Bonn Econon Discussion Papers

Discussion Paper 04/2015

On Two-Period Committee Voting: Why Straw Polls Should Have Consequences

By

Tim Julius Frommeyer

April 2015

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Kaiserstrasse 1
D-53113 Bonn

Financial support by the
Deutsche Forschungsgemeinschaft (DFG)
through the
Bonn Graduate School of Economics (BGSE)
is gratefully acknowledged.

Deutsche Post Stiftung is a sponsor of the BGSE.

On Two-Period Committee Voting: Why Straw Polls Should Have Consequences.

Tim J. Frommeyer*

April, 2015

Abstract

We consider a committee voting setup with two rounds of voting where committee members, who possess private information about the state of the world, have to make a binary decision. We investigate incentives for truthful revelation of their information in the first voting period. Coughlan (2000) shows that members reveal their information in a straw poll only if their preferences are in fact homogeneous. By taking costs of time into account, we demonstrate that committees have strictly higher incentives to reveal information if a decision can be made for high levels of consensus in the straw poll already. In such scenarios, members of all homogeneous and some heterogeneous juries are strictly better off when the requirement for early decisions is chosen carefully.

JEL classification: D72, D82, D83

Keywords: Communication, Committees, Voting

*Bonn Graduate School of Economics, Department of Economics, University of Bonn, Kaiserstraße 1, 53113 Bonn, Germany, timfrommeyer@uni-bonn.de. The author wihes to thank Mark Le Quement, Benjamin Schickner, Deszö Szalay and workshop participants In Bonn for helpful comments and suggestions.

1 Introduction

Information is particularly valuable when important and difficult decisions are pending. In many situations those decisions are not made by a single person but rather in groups, so called *committees*, in a voting procedure. University faculties typically delegate their decision on an applicant's employment to a committee and many companies, such as *Google*, proceed similarly[1]. In politics, party factions and parliaments form committees for investigations or in order to work out recommendations, and in jury trials a committee composed of representatives of the society has to decide whether a defendant is guilty or innocent of the accused crime. In any of these cases, committee members have a distinct evaluation of the subject for debate. Therefore, a committee has a richer pool of information at its disposal than an individual decision maker and could potentially make more accurate decisions. This information, however, is dispersed. A widely used approach for coordination are *straw polls* as non-binding communication devices before the decisive vote[2].

When the United Nations Security Council discussed the candidates for the position of Secretary-General of the United Nations in 2006, it conducted a series of straw polls to determine the members opinions on the candidates. Although none of these votes had decisional power, all candidates withdrew their candidacy afterwards except for Ban Ki-moon who received the most votes in each straw poll[3]. The famous movie *12 Angry Men* from 1957 covers the deliberation process of a jury. At the beginning of their deliberation, the jurors conduct a straw poll in order to collect their initial attitudes towards the underlying case and use the result as a starting point for the subsequent debate.

The motivation for conducting straw polls stems from the intuition that voters can harmlessly reveal their information, not being at risk of unintentionally causing a decision already. Additionally, from a series of experiments on decision making in groups, Goeree and Yariv (2011) conclude that members strongly appreciate information revealed by others. Although performed frequently and in many situations, the benefit of straw polls is disputed. According to Robert III. et al. (2000, p. 415)[4], conducting "an informal straw poll to 'test the water' (...) neither adopts nor rejects a measure and hence is meaningless and dilatory."

Condorcet (1785) was the first to formally argue in his *Jury Theorem* that voting decisions made in groups outperform those of individuals. In his setup, a jury votes once on a binary decision, each of them being objectively best in one of two possible states of the world, and the alternative with more votes is

[1] "An independent committee of Googlers review feedback from all of the interviewers. This committee is responsible for ensuring our hiring process is fair and that we're holding true to our 'good for Google' standards as we grow.", Google, How We Hire.

[2] Especially in large committees straw polls are a simple and swiftly conducted communication tool when verbal deliberation is tedious and confusing. Moreover, an anonymous poll can circumvent privacy issues and still provide some communication among committee members.

[3] More importantly, Ban was the only candidate who received votes from all permanent members of the Security Council which have veto power in the decisive election process. Ban then was elected by the general Assembly on October 13, 2006.

[4] *Robert's Rules of Order Newly Revised* is widely used as parliamentary authority, e.g., by the US Congress, and guide for meetings and assemblies.

implemented. Jury members condition their voting decision only on their private information about the state, which is correct with probability higher than 0.5. Moreover, the jury members' preferences are aligned, in the sense that they would undoubtedly prefer the same (objectively best) decision if the state of the world was commonly known, e.g. to convict a guilty defendant and acquit the innocent. As the vote aggregates the individuals' information, the implemented decision is more likely to be correct than the one from an imperfectly informed individual.[5]. Condorcet assumes that agents only consider their private information when voting in a group just as they would when deciding all alone. In strategic games, however, this assumption is not necessarily satisfied. Austen-Smith and Banks (1996) as well as Feddersen and Pesendorfer (1998) set up a standard model formally and consider a homogeneous committee with commonly known and aligned preferences. Preferences can be interpreted as levels of reasonable doubt which are identical for homogeneous committee members. In other words, agents would unanimously prefer the same decision not only if the state was known but also if they faced the same information. They demonstrate that the voting behavior assumed by Condorcet constitutes an equilibrium only if the voting rule is adjusted to the agents' preferences appropriately.
Coughlan (2000) considers an extension of the standard model with committees whose members are heterogeneous with respect to their levels of reasonable doubt. Preferences are aligned but agents assess the same information differently and might disagree on the preferred decision. He studies the role of a preliminary non-binding straw poll and demonstrates that information is aggregated in equilibrium only if the agents' preferences are in fact homogeneous. If an agent considers revealing his information in a straw poll, he conditions on the case where the information he discloses tips the balance in the subsequent decisive vote, that is when his vote is *pivotal*. If the committee is heterogeneous and other agents draw opposite conclusions from revealed information, disclosing information truthfully can lead to a unfavorable voting outcome for some agents. In this case, these agents can be better off with providing misinformation in the straw poll in order to manipulate the decisive vote in their favor.

In this paper, we study a modified version of Coughlan (2000)'s heterogeneous committee model in which agents additionally have preferences over the length of deliberation. More specifically, committee members incur costs of time from each round of voting and, hence, prefer to make decisions earlier. We hereby account for the deferring feature of straw polls pointed out by Robert III. et al. (2000). When committee members engage in a non-binding straw poll they always vote twice. If agents have the opportunity to circumvent a second poll whenever there is broad agreement in the first vote already, new strategic effects arise. Unlike in a non-binding straw poll, agents are then not only pivotal when their disclosed information tips the balance in the subsequent vote but also when there is a high level of consensus for one of the two alternatives within the other agents' information. In these cases an agent can maximize the probability to prevent a redundant second vote by revealing his information truthfully. As agents form beliefs about the probability of each pivotal case from their private information, consensus among

[5] In addition, by a law of large numbers the probability of a correct decision approaches 1 as the jury size grows large. See, for example, Piketty (1999) for an overview on the Jury Theorem. For a discussion and extensions see Ladha (1992), Miller (1986) and Young (1988).

the other agents for the agent's initially preferred decision is more likely than for the opposite. Hence, committee members have better incentives to reveal their information compared to a straw poll that has no consequences. In particular, even committees that are in fact heterogeneous are able to aggregate information perfectly if the straw poll is modified to implement a decision for high levels of consensus straightaway without a second vote.

In addition, we consider the committee's welfare under both a non-binding straw poll and a potentially consequential first vote. When we allow for a decision already in the first vote, we can always identify a requirement on the implementation of this vote that strictly improves the welfare of each juror of any homogeneous committee. If the requirement is strict enough, the jury will make the same decision in both setups but the ability to save costs causes a strict welfare improvement. Naturally, this positive welfare effect from saving costs is also present for heterogeneous juries. The made decisions, however, are always suboptimal for some members of heterogeneous committees. This provides a negative welfare effect for those agents. The overall welfare effect is positive for every agent as long as heterogeneity is manageable. hence, we can show that there is always a decision threshold for the first vote for which every member of committees with bounded heterogeneity is strictly better off than with a straw poll in the first period.

Finally, we are concerned with a designer who wants to optimally set a threshold for agreement in the first period. Hereby we distinguish the case where the designer observes the committee's preferences to that where she must commit to a threshold before getting to know the agents. Facilitating early decisions saves costs of time but potentially impacts the optimality of the jury decision. The designer must solve this trade-off with her choice of the decision rule. We provide conditions for her choice to be optimal from all candidate decision rules that pareto-dominate any equilibrium outcome of setups with straw polls.

This paper belongs of the literature on committee voting with commonly known preferences. As standard, we use the terminology of jury trials in the following. A committee is referred to as *jury* and a committee members is called a *juror*. The jury has to decide between *conviction* or *acquittal* of a defendant who is either *guilty* or *innocent*.

Deimen et al. (2013) show that Coughlan's impossibility result can be mitigated when the information structure is enhanced and allows jurors to examine consistency of information. Hummel (2012) as well as Thordal-Le Quement and Yokeeswaran (2013) demonstrate that information aggregation can be achieved if heterogeneous committees deliberate in homogeneous subgroups first.

For the case of privately known preferences, information is aggregated in heterogeneous juries under certain conditions on the voting rule (Austen-Smith and Feddersen, 2006), the jurors' preferences and the jury size (Meirowitz, 2007; Thordal-Le Quement, 2013).

Our work is also related to information aggregation in elections. Morgan and Stocken (2008) study the informative substance of straw polls that are held prior to elections. In their setup the result of the poll influences the subsequent policy choice which constituents take into consideration strategically. They find that information can be aggregated in small polls if the electorate is sufficiently homogeneous. Piketty (2000) investigates the effect on information aggregation if jurors use elections to communicate their

preference in order to influence future decisions.
Furthermore, our work is partly related to the literature on debates. Austen-Smith (1990) analyzes if debate can influence a later decision as well as its informational role. Bognar et al. (2013) study information aggregation over the course of repeated and costly negotiations whereas Damiano et al. (2010, 2012b) investigate the role of costly delay in repeated negotiations with asymmetric information. Damiano et al. (2012a) derive an optimal deadline on repeated negotiations.

The following chapter introduces the standard model of committee voting extended with costs of time from the voting process. In Chapter 3 we compare the results of Coughlan (2000) on non-binding straw polls followed by a decisive vote to a two-period setup with the opportunity for agreement in the first vote for broad agreement already. We show that the conditions on the juries heterogeneity for equilibria with information aggregation in the first period are weaker for the latter setup. In Chapter 4 we argue that heterogeneous juries can be strictly better off with the potential for early agreement compared to a pure straw poll and we provide conditions for a designer to set the decision rule for the first period optimally. Finally, Chapter 5 concludes. All proofs are relegated to the Appendix.

2 The Two-period Committee Voting Model

2.1 Agents/Jurors

N denotes the number and, with slight abuse of notation, the set of agents who make a binary decision. We will interpret the problem as the *conviction* (C) or *acquittal* (A) of a defendant in a jury trial. Therefore, we call an agent from now on a *juror* and the set N is a *jury* consisting of N jurors. Jurors maximize their expected utility and we assume common knowledge of rationality.

2.2 States and Preferences

There are two states of the world, $\omega \in \{G, I\}$. The defendant is either *guilty* ($\omega = G$) or *innocent* ($\omega = I$). For simplicity, we assume that both state are equally likely, that is $\Pr[\omega = G] = \Pr[\omega = I] = 1/2$. Juror j's preferences are state dependent and normalized to

$$\begin{aligned} U_j(C \mid \omega = G) &= U_j(A \mid \omega = I) = 0, \\ U_j(C \mid \omega = I) &= -q_j, \\ U_j(A \mid \omega = G) &= -(1 - q_j), \end{aligned}$$

where $q_j \in (0, 1)$. This normalization allows us to interpret q_j as a threshold probability of guilt of juror j, who prefers decision C over A if and only if his perceived probability of state G is larger or equal to q_j. Jurors with lower thresholds q_j are comparatively more biased towards C whereas jurors with higher thresholds q_j are more biased towards A. In the context of a jury trial one can also think of q_j as j's level of *reasonable doubt*. We assume that all preferences q_j are commonly known.
Without loss of generality, we sort jurors by their preferences $q_1 \leq \ldots \leq q_N$. Ex ante jurors are solely

distinguished by their preferences and we will say *juror j* for a juror with preferences $q_j \in [q_{j-1}, q_{j+1}]$. Additionally, we impose that jurors incur additive costs of $c \geq 0$ on their utility from each round of voting, which represent costs of time or opportunity costs. Alternatively, one could think of impatient jurors who prefer earlier decisions to later ones.

2.3 Information

Prior to the decision making process each juror $j \in N$ receives an informative signal $s_j \in \{i, g\}$ about the state of the world, where

$$\Pr[s_j = i|\omega = I] = \Pr[s_j = i|\omega = G] = p \in \left(\frac{1}{2}, 1\right)$$

We refer to p as the signal's *precision*. Signals are independently drawn, privately observed and not verifiable. The signal's precision is identical and known to every juror.

2.4 Voting

We consider a two-period voting game of the following form. In both periods jurors vote for either A or C. Denoting by x_t the number of C-votes in period t, the decision in period $t \in \{1, 2\}$, d_t, is determined by a decision rule represented by threshold k_t. In $t = 1$, the defendant can be convicted or acquitted only if more than or equal to $N - k_1$ jurors vote for the corresponding alternative[6]. If such a majority turns out, the game ends and the jury's respective decision is implemented. Otherwise, the decision is delayed ($d_1 = D$). The number of votes for both alternatives is revealed and the jury votes again in period 2. Formally the decision rule in period 1 is given by,

$$d_1 = \begin{cases} A & \text{if } x_1 < k_1 \\ C & \text{if } N - k_1 < x_1 \\ D & \text{else.} \end{cases} \tag{1}$$

See Figure 1 for a graphical representation with the number of C-votes in $t = 1$ on the axis. If the game continues in second period after $d_1 = D$, x_1 is disclosed the jury votes again. The defendant is convicted if at least k_2 jurors vote for C and acquitted otherwise. Formally the decision rule in $t = 2$ is given by,

$$d_2 = \begin{cases} A & \text{if } x_2 < k_2 \\ C & \text{if } k_2 \leq x_2. \end{cases} \tag{2}$$

Figure 2 shows a graphical representation. In the following, we refer to a decision rule in period t by the corresponding threshold k_t, which are exogenously given to the jurors.

If k_1 is small, a decision can only be made if there is broad agreement in the first period, whereas a decision can also be made by a smaller majority if k_1 is higher. The case of $k_1 = 0$ represents a non-binding *straw poll* in the first period which can be interpreted as a preliminary round of communication

[6] We implicitly assume symmetry in the voting rule for period 1. This simplifies the analysis but does not impact the qualitative results. For the impact which costs of time has on incentives for informative voting in the first period, it will only be necessary that a super majority is needed for conviction or acquittal in the first period, but it does not need to be the same.

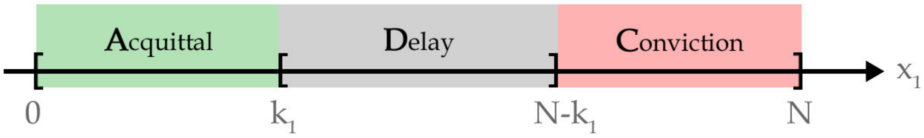

Figure 1: Decision rule in $t = 1$.

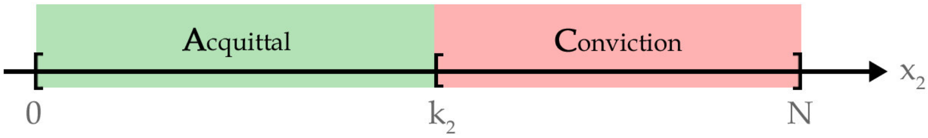

Figure 2: Decision rule in $t = 2$.

in which no decision can be made yet. In this communication, however, jurors cannot remember any information revealed but the number of jurors who prefer each alternative[7]. Straw polls are investigated in detail in Coughlan (2000). His setup is a special case of ours with $k_1 = 0$ and $c = 0$. We will use his findings, which we summarize in the following chapter, as a benchmark.

2.5 Notation

The posterior probability that the state is G if x of N signals indicate g is denoted by $\beta(x, N)$ and is computed according to Bayes' Rule. That is,

$$\beta(x, N) = \frac{(1-p)^{N-x}p^x}{(1-p)^{N-x}p^x + (1-p)^x p^{N-x}}.$$

When juror j enters period 1 or 2 his information is given by $\mathcal{I}_j^1 = (s_j)$, and $\mathcal{I}_j^2 = (s_j, x_1)$, respectively. We denote juror j's strategy by $\sigma_j = \left(\sigma_j^1, \sigma_j^2\right)$, where σ_j^t denotes the probability that j votes C in period t given his information $\mathcal{I}_j^t$. If j votes according to his signal, that is voting for conviction after receiving a g-signal and for acquittal otherwise, we say that j votes *informatively*. Formally, j's strategy prescribes informative voting in period t if

$$\sigma_j^t\left(\mathcal{I}_j^t\right) = \begin{cases} 1 & \text{if } s_j = g, \\ 0 & \text{if } s_j = i. \end{cases}$$

When juror j votes for C if and only if his perceived probability of guilt in period t, given his information, exceeds his level of reasonable doubt q_j, we say that j votes *sincerely* in period t. Formally, j's strategy prescribes sincere voting in period t if

$$\sigma_j^t\left(\mathcal{I}_j^t\right) = \begin{cases} 1 & \text{if } \Pr_j\left[G|\mathcal{I}_j^t\right] \geq q_j, \\ 0 & \text{if } \Pr_j\left[G|\mathcal{I}_j^t\right] < q_j. \end{cases}$$

Besides their own signals, jurors access information from inferring the other jurors' signals from their strategic behavior and observed aggregated voting outcomes. As all jurors receive signals of the same

[7] Even if there is no non-binding straw poll in the first stage but no decision was made, jurors can only observe how many jurors voted for each alternative but not each jurors vote individually. This assumption does not impact our results, because each jurors signal has the same precision. If this was not the case, jurors could learn about other signal's precision, such as in Bognar et al. (2013).

precision, inferring more g-signals leads to higher posterior probability of guilt. Analogously to every juror's probability threshold q_j, we can determine thresholds on the number of g-signals that a juror needs to observe in order to prefer conviction of the defendant. In this context, λ_j^N represents the *conviction threshold* of juror $j \in N$. He prefers C over A if and only if he observes more than or equal to λ_j^N of N signals indicating g.

Definition 1. Juror $j \in N$ has conviction threshold λ_j^N if

$$\beta(\lambda_j^N - 1, N) \leq q_j \leq \beta(\lambda_j^N, N).$$

Analogously to Austen-Smith and Feddersen (2006) and Thordal-Le Quement (2013), we define a jury's *minimal diversity* as follows:

Definition 2. Jury N has *minimal diversity* of m if

$$\max_{(i,j)\in N\times N} \lambda_i^N - \lambda_j^N = m.$$

A jury's minimal diversity measures its heterogeneity in terms of its jurors' conviction thresholds. If a jury N has minimal diversity of 0 its jurors have the same conviction threshold and would agree on a decision if all private information was disclosed. This is not the case for juries with minimal diversity larger than 0. In the following, we will call a jury *homogeneous* if $m = 0$, and *heterogeneous* otherwise.

2.6 Strategies and Beliefs

We are interested in conditions under which information can be aggregated and jurors in fact use revealed information in their voting strategy. Therefore, we restrict attention to the following profile of strategies. Jurors reveal their private signal by voting informatively in the first vote. If information is not congruent to make a decision already, jurors observe the outcome of the first vote x_1 and update their beliefs about the state of the world accordingly to $\beta(x_1, N)$. Then, every juror votes sincerely in the second period. More formally, we will derive conditions on the jurors' preferences for the following profile of strategies and beliefs to constitute a Perfect Bayesian Equilibrium. For all $j \in N$,

$$\sigma_j^1(s_j) = \begin{cases} 1 & \text{if } s_j = g, \\ 0 & \text{if } s_j = i, \end{cases}$$

and

$$\sigma_j^2(s_j) = \begin{cases} 1 & \text{if } \beta(x_1, N) \geq q_j, \\ 0 & \text{if } \beta(x_1, N) < q_j, \end{cases}$$

where jurors consistently (correctly) believe that the others vote informatively in the first stage and update their belief about the state of the world accordingly. We denote this profile of strategies and beliefs by $(\sigma, \mu) = (\sigma_j, \mu_j)_{j=1}^N$, for which we make the following observations.

- If juror j (hypothetically) observes x votes for conviction from the other jurors before he votes in the first stage, his belief about the defendant being guilty is $\beta(x, N)$ if $s_j = g$ and $\beta(x-1, N)$ otherwise.

- If no decision is made in the first stage, all jurors observe x_1 and update their belief about the defendant being guilty to $\beta(x_1, N)$ via Bayes' rule. Therefore, all jurors have the same posterior belief in the second stage.

Equilibria with these strategies and beliefs have the property that information is perfectly aggregated in the first stage. If the information is not congruent in the sense that informative voting does not lead to a decision in the first stage yet, the jurors make their decision in the second stage conditional on all available information.
This profile of strategies and beliefs features strategic behavior in the second period which is sensitive to available information. As we are interested in conditions for informative voting, strategies should take revealed information into account. Alternative strategies for the second period subgame would require jurors to vote predominately for one alternative independently of available information. Providing incentives for information revelation is both difficult and needless if information is not appreciated in the decision process. This is most dominant if the decision is always made in the second period, that is when the first period vote is a straw poll.

Before we proceed to the analysis, we make an assumption on jurors preferences.

Assumption 1. For any $j \in N$, $\beta(0, N) < q_j < \beta(N, N)$ or, equivalently, $\lambda_j^N \in \{1, \ldots, N\}$.

This assumption excludes jurors with extreme preferences who still prefer an alternative even if all signals were known and indicated the opposite state of the world. In other words, we exclude prejudiced jurors who vote for one alternative regardless of available information. In particular, this assumption ensures that the juror who is pivotal in the second vote, if all jurors vote sincerely, takes the available information into account.

3 Equilibrium Analysis

In this chapter, we analyze the two-period model in the following order. We start by discussing the impossibility result of Coughlan (2000) for a non-binding straw poll in the first period, that is for $k_1 = 0$. Then, we show how costs of time influence the jurors incentives to reveal information in their first period vote if $k_1 > 0$.

3.1 $k_1 = 0$: Non-binding First Period straw poll

Consider first a situation with a non-binding straw poll in the first period, i.e. $k_1 = 0$, as presented by Coughlan (2000). He shows that no jury votes informatively in the straw poll and sincerely in the decisive vote in equilibrium, unless all jurors have the same conviction thresholds. In other words, the profile of strategies and beliefs (σ, μ) can only be an equilibrium for juries with minimal diversity of 0[8].

[8] Cf. Proposition 5 in Coughlan (2000) where he differentiates three cases. Apart from the one mentioned, the other cases are ruled out by Assumption 1. Information does not influence the jury's final decision in those cases and information revelation is trivially an equilibrium behavior.

In the decisive voting period $t = 2$, jurors vote sincerely by taking revealed information into account. Each juror $j \in N$ conditions his vote on the situation in which he is pivotal, that is when his vote actually decides upon the defendant's acquittal or conviction. In any other case his vote does not affect the decision and, thus, his expected payoff from voting A or C is equal. As the jurors maximize their expect utility with their votes, they condition on the unique situation that affects their expected utility. Recall that jurors learn the outcome of the (informative) straw poll before casting their vote in $t = 2$. Fully informed, every $j \in N$ updates his posterior probability of guilt to $\beta(x_1, N)$ and votes for conviction if and only if

$$\beta(x_1, N) \geq q_j \quad \Leftrightarrow \quad x_1 \geq \lambda_j^N,$$

which coincides with sincere voting. If jurors vote informatively in the straw poll, sincere voting is straightforward part of any equilibrium strategy which is sensitive to revealed information.

Having established the equilibrium strategies in $t = 2$, we can now consider incentives for informative voting in the straw poll. The juror who is actually pivotal in the decisive vote is juror k_2. When k_2 votes for C sincerely, every $j < k_2$ will do so as well. When k_2 prefers A given the revealed information, every $j > k_2$ has the same preference. Therefore, whichever decision k_2 prefers in $t = 2$ will be implemented. Although no decision can be made in the straw poll, the vote has an impact on the information influencing the decisive vote later. Given the strategy profile, jurors anticipate that k_2 with conviction threshold $\lambda_{k_2}^N$ is pivotal in the second period. As a consequence, juror j is pivotal in the straw poll, if his vote in $t = 1$ influences k_2'th information to swing the pivotal vote in $t = 2$ to either C or A. This is the case if $\lambda_{k_2}^N - 1$ of the other $N - 1$ jurors informatively vote C in the straw poll. For j to vote informatively in $t = 1$ as well, he has to prefer C after he receives a g-signal and A after an i-signal. In the event that j is actually pivotal in the first period in the above sense, he faces $\lambda_{k_2}^N - 1$ g-signals from the other jurors, resulting in $\lambda_{k_2}^N$ g-signals in total if $s_j = g$, and $\lambda_{k_2}^N - 1$ g-signals in total if $s_j = i$. Therefore, j votes informatively if and only if

$$\beta\left(\lambda_{k_2}^N - 1, N\right) \leq q_j \leq \beta\left(\lambda_{k_2}^N, N\right) \quad \Leftrightarrow \quad \lambda_j^N = \lambda_{k_2}^N, \tag{3}$$

that is if he has the same conviction threshold as juror k_2.

Since this argument is the same for any juror of the jury, all jurors must have the same conviction threshold as k_2 to sustain an equilibrium with informative voting in the straw poll. In other words, informative voting in a straw poll can not be part of an equilibrium strategy for any heterogeneous jury. Suppose a juror j's conviction thresholds conflicts with the one of k_2, i.e., $\lambda_j^N \neq \lambda_{k_2}^N$. Given that all other jurors vote informatively, providing information truthfully in the straw poll implements an undesirable decision for j. By misinforming, however, he could improve the jury's decision from his point of view which represents a profitable deviation.

Note that costs from voting play no role for the jurors' incentives to vote informatively in a straw poll. They always vote twice and cannot avoid costs with their behavior in the first period. Therefore, this impossibility result for non-binding first period straw polls is independent of assuming costs of time.

3.2 $k_1 > 0$: Allowing for early agreements

Having seen the difficulties to incentivize information aggregation in a straw poll that is followed by a decisive vote, we turn towards two-period voting setups that allow for agreement in the first stage already. Formally, we consider $k_1 > 0$, so that the defendant can be convicted or acquitted in $t = 1$ according to decision rule (1), that is, if $x_1 > N - k_1$, or $x_1 < k_1$ respectively. For example with $k_1 = 1$, a decision in the first vote can be made unanimously and the final vote follows only if there is no unanimous agreement in the first period. Thereby jurors can avoid entering the second period and save costs of time. As a result, incentives for informative voting are influenced. Jurors trade off the influence of their vote on the information of the pivotal juror in period 2 against the opportunity that their vote causes an earlier decision and saves costs of time. This trade-off is solved in favor of informative voting in the first period. While the effect in the case of being pivotal as in a straw poll remains, voting informatively in the first period increases the probability that if an earlier decision is made it is the juror's preferred one. As a result, juries vote informatively even if their minimal diversity is larger than 0.

Proposition 3.1. *Suppose each juror j has voting costs of $c \geq 0$ and $k_1 > 0$. The profile of strategies and beliefs (σ, μ) constitutes a Perfect Bayesian Equilibrium if and only if preferences satisfy one of the following conditions:*

1. *For $\lambda_{k_2}^N \in \{1, \ldots, k_1\}$,*

$$\beta(k_1 - 1, N) - c \cdot \alpha(k_1 - 1) \leq q_j \leq \beta(k_1, N) + c \cdot \gamma(k_1) \quad \forall j \in N, \tag{4}$$

2. *for $\lambda_{k_2}^N \in \{k_1 + 1, \ldots, N - k_1\}$,*

$$\beta(\lambda_{k_2}^N - 1, N) - c \cdot \alpha(\lambda_{k_2}^N - 1) \leq q_j \leq \beta(\lambda_{k_2}^N, N) + c \cdot \gamma(\lambda_{k_2}^N) \quad \forall j \in N, \tag{5}$$

3. *for $\lambda_{k_2}^N \in \{N - k_1 + 1, \ldots, N\}$,*

$$\beta(N - k_1, N) - c \cdot \gamma(k_1) \leq q_j \leq \beta(N - k_1 + 1, N) + c \cdot \alpha(k_1 - 1) \quad \forall j \in N, \tag{6}$$

where

$$\alpha(x) = \binom{N-1}{x}^{-1} \binom{N-1}{k_1 - 1} \frac{(2p-1)\left[(1-p)^{k_1-1}p^{N-k_1} - (1-p)^{N-k_1}p^{k_1-1}\right]}{(1-p)^x p^{N-x} + (1-p)^{N-x}p^x} > 0,$$
$$\gamma(x) = \binom{N-1}{x-1}^{-1} \binom{N-1}{k_1 - 1} \frac{(2p-1)\left[(1-p)^{k_1-1}p^{N-k_1} - (1-p)^{N-k_1}p^{k_1-1}\right]}{(1-p)^x p^{N-x} + (1-p)^{N-x}p^x} > 0.$$

Proof. See Appendix. □

The terms $\alpha(x)$ and $\gamma(x)$ represent weighted conditional probabilities for those pivotal scenarios in which voting informatively saves costs without changing the jury's final decision for decision rule k_1. Note, that the impossibility result of Coughlan (2000) follows immediately for $k_1 = 0$ or $c = 0$[9].

[9] The original result Coughlan (2000) also covers the case where preferences are such that the pivotal juror in $t = 2$ always prefers C or A independently of revealed information. As discussed, we neglect these cases by Assumption 1, as informative voting is trivially equilibrium behavior for those. We focus instead on juries for which a straw poll does not always aggregate information.

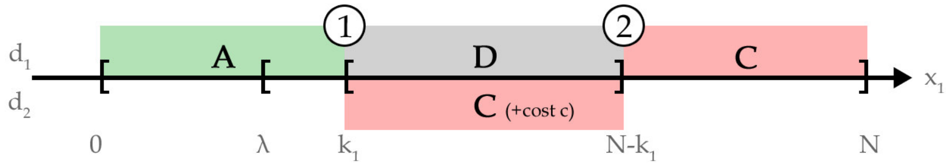

Figure 3: Pivotal scenarios in $t = 1$ for condition (4).

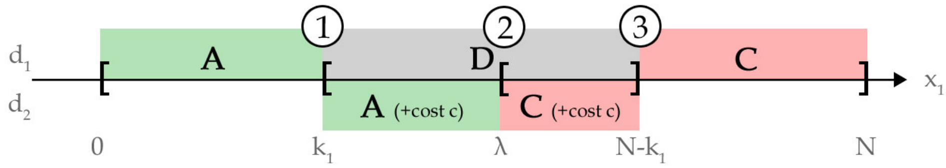

Figure 4: Pivotal scenarios in $t = 1$ for condition (5).

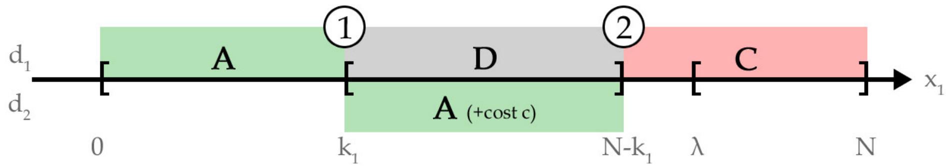

Figure 5: Pivotal scenarios in $t = 1$ for condition (6).

In order to provide intuition why heterogeneous juries aggregate information once the first vote can have consequences also, we highlight the changes on incentives compared to the previously discussed case of a straw poll in which each juror faces the same trade-off in the unique case of being pivotal. Jurors with different preferences solve this trade-off differently and some prefer to misinform the others in a straw poll in order to manipulate k_2's belief and make him implement a superior decision.

Let us treat the case where $\lambda_{k_2}^N \in \{N - k_1 + 1, \dots, N\}$ first, which corresponds to condition (5). In this case, the decision of juror k_2 in period 2 is not predetermined if the vote takes place. That is, the k_2 could decide for either A or C in $t = 2$ depending on $x_1 \in \{k_1, \dots, N - k_1\}$ that leads to a second vote. Consider some juror $j \in N$ and suppose that the other jurors vote according to (σ, μ) , i.e., informatively in the first and sincerely in the second period. When the first vote can already have consequences, 2 more scenarios arise in which jurors are pivotal compared to a straw poll. Denote by x_{-j} the number of (informative) C votes of jury N except juror j. There are now three pivotal scenarios for j who faces the following trade-offs.

1. $x_{-j} = k_1 - 1$: If j votes C, the decision is delayed and $d_2 = A$ at additional costs c. If j votes A the defendant is acquitted immediately.

2. $x_{-j} = \lambda_{k_2}^N - 1$: If j votes C, the decision is delayed and $d_2 = C$ at additional costs c. If j votes A, the decision is delayed and $d_2 = A$ at additional costs c.

3. $x_{-j} = N - k_1$: If j votes C, the defendant is convicted immediately. If j votes A the decision is delayed and $d_2 = C$ at additional costs c.

See Figure 4 for a graphical representation. Scenario 2 corresponds to the unique pivotal scenario each juror faces in a straw poll and this trade-off is solved in favor of voting C if and only if condition (3) is satisfied, i.e., j has the same conviction threshold as k_2. In Scenario 1, regardless of j's vote, the jury will decide in favor of A in the second period. By voting A in the first period, however, j can cause this decision earlier and thus save costs of time. Voting A is strictly better in Scenario 1. An analogous argument holds in Scenario 3. Here, the jury's decision will be C regardless of j's vote in the first period, which he can implement earlier by voting C. Therefore, voting C is strictly better in Scenario 3.

Now, whereas the likelihood of Scenario 2 is untouched, it is more likely for j that Scenario 1 occurs than Scenario 3 if he receives an i-signal, and Scenario 3 appears more likely if he receives a g-signal[10]. In expectation, voting A becomes more attractive after j observes an i-signal and voting C gains attraction otherwise. As these effects add up, the incentives for informative voting improve. Moreover, the intervals in which preferences q_j of all $j \in N$ have to be located to vote informatively expand linearly in c with factors $\alpha(\cdot)$ and $\gamma(\cdot)$ representing the weighted cost saving effect due to Scenario 1 and 3. As a result incentives are provided, for heterogeneous juries even, to aggregate information truthfully in the first vote for any $c > 0$.

Complementing, condition (4) and (6) cover the cases in which the pivotal juror's decision in the second period is always C, or A respectively, if reached. In these cases there are only two pivotal scenarios. If the second vote results always in C, then both decisions about an earlier agreement and about the jury's final decision coincide so that both Scenario 1 and 2 occur together if $k_1 - 1$ of the remaining jury votes C in the first period. Conversely, if the pivotal jurors always votes A in the second period, Scenarios 2 and 3 occur together by an analogous argument. See Figures 3 and 5 for a graphical representations. Besides that, the argument is similar as for the first case; if the first vote is potentially consequential, the opportunity to save costs of time improve any juries incentives to vote informatively in the first period.

The improvement of juror's incentives to vote informatively depends on the ability to avoid costs of time without changing the jury's final verdict. Making the first period vote potentially consequential as well as the presence of costs of time to jurors are inevitable for this effect. Its strength mirrors in the three conditions of Proposition 3.1 and depends on the positive factors $\alpha(\cdot)$ and $\gamma(\cdot)$ as well as on the costs parameter c. As an immediate consequence from Proposition 3.1 we can observe that juries of arbitrary heterogeneity vote informatively in equilibrium if costs of time are high enough.[11].

Before we discuss resulting welfare effects, we provide a short numerical example to stress the extend by which requirements on the juries' heterogeneity are relaxed due to potential early agreement.

4 Welfare

In the previous chapter we showed that making the first vote consequential provides incentives for heterogeneous juries to reveal their information. We are now concerned with the effect of this change on the jury's welfare.

[10] For this argument we only need that the number of votes of the other jurors to trigger Scenario 1 is smaller than $N/2$, and for Scenario 3 larger than $N/2$ respectively. This is assured by assuming symmetry in $k_t < N/2$. Symmetry, however, is not necessary but simplifies the analysis dramatically.

[11] Note as well that informative voting is trivially equilibrium behavior in any circumstances if p approaches 1. As signals are (almost) perfect, information disparity vanishes and jurors vote according to their (almost) perfect signal.

4.1 Inferiority of Straw Polls

The potential for early agreement allows jurors to avoid the costly second vote whenever agreement among them is broad enough. Albeit this new possibility drives the improvement of incentives for informative voting it is not immediately clear that jurors are better-off compared to a non-binding straw poll. In the latter setup the conviction threshold of the pivotal juror in the second vote pins down the unique pivotal scenario for all jurors. This is not the case anymore necessarily if $k_1 > 0$ is chosen relatively large. Either (4) or (6) prescribe the conditions for an informative equilibrium in this case. Both have the common feature that the minimal number of C votes required in the first period to finally convict the defendant is pinned down by the threshold for early agreement, k_1, and not by the jurors conviction thresholds. If both are too far apart, some relatively extreme homogeneous juries do not vote informatively in the first period any longer. In a straw poll, however, information is fully aggregated and the optimal decision from each jurors point of view is implemented. If saved costs are not enough to make up for this loss these juries are in fact worse off with $k_1 > 0$.

Example 4.1. As an example for such a case consider a homogeneous jury of $N = 20$ jurors with a common conviction threshold of $\lambda = 1$. This jury vote informatively in a non-binding straw poll but might not if early agreement is possible with $k_1 = 5$. As $\lambda < k_1$, the jury will convict the defendant whenever the second vote happens. A juror who is pivotal in the first period when 4 out of the other 19 jurors vote C informatively prefers to vote C instead of A, even if he receives an i-signal. Voting A in this situation saves costs but leads to an undesirable acquittal because $\lambda < 4 = k_1 - 1$. The incentives to save costs had to be unreasonably large in order to justify voting A after receiving an i-signal.

There is, however, always a voting rule for the first period that overcomes this issue such that a consequential first period vote does not alter a homogeneous jury's behavior in equilibrium. As a result, each such jury is strictly better off when it can avoid costs of time. Moreover, we can extend this insight to some heterogeneous juries. We show that the outcome of an informative equilibrium in setups with a consequential first period vote outperforms the upper bound of any equilibrium in straw poll setups for those juries.

Proposition 4.1. *Fix some k_2 and consider a jury N with bounded heterogeneity. For some $k_1 > 0$, every juror $j \in N$ is strictly better off in the equilibrium constituted by (σ, μ) than in any equilibrium of setups with a first period straw poll.*

Proof. See Appendix. □

We proceed in four steps. First, we show that any jury that votes informatively in a straw poll in equilibrium does the same for some decision rule $k_1 > 0$ in the first period. This is ensured if k_1 is chosen small enough. Second, we argue that these homogeneous juries are strictly better-off under $k_1 > 0$. For small values of k_1, the jury's decision does not change compared to a straw poll but for some realizations of signals, it is made earlier. As this saves costs, jurors are better off in expectation. In the third step, we establish an upper bound on the jurors' utilities from a setup with a straw poll in the first period. In fact, the best jurors can achieve from a straw poll setup is their preferred decision in the second vote under full information. This, however, is the outcome of (σ, μ) for homogeneous juries. Finally, we demonstrate that there are indeed some heterogeneous juries, that can improve upon this benchmark for some $k_1 > 0$. Unlike in the case of homogeneous juries, the final decision is not optimal for all jurors in this case. Nevertheless, these inefficiencies are compensated by saved costs in expectation, if their level of heterogeneity is not too high.

Note that the extent of heterogeneity that is compatible with this pareto improvement depends on the cost saving effect that is identical to each juror. A higher value of c strictly improves welfare for any $k_1 > 0$. Hence, a stronger costs saving effect can compensate higher inefficiencies from imperfect decisions caused by higher levels of heterogeneity. Additionally, note that the welfare criterion used is very strict. We require juries to improve upon the upper bound of each juror from a straw poll. In considering weaker welfare measures more heterogeneous juries could improve upon straw poll setups as well.

A non-binding straw poll in the first vote fails to provide the right incentives for heterogeneous juries to reveal their information truthfully, as Coughlan (2000) pointed out. We showed in the previous chapter that this can be accomplished if the first period vote can have consequences. Now, we establish that straw polls not only hurt incentives for information aggregation in a two-period voting setup but also lower the jurors expected utility compared to setups in which juries can make a decision in the first period already. In fact, homogeneous juries that reveal their information truthfully in a straw poll in equilibrium are strictly better off if the first vote allows for an earlier decision. One can conclude that straw polls rarely achieve information aggregation and if they do, a setup with opportunities for earlier agreement performs strictly better.

4.2 Designer's Choice of $k_1 > 0$

As a final step we consider a designer who can change the voting rule from a straw poll in the first vote to $k_1 > 0$. In many situations there are legal requirements or regulations on the majority that is needed for a decision which a designer cannot influence. We account for that by treating k_2, the majority rule of the final vote, as given and consider $k_1 < k_2$. If a decision is made early, the majority in favor is at least as large as legally required. In the following, we discuss which threshold for agreement the designer should set when the jury's preferences are known or unknown to her.

Let us consider the latter case first, where either the designer does not know the jurors' preferences or she has to commit to a decision rule for the first vote before the jury is announced. In this case, the designer cannot observe the exact value of c either, which is part of the jurors preferences. However, it is known that some costs $c > 0$ are present. For the reasons discussed previously, allowing for a decision in the first period increases the jurors welfare but if k_1 is chosen too high some relative extreme homogeneous juries are worse off. The following result follows immediately from Proposition 4.1.

Corollary 4.1. *Suppose the designer cannot observe the jurors preferences. She can increase the welfare of any jury that votes informatively in a straw poll by setting* $k_1 = 1$.

Proof. See Appendix. □

As discussed, a straw poll accomplishes information aggregation only if the jury is homogeneous. For this case, however,

Now suppose that the designer knows the jurors' preferences and can adjust the voting rule in the first period to increase their welfare accordingly. In order to do so she considers all $k_1 > 0$ that induce informative voting in the first period for this jury and we call the set of all such values $\mathcal{K}_1^N$. In the next step we isolate a subset of $\mathcal{K}_1^N$, which we call $\widetilde{\mathcal{K}}_1^N$. For any $k_1 \in \widetilde{\mathcal{K}}_1^N$, the jury votes informatively in $t = 1$ and it's welfare is strictly higher compared to the upper bound of setups with a straw poll in the first round. Let us assume that the jury's heterogeneity is bounded from above such that $\widetilde{\mathcal{K}}_1^N$ is non-empty. From Proposition 4.1, we know that this in indeed the case for small degree of heterogeneity. Finally, we provide conditions to identify a $k_1 \in \widetilde{\mathcal{K}}_1^N$ which maximizes the jury's welfare.

Proposition 4.2. *Suppose the designer can observe the jurors preferences whose degree of heterogeneity is bounded from above. Then, the optimal k_1 for which (σ, μ) is an equilibrium is given as follows:*

1. *If $\lambda_{k_2}^N \in \{k_1, \ldots, N - k_1 + 1\}$, set*

$$k_1 = \max \widetilde{\mathcal{K}}_1^N.$$

2. *If $\lambda_{k_2}^N < \frac{N}{2}$ and $\left\{k_1 \mid \lambda_{k_2}^N < k_1\right\} \subseteq \widetilde{\mathcal{K}}_1^N$, set*

$$k_1 = \max \left\{k_1 \in \widetilde{\mathcal{K}}_1^N \mid q_1 \geq \beta(k_1 - 1, N) - 2c \vee \lambda_{k_2}^N \in \{k_1, \ldots, N - k_1 + 1\}\right\}$$

3. *If $\lambda_{k_2}^N > \frac{N}{2}$ and $\left\{k_1 \mid \lambda_{k_2}^N > N - k_1 + 1\right\} \subseteq \widetilde{\mathcal{K}}_1^N$, set*

$$k_1 = \max \left\{k_1 \in \widetilde{\mathcal{K}}_1^N \mid q_N \leq \beta\left(N - (k_1 - 1), N\right) + 2c \vee \lambda_{k_2}^N \in \{k_1, \ldots, N - k_1 + 1\}\right\}$$

Proof. See Appendix. □

Intuitively, whenever the choice from all candidates $k_1 \in \widetilde{\mathcal{K}}_1^N$ does not influence the final decision, then increasing k_1 as high as possible saves the most costs and, hence, is optimal to choose. If an additional increase in k_1 impacts the jury's decision, the designer must trade-off the saved costs against the repercussions from interfering with the decision. The conditions above reflect this trade-off: The designer wants to choose a high $k_1 \in \widetilde{\mathcal{K}}_1^N$ in order to decrease expected costs for the jurors but not impact the jury's decision to an extend that outweighs the saved costs.

5 Concluding Remarks

This paper contributes to the literature on two period committee voting by additionally considering costs of time. We show that it is beneficial to grant an option for early agreement to the jurors as opposed to letting them engage in a straw poll first. Unlike in the straw poll setup of Coughlan (2000), even heterogeneous juries aggregate their information perfectly in the first vote once if it can already have consequences. Moreover, not only every homogeneous jury but also some heterogeneous juries are strictly better off in terms of expected utility when the jurors can make the decision earlier. We demonstrated how the ability to save costs of time positively influences the jurors incentives to aggregate information. Finally, we showed how a designer can profit from our insights when she can set the voting rule for the first period based on her information on the jurors' preferences.

The results are derived in a simple framework. There are numerous possible ways of identifying robustness of the effects on incentives for information aggregation from making the first vote consequential. Naturally, introducing a more sophisticated information structure or uncertainty about other players' types immediately come to mind. Moreover, the discussion on the designer's optimal choice gives rise to the consideration of endogenous costs. Higher costs increase incentives to vote informatively in setups which allow for agreement early. However, they negatively impact the jurors welfare. Even now the simple framework from this paper allows us to identify interesting effects that arise from introducing consequences in the first vote already.

This considerations can lead to interesting policy implications. We show that straw polls are not only dilatory, the reason why Robert III. et al. (2000) judged them "*meaningless*", but also can be easily outperformed in the ability to aggregate information if they can have consequences.

A Appendix

Proof of Proposition 3.1

Proof. We proceed by backward induction.

$t = 2$: Suppose jurors vote informatively in $t = 1$ and x_1 is the revealed number of C-votes from that period. Consistent with informative voting, any $j \in N$ believes with probability 1 that x_1 represents the amount of g-signals among the jurors. Accordingly, all $j \in N$ update their beliefs with Bayes' rule about the state of the world being G consistently to $\beta(x_1, N)$ for all $j \in N$, and for any signal s_j. Given the other jurors vote sincerely, j conditions his vote on being pivotal and votes C if and only if

$$\begin{aligned} -\left(1-\beta(x_1,N)\right)q_j &\geq -\beta(x_1,N)\left(1-q_j\right) \\ \Leftrightarrow q_j &\leq \beta(x_1,N). \end{aligned}$$

That is, j votes sincerely in $t = 2$.

$t = 1$: The received signal determines the jurors' prior beliefs. Any $j \in N$ attaches probability p to the state being G whenever $s_j = g$, and $1 - p$ otherwise. The jurors anticipate the outcome of the vote in $t = 2$, if it is reached, for any realization of x_1. There are three cases to distinguish.

1. Suppose $\lambda_{k_2}^N \in \{1, \ldots, k_1\}$. In this case $d_2 = C$ because $\lambda_{k_2}^N \leq k_1$. Denote by $\widetilde{\sigma} = \left(\sigma_{-j}, \left(\widetilde{\sigma}_j^1, \sigma_j^2\right)\right)$ the jury's strategy profile where all jurors vote sincerely in $t = 2$, and in $t = 1$ all but j vote informatively and j deviates to vote contrarily to his received signal, i.e., A, if $s_j = g$ and C, if $s_j = i$. We compute j's expected utilities in $t = 1$ from voting informatively, that is sticking to σ, as well as from deviating to $\widetilde{\sigma}$, that is providing misinformation.

$$\begin{aligned} EU_j\left[\sigma \mid \mu, s_j = g\right] = &\sum_{x=0}^{k_1-2}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}\left(-(1-q_j)\right) + \sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}(-c) \\ &+ \sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-c-q_j) + \sum_{x=N-k_1}^{N-1}\binom{N-1}{x}(1-p)^{x-1}p^{N-x+1}(-q_j). \\ EU_j\left[\widetilde{\sigma} \mid \mu, s_j = g\right] = &\sum_{x=0}^{k_1-1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}\left(-(1-q_j)\right) + \sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}(-c) \\ &+ \sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}\left(-c-q_j\right) + \sum_{x=N-k_1+1}^{N-1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-q_j). \end{aligned}$$

Juror j prefers to vote informatively after receiving a g-signal if and only if

$$EU_j\left[\sigma \mid \mu, s_j = g\right] - EU_j\left[\widetilde{\sigma} \mid \mu, s_j = g\right] \geq 0.$$

This condition is equivalent to

$$\begin{aligned} &(1-p)^{N-k_1}p^{k_1}(-c) + (1-p)^{k_1}p^{N-k_1}(-c-q_j) + (1-p)^{N-k_1+1}p^{k_1-1}(-q_j) \\ \geq &(1-p)^{N-k_1}p^{k_1}\left(-(1-q_j)\right) + (1-p)^{k_1-1}p^{N-k_1+1}(-c) + (1-p)^{N-k_1+1}p^{k_1-1}(-c-q_j) \qquad (7) \\ \Leftrightarrow &(1-p)^{k_1}p^{N-k_1}q_j - (1-p)^{N-k_1}p^{k_1}(1-q_j) \leq c(2p-1)\left[(1-p)^{k_1-1}p^{N-k_1} - (1-p)^{N-k_1}p^{k_1-1}\right] \qquad (8) \end{aligned}$$

However, (8) is an equivalent reformulation of $q_j \leq \beta(k_1, N) + c \cdot \gamma(k_1 - 1)$. As a result, all jurors vote informatively after receiving a g-signal if and only if $q_j \leq \beta(k_1, N) + c \cdot \gamma(k_1 - 1)$ for any $j \in N$. Analogously, given the other jurors vote informatively, a juror who receives an i-signal has the following expected utilities:

$$\begin{aligned}
EU_j\left[\sigma \mid \mu, s_j = i\right] = &\sum_{x=0}^{k_1-2} \binom{N-1}{x} (1-p)^{N-x} p^x \left(-(1-q_j)\right) + \sum_{x=k_1-1}^{N-k_1-1} \binom{N-1}{x} (1-p)^{N-x} p^x (-c) \\
&+ \sum_{x=k_1-1}^{N-k_1-1} \binom{N-1}{x} (1-p)^x p^{N-x} (-c-q_j) + \sum_{x=N-k_1}^{N-1} \binom{N-1}{x} (1-p)^x p^{N-x} (-q_j) \\
EU_j\left[\widetilde{\sigma} \mid \mu, s_j = i\right] = &\sum_{x=0}^{k_1-1} \binom{N-1}{x} (1-p)^{N-x} p^x \left(-(1-q_j)\right) + \sum_{x=k_1}^{N-k_1} \binom{N-1}{x} (1-p)^{N-x} p^x (-c) \\
&+ \sum_{x=k_1}^{N-k_1} \binom{N-1}{x} (1-p)^x p^{N-x} (-c-q_j) + \sum_{x=N-k_1+1}^{N-1} \binom{N-1}{x} (1-p)^x p^{N-x} (-q_j)
\end{aligned}$$

Analogously to the above case, $q_j \leq \beta(k_1, N) + c \cdot \gamma(k_1 - 1)$ is equivalent to

$$EU_j\left[\sigma \mid \mu, s_j = i\right] - EU_j\left[\widetilde{\sigma} \mid \mu, s_j = i\right] \geq 0$$

for all $j \in N$. Thus, jurors vote informatively after receiving an i-signal if and only if $q_j \leq \beta(k_1, N) + c \cdot \gamma(k_1 - 1)$ for all $j \in N$.

2. Suppose $\lambda_{k_2}^N \in \{k_1 + 1, \ldots, N - k_1\}$. This implies $d_2 = C$ if $x_1 \geq \lambda_{k_2}^N$, and $d_2 = A$ otherwise. Juror j's expected utilities are computed as follows in this case:

$$\begin{aligned}
EU_j\left[\sigma \mid \mu, s_j = g\right] = &\sum_{x=0}^{k_1-2} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} \left(-(1-q_j)\right) + \sum_{x=k_1-1}^{\lambda_{k_2}^N-2} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} \left(-(1-q_j) - c\right) \\
&+ \sum_{x=\lambda_{k_2}^N-1}^{N-k_1-1} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} (-c) + \sum_{x=k_1-1}^{\lambda_{k_2}^N-2} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-c) \\
&+ \sum_{x=\lambda_{k_2}^N-1}^{N-k_1-1} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-q_j - c) + \sum_{x=N-k_1}^{N-1} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-q_j) \\
EU_j\left[\widetilde{\sigma} \mid \mu, s_j = g\right] = &\sum_{x=0}^{k_1-1} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} \left(-(1-q_j)\right) + \sum_{x=k_1}^{\lambda_{k_2}^N-1} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} \left(-(1-q_j) - c\right) \\
&+ \sum_{x=\lambda_{k_2}^N}^{N-k_1} \binom{N-1}{x} (1-p)^{N-x-1} p^{x+1} (-c) + \sum_{x=k_1}^{\lambda_{k_2}^N-1} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-c) \\
&+ \sum_{x=\lambda_{k_2}^N}^{N-k_1} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-q_j - c) + \sum_{x=N-k_1+1}^{N-1} \binom{N-1}{x} (1-p)^{x+1} p^{N-x-1} (-q_j)
\end{aligned}$$

As before, we reformulate $q_j \leq \beta(\lambda_{k_2}^N, N) + c \cdot \gamma(\lambda_{k_2}^N)$ to the equivalent expression

$$\begin{aligned}
&\binom{N-1}{k_1-1} (1-p)^{N-k_1} p^{k_1} (-c) + \binom{N-1}{k_1-1} (1-p)^{k_1} p^{N-k_1} (-c) + \binom{N-1}{\lambda_{k_2}^N - 1} (1-p)^{\lambda_{k_2}^N} p^{N-\lambda_{k_2}^N} (-q_j) \\
\geq &\binom{N-1}{\lambda_{k_2}^N - 1} (1-p)^{N-\lambda_{k_2}^N} p^{\lambda_{k_2}^N} \left(-(1-q_j)\right) + \binom{N-1}{N-k_1} (1-p)^{k_1-1} p^{N-k_1+1} (-c) \\
&+ \binom{N-1}{N-k_1} (1-p)^{N-k_1+1} p^{k_1-1} (-c), \qquad (9)
\end{aligned}$$

which in turn is equivalent to

$$EU_j\left[\sigma \mid \mu, s_j = g\right] - EU_j\left[\widetilde{\sigma} \mid \mu, s_j = g\right] \geq 0.$$

Therefore, no juror has a profitable deviation from strategy profile σ after receiving a g-signal if and only if $q_j \leq \beta(\lambda^N_{k_2}, N) + c \cdot \gamma(\lambda^N_{k_2})$ for all $j \in N$.

Analogously, a juror who receives an i-signal has the following expected utilities:

$$\begin{aligned}
EU_j\left[\sigma \mid \mu, s_j = i\right] = &\sum_{x=0}^{k_1-2}\binom{N-1}{x}(1-p)^{N-x}p^x\left(-(1-q_j)\right) + \sum_{x=k_1-1}^{\lambda^N_{k_2}-2}\binom{N-1}{x}(1-p)^{N-x}p^x\left(-(1-q_j)-c\right)\\
&+\sum_{x=\lambda^N_{k_2}-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{N-x}p^x(-c) + \sum_{x=k_1-1}^{\lambda^N_{k_2}-2}\binom{N-1}{x}(1-p)^x p^{N-x}(-c)\\
&+\sum_{x=\lambda^N_{k_2}-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^x p^{N-x}(-q_j-c) + \sum_{x=N-k_1}^{N-1}\binom{N-1}{x}(1-p)^x p^{N-x}(-q_j)
\end{aligned}$$

$$\begin{aligned}
EU_j\left[\widetilde{\sigma} \mid \mu, s_j = i\right] = &\sum_{x=0}^{k_1-1}\binom{N-1}{x}(1-p)^{N-x}p^x\left(-(1-q_j)\right) + \sum_{x=k_1}^{\lambda^N_{k_2}-1}\binom{N-1}{x}(1-p)^{N-x}p^x\left(-(1-q_j)-c\right)\\
&+\sum_{x=\lambda^N_{k_2}}^{N-k_1}\binom{N-1}{x}(1-p)^{N-x}p^x(-c) + \sum_{x=k_1}^{\lambda^N_{k_2}-1}\binom{N-1}{x}(1-p)^x p^{N-x}(-c)\\
&+\sum_{x=\lambda^N_{k_2}}^{N-k_1}\binom{N-1}{x}(1-p)^x p^{N-x}(-q_j-c) + \sum_{x=N-k_1+1}^{N-1}\binom{N-1}{x}(1-p)^x p^{N-x}(-q_j)
\end{aligned}$$

We reformulate $\beta(\lambda^N_{k_2} - 1, N) - c \cdot \alpha(\lambda^N_{k_2} - 1) \leq q_j$ equivalently to

$$\begin{aligned}
&\binom{N-1}{k_1-1}(1-p)^{N-k_1+1}p^{k_1-1}(-c) + \binom{N-1}{k_1-1}(1-p)^{k_1-1}p^{N-k_1+1}(-c)\\
+&\binom{N-1}{\lambda^N_{k_2}-1}(1-p)^{\lambda^N_{k_2}-1}p^{N-\lambda^N_{k_2}+1}(-q_j)\\
\leq&\binom{N-1}{\lambda^N_{k_2}-1}(1-p)^{N-\lambda^N_{k_2}+1}p^{\lambda^N_{k_2}-1}\left(-(1-q_j)\right) + \binom{N-1}{N-k_1}(1-p)^{k_1}p^{N-k_1}(-c)\\
+&\binom{N-1}{N-k_1}(1-p)^{N-k_1}p^{k_1}(-c)
\end{aligned} \tag{10}$$

which is in turn equivalent to

$$EU_j\left[\sigma \mid \mu, s_j = i\right] - EU_j\left[\widetilde{\sigma} \mid \mu, s_j = i\right] \geq 0.$$

We can now follow that all jurors prefer to vote informatively in the first period after receiving an i-signal if and only if $\beta(\lambda^N_{k_2} - 1, N) - c \cdot \alpha(\lambda^N_{k_2} - 1) \leq q_j$ for all $j \in N$.

3. Suppose $\lambda^N_{k_2} \in \{N - k_1, \ldots, N\}$. That implies $d_2 = A$ because $\lambda^N_{k_2} \geq N - k_1 + 1$. We compute j's expected utilities from playing the equilibrium strategy σ and the previously defined deviation $\widetilde{\sigma}$ as follows:

$$\begin{aligned}
EU_j\left[\sigma \mid \mu, s_j = g\right] = &\sum_{x=0}^{k_1-2}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}\left(-(1-q_j)\right) + \sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}\left(-(1-q_j)-c\right)\\
&+\sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-c) + \sum_{x=N-k_1}^{N-1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-q_j)
\end{aligned}$$

$$EU_j\left[\widetilde{\sigma}\mid\mu,s_j=g\right]=\sum_{x=0}^{k_1-1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}(-(1-q_j))+\sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{N-x-1}p^{x+1}\left(-(1-q_j)-c\right)$$
$$+\sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-c)+\sum_{x=N-k_1+1}^{N-1}\binom{N-1}{x}(1-p)^{x+1}p^{N-x-1}(-q_j)$$

Reformulating $q_j\leq\beta(N-k_1+1,N)+c\cdot\alpha(k_1-1)$ yields

$$\begin{aligned}&(1-p)^{N-k_1}p^{k_1}\left(-(1-q_j)-c\right)+(1-p)^{k_1}p^{N-k_1}(-c)+(1-p)^{N-k_1+1}p^{k_1-1}(-q_j)\\ \geq&(1-p)^{N-k_1}p^{k_1}\left(-(1-q_j)\right)+(1-p)^{k_1-1}p^{N-k_1+1}\left(-(1-q_j)-c\right)+(1-p)^{N-k_1+1}p^{k_1-1}(-c)\end{aligned}\tag{11}$$

Analogously to the previous cases, (11) is equivalent to

$$EU_j\left[\sigma\mid\mu,s_j=g\right]-EU_j\left[\widetilde{\sigma}\mid\mu,s_j=g\right]\geq 0.$$

Therefore, all jurors vote informatively after receiving a g-signal if and only if $q_j\leq\beta(N-k_1+1,N)+c\cdot\alpha(k_1-1)$ for all $j\in N$.
Analogously, we compute expected utilities for the case where $s_j=i$.

$$EU_j\left[\sigma\mid\mu,s_j=i\right]=\sum_{x=0}^{k_1-2}\binom{N-1}{x}(1-p)^{N-x}p^{x}(-(1-q_j))+\sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{N-x}p^{x}\left(-(1-q_j)-c\right)$$
$$+\sum_{x=k_1-1}^{N-k_1-1}\binom{N-1}{x}(1-p)^{x}p^{N-x}\left(-c\right)+\sum_{x=N-k_1}^{N-1}\binom{N-1}{x}(1-p)^{x}p^{N-x}\left(-q_j\right)$$
$$EU_j\left[\widetilde{\sigma}\mid\mu,s_j=i\right]=\sum_{x=0}^{k_1-1}\binom{N-1}{x}(1-p)^{N-x}p^{x}(-(1-q_j))+\sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{N-x}p^{x}\left(-(1-q_j)-c\right)$$
$$+\sum_{x=k_1}^{N-k_1}\binom{N-1}{x}(1-p)^{x}p^{N-x}(-c)+\sum_{x=N-k_1+1}^{N-1}\binom{N-1}{x}(1-p)^{x}p^{N-x}(-q_j)$$

From $q_j\geq\beta(N-k_1,N)-c\cdot\gamma(k_1-1)$ we derive the equivalent formulation

$$\begin{aligned}&(1-p)^{N-k_1+1}p^{k_1-1}\left(-(1-q_j)-c\right)+(1-p)^{k_1-1}p^{N-k_1+1}(-c)+(1-p)^{N-k_1}p^{k_1}\left(-q_j\right)\\ \leq&(1-p)^{N-k_1+1}p^{k_1-1}(-(1-q_j))+(1-p)^{k_1}p^{N-k_1}\left(-(1-q_j)-c\right)+(1-p)^{N-k_1}p^{k_1}(-c)\end{aligned}\tag{12}$$

Again, (12) is equivalent to $q_j\geq\beta(N-k_1,N)-c\cdot\gamma(k_1-1)$. As

$$EU_j\left[\sigma\mid\mu,s_j=i\right]-EU_j\left[\widetilde{\sigma}\mid\mu,s_j=i\right]\geq 0$$

is equivalent to (12) in turn, jurors vote informatively after receiving an i-signal if and only if $q_j\geq\beta(N-k_1,N)-c\cdot\gamma(k_1-1)$ for all $j\in N$.

Summing up, (4) - (6) characterize sufficient and necessary conditions for the profile $(\sigma_j,\mu_j)_{j=1}^{N}$ to constitute an equilibrium. □

Proof of Proposition 4.1

Proof. Fix some k_2 for each of the following steps.

Step 1: Any jury that votes informatively in a straw poll in equilibrium does so as well in a setup with some $k_1 > 0$.

Consider $k_1 = 1$.

We show that if jury N votes informatively in $t = 1$ in equilibrium under $k_1 = 0$ then it votes informatively in $t = 1$ in equilibrium as well for $k_1 = 1$.

If jury N votes informatively under $k_1 = 0$ their preferences have minimal diversity of 0. That is, for some $\lambda \in \{1, \ldots, N\}$,

$$\beta(\lambda - 1, N) \leq q_j \leq \beta(\lambda, N) \quad \forall j \in N. \tag{13}$$

Depending on λ there are three cases to consider:

(i) $\lambda \in \{2, \ldots, N-1\}$: For $k_1 = 1$, (σ, μ) constitutes an equilibrium if condition (5) holds. This is implied by (13).

(ii) $\lambda = 1$: We know that

$$\beta(0, N) - c \cdot \alpha(0) \leq q_j \quad \forall j \in N, \tag{14}$$

by Assumption 1. Moreover, from (13) we know

$$q_j \leq \beta(1, N) \leq \beta(1, N) + c \cdot \gamma(1) \quad \forall j \in N. \tag{15}$$

The bounds (14) and (15) coincide with those from (4) which establish informative voting in equilibrium in $t = 1$ for jury N.

(iii) $\lambda = N$: Analogously to (b), we establish the bounds from (6). By Assumption 1 we have

$$q_j \leq \beta(N, N) \leq \beta(N, N) + c \cdot \alpha(0) \quad \forall j \in N, \tag{16}$$

and (13) yields the lower bound

$$\beta(N-1, N) - c \cdot \gamma(1) \leq q_j \quad \forall j \in N. \tag{17}$$

By combining both we establish condition (6) for all $j \in N$ so that voting informatively in $t = 1$ is an equilibrium behavior.

Step 2: Any juror of a jury that votes informatively in a straw poll in equilibrium is strictly better off in the equilibrium constituted by (σ, μ) with some $k_1 > 0$.

We compare the jurors' ex-ante expected utilities from a non-binding straw poll ($k_1 = 0$) and a potentially consequential first period vote with $k_1 = 1$. Note that in both cases the same juror k_2 is pivotal in the second vote. Also by Assumption 1, $\lambda_{k_2} \notin \{0, N+1\}$. For any λ_{k_2}, the final decision is the same for both $k_1 = 0$ and $k_1 = 1$. But in some cases, namely if all jurors receive the same signal, the process in terminated is the first period already under $k_1 = 1$, whereas jurors have to vote again after a straw poll at additional costs c. Therefore, any juror of that homogeneous jury is better off in the equilibrium (σ, μ) with $k_1 = 1$ than in a straw poll setup.

Step 3: With a straw poll in the first period, no juror can get higher expected utility than in the equilibrium with an informative straw poll, i.e., (σ, μ) with $k_1 = 0$.

Informative voting in a straw poll requires a homogeneous jury, where every juror has the same conviction threshold. In equilibrium, jurors reveal their information truthfully and agree on a decision under full information unanimously, because the decision is optimal for each (homogeneous) juror. There are no other sources that impact utilities, in particular jurors cannot agree earlier and save costs. A setup that ensures jurors always the optimal decision from their points of view given full information can not be improved upon. Therefore, the jurors' utility levels in this equilibrium will serve in the following as an upper bound on the jurors' expected utilities in any equilibrium in a straw poll setup.

This upper bound on jurors' expected utilities for any equilibrium of setups with straw polls is given for each $j \in N$ by

$$\sum_{x=0}^{\lambda_j^N-1} \binom{N}{x} (1-p)^{N-x} p^x \left(-\left(1-q_j\right)\right) + \sum_{x=\lambda_j^N}^{N} \binom{N}{x} (1-p)^x p^{N-x} \left(-q_j\right) - 2c. \tag{18}$$

Step 4: Even some heterogeneous juries are strictly better off in the informative equilibrium of a voting setup that admits agreement in the first period than in any equilibrium of a straw poll setup.

Note that his statement is true for homogeneous juries by Step 2 and 3.

Now consider a heterogeneous jury which satisfies one of the conditions of Proposition 3.1 for $k_1 = 1$, so that the jurors vote informatively in the first period. The pivotal juror in the second vote is k_2 and the other jurors' conviction thresholds differ from $\lambda_{k_2}^N$ at most by 1. We show that any juror's expected utility is strictly higher in the informative equilibrium under $k_1 = 1$ than the upper bound (18) established in Step 3.

(i) $\lambda_j^N = \lambda_{k_2}^N$: In equilibrium, j's preferred decision given full information is always implemented by juror k_2 and in some cases he saves costs of time. This is strictly better than the upper bound of any equilibrium in a straw poll setup.

(ii) $\lambda_j^N = \lambda_{k_2}^N - 1$: For a simpler notation we set $\lambda_{k_2}^N \equiv \lambda$ and consider jurors with preferences q_j such that

$$q_j > \beta\left(\lambda-1, N\right) - c \cdot \min\left\{\alpha\left(\lambda-1\right), 2\binom{N}{\lambda-1}^{-1} \frac{(1-p)^N + p^N}{(1-p)^{N-\lambda+1}p^{\lambda-1} + (1-p)^{\lambda-1}p^{N-\lambda+1}}\right\},$$

where $\alpha\left(\lambda-1\right)$ is defined as in Proposition 3.1 for $k_1 = 1$. Note that these jurors vote informatively in equilibrium in the first vote. We compare the expected utility in the informative equilibrium with $k_1 = 1$ to the upper bound of any straw poll setup (18), in which every juror votes twice but his preferred decision given full information is made in the second vote for sure. When early agreement is possible, jurors might save costs of time but the jury's decision is sub optimal for jurors with $\lambda_j < \lambda_{k_2}^N$ if the number of g-signals among all jurors is between λ_j and $\lambda_{k_2}^N$.

The net effect on expected utilities from early agreement of the gain by cost saving and the loss by a sub optimal decision is positive for jurors with preferences as specified above, because

$$2c\cdot\left[(1-p)^N+p^N\right]+\binom{N}{\lambda-1}\left[(1-p)^{N-\lambda+1}p^{\lambda-1}(-(1-q_j))-(1-p)^{\lambda-1}p^{N-\lambda+1}(-q_j)\right]>0$$

$$\Leftrightarrow q_j>\beta\left(\lambda-1,N\right)-2c\cdot\binom{N}{\lambda-1}^{-1}\frac{(1-p)^N+p^N}{(1-p)^{N-\lambda+1}p^{\lambda-1}+(1-p)^{\lambda-1}p^{N-\lambda+1}}.$$

(iii) $\lambda_j^N=\lambda_{k_2}^N+1$: Consider jurors with preferences q_j such that

$$q_j<\beta\left(\lambda,N\right)+c\cdot\min\left\{\gamma\left(\lambda\right),2\binom{N}{\lambda}^{-1}\frac{(1-p)^N+p^N}{(1-p)^{N-\lambda}p^{\lambda}+(1-p)^{\lambda}p^{N-\lambda}}\right\},$$

where $\gamma\left(\lambda\right)$ is defined as in Proposition 3.1 for $k_1=1$. Analogously to the previous case, the net effect on expected utilities from early agreement of the gain by cost saving and the loss by a sub optimal decision is positive for these jurors, because

$$2c\cdot\left[(1-p)^N+p^N\right]+\binom{N}{\lambda}\left[(1-p)^{\lambda}p^{N-\lambda}(-q_j)-(1-p)^{N-\lambda}p^{\lambda}(-(1-q_j))\right]>0$$

$$\Leftrightarrow q_j<\beta\left(\lambda,N\right)-2c\cdot\binom{N}{\lambda}^{-1}\frac{(1-p)^N+p^N}{(1-p)^{N-\lambda}p^{\lambda}+(1-p)^{\lambda}p^{N-\lambda}}.$$

□

Proof of Corollary 4.1

Proof. We know from the proof Proposition 4.1 that juries which vote informatively with $k_1=1$ are strictly better off than with a straw poll in the first period. In addition, any higher value of k_1 does not provide incentives to vote informatively to any homogeneous jury as a straw poll does. Suppose the designer sets $k_1=2$ and consider a homogeneous jury with common conviction threshold of $\lambda_j^N=1$ for all $j\in N$, i.e.,

$$\beta(0,N)\leq q_j\leq\beta(1,N)\quad\forall j\in N.$$

By (4), jurors vote informatively in the first vote with $k_1=2$ if and only if

$$\beta(1,N)-c\cdot\alpha(1)\leq q_j\leq\beta(2,N)+c\cdot\gamma(2)\quad\forall j\in N.$$

As c is unknown to the designer, she cannot rule out that

$$q_j\in\left[\beta(0,N),\beta(1,N)-c\right),$$

in which case this jury is worse off compared to a straw poll. This argument holds for any higher $k_1>0$ as well. □

Proof of Proposition 4.2

Proof. We proceed in three steps.

Step 1: Definition of $\mathcal{K}_1^N$.
Denote by $\mathcal{K}_1^N$ the set of all $k_1<N/2$ for which jury N votes informatively in the first vote.

Definition 3. $k_1 \in \mathcal{K}_1^N$ if and only if for all $j \in N$, q_j satisfies one of the conditions (4), (5) or (6) of Proposition 3.1 for k_1.

The set $\mathcal{K}_1^N$ is non-empty if the jury's heterogeneity is bounded from above by an according value which is assumed.

Step 2: Definition of $\widetilde{\mathcal{K}}_1^N$.

Denote by $\widetilde{\mathcal{K}}_1^N$ the set of all $k_1 \in \mathcal{K}_1^N$ which make all jurors $j \in N$ (weakly) better off compared to the best equilibrium outcome of a straw poll in the first period. Recall from the proof of Proposition 4.1, the best equilibrium for any juror in a setup with a straw poll is that of a homogeneous jury, where information is aggregated in the straw poll and every jury agrees to the jury's decision in $t = 2$. The expected utility in that equilibrium for any $j \in N$ is given by (18).

The expected utility of any $j \in N$ from (σ, μ) and $k_1 > 0$ depends on the relative position of k_1 and $\lambda_{k_2}^N$. We have to distinguish three cases.

(i) For $\lambda_{k_2}^N < k_1$, j's expected utility from (σ, μ) is given by

$$\sum_{x=0}^{k_1-1} \binom{N}{x} (1-p)^{N-x} p^x \left(-(1-q_j)\right) + \sum_{x=k-1}^{N} \binom{N}{x} (1-p)^x p^{N-x} \left(-q_j\right)$$
$$- \left(1 + \sum_{x=k_1}^{N-k_1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)\right) c. \tag{19}$$

(ii) For $\lambda_{k_2}^N \in \{k_1, \ldots, N-k_1\}$, j's expected utility from (σ, μ) is given by

$$\sum_{x=0}^{\lambda_{k_2}^N-1} \binom{N}{x} (1-p)^{N-x} p^x \left(-(1-q_j)\right) + \sum_{x=\lambda_{k_2}^N}^{N} \binom{N}{x} (1-p)^x p^{N-x} \left(-q_j\right)$$
$$- \left(1 + \sum_{x=k_1}^{N-k_1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)\right) c. \tag{20}$$

(iii) For $\lambda_{k_2}^N > N-k_1$, j's expected utility from (σ, μ) is given by

$$\sum_{x=0}^{N-k_1} \binom{N}{x} (1-p)^{N-x} p^x \left(-(1-q_j)\right) + \sum_{x=N-k+1}^{N} \binom{N}{x} (1-p)^x p^{N-x} \left(-q_j\right)$$
$$- \left(1 + \sum_{x=k_1}^{N-k_1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)\right) c. \tag{21}$$

We can now derive conditions for which the equilibrium (σ, μ) with $k_1 > 0$ is (weakly) better for any $j \in N$ than the best equilibrium with a straw poll.

(i) For $\lambda_{k_2}^N < k_1$, (19) $\geq$ (18) for all $j \in N$ if and only if

$$q_j \geq \frac{\sum_{x=\lambda_j^N}^{k_1-1} \binom{N}{x} (1-p)^{N-x} p^x}{\sum_{x=\lambda_j^N}^{k_1-1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)} - 2c \cdot \frac{\sum_{x=0}^{k_1-1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)}{\sum_{x=\lambda_j^N}^{k_1-1} \binom{N}{x} \left((1-p)^{N-x} p^x + (1-p)^x p^{N-x}\right)}$$
$$\forall j \in N. \tag{22}$$

(ii) For $\lambda_{k_2}^N \in \{k_1, \dots, N - k_1\}$, (20) $\geq$ (18) for all $j \in N$ if and only if

$$q_j \geq \frac{\sum_{x=\lambda_j^N}^{\lambda_{k_2}^N - 1} \binom{N}{x}(1-p)^{N-x}p^x}{\sum_{x=\lambda_j^N}^{\lambda_{k_2}^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)} - 2c \cdot \frac{\sum_{x=0}^{k_1 - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}{\sum_{x=\lambda_j^N}^{\lambda_{k_2}^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}, \tag{23}$$

for $j \in N$ with $\lambda_j^N < \lambda_{k_2}^N$, and

$$q_j \leq \frac{\sum_{\lambda_{k_2}^N}^{\lambda_j^N - 1} \binom{N}{x}(1-p)^{N-x}p^x}{\sum_{\lambda_{k_2}^N}^{\lambda_j^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)} + 2c \cdot \frac{\sum_{x=N+k_1+1}^{N} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}{\sum_{\lambda_{k_2}^N}^{\lambda_j^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}, \tag{24}$$

for $j \in N$ with $\lambda_j^N > \lambda_{k_2}^N$.

(iii) For $\lambda_{k_2}^N > N - k_1$, (21) $\geq$ (18) for all $j \in N$ if and only if

$$q_j \leq \frac{\sum_{x=N-k_1+1}^{\lambda_j^N - 1} \binom{N}{x}(1-p)^{N-x}p^x}{\sum_{x=N-k_1+1}^{\lambda_j^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)} + 2c \cdot \frac{\sum_{x=N+k_1+1}^{N} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}{\sum_{x=N-k_1+1}^{\lambda_j^N - 1} \binom{N}{x}\left((1-p)^{N-x}p^x + (1-p)^x p^{N-x}\right)}$$

$$\forall j \in N. \tag{25}$$

Note that (20) $\geq$ (18) holds for all $j \in N$ with $\lambda_j^N = \lambda_{k_2}^N$.

From here we can define the set $\widetilde{\mathcal{K}}_1^N$ formally.

Definition 4. $k_1 \in \widetilde{\mathcal{K}}_1^N$ if and only if

- $k_1 \in \mathcal{K}_1^N$, and
- either (22), or (23) and (24), or (25) holds for k_1.

Note, that $\widetilde{\mathcal{K}}_1^N$ is non-empty whenever the jury's heterogeneity is bounded from above sufficiently as assumed.

Step 3: Conditions for the optimal choice of k_1 from $\widetilde{\mathcal{K}}_1^N$.

For any $k_1 \in \widetilde{\mathcal{K}}_1^N$, (σ, μ) is an equilibrium and all jurors $j \in N$ are better off than in any equilibrium in setups with a straw poll in the first period. We now prove that the optimal $k_1 \in \widetilde{\mathcal{K}}_1^N$ is determined as in the Proposition.

1. If for all $k_1 \in \widetilde{\mathcal{K}}_1^N$ it holds that $\lambda_{k_2}^N \in \{k_1, \dots, N - k_1\}$, it is straightforward from comparing (20) and (18), that it is optimal to set

$$k_1 = \max \widetilde{\mathcal{K}}_1^N.$$

2. If $\lambda_{k_2}^N < \frac{N}{2}$ and $\lambda_{k_2}^N < k_1$ for some $k_1 \in \widetilde{\mathcal{K}}_1^N$ then the designer faces a trade-off. By the previous argument a natural candidate is

$$k_1 = \max\left\{k_1 \mid \lambda_{k_2}^N \in \{k_1, \dots, N - k_1\}\right\} \equiv k^m.$$

Increasing k_1 by 1 does decrease the expected costs but changes the jury's decision. For any $k_1 \geq \lambda_{k_2}^N$, an increase to k_1+1 leads to $d_1 = A$ instead of $d_1 = D$ and $d_2 = C$ for $x_1 = k_1$. Therefore, in this case each juror saves expected costs of

$$2c \cdot \binom{N}{k_1} \left((1-p)^{N-k_1} p^{k_1} + (1-p)^{k_1} p^{N-k_1}\right)$$

and, because a false judgment can be avoided if $\omega = I$, each juror additionally saves in expectation

$$q_j \cdot \binom{N}{k_1} (1-p)^{N-k_1} p^{k_1}.$$

On the other hand, if $\omega = I$ a false judgment is enacted which yields in expectation a loss of

$$(1-q_j) \cdot \binom{N}{k_1} (1-p)^{k_1} p^{N-k_1}.$$

Expected gains are higher than expected losses from an increase of k_1 to k_1+1 for each juror if and only if

$$q_j \geq \beta(k_1, N) - 2c \quad \forall j \in N. \tag{26}$$

Moreover, note that it can not be the case that an increase from k_1+1 to k_2+2 is profitable to all jurors but not from k_1 to k_1+1. Suppose to the contrary that this was the case. Using (26) would yield the contradiction $\beta(k_1, N) - 2c \geq q_j \geq \beta(k_1+1, N) - 2c$.

Therefore, the following choice of $k-1$ is optimal. Set the highest k_1 for which $\lambda_{k_2}^N \in \{k_1, \ldots, N-k_1\}$ if (26) is not satisfied for this value. If it is, however, set the highest k_1 for which $k_1 - 1$ does satisfy (26). Formally, set

$$k_1 = \max\left\{k_1 \in \widetilde{\mathcal{K}}_1^N \mid q_1 \geq \beta(k_1-1, N) - 2c \vee \lambda_{k_2}^N \in \{k_1, \ldots, N-k_1+1\}\right\}.$$

As condition (26) is most binding for q_1, it suffices to consider juror 1 only.

3. If $\lambda_{k_2}^N > \frac{N}{2}$ and $\lambda_{k_2}^N > N - k_1$ for some $k_1 \in \widetilde{\mathcal{K}}_1^N$ then an analogous argumentation to the previous case applies. The difference is that an increase from $k_1 \geq N - \lambda_{k_2}^N$ to k_1+1 now changes the decision for $x_1 = N - k_1$. As a result the designer faces an adjusted trade-off for each juror. An increase from k_1 to k_1+1 yields expected gains of

$$2c \cdot \binom{N}{k_1} \left((1-p)^{N-k_1} p^{k_1} + (1-p)^{k_1} p^{N-k_1}\right) + (1-q_j) \cdot \binom{N}{k_1} (1-p)^{k_1} p^{N-k_1},$$

whereas expected losses are

$$q_j \cdot \binom{N}{k_1} (1-p)^{N-k_1} p^{k_1}.$$

Expected gains are higher than expected loss if and only if

$$q_j \leq \beta(N-k_1, N) - 2c \quad \forall j \in N. \tag{27}$$

By the same argument as before, it is therefore optimal for the designer to set

$$k_1 = \max\left\{k_1 \in \widetilde{\mathcal{K}}_1^N \mid q_N \leq \beta(N-(k_1-1), N) + 2c \vee \lambda_{k_2}^N \in \{k_1, \ldots, N-k_1+1\}\right\}$$

□

References

Austen-Smith, D. (1990). Information transmission in debate. *American Journal of Political Science*, 34(1):pp. 124–152.

Austen-Smith, D. and Banks, J. S. (1996). Information aggregation, rationality, and the condorcet jury theorem. *The American Political Science Review*, 90(1):34–45.

Austen-Smith, D. and Feddersen, T. J. (2006). Deliberation, preference uncertainty, and voting rules. *American Political Science Review*, null:209–217.

Bognar, K., ter Vehn, M. M., and Smith, L. (2013). A conversational war of attrition.

Condorcet, M. d. (1785). Essay on the application of analysis to the probability of majority decisions. *Paris: Imprimerie Royale*.

Coughlan, P. J. (2000). In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *The American Political Science Review*, 94(2):375–393.

Damiano, E., Li, H., and Suen, W. (2010). Delay in strategic information aggregation.

Damiano, E., Li, H., and Suen, W. (2012a). Optimal deadlines for agreements. *Theoretical Economics*, 7(2):357–393.

Damiano, E., Li, H., and Suen, W. (2012b). Optimal delay in committees.

Deimen, I., Ketelaar, F., and Thordal-Le Quement, M. (2013). Consistency and communication in committees.

Feddersen, T. J. and Pesendorfer, W. (1998). Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *The American Political Science Review*, 92(1):23–35.

Goeree, J. K. and Yariv, L. (2011). An experimental study of collective deliberation. *Econometrica*, 79(3):893–921.

Hummel, P. (2012). Deliberation in large juries with diverse preferences. *Public Choice*, 150(3-4):595–608.

Ladha, K. K. (1992). The condorcet jury theorem, free speech, and correlated votes. *American Journal of Political Science*, 36(3):pp. 617–634.

Meirowitz, A. (2007). In defense of exclusionary deliberation: Communication and voting with private beliefs and values. *Journal of Theoretical Politics*, 19(3):301–327.

Miller, N. R. (1986). Information, electorates, and democracy: some extensions and interpretations of the condorcet jury theorem. *Information pooling and group decision making*, 2:173–192.

Morgan, J. and Stocken, P. C. (2008). Information aggregation in polls. *The American Economic Review*, 98(3):864–896.

Piketty, T. (1999). The information-aggregation approach to political institutions. *European Economic Review*, 43(4–6):791 – 800.

Piketty, T. (2000). Voting as communicating. *The Review of Economic Studies*, 67(1):169–191.

Robert III., H. M., Evans, W. J., Honemann, D. D., and Balch, T. J. (2000). *Robert's Rules of Order Newley Revised*. Robert's Rules of Order. Perseus Publishing, Cambridge, Massachusetts, 10th edition.

Thordal-Le Quement, M. (2013). Communication compatible voting rules. *Theory and Decision*, 74(4):479–507.

Thordal-Le Quement, M. and Yokeeswaran, V. (2013). Subgroup deliberation and voting.

Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review*, 82:1231–1244.