Discussion Paper Series A No.674

Collective Mistakes: Intuition Aggregation for a Trick Question under Strategic Voting

Tomoya Tajika

July 2018

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan

Collective mistakes: Intuition aggregation for a trick question under strategic voting*

Tomoya Tajika†

July 24, 2018

We consider a situation in which voters collectively answer a binary question. Each voter obtains an intuition about the answer to the question, but whether the question is intuitive or counterintuitive is not known to any voter. If each voter receives an independent signal on whether the question is intuitive or not, the majority rule under sincere voting correctly aggregates the intuitions with a large electorate; however, it is not an equilibrium. We show that in a unique pure-strategy equilibrium with a large electorate, majority voting makes an incorrect decision with a probability that can be sufficiently close to 1.

Keyword: Information aggregation; inefficiency; counterintuitive question; strategic voting
JEL Classification: C72; D72.

*The author is very grateful to Naruto Nagaoka for his valuable comments. The author also thanks Eiichi Miyagawa, Norihito Sakamoto, and Reiko Gotoh. The usual disclaimer applies.
†Institute of Economic Research, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo, 186-8603. Email: `tajika@ier.hit-u.ac.jp`

1. Introduction

In mathematics and science, counterintuitive questions are ubiquitous. The Monty Hall problem is a well-known example (Selvin, 1975).[1] An American public figure, Marilyn vos Savant published this puzzle in her column of a magazine (vos Savant, 1990). After she gave the correct answer, 10,000 letters were received, the great majority disagreeing with the correct answer. The number of critical letters from PhDs (mostly in mathematics or science departments) were close to 1,000 (Tierney, 1991). For such questions, our intuition is likely to be wrong, and, more unfortunately, we are not sure in advance whether the answer is counterintuitive or not. A majority of the people would produce the wrong answer; thus, answering by majority voting seems to lead to the wrong answer. On the contrary, a vast body of studies on information aggregation argues the effectiveness of collective wisdom. The well-known Condorcet jury theorem shows that if each voter obtains independent information, the majority voting makes the correct decision asymptotically.

Given these two conflicting views, the current study models the situation and examines whether majority voting correctly aggregates voters' information. In our model, voters face a binary question that has a unique answer. Each voter has an independent intuition about the answer to the question. The answer to the question can be intuitive or counterintuitive. While intuition is likely to be correct for an intuitive question, it is likely to be incorrect for a counterintuitive question. We also assume that when the question is counterintuitive, the informativeness of the intuition is weaker than that when the question is intuitive. Although each voter is unaware of whether the question is counterintuitive or not, each voter independently receives an informative signal on whether the question is counterintuitive or not. We call this signal a "guess" about the complexity of the question. Therefore, each voter receives an informative signal about the answer.

In our framework, we show that with a large electorate, although the majority rule under sincere voting chooses the correct answer with certainty, it is not an equilibrium. In a unique

[1] The problem is the following one. One of three boxes, A, B and C contains a key and the others are empty. I know which box contains the key. If you choose the box that contains the key, you win. Suppose that you choose B. Then, I open a remaining empty box, say C. The question is "Should you change your choice to A?" The answer is yes because now the probability with which box A contains the key is 2/3.

pure-strategy nontrivial equilibrium, each voter votes against his gained intuition irrelevant to his guess. Although voting against intuition is the optimal behavior for a counterintuitive question, it is the worst possible behavior for an intuitive question. This result remains robust even when the probability of the question being intuitive is sufficiently close to 1. Therefore, in such a situation, the majority voting chooses the incorrect alternative with a probability that is sufficiently close to 1. Surprisingly, this result is also independent of the informativeness of the signal on whether the question is intuitive or not. Note that the standard Condorcet's setting is the limit of our model with respect to the informativeness of the signal. Hence, our result implies a fragility in the Condorcet jury theorem.

The intuition underlying our result is the following. Note that under strategic voting equilibria, each voter cares only about the event in which he or she is pivotal. Under the voters' responsive behaviors to their intuition, as the intuition is more informative when the question is intuitive, for the intuitive question, the alternative is more likely to win by a large margin. Therefore, when the question is intuitive, a given voter is less likely to be pivotal, and in turn, each voter cares more about the case when the question is counterintuitive. Therefore, such a contrarian behavior is the best response.

Although the voting equilibrium is inefficient under pure strategy, by allowing mixed strategy, we can construct an efficient equilibrium under which the correct answer is chosen with probability one with a large electorate. Owing to this equilibrium multiplicity, an issue of equilibrium selection arises. On this point, while voting against intuition is robust against any small perturbation, this efficient equilibrium is not. In this respect, voting against intuition is more plausible than this efficient equilibrium.

Our setting is suitable for highly professional problems such as scientific or technical questions because it is difficult for a layman to judge with certainty whether such questions are intuitive or not. Some of managerial decisions such as choices of product design, production process, price and sales channel are also examples of such problems. Our result implies the possibility of an incorrect decision by majority rule that occurs with a sufficiently high probability, in which case, the majority rule has no superiority over an individual's decision. Therefore, for such a question, majority voting is a bad choice.

Contribution to the literature In the literature, many studies extend the Condorcet jury theorem to game theoretic environments (Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1996, 1997; McLennan, 1998; Wit, 1998; Duggan and Martinelli, 2001; McMurray, 2013). They examine whether voting rules satisfy *full information equivalence (FIE)*. This means that as the number of voters goes to infinity, the probability of the voting outcome under a strategic voting equilibrium coinciding with that under complete information converges to 1.

On the contrary, there are many studies that also show sufficient conditions under which a majority rule violates FIE. While many such studies consider the case with heterogeneous preferences (Kim and Fey, 2007; Bhattacharya, 2013a,b; Acharya, 2016; Ali, Mihm and Siga, 2017; Tajika, 2018), as in our model, a few studies consider the case with homogeneous preferences (Martinelli, 2006; Mandler, 2012; Ellis, 2016; Barelli, Bhattacharya, Siga, 2018). Among these studies, Mandler (2012) is the closest study to ours in the following respect. In our model, we consider the precision of intuition to be uncertain, as whether a question is intuitive or not is uncertain. Mandler (2012) proposes a model describing precision uncertainty of signals in a way that is different from our method. Mandler (2012) also shows the possibility of majority voting violating FIE. However, in contrast to our result, Mandler (2012) does not discuss the probability of majority voting choosing the wrong alternative. In our model, we show that the probability of the incorrect alternative being chosen can be sufficiently close to 1. This is the main contribution of our study.

Moreover, in Mandler (2012)'s model, each voter can receive only one signal, which means that each voter receives no information about the signal precision. Related to this point, Barelli, Bhattacharya, Siga (2018) provide a necessary and sufficient condition for FIE in a general common value voting game. In their environment, when the majority rule violates FIE, voters' information are necessary to be insufficient. In our model, in contrast to their setting, voters receive sufficiently informative signals such that if they are correctly aggregated, majority voting makes the correct decision. Even in this setting, we find an equilibrium in which the decision by majority voting leads to the wrong answer with sufficiently high probability.

Except for Ali, Mihm and Siga (2017) and Tajika (2018), most of the above studies only show the possibility of a violation in FIE. Although Ali, Mihm and Siga (2017) and Tajika (2018) show that majority voting chooses a "bad" alternative with a probability that is sufficiently close to 1, they assume heterogeneity in the voters' preferences.

2. Model

An odd number of voters face a binary question that has one correct answer. N denotes the set of voters and $|N| = 2n+1$, $n \in \mathbb{N}$. The *answer* of the question is denoted by $\omega \in \{A, B\}$, which is unknown to any voter. If the voters choose the correct answer, they receive a payoff of 1. Otherwise, their payoff is 0. By observing the question, each voter i independently obtains an *intuition* about the answer to the question, which is denoted by $\sigma_i \in \{A, B\}$. This intuition is an informative signal about the answer. However, the question can be intuitive or counterintuitive. If the question is *straightforward* (denoted by $\psi = S$), the answer is intuitive and, thus, the intuition is likely to be right. Formally, for each $i \in N$, $\Pr(\sigma_i = \omega \mid \psi = S) = \theta^* > 1/2$. On the contrary, if the question is *complicated* (denoted by $\psi = C$), the answer is counterintuitive. Formally, for each $i \in N$, $\Pr(\sigma_i = \omega \mid \psi = C) = \theta_* < 1/2$. The term ψ is referred to as the *complexity* of the question. We assume that $\theta^* > 1 - \theta_* > 1/2$. This inequality implies that the intuition is less informative for the complicated question. Although the complexity of the question is unknown to any voter, each voter makes a *guess* about the complexity of the question, denoted by $\gamma_i \in \{C, S\}$. The guess is an informative signal for the complexity of the question. For each $i \in N$, $\Pr(\gamma_i = \psi) = q \in [1/2, 1)$. We refer to the tuple $(\omega, \psi) \in \{A, B\} \times \{C, S\}$ as the *state* and the tuple $(\sigma_i, \gamma_i) \in \{A, B\} \times \{C, S\}$ as the *signal* of voter $i \in N$. Last, let $p_\omega = \Pr(\omega = A) \in (0, 1)$ and $p_\psi = \Pr(\psi = S) \in (0, 1)$ be the priors. Assume that ω and ψ have no correlation. We also assume that $p_\omega = 1/2$ for simplicity.

The collective decision is made by a simple majority vote. No abstention is allowed.

3. Equilibrium

Regarding the equilibrium concept, we employ the type-symmetric Bayesian Nash equilibria (abbreviated by BNE).

3.1. Best responses

Consider voter $i \in N$'s behavior. Let n_A be the number of other voters voting for A. Now the best response is summarized in the following lemma.

Lemma 1. *For voter i who receives (σ_i, γ_i), voting for alternative A is optimal if and only if $Q(\sigma_i, \gamma_i) \geqslant 1$, where*

$$Q(A, C) := \frac{(1-q)p_\psi \theta^* \Pr(n_A = n \mid A, S) + q(1-p_\psi)\theta_* \Pr(n_A = n \mid A, C)}{(1-q)p_\psi(1-\theta^*)\Pr(n_A = n \mid B, S) + q(1-p_\psi)(1-\theta_*)\Pr(n_A = n \mid B, C)}, \tag{1}$$

$$Q(A, S) := \frac{qp_\psi \theta^* \Pr(n_A = n \mid A, S) + (1-q)(1-p_\psi)\theta_* \Pr(n_A = n \mid A, C)}{qp_\psi(1-\theta^*)\Pr(n_A = n \mid B, S) + (1-q)(1-p_\psi)(1-\theta_*)\Pr(n_A = n \mid B, C)}, \tag{2}$$

$$Q(B, C) := \frac{(1-q)p_\psi(1-\theta^*)\Pr(n_A = n \mid A, S) + q(1-p_\psi)(1-\theta_*)\Pr(n_A = n \mid A, C)}{(1-q)p_\psi \theta^* \Pr(n_A = n \mid B, S) + q(1-p_\psi)\theta_* \Pr(n_A = n \mid B, C)}, \tag{3}$$

$$Q(B, S) := \frac{qp_\psi(1-\theta^*)\Pr(n_A = n \mid A, S) + (1-q)(1-p_\psi)(1-\theta_*)\Pr(n_A = n \mid A, C)}{qp_\psi \theta^* \Pr(n_A = n \mid B, S) + (1-q)(1-p_\psi)\theta_* \Pr(n_A = n \mid B, C)}. \tag{4}$$

Proof. Omitted. □

Let $v_A(\omega, \psi)$ be the probability that a given voter votes for alternative A under state $(\omega, \psi) \in \{A, B\} \times \{C, S\}$. Let $w_A(\sigma_i, \gamma_i)$ be the probability that a given voter who receives signal $(\sigma_i, \gamma_i) \in \{A, B\} \times \{C, S\}$ votes for alternative A. Then, each voting probability is calculated as follows:

$$v_A(A, C) = q\theta_* w_A(A, C) + q(1-\theta_*)w_A(B, C) + (1-q)\theta_* w_A(A, S) + (1-q)(1-\theta_*)w_A(B, S),$$

$$v_A(A, S) = (1-q)\theta^* w_A(A, C) + (1-q)(1-\theta^*)w_A(B, C) + q\theta^* w_A(A, S) + q(1-\theta^*)w_A(B, S),$$

$$v_A(B, C) = q(1-\theta_*)w_A(A, C) + q\theta_* w_A(B, C) + (1-q)(1-\theta_*)w_A(A, S) + (1-q)\theta_* w_A(B, S),$$

$$v_A(B, S) = (1-q)(1-\theta^*)w_A(A, C) + (1-q)\theta^* w_A(B, C) + q(1-\theta^*)w_A(A, S) + q\theta^* w_A(B, S).$$

3.2. Sincere voting

We first consider the sincere voting strategy in which each voter acts as if there is only one voter (Austen-Smith and Banks, 1996). As the first case, assume that $p_\psi = 1/2$. In this case, each voter's behaves in the following manner: If $\gamma_i = S$, the voter votes for A if $\sigma_i = A$; else, the voter votes for B. If $\gamma_i = C$, the voter behaves conversely. That is, if a voter makes a guess considering the question to be intuitive, he votes according to the intuition. If a voter makes a guess considering the question to be counterintuitive, he votes against the intuition. Then, the voting probability is calculated as follows:

$$\begin{aligned}
v_A(A, C) &= q(1-\theta_*) + (1-q)\theta_*, \\
v_A(A, S) &= (1-q)(1-\theta^*) + q\theta^*, \\
v_A(B, C) &= q\theta_* + (1-q)(1-\theta_*), \\
v_A(B, S) &= (1-q)\theta^* + q(1-\theta^*).
\end{aligned}$$

If $q > 1/2$, $v_A(A, S) = 1 - v_A(B, S) > 1 - v_A(B, C) = v_A(A, C) > 1/2$. Therefore, under sincere voting, information is correctly aggregated for sufficiently large n. Indeed, since $v_A(A, \psi) > 1/2 > v_A(B, \psi)$ for each $\psi \in \{C, S\}$, $\lim_{n\to\infty} \Pr(d = \omega) = 1$ ($d \in \{A, B\}$ is the decision by the majority voting).

Let us check whether the strategy profile is a BNE. Note that for each ω and ψ,

$$\Pr(n_A = n \mid \omega, \psi) = \binom{2n}{n}[v_A(\omega, \psi)(1 - v_A(\omega, \psi))]^n.$$

Under the sincere voting strategy, $\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S)$. Also note that in this case, for each $(j, k) \in \{A, B\}^2$, $\lim_{n\to\infty} \frac{\Pr(n_A = n \mid j, C)}{\Pr(n_A = n \mid k, S)} = \infty$. Therefore, for sufficiently large n,

$$\begin{aligned}
\lim_{n\to\infty} Q(A, S) = \lim_{n\to\infty} Q(A, C) &= \frac{\theta_*}{1-\theta_*} < 1, \\
\lim_{n\to\infty} Q(B, S) = \lim_{n\to\infty} Q(B, C) &= \frac{1-\theta_*}{\theta_*} > 1.
\end{aligned}$$

Then, by Lemma 1, each voter's best response is voting against the intuition, irrelevant to his or her guess, which is in contradiction to sincere voting.

Proposition 1. *Assume that $q > 1/2$ and $p_\psi = 1/2$. Then, there is $\bar{n} \in \mathbb{N}$ such that for each $n \geqslant \bar{n}$, the sincere voting is not a BNE.*

The intuition underlying Proposition 1 works as follows. Recall that the intuition is less informative if the question is complicated. Due to this assumption, by voting in response to the intuition, when the question is complicated, the voter is more likely to be pivotal than when the question is straightforward. In a strategic voting equilibrium, each voter cares only about the situation in which the voter is pivotal. Therefore, irrelevant to the guess, voting as if the question is complicated is optimal for each voter. For a complicated question, voting against the intuition is optimal.

3.3. Contrarian equilibrium

Now we examine other strategies. As observed in the previous section, voting against the intuition is the best response against sincere voting. We check whether it is a BNE. Under this strategy profile, $v_A(A, C) = 1 - \theta_*$, $v_A(A, S) = 1 - \theta^*$, $v_A(B, C) = \theta_*$, and $v_A(B, S) = \theta^*$. Then, we have that $v_A(A, S) = 1 - v_A(B, S) < 1 - v_A(A, C) = v_A(B, C) < 1/2$. This implies that $\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S)$ and therefore,

$$\lim_{n\to\infty} Q(A, S) = \lim_{n\to\infty} Q(A, C) = \frac{\theta_*}{1 - \theta_*} < 1,$$
$$\lim_{n\to\infty} Q(B, S) = \lim_{n\to\infty} Q(B, C) = \frac{1 - \theta_*}{\theta_*} > 1.$$

As in the previous case, voting against the intuition is the best response, and therefore, it is a BNE.

Proposition 2 (Contrarian equilibrium)**.** *For each $q \in [1/2, 1)$ and each $p_\psi \in (0, 1)$, there is $\bar{n} \in \mathbb{N}$ such that for each $n \geqslant \bar{n}$, voting against the intuition is a BNE.*

Under the strategy of voting against intuition, the majority voting correctly aggregates information if the question is complicated, but it fails in correct information aggregation if the question is straightforward. Indeed, under this strategy profile, $\lim_{n\to\infty} \Pr(d = \omega \mid \psi = C) = \lim_{n\to\infty} \Pr(d \neq \omega \mid \psi = S) = 1$. This implies that the full information equivalence fails if the question is straightforward, which is in contrast to the classical Condorcet jury theorem. Surprisingly, this result does not depend on the value of q. If $q = 1$, our model is reduced to a classical Condorcet's setting, wherein the jury theorem holds. This implies a fragility of the jury theorem. Proposition 2 is also independent of the value of p_ψ. In the limit, as the probability of the majority voting making a correct decision is $p_\psi \lim_{n\to\infty} \Pr(d = \omega \mid \psi = S) + (1 - p_\psi) \lim_{n\to\infty} \Pr(d = \omega \mid \psi = C) = 1 - p_\psi$, we have the following corollary:

Corollary 1. *As $p_\psi \to 1$ and $n \to \infty$, there is a sequence of BNEs such that the probability of the incorrect alternative being chosen converges to* 1.

Corollary 1 implies that the majority voting achieves the worst outcome in the limit. In this case, the majority voting is inferior not only to a single voter's decision but also to a decision made by rolling a die. More unfortunately, among the nontrivial pure-strategy BNEs, which are defined below,[2] voting against intuition is a unique equilibrium.

Definition 1. A BNE is *nontrivial* if $\Pr(n_A = n \mid \omega, \psi) > 0$ for some ω, ψ at the BNE.

Theorem 1. *For sufficiently large n and sufficiently large $p_\psi < 1$, voting against intuition is a unique pure-strategy nontrivial BNE.*

Proof. See Appendix A. □

Remark 1 (Informativeness of the intuition). One of the crucial assumptions of our main result is $\theta^* > 1 - \theta_* > 1/2$. Under this assumption, in any strategic voting equilibrium, voters care only about the case when the question is complicated and, therefore, each voter votes against the intuition. If the converse inequality holds, that is, $1 - \theta_* > \theta^* > 1/2$, the converse result holds: Voting according to the intuition is a BNE. This result also does not depend on the values of q and p_ψ. Therefore, as $p_\psi \to 0$ and $n \to \infty$, in the unique BNE, the probability that the incorrect alternative wins also converges to 1.

3.4. Efficient equilibrium

Although our focus in the previous sections is pure strategies, by allowing mixed strategies, we can show that an efficient equilibrium exists. Under such a strategy profile, the statement

[2] In the majority voting, the strategy profile such that all voters vote for the same alternative is trivially an equilibrium, in which case, $\Pr(n_A = n \mid \omega, \psi) = 0$ for each ω and ψ. We exclude such equilibria as they are trivial.

of the classical Condorcet jury theorem is valid.

Proposition 3 (Condorcet jury theorem)**.** *If $q > 1/2$, there exists a sequence of mixed strategy BNEs such that the probability that the majority rule chooses the correct answer converges to one.*

Proof of Proposition 3. We will show that there exists a BNE such that each voter mixes if $\gamma = S$ and votes against intuition if $\gamma = C$.

First, under the following conditions, $Q(A, S) = Q(B, S) = 1$.

$$\frac{\Pr(n_A = n \mid A, S)}{\Pr(n_A = n \mid B, C)} = \frac{\Pr(n_A = n \mid B, S)}{\Pr(n_A = n \mid B, C)} = \frac{1-q}{q}\frac{1-p_\psi}{p_\psi}\frac{1-2\theta_*}{2\theta^*-1}, \tag{5}$$

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C). \tag{6}$$

Then, we have that

$$Q(A, C) = \frac{(1-q)^2\theta^*(1-2\theta_*) + q^2\theta_*(2\theta^*-1)}{(1-q)^2(1-\theta^*)(1-2\theta_*) + q^2(1-\theta_*)(2\theta^*-1)}$$

$$Q(B, C) = \frac{1}{Q(A, C)}.$$

One can show that $Q(A, C) < 1 < Q(B, C)$, and thus, by Lemma 1, $w_A(B, C) = 1 - w_A(A, C) = 1$. Then, $\Pr(n_A = n \mid A, \psi) = \Pr(n_A = n \mid B, \psi)$ if $w_A(A, S) = 1 - w_B(A, S)$. Moreover, if $w_A(A, S)$ is the following value, $v_A(A, C) = v_A(A, S) > 1/2$.

$$w_A(A, S) = \frac{\theta^* - \theta_*}{q(2\theta^*-1) + (1-q)(1-2\theta_*)} \in (0, 1). \tag{7}$$

Therefore, for each n, there is a $w^n = w_A(A, S)$ that satisfies (5) and (6) and w^n converges to the value of the right hand side of (7). Thus, taking a mixed strategy is optimal for the voters such that $\gamma = S$. Now the considered strategy is a BNE.

Moreover, if $n \to \infty$, $v_A(A, C) = v_A(A, S) > 1/2$ and $v_A(A, \psi) = 1 - v_A(B, \psi)$, the probability that the correct answer is chooses converges to 1. □

The intuition for our strategy is the following. Under sincere voting that considered in Section 3.2, under state $\psi = C$, a given voter is more likely to be pivotal. In the equilibrium strategy, voter who receive $\gamma = S$ mixes voting for intuition and voting against intuition. Then, by doing so, under state $\psi = S$, the probability that a given voter is pivotal is the same that under state $\psi = C$. Now the source of incentive for voting against intuition disappears.

3.5. Instability of the efficient equilibrium

Although the mixed strategy equilibrium given in the proof of Proposition 3 is efficient, remark that this strategy profile is vulnerable to a small mistake. To see this, we define a concept of stability.

Definition 2. Let $BR\colon [0,1]^{\{A,B\}\times\{C,S\}} \to 2^{[0,1]^{\{A,B\}\times\{C,S\}}} \setminus \{\emptyset\}$ be the best response profile correspondence. Symmetric voting probability profile $w = (w_A(t))_{t\in\{A,B\}\times\{C,S\}}$ is *asymptotically stable* if there is an $\bar{\varepsilon} > 0$ such that for each $\varepsilon < \bar{\varepsilon}$ and each $w' \in B_\varepsilon(w)$,[3] there is a sequence of strategy profile $(w^k)_{k\in\mathbb{N}}$ such that $w^0 = w'$, $w^k \in BR(w^{k-1})$ and $\lim_{k\to\infty} w^k \to w$.

Therefore, under asymptotically stable strategies, even if a person makes a (small) mistake, a process of repeated adoption of best responses leads to the original strategy profile.

We show that the mixed strategy equilibrium given in the proof of Proposition 3 is not asymptotically stable. Let $w = (w_A(A,S), w_A(B,S), w_A(A,C), w_A(B,C))$ be the equilibrium voting probability. As see in the proof of Proposition 3, (5), (6) and (7) hold. Consider mistake w' that satisfies

$$w' = (w_A(A,S) - \varepsilon, w_A(B,S) + \varepsilon, w_A(A,C), w_A(B,C))$$

for small $\varepsilon > 0$. Then, while $v_A(A,\psi) = 1 - v_A(B,\psi) > 1/2$ for each ψ and sufficiently small $\varepsilon > 0$, $v_A(A,S)$ decreases and $v_A(A,C)$ increases. This implies that $\frac{P(n_A|i,S)}{P(n_A|j,C)}$ increases for each $i,j \in \{A,B\}$. Then, since under the original strategy profile, $Q(A,S) = Q(B,S) = 1$, under the new strategy profile, $Q(A,S) > 1 > Q(B,S)$. On the other hand, if ε is sufficiently small, even under the new profile $Q(A,C) < 1 < Q(B,C)$. Therefore, for such mistake, $(w_A(A,S), w_A(B,S), w_A(A,C), w_A(B,C)) = (1,0,0,1)$ is the best response. As we have seen in

[3] $B_\varepsilon(w)$ denotes the ε-neighborhood of point $w \in [0,1]^{\{A,B\}\times\{C,S\}}$.

Section 3.2, we know that this profile is not an equilibrium. Against this strategy profile, voting against intuition is the best response. Now since voting against intuition is an equilibrium, this process never leads to the original profile.

On the contrary, voting against intuition is asymptotically stable. This is because, under this strategy profile, $Q(A, S) < 1 < Q(B, S)$ and $Q(A, C) < 1 < Q(B, C)$. Since each $Q(t)$, $t \in \{A, B\} \times \{C, S\}$ is continuous in $(w_A(t'))_{t' \in \{A,B\} \times \{C,S\}}$, for each small mistake, the relation $Q(A, S) < 1 < Q(B, S)$ and $Q(A, C) < 1 < Q(B, C)$ remains to hold. Now voting against intuition is the best response. In this respect, voting against intuition is more plausible than the efficient equilibrium.

4. Conclusion

We studied the collective decision by a majority voting in responding to a question framed such that no voter knew whether it is intuitive or counterintuitive. Although sincere voting correctly aggregates the intuitions, it is not an equilibrium. Instead, voting against an intuition ignoring the signal about the complexity of the question is a unique pure-strategy equilibrium. This outcome is robust even when the prior probability of the question being complicated is sufficiently small and, therefore, it is possible that the incorrect alternative wins with a probability that is sufficiently close to 1. Our result implies the inferiority of the majority voting over an individual decision, which is in contrast to the Condorcet jury theorem. On the other hand, by allowing mixed strategies, our voting game has an efficient equilibrium. Owing to the equilibrium multiplicity, whether the equilibrium outcome is efficient or not depends on the equilibrium selection. On this point, we show that although voting against intuition is robust to any small mistake, the efficient mixed strategy equilibrium is not. In this respect, voting against intuition is a plausible equilibrium.

For highly professional questions, such as scientific or technical questions, it is difficult for a layman to infer whether the answer to the question is intuitive or counterintuitive. The finding from our contribution is that majority voting by laymen should not be used to answer such questions even when they obtain informative intuitions and make a reasonable guess about the complexity of the question.

As a remedy, one may consider a majority voting both for the intuition and for guessing about the complexity of the question. However, voters may also face uncertainty about the precision of their guess. In this case, although voters may receive informative signals about the precision of their guesses, they may also face uncertainty about the informativeness of the signal. This infinite regress may force the voters to vote for an infinite number of issues, but it is impossible to do so in reality. To deal with this problem is left for future research.

A. Omitted proof

Proof of Theorem 1. In Proposition 2, we have shown that voting against the intuition is a BNE, in which case, $\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S)$. We consider all the remaining patterns of the magnitude relations of $\Pr(n_A = n \mid \omega, \psi)$ and show that each case has a contradiction.

Case 1. $\max\{\Pr(n_A = n \mid A, C), \Pr(n_A = n \mid A, S)\} > \max\{\Pr(n_A = n \mid B, C), \Pr(n_A = n \mid B, S)\} = 0$. In this case, for sufficiently large n, $Q(\omega, \psi) > 1$ for each ω and ψ. Then, $v_A(\sigma_i, \gamma_i) = 1$ for each σ_i and γ_i, which implies that

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0.$$

This is a contradiction.

Case 2. $\max\{\Pr(n_A = n \mid A, C), \Pr(n_A = n \mid A, S)\} < \max\{\Pr(n_A = n \mid B, C), \Pr(n_A = n \mid B, S)\}$. As in case 1, we have

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0,$$

which is a contradiction.

Case 3. $\Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S) > \max\{\Pr(n_A = n \mid A, C), \Pr(n_A = n \mid B, C)\}$. In this case, for sufficiently large n, $Q(A, \psi) > 1/2 > Q(B, \psi)$ for each ψ. Therefore,

$v_A(A, C) = \theta_*$, $v_A(A, S) = \theta^*$, $v_A(B, C) = 1 - \theta_*$ and $v_A(B, S) = 1 - \theta^*$. Then, we have that

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S),$$

which is a contradiction.

Case 4. $\Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) > \max\{\Pr(n_A = n \mid A, C), \Pr(n_A = n \mid B, S)\}$. Then, as $n \to \infty$,

$$\begin{aligned} Q(A, C) &= \frac{(1-q)p_\psi \theta^*}{q(1-p_\psi)(1-\theta_*)} \\ Q(A, S) &= \frac{qp_\psi \theta^*}{(1-q)(1-p_\psi)(1-\theta_*)}, \\ Q(B, C) &= \frac{(1-q)p_\psi(1-\theta^*)}{q(1-p_\psi)\theta_*}, \\ Q(B, S) &= \frac{qp_\psi(1-\theta^*)}{(1-q)(1-p_\psi)\theta_*}. \end{aligned}$$

Then, for sufficiently large p_ψ, $Q(\omega, \psi) > 1$ for each ω and ψ. As in case 1,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0,$$

which is a contradiction.

Case 5. $\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, S) > \max\{\Pr(n_A = n \mid A, S), \Pr(n_A = n \mid B, C)\}$. As in case 4, for sufficiently large n and p_ψ, $Q(\omega, \psi) < 1$ for each ω and ψ. Then, we have

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0,$$

which is a contradiction.

Case 6. $\Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid A, C) > \Pr(n_A = n \mid B, S)$. In this case, for sufficiently large n and p_ψ, $Q(A, \psi) > 1/2 > Q(B, \psi)$ for each ψ. As in case 3,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S),$$

which is a contradiction.

Case 7. $\Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S) = \Pr(n_A = n \mid A, C) > \Pr(n_A = n \mid B, C)$. Then, for sufficiently large n and p_ψ, $Q(\omega, \psi) > 1$ for each ω and ψ. As in case 1,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0,$$

which is a contradiction.

Case 8. $\Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = \Pr(n_A = n \mid A, C) > \Pr(n_A = n \mid A, S)$. Then, for sufficiently large n and p_ψ, $Q(\omega, \psi) < 1$ for each ω and ψ. Then,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = 0,$$

which is a contradiction.

Case 9. $\Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = \Pr(n_A = n \mid A, S) > \Pr(n_A = n \mid A, C)$. Then, for sufficiently large n and p_ψ, $Q(A, \psi) > 1/2 > Q(B, \psi)$ for each ψ. As in case 3,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S),$$

which is a contradiction.

Case 10. $\Pr(n_A = n \mid B, C) = \Pr(n_A = n \mid B, S) = \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid A, C)$. Then, for sufficiently large n and p_ψ, $Q(A, \psi) > 1/2 > Q(B, \psi)$ for each ψ. As in case 3,

$$\Pr(n_A = n \mid A, C) = \Pr(n_A = n \mid B, C) > \Pr(n_A = n \mid A, S) = \Pr(n_A = n \mid B, S),$$

which is a contradiction.

□

References

Acharya, Avidit (2016) "Information aggregation failure in a model of social mobility," *Games and Economic Behavior*, 100, 257–272.

Ali, S. Nageeb, Maximilian Mihm and Lucas Siga (2017) "The perverse politics of polarization," *Working paper*.

Austen-Smith, David and Jeffery S. Banks (1996) "Information aggregation, rationality and the Condorcet jury theorem," *American Political Science Review*, 90, 34–45.

Barelli, Paulo, Sourav Bhattacharya and Lucas Siga (2018) "Full information equivalence in large election," *Working paper*.

Bhattacharya, Sourav (2013a) "Preference monotonicity and information aggregation in election," *Econometrica*, 81 1229–1247.

Bhattacharya, Sourav (2013b) "Condorcet jury theorem in a spatial model of elections," *Working paper*.

Duggan, John and César Martinelli (2001) "A Bayesian model of voting in juries," *Games and Economic Behavior,* 37, 259–294.

Ellis, Andrew (2016) "Condorcet meets Ellsberg," *Theoretical Economics*, 11, 865–895.

Feddersen, Timothy and Wolfgang Pesendorfer (1996) "The swing voter's curse," *American Economic Review*, 86, 408–424.

Feddersen, Timothy and Wolfgang Pesendorfer (1997) "Voting Behavior and Information Aggregation in Elections With Private Information," *Econometrica*, 65, 1029–1058.

Kim, Jaehoon and Mark Fey (2007) "The swing voter's curse with adversarial preferences," *Journal of Economic Theory*, 135, 236–252.

Mandler, Michael (2012) “The fragility of information aggregation in large elections,” *Games and Economic Behavior*, 74, 257–268.

Martinelli, César (2006) “Would rational voters acquire costly information?,” *Journal of Economic Theory*, 129, 225–251.

McMurray, Joseph, C. (2013) “Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses,” *Review of Economic Studies,* 80, 277–312.

McLennan, Andrew (1998) “Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents,” *American Political Science Review*, 92, 413–418.

Selvin, Steven (1975) “A problem in probability (letter to the editor),” *American Statistician*, 29, 67.

Tajika, Tomoya (2018) “Polarization and inefficient information aggregation in electoral competition,” *mimeo*.

Tierney, John (1991) “Behind Monty Hall’s Doors: Puzzle, Debate and Answer?,” *The New York Times*.

vos Savant, Marilyn (1990) “Ask Marilyn,” *Parade Magazine*, 16.

Wit, Jörgen (1998) “Rational Choice and the Condorcet Jury Theorem,” *Games and Economic Behavior*, 22, 364–376.