences on group choices through majority voting in group level choices.

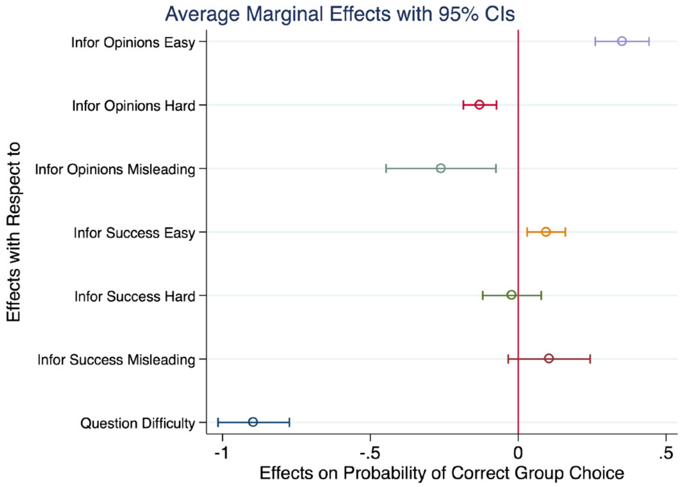

Fig. 5. Marginal effects of social information on probability of correct group choices.

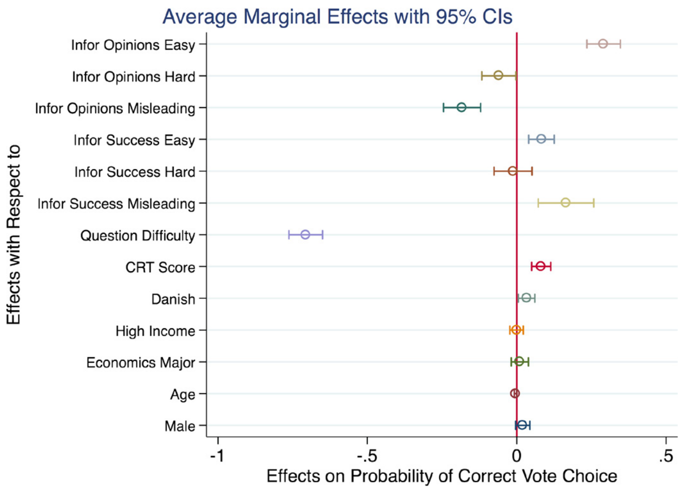

Fig. 6. Marginal effects of social information on probability of correct voter choices.

4.3. *Voter responses to social information*

Our group-level analysis suggests that social information is useful to assist groups in making better decisions when questions are easy, but not when questions are hard or misleading. Moreover, social information on success does not appear to de-bias voters and lead to better group decisions when questions are misleading, contrary to our prediction. We now turn to exploring how individuals respond to social information. Table 2 above also provides data on the percentage of correct voters choices, which, not surprisingly given that group choices are a simple aggregation of individual choices, show the same pattern of behavior in response to the treatments as found in group choices.

However, in the individual-level data we can explore the extent that the observed differences in individual choices are robust to controls for subject-specific differences as measured in our demographic survey and their CRT scores, clustering our data by subject. Table D2 in the Supplemental Online Appendix D presents the results from a similar probit estimation as reported in Table D1 for group choices adding our subject specific variables: *CRT score, Age, Male, Economics Major, Danish Nationality,* and *High Income*. We also adjusted our standard error for repeated observations of subjects. Fig. 6 presents the marginal effects of our independent variables on the probability an individual votes correctly.

Our results for individual voting are largely consistent with those found examining group choices. Voters are significantly more likely to vote correctly with social information on opinions when questions are easy but significantly less likely to do so when questions are hard or misleading. Voters are also significantly more likely to vote correctly with social information on success when questions are easy, but there is no significant effect of social information on success when questions are hard. We find one exception to our group choice results, however. We find that information on success has a positive and significant effect on the probability of voting correctly for misleading questions. Thus, we find evidence that suggests that voters are to some extent "de-biased" by receiving the information that previous groups voted incorrectly leading to more correct choices by the voters in Success.

Notably, we also find that higher CRT Scores predict a significantly higher probability of answering all types of questions correctly, even once we include controls. The marginal effect of a change in CRT score is 8%. We summarize these results below:

Result 4 *(Voter Choices with Social Information on Opinions).* Voters are significantly more likely to vote correctly with social information on opinions when questions are easy but significantly less likely to do so when questions are hard or misleading.

Result 5 *(Voter Choices with Social Information on Success).* Voters are significantly more likely to vote correctly with social information on success when questions are easy or misleading, suggesting that information on success does have a de-biasing effect on some voters even though group choices are unaffected.

Result 6 *(CRT and Voting).* We find that higher CRT scores are associated with a significantly higher probability of making a correct confirmed response in majority voting.

So far these results focus on between-subject measures of treatment effects. In addition to our measure of individual voter choices, we have two other types of voter behavior that yield within-subject measures and provide some information on voter responses by subject: (1) the extent that voters switch between their initial and final choices after receiving social information and (2) the extent social information affects voter certainty about their choices. First we turn the second measure, the extent that social information affects voter certainty about their choices.

4.3.1. Does social information affect voter certainty?

We also asked subjects to provide an estimate of how certain they were about their choices, both their initial choices and their confirmed choices on a scale of 1 to 5, where 5 represents most certainty. We created a measure, called *Certainty Difference* equal to confirmed certainty minus initial certainty with values from −4 to 4. Table D3 in the Supplemental Online Appendix D presents values of mean Certainty Difference by treatment. By examining the effects of treatments on Certainty Difference, we can measure whether social information leads to a bigger change in certainty than occurs without the social information. As with voting behavior and group level choices, within treatments we find no significant difference between the two sets of sessions in mean Certainty Difference, that is, there are no significant differences between Baseline 1 and 2, between Opinions 1 and 2, and between Success 1 and 2, even disaggregating by question type.[18] Given that we find no differences between these sets, we focus on treatment effects at the aggregate level. Comparing voting certainty in Baseline with Opinions we find that the difference in voter certainty is greater after receiving social information on opinions, by about 0.05 (t statistic = 2.21, Pr = 0.03). When we disaggregate by question type, however, we find that the effect on voter certainty is significantly different from zero only when questions are easy, an estimated difference of 0.10, but has no effect when questions are hard or misleading.[19]

In contrast to the effect of social information on opinions, we find that the difference in certainty is significantly lower after receiving information on success in Success than in Baseline, by about 0.09 (t statistic = 4.62, Pr = 0.00). When we disaggregate by question type, not surprisingly, we find that the effect on voter certainty is significantly different from zero only when questions are misleading (and voters are learning how incorrect earlier voters were on such questions), an estimated difference of 0.27, but has no effect when questions are easy or hard.[20]

Interestingly, we find that in the Baseline Treatment the mean Certainty Difference is significantly higher than in the NonVoting Treatment overall (t statistic = 3.66, Pr = 0.00). When we disaggregate by question type, we find that the main effect appears to occur when questions are misleading, as the difference is not significant for easy or hard questions, although the difference in certainty is always greater in Baseline than in NonVoting.[21]

To determine if these results are robust to including individual variables, we estimated a regression with Certainty Difference as our dependent variable as a function of the same independent variables used to explain voting choices above, which is shown in Table D4 in the Supplemental Online Appendix. The results support the analysis above: Certainty difference increases are significantly higher with social information on opinions when questions are easy and are significantly lower with social information on success when questions are misleading. There is one exception – we find that certainty increases

[18] For the comparison between Baseline 1 and 2 overall the t statistic = 0.54, Pr = 0.59; for easy questions = 0.87, Pr = 0.38; for hard questions = −0.55, Pr = 0.59; and for misleading questions = 0.63, Pr = 0.53. For the comparison of Opinions 1 and 2 the values are −0.53, Pr = 0.59; 1.14, Pr = 0.26; −1.72, Pr = 0.09; and −1.06, Pr = 0.29; respectively. For the comparison of Success 1 and Success 2 the values are 0.71, Pr = 0.48; 1.47, Pr = 0.14; 0.95, Pr = 0.34; and −0.95, Pr = 0.34; respectively.

[19] The t statistic for the comparison for easy questions = 3.49, Pr = 0.00; for hard questions = 0.20, Pr = 0.84; and for misleading questions = 0.13, Pr = 0.90.

[20] The t statistic for the comparison when questions are easy = 0.05, Pr = 0.96; when questions are hard = 1.44, Pr = 0.15; and when questions are misleading = 6.32, Pr = 0.00.

[21] For easy questions the t statistic for the comparison = 1.61, Pr = 0.11; for hard questions = 1.38, Pr = 0.17; and for misleading questions = 3.23, Pr = 0.00.

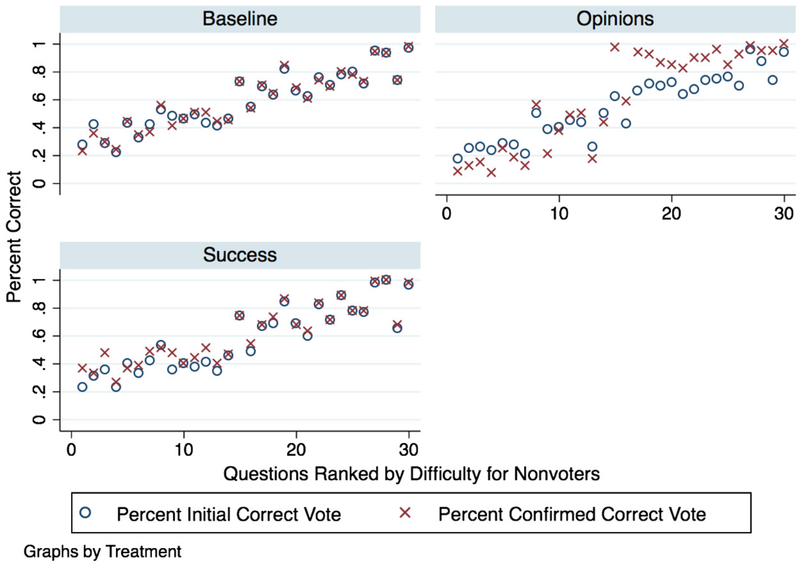

Fig. 7. Percent correct initial choices and percent correct confirmed choices by treatment and question difficulty.

are also significantly higher with social information on opinions when questions are misleading. This result makes sense: biased voters who hear that others hold the same views become even more convinced that they are doing the right thing. Thus we find that voters "think" they are more correct than they are when they receive social information on opinions and questions are misleading. These results are summarized below:

Result 7 *(Voter Certainty and Social Information on Opinions).* Social information on opinions results in a greater increase in voter certainty than without the information when questions are easy or misleading but has no effect when questions are hard.

Result 8 *(Voter Certainty and Social Information on Success).* Social information on success results in a smaller increase in voter certainty than without the information when questions are misleading but has no effect when questions are easy or hard.

4.3.2. Do voters switch answers in response to social information?

We now turn to whether voters actually change their answers in response to social information. Fig. 7 shows the relationship between the percent correct initial choices as compared to the percent correct confirmed choices by treatment by question. As can be seen, in Baseline, where voters do not receive social information, there are few differences between the percentage correct initial choices and confirmed choices. However, we see large differences in these percentages in Opinions, with confirmed correct choices more correct than initial choices for easy questions, but less correct for misleading ones. In Success, we see some suggestive evidence of switching in response to social information to more correct choices for misleading questions, although the differences are not large.

Table D5 in the Supplemental Online Appendix shows in more detail the extent that voters change their answers between initial and confirmed responses by treatment and by question type. We focus on two measures of switching behavior: (a) the percentage of switches (labeled *Percent Switches*) and (b) given an individual has switched, is he or she more likely to switch from an incorrect to correct answer or from a correct to incorrect answer, that is, the percent correct switching (labeled *Percent Correct Switches*). In general, the majority of subjects across treatments do not switch, only 10.35% change their choice between the initial and the final choices. Of those who switch, 63% switch from an incorrect answer to a correct answer, so overall we find that individuals are more likely to switch from an incorrect to a correct choice. We find no significant differences in switching behavior between the NonVoting Treatment and the Baseline voting Treatments (even though the difference in relative switching between NonVoting and Baseline appears large, it is not significant) and no significant difference in switching behavior between Baseline 1 and Baseline 2.[22]

First, as in previous behavior, we find that there are no significant differences in the amount of switching by set of sessions (i.e. comparing Baseline 1 with Baseline 2, Opinions 1 with Opinions 2, and Success 1 with Success 2), even when

[22] For the comparisons between Baseline 1 and 2: For total percent switching for easy questions the z statistic = 1.35, Pr = 0.18 and for hard questions = 0.19, Pr = 0.85. For relative percentage correct switching for easy questions = 0.15, Pr = 0.88 and for hard questions = 0.72, Pr = 0.47. For the comparisons between Baseline and NonVoting: For total percent switches for easy questions = 0.91, Pr = 0.36 and for hard questions = 0.54, Pr = 0.59. For relative percentage correct switching for easy questions = 0.46, Pr = 0.64 and for hard questions = 1.33, Pr = 0.18.

disaggregating by question type.[23] We also find that there are no significant differences in the amount of switching when we compare the subjects in NonVoting with those in Baseline overall and when disaggregating by question type.[24] Not surprisingly, we observe significant differences in switching behavior in Baseline by question type. That is, subjects are 9% likely to switch when a question is hard or misleading, but only 4% likely to do so when a question is easy.[25]

Since only a small percentage of subjects are switching, it is expected that receiving social information should result in significant more switching, which it does in the case of Opinions. Subjects in Opinions switch 17% of the time compared to only 6% of switching in Baseline (z statistic = 12.25, Pr = 0.00) and switch significantly more for each question type.[26] Hence, a significant number, albeit still a small proportion of the subjects, are influenced by social information on opinions to switch their votes.

We also find significantly more switching in Success than in Baseline overall (z statistic = 4.95, Pr = 0.00), with 10% switching overall in Success. However, when we disaggregate by question type, we see that the effect is for hard and misleading questions, not for easy ones, which is as expected.[27] Hence, we find additional evidence that voters do respond to social information on success even if we do not find that it has an effect on group voting choices.

Given that more switching is occurring under Opinions and Success, are these voters more likely to switch from incorrect to correct answers or from correct to incorrect answers? Voters who switch in Opinions are significantly more likely to switch to the correct answer than in Baseline overall (68% in Opinions compared to 50%, z statistics = 4.25, Pr = 0.00). However, the overall effect is apparently driven by easy questions, where the comparison is between (97% in Opinions compared to 53%, z statistic = 8.83, Pr = 0.00). For hard questions, the difference is not significant (54% in Opinions compared to 53%, z statistic = 0.18, Pr = 0.86) and for misleading questions, the difference is significant but in the opposite direction with voters more likely to switch to the incorrect choice (25% in Opinions compared to 45%, z statistic = 2.78, Pr = 0.01). These results support our observation of group and voter choices being heavily influenced to choose incorrectly in response to social information on opinions when questions are misleading.

When we compare group outcomes in Success with Baseline we find no significant differences suggesting that the social information on success did not influence voters. Yet in our analysis of voting behavior we find that some voters are influenced to vote more correctly in response to social information on success and in this subsection we find that voters were significantly more likely to be uncertain in Success than in the other treatments when questions were misleading, suggesting that voters did mentally respond to the social information on success in this case, the case where we would hope such information would de-bias voters. We also find above that voters in Success are significantly more likely to switch than in Baseline when questions are misleading as well. Does this switching lead them to make better choices? We find evidence that indeed, of those that switch, these voters in Success are more likely to switch to the correct answer than those in the Baseline overall and for all question types, suggesting some evidence of de-biasing among these voters. We find that this difference is significant overall (66% in Success compared to 50% in Baseline, z statistic = 3.40, Pr = 0.00) and for misleading questions (63% in Success compared to 46% in Baseline, z statistic = 2.42, Pr = 0.02), but not significant at conventional levels for easy and hard questions.[28] Hence, for a small number of voters, de-biasing does appear to occur, although not significantly enough to affect the overall outcomes.

As with the previous analyses of individual behavior in the experiment, in order to determine if these relationships are robust to controlling for individual differences, we estimate two probit equations reported on in the Supplemental Online Appendix D with the same independent variables above, one for the probability of switching (Table D6) and one for the probability that a given switch is to the correct choice (Table D7).[29] We also present the marginal results from these two estimations in Figs. 8 and 9, respectively.

As in the analysis above, we find that social information on opinions significantly increases the probability of switching for easy and hard questions and has an almost significant effect on switching for misleading questions. For easy questions, subjects are also significantly more likely then to switch to the correct choice but significantly less likely to do so when questions are misleading. Social information on opinions has an insignificant effect on the probability of switching to the correct choice when questions are hard.

[23] For the comparison of Baseline 1 and 2 overall the z statistic = 1.19, Pr = 0.23; for easy questions = 0.53, Pr = 0.60; for hard questions = 1.38, Pr = 0.17; and for missleading questions = 0.19, Pr = 0.85. The values for the comparison of Opinions 1 and 2 are 0.81, Pr = 0.42; 0.20, Pr = 0.84; 0.50, Pr = 0.62; and 0.83, Pr = 0.40; respectively. The values for the comparison of Success 1 and 2 are 1.68, Pr = 0.09; 0.28, Pr = 0.78; 1.61, Pr = 0.11; and 0.99, Pr = 0.32; respectively.

[24] For the comparison overall the z statistic = 0.42, Pr = 0.67; for easy questions = 1.18, Pr = 0.24; for hard questions = 0.11, Pr = 0.91; and for misleading questions = 0.54, Pr = 0.59.

[25] Comparing easy with hard questions or easy with misleading ones yields a z statistic of 4. 87, Pr = 0.00; comparing hard with misleading questions yields 0.13, Pr = 0.90.

[26] For easy questions the z statistic = 11.46, Pr = 0.00; for hard questions = 4.66, Pr = 0.00; and for misleading questions = 4.08, Pr = 0.00.

[27] For easy questions the z statistic = 0.38, Pr = 0.70; for hard questions = 2.89, Pr = 0.00; and for misleading questions = 4.91, Pr = 0.00.

[28] For easy questions z statistic = 1.66, Pr = 0.10 and for hard questions = 1.91, Pr = 0.06.

[29] We also estimated a selection version of these two equations with results qualitatively similar to those reported here albeit some of the controls omitted from the second equation. That is, we estimated a probit model of switching correctly with the first probit serving as the selection equation including all the controls. The χ^2 statistic from the Wald test of independent equations = 1.04, Pr = 0.31. Therefore, we report the estimations of these equations independently.

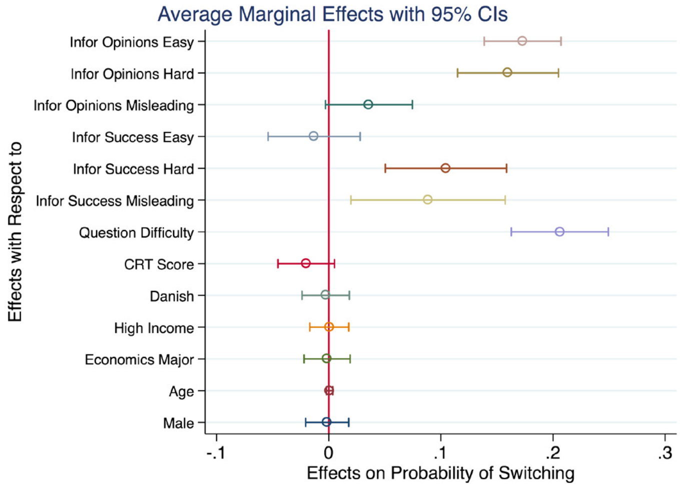

Fig. 8. Marginal effects of social information on probability of switching vote choice.

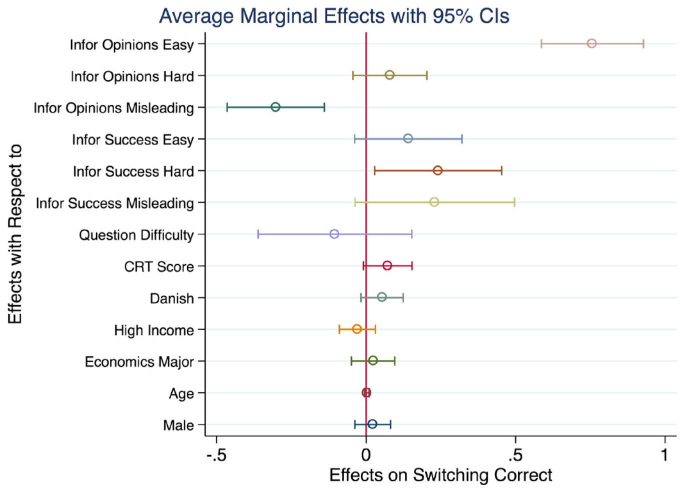

Fig. 9. Marginal effects of social information on probability of switching to correct vote choice if switching.

We find that social information on success significantly increases the probability of switching for hard and misleading questions, but has no significant effect on the probability of switching when questions are easy. Although social information on success has a positive effect on switching to the correct choice for all question types, only when questions are hard is the effect significant, hence we find that including controls reduces the effects of social information on success on the likelihood of switching correctly.

Consistent with our analysis of CRT scores above, we find that subjects with higher CRT scores are less likely to switch, but if they do switch, are more likely to switch to the correct choice. However, these effects are not significant at conventional levels. Our results are summarized below

Result 9 *(Voter Switching in Response to Social Information on Opinions).* Voters who receive social information on opinions are significantly more likely to change their votes than voters who do not receive such information for all question types. Those that switch are significantly more likely to switch to the correct choice when questions are easy but significantly more likely to switch to the incorrect choice when questions are misleading.

Result 10 *(Voter Switching in Response to Social Information on Success).* Voters who receive social information on success are significantly more likely to switch their votes than voters who do not receive such information when questions are hard or misleading. We find some evidence that those who switch are more likely to switch to the correct choice, but the effect is only robustly significant for hard questions.

Result 11 *(CRT Scores and Switching).* Voters with higher CRT scores appear to switch less often, but if they do so, switch more often to the correct choice. But the effects are not significant.

4.4. Cognitive abilities and understanding the impact of social information on voting

We find that social information either on opinions or success can be useful in information aggregation when choices are easy, but not when choices are hard. Even though logically social information on successes should de-bias voters when questions are hard, we found only minor evidence of such an effect.

The finding that cognitive abilities are predictors of how voters choose and whether they switch their votes to more correct choices may offer an explanation for these results. Consider the situation in which a voter receives social information on opinions. The voter learns either that the majority of previous voters agrees with his or her choice or does not agree with his or her choice. If the majority agrees, then it is easy to keep his or her choice the same. If the voter learns that the majority does not agree, then it is relatively easy to say, well maybe I am wrong because the majority is so consistently thinking this way. The voter does not have to think about why it is that maybe the majority disagrees, it is a relatively simple calculus. According to this logic, it makes sense that voters can easily interpret social information on opinions.

Now consider the situation in which a voter receives social information on success. The voter either learns that the majority of previous voters was correct or that the majority was incorrect. If he or she learns that the majority was correct in the past it is easy to reason, well, probably the majority will be correct again, and I do not need to do anything. The voter does not need to take the additional step to think about what this means about the voter's own inferences. But if the voter learns that on average the majority was incorrect the voter has to first figure out that this means that most people make incorrect inferences and that he or she is probably like most people and is also making an incorrect inference. The voter has to think through the implications for inferences of others and his or her own inference. So the level of reasoning for social information on success to influence voter choices is higher for this case. Again, according to this logic, it makes sense that smaller numbers of voters will be affected by social information since fewer voters will have the cognitive abilities to make the necessary inference and many of those who do, may also have been more likely to answer the hard question correctly in their initial response.

Our analysis suggests, then, that social information on opinions is particularly problematic for majority voting outcomes when voters on average make incorrect inferences in particular because the reasoning required of voters is not difficult in order to use such information, but that social information on success does not help alleviate the problems with majority voting when voters on average make incorrect inferences precisely for the same reason that voters often make incorrect inferences in the first place, because the reasoning required for such information to be influential is more cognitively taxing.

Of course there may be other explanations for why voters may not respond to social information on success as theoretically predicted. As noted, we found that very few voters changed their votes in response to social information, even when questions were hard. One possible explanation may lie in voter overconfidence in their own abilities and answers even when given information that most voters are wrong in their answers. A number of studies have found that individuals are generally overconfident in their own abilities (see for example Moore and Healy, 2008). Such overconfidence may also explain why voters do not make the inference that they too may be wrong in their view and correct their votes. The subjects in NonVoting, as we noted previously, exhibit overconfidence when incentivized to estimate their average performance on the questions. Hence, one reason so few voters change their votes is more an overconfidence in their own abilities to answer the questions rather than an inability to make the required inference implicit in the success treatments. That is, subjects may think, well most subjects are incorrect, but that is not true for me. Although, the literature suggests that men are more overconfident than women (see Bordalo et al., 2017), we do not find any evidence of less switching by men in our statistical analyses.

Finally, when voters receive information on success and learn that most voters are incorrect in the baseline sessions they may also infer that their vote is unlikely to be pivotal and thus may not be motivated to think more carefully about their choice. Our analysis depends on voters choosing as if they are pivotal. But if voters infer from the social information that their votes are not pivotal, then they may not care enough to change their votes or how they vote on the question. Finding out that the majority is vastly incorrect may not induce a single voter to sufficiently think through the implications since it also implies that there is little that he or she can do to change the outcome.

Our results do not appear to support this hypothesis as that would imply that there is little switching when questions are misleading (and thus social information implies a lack of pivotality – Opinions subjects learn that the overwhelming majority in other groups voted for one choice and in Success subjects learn that the overwhelming majority in other groups voted incorrectly). That is, we find that social information does lead to significantly more switching in Success when questions are misleading and close to significantly more switching in Opinions. Hence, we find do not find much evidence that social information affects voters' perceptions of pivotality such that they are less likely to switch.

5. Concluding remarks

To err is human. In a democracy, voters will often be uncertain about what is the right course of action, and be more or less prone to erroneously support inefficient policies. But the existence of such uncertainty and error does not imply that democracy is necessarily doomed to systematically select inefficient policies. This paper shows experimentally that majority voting is beneficial (has a "bright side") in the sense that democratic choice can be superior to the average voter's opinion if it aggregates information effectively. Majority voting has a "bright side" even when almost all voters are uncertain and when many err, as long as a majority of voters is more likely to be right than wrong about what policy to choose. Social information on opinions (the popularity of alternatives in the electorate) and on how successful democratic choice tends to be makes the "bright side" shine even brighter, i.e. further improves efficiency, or does at least not harm.

Yet, our experiment also shows that the "dark side" of the vote is looming. We find that voting is counterproductive when the average voter is biased (is more likely to be wrong than right), and that social information on opinions can further exacerbate the perverse effect of majority voting. Counter to theoretical predictions, we find that voters are not enlightened (i.e. do not make clearly better choices) when they learn about how bad choices in other electorates were on the same issue. Thus, we find that voters are not effectively de-biased by such information, probably because voters are overconfident, are not aware of their biases and de-biasing requires substantial cognitive skills.

We were able to produce these findings by virtue of a novel experimental design. Previous experimental studies of information aggregation in voting had designs that were chosen to minimize the possibility of incorrect inferences and have therefore not been able to study the dark side of the vote or how it is shaped by social information. In contrast, our design allows us to bring a series of issues before voters which all have a clear correct answer and we know (but voters do not know) for which of these issues most individuals tend to make correct inferences or systematically biased judgments.

We think our results should be read as a warning against the belief that majority voting will in all cases be beneficial in that it yields superior choices due to efficient information aggregation. Such a belief may be nurtured by theoretical accounts (based on the Condorcet Jury Theorem), but they often use psychologically unrealistic assumptions.[30] But our results should not be read as saying that democratic (majoritarian) choice is necessarily doomed (it is not, there is a "bright side") nor that majoritarian choice should be rejected in cases where it is likely to aggregate information inefficiently.

Caution in interpretation is warranted on at least two grounds. First, our paper focuses on the ability of majority voting to select the best solution when one exists (the "epistemic" quality of democracy). This issue seems relevant in situations such as when the board of a company decides on investing in product A or B, or a jury decides on whether a defendant is guilty or not. Clearly, majoritarian choice has other benefits than aggregating information,[31] and has other drawbacks than failure to aggregate information inefficiently in specific circumstances (like the exploitation of minorities by majorities, see for example Gerber et al., 1998). Thus, our results provide just one – we think an important one – aspect in the debate on the pros and cons of majoritarian choice. Second, when evaluating majoritarian choice, it needs to be compared to other (realistic) alternatives, which also have their pros and cons.

Our design to study the effects of biased choices can be used to study further aspects of the dark side of the vote or information aggregation more generally. We think interesting avenues for further research on information aggregation in majority voting are selective participation, other forms of social communication preceding voting, and other voting rules.

Selective participation and abstention are important aspects of many democratic choices, and may shape the quality of democratic choice in important ways (e.g. Bhattacharya et al., 2012). In our experiment, voting was compulsory and therefore biased and non-biased voters were equally likely to participate. Suppose that participation and voter competence is correlated. For example, biased voters may be somehow aware that they are biased and abstain while non-biased voters may participate at higher rates. If so, the dark side documented here may be mitigated. But our data suggests that is not likely to happen. Those who were wrong were not less confident than those who were right.

Other forms of social information may reduce or even eliminate the dark side of the vote. Suppose we had informed voters in our experiment about what policies other groups had chosen *and* how successful these policies were. We think it is quite likely that voters would have made near-perfect choices in this case. Such an effect is likely in our design because the groups and issues were identical, but is not guaranteed to obtain in more complicated settings, e.g. when experience with a particular policy in one state may only be a noisy predictor of success of the same policy elsewhere because states differ (see Sausgruber and Tyran, 2005 for an experimental investigation of policy emulation). Other forms of communication may or may not be helpful (Goeree and Yariv, 2010). Deliberation among non-experts or when experts cannot persuade others of their superior knowledge is not necessarily a remedy and may result in group think, i.e. on agreeing on an arbitrary policy, not necessarily the correct one (see Sausgruber and Tyran, 2011 for an experiment with free communication preceding voting on taxes in a market). However, credible experts may be game-changers. One way to become credible is to establish a proven track record of superior judgment (see Penczynski, 2016).

Alternative voting rules may mitigate the dark side of the vote. For example, point voting may restore informational efficiency if voters assign points (out of a budget of 100, say) to alternatives according to how certain they are to be right. This would be the case if the certainty correlates well with competence and voting is non-strategic (but it is well-known that voters are insincere with point voting, e.g. Nitzan et al., 1980).[32] Markets may fare better than voting in aggregating information, because the marginal rather than the median person drives the outcome. But biases also seem to beset such markets at least in some instances (e.g. Ganguly et al., 2000; Snowberg and Wolfers, 2010).

[30] We do not know to what extent such theoretical accounts shape the faith of politicians and lay people in the ability of majoritarian choice. Casual observation suggests that perceptions are rather mixed. Winston Churchill seems not to have had much faith in that ability, judging from his quip that "The best argument against democracy is a five-minute conversation with the average voter."

[31] For example, it may help to hold self-interested elites in check, allow voters to participate in decision making and to express their preferences, or increase compliance by improving the legitimacy of policy choices (see e.g. Dal Bó et al., 2010; Markussen et al., 2014 for experimental studies).

[32] See Guarnaschelli et al. (2000) for an experimental comparison of information aggregation under majority rule with unanimity.

Our experiments suggest that models of information aggregation through majority voting and associated experimental work should take the effects of biased voters on the efficiency of group choices more seriously. Our findings suggest that the dark side of the vote is real, and social information can play an important, and surprising, role in shaping it.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geb.2018.10.008.

References

Ali, S. Nageeb, Goeree, Jacob K., Kartik, Navin, Palfrey, Thomas R., 2008. Information aggregation in standing and ad hoc committees. Am. Econ. Rev. 98 (2), 181–186.

Austen-Smith, David, Banks, Jeffrey S., 1996. Information aggregation, rationality, and the Condorcet jury theorem. Am. Polit. Sci. Rev. 90 (1), 34–45, March.

Baker, Keith (Ed.), 1976. Michael de Condorcet, Marquis. Selected Writings. Bobbs-Merrill, Indianapolis.

Battaglini, Marco, Morton, Rebecca, Palfrey, Thomas, 2007. Efficiency, equity, and timing of voting mechanisms. Am. Polit. Sci. Rev. 101 (3), 409–424, August.

Battaglini, Marco, Morton, Rebecca, Palfrey, Thomas, 2008. Information aggregation and strategic abstention in large laboratory elections. Am. Econ. Rev. 98 (2), 194–200.

Battaglini, Marco, Morton, Rebecca, Palfrey, Thomas, 2010. The swing voter's curse in the laboratory. Rev. Econ. Stud. 77 (1), 61–89, January.

Bhattacharya, S., Duffy, J., Kim, S.-T., 2012. Compulsory vs. Voluntary Voting. Working paper. University of Pittsburgh Economics Department. May 14.

Bikhchandani, Sushil, Hirshleifer, David, Welch, Ivo, 1998. Learning from the behavior of others: conformity, fads, and informational cascades. J. Econ. Perspect. 12 (3), 151–170.

Blinder, Alan S., Morgan, John, 2005. Are two heads better than one? Monetary policy by committee. J. Money Credit Bank. 37 (5), 798–811.

Bordalo, Pedro, Coffman, Katherine B., Gennaioli, Nicola, Shleifer, Andrew, 2017. Beliefs About Gender. National Bureau of Economic Research, No. w22972.

Bottom, William P., Ladha, Krishna, Miller, Gary J., 2002. Propagation of individual bias through group judgment: error in the treatment of asymmetrically informative signals. J. Risk Uncertain. 25 (2), 147–163.

Charness, Gary, Sutter, Matthias, 2012. Groups make better self-interested decisions. J. Econ. Perspect. 26 (3), 157–176.

de Condorcet, Marquis, 1785. Essay on the Application of Mathematics to the Theory of Decision-Making. Reprinted in *Condorcet: Selected Writings*, pp. 33, 62 (Keith Michael Baker ed., 1976).

Coughlan, Peter, 2000. In defense of unanimous jury verdicts: mistrials, communication, and strategic voting. Am. Polit. Sci. Rev. 94 (June), 678–690.

Dal Bó, P., Foster, A., Putterman, L., 2010. Institutions and behavior: experimental evidence on the effects of democracy. Am. Econ. Rev. 100 (5), 2205–2229.

Elbittar, Alexander, Gomberg, Andrei, Martinelli, Cesar, Palfrey, Thomas R., 2018. Ignorance and bias in collective decisions. J. Econ. Behav. Organ. https://doi.org/10.1016/j.jebo.2016.12.011. In press.

Estlund, David M., 1994. Opinion leaders, independence, and Condorcet's jury theorem. Theory Decis. 36, 131–162.

Feddersen, Timothy, Pesendorfer, Wolfgang, 1998. Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. Am. Polit. Sci. Rev. 92 (1), 23–36, March.

Fehr, Ernst, Tyran, Jean-Robert, 2005. Individual irrationality and aggregate outcomes. J. Econ. Perspect. 19 (4), 43–66.

Fischbacher, Urs, 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.

Frederick, Shane, 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19 (4), 25–42.

Ganguly, Amanda R., Kagel, John H., Moser, Donald V., 2000. Do asset market prices reflect traders' judgment biases? J. Risk Uncertain. 20 (3), 219–245.

Gerber, Elisabeth, Morton, Rebecca, Rietz, Thomas, 1998. Minority representation in multimember districts. Am. Polit. Sci. Rev. 92 (1), 127–144, March.

Gneezy, Uri, Kapteyn, Ari, Potters, Jan, 2003. Evaluation periods and asset prices in a market experiment. J. Finance 58 (2), 821–837.

Goeree, Jacob, Yariv, Leeat, 2010. An experimental study of collective deliberation. Econometrica 79 (3), 893–921.

Greiner, Ben, 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1, 114–125.

Großer, Jens, Seebauer, Michael, 2016. The curse of uninformed voting: an experimental study. Games Econ. Behav. 97 (C), 205–226.

Guarnaschelli, Serena, McKelvey, Richard D., Palfrey, Thomas R., 2000. An experimental study of jury decision rules. Am. Polit. Sci. Rev. 94 (June, 2), 407–423.

Hoorens, Vera, 1993. Self-enhancement and superiority biases in social comparison, in Stroebe, W., Hewstone, Miles (Eds.). Eur. Rev. Soc. Psychol. 4 (1), 113–139.

Hung, Angela, Plott, Charles R., 2001. Information cascades: replication and an extension to majority rule and conformity-rewarding institutions. Am. Econ. Rev. 91 (5), 1508–1520.

Kerr, Norbert L., MacCoun, Robert J., Kramer, Geoffrey P., 1996. Bias in judgment: comparing individuals and groups. Psychol. Rev. 103 (4), 687–719.

Ladha, Krishna, Miller, Gary, Oppenheimer, Joseph, 1996. Information Aggregation by Majority Rule: Theory and Experiments. Washington University. Typescript.

Markussen, Thomas, Putterman, Louis, Tyran, Jean-Robert, 2014. Self-organization for collective action: an experimental study of voting on sanction regimes. Rev. Econ. Stud. 81 (1), 301–324.

McLennan, Andrew, 1998. Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents. Am. Polit. Sci. Rev. 92 (2), 413–418, June.

Mechtenberg, Lydia, Tyran, Jean-Robert, 2016. Voter Motivation and the Quality of Democratic Choice. Working paper 16-13. Department of Economics, U Copenhagen.

Moore, Don A., Healy, Paul J., 2008. The trouble with overconfidence. Psychol. Rev. 115 (2), 502.

Morton, Rebecca B., Tyran, Jean-Robert, 2011. Let the experts decide? Asymmetric information, abstention, and coordination in standing committees. Games Econ. Behav. 72, 485–509.

Morton, Rebecca B., Williams, Kenneth C., 1999. Information asymmetries and simultaneous versus sequential voting. Am. Polit. Sci. Rev. 93 (1), 51–67.

Nitzan, S., Paroush, J., Lampert, S., 1980. Preference expression and misrepresentation in point voting schemes. Public Choice 35 (4), 421–436.

Penczynski, Stefan P., 2016. Persuasion: an experimental study of team decision making. J. Econ. Psychol. 56, 244–261.

Sausgruber, Rupert, Tyran, Jean-Robert, 2005. The diffusion of policy innovations. J. Evol. Econ. 15 (4), 423–442.

Sausgruber, Rupert, Tyran, J.-R., 2011. Are we taxing ourselves? How deliberation and experience shape voting on taxes. J. Public Econ. 95, 164–176.

Shaw, Marjorie E., 1932. Comparison of individuals and small groups in the rational solution of complex problems. Am. J. Psychol. 44 (3), 491–504.

Slembeck, Tilman, Tyran, Jean-Robert, 2004. Do institutions promote rationality?: an experimental study of the three-door anomaly. J. Econ. Behav. Organ. 54 (3), 337–350.

Snowberg, Erik, Wolfers, Justin, 2010. Explaining the favorite–long shot bias: is it risk-love or misperceptions? J. Polit. Econ. 118 (4), 723–746.

Wit, Jorgen, 1998. Rational choice and the Condorcet jury theorem. Games Econ. Behav. 22 (2), 364–376, February.