



The expressive power of voting rules

Sacha Bourgeois-Gironde^{1,2,3} · João V. Ferreira^{2,4} 

Received: 10 October 2022 / Accepted: 22 August 2023
© The Author(s) 2023

Abstract

Different voting rules are commonly used to settle collective decisions. A promising way to assess voting rules, for which little is known, is to compare the expressive utility that voters derive from voting with each rule. In this paper, we first propose a simple ordinal model of expressive voting that allows us to compare voting rules in terms of the expressive utility that voters can derive from voting (their **expressive power**). Our model provides a novel testable implication according to which expected turnout increases with expressive power. We then ran an online experiment testing this implication in a controlled environment. We find that if voters are made aware of alternative voting rules, turnout is higher in voting rules with higher expressive power. Our results also show that higher expressive power is associated with a better representation of voters' actual preferences and, according to our model, higher expressive utility. This suggests that the expressive power of voting rules is a relevant criterion when choosing between voting rules for economic and political decisions.

1 Introduction

Voting rules are widely used to settle collective decisions in elections, committees, teams, among many other settings. They set the rules of how individual preferences are aggregated into a social preference, usually to select a winner.¹ Choosing which rule to use can be a consequential decision. A large literature has therefore compared voting rules across different domains and with different methods. For example,

¹ Voting rules differ from electoral systems which define a wider set of rules for elections, such as who is allowed to vote, limits on campaign spending, how proportional the system is, and so on. We limit our attention to single-winner voting rules.

✉ João V. Ferreira
j.ferreira@soton.ac.uk

¹ Université Panthéon-Assas, Paris II, Cred, France

² Institut Jean Nicod, Ecole Normale Supérieure, PSL-Research University, Paris, France

³ Faculty of Law, University of Haifa, Haifa, Israel

⁴ University of Southampton, Building 58, Highfield Campus, Southampton SO17 1BJ, UK

social choice theory has compared voting rules by proposing desirable axiomatic properties such as the ability to always elect the Condorcet winner or the ability to resist strategic manipulation (e.g. Sen 1995; Brams and Fishburn 2002; Laslier 2009), while a growing experimental literature compares voting rules in terms of their impact on voting behaviour and electoral outcomes (e.g. Forsythe et al. 1996; Van der Straeten et al. 2010a; Alós-Ferrer and Granic 2012; Baujard et al. 2013, 2021; Bouton et al. 2016, 2021; Moreno et al. 2019; Tsakas and Xefteris 2021).

A promising complementary way to assess voting rules is to compare the *expressive utility*—the intrinsic benefit derived from expressing one’s preferences/attitudes/values—that voters derive from voting with each rule. A desire to express oneself, *for its own sake*, is one of the most cited motivations to underlie turnout and voting behaviours.² However, to the best of our knowledge, little is known about the relationship between expressive utility and voting rules. Improving our understanding of this issue is important because different voting rules are widely used in companies, committees, elections, referendums and opinion polls, among many other settings, and they may have a significant impact on turnout and voters’ welfare if there are significant differences in terms of the expressive utility that voters derive from voting with different rules.

In this paper, we first propose a conceptual framework that allows us to compare voting rules in terms of the expressive utility that voters can derive from voting. Our main hypothesis is that voters’ expressive utility is highest when they express their actual preferences over alternatives, and decreases as the preferences cast in the ballot get more distant from their actual preferences. This distance can increase, in particular, because a voting rule does not allow a voter to express her actual preference. For instance, a voter that strictly prefers x to y and y to z cannot express this preference in an election decided by *plurality rule*.³ We then argue that voting rules differ in terms of their **expressive power**, in the sense that voting rule A has more expressive power than voting rule B if A allows to express more preferences than B and all preferences that are possible to express with B can be expressed with A. It follows that if voting rule A has more expressive power than voting rule B, then voters’ expressive utility with A is equal or higher than with B, all else equal. This simple model provides a novel testable implication for the relationship between voting rules, expressive utility, and turnout: if voting rule A has more expressive power than voting rule B, then expected turnout with voting rule A is equal or higher than expected turnout with voting rule B, *ceteris paribus*.⁴

To see this, consider plurality rule and *approval voting*.⁵ Approval voting allows to express all preferences that are possible to express with the plurality rule, in addition to some that are not possible to express with the latter rule. Therefore, we say

² See Igersheim et al. (2016b), Tyran and Wagner (2019) and Hamlin and Jennings (2019) for reviews.

³ With plurality rule, each voter can give one point to one option and zero to all others, and the option with the most points wins.

⁴ The *ceteris paribus* condition is relevant because other features of voting rules, such as their complexity and their ability to resist strategic voting, can influence turnout. We address these issues below.

⁵ With approval voting, each voter can give one point to one or more options and zero to all others, and the option with the most points wins.

that approval voting has more expressive power than the plurality rule. This means that—according to our model—there should be no voter for whom her expressive utility is higher with plurality rule than with approval voting, while expressive utility may be higher for some voters with approval voting than with plurality rule. It follows that we can make an unambiguous comparison in terms of expected turnout, according to which expected turnout with approval voting is equal or higher than with plurality rule, all else equal. Note, however, that this type of unambiguous comparison is not always possible. For instance, we cannot make an unambiguous comparison in terms of expressive power (and expected turnout) between plurality rule and *Borda rule*.⁶

We then ran an online experiment to test the model's implications in a controlled environment. In May 2020, we recruited 1977 UK subjects to participate in an opinion poll and asked them to vote for different policies on how to “open up” the country after the first Covid-19 lockdown. However, contrary to a standard opinion poll, participants only *registered their vote* (i.e., their vote was counted for the final public results) if they were willing to respond to a series of additional questions that were not asked in case of abstention. This feature of the opinion poll made registering one's vote (our proxy for turnout) costly. To test the model's main implication, participants were randomly assigned to three voting rules with different degrees of expressive power: plurality rule (Treatment 1, or T1), approval voting (T2), and a synthetic rule that allowed participants to express all possible preferences with approval voting and the Borda rule (hereafter A-Borda rule, T3). In two additional treatments (T4 and T5), subjects participated in the same treatments as in T1 and T3 respectively, with the only difference being that they were made aware of different voting rules before they made their turnout decision. This experimental design allows us to test if voting rules with higher expressive power lead to higher turnout, and if awareness of different voting rules is necessary for this to hold.

We find no statistically significant difference between turnout in T1, T2, and T3. This contrasts with our predictions. However, once participants are made aware of different voting rules, turnout is statistically and “economically” significantly higher with the voting rule with higher expressive power (82.99% in T5 against 76.79% in T4). Our interpretation of these results is the following. For T1 and T2, the null result seems to be explained by the low percentage of subjects who were able to better express themselves in T2 than in T1. In fact, few subjects were indifferent between options in our setting (something we discuss below). For T1/T2 and T3, we highlight three potential explanations. First, increased rule complexity in T3 can be partly to “blame”. Subjects found the A-Borda rule to be the most complex rule, and our results suggest that rule complexity negatively affects turnout in our setting. Second, strategic voting can have a similar effect. There are more opportunities to vote strategically in T3 than in T1/T2, and our results suggest that non-sincere voters (such as strategic and inattentive voters) are less likely to turn out than non-strategic

⁶ With Borda rule, each voter can give zero points to the least preferred option, one point to the penultimate ranked option, and so on. The option with the most points wins. Below, we discuss an extension of our criterion that allows us to unambiguously compare all voting rules when the empirical distribution of voters' actual preferences is known.

sincere voters. It follows that non-sincere voting can partly explain why turnout is not higher in T3 than in T1/T2.⁷ Third, there may be a bias in favour of more familiar (and/or “real”) rules. In particular, we conjectured that if voters were not able to compare the voting rule they were assigned to with other rules (i.e., without explicit reference points), the ability/inability to better express oneself would not be salient and its effect on the expressive utility of voting muted. This factor, contrary to rule complexity and non-sincere voting, can explain our findings for T4 and T5. In those treatments, we find that simple awareness of alternative voting rules is enough to change voters’ turnout behaviour and that, in the presence of reference points, they are more likely to vote with voting rules that allow them to better express themselves. In these conditions, subjects can use the rules that were presented to them but with which they did not vote as reference points to judge the voting rule with which they voted. Then, according to this interpretation, the ability/inability to express oneself becomes salient and its effect on expressive utility is magnified. Taken together, our results suggest that the expressive power of voting rules can have implications for participation as long as voters are aware of (and can compare) different voting rules.

Our results also show that subjects were better able to express their actual preferences in voting rules with higher expressive power (72% in T3, 67.5% in T5, 7.5% in T2, 3.3% in T4, and 2.3% in T1 expressed their actual preference). This difference is due to the fact that most participants had strict preferences over alternatives that they could express under T3 and T5. According to our model, this suggests that subjects derived higher expressive utility from voting with rules with higher expressive power. A series of questions at the end of the experiment support this interpretation. We find that subjects regarded that they were better able to express themselves under the rules with higher expressive power. We also find that a large majority of subjects (68.4%) would have liked to have used the A-Borda rule when they were asked to choose between plurality/approval/A-Borda *ex-post*.

These results suggest that the expressive power of voting rules is a relevant criterion when deciding which voting rule to use in economic and political decisions. In particular, these results suggest that, all else equal, the higher the expressive power of a voting rule (i) the higher the turnout is expected to be if voters are aware of different voting rules, (ii) the better the voting rule represents individuals’ actual preferences, and according to our model (iii) the higher the expressive utility derived from voting is expected to be. If this holds more generally, a voting rule with higher expressive power is desirable for the sake of higher turnout, expressive utility, and representation of voters’ *actual preferences*. These are relevant goals when considering voting and turnout behaviours, which may have significant implications in terms of voters’ welfare.

⁷ We cannot distinguish between strategic and inattentive voters in our data. Similarly, we cannot distinguish strategic voting from related behaviours such as “satisficing voting” (Kube and Puppe 2009) or “heuristic voting” (Van der Straeten et al. 2010b). However, our results are sufficient to show that the joint effect of all non-sincere voting runs contrary to the effect of expressive power. We discuss these results in Sects. 5 and 6.

1.1 Relationship to the literature

Our paper makes several contributions to the literature. First, it contributes to the growing literature that compares different voting rules (e.g. Merrill 1984; Chamberlin 1985; Nurmi 1987; Sen 1995; Forsythe et al. 1996; Brams and Fishburn 2002; Conitzer and Sandholm 2005; Brams and Fishburn 2007; Laslier 2009; Van der Straeten et al. 2010a; Alós-Ferrer and Granic 2012; Baujard et al. 2013; Plassmann and Tideman 2014; Bouton et al. 2016; Moreno et al. 2019; Baujard et al. 2021; Bouton et al. 2021; Tsakas and Xefteris 2021). Social choice theorists have studied this topic profusely, proposing many desirable axiomatic properties of voting rules (e.g. Sen 1995; Brams and Fishburn 2002; Laslier 2009). Other authors have proposed empirical criteria to evaluate these rules, such as the communication complexity and cognitive load associated with different voting rules (Conitzer and Sandholm 2005). More recently, several authors have used experimental evidence to bring new insights into this subject. For example, Bouton et al. (2021) use a lab experiment to study Duverger's prediction that the plurality rule leads to higher coordination of votes on a limited number of candidates than the majority runoff rule. We contribute to this large body of literature by providing a novel criterion to evaluate voting rules and by testing some of its implications experimentally.

Our analysis also contributes to the expressive voting literature. The central hypothesis of this literature is that individuals may vote as an act of expressive behaviour, receiving direct utility from the act of voting itself instead of, or in addition to, instrumental utility related to material self-interest (e.g. Fiorina 1976; Brennan and Buchanan 1984; Carter and Guerette 1992; Brennan and Lomasky 1993; Brennan and Hamlin 1998; Tyran 2004; Shayo and Harel 2012; Igersheim et al. 2016; Spenkuch 2018; Pons and Tricaud 2018; Ginzburg et al. 2022). Many models, however, do not make any assumptions on the non-consequential expressive element associated with voting, and therefore have no choice-based testable implications (see e.g. Green and Shapiro 1994). One exception is the *low-cost theory of expressive voting* (Brennan and Lomasky 1993). According to this theory, expressive motives become more important relative to self-interested motives when pivot probabilities are small (e.g. large elections). Assuming that expressive motives are more prosocial than self-interested ones, it follows that when pivot probabilities decrease, voters can obtain the satisfaction of "behaving charitably" much cheaper (Tullock 1971, p. 389; for experimental evidence, see e.g. Carter and Guerette 1992; Tyran 2004; Feddersen et al. 2009; Shayo and Harel 2012; Ginzburg et al. 2022).

Our contribution to this literature is four-fold. First, we put forward a conceptual framework according to which the expressive utility derived from voting depends on the ordinal distance between voters' actual preferences and the preferences that they express when voting. As explained below, this simple model has testable implications for individual turnout behaviour which we test in our data. Second, we propose a new formal criterion to compare voting rules in terms of their degree of expression. Other authors have noted that different voting rules allow for different degrees of expression (e.g. Brams and Fishburn 2005; Baujard and Igersheim 2009; Alós-Ferrer and Granic 2012). Aragonés et al. (2011), for instance, used a model of "voting as making statements" to show that approval voting allows voters

to make a richer set of statements than the plurality rule. We contribute to this body of literature by putting forward a new formal ranking of voting rules that is based on the ordinal preferences that individuals can express under different rules.⁸ Third, our study adds up to the recent experimental literature that looks at the relationship between voting rules and expressive voting. Ginzburg et al. (2022) use a theoretical model and a lab experiment to show that voting rules that reduce the probability of a committee member being pivotal may raise the share of expressive votes; Baujard et al. (2021) conducted an experiment in parallel with the 2017 French presidential election to study if voters' behaviour and the outcome of elections are sensitive to modifications of the grading scale employed in evaluative voting that allow—according to the authors' definitions—for greater expressivity. While related, our focus is significantly different. Most notably, while the other studies are interested in voting behaviour and the outcome of elections, we investigate if different voting rules offering different levels of representation of voters' actual preferences can lead to different levels of turnout. Fourth, our experimental results indicate some conditions under which differences in possibilities to better express oneself may or may not influence turnout. In particular, they suggest that awareness of alternative voting rules (i.e., the presence of explicit reference points) is necessary to increase participation. In other settings, such as committees or elections, one can easily raise awareness of alternative voting rules and provide reference points. For example, if a national voting reform increases expressive power, the government could link the reform with an informational campaign that would provide information about the old and the new rules and, in a more direct way than done in our setting, highlight the expressive benefit of the latter.

Third, our study is related to the literature devoted to voters' preferences over electoral systems and voting rules (e.g. Bosch and Orriols 2014; Blais et al. 2015; Bol et al. 2018; Müller and Jankowski 2019; Engelman et al. 2020; Weber 2020; Blais et al. 2021; Hoffmann and Renes 2022; Bol et al. 2023). Of these, the survey-based experiment by Blais et al. (2021) and the lab experiment by Bol et al. (2023) are the closest to ours. Blais et al. (2021) randomly assign participants to vote in hypothetical elections with either plurality or Borda rule, and ask them about their satisfaction with the voting rule after voting. Contrary to their hypothesis that expressing a full ranking increases satisfaction with the rule, they find little difference in overall satisfaction between treatments. In Bol et al. (2023), subjects experience elections under plurality and approval voting (first stage), and then decide which rule they want to use for extra elections (second stage). The authors find that subjects who hold more (less) egalitarian values for society are more likely to choose approval voting (plurality) in the second stage when it has produced a more (less) egalitarian distribution of payoffs than plurality (approval voting) in the first stage. In our setting, turnout is our main outcome of interest (as opposed to ex-post satisfaction or preference over different rules). All these outcomes—turnout,

⁸ Our ranking of voting rules would differ from an extended version of Aragonés et al. (2011) ranking if their ranking was extended to other rules (see fn. 5 in their paper). See Piketty (2000) for a comparison of electoral systems based on a model where voters trade off between present strategic motives and communicating their views to impact future elections via their influence on other voters' opinions.

ex-post satisfaction, and preference over rules—seem to be potentially affected by the “expressiveness” of the voting rules, but also by other factors such as the complexity of the rule (as acknowledged by Blais et al. 2021 for the case of satisfaction). In our view, these outcomes are relevant complements to understand the relationship between expressive utility and voting rules. Our conceptual framework and our results also suggest that it is important to consider people’s actual preferences over alternatives and their reference points in terms of rules when comparing the perceived expressiveness of the rules and people’s satisfaction with or preference over different voting rules.

Finally, our paper contributes to the literature in political science on the impact of alternative electoral systems on turnout (e.g. Ladner and Milner 1999; Blais 2006; Blais and Aarts 2006; Kartal 2015; Sanz 2017). Ladner and Milner (1999) is a seminal empirical study that exploits the fact that different communes in Switzerland use different electoral systems. The authors reveal that a proportional representation system entails higher turnout than a majoritarian one in this setting. This is the case whatever voting rule is associated with the representative system. Sanz (2017) offers another example. For local Spanish elections, which are held under a proportional representation system, the national law demands localities, depending on their population size, to either use a closed lists system (CLs) or an open list system (OLs). Under a CLs voters vote for a party list, while under an OLs voters can vote for candidates from a single party list but also from different party lists. Sanz (2017) reports that an OLs increases turnout compared to a CLs. One potential explanation for this effect is that the OLs captures the vote of “conflicted voters” who identify with more than one party (see Bourgeois-Gironde and Ferreira 2020). Our paper complements this literature by showing that different voting rules (an element of electoral systems) can entail different levels of turnout independently of their effective representational consequences.

2 Conceptual framework

Consider an individual i in a population of N individuals that can either cast a vote or abstain from voting. Let $P_i^1(u_i)$ represent the probability that individual i casts a vote when her utility from voting is u_i , where $P_i^1(u_i)$ is a monotonically non-decreasing function that takes values in $[0, 1]$. The higher u_i , the higher $P_i^1(u_i)$. Conversely, let $P_i^0(u_i)$ represent the probability that individual i abstains. Clearly, $P_i^0(u_i) + P_i^1(u_i) = 1$ for each $i \in N$.

Let X be the set of alternatives available for voting (e.g. candidates in an election or options in an opinion poll). A preference R_i is a collection of ordered pairs over these options. We say that (x, y) belongs to R_i if and only if x is at least as good as y from the viewpoint of R_i . Then, x is said to be indifferent to y for R_i if both (x, y) and (y, x) belong to R_i , and x is said to be strictly better than y for R_i if (x, y) belongs to R_i but (y, x) does not. For simplicity, we assume preferences to be *reflexive* [i.e., (x, x) belongs to R_i for every x in X] and *transitive* [i.e., if (x, y) and (y, z) belong to R_i then (x, z) belongs to R_i for any x, y, z in X]. Preferences are not necessarily

complete [i.e., (x, y) or (y, x) do *not* necessarily belong to R_i for every distinct x and y in X] nor *antisymmetric* [i.e., (x, y) and (y, x) may belong to R_i even if $x \neq y$].

We distinguish between two types of preferences. Individual i 's **actual preference**, denoted R_i^a , is the individual own ranking of alternatives. Individual i 's **expressed preference**, denoted R_i^e , is the ranking that i casts when voting. In other words, expressed preferences are the preference that an observer can derive from looking at the ballot. Assume that voters are attentive and sincere, in the sense that they do not make mistakes and they vote as closest as possible to their actual preference. Even so, R_i^a and R_i^e may differ. The underlying reason is that voting rules only allow i to express a subset of all possible preferences over X . To see this, let \mathcal{P}_X be the set of all possible (reflexive and transitive) preferences over X , and \mathcal{P}_A and \mathcal{P}_B be the set of preferences that voting rules A and B allow to be expressed over X respectively. It is easy to check that $\mathcal{P}_A, \mathcal{P}_B \subseteq \mathcal{P}_X$. For instance, the plurality rule only allows to express preferences in which one alternative is strictly preferred to all other alternatives and those other alternatives are indifferent between themselves; if $X = \{x, y, z\}$, the plurality rule allows i to express a preference R_i^e such that (x, y) and (x, z) belong to R_i^e (x is selected in the ballot), but it does not allow i to express a preference R_i^e such that (x, y) and (y, z) belong to R_i^e . But then, if (x, y) and (y, z) belong to R_i^a , it follows that R_i^a and R_i^e will differ. If, on the other hand, *only* (x, y) and (x, z) belong to R_i^a , then R_i^a and R_i^e will coincide.⁹ It is also easy to check that \mathcal{P}_A and \mathcal{P}_B may differ. For example, the plurality rule allows i to express R_i^e such that *only* (x, y) and (x, z) belong to R_i^e while the Borda rule does not, and the Borda rule allows i to express R_i^e such that (x, y) and (y, z) belong to R_i^e while plurality rule does not.

We assume that individual i 's utility of voting with expressed preference R_i^e , $u_i(R_i^e)$, is given as follows:

$$u_i(R_i^e) = pB_i(R_i^e) - C_i + E_i(R_i^e) \quad (1)$$

where p denotes the probability that i 's vote decides the election (i 's vote is pivotal), $B_i(R_i^e)$ denotes the material benefit (e.g. consumption) that i receives with the selection of the preferred option in R_i^e among the options for which the voter is pivotal, C_i denotes the costs associated with voting such as the time and effort to cast a vote, and $E_i(R_i^e)$ denotes i 's expressive utility of voting R_i^e . Note that $E_i(R_i^e)$ is a *non-consequential* element that is independent of whether the vote affects the outcome of the election, which means that even if p is negligible, as is the case in most large elections, it may still be in i 's best interest to vote.¹⁰

In this model, we emphasise how the expressive utility of voting $E_i(R_i^e)$ may depend on the (ordinal) distance between the preference an individual expresses when voting (R_i^e) and her actual preference (R_i^a):

⁹ Actual and expressed preferences may differ even if the voting rule allows i to express her actual preference for reasons such as inattention and strategic voting. We discuss these issues below.

¹⁰ In some voting rules, R_i^e may rank two or more options equally (e.g. approval voting). In this case, we assume that the material benefit $B_i(R_i^e)$ is the same no matter which of these options is selected. This assumption is inconsequential to our analysis. In general, our representation of i 's utility is similar to the seminal paper of Riker and Ordeshook (1968).

$$E_i(R_i^e) = \max(E_i^* - K(R_i^e, R_i^a), 0) \tag{2}$$

where $E_i^* \geq 0$ is i 's expressive utility of casting a vote with $R_i^e = R_i^a$, and $K(R_i^e, R_i^a)$ is the *Kemeny distance* between i 's expressed preference R_i^e and i 's actual preference R_i^a . The Kemeny distance between two preferences R_i and R_j is given by $K(R_i, R_j) = \#[R_i \setminus R_j \cup R_j \setminus R_i]$, where $\#$ denotes the cardinality of a set, \setminus denotes the set difference and \cup the union.¹¹ It measures the distance between two ordinal preferences as the number of “swaps” necessary to transform one preference into another. For instance, if x is strictly better than y for R_i , x is indifferent to y for R_j , and R_i and R_j are identical on all other pairs, then (y, x) will belong to R_j but it will not belong to R_i . Hence, one swap is needed to transform R_i into R_j and the Kemeny distance between the two preferences is 1. Expression (2) above captures our main hypothesis, which can be summarized as follows:

Expressive utility is highest when a voter expresses her actual preference, and decreases as the expressed preference gets more distant from her actual preference.

Following this hypothesis, we can derive testable implications for expected turnout. For individual voters, the lower the (Kemeny) distance between actual and expressed preferences the higher the expected turnout, *ceteris paribus*. This is a straightforward implication of (1) and (2) (see also the proof of Proposition 1 below). To derive testable implications for expected turnout in different voting rules, we first define what we mean by a voting rule having more expressive power than another.

Definition 1 Voting rule A is said to have more expressive power than rule B if and only if $\mathcal{P}_A \supset \mathcal{P}_B$.

In other words, voting rule A is said to have more expressive power than voting rule B if and only if A allows to express a strict superset of preferences of the subset allowed to express by B. For instance, approval voting has more expressive power than the plurality rule because approval voting allows to express preferences that cannot be expressed with the plurality rule *and* all preferences that are possible to express with the plurality rule can be expressed with approval voting. We can then show the following:

Proposition 1 *If voting rule A has more expressive power than voting rule B, then expected turnout with voting rule A is equal or higher than expected turnout with voting rule B, ceteris paribus.*

¹¹ The Kemeny (1959) distance, also known as Kendall (1938) distance, swap distance, among other names, is one of the most used distance functions for comparing ordinal preferences, rankings, orders, lists, etc. We chose it for its widespread use, tractability, and simple interpretation. However, note that our conceptual framework and its implications would hold with other distance functions between ordinal preferences (see Can 2014 and Bhattacharya and Gravel 2021 for reviews of these functions).

Proof If voting rule A has more expressive power than voting rule B then $\mathcal{P}_A \supset \mathcal{P}_B$. To prove the implication, note that it follows from (1) and (2) that $u_i(R_i^e)$ decreases with $K(R_i^e, R_i^a)$, *ceteris paribus*. By extension, $P_i^1(u_i)$, the probability that individual i casts a vote, also decreases with $K(R_i^e, R_i^a)$, *ceteris paribus*. We consider two cases: (i) There is no $i \in N$ such that $K(R_i^e, R_i^a)$ with voting rule A is lower than $K(R_i^e, R_i^a)$ with voting rule B . This is the case, for instance, if $R_i^a \in \mathcal{P}_B$ for all $i \in N$. In this case, there is no $i \in N$ such that the influence of $K(R_i^e, R_i^a)$ on $P_i^1(u_i)$ is different with voting rule A and voting rule B . Therefore, expected turnout with voting rule A is equal to expected turnout with voting rule B , *ceteris paribus*. (ii) There is $i \in N$ such that $K(R_i^e, R_i^a)$ with voting rule A is lower than $K(R_i^e, R_i^a)$ with voting rule B . This is the case, for instance, if there is at least one individual in N such that $R_i^a \in \mathcal{P}_A \setminus \mathcal{P}_B$. In this case, there is $i \in N$ such that $P_i^1(u_i)$ is higher with voting rule A than with voting rule B due to the influence of $K(R_i^e, R_i^a)$. Therefore, the expected turnout with voting rule A is higher than the expected turnout with voting rule B , *ceteris paribus*. Since $\mathcal{P}_A \supset \mathcal{P}_B$, this completes the proof. \square

Proposition 1 shows that if at least one individual can express her actual preference with voting rule A but not with voting rule B (or can decrease the distance to her actual preference changing from B to A), then the average expected turnout with voting rule A is higher than with voting rule B , *ceteris paribus*. This means that voting rules may have different impacts on expected turnout according to their expressive power.

Before proceeding, it is worth emphasising the *ceteris paribus* nature of this proposition. It only says that if a voting rule allows voters to better express their preferences than another, then we should expect turnout to be higher with the former rule than with the latter if all remains equal. However, different voting rules can change the likelihood that voters will be pivotal. Therefore, if voting rule A has more expressive power than voting rule B but the likelihood of being pivotal in A is lower than the one in B , a decisive prediction in terms of overall expected turnout cannot be made. While in theory this is not an issue for large elections and opinion polls (as is the case of our experimental design), it is a relevant consideration for elections in small groups. A similar issue arises with respect to the strategic incentives of voters. Since these can be confounding factors in our experimental setting, we explore these issues below. More generally, this means that more research is needed to understand how voting rules affect the overall utility derived from voting and turnout. The contribution of this paper is to formally identify one channel linking voting rules with voters' utility and turnout, and test this channel experimentally.

Finally, it is also worth noting that in some cases our model allows us to make unambiguous comparisons in terms of expressive power while in other cases it does not. For instance, we can unambiguously say that approval voting has more expressive power than plurality rule. At the same time, we cannot make an unambiguous comparison between the plurality and Borda rules. This means that our criterion does *not* allow us to unambiguously compare all voting rules in terms of expressive power. This issue can be addressed for collective decisions for which the empirical

distribution of voters' actual preferences is known. While this is usually not known or available, especially *prior* to the vote, it is possible to approximate this distribution based on polls or data from previous collective decisions in the same setting. In that case, one can juxtapose the observed empirical distribution of actual preferences with the preferences that alternative voting rules allow to express, and determine which voting rule allows voters to better express themselves (i.e., which voting rule leads to lower Kemeny distance between the observed distribution of actual preferences and the preferences that the voting rule allows to express). The voting rule that would allow voters to better express themselves in this sense would be the rule with higher expressive power. This alternative empirical criterion of expressive power would allow us to unambiguously compare all voting rules.

3 Experimental design

In this section, we present our experimental design following the order of the experiment: (i) the elicitation of actual preferences, (ii) the assignment to treatments and elicitation of expressed preferences, (iii) the turnout decision, and (iv) the additional questions at the end of the experiment. We end the section explaining the procedures we used to implement the experiment.

3.1 Elicitation of actual preferences

The experiment was presented to subjects as a survey on people's opinions about how the UK government should "open up" the country after the first Covid-19 lockdown. The choice of a survey on this salient topic was meant to raise participants' interest to turn out (i.e., to register their vote; see below) with low instrumental motives to do so and negligible probabilities of being pivotal. This partially shuts down these potential confounding factors, which we return to in our discussion of the results (Sect. 6).

The first stage of the experiment was meant to elicit subjects' actual preferences over potential policies to "open up" the country that were envisioned by policy-makers during the first Covid-19 lockdown. Subjects were asked to rank the following three policies in order of their preference¹²:

- (i) **Policy A:** Strict social distancing for everyone.
- (ii) **Policy B:** Strict social distancing *only* for vulnerable people.
- (iii) **Policy C:** Self-imposed social distancing.

Subjects were requested to write 1 for their preferred policy, 2 for their second-favourite, and 3 for their least favourite, and if they were indifferent between two

¹² These policies were described in more detail to subjects (see Appendix F for exact wording). Subjects were randomly assigned to two orders, one as presented and one in which the order of policies was reversed (i.e. Policy A is Policy C here, and vice versa). For each treatment, half of the subjects participated in each order. Therefore, this manipulation is inconsequential to our main results.

or more policies they could give them the same number (i.e., subjects could report all possible preferences among the 3 policies). Importantly, they were informed that this ranking would not count for the voting results. We therefore infer subjects' actual preferences from this ranking.

3.2 Treatments and elicitation of expressed preferences

In the second stage of the experiment, we randomly assigned subjects to different treatments and elicited their expressed preferences under different voting rules. Treatment 1 (T1), treatment 2 (T2) and treatment 3 (T3) are equivalent in all respects besides the voting rule that subjects use to vote:

- T1.** Subjects vote with the plurality rule. Subjects were told that they could give 1 point to their favourite policy and 0 points to the other policies *or* that they could give 0 points to all policies.
- T2.** Subjects vote with approval voting. Subjects were told that they could give 1 point to their two favourite policies and 0 points to their least favourite policy, *or* 1 point to their favourite policy and 0 points to the other policies, *or* 0 points to all policies.
- T3.** Subjects vote with the A-Borda rule, a synthetic rule that allowed participants to express all possible preferences with approval voting and the Borda rule. Subjects were told that they could give 2 points to their favourite policy, 1 point to their second favourite policy, and 0 points to their least preferred policy, *or* 1 point to their two favourite policies and 0 points to their least preferred policy, *or* 1 point to their favourite policy and 0 points to the other policies, *or* 0 points to all policies.

Three features of these treatments are worth mentioning. First, note that expressive power is strictly increasing from T1 to T3. This means that we can make unambiguous comparative predictions in terms of turnout. Second, note that the voting rule in T3 is a synthetic institution that, to the best of our knowledge, does not exist outside our experiment. The introduction of this voting rule follows the tradition of using experiments “to construct and implement institutions explicitly designed to test theory and to understand behaviour” (Kagel and Levin 1993, pp. 868–869; these authors implemented a synthetic institution to study the descriptive validity of Nash bidding theory). In our setting, the A-Borda rule allows us to make unambiguous comparisons in terms of expressive power and expected turnout that would not be possible with the Borda rule. In addition, although this rule has not yet been used to settle collective decisions outside our experiment, it is similar to other proposals to modify the Borda rule to allow for partial rankings (see Emerson 2013). Third, we note that the computer program enforced that subjects used the voting rules correctly to avoid mistakes in their use.

Finally, we introduce treatment 4 (T4) and treatment 5 (T5) which make salient to subjects the fact that there are “several common voting rules to gather people’s opinions through their vote”. The two treatments are as follows:

- T4.** Subjects vote with the plurality rule (i.e., equivalent to T1). The only difference to T1 is a screen, presented before the voting decision, in which we describe the plurality rule, approval voting, and A-Borda rule and inform subjects that they will vote using the plurality rule. See Appendix F for the exact wording.
- T5.** Subjects vote with the A-Borda rule (i.e., equivalent to T3). The only difference to T3 is a screen, analogous to the one used in T4 before the voting decision, in which we describe the plurality rule, approval voting, and A-Borda rule and inform subjects that they will vote using the A-Borda rule. See Appendix F for the exact wording.

These treatments allow us to study if awareness of different voting rules (and the contrast between those) affects the rate of turnout in our experiment. Since plurality and A-Borda rules are the voting rules with lower and higher expressive power in our experiment, the two treatments are enough to test this effect. The ability to test this effect with only two treatments is the main reason for not having run a similar treatment with approval voting.

3.3 Turnout decision

The third stage of the experiment elicits our main variable of interest: our proxy for turnout. After voting as explained in the previous section, subjects were asked if they would like “to register their vote”. Subjects were reminded that to register their vote they would “be asked some extra questions about pandemics that should take 3–5 min to answer before [they] continue the survey”, which would not be asked in case of not registering their vote.¹³ They were also reminded that if they did not register their vote, their vote would not be counted when the results of the vote would be sent to them and others through the Prolific website, newspapers, and Twitter.

This incentive mechanism makes registering one’s vote costly. To register their vote, subjects needed to perform a real effort task for no extra pay. The objective is to mimic the costs associated with turnout. As a result, subjects willing to make

¹³ Subjects that registered their vote responded to a series of 25 multiple choice questions on general issues about the Covid-19 pandemic, such as “To this date, how many people over 60 have died of Covid-19 in the UK?” or “Of the three policies we presented to you, which policy do you think would lead to higher unemployment?”. Note that these questions are asked *after* subjects’ main decisions, so they have no impact on our main results.

their vote count had to exert effort. This decision is therefore used as our proxy for turnout.¹⁴

It is worth noting that we reversed the natural order of turnout and voting in our experiment. In our experiment, subjects first cast their vote between options, and only then decide whether or not to turn out. This contrasts with most “real-life” voting situations where the order is the other way around. The main underlying reason for this design choice is that subjects are likely to be unfamiliar with some of our voting rules. Our method allows subjects to experience the voting rule and submit their vote *before* deciding whether to register their vote or not (i.e., turnout). We thereby minimize the number of subjects for whom we would not have observed an effect of expressive power on turnout *because* of a lack of familiarity with the rule. In other words, this design choice allows us to be more confident that subjects properly received our intervention (voting with a randomly assigned voting rule) before their turnout decision (our dependent variable).

3.4 Additional questions

At the end of the experiment, we asked subjects several questions about their perception of the voting rules as well as social demographics. In terms of subjects’ perception of the voting rules, we asked about the perceived expressiveness and perceived complexity of the voting rules they used to vote:

- **Perceived expressiveness:** “From 1 (absolutely) to 10 (not at all), do you feel that the way you were allowed to give points to the policies let you express your preference about the 3 policies?”
- **Perceived complexity:** “From 1 (very easy) to 10 (very complex), how difficult it was for you to understand how to attribute points to the policies?”

We also presented subjects with the three voting rules we used in our experiment and asked them their *ex-post* subjective ranking of the voting rules:

- **Preferred voting rule:** “Which of these three rules [plurality, approval, and A-Borda rule, described in the same screen] would you have most liked to have used to vote for the 3 policies?”

In terms of social demographics, we asked about gender, age, education, health status, activity status, and several questions about past exposure to Covid-19. Finally, we also included an attention question among this final set of questions and we asked subjects how difficult they found it to answer the survey.

¹⁴ For a comprehensive analysis of participants’ willingness to exert effort to turnout (or their willingness to pay money to turnout), one could use an incentive-compatible mechanism *à la* Becker et al. (1964). However, this option would have increased the complexity and duration of our online experiment and was not necessary to test our main predictions. The choice of effort over money also increases external validity, since turnout is usually associated with physical effort.

3.5 Procedures

A total of 2041 subjects were recruited online via Prolific on the 4th of May 2020. All subjects were UK nationals, currently living in the UK, with a minimum “approval rate” of 95% on Prolific.¹⁵ The experiment was run online and participation was fully anonymous.

We paid subjects 1 pound for their participation. In line with our incentive mechanism, payment was independent of the duration subjects took to complete the experiment (we present durations per treatment and per turnout decision below). Subjects were sent the main results of the poll via the Prolific message board, where participants can be contacted keeping their anonymity. We will divulge the results of the experiment via newspapers and Twitter once the paper is published.

4 Predictions

Our conceptual framework allows us to make the following prediction in terms of average expected turnout in T1, T2, and T3:

Prediction 1 Average turnout in T3 > Average turnout in T2 > Average turnout in T1.

Although our model is silent in terms of the effect of reference points and salience, it also allows us to compare the average expected turnout in T4 and T5:

Prediction 2 Average turnout in T5 > Average turnout in T4.

Evidence from psychology and behavioural economics on the effects of reference points and salience suggests that the difference in turnout between T4 and T5 should be greater than the difference between T1 and T3. Under T4 and T5, subjects are not only informed about the rule they will use to vote but also about other rules that could have been used for the vote. Subjects in these treatments can therefore contrast different voting rules. In particular, subjects can use the rules that were presented but with which they did not vote as reference points to judge the voting rule with which they voted. The influence of explicit contrasts of this sort on preference elicitation has been documented in various domains. Kahneman (1992) has seminaly shown that individuals do not perceive their utility in an absolute way but in terms of changes with respect to reference points. Other researchers have also underlined the sequential process that occurs when attitudes or choices are elicited under such contrast or saliency manipulations (e.g. Shavitt and Fazio 1991; Dhar et al. 1999). According to these authors, subjects first form a similarity judgement and then express a preference, and the first comparative process (similarity) can affect the

¹⁵ Prolific provides the option to “reject” subjects and not pay for their participation (e.g. because of *speeding*). The approval rate is the percentage of studies for which a participant has been previously approved by the person/people conducting the study.

second one (preference). To the best of our knowledge, there is no direct evidence about salience or reference points in the context of alternative voting methods. However, these indirect pieces of evidence suggest that T4 and T5 may make the relatively low (high) expressive power of the plurality rule (A-Borda rule) more salient via the comparison to the other rules. We are therefore more likely to find results consistent with Prediction 2 than with Prediction 1.

5 Results

The following analysis is based on 1977 subjects that completed the experiment (393 in T1, 398 in T2, 400 in T3, 392 in T4, and 394 in T5).¹⁶ We asked one attention question and controlled for the duration that each subject took to complete the survey. In Appendix A, we present the results for the restricted sample that did not fail the attention question and responded in a duration no lower than one-half of the treatment mean duration per turnout decision. In a nutshell, our main results are the same with the two samples.

In the final part of the experiment, we collected some social demographic measures. Table 1 summarizes subjects' characteristics. Two features of this data are worth mentioning. First, the data shows no systematic differences across treatments. Empirical tests reported in Appendix B show that, with few exceptions, the differences in observed covariates across treatments are small and the sample is mostly balanced on these variables. Still, the observed differences across treatments and the potential effect of these particular covariates on our main variable of interest (turnout) justifies a robustness check, carried out below, in which we check for the treatment effect on turnout controlling for these covariates. Second, it is worth noting that education is relatively high in our sample, with a mean of 5.5 (between post-secondary non-tertiary education and Bachelor's or equivalent level) and all subjects (with one exception) had lower secondary education or higher. We also asked subjects how difficult it was to understand the survey. From 1 (very easy) to 10 (very difficult), the sample average was 2.2 (2.1 in T1, 2.1 in T2, 2.2 in T3, 2.1 in T4, 2.4 in T5). This supports premises concerning the understanding of the instructions.

Before proceeding to our main analysis, we report the main results of the poll. The grey bars of Fig. 1 show the preferred policy of subjects that decided to register their vote and reported a unique favourite policy when voting. The results show that while Policy A was strictly preferred by almost half of the subjects that registered their vote (44.3%), the sample was far from consensual in terms of what the government should have done to "open up" the country. Policy B was strictly preferred by 19.7% of the subjects that registered their vote and Policy C by 17.8%. The black tops of each bar show the percentage of subjects that "approved" the corresponding policy alongside another policy. Several subjects voted for two preferred policies (17.1% in the treatments that allowed to report indifference), with the figure showing

¹⁶ 53 subjects started but not completed the experiment (12 in T1, 8 in T2, 6 in T3, 14 in T4, and 13 in T5). Given that the treatments are overall very similar, we have no reason to believe that dropouts are systematically correlated with the treatments.

that subjects approved Policy B alongside another policy slightly more often than Policy A and Policy C (13.3% against 11.6% and 9.3% respectively). Finally, only 1.13% of subjects voted that they were indifferent between the three policies.

5.1 Actual and expressed preferences

We start our main analysis by looking at subjects' actual and expressed preferences, and the relationship between the two. We are particularly interested in ascertaining if there is a difference between actual and expressed preferences, and if this difference is statistically significantly different across treatments.

Before voting, we asked subjects to rank the 3 policies in order of preference, allowing for indifference between two or more policies. We infer subjects' actual preferences from this ranking. Table 2 reports the Kemeny distance between participants' actual and expressed preferences, where a Kemeny distance of 0 means that subjects' actual and expressed preferences were the same.¹⁷

The results show, first, that very few subjects vote with their actual preference in T1, T2, and T4 (see row “=0” in Table 2). This is not surprising, as the voting rules in these treatments restrict the preferences that can be expressed. Namely, these voting rules exclude the possibility to vote with a complete antisymmetric preference over the 3 policies, and 93% of subjects (1838 out of 1977) across all treatments reported a complete antisymmetric actual preference. We conjecture that this high number of antisymmetric preferences (or lack of indifference between policies) was due to the particular context in which our experiment was run—the heart of the Covid-19 crisis and the daily discussion of public health policies in the media—which is likely to have induced strong opinions about this topic. Second, these results show that a large majority of subjects in T3 and T5 voted with their actual preference. Still, a significant proportion of subjects voted with a preference different than their actual preference *even though* they could have voted with their actual preference. This difference can be explained by several factors, notably (i) mistakes (“noise” common in experiments) and/or (ii) strategic voting. We discuss these issues in Sect. 6. Third, these results show that most subjects that did not vote with their actual preference voted with a Kemeny distance of 1. This is expected, given that all voting rules allowed subjects to vote within a Kemeny distance of 0 (T3 and T5), 1 (T2), or 2 (T1 and T4) to all possible actual preferences.

We are particularly interested in differences across treatments. First, we find that subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2 ($p \leq 0.01$ in both cases, Pearson Chi-square tests).¹⁸ This is explained by the presence of subjects with actual preferences with two favourite policies, that were allowed to express them in T2 but not in T1/T4. According to our model, this suggests that some subjects derived more expressive utility in T2 than in T1 and T4. However, our results show that there are few subjects for whom

¹⁷ The Kemeny distance takes values between 0 and 6 for all subjects, with 93.7% of subjects between 0 and 2. The results are for the entire sample. We find similar results if we restrict to those who decided to register their vote.

¹⁸ We use Table 2 for all Pearson Chi-square tests on Kemeny distance, unless otherwise stated.

Table 1 Subjects' characteristics

	T1	T2	T3	T4	T5	Total
Age	34.6	33.5	34.5	35.1	34.1	34.4
Female (fraction)	62.3	64.3	65.5	67.4	64.2	64.7
Education (1 lower, 8 upper)	5.4	5.5	5.5	5.5	5.5	5.5
Health status (1 excellent, 5 poor)	2.3	2.3	2.4	2.4	2.3	2.3
Covid exposure (fraction)	77.1	69.6	72.8	74.5	71.6	73.1
Activity status (fraction)						
Student	17.3	20.6	15.3	16.8	18.5	17.7
Employed	67.2	63.6	68.3	62.8	60.2	64.4
Unemployed	13.7	15.6	14.3	14.8	14.2	14.5
Inactive	5.3	5.5	6.5	9.2	11.9	7.7
Retired	4.6	2.5	4.0	2.8	4.3	3.6

Mean values for *Age*, *Education*, and *Health status*. One subject can have more than one *Activity status*. *Covid exposure* shows the fraction of subjects that self-reported to be themselves (or an immediate family member or close friend) at higher risk of Covid-19 complications, to have been hospitalized because of Covid-19, to have been on self-isolation at home with Covid-19, to have financial difficulties because of Covid-19, or to have psychological difficulties because of Covid-19 (otherwise subjects reported that none of the above applied or preferred not to answer)

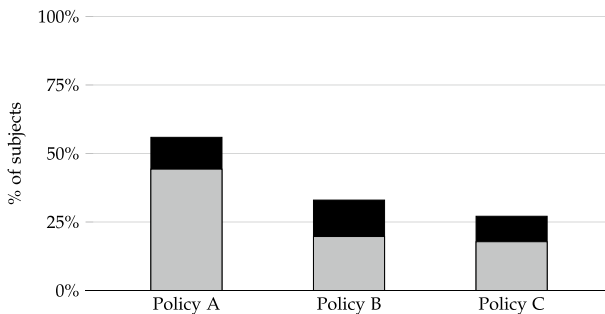


Fig. 1 Preferred policy. Percentages are based on the subjects that decided to register their vote. Grey bars are percentages of subjects that voted for a unique favourite policy. The black tops of each bar are percentages of subjects that “approved” the corresponding policy alongside another one when voting. 1.13% of subjects “approved” the three policies

this may be the case in our sample (around 5%; see Table 2). This will be important for the interpretation of our results below. Second, we find that subjects' actual and expressed preferences are statistically significantly more distant in T1/T2/T4 than in T3/T5 ($p < 0.001$ in all cases, Pearson Chi-square tests). In this case, the number of subjects for whom the Kemeny distance is 0 in T3 and T5 is much larger than the corresponding number of subjects in T1, T2 and T4 (see Table 2). This suggests that a large number of subjects derived more expressive utility in T3/T5 than in T1/T2/T4. Third, we find that the difference between T1 and T4 and the difference

Table 2 % of subjects with Kemeny distance equal to 0, 1, and equal or higher than 2

	T1 (%)	T2 (%)	T3 (%)	T4 (%)	T5 (%)	Total (%)
= 0	2.3	7.5	72.0	3.3	67.5	30.7
= 1	84.5	84.7	19.3	85.2	22.1	59.0
≥ 2	13.2	7.8	8.8	11.5	10.4	10.3

between T3 and T5 are not statistically significant. These results can be summarised as follows:

Result 1 Subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2, and in T2 than in T3/T5.

In Appendix C, we report the results of an Ordered logit regression on Kemeny distance, with dummies for each treatment and controls for gender, age, education, health status, activity status, and past exposure to Covid-19. In a nutshell, Result 1 is robust when we account for individual observed characteristics: Subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2 ($p < 0.01$ for both tests, two-tailed Wald tests), and statistically significantly more distant in T2 than in T3/T5 ($p < 0.001$ for both tests, two-tailed Wald tests).

According to our conceptual framework, Result 1 shows that our experimental manipulation led to significant differences in terms of the expressive utility that voters derived from voting with different voting rules, and that such differences are in the intended direction: Subjects derived more expressive utility in T3/T5 than in T2, and derived more expressive utility in T2 than in T1/T4.

5.2 Turnout

Of the 1977 subjects that participated in the experiment, 80% decided to register their vote (1589 subjects). Subjects that decided to register their vote spent on average 12.8 min to complete the survey, while those who abstained spent on average 8.8 min to complete the survey.¹⁹ This means that turning out entailed 4 additional minutes to complete the survey on average (an increase of 45% in duration for the same pay). This shows that 80% (20%) of subjects were willing (not willing) to exert considerable effort to turnout.

Figure 2 summarises our main turnout results. Looking first at the treatments without reference points (T1, T2, and T3), Fig. 2 shows that turnout was 80.15% in T1, 82.92% in T2, and 79.0% in T3. While the difference between T1 and T2 aligns with our main predictions, the differences between T1/T2 and T3 do not. However, the differences between these treatments are not statistically significant (Pearson Chi-square tests). These findings can be summarised as follows:

¹⁹ Average duration in minutes per treatment for those who turned out (abstained) was 12.5 (9.3) in T1, 12.8 (8.4) in T2, 13.1 (8.3) in T3, 12.8 (9.1) in T4, and 13.0 (9.1) in T5.

Result 2 Turnout is not statistically significantly different between T1, T2, and T3.

Result 2 contrasts with our predictions. In the next section, we discuss potential underlying reasons for this result. Once reference points are introduced (T4 and T5), however, results do align with our predictions. While turnout is 76.79% in T4, it is 82.99% in T5. The difference between T4 and T5 is statistically significant at 5% ($p = 0.03$, Pearson Chi-square test). A 6.2 percentage point increase in turnout seems also “economically” meaningful given the high average turnout observed across treatments. Note as well that turnout is lower in T4 than in T1 and higher in T5 than in T3. Although these differences are not statistically significant (Pearson Chi-square tests), these results are qualitatively aligned with the relevance of reference points and salience. These findings can be summarised as follows:

Result 3 Turnout is statistically significantly higher in T5 than in T4.

These results are robust when we account for possible differences in subjects’ observed characteristics across treatments. In Appendix D, we present the results of a Probit regression on turnout, with dummies for each treatment and controls for gender, age, education, health status, activity status, and past exposure to Covid-19. In a nutshell, we find no statistically significant difference for the effect of T1, T2, and T3 on turnout (two-tailed Wald tests), while the effect of T5 on turnout is statistically significantly higher than the effect of T4 when we control for individual characteristics at around 1% ($p = 0.016$, two-tailed Wald test). These results are also robust when we restrict the sample to the 1838 subjects that reported a complete antisymmetric actual preference, who can benefit from the greater expressive power of A-Borda rule (T3/T5) in relation to approval voting and plurality rule (T1/T2/T3). Taken together, these robustness results bring further support to Results 2 and 3, which suggest that differences in expressive power can have a significant impact on turnout as long as voters are aware of (and can compare) different voting rules.

5.3 Turnout and the distance between actual and expressed preferences

If the underlying mechanism of our conceptual framework is correct (our main hypothesis captured in Expression 2), we should observe, on average, subjects to be more likely to turn out if their expressed preferences are closer to their actual preferences. To test this, we look at the relationship between turnout and Kemeny distance. We find that subjects are more likely to turn out if they express their actual preference (Kemeny distance = 0) and less likely to turn out if they do *not* express their actual preference (Kemeny distance ≥ 1), and this effect is statistically significant at 2% (Pearson Chi-square test; $p = 0.002$ if we use Kemeny distances = 0, = 1, and ≥ 2). When looking at the treatments separately, the results are qualitatively aligned with the positive effect of lower Kemeny distance on turnout, although the effect is only statistically significant for T4 and T5 (T4: $p = 0.044$ if Kemeny distance = 0 and ≥ 1 , Pearson Chi-square test; $p = 0.019$ if Kemeny distance = 0, = 1, and ≥ 2)

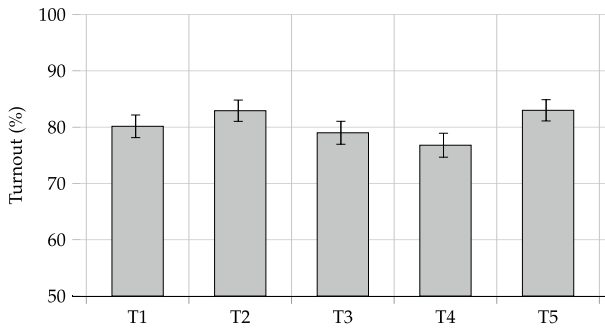


Fig. 2 Subjects' turnout in percentage. The main bars display brute percentages of turnout per treatment. The error bars display one standard error of the mean, with values based on the asymptotic standard error of the maximum likelihood estimate of the proportion who voted

(T5: $p = 0.009$ if Kemeny distance = 0, = 1, and ≥ 2 , Pearson Chi-square test). This means that, in line with our model, lower Kemeny distance is correlated with higher turnout, and this effect is particularly prominent in T4 and T5 as predicted.

However, note that this effect can be driven by non-sincere voters if these voters (for whom the Kemeny distance is higher than 0) are less likely to turn out. This is actually the case: subjects who could have voted with their actual preference but have *not* done so (around 30% in T3 and T5) were less likely to turn out than those who did vote with their actual preference ($p = 0.041$ for T3 and T5, Pearson Chi-square test).²⁰

To further probe the relationship between turnout and the Kemeny distance, we test it in a restricted sample for which we exclude all subjects who did not vote with their actual preference but could have done so. This addresses the concern raised in the previous paragraph. We also exclude all remaining subjects in T3 and T5 because—contrary to T1, T2, and T3—all subjects that remain after implementing the previous restriction have voted with their actual preference (i.e., in T3 and T5 we cannot compare voters who voted with their actual preference against voters who did not vote with their actual preference because they could not have done so). We end up with a sample of 1177 subjects (392 in T1, 394 in T2, and 391 in T4). In a nutshell, in this restricted sample we find analogous results in terms of both sign and statistical significance to the previous results on the relationship between Kemeny distance and turnout. Taken together, these findings demonstrate that lower Kemeny distance is correlated with higher turnout, bringing further support to the main hypothesis of our conceptual framework (Expression 2).

²⁰ We focus on T3 and T5 because in the other three treatments only six subjects did not vote with their actual preference when they could have done so. We discuss the relationship between non-sincere voting and our main results for turnout in Sect. 6.

5.4 Expressiveness, rule complexity, and preferred voting rule

At the end of the experiment, we asked participants several questions concerning the voting rules (see Sect. 3.3.4 for details). One of these is a direct question about the perceived “expressiveness” of the voting rule that the subject has used. We expected perceived expressiveness to increase in line with expressive power, i.e., expressiveness in T5/T3 to be higher than in T2 and expressiveness in T2 to be higher than in T1/T4. We also expected the effect of reference points to be noted here, such that the highest value would be observed in T5 and the lowest value in T4. Another relevant question is about the perceived complexity of the rule that the subject has used. We expected perceived complexity to be higher in T3/T5, followed by T2, followed by T1/T4. The results for these two questions are summarized in Table 3.

These results suggest that, in line with our predictions, T3 and T5 are considered the most expressive rules and T1 and T4 the least expressive. These differences are statistically significant ($p \leq 0.001$ for all tests, based on two-tailed Student tests and Wilcoxon signed-rank tests). T2 is also considered more expressive than T1/T4 as predicted ($p = 0.003$ for T2 and T1, $p \leq 0.06$ for T2 and T4, based on two-tailed Student tests and Wilcoxon signed-rank tests), and less expressive than T3/T5, although the difference between T2 and T3 is very small and statistically insignificant. Once reference points are introduced, the difference between the perceived expressiveness of plurality and A-Borda rules increases. This supports the hypotheses that reference points are important for perceived expressiveness (which can explain why the difference between T2 and T3 is very small) and that the observed difference in turnout between T4 and T5 is due to the difference in the expressive power of the two rules. Note as well that according to this measure, the effect of reference points is more relevant in T5 than in T4, since perceived expressiveness is statistically significantly higher in T5 than in T3 ($p \leq 0.02$, based on a two-tailed Student test and Wilcoxon signed-rank test), but the difference between T1 and T4 is not statistically significant. In Appendix E, we test if perceived expressiveness has a significant effect on turnout using Probit regressions with individual controls. Our results show that perceived expressiveness has a positive and significant effect on turnout ($p = 0.02$). Taken together, these results support the underlying mechanisms of our model, according to which higher expressive power (via higher expressive utility) leads to higher turnout.

In terms of perceived complexity, the results suggest that T3 and T5 are considered more difficult to understand than T1, T2 and T4. This accords with our predictions. Parametric and nonparametric tests suggest that the difference between T1/T4 and T3/T5 and the difference between T2 and T5 are statistically significant.²¹ The complexity of the voting rule can, in principle, impact turnout in the following way: The more complex a rule is, the lower the expected turnout is, all else equal. To probe this effect, we test if perceived complexity has a significant effect on turnout

²¹ Based on two-tailed Student tests and Wilcoxon signed-rank tests, T1 and T3 $p = 0.07$ and $p = 0.02$ respectively, T1 and T5 $p = 0.02$ and $p < 0.01$, T2 and T5 $p < 0.01$ and $p < 0.05$, and T4 and T3/T5 $p < 0.01$ for all tests. All other differences are not statistically significant, except for the two-tailed Student test for T2 and T3 ($p = 0.03$).

using Probit regressions with individual controls (see Appendix E). Our results show that, as expected, perceived complexity has a negative and significant effect on turnout ($p = 0.04$).²² Note, importantly, that this effect runs in the opposite direction to the effect of the expressive power on turnout. This can partly explain why turnout is not higher in T3 than in T1.

Finally, at the end of the experiment, we presented subjects with the three voting rules we used in our experiment and asked them which voting rule they would have most liked to have used to vote for the 3 policies. Table 4 summarizes the results. A large majority of subjects (68.39%) would like to have used the A-Borda rule. This suggests that the A-Borda rule is preferred over plurality and approval voting to express opinions over policies.

6 Discussion

In this section, we provide an interpretation for our main findings and discuss some potential confounding factors. We start with the interpretation of Result 1: Subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2, and in T2 than in T3/T5. This result shows that differences in expressive power can have a significant impact in terms of the representation of voters' actual preferences. According to our conceptual framework, it follows that differences in expressive power can have a significant impact in terms of the expressive utility that voters derive from voting. Our results in terms of the perceived expressiveness of the rules (Table 3), voters' *ex-post* preferences over voting rules (Table 4), and the relationship between turnout and Kemeny distance (Sect. 5.5.3), all support this interpretation. If this holds more generally, the expressive power of voting rules has important implications in terms of expressive utility and the representation of voters' actual preferences, and therefore the welfare of voters.

Turning to Result 2, the finding that turnout is not statistically significantly different between T1 and T2 can be explained by the low percentage of subjects who were able to better express themselves in T2 than in T1 (around 5%, see Table 2). Such a low percentage significantly reduces the impact of the treatment, and more research is needed to compare the expressive power of plurality and approval voting. In terms of T1/T2 and T3, the results reported above show that perceived complexity and non-sincere voting have a negative and statistically significant effect on turnout. Importantly, these effects run in the opposite direction to the effect of the expressive power on turnout. Therefore, this can partly explain the null result. In addition, there may be a bias in favour of more familiar (and/or even "real") rules. In particular, we conjectured that without being able to compare the voting rule being used with other rules (absence of explicit reference points), the ability/inability to better express oneself would not be salient and its effect on the expressive utility of voting muted.

²² Once we introduced perceived expressiveness and perceived complexity into the Probit regression, only perceived expressiveness has a significant effect ($p = 0.07$). This is not surprising, since perceived expressiveness and perceived complexity are significantly correlated (0.25, $p < 0.001$, pairwise correlation test).

Table 3 Perceived expressiveness and perceived complexity (averages per treatment)

	T1	T2	T3	T4	T5	Total
Perceived expressiveness (1 high, 10 low)	4.3	3.8	3.7	4.1	3.3	3.9
Perceived complexity (1 easy, 10 difficult)	2.5	2.5	2.8	2.4	2.9	2.6

The anticipation of these potential effects motivated the inclusion of T4 and T5 in our experiment. The statistically significant difference in turnout between T4 and T5 (Result 3) brings support to the relevance of reference points and salience in our setting. According to this interpretation, subjects can use the rules that were presented but with which they did not vote as reference points to judge the voting rule with which they voted. This may create an emotional reaction of elation (frustration/disappointment) with the ability (inability) to express oneself. For example, a voter may feel frustrated for not being able to use a rule with which she would better express her opinion. It follows, according to this interpretation, that the ability/inability to express oneself becomes salient. In other words, T4 and T5 make the relatively low (high) expressive power of the plurality rule (A-Borda rule) more salient. This then magnifies the effect of the expressive power of voting rules on voters' turnout likelihood.

We now turn to potential confounding factors that could, at least in principle, change across treatments and undermine the interpretation of our results: (i) non-sincere voting and (ii) instrumental sincere voting. In terms of non-sincere voting, there seem to be more possibilities for non-sincere (such as strategic voting) in T3/T5 than in T1/T4. This issue may be relevant since our results for T3 and T5—treatments in which voters could always vote with their actual preferences—suggest that around 30% of subjects did not vote sincerely (see Table 2).²³ While we cannot ascertain if the discrepancy between actual and expressed preferences is due to mistakes (“noise” common in experiments), strategic voting, or other reasons for non-sincere voting, two pieces of evidence suggest that non-sincere voting is not driving our results. First, non-sincere voting could not explain the effect of reference points observed in T4 and T5 with respect to T1 and T3. Second, it is not clear from a theoretical point of view if “non-sincere voters” are more or less likely to turn out than “sincere voters”, and our results suggest that they are less likely to do so than sincere voters. As reported above, subjects in T3 and T5 who could have voted with their actual preference but have *not* done so were less likely to turn out than those who did vote with their actual preference ($p = 0.041$, Pearson Chi-square test). Similarly, the 9% of subjects in T1 and T4 that did not vote for their actual preferred option were less likely to turn out than those that did ($p = 0.004$, Pearson Chi-square test). This evidence suggests that non-sincere voters (strategic *and* others) are less likely

²³ In T1, T2, and T4 only six subjects did not vote with their actual preference when they could have done so.

Table 4 Preferred voting rule

% of subjects who would like to have used	Plurality rule	Approval voting	A-Borda rule
	19.3%	12.3%	68.4%

to turn out than sincere voters *and* are more prevalent in T3 and T5 than in T1 and T4. It follows that this effect, if present, runs in the opposite direction to the effect of the expressive power on turnout. Then, this can partly explain why turnout is not higher in T3 than in T1, but cannot drive the significant difference observed between T4 and T5.

In terms of instrumental (sincere) voting, this effect could also change across treatments. In particular, it is possible that some subjects may have voted fully (or partially) due to instrumental motives about influencing the result of the vote and/or impacting policy-making; if these voters had the *illusion* that their vote was pivotal *and* gave instrumental weight to voting with a certain preference than another (e.g. gave instrumental weight to voting for two favourite policies instead of voting for a single preferred policy, because their least preferred alternative, according to this illusion, would be less likely chosen), then this could in principle impact turnout differently across treatments. To probe the possibility of instrumental voting, after the turnout decision, we asked subjects (in an open-end format) the reasons why they decided to register their vote/not register their vote. Therefore, we can test if subjects stated expressive or instrumental motives behind their turnout decision. To check for this, we asked a research assistant unaware of our research objectives to code the answers as follows:

- **“Instrumental”** whenever a subject refers that he/she voted to *influence* the result of the vote and/or impact policy-making (e.g. impact what the government will do). Examples include: “I think it’s good that my opinion might be able to have some impact, no matter how small”, “I think the more information you have regarding the population choices the better for the government to make an informed decision”, and “It might give me a voice in influencing a government policy which I think is in the interests of the country”. In case of abstention, whenever a subject refers they cannot influence the result of the vote and/or the vote cannot impact policy-making.
- **“Expressive”** whenever a subject refers to reasons related to one’s voice and/or expressing opinions and/or the value of democracy and/or wanting one’s vote to count. Examples include “I think it is important to share opinions”, “It is important to have my say”, “Because it is important we voice our opinion”, “Because I think people should be listened to”. In the case of abstention, whenever a subject refers there is no point in expressing one’s opinion and/or no one listens or other ideas that negate the expressive value of voting.
- **“Other”** whenever a response does not fit the two previous reasons. This includes general interest/curiosity, wanting to assist research, convenience, issues

of privacy or lack of knowledge, or whenever the reason is not clear or well-connected to the two previous reasons.

The results show that *among those who decided to register their vote*, 36% of subjects gave expressive reasons for their decision, 10% gave instrumental reasons, and 54% gave other reasons. This suggests that few voters had instrumental reasons as a primary motive to register their vote. In addition, subjects' responses suggest that differences in the frequency of instrumental reasons across treatments are small and not systematic (8.3% in T1, 11.2% in T2, 11.7% in T3, 10.63% in T4 and 8.9% in T5).²⁴ Taken together, these pieces of evidence suggest that different strengths of instrumental motives across treatments are also not driving our results.

7 Concluding remarks

In this paper, we proposed a simple ordinal model of expressive voting that allows us to compare voting rules in terms of the expressive utility that voters can derive from voting (the “expressive power” of voting rules). Our model provides a novel testable implication according to which expected turnout increases with expressive power, all else equal. We ran an online experiment to test this implication in a controlled environment and found that if voters are made aware of different voting rules, turnout is higher in voting rules with higher expressive power. This result suggests that voting rules with higher expressive power may increase participation whenever voters are aware of different voting rules (i.e., when there are explicit reference points). In other settings, such as committees or elections, one can easily raise awareness of alternative voting rules and provide reference points. For example, if a national voting reform increases expressive power, the government could link the reform with an informational campaign that would provide information about the old and the new rules and, in a more direct way than done in our setting, highlight the expressive benefit of the latter. In addition, our results show that higher expressive power is associated with a better representation of voters' actual preferences and, according to our model, higher expressive utility.

Our analysis suggests that the expressive power of voting rules is a relevant criterion when deciding which voting rule to use in economic and political decisions. It shows how it is aligned with goals for voting and turnout behaviours which are generally supported in the literature. For example, there is a wide acceptance that higher turnout is a desirable goal for modern liberal democracies (e.g. Lijphart 1997; Engelen 2007; cf. Saunders 2012). Similarly, the better representation of voters' actual preferences is often supported in the literature. For instance, Horowitz (2003) (p. 115) argues that “the best electoral system is the one that straightforwardly and most accurately reflects the preferences of voters”. Overall, this suggests

²⁴ Expressive reasons are more prominent in T3 and T5 (32.7% in T1, 32.1% in T2, 38.6% in T3, 35.2% in T4 and 38.8% in T5). However, the difference in the distribution of answers is only statistically significantly different between T1 and T3 at 10%.

that the expressive power of voting rules is a criterion with descriptive and normative appealing features that can *complement* existing criteria for the evaluation of voting rules.

Finally, our analysis points towards relevant future empirical research. For instance, are the results we find via our online experiment with an opinion poll generalised to political elections? While a similar experiment may be difficult (if not impossible) to run in political elections, it may be possible to gather indirect data as well as laboratory experimental data. While both these methods have relevant caveats in this setting that made us choose our experimental design, they would complement our analysis. It would be also interesting to compare electoral systems (the wider set of rules for elections that include but are not limited to voting rules) in terms of the expressive utility that voters derive from them. These lines of investigation are left for future work.

Appendix

A. Main results with restricted sample

In this appendix, we present our main results with a restricted sample of subjects that passed our attention checks. We impose two restrictions. First, we exclude all subjects that responded in a duration lower than one-half of the treatment mean duration per turnout decision. Second, we also exclude all subjects that responded incorrectly to the attention question. We are left with a sample of 1017.²⁵

Our main results can be summarised as follows. In this sample, the Kemeny distance between subjects' actual and expressed preferences is very similar to the whole sample for all treatments. Table 5 summarises the data. As in Result 1 above, in this sample subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2, and in T2 than in T3/T5 ($p \leq 0.01$ in all cases except for T1 and T2 for which $p = 0.015$, Pearson Chi-square tests).²⁶

In terms of turnout behaviour, 83% of subjects decided to register their vote in this sample (as opposed to 80% in the whole sample). Figure 3 shows the differences

²⁵ We exclude a total of 960 subjects. A total of 915 subjects failed the attention question, 90 subjects responded in a duration lower than one-half of the treatment mean duration per turnout decision, and 45 of these failed the attention question. It is worth noting that only 53.7% of subjects in our sample responded correctly to our attention question. This is a low percentage. One explanation for this may be the fact that the attention question was particularly tricky to spot. For one, the question was asked at the end of the experiment among a final set of questions that are easy to respond to. In addition, the attention question itself may have prompted a quick answer: it read as follows: "From 1 (very easy) to 5 (very difficult), how difficult it was to understand the policies?", with responses "1 (very easy)", "2", "3", "4", "If you are not a robot please click on this button". Since the policies were easy to understand, subjects could promptly answer "1 (very easy)" or "2" without reading the remaining part of the question. The data suggests this was the case, with 90% of subjects that failed the attention question answering "1 (very easy)" or "2" to this question. Overall, while this shows some lack of attention and led us to present our analysis again dropping the subjects that failed this attention check, it should be seen as a strict test of attention.

²⁶ We use Table 5 for all Pearson Chi-square tests on Kemeny distance.

across treatments. Turnout was 85.34% in T1, 87.5% in T2, and 84.21% in T3. The differences across these treatments are qualitatively similar to the whole sample, and, as before, these differences are not statistically significant (Pearson Chi-square tests). Concerning T4 and T5, turnout was 74.37% and 84.54% respectively. The difference between T4 and T5 is statistically significant at around 1% ($p = 0.011$, Pearson Chi-square test). In this sample, the difference between T4 and T5 seems to be explained by the lower turnout in T4 when compared to T1 ($p < 0.01$, Pearson Chi-square test).

As before, subjects are more likely to turnout if they express their actual preference (Kemeny distance = 0) and less likely to turnout if they do *not* express their actual preference (Kemeny distance ≥ 1) ($p = 0.019$ if Kemeny distance = 0 and ≥ 1 , Pearson Chi-square test; $p = 0.025$ if Kemeny distance = 0, = 1, and ≥ 2). When looking at the treatments separately, the results are in general qualitatively aligned with the positive effect of lower Kemeny distance on turnout, although the difference is only statistically significant for T3 ($p = 0.031$ if Kemeny distance = 0 and ≥ 1 , Pearson Chi-square test). For the restricted sample of T1, T2 and T4 for which we exclude all subjects who did not vote with their actual preference but could have done so (similar restriction as in the main text; $N = 579$ here), we find again a statistically significant difference across treatments ($p = 0.085$ if Kemeny distance = 0 and ≥ 1 , Pearson Chi-square test; $p = 0.063$ if Kemeny distance = 0, = 1, and ≥ 2) and, for this smaller sample, qualitatively aligned non-significant differences for each treatment.

The remaining results are also similar between the two samples: The results for the restricted sample on perceived expressiveness, perceived complexity, and preferred voting rule, are very similar to those observed in the whole sample. Overall, this robustness check brings support to the validity of our inferences.

B. Differences in subjects' characteristics across treatments

In this appendix, we present empirical tests on the differences in subjects' characteristics across treatments. For continuous variables, we follow Imbens and Wooldridge (2009) (p. 24) and report the normalised differences in means (see also Deaton and Cartwright 2018, p. 6). The results are shown in Table 6. Imbens and Rubin suggest that, as a rule of thumb, normalised differences should not exceed one quarter (see Imbens and Wooldridge 2009, p. 24). As shown in the table, this is not the case for any of our observed covariates that are continuous variables. For categorical variables, we report Chi-square homogeneity tests as an indication of distance (see Table 7). With few exceptions (e.g. Covid exposure between T1 and T2 and Active status between T1 and T5), these tests support the overall balance of observed covariates in our sample.

Table 5 % of subjects with Kemeny distance equal to 0, 1, and equal or higher than 2 (restricted sample)

	T1 (%)	T2 (%)	T3 (%)	T4 (%)	T5 (%)	Total (%)
= 0	2.6	6.8	70.6	2.5	63.3	31.0
= 1	81.7	86.5	21.5	83.9	24.6	57.9
≥ 2	15.7	6.8	7.9	13.6	12.1	11.1

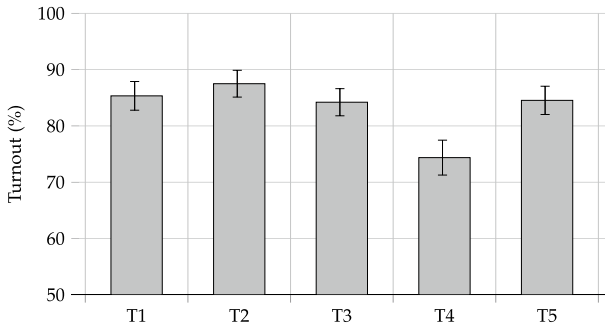


Fig. 3 Subjects' turnout in percentage (restricted sample). The main bars display brute percentages of turnout per treatment. The error bars display one standard error of the mean, with values based on the asymptotic standard error of the maximum likelihood estimate of the proportion who voted

C. Ordered logit regression analysis: Result 1

In this appendix, we provide a robustness check to Result 1 of Sect. 5.1. Table 8 reports the average values of Kemeny distances per treatment. The table shows a similar picture as above: Kemeny distance is higher in T1 and T4 and lower in T3 and T5. In accordance with Result 1, the results reported in Tables 9 and 10 show that when using an ordered logit regression and controlling for subjects' observed characteristics, subjects' actual and expressed preferences are statistically significantly more distant in T1/T4 than in T2 ($p < 0.01$ for both tests, two-tailed Wald tests), and statistically significantly more distant in T2 than in T3/T5 ($p < 0.001$ for both tests, two-tailed Wald tests). The results reported in Table 9 also show that observed individual characteristics have no significant effect on Kemeny distance.

D. Probit regression analysis: Results 2 and 3

The results reported in Tables 11 and 12 show that our main results hold when we control for gender, age, education, health status, activity status, and past exposure to Covid-19. That is, we find no statistically significant difference for the effect of T1, T2, and T3 on turnout (two-tailed Wald tests), while the effect of T5 on turnout is statistically significantly higher than the effect of T4 when we control for individual observed characteristics (two-tailed Wald test; see Table 12). This brings further confidence to our Results 2 and 3, especially to the difference between T4 and T5,

Table 6 Normalised differences in gender, education, and health

	Gender	Education	Health
Δ_{T1-T2}	0.06	-0.01	0.04
Δ_{T1-T3}	0.009	-0.03	-0.06
Δ_{T1-T4}	-0.02	-0.05	-0.04
Δ_{T1-T5}	0.03	-0.04	-0.03
Δ_{T2-T3}	-0.06	-0.01	-0.1
Δ_{T2-T4}	-0.09	-0.03	-0.08
Δ_{T2-T5}	-0.03	-0.03	-0.07
Δ_{T3-T4}	-0.03	-0.02	0.02
Δ_{T3-T5}	0.02	-0.01	0.03
Δ_{T4-T5}	0.06	0.007	0.007

This table displays normalised differences for Gender, Education, and Health as shown in Table 1, where $\Delta_{a-b} = \frac{\bar{X}_a - \bar{X}_b}{\sqrt{S_a^2 + S_b^2}}$ with \bar{X}_i representing the mean and S_i^2 the variance of $i = a, b$

Table 7 Chi-square homogeneity tests for female, Covid exposure, and active status

	Female	Covid exposure	Active status
χ^2_{T1-T2}	0.56	0.02	0.93
χ^2_{T1-T3}	0.38	0.16	0.62
χ^2_{T1-T4}	0.14	0.39	0.19
χ^2_{T1-T5}	0.59	0.08	0.08
χ^2_{T2-T3}	0.76	0.33	0.69
χ^2_{T2-T4}	0.37	0.13	0.22
χ^2_{T2-T5}	0.98	0.54	0.09
χ^2_{T3-T4}	0.55	0.58	0.4
χ^2_{T3-T5}	0.74	0.71	0.19
χ^2_{T4-T5}	0.35	0.36	0.65

This table displays Chi-square homogeneity tests for female and covid exposure as described in Table 1, and *Active status* which takes the value of 1 if the subject is active (student or employed) and 0 otherwise

Table 8 Kemeny distance (averages per treatment)

	T1	T2	T3	T4	T5	Total
Kemeny distance (0 lowest, 6 highest)	1.3	1.1	0.4	1.2	0.5	0.9

which is statistically significant at around 1% when we control for observed covariates ($p = 0.016$).

The results of the Probit regression reported in Table 11 also show that older subjects are (slightly) more likely to turn out in our setting, that female subjects

Table 9 Impact of treatments and subjects' characteristics on Kemeny distance

	(1)	(2)
T1	3.02*** (0.19)	3.05*** (0.19)
T2	2.57*** (0.18)	2.59*** (0.18)
T3	-0.25 (0.18)	-0.23 (0.18)
T4	2.90*** (0.19)	2.91*** (0.19)
T5	(omitted)	(omitted)
Age		-0.001 (0.004)
Female (fraction)		0.11 (0.10)
Education		0.04 (0.04)
Health status		-0.08 (0.05)
Covid exposure		-0.01 (0.11)
Active status		-0.15 (0.12)
N	1977	1977

*Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on Ordered logit regressions with clustered standard errors per subject. Independent variables as described in Table 1, except for *Active status* which takes the value of 1 if the subject is active (student or employed) and 0 otherwise

Table 10 Between treatment differences

	Wald tests
$H_0: T1 = T2$	< 0.001
$H_0: T1 = T3$	< 0.001
$H_0: T1 = T4$	0.231
$H_0: T1 = T5$	< 0.001
$H_0: T2 = T3$	< 0.001
$H_0: T2 = T4$	< 0.01
$H_0: T2 = T5$	< 0.001
$H_0: T3 = T4$	< 0.001
$H_0: T3 = T5$	0.191
$H_0: T4 = T5$	< 0.001

This table displays *p*-values for two-tailed Wald tests from the Ordered logit regression reported in column (2) of Table 9

are more likely to turnout than male subjects, that turnout increases with education and with Covid exposure, but that turnout is insensitive to subjects' active or self-reported health status in our sample. The evidence on age and education agrees with the patterns observed in political elections in the UK that show that turnout is higher among older people and more educated people (in particular within age groups) (see e.g. Democratic Audit 2014 and British Election Study 2021). The higher turnout of women in our setting may be explained by other factors than political engagement, since participation in UK elections is roughly the same for men and women (see e.g. Rosalind 2017). The evidence of a positive effect of Covid exposure on turnout accords with our expectations (people affected are more concerned), and the

Table 11 Impact of treatments and subjects' characteristics on turnout

	(1)	(2)
T1	0.85*** (0.07)	-0.64** (0.26)
T2	0.95*** (0.07)	-0.51** (0.26)
T3	0.81*** (0.07)	-0.68*** (0.26)
T4	0.73*** (0.07)	-0.77*** (0.26)
T5	0.95*** (0.08)	-0.52** (0.26)
Age		0.01*** (0.003)
Female (fraction)		0.16** (0.07)
Education		0.13*** (0.03)
Health status		0.004 (0.04)
Covid exposure		0.27*** (0.07)
Active status		-0.08 (0.09)
N	1977	1977

*Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on Probit regressions with clustered standard errors per subject. Independent variables as described in Table 1, except for *Active status* which takes the value of 1 if the subject is active (student or employed) and 0 otherwise

Table 12 Between treatment differences

	Wald tests
$H_0: T1 = T2$	0.214
$H_0: T1 = T3$	0.692
$H_0: T1 = T4$	0.188
$H_0: T1 = T5$	0.275
$H_0: T2 = T3$	0.101
$H_0: T2 = T4$	0.011
$H_0: T2 = T5$	0.884
$H_0: T3 = T4$	0.355
$H_0: T3 = T5$	0.136
$H_0: T4 = T5$	0.016

This table displays p -values for two-tailed Wald tests from the Probit regression reported in column (2) of Table 11

absence of evidence on active or health status may be explained by the fact that their effects are partly captured by age and Covid exposure respectively.

E. Perceived expressiveness, perceived complexity and turnout

In this appendix, we test for the effect of perceived expressiveness and perceived complexity on turnout. Results reported in Table 13 are as expected: perceived expressiveness has a positive and significant effect on turnout (the sign is negative because perceived expressiveness is highest when equal to 1 and lowest when equal

Table 13 Impact of perceived expressiveness and perceived complexity on turnout

	(1)	(2)	(3)
Perceived expressiveness	-0.03** (0.01)		-0.02* (0.01)
Perceived complexity		-0.03** (0.02)	-0.03 (0.02)
Age	0.01*** (0.003)	0.01*** (0.003)	0.01*** (0.003)
Female (fraction)	0.10 (0.06)	0.11* (0.63)	0.11* (0.06)
Education	0.09*** (0.03)	0.08*** (0.02)	0.09*** (0.03)
Health status	-0.03 (0.04)	-0.02 (0.04)	-0.02 (0.04)
Covid exposure	0.24*** (0.07)	0.24*** (0.07)	0.24*** (0.07)
Active status	-0.14* (0.08)	-0.14* (0.08)	-0.14* (0.08)
N	1977	1977	1977

*Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on Probit regressions with clustered standard errors per subject. Independent variables as described in Table 1, except for *Active status* which takes the value of 1 if the subject is active (student or employed) and 0 otherwise

to 10), while perceived complexity has a negative effect. When including both perceived expressiveness and perceived complexity [specification (3) in the table], only perceived expressiveness has a significant effect ($p = 0.07$). This is not surprising, since perceived expressiveness and perceived complexity are significantly correlated (0.25, $p < 0.001$, pairwise correlation test).

Results within treatments are qualitatively aligned with these results, but the impact of perceived expressiveness on turnout is only significant in T1 at 10% and the impact of perceived complexity is only significant in T1 at 5% with these smaller samples. Since perceived expressiveness and perceived complexity are correlated with the treatments, their effect is insignificant once we introduce dummies for treatments.

F. Instructions

In the following, we present the instructions for the main parts of the experiment (elicitation of actual preferences, treatments/elicitation of expressed preferences, and turnout decision).²⁷ All differences between treatments are presented below. Comments that are not part of the instructions are shown in square brackets.

[Introductory screen]

In the next screens, you are asked to answer an **anonymous survey**. This survey is organized by researchers at the University of Paris 2 and the University of Southampton.

In the main part of this survey, we will ask **your opinion about how the UK government should “open up” the country after the current lockdown**. You will have the option to express or to not express your opinion with a vote.

²⁷ The remaining part of the instructions consists of questions presented above and are available at <https://reshare.ukdataservice.ac.uk/856725/>.

If you want to register your vote, you will be asked some extra questions about pandemics that should take 3–5 min to answer. Otherwise, you will NOT be asked these extra questions but your vote will not be counted.

The results of the vote will be sent to you and others through the Prolific website, newspapers, and Twitter, keeping the anonymity of all responders. Only the votes of people that decided to register their vote will be counted.

Most people will take around 8 min to complete this questionnaire, not counting the extra 3 to 5 min in case you decide to register your vote. We ask you to please reflect upon the questions before answering.

Thank you for your participation!

[Elicitation of actual preferences]

As you are aware, the Coronavirus (COVID-19) crisis led the UK government to put the country under lockdown. Under current circumstances, it seems that the lockdown will be extended for more weeks, possibly months. Suppose the lockdown in the UK, as currently held, ends on the 15th of June 2020 with COVID-19 new infections under control and with a low-medium risk of a second peak of infections and deaths.

The UK government then faces a difficult decision of how to ease lockdown after the 15th of June 2020. We would like to know your preferences in terms of possible policies to “open up” the country.

Consider the three following policies:

- **Policy A (strict social distancing for everyone):** For 6 months after lockdown, limit public transportation, prohibit gatherings with more than 5 people, limit the number of people in shops/pubs/restaurants/etc., require everyone to use masks in public spaces, control borders, and impose strict social distancing between students and teachers in schools. Besides this, people and businesses can return to normal.
- **Policy B (strict social distancing *only* for vulnerable people):** For 6 months after lockdown, require vulnerable people (people over 65 and people with underlying medical conditions) to stay at home except to buy essential goods and exercise or meet a single designated person, require vulnerable people to use masks in public spaces, and encourage non-vulnerable people to limit physical contacts and exert full prophylactic hygiene (e.g. wash hands frequently, wear masks, etc.). Besides this, people and businesses can return to normal.
- **Policy C (self-imposed social distancing):** For 6 months after lockdown, encourage everyone to limit physical contacts and exert full prophylactic hygiene (e.g. wash hands frequently, wear masks, etc.). Besides this, people and businesses can return to normal.

Please rank the 3 policies in order of your preference. Write 1 for your preferred policy, 2 for your second-favourite, and 3 for your least favourite.

If you are indifferent between two or more policies, please give them the same number. For example, write 1 for two policies and 2 for one policy.

This rank will not count for the voting results.

- [] Policy A
- [] Policy B
- [] Policy C

[Treatments and elicitation of expressed preferences]

[Treatment 1]

You can now vote for the policies. You will be asked later if you want to register or not register your vote.

To vote for the policies, you have the following options:

- **You can give 1 point to your favourite policy and 0 points to the other policies, or**
- **You can give 0 points to all policies.**

	1 point	0 points
Policy A	<input type="radio"/>	<input type="radio"/>
Policy B	<input type="radio"/>	<input type="radio"/>
Policy C	<input type="radio"/>	<input type="radio"/>

[Treatment 2]

You can now vote for the policies. You will be asked later if you want to register or not register your vote.

To vote for the policies, you have the following options:

- **You can give 1 point to your two favourite policies and 0 points to your least favourite policy, or**
- **You can give 1 point to your favourite policy and 0 points to the other policies, or**
- **You can give 0 points to all policies.**

	1 point	0 points
Policy A	<input type="radio"/>	<input type="radio"/>
Policy B	<input type="radio"/>	<input type="radio"/>
Policy C	<input type="radio"/>	<input type="radio"/>

[Treatment 3]

You can now vote for the policies. You will be asked later if you want to register or not register your vote.

To vote for the policies, you have the following options:

- **You can give 2 points to your favourite policy, 1 point to your second favourite policy, and 0 points to your least favourite policy, or**
- **You can give 1 point to your two favourite policies and 0 points to your least favourite policy, or**
- **You can give 1 point to your favourite policy and 0 points to the other policies, or**
- **You can give 0 points to all policies.**

	2 points	1 point	0 points
Policy A	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Policy B	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Policy C	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Treatment 4]**[First screen]**

Before you vote, please note that there are several common voting rules to gather people's opinions through their vote. Here are some common rules that could be used to vote for the 3 policies:

- **Rule #1:** Voters can give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.
- **Rule #2:** Voters can give 1 point to their two favourite policies and 0 points to their least favourite policy. Voters can also give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.
- **Rule #3:** Voters can give 2 points to their favourite policy, 1 point to their second favourite policy, and 0 points to their least favourite policy. Voters can also give 1 point to their two favourite policies and 0 points to their least favourite policy. Voters can also give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.

In this study, **you are only allowed to use Rule #1.**

[Second screen]

You can now vote for the policies. You will be asked later if you want to register or not register your vote.

To vote for the policies, you have the following options:

- **You can give 1 point to your favourite policy and 0 points to the other policies, or**
- **You can give 0 points to all policies.**

	1 point	0 points
Policy A	<input type="radio"/>	<input type="radio"/>
Policy B	<input type="radio"/>	<input type="radio"/>
Policy C	<input type="radio"/>	<input type="radio"/>

[Treatment 5]

[First screen]

Before you vote, please note that there are several common voting rules to gather people’s opinions through their vote. Here are some common rules that could be used to vote for the 3 policies:

- **Rule #1:** Voters can give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.
- **Rule #2:** Voters can give 1 point to their two favourite policies and 0 points to their least favourite policy. Voters can also give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.
- **Rule #3:** Voters can give 2 points to their favourite policy, 1 point to their second favourite policy, and 0 points to their least favourite policy. Voters can also give 1 point to their two favourite policies and 0 points to their least favourite policy. Voters can also give 1 point to their favourite policy and 0 points to the other policies. Voters can also give 0 points to all policies.

In this study, **you are allowed to use Rule #3.**

[Second screen]

You can now vote for the policies. You will be asked later if you want to register or not register your vote.

To vote for the policies, you have the following options:

- **You can give 2 points to your favourite policy, 1 point to your second favourite policy, and 0 points to your least favourite policy, or**
- **You can give 1 point to your two favourite policies and 0 points to your least favourite policy, or**
- **You can give 1 point to your favourite policy and 0 points to the other policies, or**
- **You can give 0 points to all policies.**

	2 points	1 point	0 points
Policy A	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Policy B	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Policy C	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[The following text, which recalls the policies subjects could vote for, appeared at the bottom of the screen where subjects voted in all treatments]

Recall the possible policies:

Policy A (strict social distancing for everyone): For 6 months after lockdown, limit public transportation, prohibit gatherings with more than 5 people, limit the number of people in shops/pubs/restaurants/etc., require everyone to use masks in public spaces, control borders, and impose strict social distancing between students and teachers in schools. Besides this, people and businesses can return to normal.

Policy B (strict social distancing *only* for vulnerable people): For 6 months after lockdown, require vulnerable people (people over 65 and people with underlying medical conditions) to stay at home except to buy essential goods and exercise or meet a single designated person, require vulnerable people to use masks in public spaces, and encourage non-vulnerable people to limit physical contacts and exert full prophylactic hygiene (e.g. wash hands frequently, wear masks, etc.). Besides this, people and businesses can return to normal.

Policy C (self-imposed social distancing): For 6 months after lockdown, encourage everyone to limit physical contacts and exert full prophylactic hygiene (e.g. wash hands frequently, wear masks, etc.). Besides this, people and businesses can return to normal.

[Turnout decision]

Please choose one of the two actions:

1. **I want to register my vote.** In this case, you will be asked some extra questions about pandemics that should take 3 to 5 min to answer before you continue the survey.
2. **I don't want to register my vote.** In this case, you will NOT be asked these extra questions but your vote will not be counted.

The results of the vote will be sent to you and others through the Prolific website, newspapers, and Twitter, counting only the votes of people that decided to register their vote.

- I want to register my vote
- I don't want to register my vote

Acknowledgements We are grateful to the Associate Editor in Charge and two anonymous referees for their valuable comments and suggestions that greatly improved the paper. We are also grateful to Thomas Gall, Georg Granic, Nicolas Gravel, Max Kwiek, Jean Mercier Ythier, Benoît Tarroux, Michael Vlasopoulos, and the participants at the ESA 2021 Global Online Conference and the RES 2023 Annual Conference for valuable discussions, written comments and suggestions. We thank Thibault Vercoûtère for his excellent research assistance. Finally, the financial support of the research project “ValFree” (*The Value of Choice*, grant No. ANR-16-CE41-0002-01) and the research project “FrontCog” (grant No. ANR-17-EURE-0017) of the French National Agency for Research (ANR) are gratefully acknowledged. The experiment is approved by the Institutional Review Board of Sacha Bourgeois-Gironde’s institution (IRB No. 20152900001072, Laboratoire Institut Jean-Nicod, CNRS UMR 8129).

Data availability The dataset generated by the experiment and analyzed during the current study is publicly available in the UK Data Service’s ReShare repository at <https://reshare.ukdataservice.ac.uk/856725/>, <https://doi.org/10.5255/UKDA-SN-856725>.

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alós-Ferrer C, Granic GD (2012) Two field experiments on approval voting in Germany. *Soc Choice Welf* 39:171
- Aragones E, Gilboa I, Weiss A (2011) Making statements and approval voting. *Theor Decis* 71:461–72
- Baujard A, Gavrel F, Igersheim H, Laslier J-F, Lebon I (2013) Approval voting, evaluation voting: an experiment during the 2012 french presidential election. *Revue Économique* 64(2):345–56
- Baujard A, Igersheim H (2009) Experiment of evaluation voting and approval voting on april 22, 2007. first results. *Revue Économique* 60:189–201
- Baujard A, Igersheim H, Lebon I (2021) Some regrettable grading scale effects under different versions of evaluative voting. *Soc Choice Welf* 56:803–34
- Becker GM, DeGroot MH, Marschak J (1964) Measuring utility by a single-response sequential method. *Behav Sci* 9:226–32
- Bhattacharya M, Gravel N (2021) Is the preference of the majority representative? *Math Soc Sci* (in press)
- Blais A (2006) What affects voter turnout? *Annu Rev Polit Sci* 9:111–25
- Blais A, Aarts K (2006) Electoral systems and turnout. *Acta Polit* 41(2):180–96
- Blais A, Laslier J-F, Poinas F, Van der Straeten K (2015) Citizens' preferences about voting rules: self-interest, ideology, and sincerity. *Public Choice* 164:423–42
- Blais A, Plescia C, Högström J, Péloquin-Skulski G (2021) Do (many) voters like ranking? *Party Polit* 27(6):1223–1228
- Bol D, Blais A, Choice J-FLP (2018) A mixed- utility theory of vote choice regret. *Public Choice* 176:461–78
- Bol D, Blais A, Coulombe M, Laslier J-F, Pilet JB (2023) Choosing an electoral rule: values and self-interest in the lab. *J Econ Psychol* 95:102602
- Bosch A, Orriols L (2014) Ballot structure and satisfaction with democracy. *J Elect Public Opin Parties* 24(4):493–511
- Bourgeois-Gironde S, Ferreira JV (2020) Conflicted voters: a spatial voting model with multiple party identifications. *J Econ Behav Organ* 174:360–79
- Bouton L, Castanheira M, Llorente-Saguer A (2016) Divided majority and information aggregation: theory and experiment. *J Public Econ* 134:114–28
- Bouton L, Gallego J, Llorente-Saguer A, Morton R (2021) Run-off elections in the laboratory. *Econ J* (in press)
- Brams S, Fishburn P (2002) Voting procedures. In: Sen KAAK, Suzumura K (eds) *Handbook of social choice and welfare*. Elsevier Science, Amsterdam, pp 173–236
- Brams SJ, Fishburn PC (2005) Going from theory to practice: the mixed success of approval voting. *Soc Choice Welf* 25:457–74
- Brams SJ, Fishburn PC (2007) *Approval Voting*. Springer, New York, second edition
- Brennan G, Buchanan J (1984) Voter choice: evaluating political alternatives. *Am Behav Sci* 28(2):185–201
- Brennan G, Hamlin A (1998) Expressive voting and electoral equilibrium. *Public Choice* 95:149–75

- Brennan G, Lomasky LE (1993) *Democracy and decision: the pure theory of electoral preference*. Cambridge University Press, New York
- British Election Study (2021) Age and voting behaviour at the 2019 general election. <https://www.britishelectionstudy.com/bes-findings/age-and-voting-behaviour-at-the-2019-general-election/#.YVxhMS-ZMII>
- Can B (2014) Weighted distances between preferences. *J Math Econ* 51:109–15
- Carter JR, Guerette SD (1992) An experimental study of expressive voting. *Public Choice* 73(3):251–60
- Chamberlin JR (1985) An investigation into the relative manipulability of four voting systems. *Behav Sci* 30(4):195–203
- Conitzer V, Sandholm T (2005) Communication complexity of common voting rules. In: *Proceedings of the 6th ACM conference on electronic commerce*, pp 78–87
- Deaton A, Cartwright N (2018) Understanding and misunderstanding randomized controlled trials. *Soc. Sci. Med.* 210:2–21
- Democratic Audit (2014) Highly educated young people are less likely to vote than older people with much lower levels of attainment. <https://www.democraticaudit.com/2014/03/13/in-britain-people-with-higher-educational-attainment-are-less-likely-to-vote-because-they-are-younger/>
- Dhar R, Nowlis SM, Sherman SJ (1999) Comparison effects on preference construction. *J Consum Res* 26(3):293–306
- Emerson P (2013) The original borda count and partial voting. *Soc Choice Welf* 40:353–58
- Engelen B (2007) Why compulsory voting can enhance democracy. *Acta Polit* 42(1):23–39
- Engelman D, Gruener HP, Hoffmann T, Possajennikov A (2020) Minority protection in voting mechanisms: experimental evidence. CEPR Discussion Paper No. DP14393
- Feddersen T, Gailmard S, Sandroni A (2009) Moral bias in large elections: theory and experimental evidence. *Am Polit Sci Rev* 103(2):175–92
- Fiorina MP (1976) The voting decision: instrumental and expressive aspects. *J Polit* 38(2):390–413
- Forsythe R, Rietz T, Myerson R, Weber R (1996) An experimental study of voting rules and polls in three-candidate elections. *Int J Game Theory* 25:355–83
- Ginzburg B, Guerra JA, Lekuangfu WN (2022) Counting on my vote not counting: expressive voting in committees. *J Public Econ* 205:1–16
- Green DP, Shapiro I (1994) *Pathologies of rational choice theory. A critique of applications in political science*. Yale University Press, New Haven
- Hamlin A, Jennings C (2019) Expressive voting. In: In RD, Congleton BG, Voigt S (eds) *The Oxford handbook of public choice*, vol 1. Oxford University Press, New York, pp 333–350
- Hoffmann T, Renes S (2022) Flip a coin or vote? an experiment on the implementation and efficiency of social choice mechanisms. *Exp Econ* 25(1):624–55
- Horowitz DL (2003) Electoral systems: a primer for decision makers. *J Democr* 14(4):115–127
- Igersheim H, Baujard A, Gavrel F, Laslier J-F, Lebon I (2016) Individual behavior under evaluative voting: a comparison between laboratory and in situ experiments. In: Blais A, Laslier J-F, Van der Straeten K (eds) *Voting experiments*. Springer, Cham, pp 257–69
- Igersheim H, Baujard A, Laslier J-F (2016b) La question du vote. expérimentations en laboratoire et in situ. *L'Actualité économique* 92(1):151–189
- Imbens GW, Wooldridge JM (2009) Recent developments in the econometrics of program evaluation. *J Econ Liter* 47(1):5–86
- Kagel JH, Levin D (1993) Independent private value auctions: bidder behavior in first, second and third price auctions with varying numbers of bidders. *Econ J* 103:868–79
- Kahneman D (1992) Reference points, anchors, norms, and mixed feelings. *Organ Behav Hum Decis Process* 51(2):296–312
- Kartal M (2015) A comparative welfare analysis of electoral systems with endogenous turnout. *Econ J* 125(587):1369–92
- Kemeny J (1959) Mathematics without numbers. *Daedalus* 88(4):577–91
- Kendall M (1938) A new measure of rank correlation. *Biometrika* 30(1/2):81–93
- Kube S, Puppe C (2009) (when and how) do voters try to manipulate? Experimental evidence from borda elections. *Public Choice* 139:39–52
- Ladner A, Milner H (1999) Do voters turn out more under proportional than majoritarian systems? The evidence from swiss communal elections. *Elect Stud* 18(2):235–50
- Laslier J-F (2009) The leader rule: a model of strategic approval voting in a large electorate. *J Theor Polit* 21(1):113–36
- Lijphart A (1997) Unequal participation: democracy's unresolved dilemma. *Am Polit Sci Rev* 91(1):1–14

- Merrill S (1984) A comparison of efficiency of multicandidate electoral systems. *Am J Polit Sci* 28(1):23–48
- Moreno B, del Pino Ramos-Sosa M, Rodriguez-Lara I (2019) Conformity and truthful voting under different voting rules. *Soc Choice Welf* 53:261–82
- Müller S, Jankowski M (2019) Do voters really prefer more choice? Determinants of support for personalised electoral systems. *J Elect Public Opin Parties* 29(2):262–81
- Nurmi H (1987) Comparing voting systems. D. Reidel Publishing Company, Dordrecht, Holland
- Piketty T (2000) Voting as communicating. *Rev Econ Stud* 67(1):169–91
- Plassmann F, Tideman TN (2014) How frequently do different voting rules encounter voting paradoxes in three-candidate elections? *Soc Choice Welf* 42:31–75
- Pons V, Tricaud C (2018) Expressive voting and its cost: evidence from runoffs with two or three candidates. *Econometrica* 86(5):1621–49
- Riker WH, Ordeshook PC (1968) A theory of the calculus of voting. *Am Polit Sci Rev* 62:25–42
- Rosalind S (2017) In what ways does gender matter for voting behaviour in ge2017? <http://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/>
- Sanz C (2017) The effect of electoral systems on voter turnout: evidence from a natural experiment. *Polit Sci Res Methods* 5(4):689–710
- Saunders B (2012) The democratic turnout problem. *Polit Stud* 60:306–320
- Sen A (1995) How to judge voting schemes. *J Econ Perspect* 9(1):91–8
- Shavitt S, Fazio RH (1991) Effects of attribute salience on the consistency between attitudes and behavior predictions. *Pers Soc Psychol Bull* 17(5):507–16
- Shayo M, Harel A (2012) Non-consequentialist voting. *J Econ Behav Organ* 81(1):299–313
- Spenkuch JL (2018) Expressive vs. strategic voters: an empirical assessment. *J Public Econ* 165:73–81
- Tsakas N, Xeferis D (2021) Stress-testing the runoff rule in the laboratory. *Games Econom Behav* 128:18–38
- Tullock G (1971) The charity of the uncharitable. *Econ Inq* 9(4):379–92
- Tyran J-R (2004) Voting when money and morals conflict: an experimental test of expressive voting. *J Public Econ* 88:1645–64
- Tyran J-R, Wagner AK (2019) Experimental evidence on expressive voting. In: In RD, Congleton BG, Voigt S (eds) *The Oxford handbook of public choice*, vol 2. Oxford University Press, New York, pp 928–40
- Van der Straeten K, Laslier J-F, Sauger N, Blais A (2010) Strategic, sincere, and heuristic voting under four election rules: an experimental study. *Soc Choice Welf* 35:435–72
- Van der Straeten K, Laslier J-F, Sauger N, Blais A (2010) Strategic, sincere, and heuristic voting under four election rules: an experimental study. *Soc Choice Welf* 35(3):435–72
- Weber M (2020) Choosing the rules: preferences over voting systems for assemblies of representatives. *J Econ Behav Org* 174:420–434

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.