Liquid Democracy. Two Experiments on Delegation in Voting*

Victoria Mooers[†], Joseph Campbell[‡], Alessandra Casella[§], Lucas de Lara[¶], and Dilip Ravindran[‖]

April 16, 2024

*With the exception of the first author, all others are in alphabetical order. We thank audiences at numerous academic seminars and conferences, and in particular Tim Feddersen, Chloe Tergiman, and Richard van Weelden for comments. We are grateful to Mael Lebreton, Jonathan Nicholas, Nahuel Salem, and Camilla van Geen for their help with the second experiment. We thank the Program for Economic Research at Columbia and the Columbia Experimental Lab for the Social Sciences for their financial support. We acknowledge computing resources from Columbia University's Shared Research Computing Facility project, which is supported by NIH Research Facility Improvement Grant 1G20RR030893-01 and associated funds from the New York State Empire State Development, Division of Science Technology and Innovation (NYSTAR) Contract C090171, both awarded April 15, 2010. The experiments have been approved by Columbia University's IRB.

[†]Columbia University, v.mooers@columbia.edu

[‡]Columbia University, joseph.campbell@columbia.edu

[§]Columbia University, NBER and CEPR, ac186@columbia.edu

[¶]Columbia University, lpd2122@columbia.edu

[‖]Humboldt University of Berlin, dilip.ravindran@hu-berlin.de

Abstract

Proponents of participatory democracy praise Liquid Democracy: decisions are taken by referendum, but voters delegate their votes freely. When better informed voters are present, delegation can increase the probability of a correct decision. However, delegation must be used sparely because it reduces the information aggregated through voting. In two different experiments, we find that delegation underperforms both universal majority voting and the simpler option of abstention. In a tightly controlled lab experiment where the subjects' precision of information is conveyed in precise mathematical terms and very salient, the result is due to overdelegation. In a perceptual task run online where the precision of information is not known precisely, delegation remains very high and again underperforms both majority voting and abstention. In addition, subjects substantially overestimate the precision of the better informed voters, underlining that Liquid Democracy is fragile to multiple sources of noise. The paper makes an innovative methodological contribution by combining two very different experimental procedures: the study of voting rules would benefit from complementing controlled experiments with known precision of information with tests under ambiguity, a realistic assumption in many voting situations.

JEL codes: **C92, D70, D72, D83**

Keywords: voting rules, majority voting, information aggregation, laboratory experiments, Condorcet Jury theorem, perceptual tasks, ambiguity.

1 Introduction

The widespread sense that party politics is in a crisis has been accompanied by diverse and increasingly loud proposals for new forms of participatory democracy. Among them are digital platforms to coordinate common actions, mini-publics and citizens assemblies, participatory budgeting, and algorithms for interactive democracy. Among the latter, Liquid Democracy has caught the imagination of the young and the tech-savvy. It advocates a voting system where all decisions are submitted to referendum, but voters can delegate their votes freely. It traces its intellectual roots to Charles Dodgson (1884) and James Miller (1969), and has been adopted occasionally for internal decisions by European protest parties–the Swedish and the German Pirate parties being the most famous examples. Now it finds vocal support in the tech community, where it aligns both with the emphasis on a non-hierarchical order and with the possible use of cryptographic tools to maintain confidentiality and reliability.[1] Although the details vary, the common point of different implementations is the ease and specificity of delegation: a vote can be delegated freely, and to a different person according to the issue. According to supporters, Liquid Democracy is superior to representative democracy because representatives can be chosen according to their specific competence on each decision, and is superior to direct democracy because uninformed or uninterested voters can delegate their votes.

The desirability of delegation in Liquid Democracy has been studied under two main headings. When the emphasis is on alignment of preferences, the literature has focused on delegation as a means of completing preference rankings: individuals unable to express a full ranking over all alternatives proposed delegate to others with whom they agree on the issues on which they know their own preferences.[2] More frequently, the focus has been on common interest environments with two alternatives and an unknown "ground truth" on which experts have superior information.[3] This is the perspective we take in this paper.

Delegating to better informed experts with preferences that match our own can seem a trivially beneficial step. But note a fundamental problem: even if the experts are correctly identified, delegation deprives the electorate of the richness of noisy but abundant information distributed among all voters. Condorcet's Jury Theorem teaches us the value of a large

[1] See for example LiquidFeedback (https://liquidfeedback.com/en/), the Association for Interactive Democracy (https://interaktive-demokratie.org/association.en.html), or Democracy.Earth (https://democracy.earth/). Google ran a 3-year experiment on its internal network, implementing Liquid Democracy for decisions like food menu choices, tee-shirt designs, or logos for charitable events (Hardt and Lopes, 2015). Liquid Democracy is becoming the governance choice for cryptoworld DAOs (Decentralized Autonomous Organizations)–see for example Element Finance (https://medium.com/element-finance).

[2] See for example, Christoph and Grossi (2017), Brill and Talmon (2018), or Harding (2022).

[3] For example, Kahng, Mackenzie and Procaccia (2018), Caragiannis and Michas (2019), Armstrong and Larson (2021), Halpern et al. (2021), Ravindran (2021), Dhillon et al. (2023).

electorate with barely accurate but independent sources of information. In a binary decision, if voters receive independent signals whose accuracy exceeds 50%, the percentage of correct signals exceeds 50% with probability approaching one as the size of the electorate becomes very large. If voters vote according to their signals, majority voting delivers the correct outcome with probability that approaches one (Condorcet, 1785). Thus, unless the extent of delegation is modulated correctly, a smaller number of voters, even if more accurate, may well lead to worse decision-making. This very basic trade-off is necessarily at the core of Liquid Democracy and is the focus of this paper.

We study a canonical common interest model where voters receive independent signals, conditional on an unknown binary state. The common objective is to identify the state correctly, aggregating information via majority voting. Signals vary across individuals in the probability of being correct (their *"precision"*). Experts are publicly identified and the precision of their signals is known; for all other voters, signals' precisions are private information but known to be weakly lower than the experts'. If a voter chooses to delegate, the vote is randomly assigned to one of the experts. We begin by showing theoretically that for any finite size of the group and any number of experts, there is an equilibrium with positive delegation such that the outcome is strictly superior to majority voting without delegation. The result, summarized formally in a theorem, is the starting point of the experiment. In such an equilibrium, unless the experts' information is extremely precise, delegation must not be too frequent, given its informational cost. Consider for example one of the parametrizations we bring to the lab: a group of 15 voters of which 3 are experts; the experts' information is correct with probability 70%; the precision of non-experts' signals can take any value between 50 and 70%, with equal probability. Then only non-experts with precision lower than 53%–i.e. with signals very close to random–should delegate. And mistakes are costly: small errors towards over-delegation can lead to severe expected losses.

Before describing the experiments in detail, note that the informational benefit from overweighing voters with more precise signals can be achieved via abstention as well, as long as abstention correlates with less precise signals. Abstention differs from delegation because it increases the voting weight of *all* individuals who choose to vote, not only those targeted as delegates. Yet, we know from McMurray (2013) that, under common interest and in the absence of voting costs, it too can lead to improvements over simple majority voting, and for reasons that parallel those favoring delegation. Abstention is a familiar option, well understood and well accepted. Its performance relative to delegation is thus an important reference point for Liquid Democracy, and our experiments compare the two alternatives.

The first experiment is designed for the lab and follows the theory very closely. We study groups of either 5 voters (of which 1 is an expert) or 15 voters (of which 3 are experts).

After receiving their signals, non-experts choose whether to cast their vote or to delegate. We observe the frequency of delegation and the fraction of group decisions that yield the correct outcome. We then compare these results to a second treatment, where the option of abstention takes the place of delegation. Finally, we evaluate both treatments relative to simple majority voting with voting by all. With the experimental parameters, theory predicts that the frequency of correct decisions under delegation and abstention should be closely comparable and in both cases superior to majority voting. Under Liquid Democracy (LD), we find systematic over-delegation: delegation rates that are between two and three times the rate in the unique strict equilibrium. Under Majority Voting with Abstention (MVA), abstention rates are instead close to the theory. As a result, the frequency of correct decisions under LD is lower than under MVA. LD underperforms also relative to simple majority voting, while MVA obtains very similar efficiency levels.

Why such high rates of delegation? And why the difference with respect to abstention? In the second part of the paper, we investigate whether the results could be due to the experimental design, canonical as it is. We test the robustness of Experiment 1's findings in a very different environment: a perceptual task where individuals do not have precise information about their and others' perceptual accuracy, or, in the language of Experiment 1, about the reliability of their and others' signals.

There are two reasons for such a test. First, voting decisions take place in an ambiguous world, where voters have "some sense" of how likely to be correct they and the experts are, but such sense is vague and instinctive. Perceptual tasks capture ambiguity well and indeed have become part of economists' standard tool-kit when studying individual decision-making. They are much less common when testing group decision-making.[4] And yet, as realized by a recent theoretical literature (Ellis, 2016; Ryan, 2021; Fabrizi et al., 2021 and 2022), in voting problems, the complexity of many questions and the small marginal impact of a single vote make the ambiguity of information a particularly desirable assumption. Perceptual tasks, with the large and sophisticated literature that accompanies them, can be a useful tool for social choice scholars.

In the specific case of delegation and abstention, there is a second reason for studying a parallel experiment with ambiguous information: the possibility that the explicit numerical frame of Experiment 1 could influence the results. Consider LD first. A participant told that her precision is, for example, 55%, versus 70% for an expert, is naturally induced to compare the two numbers. Choosing to delegate seems very reasonable: if the decision involved only

[4] There is an increasing focus on strategic uncertainty. But the question is different from the lack of basic information about the distributions of relevant parameters in the population, and even about one's own parameters (precisions, for us).

the participant and the expert, delegation would always be optimal. Delegation becomes more dubious only if the participant factors in the behavior of others and the possibility that they too may delegate. But others' choices are not made salient by the question of whether to delegate to an expert. With MVA, the question is whether to abstain–that is, whether to leave the decision to all those who vote. Although the participant is again presented with the numerical values of her own and the experts' precision, abstention invokes the comparison not to the experts' precision only, but to the precisions of all those participants who do not abstain. And such precisions are not revealed. The comparison between a precision of 55% for oneself and of 70% for an expert may well be less likely to trigger an automatic reaction.[5]

We implement a classic perceptual task amply used in vision and cognitive research.[6] A number of moving dots are displayed for a very short interval (1 second); some move in a coherent direction, either Left or Right in our binary implementation, others move at random; subjects report in which direction they think coherent dots are moving. We label experts ex post as the individuals with recent performance in the highest quintile, and generate a collective decision by aggregating individual responses, with the additional option of delegation to the experts (in the LD treatments) or abstention (in the MVA treatments). We ran the experiment online with three electorate sizes: $N = 5$ and $N = 15$, as in Experiment 1, and a larger electorate of $N = 125$.

We find that the results match closely the results of Experiment 1. First, delegation remains much more frequent than abstention. Second, universal majority voting delivers the highest frequency of correct group decisions in all treatments; MVA is only slightly less efficient, while LD is dominated by both in all treatments. Interestingly, the evidence also suggests that participants, while estimating correctly their own accuracy, substantially overestimate the accuracy of the experts. This is an important point because it highlights a natural concern about LD: the possibility that experts be chosen mistakenly. In Experiment 1, the tight mathematical design sets boundaries on how unreliable the experts' signals may be; in Experiment 2, the divergence between reality and beliefs is larger.

In conclusion, then, the two experiments, different as they are in their design, complement each other. They reach the same qualitative conclusion, strengthening our confidence in its

[5]The hypothesis that precise mathematical information becomes counterproductive when it encourages a wrong mental model finds echos in the literature. In a recent paper, Esponda et al. (2023) report that subjects given precise quantitative information in a simple belief updating task are less able to learn from experience, relative to a control treatment where such precise information is withheld. The knowledge of the base parameters induces a persistent mistaken mental model.

[6]The Random Dot Kinematogram (RDK) was originally developed to study the perception of motion under noisy conditions in humans and non-human primates (e.g. van de Grind et al., 1983). In neuroscience, it has been used to study the neuronal correlates of motion perception (Newsome et al., 1989; Britten et al., 1992; Roitman and Shadlen, 2002).

robustness: evaluated on informational benefits alone, we do not find evidence in favor of LD.

One final contribution of this paper is to stress the theoretical similarity between delegation and abstention in a common interest environment. Because of such similarity, it is important to evaluate their relative performance. Note what this also implies for generalizations of the model. LD proponents point out, with good reason, that when acquiring information is costly, the advantages of delegation increase. But cost savings are possible with abstention as well. In fact, under MVA, the inability to free-ride on expert information tempers the incentive not to become informed, and the higher equilibrium quality of decision-making can more than compensate for the cost of investing in information. The two options–delegation and abstention–can also be compared when extending the model to include private values, or when studying the impact of correlated signals. We return to these points in the discussion section at the end of the paper.

Our work is related to three separate literatures. First, to the study of voting as information aggregation. The informational costs and benefits of delegation in pure common interest voting problems were the subject of early studies on the Condorcet Jury Theorem (Margolis, 1976; Grofman et al., 1983; Shapley and Grofman, 1984). These studies asked important statistical questions but did not focus on equilibrium behavior. More recent work (Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1997; McLennan, 1998; Wit, 1998) put the analysis of the Condorcet Jury Theorem on solid equilibrium grounds, but abstracted from the focus on delegation. We did not find in the literature our starting theoretical result: in a finite sample, the efficient equilibrium must allow for delegation. The literature on proxy voting in finance shares our interest in delegation and information aggregation. Bar-Isaac and Shapiro (2020) study the conditions under which a voter representing a large block of shares chooses to abstain with some of them, not to overwhelm the information dispersed among smaller but independently informed other voters. Malenko and Malenko (2019) and Buechel et al. (2023) ask whether for-profit proxy advisors — firms selling voting recommendations to share-holders — can decrease shareholders' incentives to acquire independent information, and thereby result in worse corporate decisions. Both types of questions are tailored to the corporate world, where the existence of a blockholder need not be the result of delegation, and for-profit advisors sell information at a cost.

The direct comparison between delegation and abstention, central to our work, is not the focus of these works. Information aggregation under abstention has been studied either in models with pure common interest, like ours (McMurray, 2013), or when partisan voters are present (as in Feddersen and Pesendorfer, 1996). Morton and Tyran (2011), Mengel and Rivas (2017), and Battaglini et al. (2010) test the two models experimentally in small committees. We are not aware of published experimental works that study delegation in voting as a tool for information aggregation.[7]

The second strand of related works are studies of LD, mostly in normative political theory and computer science. Green-Armytage (2015) and Blum and Zuber (2016) discuss possible advantages of Liquid Democracy, on both epistemic and egalitarian reasons: decisions are taken by better informed voters, and LD avoids the creation of a class of semi-permanent professional representatives. The focus is normative, and these studies do not analyze strategic incentives. The computer science literature is instead largely concerned with understanding how LD can work in practice. It models behavior via a priori algorithms and studies rich interactions where delegation takes place on networks (Christoff and Grossi, 2017; Kahng, Mackenzie and Procaccia, 2018; Bloembergen, Grossi and Lackner, 2019; Caragiannis and Michas, 2019). These authors connect LD to the social choice tradition, but here too strategic considerations are absent. An exception is Armstrong and Larson (2021) which discusses the informational trade-off involved in delegation and focuses on a Nash equilibrium. The paper retains the algorithmic flavor of this literature by modeling the delegation choice as sequential; the common interest nature of the problem, together with the added assumptions of complete information and costly voting, then results in the equilibrium superiority of delegation over universal majority voting. The theoretical conclusion is thus similar to ours, but the assumptions driving the result differ.

Strategic concerns are at the heart of two recent paper in economics, Ravindran (2021) and Dhillon et al. (2023). In Ravindran's model, voters' types are binary and known, with either high or low information accuracy, and the goal is the characterization of the efficient equilibrium. With a single expert, optimal delegation is defined precisely; with multiple experts, complications can arise, although, as in Armstrong and Larson, they can be solved if delegation decisions are sequential. Dhillon et al. study delegation in a model à la Feddersen and Pesendorfer (1996), with partisan voters and perfectly informed experts. Here too under complete information delegation has desirable properties: the game is dominance-solvable and delegation allows voters to coordinate on the best equilibrium. With incomplete information, multiple equilibria are more difficult to avoid and results are weaker. All of these works are theoretical or discuss numerical simulations.

[7]Revel et al. (2022) tests LD experimentally by asking groups to vote on answers to general culture questions. However, the small number of data points prevents robust conclusions, and the target metric–the percentage of votes cast in the correct direction–differs from the metric practically most relevant–the fraction of correct group decisions with and without LD. In Kawamura and Vlaseros (2017), a public statement by an "expert" conveys additional information prior to voting. The public statement moves participants' prior but there is no actual delegation of votes. Delegation is studied instead in experiments that focus on representative democracy and the aggregation of heterogeneous preferences in the electorate (see for example, Hamman et al., 2011).

Finally, a literature in social psychology studies a question that is closely related to our second experiment: if a group of individuals face, individually, a perceptual task but can then aggregate their reactions into a group decision, which decision rule for the group will reach the correct answer most frequently? In Sorkin et al. (1998) a small group of subjects are faced with a signal detection task and asked whether the display reflects noise only or signal plus noise. Although the group falls short of normative predictions, simple majority rule leads to the highest accuracy. Subsequent studies have focused on communication and confidence (Sorkin et al., 2001; Bahrami et al., 2012; Silver et al., 2021). Although it seems a natural next step, we are not aware of similar works that include the possibility of delegation.

In what follows, we begin by describing the theoretical model (Section 2) and its equilibrium properties (Section 3). We then describe our first experiment: its parametrization and treatments (Section 4); its implementation (Section 5), and its results (Section 6). Section 7 discusses the motivation and design of our second experiment; Section 8 reports its results. Section 9 sets the ground for future research, briefly discussing mechanisms that could be active under both experimental designs, and the effects of relaxing some assumptions of the model. Section 10 concludes. The Appendix collects longer proofs and some additional experimental findings.

2 A model of delegation in voting

We study the canonical problem of information aggregation through voting in a pure common interest problem. Motivated by our interest in LD, the theoretical model allows for large numbers of voters and for asymmetric information about the precision of other non-experts' signals.

N (odd) voters face an uncertain state of the world ω and must take a decision d. There are two possible states of the world, $\omega \in \{\omega_1, \omega_2\}$, and two alternative decisions $d \in \{d_1, d_2\}$. Every voter i's payoff equals 1 if the decision matches the state of the world ($d = d_s$ when $\omega = \omega_s,\ s = 1, 2$), and 0 otherwise. Voters share a common prior $\pi = Pr(\omega_1)$ and receive conditionally independent signals $\sigma_i \in \{\sigma_1, \sigma_2\}$ that recommend one of the two decisions. We call q_i the *precision* of individual i's signal, or the probability that i's signal is correct. Precision varies across individuals but is symmetric over the two possible states of the world: $q_i = Pr(\sigma_i = \sigma_1|\omega_1) = Pr(\sigma_i = \sigma_2|\omega_2)$.

The group of N voters is composed of K (odd) experts and M (even) non-experts. Whether any given voter is an expert or a non-expert is commonly known. Every expert e receives signals of known precision $q_e = p > 1/2$. The precision of a non-expert i's signal is instead private information: q_i is an independent draw from a commonly known distribution $F(q)$ everywhere continuous over support $[\underline{q}, \overline{q}]$, with $\underline{q} = 1/2$ and $\overline{q} = p$. The signals themselves are also private information, for both experts and non-experts. Each voter, whether expert or non-expert, holds a single non-divisible vote. We denote by EU individual ex ante expected payoff, before the realizations of precisions and signals. EU equals the ex ante probability that the group reaches the correct decision.

Before the election, each voter receives a signal and is informed of the signal's precision. The voter then chooses whether indeed to vote, for one or the other of the two options, or whether to delegate the vote, and in this case, whether to delegate it to an expert or to a non-expert. If delegated, the vote is assigned randomly, with equal probability, to any individual in the indicated category.[8] When counting votes, each voter who has chosen not to delegate receives a weight equal to the number of votes delegated to her, plus 1. The decision receiving more votes is chosen.

With an eye to the experimental implementation, we will select equilibria that require little coordination, and in particular those in which experts never delegate, and non-experts only delegate to experts. Hence, multi-step delegation (i delegates to j who delegates to z) will not be observed in equilibrium, and thus neither will circular delegation flows (i delegates to j who delegates to z who delegates to i). The model nevertheless needs to specify what would happen in such cases. We allow for multi-step delegation: if delegation targets a voter who has herself chosen delegation, the full packet of votes is delegated according to her instructions. However, if a set of delegation decisions results in a circular delegation flow, we specify that one link in the cycle is chosen randomly and that delegation is redirected randomly to another voter in the selected category (either expert, or non-expert).[9]

3 Equilibrium

We study an environment that matches the experimental set-up, and where, specifically, $\pi = Pr(\omega_1) = 1/2$. In this symmetric environment, conditional on voting, voting according to signal is an undominated strategy—a result that holds whether delegation is allowed, as in our model, or is not, as in traditional majority voting. Again with the goal of selecting

[8]In the absence of distinguishing characteristics across experts, random assignment of delegated votes is a natural assumption. It is also desirable because it prevents a too unbalanced accumulation of voting power (Gőlz et al., 2018. Buechel and Mechtenberg, 2019, make a similar point when studying experts' recommendations on networks.) In our model, mixing uniformly when delegating to experts is also an equilibrium strategy if voters can target individual experts.

[9]For example, suppose i and j (non-experts) delegate to z (an expert), and z delegates to i. Then one of i or z is chosen randomly; if i is chosen, all of his votes (and thus in this case all three votes) are delegated to another random expert; if z is chosen, all three votes are delegated to another random non-expert. If all voters in the target category are delegating to someone in the cycle, then a different link in the cycle is chosen. If all N voters are in the cycle, no voting occurs and the decision is taken with a coin toss.

equilibria that require minimal coordination, we focus on equilibria where the delegation decision depends on the signal's precision, but not on its message. Thus, we select semi-symmetric Perfect Bayesian equilibria in undominated strategies where, when voting, voters follow their signal, and voters of a given type (non-experts or experts) follow the same strategy, symmetric across signals. In what follows, "equilibrium" refers to such a notion.[10] We are interested in the welfare properties of delegation, and say that an equilibrium "strictly improves over majority voting" if in equilibrium the ex ante probability of reaching the decision that matches the state of the world is strictly higher than under (sincere) majority voting (MV), or $EU_{LD} > EU_{MV}$. Our starting point is the following general theoretical result:

Theorem. *Suppose $\pi = Pr(\omega_1) = 1/2$. Then for any F and for any N and K odd and finite there exists an equilibrium with delegation that strictly improves over MV.*

We prove the theorem in the Appendix, but the intuition is both straightforward and interesting. The essence of the proof is that, when delegation is possible and some voters' information may be barely better than random, there cannot be an equilibrium where delegation is excluded with probability 1: every voter casting their vote with probability 1 (and thus replicating MV) is not an equilibrium. But in this common interest problem, we know from McLennan (1998) that if a set of strategies that maximizes expected utility exists, then it must be an equilibrium. Such a set does exist in our game, and thus an equilibrium must exist that strictly improves over MV. But then such an equilibrium must include delegation.[11]

Note that, although intuitive, the result in the theorem does not follow automatically from the superior information of the experts. In its full generality, it requires a positive probability of voters with close-to-random signals ($\underline{q} = 1/2$). If even the least informed non-experts always hold useful information, then delegation can strictly improve over MV only if the difference in precision between experts and non-experts' signals is sufficiently large. For example, if there is a single expert, $\underline{q} = .65$, and $p = .75$, then for any number of non-experts majority voting is superior to delegation.[12]

The theorem does not characterize the equilibria with delegation. We do so in the Appendix and in Section 4, when we specialize the model to the parameter values we use in the experiment. Here, to ensure that the mechanisms that drive the model are intuitively

[10] To be clear: the symmetry restrictions we impose are equilibrium selection criteria, not assumptions.

[11] Because the environment is fully symmetric for all voters of a given type, the conclusion continues to apply when we restrict attention to semi-symmetric strategies.

[12] Precisely, if there is a single expert, majority voting dominates delegation if $Log[p/(1-p)] < 2Log[\underline{q}/(1-\underline{q})]$.

clear, we characterize a semi-symmetric interior equilibrium for arbitrary $F(q)$ and p but a single expert.

Proposition 1. *Suppose $\pi = Pr(\omega_1) = 1/2$ and $K = 1$. Then for any N odd and finite, there exists an equilibrium such that: (i) the expert never delegates her vote and always votes according to signal; (ii) there exists a threshold $\widetilde{q}(N) \in (\underline{q}, \overline{q})$ such that non-expert i delegates her vote to the expert if $q_i < \widetilde{q}$ and votes according to signal otherwise. Such an equilibrium strictly improves over MV and is ex ante maximal among sincere semi-symmetric equilibria where the expert never delegates and non-experts delegate to the expert only.*

The proposition is proved in the Appendix, but the structure of the equilibrium is intuitive. The key observation is that in all interior equilibria (in fact, for any number of experts), non-experts must adopt monotone threshold strategies—there must exist a precision threshold $\widetilde{q}$ such that voters with lower precision delegate, and voters with higher precision do not. The reason is immediate: if the voter delegates, expected utility does not depend on the voter's precision; but if the voter does not delegate, there is a non-zero probability that the voter is pivotal, in which case expected utility increases with the voter's precision. Given monotone threshold strategies, the Appendix shows that the delegation directions in the proposition—the expert never delegating and non-experts delegating to the expert only—are indeed best responses when all others adopt them too. Finally, we also find that the condition identifying the equilibrium threshold corresponds to the first order condition from the maximization of ex ante expected utility, over all profiles of semi-symmetric monotone threshold strategies with sincere voting and the specified directions of delegation.

It is important to note that the equilibrium threshold $\widetilde{q}$ that supports the improvement over MV is strictly interior to the range $(\underline{q}, \overline{q})$. Because $\overline{q} = p$, that means that in equilibrium there are voters who know that their precision is strictly lower than the expert's precision, and yet cast their vote, rather than delegating. In fact, since delegation decreases the aggregate information in the system, and yet the equilibrium with delegation is superior to MV, we expect the threshold $\widetilde{q}$ to be low—only voters with very imprecise information delegate in equilibrium. Indeed, this is what the solution of the model at the experimental parameters will show.

3.1 Abstention

Delegation can improve voting because it reduces the weight of less informed voters. However, abstention can do so as well, with the added advantage of being a simpler and familiar option. Note that the two mechanisms have similarities but are not equivalent: under abstention, voting weight is redistributed towards *all* voters who choose to vote; under LD, delegated votes target the experts only. One goal of our experiments is to compare the performance of the two mechanisms.

Our model is easily adapted to abstention, keeping the underlying environment constant. After voters learn, privately, the precision and the content of their personal signal, they decide, simultaneously and independently, whether to vote or to abstain. Everything else remains unchanged. The model of abstention is closely related to McMurray (2013), and its main results—the existence of an equilibrium in monotone cutpoint strategies, and its superiority to MV—carry over to our setting. As in the case of delegation, and for very similar reasons, abstention too is limited to voters with weak information.[13]

We do not replicate McMurray's results here, but report in the Appendix our model's equilibrium conditions with either delegation (LD) or abstention (which we denote by MVA: "majority voting with abstention"), for arbitrary parameter values. In the next section, we describe the equilibrium predictions for the parameter values we take to the lab. Such predictions will make clear that neither of the two systems systematically dominates the other: the comparison depends on parameter values. Two observations however are intuitive. First, because delegation concentrates power in the hands of the experts only, the benefits from higher experts' accuracy are more robust under LD than under MVA. Consider for example a limit scenario where experts are fully informed ($p = 1$). Under LD, there is a unique equilibrium with full delegation. Under MVA, the equivalent equilibrium, with full abstention, also exists. However, with N odd, the equilibrium with no abstention and everyone voting exists as well, and reaches an inferior outcome. Second, because all experts have equal precision, they should have equal voting weight. They do under MVA (with a weight equal to 1) but under LD, unbalanced numbers of delegated votes are possible. As indeed our parametrization shows, MVA benefits, relative to LD, when the number of experts is higher.

Less intuitively, in a model allowing for both delegation and abstention, whether both are used with positive probability depends on parameter values. We have verified that no such equilibrium exists for our experimental parameters, and in line with this result, and with our interest in comparing the two systems, the experiment does not allow for both.

[13] McMurray's model and ours differ in two aspects. First, we assume the existence of a known group of experts with higher, known, but not perfect precision. McMurray does not distinguish experts, but widens the support of the distribution of precisions $F(q)$ to cover the full interval $[1/2, 1]$. Second, because of our experimental aim, we assume that the size of the electorate is known and may but need not be large, deviating from McMurray's large Poisson game set-up. The logic of the two models is otherwise identical. The central intuition is that best response strategies are monotone in individual precision and thus abstaining in equilibrium shifts voting weight towards better informed individuals.

4 Experiment 1: Treatments and Parametrizations

The game we study in Experiment 1 follows very closely the theoretical model, with one simplification. Under LD, we constrain the direction of delegation: experts cannot delegate, and non-experts can only delegate to the experts. Under MVA, only non-experts can choose to abstain. Both features are equilibrium strategies under the corresponding voting mechanism and we impose them in the laboratory to simplify the subjects' tasks and reduce unnecessary noise. We are interested in three main questions: (1) We consider first the simplest setting, when decisions are taken by a small group and a single expert. How well do LD and MVA perform, relative to MV? (2) Does such performance change qualitatively in a larger group with more than one expert? (3) How do LD and MVA compare to each other, both in the small and in the larger group?

In all experiments, we set $\pi = 0.5$, $p = 0.7$, and $F(q)$ Uniform over $[0.5, 0.7]$. We study four treatments: LD and MVA for two different group sizes. The smaller group consists of 5 voters and a single expert: $N = 5$, $K = 1$. Using the number of experts as index, we call the two corresponding treatments LD1 and MVA1. Maintaining the share of experts fixed at one fifth of the group, we study a larger group of 15 voters in all, with 3 experts: $N = 15$, $K = 3$. The two treatments corresponding to the larger group are denoted by LD3 and MVA3.

4.1 Liquid Democracy

Table 1 reports the theoretical predictions under the option of delegation.[14]

Table 1: $p = 0.7$, $F(q)$ Uniform over $[0.5, 0.7]$

$LD1 : N = 5;\ K = 1$

$\widetilde{q}$	$F(\widetilde{q})$	EU_{LD}	EU_{MV}
0.7	1	0.7	0.717
0.543	0.215	0.731	

$LD3 : N = 15;\ K = 3$

$\widetilde{q}^3$	$F(\widetilde{q}^3)$	EU^3_{LD}	EU^3_{MV}
0.532	0.162	0.843	0.832

In treatment LD1, we find two semi-symmetric equilibria. For any realization of non-expert precisions, there always exists an equilibrium where every voter delegates to the expert with probability 1: no individual non-expert is ever pivotal, and delegating one's vote is a (weak) best response. The expert alone controls the outcome. With semi-symmetric strategies, such an equilibrium corresponds to $\widetilde{q} = \overline{q}$ and yields ex ante utility $EU_{LD}(\widetilde{q} = \overline{q}) = p = 0.7$. Note that the equilibrium is not strict. In addition, there is a unique

[14]The details of the derivations are in the Appendix.

strict equilibrium where $\widetilde{q}$ is strictly interior. As argued earlier, the $\widetilde{q}$ threshold is low, and the ex ante probability of delegation is only just above 20%. The ex ante probability of reaching the correct decision, equivalent to the expected utility measures, is lowest when the expert decides alone ($\widetilde{q} = \overline{q}$), intermediate under MV, and highest in the equilibrium with delegation and interior $\widetilde{q}$. However, the proportional increase in the probability that the group selects the correct option is small, about 2 percent at each step.[15]

In LD3, full delegation is not an equilibrium any longer. Intuitively, when there are multiple experts and all other non-experts delegate, voter i can be pivotal only if the experts disagree among themselves. The disagreement reduces the attraction of delegation and for sufficiently high q (still smaller than $\overline{q}$) casting a vote is preferable. As we know from the theorem, the equilibrium with interior $\widetilde{q}$ continues to exist. Equilibrium delegation, however, is rare: the expected frequency of individual delegation falls to 16%. As theory teaches, the interior equilibrium yields a higher probability of a correct decision than MV. However, with the increase in the size of the group, the Condorcet Jury Theorem effect becomes pronounced: majority voting works very well and the scope for improvement is small. The percentage gain is only 1.3%.[16]

The table conveys two main messages. First, as expected, equilibrium delegation is not frequent and concerns only voters with precisions not far from 0.5. Second, the improvement in the probability of making the correct decision is small. Both findings are surprisingly robust to parameter changes. For example, with $N = 5$, increasing p to 0.9 raises the equilibrium frequency of delegation in the interior equilibrium to 29.5% (from 21.5% with $p = 0.7$), and the gain in expected utility, relative to MV, to 3.5% (from 2% with $p = 0.7$); with $N = 15$, the equilibrium frequency of delegation rises to 21% (from 16% with $p = 0.7$), but the gain in expected utility falls to 0.8% (from 1.3% with $p = 0.7$). We maintain $p = 0.7$, and in the lab expect to see similar efficiencies for LD and for MV, leaving the data free to favor either.

[15]With a single expert, the uniqueness of the semi-symmetric equilibrium with interior $\widetilde{q}$ can be proven analytically and holds for arbitrary N. Absent either communication or repetition, asymmetric equilibria are implausible in the lab because they require coordination.

[16]Obtaining comparative statics for arbitrary parameter values is difficult. With $p = 0.7$ and q uniform over [0.5,0.7], we have verified computationally that as N increases, for any given K, equilibrium $\widetilde{q}$ approaches 0.5. As expected, the sequence of equilibria converges to the asymptotic efficiency of universal majority voting.

4.2 Abstention

Table 2 shows the equilibria under the possibility of abstention, for the experimental parametrizations.[17] We denote by $\widetilde{\alpha}$ the precision threshold below which in equilibrium a non-expert abstains, and above which a non-expert votes.

Table 2: $p = 0.7$, $F(q)$ Uniform over $[0.5, 0.7]$

$MVA1 : N = 5; K = 1$

$\widetilde{\alpha}$	$F(\widetilde{\alpha})$	EU_{MVA}	EU_{MV}
0.7	1	0.7	
0.580	0.40	0.724	0.717
0.5	0	0.717	

$MVA3 : N = 15; K = 3$

$\widetilde{\alpha}^3$	$F(\widetilde{\alpha}^3)$	EU^3_{MVA}	EU^3_{MV}
0.7	1	0.784	
0.580	0.40	0.849	0.832
0.5	0	0.832	

For both group sizes, there are three semi-symmetric equilibria. Two are boundary equilibria, with either zero ($\widetilde{\alpha} = 0.5$) or full ($\widetilde{\alpha} = 0.7$) abstention; one is an interior equilibrium where, for both group sizes, a non-expert abstains if precision is below 0.58, i.e. with ex ante probability of 40%. The boundary equilibrium with zero abstention corresponds to MV; the one with full abstention, where the decision is delegated to the experts, is inferior to MV. As in McMurray's analysis, the interior equilibrium does deliver expected gains over MV, but these remain quantitatively small.[18] The interior equilibrium threshold for abstention is higher than the threshold for delegation, and remains constant in the two group sizes. It implies a larger expected number of abstentions than delegations: for example, and rounding up to integers, when $N = 15$, in equilibrium we expect 2 non-experts to delegate under LD, but 5 non-experts to abstain under MVA.

The expected improvements over MV remain minor under MVA. Note in particular that the choice of parameters does not stack the experiment in favor of either LD or MVA. LD dominates MVA at $N = 5$, but the reverse holds at $N = 15$. Which system performs better depends on parameters. Since one of our questions is their relative performance, this is important: in theory and in the lab, LD and MVA are expected to perform similarly. [19]

[17] As in the case of delegation, we focus on semi-symmetric Perfect Bayesian Equilibria in undominated strategies where abstention strategies are invariant to signal realizations. The details of the derivations are in the Appendix.

[18] The existence of the boundary equilibria depends on N and Kodd.

[19] As expected, a higher p would favor LD over MVA, but here too the magnitude of the change is small. At $p = 0.9$, with $N = 5$, in equilibrium EU_{LD}/EU_{MVA} =1.016, v/s 1.01 with $p = 0.7$. With $N = 15$, in equilibrium MVA continues to dominate LD, although the difference is reduced: $EU_{MVA}/EU_{LD} = 1.001$ at $p = 0.9$, v/s 1.007 at $p = 0.7$.

4.3 Robustness

In the lab, as in life, some deviation from optimal strategies is highly likely. For both LD and MVA, how robust are potential improvements over MV to strategic mistakes? We consider here a particularly simple parametrization of strategic mistakes: we suppose that behavior remains symmetric, but the precision threshold for delegation or abstention is chosen incorrectly. In Figure 1, the horizontal axis is the common threshold, and the vertical axis reports gains and losses in expected utilities relative to MV (fixed at 1). Thus the plots depict the percentage changes in the probability of the group making the correct choice, relative to MV, at different delegation or abstention thresholds. LD is plotted in blue; MVA in green; the first panel corresponds to $N = 5$, $K = 1$; the second to $N = 15$, $K = 3$. The highest points on the blue and green curves coincide with the respective equilibrium thresholds. The declines away from the highest points capture the cost of strategic mistakes: delegating or abstaining too much (at too high thresholds of precisions, relative to equilibrium, or to the right of the highest points), or too little (voting at too low precision, or to the left of the highest points).

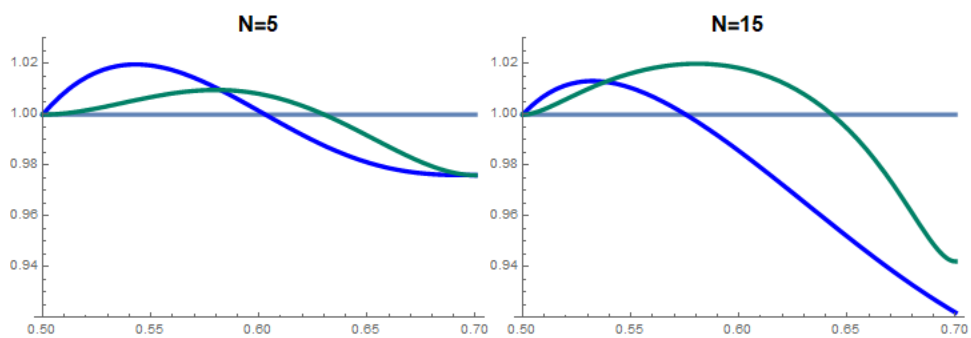

Figure 1: *Robustness to strategic errors.* The horizontal axis is $\widetilde{q}$ ($\widetilde{\alpha}$), the vertical axis is the probability of reaching the correct decision, relative to MV. Blue is LD, green is MVA, and grey is MV.

At $\widetilde{q}$ or $\widetilde{\alpha} = 1/2$, no-one delegates their vote or abstains, and all curves equal MV and coincide. At $\widetilde{q}$ or $\widetilde{\alpha} = 0.7$, all non-experts delegate or abstain, and only the expert(s) decide(s).[20] In the first panel, with a small group, the maximum potential improvement over MV from delegation (from LD) is higher than from abstention (MVA). However, this is not true in the second panel, with the larger group. The two results were already shown in Tables 1 and 2; here the figure depicts them graphically. It adds to the earlier tables the range of thresholds for which each voting rule dominates MV. Here the message is consistent across the two group sizes: in both cases, the range of thresholds that deliver improvements

[20]When $K = 3$, the blue and green curves do not coincide at $\widetilde{q} = 0.7$ because under MVA3 each expert has the same weight, while under LD3 the number of votes each of them commands is stochastic.

over MV is limited, and particularly limited for LD. When the group is larger, LD's potential for losses is evident in the figure, as is its increased fragility, relative to MVA: the range of thresholds that improve over MV is half as large under LD3 than under MVA3. With both voting schemes, but with LD in particular, while potential gains are small, there is the real danger of reaching worse decisions: under LD3, maximal potential losses are more than six times maximal potential gains.

5 Experiment 1: Implementation

We ran the experiment online over the Summer of 2021, using the Zoom videoconferencing software. Participants were recruited from the Columbia Experimental Laboratory for the Social Sciences (CELSS)' ORSEE website.[21] They received instructions and communicated with the experimenters via Zoom, and accessed the experiment interface on their personal computer's web browser. The experiment was programmed in oTree and, with the exception of a more visual style for the instructions, developed very similarly to an in-person experiment. Each session lasted about 90 minutes with average earnings of $26, including a show-up fee of $5.

Participants were asked to vote on the correct selection of a box containing a prize, out of two possible choices, a green box and a purple box. The computer selected the winning box putting equal probability on either; conditionally on the computer's random choice, participants then received a message suggesting a color, and were told the probability that the message was accurate.[22] The same screen also informed them of whether or not they were an expert (for that round). Participants were then asked to vote for one of the two boxes, if experts, or, if non-experts, to either choose one of the boxes or delegate their vote to an expert (in the LD treatments), or abstain (in the MVA treatments). After each round, all participants received the same feedback, reporting where the prize was, how many votes each box received, how the experts voted, and how many non-experts delegated their vote (in LD treatments) or abstained (in MVA treatments). Across rounds, expert/non-expert identities were re-assigned randomly, under the constraint that groups of 5 voters had a single expert, and a group of 15 had three; if the session involved multiple groups, they were re-formed randomly. A copy of the instructions is reproduced in online Appendix C.

[21] Greiner (2015). CELSS' ORSEE subjects are primarily undergraduate students at Columbia University or Barnard College.

[22] To limit decimal digits, the precision of the signal was drawn uniformly from a discrete distribution with bins of size 0.01. When comparing the experimental results to the theory, below, we compute equilibria using the corresponding discrete distribution of precisions. The differences from using a continuous distribution are minute.

We ran 10 sessions, each involving 15 subjects (150 subjects in total). Participants played 20 rounds each of two treatments (40 rounds in total), according to an experimental design that balanced order and treatment composition across sessions. We reproduce it in Table 3 below. In total, we have data for 240 rounds for LD1 and for MVA1, and 120 rounds for LD3 and for MVA3.

Table 3: *Experiment 1: Experimental Design*

Sessions	Treatments	Rounds	Subjects	Groups
1a	LD1, LD3	20, 20	15	3, 1
1b	LD3, LD1	20, 20	15	1, 3
2a	MVA1, MVA3	20, 20	15	3, 1
2b	MVA3, MVA1	20, 20	15	1, 3
3a, 3a'	LD3, MVA3	20, 20	15	1, 1
3b, 3b'	MVA3, LD3	20, 20	15	1, 1
4a	LD1, MVA1	20, 20	15	3, 3
4b	MVA1, LD1	20, 20	15	3, 3

6 Experiment 1: Results

6.1 Individual behavior

Under both voting systems, equilibrium voting strategies are monotonic in precision (if non-expert i votes at precision $q(i)$, then i votes at all $q'(i) > q(i)$). As we document in the Appendix, individual behavior in the experiment is mostly monotonic. There is weak evidence of fewer monotonicity violations under MVA, but the two treatments are effectively comparable. Just below 60% of subjects have no violations at all under LD; just above 60% under MVA, and the fractions reach 80% and above if we limit attention to the last 10 rounds of each treatment. The results are invariant to the size of the group. In all cases, it is possible to generate perfect monotonicity for at least 80% of participants by changing at most 2 of their non-expert choices.[23]

The data convey the subjects' sensitivity to the precision of the signals. The panel on the left in Figure 2 reports the empirical frequencies of delegation or abstention as functions of individual precision. In all treatments, the frequencies decline at higher signal precision. Note however the differences across treatments: for any precision the delegation rate is higher than the abstention rate, for both group sizes.

[23]With type randomly assigned, the expected number of rounds played as non-experts is 16. Given the binary choice of voting or not, the maximum possible number of monotonicity violations over 16 rounds is 8.

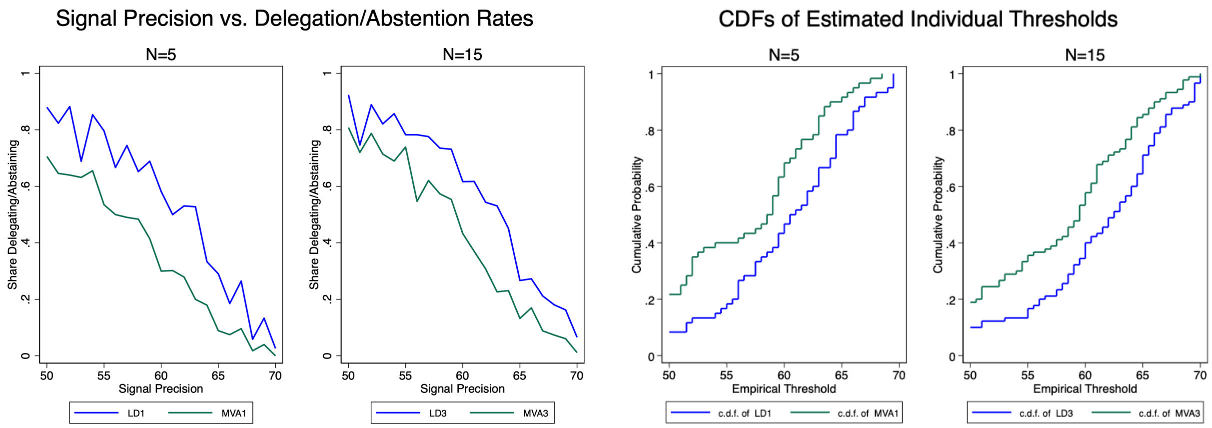

Figure 2: *Delegation and abstention frequencies.* Left: Empirical frequencies in the data. Right: CDFs of estimated individual voting thresholds.

The higher propensity towards delegation, relative to abstention, is confirmed by the panel on the right of Figure 2, where we plot the the cumulative distribution functions of estimated individual voting thresholds.[24] For both group sizes, the LD distribution, in blue, first order stochastically dominates the MVA distribution, in green: at any precision, including at the lower boundary of the support, the fraction of subjects estimated to delegate is above the corresponding fraction of abstainers (the fraction of subjects whose estimated threshold is below the threshold, and thus are voting at that precision, is lower). Two-sample Kolmogorov-Smirnov tests adjusted for discreteness confirm the visual impression: for both group sizes, the probability that the two samples of thresholds, for LD and for MVA, are drawn from the same distribution is very low ($p = 0.034$ for $N = 5$, and $p = 0.0047$ for $N = 15$). On the other hand, for given voting system, allowing delegation or abstention, the data do not show substantive differences between the two group sizes.[25]

The higher frequency of delegation cannot be imputed to the existence of an equilibrium with full delegation under LD1.[26] The number of subjects choosing to always delegate is very small and comparable across LD and MVA treatments. No subject (out of 60) has an estimated threshold of 0.7 (never vote) under either LD1 or MVA1.; the number rises to 3 (out of 90) under LD3 (where full delegation is not an equilibrium), and 1 for MVA3.

The gradual decline in delegation or abstention as precision increases (in the left panel of Figure 2) and the dispersion in estimated thresholds (in the right panel) are typical of similar

[24]Exploiting monotonicity, we can estimate individual voting thresholds $\widetilde{q}_i$ (or $\widetilde{\alpha}_i$) that, for each subject, minimize the frequency of monotonicity violations. See the Appendix for details.

[25]Comparing LD1 and LD3, the KS test yields $p = 0.3452$; comparing MVA1 and MVA3, it yields $p = 0.5332$.

[26]Note that such an equilibrium also exists under the MVA treatments. In fact, it is a weak equilibrium under LD1, but a strict equilibrium in the MVA treatments.

experiments (for example, Levine and Palfrey, 2007; Morton and Tyran, 2011). The data cast some doubt on the focus on symmetric equilibria, but each individual's precision draws are relatively few,[27] unlikely to span the whole support of precisions, and heterogeneous across subjects. Some gradual changes in delegation and heterogeneity in estimated thresholds will result by construction.

Regressions of individual voting behavior on signal quality, controlling for round and treatment order effects, confirm the sensitivity to precision and the higher propensity towards delegation. Table 4 reports linear probability regressions for LD1 and MVA1 ($N = 5$) and for LD3 and MVA3 ($N = 15$).[28] As expected, the propensity to abstain or delegate responds negatively to higher precision of the signal, similarly across the two group sizes. The coefficient of the LD dummy is positive and highly significant for both group sizes.[29]

6.2 Aggregate results

6.2.1 Frequency of delegation and abstention

Figure 3 reports the aggregate frequencies of delegation (in blue) and abstention (in green) in the data, and according to the predictions of the interior equilibrium, given realized precision draws in the experiment (in grey). Columns on the left refer to LD treatments; columns on the right to MVA. The 95% confidence intervals are calculated from standard errors clustered at the session level.

Delegation rates in the experiment are between two and three times what theory predicts for the interior equilibrium. Abstention rates on the other hand are comparable to the predictions. The conclusion is robust to all plausible ways of cutting the data: disaggregating by session, considering only the 10 final rounds, clustering standard errors at the individual level. The equilibrium with full delegation (under LD1) and the equilibria with either full or zero abstention (under both MVA1 and MVA3) are equally rejected by the data. Note in passing that the high frequency of delegation cannot be attributed to participants best responding to others' experimental choices: too high delegation rates by others depress own optimal delegation rates.

[27] As noted, 16 rounds on average are played as non-expert.

[28] "Second" indicates that the treatment appeared second in the session. "Mixed" indicates that both an LD treatment and an MVA treatment appeared in the session. In both columns, the excluded case is MVA played as first treatment in MVA-only sessions.

[29] There is some evidence of order effects, especially for $N = 5$, but the net effect of delegation remains always positive. Running the regressions on first treatments only gives a second check on the importance of these effects, at the cost of fewer data and less experience. The results for LD1 and MVA1 remain unchanged; in the larger groups, with fewer data points, the standard errors are larger and the parameters less precisely estimated, but the coefficient of the LD dummy continues to be positive. We report the regressions limited to first treatments, as well as Probit estimates of Table 4 in the Appendix.

Experiment 1: Frequency of Delegation or Abstention.

	(1) N=5	(2) N=15
LD	0.328***	0.208**
	(0.073)	(0.071)
	[0.006]	[0.022]
Signal Precision	-0.777***	-0.861***
	(0.080)	(0.047)
	[0.000]	[0.000]
LD * Signal Precision	-0.078	0.011
	(0.059)	(0.037)
	[0.248]	[0.783]
Round	0.031	0.078
	(0.057)	(0.056)
	[0.613]	[0.208]
LD * Round	-0.112	-0.051
	(0.100)	(0.067)
	[0.314]	[0.475]
Second	0.154***	-0.096**
	(0.010)	(0.035)
	[0.000]	[0.029]
LD * Second	-0.090	0.037
	(0.050)	(0.037)
	[0.134]	[0.349]
Second * Mixed	-0.129***	0.078***
	(0.002)	(0.006)
	[0.000]	[0.000]
LD * Second * Mixed	-0.025***	-0.166**
	(0.006)	(0.055)
	[0.008]	[0.019]
Constant	0.675***	0.832***
	(0.048)	(0.069)
	[0.000]	[0.000]
Observations	1,920	2,880
R-squared	0.309	0.309

*** p<0.01, ** p<0.05, * p<0.1

Table 4: *Determinants of delegation and abstention.* Linear probability models. Standard errors in parentheses, clustered at the session level. P-values in brackets. Delegation/abstention is measured as a binary 0-1 subject decision. The values for signal precision and round have been scaled to be between 0 and 1.

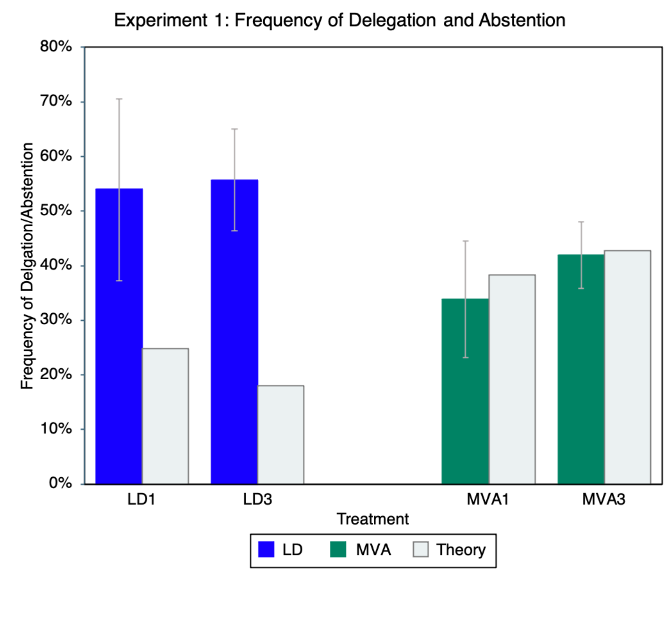

Figure 3: *Aggregate frequency of delegation and abstention.* Confidence intervals are calculated from standard errors clustered at the session level.

6.2.2 Frequency of correct choice

Beyond regularities of delegation and abstention, the real variable of interest is the frequency with which the voting system leads to the correct choice. We begin by reporting the data. Because we are studying variations of majority voting, a large share of outcomes under both LD and MVA correspond to MV. Testing the relative performance of the voting systems requires conditioning on reaching different outcomes, and we will move to that after describing the data.

Figure 4 reports the experimental data and compares them to the theoretical interior equilibrium and to MV. For subjects who delegated or abstained, the outcomes under MV are constructed with a positive probability of voting against signal. In the figure, such a probability equals the frequency observed in the treatment, but the comparison across voting systems is unchanged across robustness checks.[30]

We report results grouped by N. The vertical axis is the frequency of correct outcomes over the full data set for the corresponding treatment. The figure holds three main lessons.

[30]We replicated the figure using the subject's own observed frequency of voting against signal in the treatment; or imposing sincerity in all systems; or selecting the highest/lowest frequency of votes against signals across treatments and imposing it on all. The qualitative features of the figure are unchanged. To account for the randomness in MV data and then for consistency, all 95% confidence intervals are calculated from bootstrapping, using 100,000 simulated data sets. Each subject is drawn with a full set of 20 choices, thus allowing for within-subject correlations. All results are calculated given the experimental realizations of the state and of precision draws.

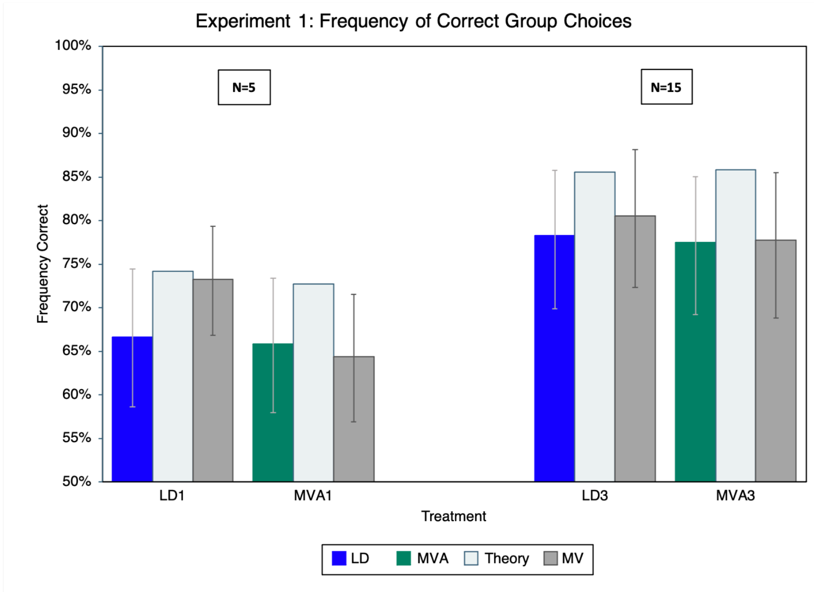

Figure 4: *Frequency of correct outcomes.* For subjects who delegated or abstained, MV data allow voting against signal with probability equal to the frequency observed in the treatment. Confidence intervals are calculated from bootstrapping, using 100,000 simulated data sets and allowing for within-subject correlation.

First, for both group sizes, LD and MVA yield very similar frequencies of correct decisions. Second, for both group sizes, both systems fall short of their possible best performance. Third, MVA outcomes are closely comparable to MV for both group sizes, but LD outcomes fall short of MV, especially for small groups.

Two main deviations could be responsible for the systems' underperformance relative to the theory.[31] The first is the erroneous choice of delegation/abstention thresholds. Figure 3 strongly supports this interpretation for LD, but not for MVA. The second is random voting in the form of voting against signal. As we document in the Appendix, the frequency of voting against signal correlates negatively and significantly with signal precision. Thus experts vote against signal more rarely than non-experts. Both because experts cast multiple votes under LD, and because subjects abstain less and thus choose to vote at lower precision under MVA, the share of votes cast against signal is lower in the LD treatments (at about 6%) than in the MVA ones (at about 10%), with little difference across group sizes. These numbers are comparable to those found in similar experiments,[32] but mean that MVA suffers

[31]A priori, a third possibility would be non-monotonicity in delegation and abstention decisions. But as we described earlier, violations of monotonicity are rare.

[32]Guarnaschelli et al. (2000) and Goeree and Yariv (2011), for example, report frequencies of voting against signal of 6-9% for juries voting under simple majority and pure common interest, in the absence of communication. In our experiment, in all treatments, 70-80% of subjects never vote against signal; for the remaining 20-30%, votes against signal are cast occasionally, with frequency negatively correlated to

from more random voting. Both systems thus fail to realize their potential gains over MV, but for different reasons.

The comparison to MV shows that the penalty is higher for LD, and particularly for LD1. The better performance of MV in the LD samples reflects the random superiority of the signal draws in those samples. Comparing Figure 4 to the model's robustness results in Figure 1 helps to make the point more precise. In LD1, the average voting threshold in the lab was 0.604 (v/s the optimal interior threshold of 0.543). According to Figure 1, if other deviations had minimal influence, at that threshold LD1 should have performed similarly to MV. The difference comes from the stochasticity of the signals: in LD1, experts' realized signals were correct only 66.7% of the times, while non-experts' were correct more frequently than expected (at 62.2%)–a bad combination with a pro-delegation bias.[33] Under LD3, the average threshold in the lab was 0.614 (v/s the optimal internal threshold of 0.532). According to Figure 1, missing other deviations, LD3's loss relative to MV should be just below 3%, and indeed it was 2.7%. For LD3, the figure provides an accurate enough prediction because the realized signals were correct with the expected frequencies–on average, 70% for the experts, and 60.8% for the non-experts.

These observations hold a lesson: under LD, for given frequency of delegation, results are very sensitive to the differential accuracy of the experts, relative to non-experts. This is particularly true when there is a propensity to over-delegate, as we find in the lab. MVA, on the other hand, is a more robust system by design, because it relies on votes by both experts and the most accurate of the non-experts. The robustness to incorrect voting shown in Figure 1 is mirrored in the lab.[34]

6.2.3 Comparing LD and MVA to MV

Evaluating the significance of the disparities observed between LD or MVA on one side, and MV on the other is not immediate. One difficulty is the complexity of the correlation structure,[35] but the fundamental problem is simpler: as mentioned above, outcomes coincide in a large majority of cases.[36] Restricting the data sample to those elections in which outcomes differ leaves us with little information. To overcome this difficulty, we use bootstrapping

precision. One outlier subject in session 2a always voted against signal.

[33]The better than expected signals for non-experts explain the superior performance of MV in the treatment.

[34]Under MVA1 (MVA3), experts' realized signals were correct 72% (71%) of the times, while non-experts' were correct 58% of the times on average in both MVA treatments.

[35]Individuals are observed over multiple rounds; the frequency with which they are assigned the role of experts is random and variable; the imputation of missing votes under MV creates randomness in the MV outcomes.

[36]More than 70% of all outcomes under LD, and more than 80% for MVA

methods to simulate a large number of elections in a population for which our data are representative. By simulating many elections, conditioning on different outcomes becomes feasible.

The procedure we implement allows for correlation across an individual's multiple decisions, and uses randomization to generate the correct balance of experts and non-experts. For each voting system and group size, we generate outcomes by drawing subjects, with replacement, each with their full set of 20-round decisions, and matching them randomly into groups. We then study the outcomes corresponding to 100,000 replications of the experiment for each treatment, using the population of subjects for that voting system and group size. Thus we analyze 100,000 replications of 240 decisions for LD1 and MVA1, and 120 decisions for LD3 and MVA3. We describe the procedure in more detail in the online Appendix. Figure 5 shows the distributions of the differential frequency of correct decisions between the voting systems we are studying and MV, for each group size, conditioning on the decisions being different. Consider for example LD1. For each of the 100,000 simulations, we focus on the subset of elections D_{LD1} such that LD1 and MV reach a different outcome. Call $\gamma_{LD1}(D_{LD1})$ ($\gamma_{MV}(D_{LD1})$) the frequency with which LD1 (MV) is correct over subset D_{LD1}, a variable that ranges from 0 to 1. We are interested in $\gamma_{LD1}(D_{LD1}) - \gamma_{MV}(D_{LD1})$, where, by construction, $\gamma_{MV}(D_{LD1}) = 1 - \gamma_{LD1}(D_{LD1})$. Hence $\gamma_{LD1}(D_{LD1}) - \gamma_{MV}(D_{LD1}) = 2\gamma_{LD1}(D_{LD1}) - 1$. Our measure then ranges from 1—when, conditional on disagreement, LD1 always reaches the correct outcome, and MV the incorrect outcome—to -1, when the opposite holds; a value of zero indicates that the two rules are correct with equal frequency, conditioning on disagreement. The first panel of Figure 5 plots, in blue, the distribution of such variable over the 100,000 replications. The equivalent distribution for MVA is plotted in the same panel in green. The second panel reports the results for groups of size 15.[37]

For both group sizes, the blue distribution is shifted to the left, relative to the zero point indicated by the vertical black line: when LD and MV differ, the correct decision is more likely to be the one reached by MV. The asymmetry is more pronounced for $N = 5$, where the blue mass to the left of zero—the probability that MV is superior to LD1, conditional on disagreement—is 85%, versus 67% for LD3. MVA on the other hand, when disagreeing with MV, is more likely to be right than wrong: only barely when $N = 5$ and the probability that MV is superior to MVA1 is just below 50% (48%), but more substantially when $N = 15$ and the probability that MV is correct, conditional on disagreement with MVA3, falls to 26%. The distributions are also informative of the quantitative gap in the probability of being correct, relative to MV. In the panel on the left, for example, the mode of the blue

[37]Averaging over all replications, the share of elections in which the outcome differs from MV is 23.4% for LD1, 15.3% for MVA1, 20.1% for LD3, and 15% for MVA3.

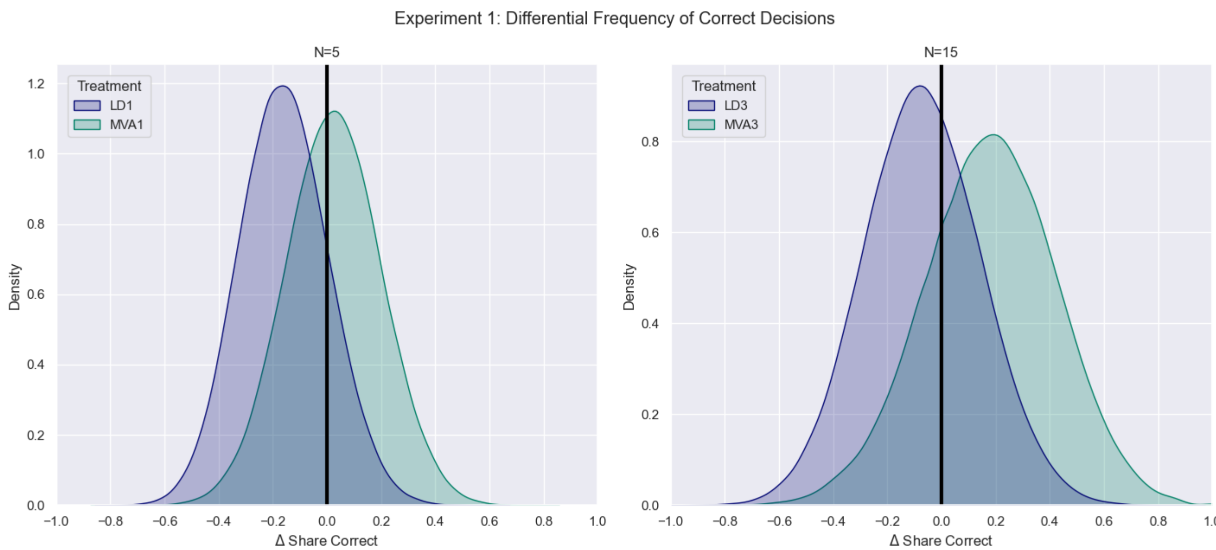

Figure 5: *Differential frequency of correct decisions, relative to MV, conditional on different outcomes.* Distributions over 100,000 bootstrap replications.

distribution at -16% tells us that over the 100,000 replications of 240 decisions, conditional on disagreement, the highest probability mass is around a frequency of correct decisions of about 42% for LD1, versus 58% for MV (or about 101 correct decisions under LD and 139 under MV) [38]

In our first experiment then, LD falls short of the hopes of its supporters, even in a streamlined environment where experts are correctly identified. Like delegation, abstention allows voters with weak information not to influence the final choice, but is simpler and performs better. In our data, its efficiency is either comparable or somewhat superior to universal majority voting, contrary to what we see for delegation.

But do these results reflect some core feature of the systems we are studying, or are they artifacts of the lab? We analyze this question in our second experiment.

7 Experiment 2: The Random Dot Kinematogram

There are three possible concerns. First, could the design, canonical as it is, be inducing an incorrect mental model? As already described, the precise numerical values make the difference in precision between one's own signal and the experts' very salient. Under LD, the design draws attention to the bilateral relationship between the subject and an expert, and

[38] The conclusions remain qualitatively similar if we construct the bootstrap ignoring the possibility of correlation in individual behavior across rounds, and thus draw each individual choice from the full data set for that treatment.

away from equilibrium considerations that include the behavior of other non-experts. Under MVA, the distortion, if it exists, is likely to be reduced because abstention automatically invokes the behavior of others: "If I abstain, who votes?"

Second, could the votes against signals, at low signal precision, also stem from the precise numerical format in which such precision is conveyed? After all, thinking that a signal of precision little above 50% should be disobeyed about half the time is a plausible thought. Outside the lab, on the other hand, individuals would be very unlikely to vote against their best estimate of the superior alternative.

Finally, by imposing that all signals are correct with probability larger than 50% our formulation could be biasing the analysis in favor of MV. The assumption is natural in a model with binary choices and known own precision, and is routinely made in experiments: conveying information with less than random accuracy would require adding a second level of noise, confusing participants. But is it a plausible assumption? The ambiguity of a perceptual task gives us a simple entry into a more realistic set-up.[39]

With this in mind, we chose to run a second experiment, where the information about the precision of the signals is ambiguous. The experiment consists of a perceptual task—the Random Dot Kinematogram (RDK)—where individual signals correspond to the accuracy of individual perceptions, and neither own nor others' accuracies are described or known in precise probabilistic terms. Because the task may be unfamiliar, we describe it in some detail.[40]

We ran the experiment via Cloud Research, with prescreening of subjects recruited from Amazon Mechanical Turk. We used three electorate sizes: $N = 5$ and $N = 15$, as in Experiment 1, but also $N = 125$, i.e. a larger size than we could run in the lab or conveniently on Zoom. In our implementation, 300 moving dots appear in each subject's screen for 1 second only; a small fraction of them (dependent on treatment) moves in a coherent horizontal direction, either Left or Right with equal ex ante probability; the rest move randomly and independently, and thus can each move in any direction over 360 degrees. The "coherence" of a task is the fraction of dots moving in the coherent horizontal direction. After one second, the image disappears and each participant reports whether the perceived coherent direction was Left or Right. We report the precise parameters in the online Appendix (the

[39] In a still current analysis of the Condorcet Jury theorem, suggestively titled "A Note on Incompetence," Margolis (1976) discussed the tension between the asymptotic efficiency promised by the theorem and political reality. What if, over some questions and for some voters, information is actually correct with probability *lower* than 1/2? As we learnt after conducting the present study, Margolis went on to advocate understanding judgment, including judgment in voting and political reasoning, through the lens of pattern recognition, starting with perception biases (Margolis, 1987).

[40] Additional information is in the online Appendix B, reproduced for convenience at the end of the current file. Experimental instructions are online at Appendix C.

size and color of the dots, the movements per frame, the random process for the dots moving randomly, etc.), but it should be clear that our experiment does not aim at measuring perception per se—for example, we cannot control the ambient light, screen size, or contrast of the monitors our subjects use. Our focus remains on collective decision-making.[41]

We divide the experiment into two parts. Each part is preceded by a few practice tasks and is subdivided into six blocs, with a bloc consisting of 20 tasks of equal coherence. Part 1 plays the role of extended training: subjects are rewarded on the basis of their individual accuracy only and experience decreasing rates of coherence across blocs, reaching the same coherence used in Part 2 for the final two blocs. At the end of Part 1, each subject is informed of her fraction of correct answers in each bloc.

In part 2, each task has both an individual component ("Choose the coherent direction"), and a subsequent group decision with the possibility of delegation (under LD), or abstention (under MVA). ("You said Left. Do you want to Vote or to Delegate (Abstain)?"). When delegation is chosen, the vote is assigned randomly to an "expert," that is, one of the participants whose accuracy is in the top 20% of the group over the last 2 blocs (40 tasks); experts are not allowed to delegate (under LD) or to abstain (under MVA). Thus, in line with Experiment 1, groups of 5 have 1 expert, and groups of 15 have 3. The group of 125 has 25 experts, and, following our standard notation, we denote the two treatments with the larger group by LD25 and MVA25. The group decision corresponds to the majority of votes cast, and individuals are rewarded both for their individual accuracy and for the accuracy of the group. As in Part 1, feedback about average individual accuracy in each bloc is provided at the end of Part 2.[42] In Part 2, coherence is kept constant across all blocs. We chose its value according to two main criteria: the task should not be so difficult that subjects are discouraged and act randomly, and should not be so easy that MV accuracy, especially in the large group, leaves effectively no room for possible improvement. Based on the results of two preliminary pilots, we fixed coherence in Part 2 at 5% for electorates of sizes 5 and 15, and at 3% for the electorate of size 125. The task is objectively hard, as the reader can verify at the following link: `https://blogs.cuit.columbia.edu/ac186/files/2022/05/rdk-video.gif`

The experiment used the RDK plugin in jsPsych (Rajananda et al., 2018) and was hosted on cognition.run. For each of LD and MVA, we recruited 60 subjects divided into 12 groups

[41] Heer and Bostock (2010) and Woods et al. (2015) report on the replication successes and challenges of conducting research on perceptual stimuli online.

[42] Feedback over group accuracy cannot be provided because it depends on choices made by others and is calculated ex post. Recall that participants are online and come to the experiment at different times. Participants were randomly divided into groups after completing the experiment, and assigned to the same group for the whole experiment.

for the $N = 5$ treatment and 90 subjects divided into 6 groups for $N = 15$ (thus replicating the corresponding number of subjects and groups in Experiment 1), and an additional 125 subjects for the largest group. There were then 275 subjects for each voting system, or 550 in total. The group size and the relevant number of experts were always made public. The experiment lasted about 20 minutes. Subjects earned $1 for participation and a bonus proportional to the number of correct responses, for a total average compensation of $4.92, or just below $15 an hour.

Relative to Experiment 1, Experiment 2 differs under a few dimensions. First, choices are made under ambiguity, and neither one's own precision nor the difference in precision between oneself and the experts are known. This is for us the experiment's major characteristic, and the core motivation for running it. Note that the exposure to the stimulus is extremely brief–1 second–and although a subject can withdraw attention, above a low threshold it seems practically very difficult to increase accuracy through increased effort. In line with psychologists' use of the task, differences in accuracy are more likely to mirror innate differences in perception.[43] Second, not only are probabilities not known, but neither we, the experimenters, nor the subjects know others' beliefs. We cannot formulate a tight theoretical prediction to which the data can be compared.

8 Experiment 2: Results

8.1 Accuracy

We define an individual's accuracy as the fraction of correct responses. Figure 6 reports the distributions of accuracy in Part 2 calculated over each of the 6 blocs for each subject, that is, over 20 tasks. The two panels correspond to the two levels of coherence used in the experiment (0.05 on the left; 0.03 on the right).

For given coherence, the distributions are very similar across treatments. In all cases, the spread in the distribution of accuracies is large, ranging from about 25% all the way to 95%. Mean accuracy over all participants is 59% in treatments with 0.05 coherence, and 56% in treatments with 0.03 coherence. Experts' accuracy is higher than non-experts': average accuracy per bloc is 63% for experts (v/s 58% for non-experts) in treatments with 0.05 coherence, and 59% (v/s 55.5%) in treatments with 0.03 coherence.[44]

[43] As for the impact of effort on the delegation/abstention decision, recall that subjects always receive payment for the accuracy of their individual answers. The incentive not to vote in the collective task is mitigated by the incentive for individual performance. In addition, here too, our focus is the comparison between delegation and abstention, both of which allow for not voting.

[44] In regressions of individual accuracy, the coefficient of an expert dummy is positive and strongly significant in all cases; the coefficient of an LD treatment dummy is barely negative and insignificant for $N = 5$

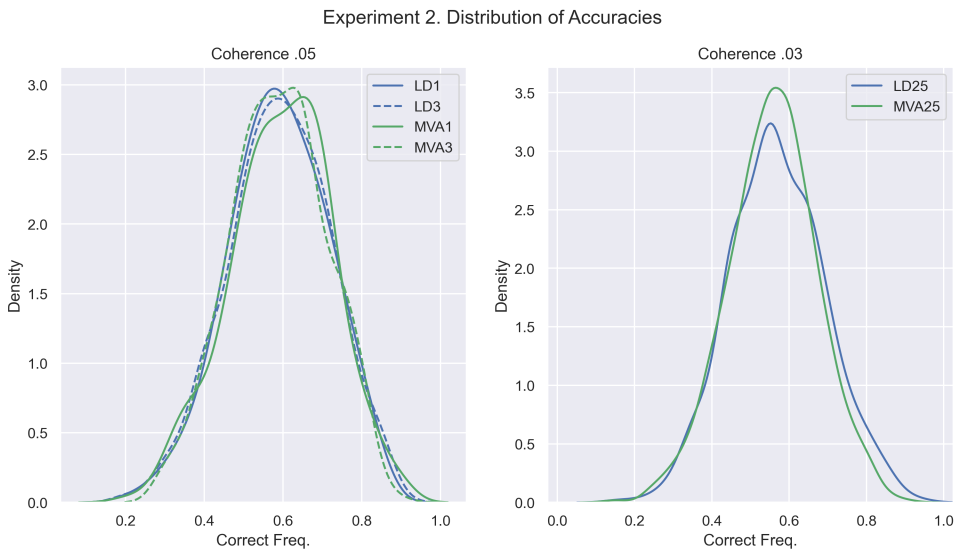

Figure 6: *Accuracies per bloc per subject. Distributions.* A bloc consists of 20 tasks.

The frequency of blocs with accuracy below 50% is non-negligible (18% for coherence of 0.05, just above 23% for coherence of 0.03) and, surprisingly, persists when we aggregate over a larger number of tasks. Averaging at the subject level over all 120 tasks, 9% of subjects have accuracy below randomness with coherence 0.05, and 12% with coherence 0.03.[45] If we want to study voting and information aggregation when information may be faulty, perceptual tasks can provide a very useful tool.

8.2 Frequency of delegation and abstention

Absent knowledge of the distributions of subjects' beliefs, we do not have theoretical predictions for delegation and abstention. We can however compare the two systems, under the plausible assumption, supported by Figure 6, that accuracies and beliefs about accuracies are comparable across the LD and MVA samples. Figure 7 plots the frequencies of delegation and abstention for each group size, calculated over non-experts only for possible comparison to Experiment 1.[46] The 95% confidence intervals are calculated from standard errors clustered at the individual level.

If we cannot talk rigorously of overdelegation, we can however see that in Experiment 2, delegation remains much more common than abstention, for all group sizes. In groups of 5, where the disparity is largest, delegation is more than twice as frequent; in groups of

and $N = 15$, and small and positive with p-values around 0.09 for $N = 25$.

[45] Individual subjects' accuracies show high variability across blocs, evidence of random noise in perceiving and recording the stimulus in the brain, as formalized in psychophysics research.

[46] The figure is almost identical if frequencies are calculated over the full sample.

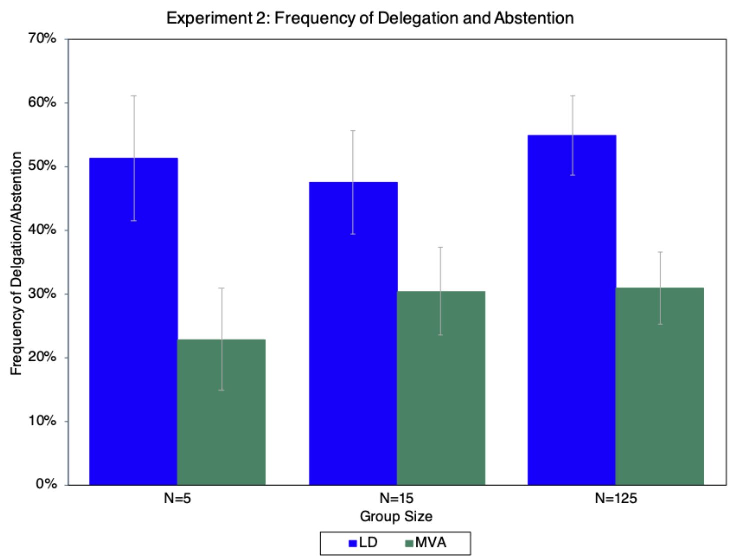

Figure 7: *Aggregate frequency of delegation and abstention (non-experts).* 95% confidence intervals calculated from standard errors clustered at the individual level.

15, where we see the least disparity, delegation is still 60% more common. The decline in coherence, from $N = 5$ or 15 to $N = 125$, has small effects on the data.

The higher frequency of delegation is confirmed in the regressions reported in Table 5. The unit of analysis is the bloc at the individual subject level (hence 6 blocs per subject), with data grouped by coherence level. The regressions reported below confirm the results of the figure: in all treatments, delegation is significantly more frequent than abstention. In Experiment 2, accuracy is at best a very weak predictor of participation in voting, never significant at conventional levels, confirming the high uncertainty in subjects' evaluation of their own accuracy.[47]

8.3 Frequency of correct outcomes

How well did the three voting systems do in Experiment 2? Figure 8 reports the frequency of correct group decisions, aggregated over all groups and tasks for given treatment.

As expected, for all three systems, the fraction of correct decisions increases with the size of the group, ranging from about 65% at $N = 5$ to 90-95% at $N = 125$. The aggregation of independent signals remains very powerful, even in the presence of weak accuracies. Relative to MV, MVA performs better than LD.

As in the case of Experiment 1, the relative performance of LD and MVA can be evaluated precisely only when conditioned on disagreement with MV.[48] We again bootstrap the data

[47]Corresponding probit regressions are reported in the Appendix.

[48]The fraction of outcomes that coincide with MV outcomes ranges from a minimum of 75% under LD1

Experiment 2: Frequency of Delegation or Abstention.

	(1) N=5 & N=15	(2) N=125
Accuracy	-0.122	0.004
	(0.084)	(0.102)
	[0.146]	[0.970]
LD	0.226***	0.223***
	(0.038)	(0.042)
	[0.000]	[0.000]
N=15	0.004	
	(0.038)	
	[0.908]	
Keys: [E][Y]	-0.009	0.029
	(0.038)	(0.041)
	[0.816]	[0.484]
Bloc	0.001	0.010
	(0.012)	(0.013)
	[0.965]	[0.459]
Constant	0.343***	0.295***
	(0.066)	(0.067)
	[0.000]	[0.000]
Observations	1,800	1,500
R-squared	0.100	0.097

*** p<0.01, ** p<0.05, * p<0.1

Table 5: *Frequency of delegation or abstention*. Linear probability models. Standard errors are clustered at the individual level. P-values in brackets. Delegation/abstention is measured as the share of rounds in a given bloc in which a subject chose to delegate/abstain (with a range from 0 to 1). Accuracy is the share of rounds in the bloc that subject answered correctly. Subjects randomly use either keys [V] and [N] or [E] and [Y] to decide whether to vote; a dummy for being assigned [E][Y] is included. The values for bloc have been scaled to be between 0 and 1; the coefficient for "bloc" thus indicates the effect of going from the first to last bloc.

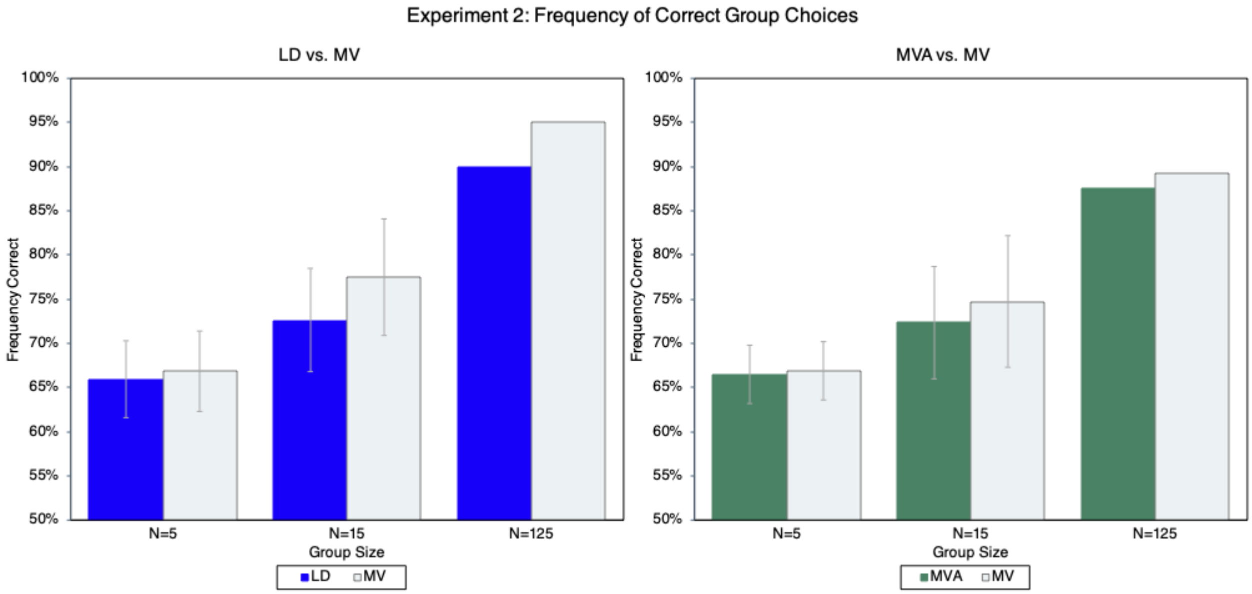

Figure 8: *Frequency of correct outcomes.* 95% confidence intervals are calculated from standard errors clustered at the group level. There is no confidence interval for $N = 125$ because there is a single group.

and replicate the group decisions a large number of times, generating a large sample of decisions over which the voting results differ, as we did with the data from Experiment 1.

Figure 9 reports the results of 100,000 simulations, with three panels corresponding, in order, to $N = 5$, $N = 15$, and the single large group at $N = 125$. As in Figure 5, we plot the distributions of the differential frequency of correct decisions under LD (in blue) or MVA (in green), relative to MV. Recall that, if the distribution is skewed to the left of the vertical line at zero, then conditional on disagreement, the correct decision is more likely to be the one reached by MV; and vice-versa if the distribution is skewed to the right.

In all three panels, the blue mass is shifted to the left: in Experiment 2 as well, LD underperforms, relative to both MV and MVA. Conditional on disagreement, the share of simulated experiments in which MV is more likely than LD to yield the correct outcome is 87% for LD1, 97% for LD3, and 95% for LD25. MVA (green in the figure) fares better: the corresponding numbers are 58% for MVA1, 69% for MVA3, and 48% for MVA25, when MVA is just barely more likely to be correct than MV, conditional on disagreement. As in Figure 5, the shapes of the distributions tell us the frequencies with which the two systems are correct, relative to MV. The proximity to zero indicates that even if the voting rule performs less well than MV, the difference need not be large. When $N = 5$ for example, the mode of the blue distribution at -0.08 says that over the 100,000 simulations the most likely result is

to a maximum of 92% under MVA25.

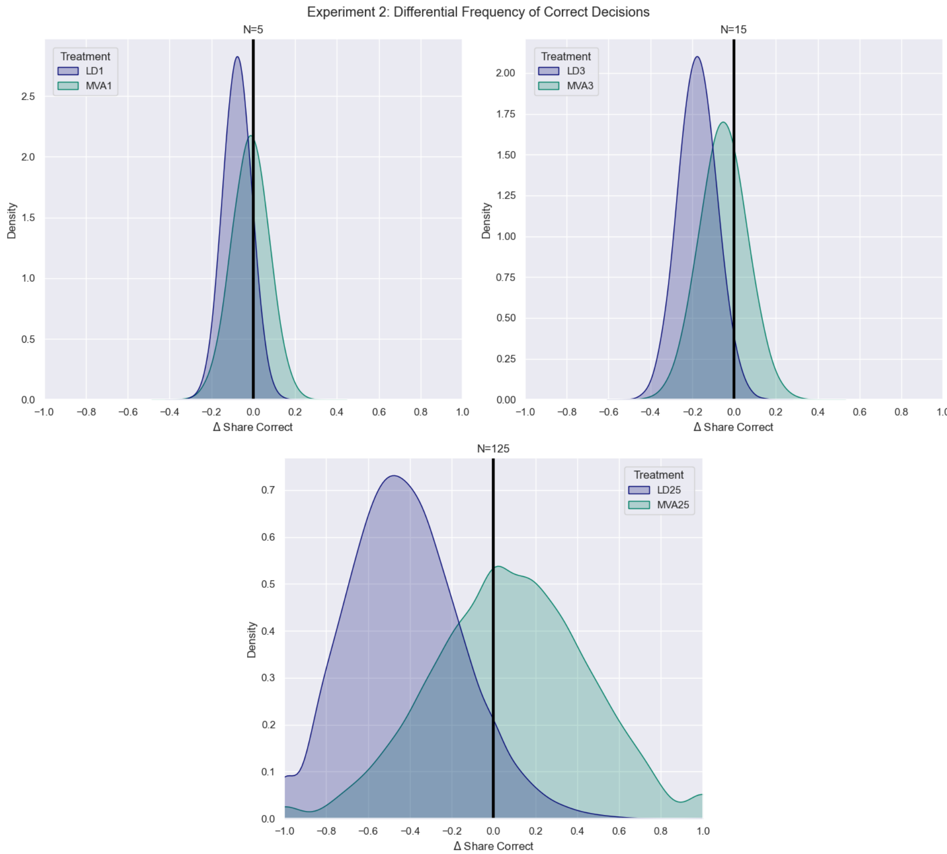

Figure 9: *Differential frequency of correct decisions, relative to MV, conditional on different outcomes.* Distributions over 100,000 bootstrap replications.

a frequency of correct decision of 46% for LD1 versus 54% for MV.[49]

In perceptual research, the focus is the immediate, apparently unconscious reaction to the stimulus, and the experimental design minimizes interference with such a process. With this in mind, we chose not to ask participants for their beliefs within each task. We did, however, add two summary questions at the end of the experiment. We asked: "On average, what percentage of trials in the second part do you think you got right?" and "On average, what percentage of trials in the second part do you think the experts got right?" With these two questions only, barely incentivized,[50] our information on beliefs can only be very noisy.

[49]The distributions are more concentrated than in Figure 5 because the number of group decisions is much higher. For $N = 5$, Experiment 2 had 12 groups, each completing 120 elections; thus each of the 100,000 data points of the bootstrapping exercise reports the differential frequency of correct decisions over 1,440 elections. In Experiment 1, each data point in the corresponding exercise reports the results of 240 elections. Similarly, for $N = 15$, the number of elections in each simulation is 720 (6 groups, each completing 120 elections), six times the 120 elections represented by each data point in Figure 5. For $N = 125$, we had a single group, completing 120 elections; thus we simulated 120 elections 100,000 times.

[50]We rewarded replies with 25 cents if the answer was within 5% of the observed percentage for the group the participant was assigned to.

Yet, we learnt three lessons. First, in all treatments, beliefs about own accuracy track actual accuracy surprisingly well: for treatments with 0.05 coherence, average believed accuracy is 58% (v/s 59% for average realized accuracy); with coherence 0.03, average believed accuracy is 55% (v/s 56% for average realized accuracy). Second, beliefs about experts' accuracy are inflated by about 15% at both coherence levels (71% v/s 63% at 0.05 coherence, and 70% v/s 59% at 0.03 coherence). Third, while, as we saw, accuracy is a very weak predictor of choosing to vote, beliefs about own and the experts' accuracy are instead strong predictors of voting behavior. Beliefs, however, cannot explain the difference in the frequency of delegation and abstention: controlling for beliefs, delegation remains substantially and significantly higher than abstention.[51]

Together, these findings suggest an interesting hypothesis. As we told participants, experts were identified as being in the top quintile of accuracies *in the two preceding blocs*. We did so to capture how a person's acknowledged expertise is built from the person's track record, exactly as in Experiment 1 the high expert's precision does not prevent bad signal realizations. But perceptual accuracy, and indeed accuracy in general, is stochastic, and the same subjects who scored in the two top quintiles in preceding blocs need not be in the top quintile in the *current* bloc. Interestingly, while beliefs about experts are biased upward, the same beliefs track very well average accuracy of the top quintile. In other words, beliefs about experts may be inaccurate because subjects neglect the reversion to the mean that accompanies the stochasticity of individual accuracies.[52]

9 Discussion

We conclude the discussion of our results with two questions. Both are introduced here as possible guidelines for future research. The first question asks what could be the reasons for the consistently high frequency of delegation, in absolute terms and relative to abstention, across the two experiments. The second question addresses the maintained assumptions of our model. The information aggregation model is standard, but could its assumptions be biased against the specific advantages of delegation?

[51]See the regressions in the online Appendix, where we also plot distributions of errors in beliefs.

[52]The average realized accuracy of the top quintile of a bloc with 0.05 coherence is 74%, and with 0.03 coherence is 73%. Notice, however, that conditioning the status of expert on ex post accuracy is both practically infeasible and logically weak.

9.1 Why is delegation so frequent, and more frequent than abstention?

Why were participants drawn to delegation more than to abstention? The second experiment shows that the effect does not disappear in the absence of precise numerical information. Other factors must play a role. One plausible answer is that abstention, but not delegation, suffers from moral stigma: not taking a position can be seen as cowardly, as opposed to actively delegating to a more knowledgeable expert. Note however two arguments to the contrary. First, the explanation suggests too little abstention, as opposed to our finding of excessive delegation. Second, in our experiments the moral argument is weak: without either costs of voting or costs of information acquisition, abstention does not deliver any private gain. Nor is voting with weak information virtuous: it worsens everybody's prospects.[53]

A different conjecture focuses on low confidence in other non-expert voters. If a participant believes that too many poorly informed non-experts are voting, she will wish to increase the power of the experts and delegate, but, on the contrary, will refrain from abstaining. The logic, whose validity we verified under our experimental parameters, is intuitively plausible and suggests an interesting contrast between delegation and abstention. It is challenged, however, by the feed-back participants receive in Experiment 1: after voting, each participant is informed of how many other votes are delegated under LD (and how many are cast under MVA). Even lacking information about other non-experts' precisions, if voting by others is a concern, such feedback should be salient, and the misperception that too many others are voting should be corrected: the probability of delegating/abstaining should respond to the number of other subjects delegating/abstaining in the previous round. As we show in the online Appendix, we find no such evidence.

We return to our conjecture that under LD subjects focus on the pairwise relationship between the voter and the expert, ignoring equilibrium effects. We can try a different test: if "equilibrium neglect"[54] is the correct hypothesis, the frequency of delegation should be invariant to the size of the group because the group is ignored; the frequency of abstention could instead respond to such a size because the group is not ignored. Regressions reported in the online appendix are in line with the prediction.

[53]In an environment similar to ours, Morton and Tyran (2011) find excessive abstention. Mengel and Rivas (2017) introduce asymmetric priors and in this complex environment find too little abstention, but also repeated evidence of random behavior.

[54]We define "equilibrium neglect" as neglecting the number of other votes delegated to the experts, and hence the loss of independent information. The term "correlation neglect" is misleading in this instance because once a subject recognizes that an expert casts multiple votes, the realization that all are cast according to a single signal is immediate.

9.2 Are the advantages of LD undermined by the model's simplifying assumptions?

In both of our experiments, LD underperforms both MV and MVA. But is LD handicapped by the maintained assumptions of our model? We discuss here very briefly three possible extensions: (1) introducing private values; (2) making information costly; (3) allowing for correlation in signals. How would such extensions affect the comparative performance of delegation and abstention? We study a simple example and find that in all three cases the effect is ambiguous. There are two non-experts and one expert; the expert cannot delegate or abstain, and all non-experts have equal precision of information $q \in (1/2, p)$. We describe the results informally; the model is solved in the online Appendix.

Consider first introducing private values. Suppose that voters can be of two types, A and B; type A derives utility 1 from matching the decision to the state (and 0 from mismatching it); type B derives utility 1 from mismatching the decision to the state (and 0 from matching it). The expert's type is publicly known, while non-experts' types are private information, but it is known that each non-expert's type is drawn independently and can be either A or B with equal probability. Delegation to the expert can be valuable not only because of the expert's superior information, but also because of the alignment of private interests.[55] How does LD compare to MVA in such a model?

Under both LD and MVA, there exists an equilibrium in which a non-expert delegates/abstains if she shares the expert's preferences, and votes otherwise.[56] An individual who shares the expert's preferences prefers LD to MVA because she can make the expert pivotal by delegating; but an individual whose preferences differ from the expert's prefers MVA for exactly the same reason: she retains some decision power even if the other non-expert shares the expert's type. In this simple model, non-experts' ex ante expected utility, before the realization of preferences, is identical under the two voting systems. Clearly the example is special. But the comparison between delegation and abstention may well remain ambiguous more broadly: abstention is a supple means of giving some weight to both aligned and misaligned preferences. Notice also that multiple experts need not solve the problem: with opposite private preferences, experts will be more likely to disagree, reducing the value of their expertise.

Similar considerations are relevant if we relax a different simplifying assumption and allow for the possibility that better information can be acquired at a cost. Consider our

[55]In contrast to the seminal model of partisans and independent voters in Feddersen and Pesendorfer (1996), our model preserves uncertainty of the outcome under knowledge of the state of the world.

[56]We focus on symmetric pure equilibria with sincere voting. The equilibrium we describe exists for all q values under LD, and for q not too high under MVA. See the online Appendix for a full discussion.

simple example, brought back to common interest: everyone derives utility 1 from matching the decision to the state (and 0 otherwise). Now, knowing q but before receiving a signal, each of the two non-experts decides independently whether or not to invest in acquiring better information. Investment in information costs c and yields an informative signal of precision p, as precise as the expert's signal. Thus delegation to the expert gains the added value of saving on information costs. A voter can spare the cost of acquiring information by abstaining, but abstention does not provide vicarious access to better information. Which one of the two systems, LD or MVA, leads to better decision-making?[57]

It is not difficult to verify that if the parameters $\{p, q, c\}$ are such to support investment in information under LD, then they also support such investment under MVA. However, the reverse is not true: there are parameter values for which non-experts invest in information with positive probability under MVA, but not under LD. As a result, there are values of $\{p, q, c\}$ such that the probability of reaching the correct decision is strictly higher under MVA than under LD, while the reverse is not true. Under LD, the possibility to free-ride on the expert's superior information blunts the incentive to invest in costly information, reducing the aggregate information in the system, relative to MVA, and thus the probability of a correct decision. There are savings in information costs, but more faulty decisions. Examples show that the latter can dominate the former, leading to lower ex ante utility under LD.

We can argue along similar lines about the effect of introducing correlation in the voters' signals. We build again on our example, with pure common interest and without the option of information acquisition. Now however we suppose that all signals are (conditionally) independent with probability $\alpha < 1$; with probability $1-\alpha$, the signals of the non-experts are fully correlated and only the expert's signal is (conditionally) independent. The probability α is commonly known, but non-experts do not know whether their signals are correlated. The expert here is favored not only by higher precision but also by the possibility of an independent signal.[58] Does LD then outperform MVA? The answer remains ambiguous. For any $\{p, q, \alpha\}$, both LD and MVA support an equilibrium where only the expert votes–both non-experts delegate (under LD) or abstain (under MVA). Both LD and MVA also support, for high enough values of q, an equilibrium where all vote. From a welfare perspective, both

[57]The impact of different voting procedures on the acquisition of costly information has been studied in a number of influential papers. Among others, see Persico (2004) and Martinelli (2006) for theory; Grosser and Sebauer (2016), Elbittar et al. (2017), Bhattacharya et al. (2017), and Mechtenberg and Tyran (2019), for experiments. To our knowledge, none has studied delegation, either alone or in comparison to abstention.

[58]Note that allowing the expert's signal to be fully correlated with one or both of the non-experts' signals does not affect the results. The reason is that it only adds the same probability of reaching the correct decision whenever the expert's signal is not independent, and because the expert always votes, it does so equally in all scenarios, canceling out of expected utility comparisons.

LD and MVA can support, at any value of $\{q, p, \alpha\}$, the best equilibrium at those parameter values. In this case, too, we cannot conclude that the model's original assumptions are stacked against delegation.

10 Conclusions

Liquid Democracy is a computer-mediated voting system such that all decisions are subject to popular referendum but voters can delegate their votes freely. The option of delegation to better informed experts seems intuitively valuable, and indeed, theory shows that if experts are correctly identified, delegation improves the chances of reaching the correct decision. However, delegation must be used sparingly because, by reducing the number of independent voices, it also reduces the aggregate amount of information expressed by the electorate. This paper reports the results of two very different experiments that measure participants' propensity to delegate their vote and compare the decisions of the group when delegation is possible, when abstention is possible, and under universal majority voting. In line with Condorcet's message, we find that in both experiments universal majority voting leads to the highest frequency of correct outcomes, even in small groups; abstention is closely comparable, but delegation is inferior to both, even in our simplified world where experts are indeed better informed. The weak performance of Liquid Democracy in our experiments has two causes. First, in both experiments, participants delegate with very high frequency, much more frequently than they abstain and two to three times more frequently than optimal in the experiment for which we have precise predictions. Second, the quality of decision-making under delegation depends crucially on the accuracy of the experts. Small downward deviations, resulting from the stochastic realization of accuracies, exact a high cost and do so especially when matched to high rates of delegation. Abstention, in practice a form of delegation to all those who cast their votes, is a more robust system, exactly because it aggregates the information of more voters.

Our results are obtained in experimental environments where experts are objectively better informed than non-experts. In the field, the identification of the experts is likely to be much noisier and contested. Liquid Democracy should be tested outside the lab, but our tentative conclusion is that, on informational grounds alone, the arguments in favor of Liquid Democracy should be considered with caution. Other rationales for Liquid Democracy should be given weight and studied—the advantages, and weaknesses, of direct democracy; the heightened sense of responsibility of an empowered electorate in this more participatory form of democracy.

The paper also makes a methodological contribution. We match a canonical, fully controlled lab experiment on voting with a perceptual experiment where information is ambiguous: participants do not know the probabilities with which either their own or the experts' information is correct. We use such a design because there are plausible concerns that the precise mathematical framing of the canonical experiment may affect the results. In addition, the ambiguity of the information in the perceptual experiment seems closer to realistic conditions of voting for political decisions. We find that the results of the first experiment replicate closely in the second and consider the robustness of the conclusions a central contribution of our study. Beyond the specific results, one of our goals is to stress that perceptual experiments are a very useful tool for the study of group decision-making under ambiguity, especially in conjunction with the precise and explicit information structure of traditional lab experiments. Social psychologists have been studying them for decades; it is time economists interested in social choice added them to their tool box as well.

References

[1] Armstrong, Ben and Kate Larson, 2021, "On the Limited Applicability of Liquid Democracy". *3rd Games, Agents, and Incentives Workshop (GAIW2021), 20th International Conference on Autonomous Agents and Multiagent Systems*, London, UK, May 2021, IFAAMAS, 7 pages.

[2] Austen-Smith, David, and Jeffrey Banks, 1996, "Information aggregation, rationality, and the Condorcet Jury theorem", *American Political Science Review*, 90:34-45.

[3] Bar-Isaac, Heski and Joel Shapiro, 2020, "Blockholder voting", *Journal of Financial Economics* 136:695-717.

[4] Bhattacharya, Sourav, John Duffy, and Sun-Tak Kim, 2017, "Voting with endogenous information acquisition: Experimental evidence", *Games and Economic Behavior,* 102:316-338.

[5] Battaglini, Marco, Rebecca Morton, and Thomas Palfrey, 2010, "The swing voter's curse in the laboratory", *Review of Economic Studies* 77:61-89.

[6] Bahrami, B., K. Olsen, D. Bang, A Roepstorff, G Rees and C. Frith, 2012, "What failure in collective decision-making tells us about metacognition", *Philosophical Transactions of the Royal Society* 367:1350-1365.

[7] Blum, Christian and Christina Zuber, 2016, "Liquid Democracy: Potentials, Problems, and Perspectives", *Journal of Political Philosophy* 24:162–182.

[8] Brill Markus and Nimrod Talmon, 2018, "Pairwise liquid democracy", *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18)*, 137–143.

[9] Britten K.H., M. Shadlen, W. Newsome, and J.A. Movshon, 1992, "The analysis of visual motion: a comparison of neuronal and psychophysical performance", *Journal of Neuroscience* 12: 4745–4765.

[10] Buechel, B. and L. Mechtenberg, 2019, "The swing voter's curse in social networks", *Games and Economic Behavior* 118: 241–268.

[11] Buechel, Berno, Lydia Mechtenberg, and Alexander F. Wagner, 2023, "When do proxy advisors improve corporate decisions?", SSRN: https://ssrn.com/abstract=4123989.

[12] Caragiannis, Ioannis. and Evi Michas, 2019, "A Contribution to the Critique of Liquid Democracy", *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI-19),* 116-122.

[13] Christoff Zoé and Davide Grossi, 2017, "Binary voting with delegable proxy: an analysis of liquid democracy", *Proceedings of the Sixteenth Conference on Theoretical Aspects of Rational Knowledge (TARK-17)*,134–150.

[14] Condorcet, N. C., 1785, *Essai sur I'Application de 1'Analyse à la Probabilité des Decisions Rendues à la Pluralité des Voix*, Paris: De l'Imprimerie Royale.

[15] van de Grind, W.A., A.J. van Doorn, and J.J. Koenderink, 1983, "Detection of coherent movement in peripherally viewed random-dot patterns," *Journal of the Optical Society of America* 73: 1674-1683.

[16] Dhillon, Amrita, Grammateia Kotsialou, Dilip Ravindran and Dimitrios Xefteris, 2023, "Information Aggregation with Delegation of Votes", *arXiv:2306.03960*.

[17] Dodgson, Charles, 1884, *The Principles of Parliamentary Representation*, London: Harrison and Sons.

[18] Elbittar, A., A. Gomberg, C. Martinelli, and T. R. Palfrey, 2020, "Ignorance and bias in collective decisions", *Journal of Economic Behavior and Organization* 174: 332-359.

[19] Ellis, Andrew, 2016, "Condorcet meets Ellsberg", *Theoretical Economics*, 11:866-895.

[20] Esponda, Ignacio, Emanuel Vespa and Sevgi Yuksel, 2023, "Mental Models and Learning: The Case of Base-Rate Neglect", https://faculty.econ.ucsb.edu/˜sevgi/EVY.pdf.

[21] Fabrizi, Simona, Steffen Lippert, Addison Pan and Matthew Ryan, 2022, "A Theory of Unanimous Jury Voting with an Ambiguous Likelihood", *Theory and Decision*, 93:399-425.

[22] Fabrizi, Simona, Steffen Lippert, Addison Pan and Matthew Ryan, 2021, "Unanimity under ambiguity", AUT Economics W.P. No. 2021/07.

[23] Feddersen, Timothy and Wolfgang Pesendorfer, 1996, "The swing voter's curse", *American Economic Review* 86:408–424.

[24] Feddersen, Timothy and Wolfgang Pesendorfer, 1997, "Voting Behavior and Information Aggregation in Elections With Private Information", *Econometrica* 65:1029-1058.

[25] Goeree Jacob and Leeat Yariv, 2011, "An Experimental Study of Collective Deliberation", *Econometrica*, 79:893–921.

[26] Gölz, Paul, Anson Kahng, Simon Mackenzie, and Ariel D. Procaccia, 2018, "The Fluid Mechanics of Liquid Democracy", *arXiv:1808.01906v1* [cs.GT].

[27] Green-Armytage, James, 2015, "Direct voting and proxy voting", *Constitutional Political Economy* 26:190–220.

[28] Grofman, Bernard, Guillermo Owen and Scott Feld. 1983. "Thirteen theorems in search of the truth", *Theory and Decision* 15:261-278.

[29] Grofman, Bernard, Guillermo Owen and Scott Feld. 1982. "Evaluating the competence of experts, pooling individual judgments into a collective choice, and delegating decision responsibility to subgroups". In Felix Geyer and Hans van der Zouwen (Eds.), *Dependence and Inequality.* NY: Pergamon Press, 221-238.

[30] Grosser, Jens and Michael Seebauer, 2016, "The curse of uninformed voting: An experimental study", *Games and Economic Behavior*, 97: 205-226.

[31] Guarnaschelli, Serena, Richard McKelvey and Thomas Palfrey, 2000, "An Experimental Study of Jury Decision Rules", *The American Political Science Review* 94:407-423.

[32] Halpern, D., J. Halpern, A. Jadbabaie, E. Mossel, A. Procaccia and M. Revel, 2021, "In Defense of Liquid Democracy", *arXiv:2017.11868v2.*

[33] Hamman, John, Roberto Weber, and Jonathan Woon, 2011, "An Experimental Investigation of Electoral Delegation and the Provision of Public Goods", *American Journal of Political Science* 55:738-752.

[34] Harding, Jacqueline, 2022, "Proxy Selection in Transitive Proxy Voting", *Social Choice and Welfare*, 58:69–99.

[35] Hardt, Steve and Lia Lopes, 2015, "Google Votes: A Liquid Democracy Experiment on a Corporate Social Network", *Technical Disclosure Commons*, http://www.tdcommons.org/dpubs_series/79

[36] Heer, Jeffrey and Michael Bostock, 2010, "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design", *CHI 2010*: April 10–15, 2010, Atlanta, Georgia, USA: 203-212.

[37] Kahng, Anson, Simon MacKenzie and Ariel Procaccia, 2018, "Liquid Democracy: An Algorithmic Perspective", *Proceedings of the 32nd AAAI Conference in Artificial Intelligence (AAAI)*, 1095-1102.

[38] Malenko, Andrey and Nadya Malenko, 2019, "Proxy Advisory Firms: The Economics of Selling Information to Voters", *Journal of Finance*, 74:2441-2490.

[39] Margolis, Howard, 1976, "A Note on Incompetence", *Public Choice* 26:119-127.

[40] Margolis, Howard, 1987, *Patterns, Thinking, and Cognition*, The University of Chicago Press: Chicago, IL.

[41] Martinelli, César, 2006, "Would rational voters acquire costly information?", *Journal of Economic Theory* 129: 225-251.

[42] McMurray, Joseph, 2013, "Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses*", Review of Economic Studies* 80:277-312.

[43] McLennan, Andrew, 1998, "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents", *American Political Science Review* 92: 413-418.

[44] Mechtenberg,Lydia and Jean-Robert Tyran, 2019, "Voter Motivation and the Quality of Democratic Choice", *Games and Economic Behavior*, 116: 241-259.

[45] Mengel, Friederike and Javier Rivas, 2017, "Common value elections with private information and informative priors: theory and experiments", *Games and Economic Behavior* 104:190-221.

[46] Morton, Rebecca and Jean-Robert Tyran, 2011, “Let the experts decide? Asymmetric information, abstention, and coordination in standing committees”, *Games and Economic Behavior* 72:485–509.

[47] Nitzan, Shmuel and Jacob Paroush, 1982, “Optimal decision rules in uncertain dichotomous choice situations”, *International Economic Review* 23:289-297.

[48] Newsome William, K.H. Britten, and J. Anthony Movshon, 1989, “Neuronal correlates of a perceptual decision”, *Nature* 341:52–54.

[49] Persico, Nicola, 2004, “Committee Design with Endogenous Information”, *Review of Economic Studies*, 71: 165-191.

[50] Pilly, Praveen and Aaron R. Seitz, 2009, “What a difference a parameter makes: A psychophysical comparison of random dot motion algorithms“, *Vision Research* 49:1599-1612.

[51] Rajananda, S., Lau, H. & Odegaard, B., 2018, “A Random-Dot Kinematogram for Web-Based Vision Research”, *Journal of Open Research Software* 6(1), p.6.

[52] Ravindran, Dilip, 2021, “Liquid Democracy”, in *Essays in information and behavioral economics*, PhD Dissertation, Columbia University, New York, NY.

[53] Revel, M., D. Halpern, A Berinsky and A. Jadbabaie, 2022, “Liquid Democracy in Practice: An Empirical Analysis of its Epistemic Performance”, https://danielhalpern.com/files/liquid-in-practice.pdf.

[54] Roitman, Jamie and Michael Shadlen, 2002, “Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task”, *Journal of Neuroscience* 22: 9475–9489.

[55] Ryan, Matthew, 2021, “Feddersen and Pesendorfer meet Ellsberg”, *Theory and Decision*, 90:543-577.

[56] Schütz, Alexander, Doris Braun, J. Anthony Movshon, and Karl Gegenfurtner, 2010, “Does the noise matter? Effects of different kinematogram types on smooth pursuit eye movements and perception”, *Journal of Vision*, 10(13):26..

[57] Shapley, Lloyd and Bernard Grofman, 1984, “Optimizing group judgmental accuracy in the presence of interdependencies”, *Public Choice* 43:329-343.

[58] Silver, Ike, Barbara Mellers and Philip Tetlock, 2021, "Wise teamwork: collective confidence calibration predicts the effectiveness of group discussion", *Journal of Experimental Social Psychology*, 96: Case Report 104157.

[59] Sorkin, Robert, Ryan West and Donald Robinson, 1998, "Group Performance Depends on the Majority Rule", *Psychological Science* 9:456-463.

[60] Sorkin, Robert, Christopher Hays, and Ryan West, 2001, "Signal-Detection Analysis of Group Decision Making", *Psychological Review* 108:183-203.

[61] Wit, Jőrgen, 1998, "Rational Choice and the Condorcet Jury Theorem", *Games and Economic Behavior* 22: 364-376.

[62] Woods Andy et al., 2015, "Conducting perception research over the internet: a tutorial review", *PeerJ 3:e1058*; DOI 10.7717/peerj.1058

A Appendix

A.1 Theoretical results

Theorem. *Suppose $\pi = Pr(\omega_1) = 1/2$. Then for any F and for any N and K odd and finite, there exists an equilibrium with delegation that strictly improves over MV.*

The proof proceeds in four steps. We begin with Lemma 1:

Lemma 1. *Consider a profile of strategies $\Sigma = \{\Sigma_{ne}, \Sigma_e\}$ symmetric for voters of each class, and such that votes are cast according to signal and delegation decisions are symmetric with respect to signals' realization. There exists a profile $\Sigma^* = \{\Sigma^*_{ne}, \Sigma^*_e\} \in \{\Sigma_{ne}, \Sigma_e\}$ such that $EU(\Sigma^*) \geq EU(\Sigma)$ for all $\Sigma \neq \Sigma^*$.*

Proof. If all votes cast are cast according to signal, strategic choices are limited to delegation. Because we are focusing on delegation strategies that are symmetric with respect to signals' realizations, such strategies depend on signals' precisions only. Consider first non-expert i. Keeping in mind that all q_i's are independent draws from $F(q)$, expected utility conditional on delegation—$EUD_i^{(e)}$ for delegation to an expert, or $EUD_i^{(ne)}$ for delegation to a non-expert—does not depend on q_i (because i would not be voting). On the other hand, i's expected utility when not delegating, $EUND$, must be weakly increasing in q_i (because the probability of reaching the correct outcome must be weakly increasing in q_i), and strictly increasing if i's probability of being pivotal is positive. It follows that in any equilibrium there must exist a $\widetilde{q} \in [\underline{q}, \overline{q}]$ such that individual i votes (according to signal) if $q_i \geq \widetilde{q}$, and delegates otherwise. If $q_i < \widetilde{q}$, non-expert i will delegate and may delegate to either an expert (with probability $\delta^{(e)}$) or to a non-expert (with probability $\delta^{(ne)}$). These probabilities may correspond to realizations of q_i in different subintervals, but the precise characterization of such sub-intervals is irrelevant because upon delegation q_i has no effect on expected utility. Hence symmetric non-expert i's delegation strategies are summarized by a set of three numbers $\Sigma_{ne} = \{\delta^{(e)}, \delta^{(ne)}, \widetilde{q}\}$, with $\delta^{(e)} + \delta^{(ne)} = F(\widetilde{q}) \in [0, 1]$ and $\widetilde{q} \in [\underline{q}, \overline{q}]$. In the case of expert individuals, precision is fixed at p. Hence, denoting by $\xi^{(e)}(\xi^{(ne)})$ an expert's probability of delegating to another expert (non-expert), strategies are given by $\Sigma_e = \{\xi^{(e)}, \xi^{(ne)}\}$, with $\xi^{(e)} + \xi^{(ne)} \in [0, 1]$. Although the set of strategies Σ is infinite, it is compact. Because EU is a continuous function of Σ, we can then apply Weierstrass's Theorem, and the conclusion follows: there exists a profile $\Sigma^* = \{\Sigma^*_{ne}, \Sigma^*_e\} \in \{\Sigma_{ne}, \Sigma_e\}$ such that $EU(\Sigma^*) \geq EU(\Sigma)$ for all $\Sigma \neq \Sigma^*$. □

Lemma 2. *The profile of strategies Σ^* is an equilibrium.*

Proof. The result is established in two steps, both derived from McLennan (1998). First, restrict attention to profiles of semi-symmetric strategies with sincere voting and symmetric delegation with respect to signals' realization. Within such profiles, Σ^* is maximal (i.e. $\Sigma^* = argmax_{\Sigma} EU(\Sigma)$) by construction. But in this pure common interest game, it then follows that Σ^* must be an equilibrium (McLennan, Theorem 1). Second, the environment is fully symmetric for each class of individuals, experts and non-experts, and at any such type-symmetric strategy profile, with sincere voting and symmetric delegation, no voter wants to deviate to asymmetric delegation or an insincere voting strategy. A priori the two states are equally likely, and the precision of the signals does not depend on the state. All voters have identical preferences and are endowed with a single vote. All experts have equal precision and equal probability of receiving any delegated vote, and thus, for any delegation strategy by non-experts, each expert's vote has equal expected weight on the final decision. Non-experts will have heterogeneous realized precisions, and the equilibrium action will depend on individual precision, but each precision q_i is an independent draw from the same distribution F. Hence any permutation of realized precisions to different non-expert voters is assigned equal probability, and each voter holds equal beliefs about the others' precisions. For each type of voter, these symmetry conditions satisfy the requirements of McLennan's Theorem 2: if Σ^* is maximal with respect to semi-symmetric strategies with sincere voting and symmetric delegation, then it is an equilibrium over all profiles of strategies. Note that asymmetric equilibria may exist, and be superior to Σ^*. □

Lemma 3. *Sincere majority voting is not an equilibrium.*

Proof. Sincere majority voting is feasible within the set of strategies Σ (semi-symmetric strategies with delegation symmetric across signals and sincere voting). It corresponds to $\widetilde{q} = \underline{q}$ (and thus $\delta^{(e)} = \delta^{(ne)} = 0$), and $\xi^{(e)} = \xi^{(ne)} = 0$. Can such set of strategies be an equilibrium? If the answer is negative, then we know, by the previous results, that there must exist an equilibrium of the LD voting game that strictly dominates MV and involves delegation.

Consider the perspective of non-expert voter i, with q_i in the neighborhood of $\underline{q}$. Suppose no-one else delegates. We show in what follows that i's best response is to delegate his vote. Note first that if no-one delegates, all non-i voters cast a single vote and have equal weight on the group decision. Hence if i delegates, it is optimal to delegate the vote to an expert, with precision $p \geq q_j$ for all j.

We need to calculate i's interim expected utility from non-delegating ($EUND(q_i)$) or delegating (EUD). The expressions are somewhat cumbersome but conceptually straightforward. We find:

$$EUND(q_i) = \sum_{c_n=0}^{M-1} \binom{M-1}{c_n} \mu^{c_n}(1-\mu)^{M-1-c_n} \times$$
$$\times \left[\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e}(1-p)^{K-c_e} \left(q_i I_{c_n+c_e+1>\frac{(M+K)}{2}} + (1-q_i) I_{c_n+c_e>\frac{(M+K)}{2}} \right) \right]$$

$$EUD = \sum_{c_n=0}^{M-1} \binom{M-1}{c_n} \mu^{c_n}(1-\mu)^{M-1-c_n} \times$$
$$\times \left[\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e}(1-p)^{K-c_e} \left(\left(\frac{c_e}{K}\right) I_{c_n+c_e+1>\frac{(M+K)}{2}} + \left(\frac{K-c_e}{K}\right) I_{c_n+c_e>\frac{(M+K)}{2}} \right) \right],$$

where M is the number of non-experts, $c_n(c_e)$ indexes the number of non-experts other than i (experts) whose signals are correct, μ is the expected precision of non-experts who choose to vote, and thus in this conjectured scenario, $\mu = \int_{\underline{q}}^{\overline{q}} q dF(q)$, and I_C is an indicator function that takes value 1 if condition C is satisfied and 0 otherwise. For each realized c_n and c_e, i's expected utility always equals 1 if $(c_n + c_e) > (M + K)/2$, i.e. if the other voters with correct signals constitute a majority of the electorate. For i, the choice to delegate or not matters when $(c_n + c_e)$ falls short of the majority by one vote. In such a case, $EUND(q_i)$ equals 1 if i's own signal is correct (with probability q_i) and zero otherwise; EUD equals 1 if i's vote is delegated to an expert with a correct signal (with probability c_e/K) and zero otherwise.

Voter i, with q_i in the neighborhood of $\underline{q}$, strictly prefers delegation if it yields higher expected utility, or:

$$lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) < 0$$

Denote by r the number of additional correct votes required to reach a majority, given the votes of the non-experts, excluding i, or $r \equiv (M + K + 1)/2 - c_n$. After some simplifications, we can write:

$$lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) =$$
$$= \sum_{r=1}^{K+1} \binom{M-1}{\frac{M+K+1}{2} - r} \mu^{\frac{M+K+1}{2} - r} (1-\mu)^{\frac{M-(K+3)}{2} + r} \binom{K}{r-1} p^{r-1} (1-p)^{K-(r-1)} \left(\underline{q} - \frac{r-1}{K} \right) \tag{1}$$

Signing this expression is not immediate because the sign depends on the last term. However, the problem is simplified by noticing that:

$$\binom{M-1}{\frac{M+K+1}{2} - r} = \binom{M-1}{\frac{M+K+1}{2} - (K+2-r)}.$$

Equation (1) can then be written as:

$$lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) = \sum_{x=1}^{(K+1)/2} \binom{M-1}{\frac{M+K+1}{2} - x} (\mu(1-\mu))^{\frac{M+K+1}{2} - (K+2-x)} \binom{K}{x} p^x (1-p)^x \times$$
$$\times \left\{ (\mu(1-p))^{K+2-2x} \left[\underline{q} - \frac{x-1}{K} \right] + ((1-\mu)p)^{K+2-2x} \left[\frac{x-1}{K} - (1-\underline{q}) \right] \right\}$$

or, with $\underline{q} = 1 - \underline{q} = 1/2$:

$$lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) = \sum_{x=1}^{(K+1)/2} \binom{M-1}{\frac{M+K+1}{2} - x} (\mu(1-\mu))^{\frac{M+K+1}{2} - (K+2-x)} \binom{K}{x} p^x (1-p)^x \times$$
$$\times \left[\underline{q} - \frac{x-1}{K} \right] \left\{ (\mu(1-p))^{K+2-2x} - ((1-\mu)p)^{K+2-2x} \right\}$$

With $\underline{q} = 1/2$, $\left[\underline{q} - (x-1)/K \right] > 0$ for all $x < (K+2)/2$, and thus for all relevant x values. It follows that:

$$(\mu(1-p))^{K+2-2x} < ((1-\mu)p)^{K+2-2x} \Rightarrow lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) < 0$$

With all $x < (K+2)/2$, the exponent on both sides is positive, and we can compare the roots:

$$\mu(1-p) < (1-\mu)p \Rightarrow lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) < 0,$$

a condition that reduces to:

$$\mu < p$$

and is always satisfied. Hence $lim_{q_i \to \underline{q}} (EUND(q_i) - EUD) < 0$: delegation is the best response. A profile of strategies such that all non-experts cast their vote with probability 1 cannot be an equilibrium.

Assuming $\underline{q} = 1/2$ is necessary to establish the result in its full generality, but is also a natural assumption: $\underline{q}$ cannot be inferior to 1/2, and thus 1/2 is the natural lower boundary of the support of the precision distribution.[59] The probability of realizations near this lower bound needs to be positive, but can be arbitrarily small. □

Having established Lemmas 1, 2 and 3, the theorem follows. There exists an equilibrium in semi-symmetric strategies and sincere voting that dominates sincere majority voting. Such an equilibrium must then include a strictly positive probability of delegation.

Proposition 1. *Suppose $\pi = Pr(\omega_1) = 1/2$ and $K = 1$. Then for any N odd and finite, there exists an equilibrium such that: (i) the expert never delegates her vote and always votes according to signal; (ii) there exists a threshold $\widetilde{q}(N) \in (\underline{q}, \overline{q})$ such that non-expert i delegates her vote to the expert if $q_i < \widetilde{q}$ and votes according to signal otherwise. Such an equilibrium strictly improves over MV and is maximal among sincere semi-symmetric equilibria where the expert never delegates and non-experts delegate to the expert only.*

We prove the proposition in two lemmas. Recall that we focus on sincere semi-symmetric strategies invariant to signal content. The first lemma shows that ex ante expected utility, defined over such strategies and allowing delegation to the expert only, has a maximum at some probability of delegation strictly between 0 and 1. By the logic of common interest games, the lemma implies that when strategies are restricted as stated, there exists $\widetilde{q} \in (\underline{q}, \overline{q})$ such that delegating to the expert if $q_i < \widetilde{q}$ and voting according to signal otherwise constitute a set of mutual best responses. The second lemma shows that the strategies described, including the directions of delegation and the threshold of precision $\widetilde{q}$, are an equilibrium within the set of all possible strategies. Hence the equilibrium exists and maximizes expected utility over sincere semi-symmetric strategies with the described directions of delegation.

[59]As noted earlier, with binary states, a signal correct with probability inferior to 1/2 is equivalent to the opposite signal being correct with probability higher than 1/2.

Lemma 4. *Suppose $\pi = Pr(\omega_1) = 1/2$, $K = 1$, players are restricted to sincere semi-symmetric strategies invariant to signal content, non-experts can only delegate to the expert, and the expert does not delegate. Then ex ante expected utility EU is maximized at some $\widetilde{q}$ strictly between 0 and 1.*

Proof. The set of strategies over which expected utility is maximized is restricted to a threshold $\widetilde{q} \in [\underline{q}, \overline{q}]$ such that non-experts delegate to the expert if $q_i < \widetilde{q}$ and vote according to signal otherwise. Ex ante expected utility is thus $EU(\widetilde{q})$. Because the set is compact and and expected utility is continuous over $\widetilde{q}$, a maximum exists. By the argument in the proof of the theorem, $\widetilde{q} = \underline{q}$, or no delegation, cannot be a maximum (because $\widetilde{q} = \underline{q}$ is not an equilibrium over the set of strategies considered here). The upper boundary of the strategy set, $\widetilde{q} = \overline{q}$, however is an equilibrium: at $\widetilde{q} = \overline{q}$ all non-experts delegate to the expert; the expert is dictator, and no deviating non-expert can be pivotal. Hence there can be no strict gain from deviation. To prove the proposition, we need to show that $EU(\widetilde{q})$ is decreasing in the neighborhood of $\overline{q}$ and thus the maximal $\widetilde{q}$ is interior.

Because the game is common interest, $EU(\widetilde{q})$ equals the ex ante expected utility of a non-expert i when all non-experts, including i, adopt threshold $\widetilde{q}$. Consider $EU(\widetilde{q}) - EU(\overline{q})$ as $\widetilde{q} \to \overline{q}$. The difference is not 0 only if there exist realizations of delegation decisions and signals at which a non-expert vote can be pivotal, hence only if the expert has been delegated fewer than $\frac{M}{2}$ votes. Let $\mu(\widetilde{q}) = \mathbb{E}[q_i | q_i \geq \widetilde{q}]$ be the conditional expected precision of a non-expert i with precision above the threshold. Consider the ratio of probabilities of the event (i is pivotal and the expert was delegated $z < \frac{M}{2} - 1$ votes) to the event (i is pivotal and the expert was delegated $\frac{M}{2} - 1$ votes):

$$\frac{Pr(\text{piv}_i | z \text{ votes delegated}) Pr(z \text{ votes delegated})}{Pr(\text{piv}_i | \frac{M}{2} - 1 \text{ votes delegated}) Pr(\frac{M}{2} - 1 \text{ votes delegated})} =$$

$$\frac{\binom{M-1-z}{\frac{M}{2}} [p\mu(\widetilde{q})^{\frac{M}{2}-z-1}(1-\mu(\widetilde{q}))^{\frac{M}{2}} + (1-p)(1-\mu(\widetilde{q}))^{\frac{M}{2}-z-1}\mu(\widetilde{q})^{\frac{M}{2}}] \binom{M-1}{z} F(\widetilde{q})^z (1-F(\widetilde{q}))^{M-z-1}}{\binom{\frac{M}{2}}{\frac{M}{2}} [p(1-\mu(\widetilde{q}))^{\frac{M}{2}} + (1-p)\mu(\widetilde{q})^{\frac{M}{2}}] \binom{M-1}{\frac{M}{2}-1} F(\widetilde{q})^{\frac{M}{2}-1} (1-F(\widetilde{q}))^{M-(M/2-1)-1}} =$$

$$\binom{\frac{M}{2}-1}{z} \frac{[p\mu(\widetilde{q})^{\frac{M}{2}-z-1}(1-\mu(\widetilde{q}))^{\frac{M}{2}} + (1-p)(1-\mu(\widetilde{q}))^{\frac{M}{2}-z-1}\mu(\widetilde{q})^{\frac{M}{2}}]}{p(1-\mu(\widetilde{q}))^{\frac{M}{2}} + (1-p)\mu(\widetilde{q})^{\frac{M}{2}}} \; \frac{F(\widetilde{q})(1-F(\widetilde{q}))^{M-z-1}}{F(\widetilde{q})^{\frac{M}{2}-1}(1-F(\widetilde{q}))^{\frac{M}{2}}}$$

As $\widetilde{q} \to \overline{q}$, $\mu(\widetilde{q}) \to \overline{q}$ and $F(\widetilde{q}) \to 1$. Hence the limit of the first two terms above is a strictly positive number, but the numerator and denominator of the last ratio both go to 0. Using L'Hôpital's rule, we find

$$\lim_{F(\widetilde{q}) \to 1} \frac{F(\widetilde{q})(1-F(\widetilde{q}))^{M-z-1}}{F(\widetilde{q})^{\frac{M}{2}-1}(1-F(\widetilde{q}))^{\frac{M}{2}}} = 0$$

As $\widetilde{q} \to \overline{q}$, conditional on i being pivotal, the probability that the expert was delegated $\frac{M}{2} - 1$ votes, rather than any fewer, becomes arbitrarily high, while i's ex-interim expected utility difference between delegating and keeping her vote remains bounded away from 0 as $\widetilde{q} \to \overline{q}$. Thus i conditions the decision to delegate on such an event. But, conditional on the expert being delegated $\frac{M}{2} - 1$ votes and i being pivotal, it must be that $\sigma_e \neq \sigma_j$ for all non-experts $j \neq i$ who have not delegated: the expert disagrees with $\frac{M}{2}$ non-experts other than i who all have precisions (weakly) greater than $\widetilde{q}$. As $\widetilde{q} \to \overline{q}$, it then follows that not delegating is superior to delegating at $\widetilde{q}$,[60] and thus $EU(\widetilde{q}) > EUD = EU(\overline{q})$. Or $EU(\widetilde{q}) - EU(\overline{q}) > 0$ as $\widetilde{q} \to \overline{q}$. Thus $argmax_{\widetilde{q}} EU(\widetilde{q}) \in (\underline{q}, \overline{q})$: $EU(\widetilde{q})$ reaches a maximum at a strictly interior probability of delegation. □

By McLennan's argument and the proof of the Theorem, Lemma 4 implies that, when restricting attention to semi-symmetric strategies invariant to signals' content and with the specified directions of delegation, there exists equilibrium $\widetilde{q} \in (\underline{q}, \overline{q})$ such that every non-expert i delegates to the expert if $q_i < \widetilde{q}$, and votes sincerely otherwise. Lemma 5 shows that such strategies are equilibrium strategies in the space of all strategies.

Lemma 5. *Suppose $\pi = Pr(\omega_1) = 1/2$ and $K = 1$. Then there exists a threshold strategy $\widetilde{q}$ such that: (i) If all non-experts adopt strategy $\widetilde{q}$ and only delegate to the expert, and if all votes cast by non-experts are cast sincerely, then it is optimal for the expert never to delegate and to vote sincerely. (ii) Consider non-expert i. If the expert never delegates and always votes sincerely; if every other non-expert $j \neq i$ delegates to the expert if $q_j < \widetilde{q}$ and votes sincerely otherwise, then it is optimal for i to delegate to the expert if and only if $q_i < \widetilde{q}$, and vote sincerely otherwise. Hence the strategies are mutual best responses.*

Proof. Given the uniform prior, the symmetry of precisions across signals, and the signals' conditional independence, it is well-known that sincere voting and delegation strategies symmetric across signals' content are mutual best responses. The contribution of the lemma is in proving that, given such strategies and optimal symmetric non-experts' threshold strategies, the directions of delegation—the expert never delegating and non-experts delegating to the expert only—are mutual best responses. We begin by proving claim (i): delegation from the expert to a non-expert cannot be optimal. First consider expert e delegating to some non-expert j when j does not delegate to e (and thus there is no cycle). Expected utility when all M non-experts use threshold $\widetilde{q}$ is higher than expected utility when $M - 1$ non-experts use

[60] The claim holds if $lim_{\widetilde{q} \to \overline{q}} \left(EU_{ND}(q_i)\Big|_{z=M/2-1} - EU_D\Big|_{z=M/2-1} \right) > 0$. It is not difficult to verify that the condition corresponds to $(1-p)q_i p^{M/2-1} > p(1-q_i)(1-p)^{M/2-1}$, which is satisfied for all p and q_i in $(0.5, 1)$.

cutoff $\widetilde{q}$ and one non-expert i delegates to the expert for all q_i (because i delegating when $q_i > \widetilde{q}$ strictly decreases expected utility). In turn, expected utility in this latter case is higher than expected utility from the same actions if the expert's precision were drawn from $[\widetilde{q}, \overline{q}]$ according to distribution F, rather than being p. But the expected utility from this last scenario is identical to the expected utility from e delegating to j: $M-1$ non-experts delegate using cutoff $\widetilde{q}$, but all their delegated votes are turned over to non-expert j, with precision randomly drawn from $[\widetilde{q}, \overline{q}]$, and e, who always delegates regardless of precision, is the analogue of voter i in the constructed scenario. Now suppose that when e delegates to j, j also delegates to e. This creates a cycle, and by the rules of the game, e's vote is reassigned randomly to a different non-expert j'. If j' does not delegate, the argument above applies. If j' delegates to e as well, another non-expert j'' is randomly chosen to receive e's vote, and so forth until a non-expert who does not delegate is chosen, and the argument above applies. If all non-experts delegate to e, then, again by the rules of the game, the decision is taken by a coin toss. But this leads to $EU = 1/2$, inferior to expected utility when e does not delegate. Hence claim (i) applies.

Consider now claim (ii). Given threshold $\widetilde{q}$, consider the difference in expected utility for a non-expert i between delegating to expert e or instead delegating to some non-expert j who has not herself delegated her vote to e (the choice would otherwise be irrelevant). Expert e has precision p; non-expert j has expected precision $\mu(\widetilde{q}) \equiv \mathbb{E}_F[q_j|q_j > \widetilde{q}] < p$. Voter i's expected utility from the two forms of delegation can differ only if e and j's signals differ ($\sigma_e \neq \sigma_j$), and i's vote is pivotal: for any number of delegated votes $z \leq (\frac{M}{2}-1)$, the expert agrees with $\frac{M}{2}-(z+1)$ non-experts and disagrees with $\frac{M}{2}$ (j included, i not included). The delegation choice is not trivial because, although $\mu(\widetilde{q}) < p$, when i is pivotal and $\sigma_e \neq \sigma_j$, there must be fewer independent signals agreeing with e than with j. Let $piv_i(z)$ be the event corresponding to the set of signal realizations at which i's vote is pivotal, conditional on the expert being delegated z votes, and $pivj_i(z)$ be the same event additionally conditioning on $\sigma_e \neq \sigma_j$. Note that in both events the expert agrees with $\frac{M}{2}-(z+1)$ non-experts and disagrees with $\frac{M}{2}$: the events contain the same information content and $Pr(\sigma_e = \omega|piv_i(z)) = Pr(\sigma_e = \omega|pivj_i(z))$ for all z. Then, noting that the summation below is to $M-2$ to exclude i and j:

$$EUD(i \text{ delegate to } e) - EUD(i \text{ delegate to } j) =$$

$$= \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1-F(\widetilde{q}))^{M-2-z} Pr(pivj_i(z)|z)[Pr(\sigma_e = \omega|pivj_i(z)) - Pr(\sigma_j = \omega|pivj_i(z))]$$

Define $r(z) \equiv Pr(\sigma_e = \omega | pivj_i(z)) = Pr(\sigma_e = \omega | piv_i(z))$. Then, for $z \leq (\frac{M}{2} - 1)$:[61]

$$r(z) = \frac{Pr(pivj_i(z)|\sigma_e = \omega)Pr(\sigma_e = \omega)}{Pr(pivj_i(z))} =$$

$$= \frac{p(\mu(\widetilde{q}))^{\frac{M}{2}-(z+1)}(1-\mu(\widetilde{q}))^{\frac{M}{2}}}{p(\mu(\widetilde{q}))^{\frac{M}{2}-(z+1)}(1-\mu(\widetilde{q}))^{\frac{M}{2}} + (1-p)(1-\mu(\widetilde{q}))^{\frac{M}{2}-(z+1)}(\mu(\widetilde{q}))^{\frac{M}{2}}}$$

As $Pr(\sigma_j = \omega | pivj_i(z)) = 1 - Pr(\sigma_e = \omega | pivj_i(z))$, we can rewrite:

$$EUD(i \text{ delegate to } e) - EUD(i \text{ delegate to } j) =$$

$$= \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1 - F(\widetilde{q}))^{M-2-z} Pr(pivj_i(z)|z)[2r(z) - 1]$$

We can sign this expression by exploiting the equilibrium condition for $\widetilde{q}$. Consider the difference in expected utility between i delegating to e and i voting when her precision is $q_i = \widetilde{q}$:

$$EUD(i \text{ delegate to } e) - EUND(q_i = \widetilde{q}) =$$

$$\sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1 - F(\widetilde{q}))^{M-2-z} Pr(piv_i(z)|z)[Pr(\sigma_e = \omega | piv_i(z)) - \widetilde{q}] =$$

$$= \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1 - F(\widetilde{q}))^{M-2-z} Pr(piv_i(z)|z)[r(z) - \widetilde{q}]$$

Note that $Pr(pivj_i(z)|z) = \frac{\frac{M}{2}}{M-1} Pr(piv_i(z)|z)$ (i.e. j must be part of the $M/2$ non-experts who disagree with e, out of $M-1$ non-experts, ignoring i). For equilibrium $\widetilde{q}$, $EUD(i \text{ delegate to } e) - EUND(q_i = \widetilde{q}) = 0$ which implies:

$$\frac{M-1}{\frac{M}{2}} \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1 - F(\widetilde{q})^{M-2-z} Pr(pivj_i(z)|z) r(z) =$$

$$= \frac{M-1}{\frac{M}{2}} \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1 - F(\widetilde{q}))^{M-2-z} Pr(pivj_i(z)|z) \widetilde{q}$$

[61]Note that $r(z)$ is strictly decreasing in z for all $z \leq (\frac{M}{2} - 1)$.

or:

$$r(z) = \widetilde{q}.$$

Hence:

$$EUD(i \text{ delegate to } e) - EUD(i \text{ delegate to } j) =$$
$$= \sum_{z=0}^{M-2} \binom{M-2}{z} F(\widetilde{q})^z (1-F(\widetilde{q}))^{M-2-z} Pr(pivj_i(z)|z)[2\widetilde{q}-1]$$

But $(2\widetilde{q}-1) > 0$ for all $\widetilde{q} \in (\underline{q}, \overline{q})$, and thus $EUD(i \text{ delegate to } e) - EUD(i \text{ delegate to } j) > 0$. □

A.1.1 LD: K (odd) experts; M (even) non-experts.

We report here, for generic parameter values, the formulas we used to derive the equilibria for the experimental parametrizations. As always, $EUND(q_i)$ is interim expected utility for a voter with realized precision q_i; the equilibrium threshold is denoted $\widetilde{q}$ and solves $EUND(q_i = \widetilde{q}) = EUD$, and $\mu_v(\widetilde{q}) \equiv \mathbb{E}_F[q_j|q_j > \widetilde{q}]$. We find:

$$EUND(q_i, \widetilde{q}) = \sum_{z=0}^{M-1} \binom{M-1}{z} (1-F(\widetilde{q}))^z \, F(\widetilde{q})^{M-1-z} \sum_{c_n=0}^{z} \binom{z}{c_n} (\mu_v(\widetilde{q}))^{c_n} (1-\mu_v(\widetilde{q}))^{z-c_n} \times$$
$$\times \left\{ \left(\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e} (1-p)^{K-c_e} \sum_{x_1=0}^{M-z-1} \sum_{x_2=0}^{M-z-1-x_1} \cdots \sum_{x_{K-1}=0}^{M-z-1-\sum_{k=1}^{K-2} x_k} \frac{(M-z-1)!}{\prod_{k=1}^{K} x_k!} \right) \times \right.$$
$$\times \left(q_i \left((1/K)^{M-z-1} I_{c_n+1+c_e+\sum_{k=1}^{c_e} x_k > (M+K)/2} \right) + \right.$$
$$\left. \left. (1-q_i) \left((1/K)^{M-z-1} I_{c_n+c_e+\sum_{k=1}^{c_e} x_k > (M+K)/2} \right) \right) \right\}$$

,

where $x_K \equiv M - z - 1 - \sum_{k=1}^{K-1} x_k$, and I_C is an indicator function that equals 1 if condition C is realized and 0 otherwise. Similarly:

$$EUD(\widetilde{q}) = \sum_{z=0}^{M-1} \binom{M-1}{z} (1-F(\widetilde{q}))^z F(\widetilde{q})^{M-1-z} \sum_{c_n=0}^{z} \binom{z}{c_n} (\mu_v(\widetilde{q}))^{c_n} (1-\mu_v(\widetilde{q}))^{z-c_n} \times$$
$$\times \left\{ \left(\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e} (1-p)^{K-c_e} \sum_{y_1=0}^{M-z} \sum_{y_2=0}^{M-z-y_1} \ldots \sum_{y_{K-1}=0}^{M-z-\sum_{k=1}^{K-2} y_k} \frac{(M-z)!}{\prod_{k=1}^{K} y_k!} \right) \times \right.$$
$$\left. \times \left((1/K)^{M-z} I_{c_n + c_e + \sum_{k=1}^{c_e} y_k > (M+K)/2} \right) \right\}$$

where $y_K \equiv M - z - \sum_{k=1}^{K-1} y_k$.

We use as welfare criterion ex ante expected utility, i.e. expected utility before the realization of q_i (but under the correct expectation of $\widetilde{q}$). Hence:

$$EU(\widetilde{q}) = \int_{\underline{q}}^{\widetilde{q}} EUDf(q)dq + \int_{\widetilde{q}}^{p} EUND(q_i)f(q)dq$$

Under MV, ex ante expected utility is given by:

$$EU_{MV} = \sum_{c_n=0}^{M} \binom{M}{c_n} \mu^{c_n} (1-\mu)^{z-c_n} \sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e} (1-p)^{K-c_e} I_{c_n + c_e > \frac{(M+K)}{2}}$$

where, as in earlier use, $\mu \equiv \mathbb{E}_F(q_i)$.

A.1.2 MVA: K (odd) experts; M (even) non-experts.

Under the possibility of abstention as well, all equilibria are in monotone threshold strategies. We denote by $\widetilde{\alpha}$ the equilibrium threshold such that all i with $q_i < \widetilde{\alpha}$ choose to abstain, and all i with $q_i > \widetilde{\alpha}$ choose to vote. With a known and finite electorate size, the equilibria are sensitive to whether K and M are odd or even. In particular, an equilibrium with $\widetilde{\alpha} = \underline{q}$ (all voters cast their vote) exists if and only if N is odd. An equilibrium with $\widetilde{\alpha} = p$ (all experts vote, and none of the other voters do) exists if and only if K is odd. Thus both equilibria exist in our experimental parametrizations. In addition, there are interior equilibria where $\widetilde{\alpha} \in (\underline{q}, p)$. Denoting by $EUV(q_i, \widetilde{\alpha})$ interim expected utility from voting, given q_i, and by $EUA(\widetilde{\alpha})$ interim expected utility from abstaining (which does not depend on q_i), $\widetilde{\alpha}$ must solve $EUV(q_i, \widetilde{\alpha}) = EUA(\widetilde{\alpha})$, where:

$$EUV(q_i,\widetilde{\alpha}) = \sum_{v=0}^{M-1} \binom{M-1}{v} (1-F(\widetilde{\alpha}))^v F(\widetilde{\alpha})^{M-1-v} \sum_{c_n=0}^{v} \binom{v}{c_n} (\mu_v(\widetilde{\alpha}))^{c_n} (1-\mu_v(\widetilde{\alpha}))^{v-c_n} \times$$

$$\times \left\{ \left(\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e} (1-p)^{K-c_e} \right) \times \right.$$

$$\times \left(q_i \left(I_{c_n+1+c_e>(v+1+K)/2} + (1/2) I_{c_n+1+c_e=(v+1+K)/2} \right) + \right.$$

$$\left. \left. (1-q_i) \left(I_{c_n+c_e>(v+1+K)/2} + (1/2) I_{c_n+c_e=(v+1+K)/2} \right) \right) \right\},$$

and:

$$EUA(\widetilde{\alpha}) = \sum_{v=0}^{M-1} \binom{M-1}{v} (1-F(\widetilde{\alpha}))^v F(\widetilde{\alpha})^{M-1-v} \sum_{c_n=0}^{v} \binom{v}{c_n} (\mu_v(\widetilde{\alpha}))^{c_n} (1-\mu_v(\widetilde{\alpha}))^{v-c_n} \times$$

$$\times \left(\sum_{c_e=0}^{K} \binom{K}{c_e} p^{c_e} (1-p)^{K-c_e} \right) \left(I_{c_n+c_e>(v+K)/2} + (1/2) I_{c_n+c_e=(v+K)/2} \right).$$

Ex ante expected utility, before the realization of q_i but under the correct expectation of $\widetilde{\alpha}$, is given by:

$$EU_{MVA}(\widetilde{\alpha}) = \int_{\underline{q}}^{\widetilde{\alpha}} EUA(\widetilde{\alpha}) f(q) dq + \int_{\widetilde{q}}^{p} EUV(q_i,\widetilde{\alpha}) f(q) dq$$

A.2 Additional empirical results

A.2.1 Frequency of delegation and abstention, including order and session treatments controls

Table A.1 reports the regressions on the determinants of delegation and abstention in Experiment 1 with the set of controls for order and treatment composition of the relevant session. In both tables, the excluded case is MVA played as first treatment in MVA-only sessions. Delegation is lower in sessions where LD is experienced after MVA, regardless of group size, but the net effect of delegation remains positive. Abstention responds to order in MVA-only sessions, and the effect depends on group size.

Table A.2 presents the corresponding fractional probit estimates for Experiment 2. These estimates confirm the results of the linear probability model in the text. Accuracy is not a

Experiment 1: Frequency of Delegation or Abstention.
Probit Model.

	(1) N=5	(2) N=15
LD	0.938**	0.677**
	(0.274)	(0.249)
	[0.001]	[0.007]
Signal Precision	-2.624***	-2.691***
	(0.307)	(0.208)
	[0.000]	[0.000]
LD * Signal Precision	0.065	0.102
	(0.225)	(0.144)
	[0.772]	[0.477]
Round	0.085	0.274
	(0.209)	(0.197)
	[0.684]	[0.164]
LD * Round	-0.345	-0.174
	(0.371)	(0.239)
	[0.352]	[0.466]
Second	0.534***	-0.341**
	(0.026)	(0.126)
	[0.000]	[0.007]
LD * Second	-0.332*	0.125
	(0.160)	(0.132)
	[0.038]	[0.343]
Second * Mixed	-0.451***	0.295***
	(0.025)	(0.030)
	[0.000]	[0.000]
LD * Second * Mixed	-0.033	-0.577**
	(0.022)	(0.183)
	[0.140]	[0.002]
Constant	0.582***	0.992***
	(0.128)	(0.237)
	[0.000]	[0.000]
Observations	1,920	2,880

*** p<0.01, ** p<0.05, * p<0.1

Table A.1: *Determinants of delegation and abstention.* Probit models. Standard errors in parentheses, clustered at the session level. P-values in brackets. Delegation/abstention is measured as a binary 0-1 subject decision. The values for signal precision and round have been scaled to be between 0 and 1.

significant predictor of participation in voting. The regressions detect a decline in delegation and abstention as blocs proceed, in line with increased familiarity with the task

A.2.2 Frequency of delegation and abstention. First treatments only

We report here (Tables A.3 and A.4) linear probability and probit regressions of the individual decision to delegate or abstain, selecting data from the first treatment played in each session only. That is, corresponding to an in-between subjects design with 20 rounds only.

Standard errors are clustered at the session level. As discussed in the text, the results are effectively unchanged for the small groups ($N = 5$), but there is a loss of precision in the $N = 15$ regressions. The most striking result is the null effect of signal quality on the decision to abstain under MVA3 (while signal precision continues to affect negatively delegation under LD3).

Experiment 2: Frequency of Delegation or Abstention.
Fractional Probit Model.

	(1) N=5 & N=15	(2) N=125
Accuracy	-0.339	0.013
	(0.228)	(0.269)
	[0.138]	[0.962]
LD	0.604***	0.577***
	(0.103)	(0.110)
	[0.000]	[0.000]
N=15	0.017	
	(0.105)	
	[0.871]	
Keys: [E][Y]	-0.025	0.075
	(0.104)	(0.110)
	[0.809]	[0.496]
Bloc	0.001	0.025
	(0.032)	(0.034)
	[0.981]	[0.460]
Constant	-0.414**	-0.534***
	(0.180)	(0.180)
	[0.022]	[0.003]
Observations	1,800	1,500

*** p<0.01, ** p<0.05, * p<0.1

Table A.2: *Frequency of delegation or abstention.* Fractional probit models. Standard errors are clustered at the individual level. P-values in brackets. Delegation/abstention is measured as the share of rounds in a given bloc in which a subject chose to delegate/abstain (with a range from 0 to 1). Accuracy is the share of rounds in the bloc that subject answered correctly. Subjects randomly use either keys [V] and [N] or [E] and [Y] to decide whether to vote; a dummy for being assigned [E][Y] is included. The values for bloc have been scaled to be between 0 and 1; the coefficient for "bloc" thus indicates the effect of going from the first to last bloc.

Experiment 1: Frequency of Delegation or Abstention, First Treatments. N=5.		
	(1) Linear Probability	(2) Probit
LD	0.459**	1.450***
	(0.081)	(0.288)
	[0.011]	[0.000]
Round	-0.005	-0.022
	(0.020)	(0.079)
	[0.816]	[0.778]
Signal Precision	-0.668***	-2.295***
	(0.107)	(0.423)
	[0.008]	[0.000]
LD * Round	-0.127**	-0.459***
	(0.024)	(0.117)
	[0.013]	[0.000]
LD * Signal Precision	-0.322*	-0.890*
	(0.111)	(0.456)
	[0.062]	[0.051]
Constant	0.639***	0.506***
	(0.056)	(0.172)
	[0.001]	[0.003]
Observations	960	960
R-squared	0.330	

*** p<0.01, ** p<0.05, * p<0.1

Table A.3: *Determinants of delegation and abstention; N=5. First treatments only.* Standard errors are clustered at the session level.

Experiment 1: Frequency of Delegation or Abstention, First Treatments. N=15.

	(1) Linear Probability	(2) Probit
LD	0.174	0.610
	(0.105)	(0.375)
	[0.160]	[0.104]
Round	0.078	0.263
	(0.075)	(0.261)
	[0.348]	[0.314]
Signal Precision	-0.868***	-2.648***
	(0.045)	(0.129)
	[0.000]	[0.000]
LD * Round	-0.048	-0.162
	(0.077)	(0.270)
	[0.557]	[0.548]
LD * Signal Precision	0.077	0.202
	(0.074)	(0.336)
	[0.351]	[0.547]
Constant	0.835***	0.977***
	(0.094)	(0.297)
	[0.000]	[0.001]
Observations	1,440	1,440
R-squared	0.301	

*** p<0.01, ** p<0.05, * p<0.1

Table A.4: *Determinants of delegation and abstention; N=15. First treatments only.* Standard errors are clustered at the session level.

A.2.3 Monotonicity violations and individual thresholds

As described in the text, for both group sizes, just below 60% of subjects have no violations at all under LD; just above 60% under MVA. Figure A.1 reports histograms of the frequency of monotonicity violations, by subject.

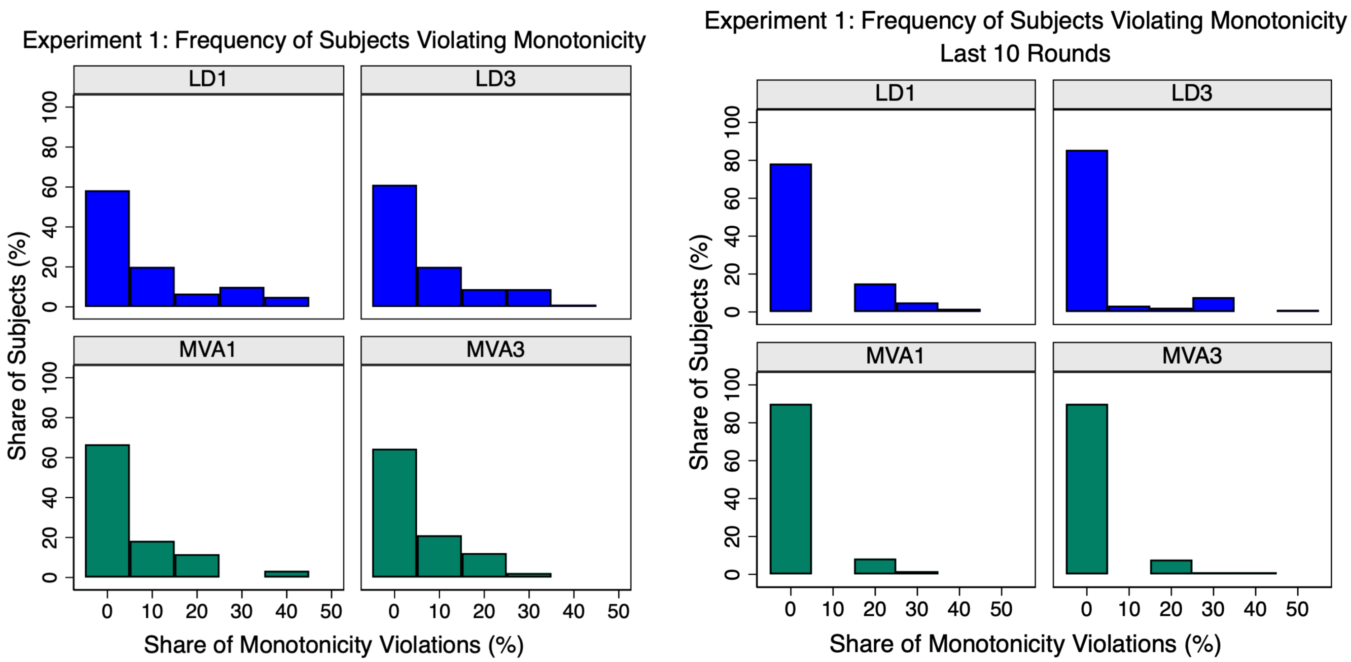

Figure A.1: *Monotonicity violations. Histograms.* For each subject, we calculate the frequency of violations given the number of rounds played as non-expert. The maximum possible frequency is 50%.

We can use monotonicity to estimate individual precision thresholds for delegation and abstention–the thresholds below which each participant delegates or abstains. Figure A.2 reports, for each participant, the mean of the range of thresholds that are consistent with minimal monotonicity violations; the size of the dots is proportional to the number of participants at that threshold. The dark blue (for LD) and dark green (for MVA) diamonds correspond to the average empirical thresholds, and the respective light ones to the theory. The figure confirms the over-delegation that characterizes LD, while again average values for abstention are close to the theoretical predictions. The dispersion in estimated thresholds is typical of similar experiments (for example, Levine and Palfrey, 2007; Morton and Tyran, 2011), but is in clear tension with the focus on symmetric equilibria.

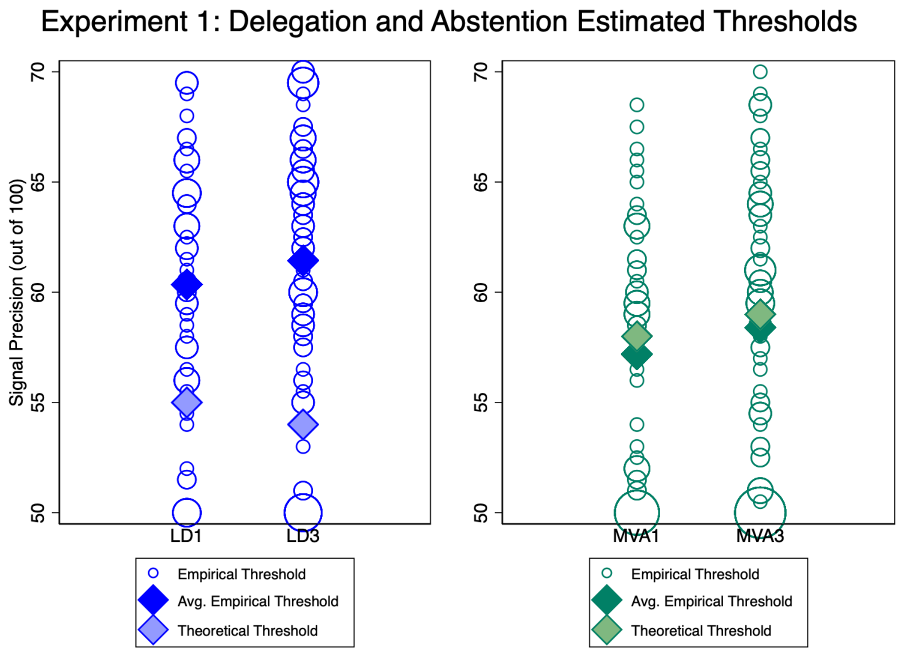

Figure A.2: *Individual delegation and abstention thresholds.*

A.2.4 Voting against signal and signal quality

Table A.5 reports linear probability and probit regressions studying the determinants of voting against signal. The only significant variable is signal quality (precision). Neither the voting system nor the size of the group matter, but a decrease in signal quality is strongly correlated with an increase in voting against signal. Although voting against signal is an inferior action, the loss from doing so is indeed increasing in signal quality. There could be a small learning effect (voting against signal is less frequent in the treatments run second), but it is quite noisy and counterbalanced by a similarly sized, if statistically insignificant, increase in voting against signal as rounds proceed, possibly from fatigue.

Experiment 1: Frequency of Voting Against Signal.

	(1) Linear Probability	(2) Probit
Signal Precision	-0.407***	-2.058***
	(0.038)	(0.159)
	[0.000]	[0.000]
Round	0.005	0.031
	(0.006)	(0.050)
	[0.457]	[0.535]
LD	-0.006	-0.057
	(0.009)	(0.066)
	[0.504]	[0.390]
N = 15	-0.01	-0.101
	(0.022)	(0.160)
	[0.648]	[0.528]
Second	-0.012	-0.098*
	(0.009)	(0.051)
	[0.216]	[0.056]
Second * Mixed	-0.034	-0.230
	(0.021)	(0.177)
	[0.136]	[0.193]
N = 15 * Second	0.037	0.282
	(0.023)	(0.181)
	[0.132]	[0.120]
N = 15 * Mixed	-0.010	-0.066
	(0.031)	(0.248)
	[0.744]	[0.789]
Constant	0.424***	0.174
	(0.036)	(0.116)
	[0.000]	[0.134]
Observations	2,552	2,552
R-squared	0.154	

*** p<0.01, ** p<0.05, * p<0.1

Table A.5: *Frequency of voting against signal.* Standard errors are clustered at the session level.

B Online Appendix

B.1 The bootstrapping procedure: allowing for individual correlation across rounds

Replicating what happens in an individual session, we draw with replacement 15 subjects from the relevant treatment, each with all choices made over the 20 rounds. Among these 15 subjects, we draw, with replacement, 3 subjects, assigning to each of them one choice they made as expert,[62] and 12 subjects, assigning to each one choice made as non-expert. In $N = 15$, that constitutes the group and yields one group decision; in $N = 5$ treatments, we divide the 12 subjects randomly into three groups of 4 and assign to each group one of the experts drawn earlier, generating three group decisions. We repeat new draws of 3 experts and 12 non-experts as above 20 times, generating 20 decisions from the same sample of 15 subjects if the treatment has $N = 15$, and 60 if the treatment has $N = 5$, thus simulating one experimental session. We then draw, always with replacement, a new group of 15 subjects, and repeat the procedure, each time generating 20 (60) decisions from the same group of 15 subjects, depending on the size of the group in the treatment. We repeat the whole procedure 4 times ($N = 5$), or 6 times ($N = 15$) generating 240 (120) decisions, as in our data from each of the $N = 5$ ($N = 15$) treatments. We then calculate the frequency of correct decisions, and consider that one data point for that treatment. We repeat the whole process 100,000 times and generate a distribution of the frequency with which the correct decision was reached.

B.2 A short note on beliefs

As reported in the text, at the end of the RDK experiment, we asked: "On average what percentage of trials in the second part do you think you got right?" and "On average what percentage of trials in the second part do you think the experts got right?" Participants earned 25 cents for replies the fell within 5% of the observed percentage in the group to which the participant was assigned. Figure B.1 plots distributions of the difference between actual own accuracy and the corresponding reported belief, in the left panel, and between actual average experts' accuracy and the corresponding reported belief, in the panel on the right, for the two different coherence levels. Both panels rely on a single measure of beliefs per subject, and thus the unit of analysis is the subject.

As the figure shows, accuracy of beliefs varies little with coherence. On average, beliefs about own accuracy are remarkably correct, with barely detectable underestimation. The

[62] If the subject was never an expert, the subject is dropped and another one is drawn.

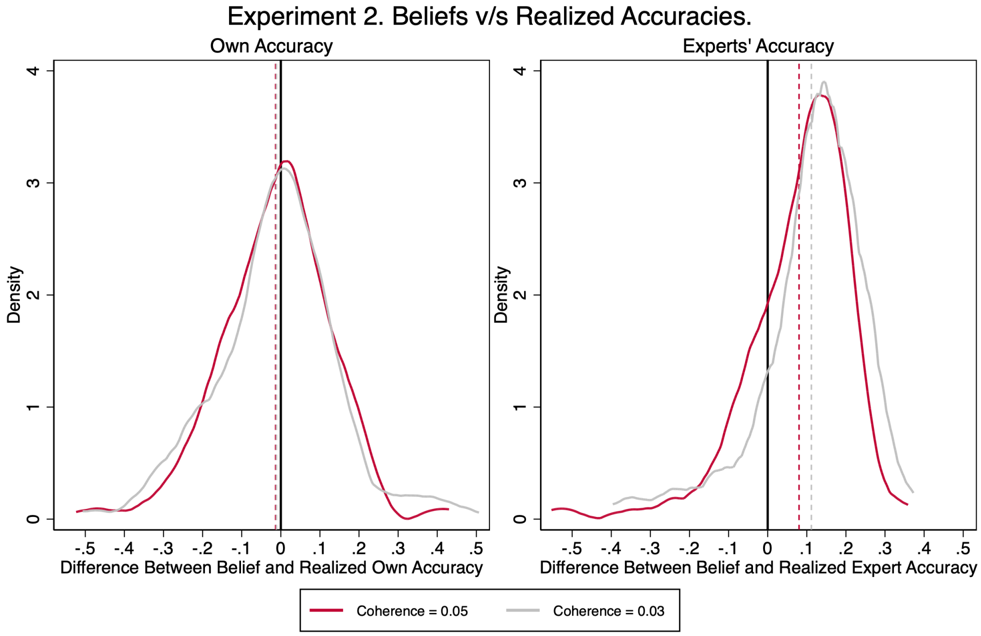

Figure B.1: *Difference between actual and reported accuracies.* The left panel refers to own accuracy; the right panel to the average of the experts' accuracy. The dashed vertical lines correspond to the means.

distribution is instead shifted to the right in the case of experts, with about 15% average overestimation, imputable, as we say in the text, to stochasticity in accuracies and reversion to the mean, neglected in the formation of beliefs. Overestimation is somewhat larger at smaller coherence, when the task is more difficult.

Table B.1 reports variants of the regressions in Table 5 in the text, now including measures of beliefs. The dependent variable is the frequency of delegation or abstention at the subject level (120 tasks per subject), with data aggregated by coherence level.

As described in the text, beliefs are strongly significant and have the expected signs but cannot explain the higher tendency towards delegation rather than abstention.

B.3 The Random Dot Kinematogram

In a Random Dot Kinematogram (RDK), the perceptual stimulus consists of a number of dots being displayed on a screen. A proportion of these dots are determined to be signal dots, while the remaining are noise dots. Signal dots all move in a determined direction, while noise dots move at random according to an algorithm. The task consists in reporting the direction in which the signal dots are moving. This direction is called the coherent direction and the proportion of signal dots, the coherence, is the main factor in determining

Experiment 2: Frequency of Delegation or Abstention.

	(1) Coherence = 0.05	(2) Coherence = 0.03
Own Accuracy	-0.250	-0.014
	(0.288)	(0.334)
	[0.387]	[0.968]
Experts' Accuracy	-0.282	
	(0.389)	
	[0.469]	
Belief About Own Accuracy	-0.714***	-0.607***
	(0.144)	(0.161)
	[0.000]	[0.000]
Belief About Experts' Accuracy	0.337**	0.615***
	(0.160)	(0.181)
	[0.037]	[0.001]
LD	0.249***	0.230***
	(0.037)	(0.041)
	[0.000]	[0.000]
Constant	0.757***	0.223
	(0.269)	(0.194)
	[0.005]	[0.252]
Observations	300	250
R-squared	0.184	0.163

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B.1: *Beliefs and the decision to delegate/abstain.* Linear probability models. Expert accuracy is excluded from coherence 0.03 because, with a single group per treatment, it is collinear with LD.

the difficulty of the task.

The task can be programmed in various ways, using a variety of parameters (e.g. color, duration, algorithm, number of dots). Research has been done to study the various effects of using different combinations of these parameters (Pilly and Seitz, 2009; Schütz et al., 2010). We take advantage of the recent development of a customizable version of the RDK (Rajananda et al., 2018) which can be implemented as a plugin in jsPsych. This version allows for the configuration of various parameters in order to adjust the task as desired by the researchers. We report in the table below the parameters that we used in our experiment. The reader can find details about how they affect the task in Rajananda et al., 2018 and in the following link: https://www.jspsych.org/6.3/plugins/jspsych-rdk/. It is important to emphasize again that our objective in using the RDK was not to study perception in itself, but rather to create a common task that is reasonably well controlled and calibrated and where, nevertheless, the information about the accuracy of the signals remains ambiguous.

Duration:	1 second
Directions:	Left/Right
Number of dots:	300
Background color:	Black
Color of dots:	White
Dot radius:	2 pixels
Dot movement per frame:	1 pixel
Aperture width:	600 pixels
Aperture height:	400 pixels
Signal selection:	Same
Minimal screen resolutions	1000x600 pixels
Noise type:	Random direction
Aperture shape:	Ellipse
Reinsertion:	Dots reappear randomly when hitting edge
Fixation cross:	No
Aperture border:	No
Coherence:	20% to 3% (according to treatment)

Table B.2: *Experiment 2: RDK parameters*

B.4 Discussion: Some additional regressions

We report below regressions discussed in Section 9.1. The first set tests the sensitivity of delegation/abstention to one-period lagged delegation/abstention by others. As discussed in the text, if participants believe that others are voting at too low precisions, delegation should respond negatively to lagged delegations by others. We see no convincing evidence in the data.[63]

The second set of tables tests the sensitivity of delegation/abstention decisions to group size. As discussed in the text, if participants neglect the behavior of others when thinking whether to delegate but not when thinking whether to abstain, no such sensitivity should be detected under LD, while group size could affect abstention under MVA. As the tables show, there is no effect of group size on the frequency of delegation, while the frequency of abstention is significantly higher in MVA3 than in MVA1 in Experiment 1, and higher, if only marginally significant, in MVA25 than in the two smaller groups in Experiment 2.[64]

B.5 Discussion: Three simple models

B.5.1 Introducing private values

As in our main model, we focus on symmetric pure equilibria with sincere voting where the decision to delegate or abstain is symmetric across signals. Because the preferences of the expert are publicly known, such a decision typically conveys information about the voter's preferences. But note that we have introduced private values while preserving the symmetry of the model. Symmetric equilibria with sincere voting and delegation/abstention decisions symmetric across signals continue to exist.

We say that two individuals are *congruent* if they share the same private preferences, and denote by s the state of being congruent with ("similar" to) the expert. The set of strategies is $\{x, x(s), x(-s), d/a\}$, where x stands for voting unconditionally on congruence, $x(s)$ for voting if congruent and delegating or abstaining if not; $x(-s)$ stands for voting if non-congruent and delegating or abstaining otherwise, and d/a stands for delegating or abstaining always. We begin by analyzing LD.

LD As always, for any $\{p, q\}$, there is an equilibrium in which both non-experts delegate. Is there an equilibrium with voting always?

Consider the expected utility of non-expert voter i who is congruent with the expert.

[63] We do not estimate probit regressions to avoid biases due to lagged dependent variables.

[64] The positive coefficient on size under MVA suggests free-riding, but without further assumptions we are not positing a specific sign. Recall that the RDK coherence is lower in the largest group, which would naturally induce less willingness to vote. But what we emphasize here is the differential response of abstention (that increases) and delegation (that does not).

Experiment 1: Frequency of Delegation or Abstention. N=5 and N=15.

	(1) LD & N=5	(2) LD & N=15	(1) MVA & N=5	(2) MVA & N=15
# Delegations/Abstentions by Others in Prev. Round	-0.027	-0.013	0.008	-0.012
	(0.033)	(0.015)	(0.021)	(0.010)
	[0.481]	[0.411]	[0.737]	[0.267]
Signal Precision	-0.866***	-0.849***	-0.782***	-0.865***
	(0.079)	(0.046)	(0.085)	(0.049)
	[0.002]	[0.000]	[0.003]	[0.000]
Second	0.071	-0.065**	0.144***	-0.108*
	(0.068)	(0.019)	(0.010)	(0.045)
	[0.375]	[0.018]	[0.001]	[0.063]
Second * Mixed	-0.159***	-0.100	-0.118***	0.090***
	(0.026)	(0.068)	(0.010)	(0.011)
	[0.009]	[0.200]	[0.001]	[0.000]
Round	-0.087	0.022	0.018	0.092
	(0.051)	(0.026)	(0.065)	(0.068)
	[0.183]	[0.433]	[0.805]	[0.235]
Constant	1.055***	1.131***	0.681***	0.883***
	(0.039)	(0.089)	(0.056)	(0.100)
	[0.000]	[0.000]	[0.001]	[0.000]
Observations	912	1,368	912	1,368
R-squared	0.296	0.290	0.267	0.302

*** p<0.01, ** p<0.05, * p<0.1

Table B.3: *Response of delegation/abstention to number of others who delegated/abstained in group in previous round; N=5 and N=15.* Linear probability models. Standard errors in parentheses, clustered at the session level.

Experiment 1: Frequency of Delegation. N=5 and N=15.

	(1) Linear Probability	(2) Probit
N=15	0.093	0.326*
	(0.053)	(0.179)
	[0.123]	[0.069]
Signal Precision	-0.854***	-2.580***
	(0.042)	(0.165)
	[0.000]	[0.000]
Second	0.046	0.138
	(0.054)	(0.181)
	[0.418]	[0.448]
Second * Mixed	-0.119***	-0.363***
	(0.027)	(0.088)
	[0.003]	[0.000]
N = 15 * Second	-0.085	-0.300
	(0.058)	(0.200)
	[0.187]	[0.133]
N = 15 * Mixed	0.004	-0.007
	(0.023)	(0.079)
	[0.873]	[0.925]
Round	-0.016	-0.045
	(0.024)	(0.084)
	[0.526]	[0.589]
Constant	0.969***	1.419***
	(0.055)	(0.197)
	[0.000]	[0.000]
Observations	2,400	2,400
R-squared	0.290	

*** p<0.01, ** p<0.05, * p<0.1

Table B.4: *Determinants of delegation; N=5 and N=15.* Standard errors in parentheses, clustered at the session level.

Experiment 1: Frequency of Abstention. N=5 and N=15.		
	(1) Linear Probability	(2) Probit
N=15	0.138**	0.479***
	(0.047)	(0.157)
	[0.021]	[0.002]
Signal Precision	-0.827***	-2.636***
	(0.044)	(0.170)
	[0.000]	[0.000]
Second	0.096*	0.325**
	(0.047)	(0.164)
	[0.080]	[0.047]
Second * Mixed	-0.011	-0.023
	(0.051)	(0.183)
	[0.833]	[0.901]
N = 15 * Second	-0.132	-0.451*
	(0.073)	(0.255)
	[0.112]	[0.077]
N = 15 * Mixed	0.000	-0.002
	(0.048)	(0.166)
	[0.994]	[0.989]
Round	0.060	0.203
	(0.044)	(0.157)
	[0.215]	[0.195]
Constant	0.685***	0.525***
	(0.039)	(0.127)
	[0.000]	[0.000]
Observations	2,400	2,400
R-squared	0.285	

*** p<0.01, ** p<0.05, * p<0.1

Table B.5: *Determinants of abstention; N=5 and N=15.* Standard errors in parentheses, clustered at the session level.

Experiment 2: Frequency of Delegation. N=5, N=15, and N=125.				
	(1)	(2)	(3)	(4)
	Linear Probability	Linear Probability	Fractional Probit	Fractional Probit
Accuracy	-0.074	-0.069	-0.187	-0.174
	(0.093)	(0.094)	(0.235)	(0.236)
	[0.427]	[0.463]	[0.426]	[0.462]
N=15	-0.059	-0.057	-0.147	-0.144
	(0.059)	(0.060)	(0.149)	(0.149)
	[0.324]	[0.336]	[0.323]	[0.336]
N=125	0.008	0.006	0.019	0.0152
	(0.055)	(0.056)	(0.139)	(0.140)
	[0.890]	[0.915]	[0.890]	[0.913]
Keys: [E][Y]		0.040		0.100
		(0.042)		(0.106)
		[0.344]		[0.342]
Bloc		0.017		0.042
		(0.014)		(0.035)
		[0.229]		[0.228]
Constant	0.575***	0.545***	0.189	0.113
	(0.073)	(0.076)	(0.184)	(0.191)
	[0.000]	[0.000]	[0.304]	[0.554]
Observations	1,650	1,650	1,650	1,650
R-squared	0.008	0.011		

*** p<0.01, ** p<0.05, * p<0.1

Table B.6: *Determinants of delegation; N=5, N=15, and N=125.* Standard errors in parentheses, clustered at the individual level.

Experiment 2: Frequency of Abstention. N=5, N=15, and N=125.

	(1) Linear Probability	(2) Linear Probability	(3) Fractional Probit	(4) Fractional Probit
Accuracy	-0.061	-0.061	-0.187	-0.188
	(0.089)	(0.089)	(0.259)	(0.260)
	[0.494]	[0.493]	[0.470]	[0.470]
N=15	0.068	0.069	0.210	0.212
	(0.049)	(0.048)	(0.152)	(0.151)
	[0.162]	[0.158]	[0.166]	[0.162]
N=125	0.086*	0.086*	0.260*	0.262*
	(0.047)	(0.047)	(0.146)	(0.146)
	[0.068]	[0.067]	[0.075]	[0.073]
Keys: [E][Y]		-0.025		-0.073
		(0.037)		(0.108)
		[0.498]		[0.497]
Bloc		-0.008		-0.023
		(0.010)		(0.030)
		[0.431]		[0.434]
Constant	0.265***	0.280***	-0.633***	-0.588***
	(0.070)	(0.073)	(0.212)	(0.220)
	[0.000]	[0.000]	[0.003]	[0.007]
Observations	1,650	1,650	1,650	1,650
R-squared	0.012	0.013		

*** p<0.01, ** p<0.05, * p<0.1

Table B.7: *Determinants of abstention; N=5, N=15, and N=125.* Standard errors in parentheses, clustered at the individual level.

Suppose first that non-expert j always votes, whether congruent or not. Then:

$$EU_i(x|i_s, x_j) = pq + [(1-p)q + (1-q)p][q/2 + (1-q)/2] = (p+q)/2$$
$$EU_i(d) = p$$

Consider $EU_i(x|i_s, x_j)$. Because i is s, from $i's$ perspective, the correct decision is reached if both i and the expert have the correct signal (with probability pq), or if one of them has the wrong signal (with probability $p(1-q) + q(1-p)$) but the desired decision wins thanks to j's vote–which happens if j is s and has the correct signal (with probability $(1/2)q$), or if j is $\not s$ and has the wrong signal (with probability $(1/2)(1-q)$). (All derivations below follow a similar logic).

With $p > q$, delegation dominates voting for the congruent non-expert. Hence there is no pure symmetric equilibrium such that the two non-experts always vote, whether congruent or not.

Suppose then that j delegates if congruent, and votes if not. It follows that i can be pivotal only if j votes, and thus if j is $\not s$. Conditioning on pivotality, if i is congruent:

$$EU_i(x|i_s, x(-s)_j) = pq + p(1-q)^2 + q(1-p)(1-q) = p - q(1-q)(2p-1)$$
$$EU(d|i_s) = p$$

Given $p > 1/2$, if i is congruent, i's best response to j voting when non-congruent is to delegate.

If i is not congruent, again conditioning on pivotality:

$$EU_i(x|i_{\not s}, x(-s)_j) = q^2 + 2(1-q)q(1-p)$$
$$EU(d|i_{\not s}) = 1-p$$

For all $p > q$, conditioning on pivotality, $EU(d|i_{\not s}) < EU_i(x|i_{\not s} x(-s)_j)$. Thus indeed there is an equilibrium in which a non-expert delegates if congruent, and votes if non congruent.

Finally, as expected, there is no equilibrium such that a non-expert votes if congruent and delegates if non-congruent. Suppose j follows such a strategy, and consider i's best response when i is not congruent. Again, i can be pivotal only if j does not delegate, and thus if j votes and is congruent. Then, if j votes and is congruent:

$$EU_i(x|i_{\not s}, (xs)_j) = q(1-q)(1-p) + q[(1-q)p + q(1-p)] + (1-q)^2(1-p)$$
$$EU(d|i_{\not s}) = 1-p$$

For all $p > q$, conditioning on pivotality, $EU(d|i_{\not s}) < EU_i(x|i_{\not s}, x(s)_j)$: if i is not congruent, voting dominates delegating when j follows strategy $x(s)$. Voting when congruent and delegating when not is not an equilibrium.

Summarizing, then, under LD there are two pure equilibria with symmetric strategies: in one equilibrium, non-experts always delegate; in the second equilibrium, the equilibrium strategy is to delegate if congruent, and to vote if not congruent.

MVA Consider now the problem when the choice for non-experts is between voting and abstaining. We begin by establishing that there is no equilibrium in which both non-experts always abstain. Suppose i is not congruent and j abstains. Then:

$$\begin{aligned}
EU_i(x|i_{\not s}, a_j) &= q(1-p) + (1/2)qp + (1/2)(1-p)(1-q) \\
EU(a|i_{\not s}, a_j) &= 1-p
\end{aligned}$$

where we denote by a_j j's choice to abstain.

For all $p > q$, $EU(x|i_{\not s}, a_j) > EU_i(a|i_{\not s}, a_j)$: if j abstains, non-congruent i prefers to vote. Hence indeed there is no equilibrium where both non-experts always abstain.

There is always, however, an equilibrium where both non-experts always vote. Suppose j always votes, and suppose first that i is congruent. Then:

$$\begin{aligned}
EU_i(x|i_s, x_j) &= pq(q/2 + (1-q)/2) + pq((1-q)/2 + q/2) + \\
&\quad + (1-q)p(q/2 + (1-q)/2) + (1-p)q(q/2 + (1-q)/2) = \\
&= pq + (1/2)(p(1-q) + q(1-p)) \\
EU(a|i_s, x_j) &= p(q/2 + (1-q)/2) + p((1-q)/2 + \\
&\quad + q/2)(1/2) + (1-p)(q/2 + (1-q)/2)(1/2) = p/2 + 1/4
\end{aligned}$$

For all $q \in (1/2, p)$ and all $p < 1$, $EU_i(x|i_s, x_j) > EU_i(a|i_s, x_j)$. Suppose now that i is not congruent. Then:

$$\begin{aligned}
EU_i(x|i_{\not s}, x_j) &= (1-p)q((1-q)/2 + q/2) + (1-p)q(q/2 + (1-q)/2) + \\
&\quad + (1-p)(1-q)((1-q)/2 + q/2) + pq((1-q)/2 + q/2) = \\
&= (1 + q - p)/2 \\
EU(a|i_{\not s}, x_j) &= (1-p)((1-q)/2 + q/2) + (1-p)(q/2 + (1-q)/2) + \\
&\quad + p((1-q)/2 + q/2)(1/2) = \\
&= (1-p)/2 + 1/4 = (1 + 1/2 - p)/2
\end{aligned}$$

Indeed, for all $q \in (1/2, p)$ and all $p < 1$, $EU_i(x|i_{\not s}, x_j) > EU_i(a|i_{\not s}, x_j)$. For all relevant parameter values, an equilibrium exists in which both non-experts always vote.

Consider now a candidate equilibrium in which non-experts abstain if congruent and vote if not congruent. Suppose j follows such a strategy. What is i's best response?

Suppose first that i is congruent. Then:

$$\begin{aligned} EU_i(x|i_s, x(-s)_j) &= (1/2)(pq + p(1-q)/2 + q(1-p)/2) + \\ &\quad + (1/2)(pq(1-q) + pq^2 + p(1-q)^2 + (1-p)q(1-q) = \\ &= (p+q)/4 + (1/2)[pq + (1-q)\,(p(1-q) + q(1-p))] \\ EU(a|i_s, x(-s)_j) &= (1/2)p + (1/2)(pq/2 + p(1-q) + \\ &\quad +(1-p)(1-q)/2 = \\ &= p(2-q)/2 + (pq + (1-p)(1-q))/2 \end{aligned}$$

If instead i is non-congruent, expected utilities are:

$$\begin{aligned} EU_i(x|i_{\not s},, x(-s)_j) &= (1/2)((1-p)q + pq/2 + (1-q)(1-p)/2) + \\ &\quad + (1/2)((1-p)q^2 + 2(1-p)q(1-q) + pq^2 \\ EU(a|i_{\not s},, x(-s)_j) &= (1/2)(1-p) + (1/2)((1-p)q + pq/2 + (1-p)(1-q)/2) = \\ &= (1-p)(1+q)/2 + (pq + (1-p)(1-q))/4 \end{aligned}$$

For any $q \in (1/2, p)$ and all $p < 1$, $EU_i(x|i_{\not s}, x(-s)_j) > EU_i(a|i_{\not s}, x(-s)_j)$: a non-congruent i indeed prefers to vote. As for a congruent i, i's best response is less straightforward and depends on parameters: for any $p \in (1/2, 1)$, there exists $\overline{q}(p) \in (1/2, p)$ such that $EU(a|i_s, x(-s)_j) > EU_i(x|i_s, x(-s)_j)$ if $q < \overline{q}(p)$, and not otherwise.[65] Thus an equilibrium such that non-experts prefer to vote if non-congruent and abstain if congruent exists for $q \leq \overline{q}(p)$.

Finally, we want to verify that voting when congruent and abstaining when non-congruent cannot be an equilibrium strategy. Suppose j follows such a strategy, and consider i's best

[65] The precise expression is

$$\overline{q}(p) = \frac{-1 + p + \sqrt{1/2 - p + p^2}}{-1 + 2p}.$$

For example, if $p = 0.7$,then $\overline{q}(0.7) = 0.596$.

response when i is non-congruent. Then:

$$\begin{aligned}
EU_i(x|i_{\not s},,x(s)_j) = {} & (1/2)\left((1-p)(1-q)q + (1-p)q(1-q) + \right.\\
& \left. + (1-p)(1-q)^2 + p(1-q)q\right) + \\
& +(1/2)\left((1-p)q + pq/2 + (1-p)(1-q)/2\right) \\
EU(a|i_{\not s},, x(s)_j) = {} & (1/2)(1-p) + (1/2)\left((1-p)(1-q) + (1-p)q/2 + p(1-q)/2\right)
\end{aligned}$$

It is readily verified that non-congruent i prefers to deviate and vote.

Summarizing then, under MVA there are two pure symmetric sincere equilibria: for all relevant parameter values, there is an equilibrium in which both non-experts always vote; given $p \in (1/2, 1]$, there exists a second equilibrium in which non-experts vote if non-congruent and abstain if congruent if q is not too large (i.e. there exists $\overline{q}(p) \in (1/2, p)$ such that the equilibrium exists if $q \leq \overline{q}(p)$).

Comparing LD and MVA Can we rank the different equilibria and the two voting systems? We begin by focussing on non-experts only. We compare equilibria by comparing ex ante expected utility (before the realization of the voter's types).

Under LD, we denote by $EU_{D,all}$ ex ante expected utility in the delegation equilibrium, and by $EU_{D,s}$ ex ante expected utility in the equilibrium where delegation is only chosen if congruent. Thus, keeping in mind that the ex ante probability of agreeing with the expert is 1/2:

$$\begin{aligned}
EU_{D,all} &= (1/2)p + (1/2)(1-p) = 1/2 \\
EU_{D,s} &= (1/4)p + (1/4)p + (1/4)(1-p) + (1/4)[q^2(1-p) + 2(1-p)q(1-q) + q^2 p]
\end{aligned}$$

As we state in the text, the equilibrium with full delegation does poorly because of the ex ante uncertainty on the preferences of the expert. With $EU_{D,all} = 1/2$, it is dominated by the equilibrium with congruent delegation for all $q \in (1/2, p)$: $EU_{D,s} > EU_{D,all}$.

Under MVA, denoting by $EU_{A,none}$ ex ante expected utility when non-experts always vote, and $EU_{A,s}$ ex ante expected utility in the equilibrium with congruent abstention (and

$q \leq \overline{q}(p)$), we find:

$$\begin{aligned}
EU_{A,none} = & (1/4)\left(pq^2 + 2pq(1-q) + (1-p)q^2\right) + \\
& + (1/4)\left(pq(1-q) + pq^2 + p(1-q)^2 + (1-p)q(1-q)\right) + \\
& + (1/4)\left((1-p)q(1-q) + (1-p)q^2 + (1-p)(1-q)^2 + q(1-q)p\right) + \\
& + (1/4)\left((1-p)q^2 + 2(1-p)q(1-q) + pq^2\right) = \\
& = (1/4)(1+2q) \\
EU_{A,s} = & (1/4)p + (1/4)\left(p(1-q) + pq/2 + (1-p)(1-q)/2\right) + \\
& + (1/4)\left((1-p)q + (1-p)(1-q)/2 + pq/2\right) + \\
& + (1/4)\left((1-p)q^2 + 2(1-p)q(1-q) + pq^2\right)
\end{aligned}$$

When the equilibrium with congruent abstention exists, i.e. if $q \leq \overline{q}(p)$, it dominates the no abstention equilibrium: $EU_{A,s} > EU_{A,none}$.

Comparing ex ante expected utilities across the two systems, LD and MVA, $EU_{A,none} > EU_{D,all}$ for all $q \in [1/2, p]$; $EU_{A,s} > EU_{D,all}$ whenever the equilibrium with congruent abstention exists, i.e. for all $q \in (1/2, \overline{q}(p)]$; finally, there exists $q'(p) > \overline{q}(p)$ such that $EU_{A,none} > EU_{D,s}$ if $q > q'(p)$, but the reverse holds otherwise.[66]

Particularly instructive is the comparison between $EU_{D,s}$ and $EU_{A,s}$, that is, between congruent delegation and congruent abstention. In this model, the two expressions are identical. Comparing $EU_{D,s}$ and $EU_{A,s}$ above, note that they must be identical if both non-experts are congruent (with probability 1/4)—in which case the expert alone decides and each member obtains utility 1 with probability p; or if both non-experts are not-congruent (with probability 1/4)—in which case all vote and each obtains utility 1 if all three voters vote in the same direction (with probability $q^2(1-p)$), or if two of the three voters vote in the direction preferred by the non-experts (with probability $2(1-p)q(1-q) + pq^2$). Where the expressions differ is when the two non-experts disagree with each other. Suppose first $\{i = s, j \neq s\}$(an event with probability (1/4)). Under congruent delegation, i delegates and obtains utility 1 with probability p; under congruent abstention, i abstains and obtains utility 1 with probability $(1+p-q)/2 < p$. Conditional on being the voter who is congruent with the expert, congruent delegation is preferred to congruent abstention because it makes it possible for the voter to shift all decision power to the expert, something abstention cannot achieve. But suppose now $\{i \neq s, j = s\}$(again with probability (1/4)). Under congruent

[66] $q'(p) = (p - \sqrt{(p-p^{2)}})/(2p-1)$. For example, if $p = 0.7$, $\overline{q}(0.7) = 0.596$, and $q'(0.7) = 0.604$. If $q = 0.6$, the equilibrium with congruent abstention does not exist, and we find $EU_{D,all} = 0.5$, $EU_{D,s} = 0.551 > EUA_{none} = 0.55$. If $p = 0.8$, $\overline{q}(0.8) = 0.638$ and $q'(0.8) = 0.667$. If $q = 0.6$, $EU_{D,all} = 0.5$, $EU_{D,s} = 0.564 = EU_{A,s} = 0.564 > EUA_{none} = 0.55$; if $q = 0.7$, $EU_{D,all} = 0.5$ $EU_{D,s} = 0.593 < EUA_{none} = 0.6$.

delegation, j delegates and thus i achieves her preferred outcome only with probability $(1-p)$. Under congruent abstention, j cannot make the expert dictator, and i's probability of the preferred outcome is $(1+q-p)/2 > 1-p$: conditional on being the voter who is not congruent with the expert, congruent abstention is better exactly because it leaves the voter with some decision power. In this simple model, the two effects, based on equally probable events, exactly cancel each other: $(1/4)(p+1-p) = (1/4)(1+p-q+1+q-p) = (1/4)$.

Contrary to what one may have expected, when private values are included, LD need not dominate MVA: if we focus on the worst equilibrium, it corresponds to LD with full delegation; if we focus instead on the best equilibrium, at $q \leq \overline{q}(p)$ it corresponds to congruent delegation or abstention, and LD and MVA are equivalent; at $q > q'(p)$, the best equilibrium requires voting by all and can only be supported under MVA. LD dominates only over the range $(\overline{q}(p), q'(p))$, when congruent abstention is not an equilibrium and voting by all is dominated. The range $(q'(p)-\overline{q}(p))$ is increasing in p, but typically not large; for all $p \leq 0.8$, for example, it covers less than 10% of the possible range of q values.

How about the expert then? Here the answer is unambiguous and expected: the expert always prefers LD with full delegation because in such an equilibrium the expert has full control and achieves her preferred outcome with probability p. For completeness, using the notation EUe to indicate the expert's ex ante expected utility (with indices specifying the voting regime and the equilibrium), we report here the expert's ex ante utilities under the different equilibria:

$$\begin{aligned}
EUe_{D,all} &= p \\
EUe_{D,s} &= (1/4)p + 2(1/4)p + (1/4)[(1-q)^2 p + 2q(1-q)p + (1-q)^2(1-p)] \\
EUe_{A,none} &= (1/4)[q^2 p + 2(1-q)qp + (1-p)q^2] + \\
&\quad +(1/4)[(1-q)^2 p + 2(1-q)qp + (1-p)(1-q)^2] + \\
&\quad +(1/2)[(1-q)qp + (1-q)^2 p + q^2 p + q(1-q)(1-p)] = \\
&\quad =(1/4)(1+2p) \\
EUe_{A,s} &= (1/4)p + (1/2)[(1-q)p + pq/2 + (1-p)(1-q)/2] + \\
&\quad +(1/4)[(1-q)^2 p + 2q(1-q)p + (1-q)^2(1-p)]
\end{aligned}$$

It is not difficult to verify that $EUe_{D,s}$, $EUe_{A,none}$, and $EUe_{A,s}$, all are inferior to $EUe_{D,all} = p$. If we construct a measure of welfare that includes the expert's, introducing private values as we did in this model can lead to LD being preferred. But the reason is the expert's preference for full control, not the non-experts' option of delegating to someone aligned with their own preferences. And that is because, as we point out in the text, if we deviate from a common interest scenario and recognize the existence of heterogeneous private values, we also need to acknowledge that some voters will feel misrepresented by an expert with whose preferences they disagree.

B.5.2 Introducing costly information acquisition

We focus on pure equilibria with sincere voting but now it is natural to allow for asymmetric equilibria: with pure common interest, free-riding on other voters' investment in information can be an attractive and intuitive strategy. The expert has no choice to make; a non-expert voter must choose whether to invest in information or not, and whether to vote or not. A non-expert who delegates (under LD) or abstains (under MVA) will not invest. We denote the three relevant strategies $\{xc, x, \{d \text{ or } a\}\}$where xc stands for the decision to invest and vote; x for voting without investing, and d or a for not voting and delegating (under LD) or abstaining (under MVA), and not investing. Consider LD first.

LD As always, for any $\{p, q, c\}$, there is an equilibrium where both non-experts delegate. Are there equilibria with voting?

Consider the expected utility of non-expert voter i. Suppose first that non-expert j invests and votes. Then:

$$EU_i(xc|xc) = p^3 + 3p^2(1-p) - c$$
$$EU_i(x|xc) = p^2 + 2pq(1-p)$$
$$EU_i(d|xc) = p$$

Note that $EU_i(x|xc) > EU_i(d|xc)$ always. Hence $i's$ best response to j investing and voting is always to vote, with investment in information if $c \leq \widetilde{c} \equiv 2p(1-p)(p-q)$. For any p, there is an equilibrium such that both non-experts invest in information and vote if $c \leq \widetilde{c}$.

Suppose now that non-expert j votes without investing. Then:

$$EU_i(xc|x) = p^2 + 2pq(1-p) - c$$
$$EU_i(x|x) = q^2 + 2pq(1-q)$$
$$EU_i(d|x) = p$$

There are two possibilities. If $p \leq p' = q^2/[1-2q(1-q)]$, i's best response is to vote, with investment in information if $c \leq c'_{vote} \equiv (p-q)(p+q-2pq)$. If instead $p > p'$, i's best response is to delegate if $c > c'_{del} \equiv p(1-p)(2q-1)$, and invest in information and vote otherwise.

Note that if $p \leq p'$, then $\widetilde{c} \leq c'_{vote}$; if $p > p'$, the ranking between $\widetilde{c}$ and c'_{del} depends on $\{p, q\}$. Consider first $p \leq p'$. If $c \in (\widetilde{c}, c'_{vote})$, there is an equilibrium in which one non-expert invests and one does not. If $c \geq c'_{vote}$, both non-experts vote but neither invests. Suppose now $p > p'$. Then: (i) if $c \geq c'_{del}$ neither non-expert invests, and the only equilibrium is the delegation equilibrium. (ii) If $c_{del} > \widetilde{c}$ and $c \in (\widetilde{c}, c'_{del})$, then there is an equilibrium in which one non-expert invests and one does not. As an example, suppose $p = 0.7$ and $q = 0.6$. Then $p > p'$, $\widetilde{c} = c'_{del} = 0.042$. There exists an equilibrium where both non-experts vote; both invest in information if $c \leq 0.042$, and neither invests otherwise. The probability of a correct outcome is 0.784 when both experts invest and vote, 0.696 when both vote without investing, and 0.7 when both delegate to the expert.

MVA Under MVA, a non-expert who does not vote abstains. Suppose first that non-expert j invests and votes. Then:

$$EU_i(xc|xc) = p^3 + 3p^2(1-p) - c$$
$$EU_i(x|xc) = p^2 + 2pq(1-p)$$
$$EU_i(a|xc) = p^2 + 2p(1-p)/2$$

Since $q > 1/2$, voting without investing again dominates not voting whenever the other non-expert invests and votes. Thus, as in the case of delegation, i's best response is to invest and vote if $c \leq \widetilde{c}$, and vote without investing otherwise. Thus, for any p, under MVA as well there is an equilibrium such that both non-experts invest in information and vote if $c \leq \widetilde{c}$.

Suppose now that non-expert j votes without investing. Then:

$$EU_i(xc|x) = p^2 + 2pq(1-p) - c$$
$$EU_i(x|x) = q^2 + 2pq(1-q)$$
$$EU_i(a|x) = pq + p(1-q)/2 + q(1-p)/2 = (p+q)/2$$

For all $p < 1$, voting without investing dominates abstaining. Hence i's best response is to invest and vote if $c \leq c'_{vote}$, vote without investing otherwise.

Finally, suppose that j abstains. Then:

$$EU_i(xc|a) = p^2 + 2p(1-p)/2 - c$$
$$EU_i(x|a) = pq + p(1-q)/2 + q(1-p)/2 = (p+q)/2$$
$$EU_i(a|a) = p$$

With $q < p$ and $c > 0$, abstaining is the best response to j abstaining.

It then follows that: (1) for any $\{p, q, c\}$ in the relevant range ($q \in [1/2, p]$,$p < 1$,$c > 0$) there is an equilibrium where both non-experts abstain. As in the case of LD, there is also an equilibrium where both non-experts vote, investing or not investing in information depending on c. If $c \leq \widetilde{c}$, both non-experts invest and vote; if $c \in (\widetilde{c}, c'_{vote})$, in equilibrium one non-expert invests and one does not; if $c \geq c'_{vote}$, neither invests.

Comparing LD and MVA Under both LD and MVA, for all parameter values, there is an equilibrium in which only the expert votes, and no-one invests in information. More interesting is the comparison of the equilibria with voting. For all $\{p, q\}$, under both LD and MVA there is an equilibrium in which both non-experts invest and vote if $c \leq \widetilde{c}$. Under both LD and MVA, there is an asymmetric equilibrium in which both non-experts vote but only one invests if $c \in (\widetilde{c}, c'_{vote})$, but this equilibrium exists only for $p < p'$ under LD, while it exists for all p under MVA. If $p > p'$, under LD the asymmetric equilibrium requires $c'_{del} > \widetilde{c}$ and $c \in (\widetilde{c}, c'_{del})$. However, $c'_{del} < c'_{vote}$ for all $p > p'$:[67] the asymmetric equilibrium exists for a strictly larger range of parameter values under MVA than under LD. This is the statement reported in the text: whenever parameter values are such that investment in information (either by both non-experts or by one only) can be supported in equilibrium under LD, then it can be supported under MVA. But there exists parameters ($p > p'$ and $c \in (c'_{del}, c'_{vote})$) such that there exists an equilibrium in which one non-expert invests in information under MVA but none does under LD. If $p = 0.7$ and $q = 0.6$, for example, there is an equilibrium where both non-experts invest and vote if $c \leq 0.042$ under both LD and MVA. However, if $c \in$ (0.042,0.046), under MVA there is an equilibrium in which both non-experts vote and one of them invests in information; under LD in the only equilibrium the expert is dictator and none of the non-experts invests. Expected welfare (the sum of expected utilities) is higher under MVA.

Note also, however, that if $p > p'$ and $c > c'_{del}$, under LD, delegation is the only equilibrium. Under MVA, the equilibrium in which both non-experts abstain exists as well, but

[67] The threshold c'_{del} is decreasing in p, while c'_{vote} is increasing in p. It then follows that $(c'_{vote} - c'_{del})$ is minimal at minimum p, that is, at $p = p'$. But $c'_{vote} = c'_{del} = ((1-q)^2 q^2 (2q-1))/(1-2(1-q)q)^2$ at $p = p'$. The statement then follows.

so does a second equilibrium in which both vote without investing. The range of parameter values supporting the equilibrium with voting without information acquisition is strictly larger under MVA than under LD.

B.5.3 Introducing correlated signals

Once again, we study the pure common interest model with $N = 3$. We now suppose that all signals are conditionally independent with probability $\alpha < 1$; with probability $(1 - \alpha)$, the two non-experts receive the same signal, of precision q, while the expert alone has a conditionally independent signal, of precision $p \in (q, 1)$. The expert thus benefits both from higher signal precision, and from the signal's (conditional) independence. The probability α is known, but it is not known whether the realized non-experts' signals are correlated (fully, in this example). We focus on pure symmetric equilibria with no communication and with sincere voting, where all votes are always cast in line with the received signal, and non-experts' strategies are $\{x, d/a\}$–that is, whether to vote or to delegate/abstain. Consider first LD.

LD As always, for all parameter values there is an equilibrium in which both non-experts delegate, and $EU_i(d|d) = p$. Is there an equilibrium in which the non-experts do not delegate? Suppose j does not delegate. Then:

$$EU_i(d|x_j) = p$$
$$EU_i(x|x_j) = (1 - \alpha)q + \alpha[pq^2 + 2pq(1 - q) + q^2(1 - p)]$$

Note that if the non-experts' signals are fully correlated (with probability $(1 - \alpha)$) and votes are sincere, the non-experts always vote in the same direction and always prevail; the correct decision is taken if the non-experts' common signal is correct, that is, with probability q.

From the expressions above, non-expert i's best response to j voting is to vote if:

$$\alpha(q^2 + 2pq(1 - q) - p) \geq (1 - \alpha)(p - q)$$

and delegate otherwise. It then follows that there exists $\widetilde{q}(p, \alpha)$ such that if $q < \widetilde{q}$ LD admits a unique equilibrium where both non-experts always delegate; if $q \geq \widetilde{q}$, then under LD there are two equilibria: the equilibrium with full delegation and a second equilibrium in which the two non-experts cast their votes.[68]

MVA Consider now MVA. Suppose first that j abstains. Then:

$$EU_i(a|a_j) = p$$
$$EU_i(x|a_j) = [pq + 1/2[(1-p)q + p(1-q)] = (1/2)(p+q)$$

With $p > q$, abstaining is the superior response. Thus, for all relevant parameter values, there is always an equilibrium in which both non-experts abstain.

Suppose now that j votes. Then:

$$EU_i(a|x_j) = (1/2)(p+q)$$
$$EU_i(x|x_j) = (1-\alpha)q + \alpha[pq^2 + 2pq(1-q) + q^2(1-p)]$$

If both non-experts vote, the probability of reaching the correct decision, and thus $EU_i(x|x_j)$, must be identical under LD or MVA; if only one votes, however, such probability is lower under MVA than under LD because abstention does not benefit as much as delegation from the expert's superior signal. It follows that supporting the equilibrium in which both non-experts vote is easier under MVA than under LD–that is, there is a larger range of parameter values for which $\{x, x\}$ is an equilibrium. Under MVA, non-expert i's best response to j voting is to vote if:

$$\alpha\left(q^2 + 2pq(1-q) - \frac{(p+q)}{2}\right) \geq (1-\alpha)\left(\frac{p-q}{2}\right)$$

and abstain otherwise. There exists $q'(p,\alpha)$ such that if $q<q'(p,\alpha)$ MVA admits a unique equilibrium where both non-experts always abstain; if $q \geq q'$, there are two equilibria: the equilibrium with full abstention and a second equilibrium in which the two non-experts cast their votes. Note that for any given (p,α), $q' < \widetilde{q}$.[69]

[68] The threshold $\widetilde{q}(p,\alpha)$ is increasing in p and decreasing in α. It is given by:

$$\widetilde{q}(p,\alpha) = \frac{1 + \alpha(2p-1) - \sqrt{1-(2-\alpha)\alpha(2p-1)^2}}{2\alpha(2p-1)}.$$

[69] The threshold $q'(p,\alpha)$, again increasing in p and decreasing in α, is given by:

$$q'(p,\alpha) = \frac{1 + 2\alpha(2p-1) - \sqrt{1-4(1-\alpha)\alpha(2p-1)^2}}{4\alpha(2p-1)}.$$

Comparing LD and MVA Thus, both LD and MVA support, for all parameter values, an equilibrium in which the expert alone dictates the decision (non-experts do not vote). Both voting systems also support an equilibrium in which both experts vote, but only for q high enough, higher than $\widetilde{q}(p, \alpha)$ in the case of LD, and higher than $q'(p, \alpha)$ in the case of MVA, with $q' < \widetilde{q}$. Comparing the two equilibria, the first equilibrium, with the expert dictating, is superior if $q < \widetilde{q}(p, \alpha)$, while the second, with the non-experts voting, is superior for $q > \widetilde{q}(p, \alpha)$. It follows, as stated in the text, that both voting systems can support the superior equilibrium for any parameter values. However, both systems also can support delegation/abstention when the probability of reaching the correct decision would be higher if the non-experts voted. Finally, there exists a range of parameters ($q \in (q'(p, \alpha), \widetilde{q}(p, \alpha))$ such that MVA but not LD can support voting by non-experts when the expert alone dictating would be superior.