Discussion Paper Series – CRC TR 224

Discussion Paper No. 328
Project B 04

Voting With Endogenous Timing

Finn Schmieter[1]

January 2022

[1] Department of Economics , University of Bonn,
Email: finn.schmieter@uni-bonn.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Voting with Endogenous Timing *

Finn Schmieter †

December 21, 2021

Abstract

This paper analyses the role of timing in common-value elections. There are two voting periods where voters can decide for themselves when to publicly cast their votes after receiving private signals. In welfare-optimal equilibria, agents use their timing to communicate the strength of their private information to the other voters. This communication allows for better information aggregation than simultaneous voting or voting with exogenously fixed timing. In the case of a simple majority voting rule, a second voting period mitigates the Swing Voter's Curse more effectively than abstention.

JEL Classification: D72, D82, D83

Keywords: Elections, Pivotal Voting, Communication, Information

*I am grateful to Stephan Lauermann and Francesc Dilmé for their guidance, support, and helpful comments. I also thank Tobias Rachidi, Axel Niemeyer, Kailin Chen, Justus Preusser, Thomas Kohler, Christina Luxen, Andre Speit, and seminar participants at the University of Bonn. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B04) is gratefully acknowledged. I also benefited from the infrastructure and financial support of the Bonn Graduate School of Economics.

†University of Bonn, Department of Economics, finn.schmieter@uni-bonn.de

1 Introduction

In many collective decision processes agents can and do preemptively announce their positions. For instance, when a parliament votes on a bill, participating politicians often publicly disclose their stances beforehand. An example was the second impeachment trial of Donald Trump, where various senators spoke out in favor or against an impeachment before the official vote even began.[1] This kind of information disclosure can affect the decision-making. It discloses information to the other politicians about the own opinion on the bill. The other politicians can react to this information and adjust their behavior accordingly. Additionally, voters can anticipate this effect and may use their own vote to influence the other voters.

In this chapter, we analyze how the possibility to disclose the own action and thereby inform the other voters can affect a voting procedure and the associated information aggregation.[2] We provide a stylized model of sequential voting with common values, two voting periods, and endogenous timing. The voters have to decide between two options, and every voter receives a private signal about which option is more preferable. We model a prior announcement of the own vote as binding.[3] We restrict the analysis to homogeneous preferences throughout this chapter. The voters all agree on the best decision for each state of the world and get the same utility, but their private information about the state have different realizations.

The main trade-off for each voter arises from the timing decision: The voter can vote early and disclose her vote to the other voters. This informs others and allows them to make better-informed voting decisions but the voter cannot observe other votes herself. Alternatively, the voter can vote late and first observe the other voters' early votes. This provides additional information to the voter and allows her to make a better-informed voting decision but in return she cannot inform others. The more informative the votes are in period one, the higher is the incentive to vote in period two.

We start by showing the existence of a welfare-optimal equilibrium. Due to the infinite type space and the sequential voting structure, we construct a non-standard metric on the strategy-space for this. Then, we characterize the welfare-optimal equilibria of the two-period voting game. Similarly to simultaneous voting studied by Duggan and Martinelli (2001), the strategies of a welfare-optimal equilibrium follow a cutoff rule. In the first period, voters with more informative signals cast an early vote to influence other voters in their direction. Voters with less informative

[1] The senators' positions were prominently announced in the media at that time. See for example CNN Politics (2021) or Zurcher, Anthony (2021).

[2] There are various other effects associated to a preemptive disclosure. For example, a prior announcement of the own vote informs the citizens about the political agenda and increases transparency. Politicians can use this for reputation-building as described by Keefer and Vlaicu (2007) who analyze the role of credibility and reputation in democracies.

[3] Even though a public disclosure of the own vote is only a partial commitment, it is strong in the sense that politicians generally care about their reputation, and deviating from an announcement may lead to a loss of reputation.

signals wait for period two to get more information before voting.

We show that the welfare-optimal equilibria of our sequential voting model with endogenous timing welfare-dominate all equilibria of simultaneous voting games and voting games with exogenously fixed voting sequences. More precisely, it is the combination of the timing decision and the voting decision that conveys useful information to the other voters.

In setups where the swing voter's curse[4] occurs, voting with endogenous timing mitigates its negative effect on welfare, even outperforming simultaneous voting with abstention.

Moreover, information is aggregated even under assumptions for which the simultaneous voting model fails to do so. In particular, even in a setting with bounded signals and under the unanimity voting rule, the probability of the correct decision under a welfare-optimal equilibrium of our two-period voting game converges to one as the number of voters grows large.

Our results contrast the result of Dekel and Piccione (2000) who show that in general if the timing is exogenous, the disclosure of the votes alone does not improve the information aggregation of a voting procedure compared to simultaneous voting. The reason is that learning the other agents' votes only changes the probability of being pivotal but not the optimal action upon being pivotal. Instead, if the timing is endogenous, agents cannot only use the vote itself but also the timing of the vote to convey information about the strength of the own signal to the other voters. As a result, endogenizing timing improves the outcome of a voting procedure.

The rest of this chapter is organized as follows. Section 1.1 gives an overview over the related literature. Section 2 lays out the model with two periods. In Section 3, an example illustrates the model and the voter's behavior. Section 4 contains the main analysis and characterizes the welfare-optimal equilibria. Section 5 covers information aggregation and Section 6 relates sequential voting to the swing voter's curse under the simple majority voting rule. Section 7 shows that voting with endogenous timing welfare-dominates a voting procedure with a fixed voting sequence and Section 8 concludes. The proofs of the results can be found in Appendix A.

1.1 Related Literature

Our model is related to the Condorcet Jury Theorem and information aggregation in large elections. Condorcet (1785) suggested that for homogeneous preferences, a decision made by a large group of "sincere" voters yields better results than a decision made by an individual alone. This result was later reproduced for strategic voters in simultaneous voting procedures.[5]

[4] If ties are randomly broken, less informed voters may strictly prefer to abstain rather than to vote. See Feddersen and Pesendorfer (1996) for more details on the swing voter's curse.

[5] Among others, Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1996, 1997, 1998) and Duggan and Martinelli (2001) analyze strategic voting.

Feddersen and Pesendorfer (1998) show that even for large electorates information is not aggregated under the unanimity voting rule for binary signals due to the bounded informativeness. In contrast to their result, the probability of choosing the optimal decision converges to one in our two-period voting model with endogenous timing.

Previous work on sequential voting has mainly focused on exogenously fixed voting sequences. We contribute to this line of research by endogenizing the timing decision in a sequential voting game. This work builds on Schmieter (2019), where the welfare-optimality of cutoff rules is shown for the special case of the unanimity voting rule.

Dekel and Piccione (2000) consider sequential voting with two alternatives where the order of voting is exogenously fixed. In their setting, voters cast their vote in a given order, and every voter observes all actions that have been made prior to her vote. They show that each symmetric equilibrium of the corresponding simultaneous voting game is also an equilibrium of any sequential voting game, regardless of the voting sequence. Furthermore, they prove that under the unanimity voting rule, the set of equilibria of any sequential voting game, regardless of the voting sequence, is equal to the set of equilibria of the corresponding simultaneous voting game. An important implication from their work is that observing the votes of the other voters does not improve the aggregation of information. This is due to the fact that voters condition on the event of being pivotal. Since in their model there exists exactly one event for which a voter is pivotal, this conditioning is equivalent to observing the other agents' votes directly. Thus, learning the votes of other agents does not convey useful information. In particular, learning the earlier voters' actions does not change the behavior of the later voters. However, except for the unanimity voting rule, they do not show whether a new equilibrium of the sequential voting game might welfare-dominate the equilibria of the simultaneous voting game. Also, their equivalence result under the unanimity voting rule relies on the exogeneity of the voting sequence.

One crucial aspect of Dekel and Piccione (2000) is that they do not allow for any tie-breaking in their model. Instead, they restrict their analysis to n_p-voting rules, where alternative one is adopted if and only if at least n_p voters vote for it and alternative two is chosen otherwise. Their result does for example not carry over to a simple majority voting rule with tie-breaking by a fair coin toss. In particular, their analysis excludes settings where the so-called swing voter's curse occurs: Feddersen and Pesendorfer (1996) show that under the simple majority voting rule with an even number of voters and tie-breaking by a fair coin toss, less informed voters strictly prefer to abstain. As a result, allowing abstention in such simultaneous voting settings increases welfare. We show that the welfare-optimal equilibrium of our two-period model without abstention welfare-dominates all equilibria of the simultaneous voting model with abstention.

Dekel and Piccione (2014) analyze voting with an endogenous timing decision. Compared to our model, they cover three alternatives, private values, and voters have to decide for a voting period before they learn their preferences. In particular, voters in their model have a conflict of interest, and, in contrast to our model, revealing information to other voters can have a negative effect for oneself.

There are various other papers related to sequential voting. Battaglini (2005) adds abstention and costs of voting to the model of Dekel and Piccione (2000) and shows that even arbitrarily small voting costs can break the equivalence of equilibria. Another strand of literature analyzes herding behavior (see for example Fey (1998)), where herding hinders full information aggregation. The difference to sequential voting is that herding features an individual payoff relevant choice for each agent instead of a collective decision. Callander (2002) relates herding to sequential voting and shows that if voters want to vote for the winning candidate, herding occurs with probability one. Eyster and Rabin (2005) introduce the concept of cursed equilibria, where agents underestimate the correlation of other players' information. Piketty (2000) considers two-period voting, where both periods are payoff-relevant. There, agents of different types use the first period to signal information and influence the outcome of the second voting period. In contrast to our model, there are three competing candidates and the voters are confronted with a coordination problem rather than a problem of information aggregation. McLennan (1998) shows that for common interest games, a symmetric strategy that maximizes the expected welfare is a Nash equilibrium. We use this finding multiple times to prove our results.

2 The Model

In this section, we introduce our model of sequential voting with two voting periods. There are $N \geq 2$ jurors who vote on whether to convict or acquit a defendant.[6] An unknown state ω describes whether the defendant is innocent, I, or guilty, G. The realization of ω is randomly drawn according to a commonly known prior $q := P(\omega = I)$ and $1 - q = P(\omega = G)$ with $q \in (0, 1)$.

Each agent $i \in \{1, \ldots, N\}$ receives a private signal s_i about ω from the closed interval $S := [\underline{s}, \bar{s}] \subseteq \mathbb{R}$. Conditional on the state, the signals are drawn independently from each other, according to the cumulative distribution function $F(\cdot|I)$ if the defendant is innocent or $F(\cdot|G)$ if the defendant is guilty. The distribution functions $F(\cdot|I)$ and $F(\cdot|G)$ are absolutely continuous and have piecewise continuous densities $f(\cdot|I)$ and $f(\cdot|G)$ which are strictly positive on S.

We assume that the *likelihood ratio of the signals*, $f(s|I)/f(s|G)$, is weakly decreasing on S. This implies that low signals indicate innocence, while high signals

[6] To simplify the exposition, we frame the model as if it were about a decision at court, but it is in no way restricted to this particular application. The notation mainly follows Duggan and Martinelli (2001).

are indicators of guilt. We let the signals be sufficiently informative by assuming that both events

$$\left\{s \in S \;\middle|\; \frac{f(s|I)}{f(s|G)} < \frac{1-q}{q}\right\} \text{ and } \left\{s \in S \;\middle|\; \frac{f(s|I)}{f(s|G)} > \frac{1-q}{q}\right\}$$

occur with positive probability. That is, the likelihood ratio of a single signal can dominate the likelihood ratio of the prior in either direction.

Preferences and Timing The defendant can be either acquitted, A, or convicted, C. The agents have common preferences and want to match the outcome with the state. They get a utility of 1 if C is implemented in state G or if A is implemented in state I and a utility of 0 otherwise.

The outcome is determined by a voting procedure with two voting periods, period one (*early*) and period two (*late*). In period one, the agents can either vote for A or C or choose to wait, denoted by W. In period two, the agents who waited observe the aggregated votes from period one and now have to vote for either A or C themselves. Abstention is not allowed in period two. Agents who already voted in period one cannot change their decision anymore and are not allowed to vote a second time.

The *voting rule* is parameterized by a pair $(K, p) \in \{1, 2, \ldots, N-1\} \times [0, 1]$. If strictly less than K voters vote for conviction, then the defendant is acquitted and if strictly more than K voters vote for conviction, then the defendant is convicted. If the number of C-votes is exactly K, then conviction occurs with probability p.[7] This captures all standard (anonymous) voting rules such as the unanimity voting rule, all super-majority voting rules, and the simple majority voting rule with and without random tie-breaking. For example, for the parameters $(N-1, 0)$, the defendant is only convicted if all N voters vote unanimously for C, i.e., we have the unanimity voting rule. With an even number of voters N, the voting rule $(\frac{N}{2}, \frac{1}{2})$ represents the simple majority voting rule where a tie is broken by a fair coin flip.

Histories, Strategies, and Equilibria A (public) *history* h specifies the past voting actions. Let $h = \emptyset$ denote the empty history at the beginning of period one. A history in period two can be characterized by a pair $h = (n_A, n_C)$ that specifies the number n_A of early A-votes and the number n_C of early C-votes. Let H be the set of all histories.[8]

A *mixed strategy* for voter i is given by the probabilities of voting for A, waiting W, and voting for C for every private signal $s \in S$ and every history $h \in H$.

[7] Note that the voting rules $(K, 0)$ and $(K+1, 1)$ are equivalent.

[8] Formally, $H = \left\{(n_A, n_C) \in \{0, 1, \ldots, N-1\}^2 \mid n_A + n_C < N,\ n_C \leq K,\ n_A \leq N-K\right\} \cup \{\emptyset\}$.

Formally, a mixed strategy is a measurable[9] function

$$\sigma^i : S \times H \to \{(p_A, p_W, p_C) \in [0,1]^3 \mid p_A + p_W + p_C = 1\},$$

with $p_W = 0$ for every history of period two. The triple $\sigma^i(s_i, h) = (p_A, p_W, p_C)$ specifies the probabilities p_Y of playing action Y for each $Y \in \{A, W, C\}$ for every signal $s_i \in S$ and at every history $h \in H$. Let σ^i_A, σ^i_W and σ^i_C be the marginals of σ^i, i.e., the maps to p_A, p_W and p_C, respectively. For convenience, let $\sigma^i(s_i, h) = A$, $\sigma^i(s_i, h) = W$, and $\sigma^i(s_i, h) = C$ denote that the corresponding actions are played with probability 1.

Fix a single voter i, fix a strategy σ^i and the strategies σ^{-i} of the other voters. For these strategies, the expected utility for voter i is given by

$$U(\sigma^i, \sigma^{-i}) = qP(A|I, \sigma^i, \sigma^{-i}) + (1-q)P(C|G, \sigma^i, \sigma^{-i}),$$

where $P(Y|\omega, \sigma^i, \sigma^{-i})$ denotes the probability of outcome Y under state ω given the strategies σ^i and σ^{-i}, i.e., the expected utility is the ex-ante probability of choosing the correct outcome. The strategies $(\sigma^1, \ldots, \sigma^N)$ constitute a mixed Bayesian Nash equilibrium if for every voter i, the strategy σ^i maximizes $U(\sigma^i, \sigma^{-i})$ for fixed σ^{-i}. We restrict attention to symmetric strategies and omit the index i to write σ/σ_Y for the strategies/marginals instead of σ^i/σ^i_Y.

As the expected utility in an equilibrium is identical for every voter, we consider welfare on a per-capita level and call it the *expected welfare* $U(\sigma)$. A *welfare-optimal equilibrium* is an equilibrium that maximizes the welfare, or equivalently, the ex-ante probability of a correct decision. Unless stated otherwise, "equilibrium" refers to symmetric mixed Bayesian Nash equilibrium.[10]

Assumptions For some results, we additionally assume that the following properties hold. Their usage is explicitly stated each time. The first assumption says that the likelihood ratio is strictly decreasing instead of weakly decreasing. That is, no two signals induce the same belief.

Strictly monotone likelihood ratio property (MLRP$_<$). The likelihood ratio of the signals, $f(s|I)/f(s|G)$, is strictly decreasing on S.

The second assumption states that the informativeness of the signals is un-

[9]Let $\mathcal{B}(S)$ and $\mathcal{B}([0,1]^3)$ denote the Borel σ-algebras on S and $[0,1]^3$, respectively. Consider the power set $\mathcal{P}(H)$, which is a σ-algebra on the finite set H. A strategy σ^i is required to be measurable with respect to the product σ-algebra $\Sigma = \mathcal{B}(S) \times \mathcal{P}(H)$ and $\mathcal{B}([0,1]^3)$.

[10]We will later see that in a welfare-optimal Bayesian Nash equilibrium $p_\sigma(Y|\emptyset, \omega) \in (0,1)$ holds for all $Y \in \{A, W, C\}$ and $\omega \in \{I, G\}$, i.e., agents wait with positive probability. Therefore, all public histories are reached with strictly positive probability, and the beliefs are determined by Bayes' rule. For improved readability, we omit the beliefs and consider Bayesian Nash equilibria instead of perfect Bayesian equilibria throughout this chapter.

bounded. This is for convenience only and ensures that all cutoffs are in the interior of S, thus avoiding the need to consider corner cases.

Unbounded likelihood ratio (ULR). The likelihood ratio of the signals is unbounded, i.e.,

$$\lim_{s \to \underline{s}} \frac{f(s|I)}{f(s|G)} = \infty$$

$$\lim_{s \to \overline{s}} \frac{f(s|I)}{f(s|G)} = 0.$$

Monotonicity A strategy profile is *monotone* if voting A or C in period one increases the probability of the respective outcome regardless of the state ω. We will see that this class of strategy profiles has multiple desirable properties. First, in monotone equilibria, the agents' votes and beliefs are aligned: If an agent knew which outcome is correct, then she would vote for that outcome. Second, for the number of voters being large, there are monotone equilibria that implement the correct outcome with a probability close to one. Therefore, under the viewpoint of information aggregation, it is without loss to restrict attention to monotone equilibria. Third, monotone equilibria are a generalization of cutoff equilibria of the simultaneous voting game and they will be particularly straightforward to work with.

Formally, monotonicity is defined as follows. Fix a strategy profile σ of the two-period game. Let $P(Y, \omega)$ denote the probability that the defendant is convicted given that the state is ω and given that a voter i votes for $Y \in \{A, W, C\}$ in the first stage and that the remaining $N-1$ voters follow strategy σ. If voter i waits in period one, then she also follows strategy σ in period two. Now, the strategy profile σ is called monotone if the inequalities

$$P(A, I) \leq P(W, I) \leq P(C, I)$$

$$P(A, G) \leq P(W, G) \leq P(C, G)$$

hold, i.e., the probability of conviction is monotone increasing in the actions A, W, and C.

An illustration of a monotone strategy profile is the following strategy profile where agents follow cutoff rules: Agents with a strong signal towards innocence vote for acquittal and agents with a strong signal towards guilt vote for conviction in period one. Agents with intermediate signals wait in period one and vote in period two, conditioning on the own signal and the observed votes (see Figure 1).

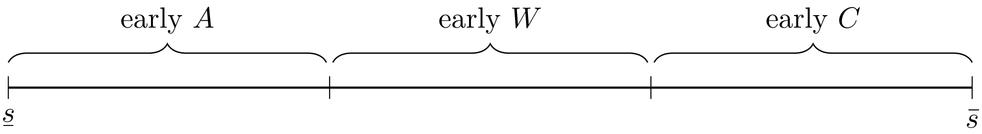

Figure 1: Example of a monotone strategy

Monotonicity rules out strategy profiles where the meanings of the votes are reversed, e.g., strategy profiles where voting early for conviction actually decreases the probability of conviction.

One example of a strategy profile that is not monotone is illustrated in Figure 2. Voters with low/intermediate signals vote for acquittal/conviction in period one and voters with high signals wait in period one. Voters understand waiting as a strong signal towards guilt and vote in period two according to their updated beliefs. For a given number of early A-votes, the lower the number of early C-votes is, the more the agents update their beliefs towards G. Therefore, voting early for C can actually decrease the probability of C being the outcome. Depending on the parameters, there can exist strategy profiles of this form that constitute equilibria.

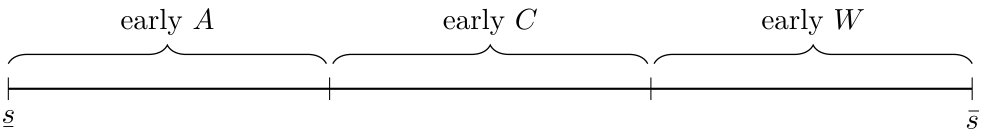

Figure 2: Example of a non-monotone strategy

Note that non-monotone equilibria can only exist due to the sequential nature of the voting procedure and are not possible in related simultaneous voting games. For the remainder of this chapter, we restrict our attention to the class of monotone equilibria.

Derived Terms We conclude this section by defining and deriving some technical terms for later use. Fix some strategy σ^i. The probability that voter i votes for $Y \in \{A, W, C\}$ in period one, conditional on the state ω, is obtained by integrating the marginal σ^i_Y over all signals, i.e.,

$$p_{\sigma^i}(Y|\emptyset, \omega) := \int_S \sigma^i_Y(s, \emptyset) dF(s|\omega).$$

Now, assume that waiting occurs with positive probability and consider an agent i who waited in period one. Then, the probability $p_{\sigma^i}(Y|h, \omega)$ of agent i voting for $Y \in \{A, C\}$ at history $h \neq \emptyset$ given that the state is ω is

$$p_{\sigma^i}(Y|h, \omega) := \frac{\int_S \sigma^i_Y(s, h)\sigma^i_W(s, \emptyset) dF(s|\omega)}{p_{\sigma^i}(W|\emptyset, \omega)}.$$

Furthermore, let

$$G_{\sigma^i}(s|\omega) := \frac{\int_{\underline{s}}^{s} \sigma_W^i(s', \emptyset) dF(s'|\omega)}{p_{\sigma^i}(W|\emptyset, \omega)}$$

denote the conditional distribution of signals of agents who waited in period one.

Fix a state ω and a strategy profile σ. Then, the history after period one is trinominally distributed with parameters N and $p_\sigma(Y|\emptyset, \omega)$ for $Y \in \{A, W, C\}$. More precisely, the probability that history $h = (n_A, n_C)$ occurs is

$$P(h|\omega, \sigma) = \frac{N!}{n_A! n_C! (N - n_A - n_C)!} p_\sigma(A|\emptyset, \omega)^{n_A} p_\sigma(C|\emptyset, \omega)^{n_C} p_\sigma(W|\emptyset, \omega)^{N - n_A - n_C}.$$

The period two vote count after history h is then binominally distributed with parameters $N - n_A - n_C$ and $p_\sigma(C|h, \omega)$. The probability that after history h, the total number of C-votes is equal to $k \geq n_C$ is

$$P(k|h, \omega, \sigma) = \binom{N - n_A - n_C}{k - n_C} p_\sigma(C|h, \omega)^{k - n_C} p_\sigma(A|h, \omega)^{N - k - n_A}.$$

Taking the sum over all possible histories yields the ex-ante probability $P(k|\omega, \sigma)$ of a vote count k

$$P(k|\omega, \sigma) = \sum_{h = (n_A, n_C)} P(h|\omega, \sigma) P(k|h, \omega, \sigma).$$

From this, we obtain the probability of conviction in state ω under strategy profile σ. It is given by the sum of the probabilities of all vote counts where the defendant is convicted

$$P(C|\omega, \sigma) = \sum_{k = K+1}^{N} P(k|\omega, \sigma) \; + \; pP(K|\omega, \sigma).$$

This includes the event of exactly K votes for conviction where the outcome is a conviction with probability p.

3 Example with Two Voters

To illustrate the model, we present an example with $N = 2$ voters under the voting rule $(1, 0)$, i.e., under the unanimity voting rule, and solve it for one and two periods, respectively.

Let the prior be $q = \frac{1}{2}$ and let the signals be distributed on the unit interval $[0, 1]$ according to the conditional density functions

$$f(s|I) = 2 - 2s, \qquad f(s|G) = 2s.$$

Figure 3 displays the signal distributions and the likelihood ratio.

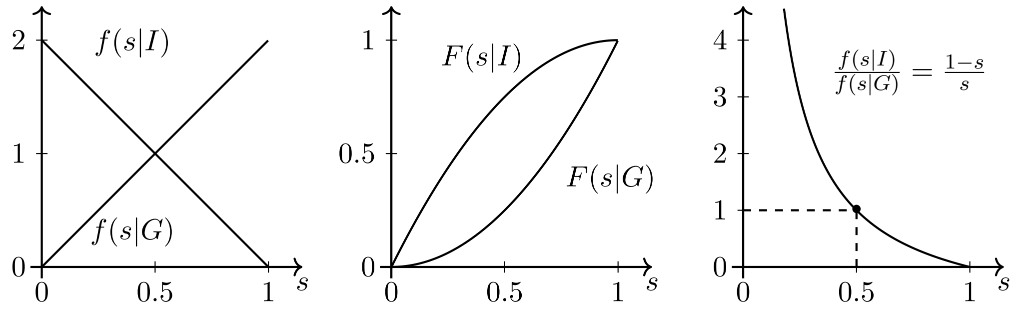

Figure 3: Density functions, c.d.f.'s and likelihood ratio

The densities are symmetric in the sense that $f(s|I) = f(1-s|G)$ holds. However, due to the unanimity voting rule, the setup is asymmetric in I and G. A single vote for A suffices for acquittal, while two votes are necessary for conviction. Strategic voters take the voting rule into account and adjust their voting behavior accordingly.

Example 1a: One Period First, consider a single voting period with simultaneous voting. We use the results from Duggan and Martinelli (2001) who show that in their one-period model there is a unique responsive[11] equilibrium. The equilibrium follows a cutoff rule, i.e., there is a unique cutoff $\hat{s}$ such that the strategies are almost everywhere equal to

$$\sigma(s) = \begin{cases} A, & \text{for } \; s \in [0, \hat{s}] \\ C, & \text{for } \; s \in (\hat{s}, 1]. \end{cases}$$

To calculate the cutoff $\hat{s}$, one has to condition on the event that a voter is pivotal, i.e., the event that a voter's decision could change the outcome. In this example, a voter is pivotal if and only if the other voter votes C. Conditioning on this event, a voter with signal $\hat{s}$ is indifferent between voting for A and voting for C if and only if

$$\frac{f(\hat{s}|I)}{f(\hat{s}|G)} \; \frac{1-F(\hat{s}|I)}{1-F(\hat{s}|G)} \; \frac{q}{1-q} = 1$$

holds. Solving this for $\hat{s}$ yields $\hat{s} = \frac{1}{3}$ and the ex-ante expected pay-off in this equilibrium is $U_1 \approx 0.796$. Voters with low signals vote for A, while voters with high signals vote for C. Although voters with a signal $s \in (\frac{1}{3}, \frac{1}{2})$ assign a higher probability to state I than to state G, they still vote for C in equilibrium as they try to counteract the bias of the voting rule.

[11]Duggan and Martinelli (2001) call an equilibrium in their simultaneous voting model a responsive equilibrium if there is no σ_i that chooses one action with probability 1, i.e., for all σ_i, $0 < \int \sigma_i(s) dF(s|G) < 1$ and $0 < \int \sigma_i(s) dF(s|I) < 1$ hold.

The probabilities of having a signal in the respective intervals conditional on the state are depicted in Figure 4.

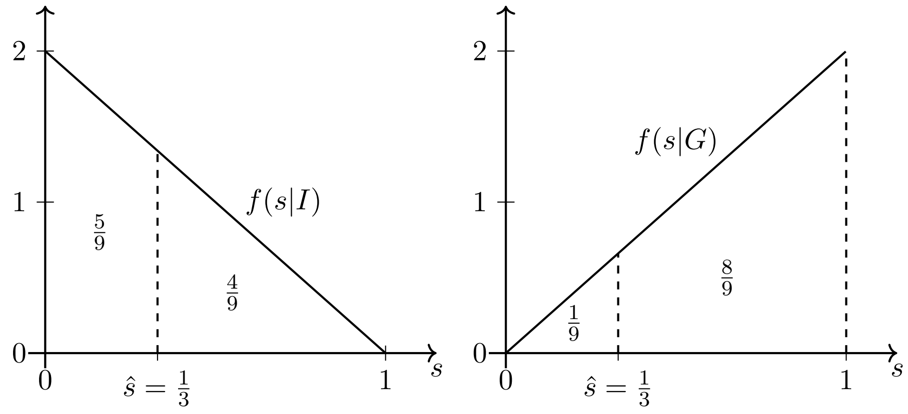

Figure 4: Conditional probabilities for Example 1a

Example 1b: Two Periods Now, consider the same example within our two-period model. For this setup, there exist various equilibria. We present a welfare-optimal equilibrium. Recall that $h = (0, 0)$ and $h = (0, 1)$ denote the possible histories in period two with 0 and 1 early C-votes, respectively.

Claim 1. A welfare-optimal equilibrium is given by the strategies

$$\sigma(s, \emptyset) = \begin{cases} A, & \text{for } s \in [0, \hat{x}] \\ W, & \text{for } s \in (\hat{x}, \hat{z}] \\ C, & \text{for } s \in (\hat{z}, 1] \end{cases}$$

$$\sigma(s, (0,0)) = \begin{cases} A, & \text{for } s \in [0, \hat{y}] \\ C, & \text{for } s \in (\hat{y}, 1] \end{cases}$$

$$\sigma(s, (0,1)) = \begin{cases} A, & \text{for } s \in [0, \hat{x}] \\ C, & \text{for } s \in (\hat{x}, 1] \end{cases}$$

with the cutoffs $\hat{x} = \frac{1}{7}$, $\hat{y} = \frac{3}{7}$ and $\hat{z} = \frac{5}{7}$.

The strategies are graphically illustrated in Figure 5. There, "late A/C" labels the signals for which a voter votes either A or C in period two depending on the other voter's action as described below.

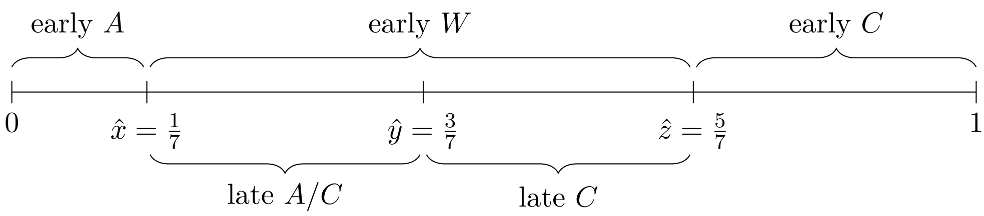

Figure 5: Strategy for Example 1b

The probabilities of having a signal in the respective intervals conditional on the state are the integrals of the corresponding densities and they are displayed in Figure 6.

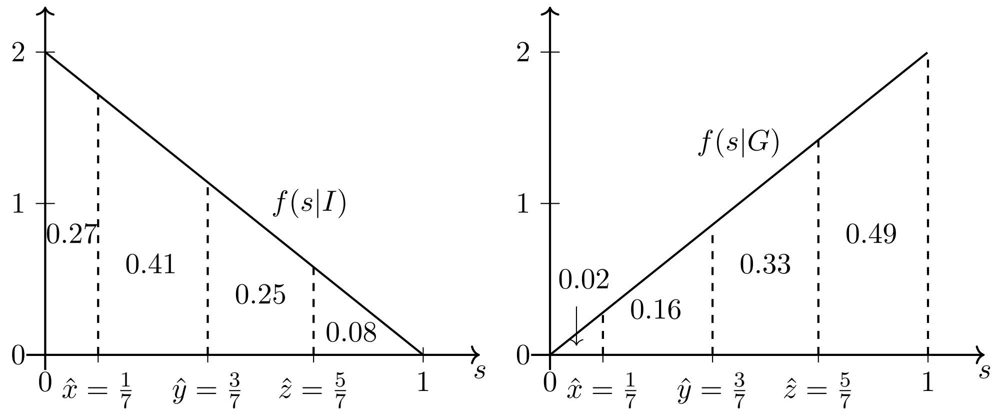

Figure 6: Conditional probabilities for Example 1b

In this equilibrium, a voter with a signal $s \leq \frac{1}{7}$ immediately votes for A and ends the game. A voter with a signal $s > \frac{5}{7}$ votes for C in period one. This can be understood as a message for the other agent about the strength of the private signal. A voter i with a signal $s \in (\frac{1}{7}, \frac{3}{7}]$ waits in period one and then votes depending on the other voter j's behavior. If j has voted for C in period one, then i also votes for C in period two. If j has instead waited in period one, then i votes for A. A voter with a signal $s \in (\frac{3}{7}, \frac{5}{7}]$ always waits and then votes C in period one. This way, she votes for C but ensures that the other voter does not misinterpret her voting as a strong indicator of guilt.

Using this voting structure allows the agents to communicate with each other. An agent with a strong signal votes early, and by doing so, she informs the other voter that her signal is highly informative. A voter with a weak signal waits for the other agent to vote and updates her beliefs depending on the outcome of period one.

The values of the cutoffs $\hat{x}, \hat{y}$ and $\hat{z}$ are determined by the likelihood ratios

$$\frac{f(\hat{x}|I)}{f(\hat{x}|G)} \frac{1-F(\hat{z}|I)}{1-F(\hat{z}|G)} = 1 \tag{1a}$$

$$\frac{f(\hat{y}|I)}{f(\hat{y}|G)} \frac{F(\hat{z}|I)-F(\hat{y}|I)}{F(\hat{z}|G)-F(\hat{y}|G)} = 1 \tag{1b}$$

$$\frac{f(\hat{z}|I)}{f(\hat{z}|G)} \frac{F(\hat{y}|I)-F(\hat{x}|I)}{F(\hat{y}|G)-F(\hat{x}|G)} = 1. \tag{1c}$$

Setting the likelihood ratios equal to 1 identifies the signal strength at which a strategic voter, who conditions on the event of being pivotal, is indifferent between two actions. Consider an agent i with a signal equal to the cutoff $\hat{x}$ who considers voting early A or voting late A/C. She is pivotal with her choice, if and only if the other agent votes for C in period one. If the other agent has a signal lower than $\hat{z}$, then i will vote for A in period two either way. Similarly, an agent with signal $\hat{y}$ is only pivotal if the other agent has a signal $s \in (\hat{y}, \hat{z}]$ and an agent with signal $\hat{z}$ is only pivotal if the other agent has a signal $s \in (\hat{x}, \hat{y}]$. The cutoffs are derived in detail in the appendix.

The defendant is acquitted if at least one voter has a signal below $\hat{x}$ or both voters have signals in $(\hat{x}, \hat{z}]$ with at least one of them being in $(\hat{x}, \hat{y}]$. Otherwise, the defendant is convicted. As a result, with two voting periods, the ex-ante expected payoff, which is the probability of a correct choice, is $U_2 \approx 0.8265$ and it is larger than the ex-ante expected payoff with only one period. In this example, introducing a second period results in a strict welfare improvement.

At least a weak welfare improvement was to be expected since the outcome of the equilibrium of the model with one period can also be implemented by an equilibrium in the model with two periods. To see this claim, note that if all agents vote in period one, then no agent has a strict incentive to wait.[12] We show in the next section that it holds generally that the introduction of the second voting period implies a strict welfare gain for all parameters.

4 Welfare-Optimal Equilibrium

In this section, we show the existence of a welfare-optimal equilibrium, we characterize the structure of welfare-optimal equilibria, and we show that there is a strict welfare improvement to a standard voting procedure with only one period. The results hold for all (K, p)-voting rules. In particular, they also apply to the unanimity voting rule and the simple majority voting rule.

First, we formalize the notion of a cutoff equilibrium in the two-period model.

[12] An agent who waits only learns whether or not she is pivotal. Since agents already condition on the event of being pivotal, waiting does not increase the expected payoff of an agent in this situation.

An equilibrium is called a cutoff equilibrium if, at every history, the agents' strategies follow (monotone) cutoff rules. More precisely, an equilibrium σ *follows a cutoff rule in period one* if there are cutoffs $\hat{s}_A, \hat{s}_C \in [\underline{s}, \bar{s}]$ with $\hat{s}_A \leq \hat{s}_C$ such that

$$\sigma(s, \emptyset) = \begin{cases} A, & \text{for } s \leq \hat{s}_A \\ W, & \text{for } \hat{s}_A < s \leq \hat{s}_C \\ C, & \text{for } s > \hat{s}_C \end{cases}$$

holds for almost all signals $s \in S$.

Fix a history $h \in H \backslash \{\emptyset\}$ from period two. Then, an equilibrium σ *follows a cutoff rule at history* h if there is some cutoff $\hat{s}_h \in [\underline{s}, \bar{s}]$ such that

$$\sigma(s, h) = \begin{cases} A, & \text{for } s \leq \hat{s}_h \\ C, & \text{for } s > \hat{s}_h \end{cases}$$

holds for almost all signals $s \in S$.

We call an equilibrium σ a *cutoff equilibrium*, if it follows a cutoff rule in period one and at every history $h \neq \emptyset$.

Our main result shows that (i) there exists an equilibrium that maximizes welfare (in the class of all symmetric monotone equilibria) and (ii) all equilibria that maximize welfare (in the class of all symmetric monotone equilibria) follow cutoff rules. This result generalizes the findings of Duggan and Martinelli (2001) for the simultaneous voting model to the model with two voting periods.

Theorem 1. *There exists a welfare-optimal equilibrium. Under (MLRP$_<$) and (ULR), every welfare-optimal equilibrium is a cutoff equilibrium.*

Note that there is a multiplicity of welfare-optimal equilibria. First, changing σ on a set of measure zero does not change the welfare and does still constitute an equilibrium. Moreover, for some parameters, there also exist welfare-optimal cutoff equilibria with different cutoffs simultaneously. Theorem 1 uses the assumptions (MLRP$_<$) and (ULR) to ensure that every action A, W and C is played in period one with positive probability. Relaxing these assumptions allows for setups where degenerate[13] equilibria can be welfare-optimal.

The first part of the theorem states the existence of a welfare-optimal equilibrium. This is proven by the maximality principle. While the strategy-space is not compact under the usual metrics, we construct a specific metric on the set $\mathcal{S}$ of the symmetric monotone strategy profiles. Under this metric, $\mathcal{S}$ is compact, and the

[13]Here, with "degenerate", we mean that not all actions are used. For example, without (ULR), there exist parameters for which it is never optimal to vote for A. Similarly, without (MLRP$_<$), there exist parameters for which the welfare-optimal equilibria do not use both periods (e.g. settings with binary signals and a small number of voters). In such settings, there can exist welfare-optimal equilibria that yield the same outcome as a degenerate cutoff equilibrium but do not follow cutoff rules themselves.

function $\Psi : \mathcal{S} \to [0, 1]$ that maps a strategy profile to its induced welfare is continuous. By the maximality principle, there exists a strategy profile which maximizes welfare. McLennan (1998) shows that such a welfare-optimal strategy profile always constitutes an equilibrium.

After having established existence, the remainder of this section is dedicated to proving the second part of Theorem 1: First, in Lemma 1 we show that in every welfare-optimal equilibrium, both voting periods are used. Then, in Lemma 2 and Lemma 3 we analyze the equilibrium strategies in period one and two, respectively.

Lemma 1. *Assume that (MLRP$_<$) holds and let σ be a welfare-optimal equilibrium. Then, agents wait with a probability strictly between* 0 *and* 1*, i.e.,*

$$0 < p_\sigma(W|\emptyset, \omega) < 1$$

holds for $\omega \in \{I, G\}$.

The first period can be used to differentiate between agents with more informative and less informative signals. This way, any agent who waits can update her prior accordingly and make a better-informed decision, increasing the expected payoff. Equilibria that do not use both periods forfeit this opportunity of communication and, as a result, cannot be welfare-optimal.

For the proof, we start with an equilibrium σ in which only one period is used and construct an equilibrium with higher welfare over several steps. First, we construct a strategy profile σ' with the same welfare as σ: The agents' time of voting can be split between both periods without changing the outcome. To do so, we let voters with more informative signals vote in period one and voters with less informative signals vote in period two. The resulting strategy profile σ' is not necessarily an equilibrium but yields the same welfare as σ by construction. Now, starting from σ', we construct a strategy profile σ'' with strictly higher welfare as follows. For some signals, there is a profitable deviation from σ' for an individual voter in period two: As the voting of period one reveals information about the signal strengths of the other voters, a voter in period two can update her prior accordingly and deviate to a more profitable strategy. By McLennan (1998), this individual profitable deviation shows the existence of a symmetric profitable deviation σ'' where every voter plays the individual profitable deviation with a small probability ε. Hence we have shown that an equilibrium in which only one voting period is used is not a welfare-optimal (symmetric) strategy profile. For the second step, we use again an argument by McLennan (1998) who shows that a welfare-optimal equilibrium is also a welfare-optimal symmetric strategy profile. Since σ is not the latter, it can also not be a welfare-optimal equilibrium.

As an immediate implication, there exists an equilibrium of the two-period voting game that yields a strictly higher welfare than all equilibria of the voting game with only one period.

Corollary 1. *Under assumption (MLRP$_<$), there exists an equilibrium of the two-period voting game that strictly welfare-dominates all equilibria of the simultaneous voting game.*

Another direct implication of Lemma 1 is that in a welfare-optimal equilibrium, the inequalities of the monotonicity conditions are strict, i.e., the probabilities of conviction given that one voter votes A, W, or C, respectively, cannot be equal.

Corollary 2. *Under assumption (MLRP$_<$), in any welfare-optimal equilibrium*

$$P(A,\omega) \ < \ P(W,\omega) \ < \ P(C,\omega)$$

holds for $\omega \in \{I, G\}$.

If, in a welfare-optimal equilibrium, waiting led to the same probability of conviction as any other action, then the equilibrium would be outcome-equivalent to an equilibrium without waiting. By Lemma 1, this cannot be true for a welfare-optimal equilibrium.

Strategies in Period One Now, we analyze the equilibrium strategies in period one. We show why agents follow cutoff strategies, and we establish equations that characterize these cutoffs.

Fix a strategy profile σ and assume that (MLRP$_<$) holds. Recall that $P(Y, \omega)$ denotes the probability that the defendant is convicted given that the state is ω and given that all voters follow strategy σ except for one voter who instead votes for $Y \in \{A, W, C\}$ in the first stage.

First, we focus on comparing voting C with waiting in period one. The probability that one individual voter changes the outcome with voting C instead of waiting is

$$P(C,\omega) - P(W,\omega).$$

For a given signal s, the conditional probability of being in state G is

$$(1-q)\ \frac{f(s|G)}{qf(s|I) + (1-q)f(s|G)}$$

and therefore, the probability of changing the outcome for the better with voting C instead of waiting is

$$(1-q)\ \frac{f(s|G)}{qf(s|I) + (1-q)f(s|G)}\ \left(P(C,G) - P(W,G)\right). \tag{2}$$

Analogously, the probability of changing the outcome for the worse is

$$q\ \frac{f(s|I)}{qf(s|I) + (1-q)f(s|G)}\ \left(P(C,I) - P(W,I)\right). \tag{3}$$

The net effect of voting C instead of waiting is strictly positive if and only if the ratio of term (3) divided by term (2),

$$\frac{q}{1-q}\,\frac{f(s|I)}{f(s|G)}\,\frac{P(C,I)-P(W,I)}{P(C,G)-P(W,G)}, \tag{4}$$

is strictly smaller than 1. If the ratio is strictly larger than 1, then an agent stricly prefers waiting to voting C in period one. For a fixed strategy in period two and fixed strategies of the other voters, because of $\frac{f(s|I)}{f(s|G)}$ being strictly monotone, setting term (4) equal to 1 yields a unique cutoff $\hat{s}_C$ for which an agent is indifferent between voting C in period one and waiting.

For the second cutoff, $\hat{s}_A$, we analogously get that an agent is indifferent between voting A in period one and waiting if and only if

$$\frac{q}{1-q}\,\frac{f(s|I)}{f(s|G)}\,\frac{P(W,I)-P(A,I)}{P(W,G)-P(A,G)} = 1 \tag{5}$$

holds. Note that for general strategy profiles $\hat{s}_A \leq \hat{s}_C$ does not need to hold. However, for welfare-optimal equilbria, agents wait with strictly positive probability and we get that the cutoffs for the first period are ordered as in Figure 7.

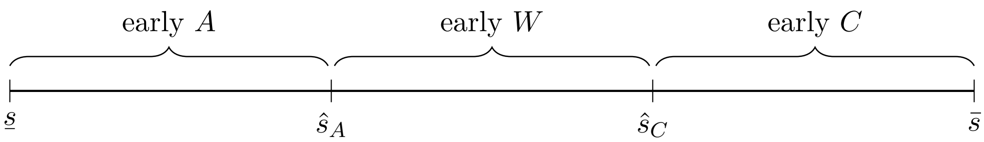

Figure 7: Cutoff strategies in period one

This observation is formally derived by the following lemma.

Lemma 2. *Under (MLRP$_<$) and (ULR) for every monotone equilibrium σ with a probability of waiting strictly between* 0 *and* 1*, the equilibrium follows a cutoff rule in period one.*

Strategies in Period Two We continue by analyzing the equilibrium strategies in period two. To understand the equilibrium behavior, first, note that every history h in period two induces a single-period game for all agents who waited. In the induced game, the votes of all agents who voted in period one are observed by the remaining agents and thus result in an updated prior.

Fix an equilibrium σ in which agents wait with probability strictly between 0 and 1. Then, the *updated prior at history* $h = (n_A, n_C)$ is given by

$$\rho_{\sigma,h} = \frac{q}{1-q}\left(\frac{p_\sigma(A|\emptyset,I)}{p_\sigma(A|\emptyset,G)}\right)^{n_A}\left(\frac{p_\sigma(C|\emptyset,I)}{p_\sigma(C|\emptyset,G)}\right)^{n_C}\left(\frac{p_\sigma(W|\emptyset,I)}{p_\sigma(W|\emptyset,G)}\right)^{N-n_A-n_C-1}. \tag{6}$$

It consists of the ex-ante prior $\frac{q}{1-q}$, the likelihood ratio of n_A agents voting for A in period one,

$$\left(\frac{p_\sigma(A|\emptyset, I)}{p_\sigma(A|\emptyset, G)}\right)^{n_A}, \tag{6a}$$

the likelihood ratio of n_C agents voting for C in period one,

$$\left(\frac{p_\sigma(C|\emptyset, I)}{p_\sigma(C|\emptyset, G)}\right)^{n_C}, \tag{6b}$$

and the likelihood ratio of the remaining $N - n_A - n_C - 1$ voters waiting in period one,

$$\left(\frac{p_\sigma(W|\emptyset, I)}{p_\sigma(W|\emptyset, G)}\right)^{N-n_A-n_C-1} \tag{6c}$$

(excluding one voter, since every voter knows her own signal).

At history h, the one-period game is played with the induced conditional distribution function

$$G_\sigma(s|\omega) = \frac{\int_{\underline{s}}^{s} \sigma_W(s', \emptyset) dF(s'|\omega)}{p_\sigma(W|\emptyset, \omega)}$$

replacing $F(s|\omega)$ and with the updated prior $\rho_{\sigma,h}$ replacing the ex-ante prior $\frac{q}{1-q}$. We call this the *at h induced game* G_h and use σ_{G_h} for the strategy profile played at G_h that is induced by σ.

Under voting rules without tie-breaking[14] and with the assumption of a sufficiently weak prior, Duggan and Martinelli (2001) show that in the simultaneous voting model, there is an almost everywhere unique responsive equilibrium[15] and it follows a cutoff rule.

However, in the two-period voting game, the updated prior $\rho_{\sigma,h}$ is an endogenous object. For fixed parameters, there may be some histories for which the prior is sufficiently weak and other histories for which it is not. Therefore, there can exist induced games G_h with a responsive equilibrium and other induced games without one. Our next result characterizes the period two equilibrium strategies of a welfare-optimal equilibrium. To calculate the cutoffs, we first need the likelihood ratio of being pivotal. If the (K, p)-voting rule allows for tie-breaks, i.e., $p \in (0, 1)$, then there are two events where a single voter is pivotal. In the first event, the voter changes the outcome from A to a tie-break, and in the second event, the voter changes the outcome from a tie-break to C. Without random tie-breaks, i.e., with $p \in \{0, 1\}$, exactly one of these events can occur. For a general (K, p)-voting rule,

[14] In our model, these are voting rules of the form $(K, 0)$ or $(K, 1)$.

[15] An equilibrium of the one-period voting game is called *responsive* if both actions, A and C, are played with positive probability.

the likelihood ratio $l_h(\hat{s}_h)$ of a single voter to be pivotal at history h is

$$\frac{pG_\sigma(\hat{s}_h|I)^{(N-K)-n_A}\left(1-G_\sigma(\hat{s}_h|I)\right)^{K-n_C-1}+(1-p)G_\sigma(\hat{s}_h|I)^{(N-K)-n_A-1}\left(1-G_\sigma(\hat{s}_h|I)\right)^{K-n_C}}{pG_\sigma(\hat{s}_h|G)^{(N-K)-n_A}\left(1-G_\sigma(\hat{s}_h|G)\right)^{K-n_C-1}+(1-p)G_\sigma(\hat{s}_h|G)^{(N-K)-n_A-1}\left(1-G_\sigma(\hat{s}_h|G)\right)^{K-n_C}}, \tag{7}$$

if all other voter follow a cutoff rule with cutoff $\hat{s}_h$.

Now, we are ready to characterize the period-two strategies in a welfare-optimal equilibrium.

Lemma 3. *Assume that* $(MLRP_<)$ *and* (ULR) *hold and fix a welfare-optimal equilibrium* σ *of the two-period voting game. Then, at every history* $h \in H\backslash\{\emptyset\}$, σ *follows a cutoff rule with a cutoff* $\hat{s}_h$ *which is equivalent*[16] *to the unique solution for* s' *of the equation*

$$\rho_{\sigma,h} \cdot \frac{f(s'|I)}{f(s'|G)} \cdot l_h(s') = 1. \tag{8}$$

Equation (8) consists of the updated prior $\rho_{\sigma,h}$, the likelihood ratio of the own signal $\frac{f(s'|I)}{f(s'|G)}$, and the likelihood ratio of the event of being pivotal in period two conditional on the observations in period one. A voter in period two is indifferent between voting for A or C if and only if the product of these three is equal to 1. If the product is strictly larger than 1, then conditioning on the event of being pivotal, the voter reasons that the state is more likely to be I and, therefore, strictly prefers voting for A in period two. Analogously, if the product is strictly smaller than 1, then the voter strictly prefers voting for C.

We have now seen that the voters follow cutoff strategies in both periods and how these cutoffs are calculated. This concludes the analysis of the structure of the welfare-optimal equilibria.

5 Information Aggregation

We now show that the two-period voting procedure aggregates information when the number of voters grows large. Consider a sequence $(a_N)_{N\in\mathbb{N}}$ of voting setups where every a_N has exactly N voters. We say that $(a_N)_{N\in\mathbb{N}}$ *allows information aggregation* if there exists a sequence of equilibria σ_N for a_N, respectively, such that the probability of the correct decision under σ_N converges to 1 as N tends to infinity.

We show that our model with two periods allows for information aggregation even in settings where simultaneous voting and sequential voting with an exogenous voting sequence fail to do so. Feddersen and Pesendorfer (1998) analyze the unanimity voting rule in a simultaneous voting model with binary signals. They prove that

[16] More precisely, if s' lies between the first-period cutoffs $\hat{s}_A$ and $\hat{s}_C$, then the cutoff $\hat{s}_h$ is equal to s'. If s' is smaller than $\hat{s}_A$ or larger than $\hat{s}_C$, then the induced game G_h has an unresponsive equilibrium that maximizes its welfare and every cutoff $\hat{s}_h < \hat{s}_A$ or $\hat{s}_h > \hat{s}_C$, respectively, yields the same outcome.

even for large electorates, the probability of the correct decision is bounded away from 1. There, an increase in the jury size does not lead to information aggregation. Dekel and Piccione (2000) show that in a sequential voting model with exogenous timing, the equilibria are equivalent to the equilibria of the simultaneous voting model. Therefore, an exogenous voting sequence does also not allow for information aggregation under the unanimity voting rule as long as the likelihood ratio of the signals is bounded. In our two-period voting model, this observation does not hold anymore. Theorem 2 shows that, even with only binary signals, information is aggregated for a large jury size regardless of the voting rule.

Theorem 2. *Fix a sequence $(a_N)_{N\in\mathbb{N}}$ of voting setups that share the same parameters and only differ in the number of voters N and the voting rule (K,p). Then, regardless of the voting rules along the sequence, there exists a sequence of equilibria for which the probability of a correct decision converges to 1.*

The idea of the proof is to construct a strategy profile σ as follows. For every voting rule there is at least one of the two alternatives that needs a vote share of at least $\frac{1}{2}$ to win. Without loss of generality, let it be C. Fix a cutoff z with $F(z|I) + F(z|G) = 1$ and let voters with signals above the cutoff z vote early for C and the remaining voters wait for period two. The weakly monotone likelihood ratio implies that $F(z|I) > \frac{1}{2}$ and $F(z|G) < \frac{1}{2}$ hold. By the strong law of large numbers, the realized vote share of C-votes converges to the expected vote share. The probability that the game ends in period one with a wrong decision converges to 0. The expected vote share of early C-votes is different for both states, and therefore, the late voters learn the correct state with probability converging to 1. Thus, we have constructed a sequence of strategy profiles for which the probability of a correct decision converges to 1. Now, the result of Theorem 2 follows by using an argument by McLennan (1998) that says that under homogeneous preferences, the welfare-optimal symmetric strategy profile is an equilibrium. Therefore, for a sequence of welfare-optimal equilibria, the probability of the correct decision also converges to 1.

Assumptions (MLRP$_<$) and (ULR) are not needed for Theorem 2. The result also holds if the informativeness of the signals is bounded. In particular, Theorem 2 also applies to the setup of Feddersen and Pesendorfer (1998) who consider a binary signal space.

We conclude this section with Lemma 4 giving a bound on the speed of convergence.

Lemma 4. *The rate of convergence of the probability of a correct decision in the welfare-optimal equilibrium of the two-period game is at least N^{-1}.*

6 The Swing Voter's Curse

In this section, we analyze the so-called swing voter's curse, which occurs under the simple majority voting rule for an even number of voters.[17] In this situation, in the simultaneous voting game, less informed voters strictly prefer to abstain rather than to vote (see Feddersen and Pesendorfer (1996)). The reason for this swing voter's curse is that there exist two different voting situations where a single voter i's decision is pivotal. That is, if the aggregated number of A-votes of the other voters is either one vote more or one vote less than the aggregated number of C-votes of the other voters. In a simultaneous voting game, the swing voter's curse reduces welfare because less informed agents who strictly prefer not to vote have to vote and may be pivotal, changing the outcome to the wrong alternative.

In the simultaneous voting game, this effect can be mitigated, and the welfare can be improved by allowing agents to abstain. However, this way, the information of the less informed voters is lost. We show that the introduction of a second voting period (without allowing for abstention in period two) can utilize the information of such voters and leads to a greater welfare improvement than the possibility to abstain.

First, note that voters in the two-period voting game can mimic abstention of the simultaneous voting model. Voters can effectively abstain by waiting in period one and then voting for the majority outcome from period one in period two (or randomizing with probability $\frac{1}{2}$ if the outcome of the first period is a tie). Therefore, the two-period voting game can achieve the welfare of the welfare-optimal equilibrium of the simultaneous voting game with abstention. The following theorem states that there even exists a strict welfare improvement.

Theorem 3. *Assume that N is even and that (MLRP$_<$) and (ULR) hold. Then, under the simple majority voting rule, the welfare-optimal equilibrium of the model with two periods (without abstention) strictly welfare-dominates all equilibria of the simultaneous voting game with abstention.*

The first part of the proof is to construct a strategy profile of the two-period voting game that yields the same welfare as the welfare-optimal equilibrium of the simultaneous voting game with abstention. Then, we show that there is a profitable deviation in the two-period voting game. Using McLennan, 1998, this shows that there is an equilibrium in the two-period voting game, which yields a strictly higher welfare than the simultaneous voting game with abstention.

As a result, the welfare of the welfare-optimal equilibria in the different voting

[17] We follow the literature by concentrating our analysis on the swing voter's curse under the simple majority voting rule. Note that the swing voter's curse occurs in our setup also under other voting rules as long as random tie-breaks can occur.

procedures can be ranked as follows:

$$U_{\text{Simultaneous}} \preccurlyeq U_{\text{Abstain}} \preccurlyeq U_{\text{Twoperiods}} \tag{9}$$

If (MLRP$_<$) and (ULR) hold, then the inequalities are strict. Without these assumptions, there exist parameters for which the welfare of all three voting procedures is equal: A voting setting with binary signals, symmetric likelihood ratios, and a prior of $\frac{1}{2}$ is an example.

7 Endogenous Timing Compared to a Fixed Sequence

In this section, we compare our voting model with endogenous timing to a voting procedure with an exogenously fixed voting sequence. More precisely, we compare it to a setup with two voting periods where for each voter it is exogenously given (and common knowledge) in which period this voter casts her vote. For a more detailed analysis of voting with an exogenously fixed sequence, see Dekel and Piccione (2000).

One substantial difference between an exogenously fixed voting sequence and voting with endogenous timing is the asymmetry between voters that is induced by the fixed timing of voting. Naturally, voters that vote in period one and voters that vote in period two are not ex-ante equal. We show that if we allow for asymmetric strategies in our voting model with endogenous timing, a strict welfare improvement is gained over voting with a fixed sequence.

Theorem 4. *Under assumptions (MLRP$_<$) and (ULR), there exists a (potentially) asymmetric equilibrium of the two-period voting game with endogenous timing that strictly welfare-dominates all equilibria of the two-period voting game with an exogenously fixed voting sequence.*

The strict welfare gain is obtained by constructing a profitable deviation. We start with a welfare-optimal equilibrium of the voting game with a fixed sequence. The outcome of this equilibrium can be replicated with endogenous timing. Now, we let a single voter in period one deviate and instead vote in period two with a positive probability. This discloses additional information to the voters in period two and subsequently allows for a profitable deviation. Therefore, an endogenous timing decision yields a strict welfare improvement over a fixed voting sequence.

8 Conclusion

In this chapter, we explore a voting model with an endogenous timing decision. We show the existence and characterize the structure of welfare-optimal equilibria. We generalize the well-known result from simultaneous voting models that responsive strategies follow cutoff rules. Moreover, the welfare-optimal equilibria of our model with endogenous timing welfare-dominate the equilibria from simultaneous voting models and voting procedures with a fixed voting sequence. Information is aggregated even with bounded informativeness of the signals under the unanimity voting rule. In the case of a possible random tie-break, sequential voting mitigates the swing voter's curse more effectively than abstention. The endogenous sorting into the two voting periods allows the voters to convey the strength of their private information to each other and ultimately make a better-informed collective decision.

There are various extensions to our model that can be pursued for future research. First, the two periods can be generalized to an arbitrary finite number or a countable infinite number of periods. Adding more periods makes the information transmission of the agents more efficient, resulting in a higher probability of choosing the correct outcome. However, we have shown in this chapter that even under the unanimity voting rule with bounded signals, two voting periods suffice for information aggregation.

Another possible extension is the generalization to a continuous time interval. Going from two periods to continuous time allows for finer communication between the voters. Depending on the modeling of the strategies, continuous time can allow the agents to perfectly communicate their signals and solves the collective coordination problem completely. Note that allowing the set of possible voting times to be as rich as a real interval is a particularly strong assumption.

Possible other extensions in this line of research could be the addition of voting costs that induce a free-riding problem, making waiting costly, or considering a private value component such that the voter's interests are not perfectly aligned anymore.

A Proofs

A.1 Proofs for Section 3

Proof of Claim 1. First, we show that if an equilibrium follows such cutoff rules, then the cutoffs $\hat{x}$, $\hat{y}$, and $\hat{z}$ solve the following system of equations:

$$2\hat{x} + \hat{z} = 1 \tag{10}$$

$$2\hat{z} - \hat{z}^2 - 2\hat{y} - 2\hat{y}\hat{z} + 3\hat{y}^2 = 0 \tag{11}$$

$$2\hat{y} - \hat{y}^2 - 2\hat{x} + \hat{x}^2 - 2\hat{z}\hat{y} + 2\hat{z}\hat{x} = 0. \tag{12}$$

Solving the system numerically then yields the unique solution $\hat{x} = \frac{1}{7}$, $\hat{y} = \frac{3}{7}$ and $\hat{z} = \frac{5}{7}$.

Equation (10) is given by setting a voter with signal $\hat{x}$ to be indifferent between voting A and waiting for period two:

$$\begin{aligned}
& \frac{f(\hat{x}|I)}{f(\hat{x}|G)} \frac{P(\text{Piv}_{AW}\,|I)}{P(\text{Piv}_{AW}\,|G)} = 1 \\
\iff & \frac{f(\hat{x}|I)}{f(\hat{x}|G)} \frac{1-F(\hat{z}|I)}{1-F(\hat{z}|G)} = 1 \\
\iff & \frac{2-2\hat{x}}{2\hat{x}} \frac{1-2\hat{z}+\hat{z}^2}{1-\hat{z}^2} = 1 \\
\iff & (1-\hat{x})(1-\hat{z}) = \hat{x}(1+\hat{z}) \\
\iff & 1-\hat{x}-\hat{z}+\hat{x}\hat{z} = \hat{x}+\hat{x}\hat{z} \\
\iff & 1 = 2\hat{x}+\hat{z},
\end{aligned}$$

where $P(\text{Piv}_{AW}\,|\omega)$ is the probability of being pivotal with the decision of voting A or waiting. It is equal to the probability that the other voter votes C. If the other voter votes A or waits, then A will be the outcome even if i waits.

Equation (11) is given by setting a voter with signal $\hat{y}$ who waited to be indifferent for the case that the other voter also waited. Let $P(\text{Piv}_{AC})$ be the probability that a voter is pivotal with deciding between voting late A or C. Then, one gets

$$\begin{aligned}
& \frac{f(\hat{y}|I)}{f(\hat{y}|G)} \frac{P(\text{Piv}_{AC}\,|I)}{P(\text{Piv}_{AC}\,|G)} = 1 \\
\iff & \frac{f(\hat{y}|I)}{f(\hat{y}|G)} \frac{F(\hat{z}|I)-F(\hat{y}|I)}{F(\hat{z}|G)-F(\hat{y}|G)} = 1 \\
\iff & \frac{2-2\hat{y}}{2\hat{y}} \frac{2\hat{z}-\hat{z}^2-2\hat{y}+\hat{y}^2}{\hat{z}^2-\hat{y}^2} = 1
\end{aligned}$$

$$\begin{aligned}
\iff & (1-\hat{y})(2\hat{z}-\hat{z}^2-2\hat{y}+\hat{y}^2) = \hat{y}(\hat{z}^2-\hat{y}^2) \\
\iff & 2\hat{z}-\hat{z}^2-2\hat{y}+\hat{y}^2-2\hat{y}\hat{z}+\hat{y}\hat{z}^2+2\hat{y}^2-\hat{y}^3 = \hat{y}\hat{z}^2-\hat{y}^3 \\
\iff & 2\hat{z}-\hat{z}^2-2\hat{y}-2\hat{y}\hat{z}+3\hat{y}^2 = 0.
\end{aligned}$$

Equation (12) is given by setting a voter with signal $\hat{z}$ to be indifferent between waiting and voting early C:

$$\begin{aligned}
& \frac{f(\hat{z}|I)}{f(\hat{z}|G)} \frac{P(\mathrm{Piv}_{WC}|I)}{P(\mathrm{Piv}_{WC}|G)} = 1 \\
\iff \quad & \frac{f(\hat{z}|I)}{f(\hat{z}|G)} \frac{F(\hat{y}|I) - F(\hat{x}|I)}{F(\hat{y}|G) - F(\hat{x}|G)} = 1 \\
\iff \quad & \frac{2-2\hat{z}}{2\hat{z}} \frac{2\hat{y} - \hat{y}^2 - 2\hat{x} + \hat{x}^2}{\hat{y}^2 - \hat{x}^2} = 1 \\
\iff \quad & (1-\hat{z})(2\hat{y} - \hat{y}^2 - 2\hat{x} + \hat{x}^2) = \hat{z}(\hat{y}^2 - \hat{x}^2) \\
\iff \quad & 2\hat{y} - \hat{y}^2 - 2\hat{x} + \hat{x}^2 - 2\hat{z}\hat{y} + 2\hat{z}\hat{x} = 0.
\end{aligned}$$

The probability $P(\mathrm{Piv}_{WC}|\omega)$ of being pivotal between waiting and voting C in period one is given by the probability that the other voter has a signal in the interval $(\hat{x}, \hat{y}]$, which means that for voter i voting C early changes the outcome compared to waiting.

To finish the proof of Claim 1, it is left to show that there exists a welfare-optimal equilibrium that follows cutoff strategies. This part is deferred to Section 4, Theorem 1, which shows for the two-period voting model that there is an equilibrium that maximizes welfare and follows such cutoff rules. □

A.2 Proofs for Section 4

Proof of Theorem 1. First, we show the existence of a welfare-optimal equilibrium. We construct a metric $d_{\mathcal{S}}$ on the set $\mathcal{S}$ of the symmetric monotone strategy profiles. Let $Z := H \times \{A, W, C\} \times \{I, G\}$. The distance of two strategies under $d_{\mathcal{S}}$ is given by the sum of the differences of the induced ex-ante probabilities $p_\sigma(Y|h,\omega)$ of playing certain actions:

$$d_{\mathcal{S}}(\sigma_1, \sigma_2) = \sum_{(h,Y,\omega)\in Z} \left| p_{\sigma_1}(Y|h,\omega) - p_{\sigma_2}(Y|h,\omega) \right|.$$

Next, we show that the metric space $(\mathcal{S}, d_{\mathcal{S}})$ is compact. To show sequentially compactness, we start with a sequence $(\sigma_n)_{n\in\mathbb{N}}$ of symmetric monotone strategy profiles. Let φ denote the function that maps such a strategy into the space of induced ex-ante probabilities, i.e.,

$$\begin{aligned}
\varphi : \mathcal{S} &\to [0,1]^{6|H|} \\
\sigma &\mapsto \left(p_\sigma(Y|h,\omega)\right)_{(h,Y,\omega)\in Z}
\end{aligned}$$

Then, as $[0,1]^{6|H|}$ together with the taxicab distance d_1 is a compact space, the sequence of the induced probabilities $(\varphi(\sigma_n))_{n\in\mathbb{N}}$ has a convergent subsequence

$(\varphi(\sigma_{n_k}))_{k\in\mathbb{N}}$. Its limit is induced by a strategy profile σ^*.[18] By construction, the subsequence $(\varphi(\sigma_{n_k}))_{k\in\mathbb{N}}$ converges to σ^*. Therefore, in $\mathcal{S}$, every sequence has a converging subsequence and $(\mathcal{S}, d_{\mathcal{S}})$ is a compact metric space.

Now, the function φ is continuous with respect to the metrics $d_{\mathcal{S}}$ and d_1. Furthermore, the function that maps the probabilities of actions to the expected welfare

$$\begin{aligned} \psi : [0,1]^{6|H|} &\to [0,1] \\ \left(p_\sigma(Y|x,\omega)\right)_{(h,Y,\omega)\in Z} &\mapsto U \end{aligned}$$

is continuous with respect to the metrics d_1 on $[0,1]^{6|H|}$ and d_1 on $[0,1]$.

By the continuity of the composition $\psi \circ \varphi$ and the compactness of the strategy space, there exists a welfare-optimal symmetric strategy profile. McLennan (1998) shows that such a welfare-optimal strategy profile constitutes an equilibrium.

The second statement of the theorem says that all welfare-optimal equilibria follow cutoff rules. We prove this through the other results established in Section 4. The overview is as follows: First, we fix a welfare-optimal equilibrium. Lemma 1 shows that in a welfare-optimal equilibrium, both periods are used. Lemma 2 shows that agents in a welfare-optimal equilibrium follow cutoff strategies in period one for almost all signals. Lemma 3 shows that agents follow a strategy in period two that is equivalent to a cutoff strategy. Together, these results show that in a welfare-optimal equilibrium, agents follow a cutoff strategy for all signals except for a subset with probability measure zero. Therefore, every welfare-optimal equilibrium is almost everywhere equal to a cutoff equilibrium. □

Proof of Lemma 1. Every equilibrium σ with $p_\sigma(W|\emptyset, I) = 1$ yields the same payoff as a corresponding equilibrium with $p_\sigma(W|\emptyset, I) = 0$, i.e., it is of no importance whether all agents wait or no agent waits. Thus, it suffices to fix an equilibrium σ with $p_\sigma(W|\emptyset, I) = 0$, which is optimal in the class of such equilibria and to show that there exists an equilibrium with higher welfare. We show that there exists an equilibrium with higher welfare by dividing the agents who vote for one action such that some agents with specific signals vote in period one and the agents with other signals vote in period two. Then, a single agent can profitably deviate due to her updated information. Using a result by McLennan (1998), the existence of a welfare-better strategy profile implies the existence of an equilibrium with higher welfare.

First, we construct an equilibrium with $p_\sigma(W|\emptyset, I) = 0$ that is optimal in the class of all such equilibria. As an equilibrium with $p_\sigma(W|\emptyset, I) = 0$ is equivalent to an

[18] To construct such a strategy profile σ^* for a given limit, define the strategy separately for every history h. For a given h, start with cutoff strategies that induce the correct probabilities for state $\omega = I$. Then, adjust the strategy by shifting the probability mass between the actions to obtain the probabilities for state $\omega = G$ without changing the probabilities for $\omega = I$. As the probabilities are the limit probabilities induced by monotone strategy profiles, the limit strategy profile σ^* is also monotone.

equilibrium of the simultaneous voting model, we can apply the results from Duggan and Martinelli (2001) to our (more general) (K,p)-voting rule. A welfare-optimal equilibrium is given by the strategy profile

$$\sigma(s,h) = \begin{cases} A, & \text{for } \; s \in [\underline{s}, \hat{s}] \\ C, & \text{for } \; s \in (\hat{s}, \bar{s}] \end{cases}$$

with $\hat{s}$ being the solution of the equation

$$\frac{q}{1-q}\frac{f(\hat{s}|I)}{f(\hat{s}|G)}\; l_h(\hat{s}) = 1$$

where $l_h(\hat{s})$ denotes the term

$$\frac{pF(\hat{s}|I)^{N-K}\big(1-F(\hat{s}|I)\big)^{K-1} + (1-p)F(\hat{s}|I)^{N-K-1}\big(1-F(\hat{s}|I)\big)^{K}}{pF(\hat{s}_h|G)^{N-K}\big(1-F(\hat{s}|G)\big)^{K-1} + (1-p)F(\hat{s}|G)^{N-K-1}\big(1-F(\hat{s}|G)\big)^{K}},$$

which is the likelihood ratio of being pivotal if all other voters follow the cutoff rule with cutoff $\hat{s}$. Now, we modify the strategies without changing the outcome by letting a small part of C-voters vote in period two instead of period one. Fix a positive $\varepsilon < 1 - \hat{s}$ and define a new strategy profile σ' by

$$\sigma'(s,\emptyset) = \begin{cases} A, & \text{for } \; s \in [\underline{s}, \hat{s}] \\ W, & \text{for } \; s \in (\hat{s}, \hat{s}+\varepsilon] \\ C, & \text{for } \; s \in (\hat{s}+\varepsilon, \bar{s}] \end{cases}$$

$$\sigma'(s,h) = \begin{cases} A, & \text{for } \; s \in [\underline{s}, \hat{s}] \\ C, & \text{for } \; s \in (\hat{s}, \bar{s}] \end{cases}, \qquad \text{for all } \; h \neq \emptyset.$$

Now, we fix an agent i, the threshold $\hat{s}$ and the strategies of all other agents. For the case $p \neq 1$, we construct a payoff increasing strategy profile σ''_i for agent i by letting her wait in period one and updating her prior at one particular history $h = (N-K-1, K)$. At any other history, i follows the strategy σ. This is given by

$$\sigma''(s,\emptyset) = \quad W$$

$$\sigma''(s,(N-K-1,K)) = \begin{cases} A, & \text{for } \; s \in [\underline{s}, \hat{s}'] \\ C, & \text{for } \; s \in (\hat{s}', \bar{s}]. \end{cases}$$

$$\sigma''(s,h) = \begin{cases} A, & \text{for } \; s \in [\underline{s}, \hat{s}] \\ C, & \text{for } \; s \in (\hat{s}, \bar{s}]. \end{cases} \qquad \text{for all } h \neq \emptyset, \; h \neq (N-K-1,K)$$

with $\hat{s}'$ being the unique solution of the equation

$$\frac{q}{1-q}\,\frac{f(\hat{s}'|I)}{f(\hat{s}'|G)}\,\left(\frac{F(\hat{s}|I)}{F(\hat{s}|G)}\right)^{N-K-1}\,\left(\frac{1-F(\hat{s}+\varepsilon|I)}{1-F(\hat{s}+\varepsilon|G)}\right)^{K}=1. \tag{13}$$

Duggan and Martinelli (2001) show that the inequality

$$\frac{1-F(\hat{s}|I)}{1-F(\hat{s}|G)}>\frac{1-F(\hat{s}+\varepsilon|I)}{1-F(\hat{s}+\varepsilon|G)}$$

follows from ($\mathrm{MLRP}_<$). As an immediate consequence, the likelihood ratio of being pivotal is different for $\hat{s}$ and $\hat{s}'$. This implies that

$$\hat{s}'<\hat{s}$$

holds, i.e., the cutoffs of σ' and σ'' are different at history $h=(N-K-1,K)$. Since $\hat{s}'$ solves equation (13), it is the optimal strategy for agent i given that she observes that exactly K voters vote C in period one. Hence, σ'' is a profitable deviation for player i. For the case $p=1$, the analogue construction for history $h=(N-K,K-1)$ instead of $h=(N-K-1,K)$ yields the same result.

By a result of McLennan (1998), this implies that there exists a symmetric equilibrium with higher welfare. □

Proof of Corollary 1. Every equilibrium of the simultaneous voting game is outcome-equivalent to an equilibrium of the two-period model with $p_\sigma(W|\emptyset,I)=0$. By Lemma 1, there exists an equilibrium with strictly higher welfare. □

Proof of Corollary 2. Suppose for contradiction that there exists a welfare-optimal equilibrium σ^* with one of the inequalities being an equality. Without loss of generality, let

$$P(A,\omega)=P(W,\omega)$$

be true. Then, the strategy profile where in the first period all probability mass from waiting is put onto A instead, yields the same expected welfare. By Lemma 1, there exists an equilibrium with strictly higher welfare, which contradicts welfare-optimality of σ^*. □

Proof of Lemma 2. We show that the best response to any symmetric strategy profile σ follows cutoff rules in period one. Recall that $P(Y,\omega)$ is the probability that the defendant is convicted given that the state is ω and given that one voter votes for $Y\in\{A,W,C\}$ in the first period and all other voters follow strategy σ. The expected payoff of voting A early after receiving signal s is now given by

$$U(A,s) = \frac{f(s|I)}{f(s|I)+f(s|G)}(1-P(A,I)) + \frac{f(s|G)}{f(s|I)+f(s|G)}P(A,G). \tag{14}$$

Similarly, let

$$U(W,s) = \frac{f(s|I)}{f(s|I)+f(s|G)}(1-P(W,I)) + \frac{f(s|G)}{f(s|I)+f(s|G)}P(W,G) \tag{15}$$

and

$$U(C,s) = \frac{f(s|I)}{f(s|I)+f(s|G)}(1-P(C,I)) + \frac{f(s|G)}{f(s|I)+f(s|G)}P(C,G) \tag{16}$$

denote the respective expected payoffs. To see when a voter is indifferent between two options, let x, $a_l < a_h$ and $b_l < b_h$ be real numbers and consider the equation

$$xa_h + (1-x)b_l = xa_l + (1-x)b_h, \tag{17}$$

which is uniquely solved by

$$x = \frac{b_h - b_l}{(a_h - a_l) + (b_h - b_l)} \in [0,1].$$

Set any two of the three utility functions (14) , (15) and (16) equal to each other. Then, the resulting equation has the form of equation (17) with $x = \frac{f(s|I)}{f(s|I)+f(s|G)}$. Thus, for every pair of utility functions this gives a unique solution for

$$\frac{f(s|I)}{f(s|I)+f(s|G)} \in [0,1].$$

Furthermore, we know that it lies in the interior $(0,1)$ by Corollary 2.

Let x_{AW} denote the value obtained by setting $U(A,s)$ and $U(W,s)$ to be equal. Then, the utility of voting for A is strictly higher than the utility of voting for W for all signals s with $\frac{f(s|I)}{f(s|I)+f(s|G)} > x_{AW}$ and strictly lower for all signals s with $\frac{f(s|I)}{f(s|I)f(s|G)} < x_{AW}$. In particular a voter is indifferent with a signal s with $\frac{f(s|I)}{f(s|I)+f(s|G)} = x_{AW}$. The same holds for x_{WC} and x_{AC}, which are defined the same way.

By monotonicity, one can rewrite

$$\begin{aligned}
P(W,I) &= P(A,I) + \varepsilon_1 \\
P(C,I) &= P(A,I) + \varepsilon_1 + \varepsilon_2 \\
P(W,G) &= P(A,I) + \delta_1 \\
P(C,G) &= P(A,I) + \delta_1 + \delta_2
\end{aligned}$$

for $\varepsilon_1, \varepsilon_2, \delta_1, \delta_2 > 0$. Thus, one gets

$$x_{AW} = \frac{\delta_1}{\varepsilon_1 + \delta_1}$$
$$x_{AC} = \frac{\delta_1 + \delta_2}{\varepsilon_1 + \varepsilon_2 + \delta_1 + \delta_2}$$
$$x_{WC} = \frac{\delta_2}{\varepsilon_2 + \delta_2}.$$

In particular, x_{AC} is a convex combination of x_{AW} and x_{WC}. By (MLRP$_<$), the term $\frac{f(s|I)}{f(s|I)+f(s|G)}$ is strictly decreasing in s. By (ULR), for all $x \in (0,1)$ there exists a unique

$$s_x = \sup_s \Big\{ \frac{f(s|I)}{f(s|I) + f(s|G)} \leq x \Big\}.$$

Now, exactly one of the three (in-)equalities

$$s_{x_{AW}} < s_{x_{AC}} < s_{x_{WC}}, \tag{18}$$
$$s_{x_{AW}} > s_{x_{AC}} > s_{x_{WC}}, \tag{19}$$
$$s_{x_{AW}} = s_{x_{AC}} = s_{x_{WC}} \tag{20}$$

holds. If either (19) or (20) holds, then the equilibrium cannot be welfare-optimal by Lemma 1 since there is no set of signals with a positive measure for which W is a strictly best response. Thus, inequality (18) holds, which implies that the equilibrium follows a cutoff rule in period one. □

Proof of Lemma 3. Consider the history $h = (n_A, n_C)$ in period two after n_A voters voted early A and n_C voters voted early C. Let

$$A_{\sigma,n} = \left\{ s \;\middle|\; \left(\frac{f(s|I)}{f(s|G)}\right)^N > \frac{1}{\rho_{\sigma,n}} \right\}$$

denote the set of all signals s such that the likelihood ratio raised to the power of N overcomes the updated prior. Similarly, define

$$B_{\sigma,n} = \left\{ s \;\middle|\; \frac{f(s|I)}{f(s|G)} < \frac{1}{\rho_{\sigma,n}} \right\}$$

to be the set of all signals whose likelihood ratio is smaller than the updated prior. Taking the idea from the proof of Lemma 2 in Duggan and Martinelli (2001), there exists a responsive equilibrium in the induced game G_h if and only if the inequalities

$$\int_{A_{\sigma,n}} \sigma_W(s|\emptyset)\mu(ds) > 0$$

and

$$\int_{B_{\sigma,n}} \sigma_W(s|\emptyset)\mu(ds) > 0$$

hold, i.e., if the probability that an agent has a signal which is stronger than the prior in either direction is positive.

Consider now the case that this condition is satisfied at h. Even though assumption (A4) in Duggan and Martinelli (2001) does not necessarily hold in our two-period model, the assumptions necessary for their Theorem 1 are fulfilled and its conclusion applies to the induced game G_h. Hence, there exists an almost everywhere unique responsive strategy profile that is an equilibrium of G_h with cutoff s' given as the solution of

$$\rho_{\sigma,h} \cdot \frac{f(s'|I)}{f(s'|G)} \cdot l_h(s') = 1.$$

By Lemma 1 all histories are reached with positive probability. As an unresponsive equilibrium in a one-period voting game yields a lower welfare than the unique responsive equilibrium, we get that in a welfare-optimal equilibrium, the unique responsive equilibrium is played in every induced game of periode two where one exists.

At the histories where no responsive equilibrium exists, the welfare-optimal unresponsive equilibrium is played in period two, i.e., either all voters vote A or all voters vote C. □

A.3 Proofs for Section 5

Proof of Theorem 2. As a consequence of McLennan (1998), it is sufficient to show that there exists a sequence of strategy profiles for which the probability of a correct decision converges to one.

For our construction, let z be a cutoff with the symmetric property $F(z|I) + F(z|G) = 1$. By the intermediate value theorem, such a z exists. Let $r := F(z|I) = 1 - F(z|G)$. Intuitively, treating the two intervals $[0, z]$ and $(z, 1]$ like two discrete signals that indicate innocence/guilt, respectively, r is the probability that an agent receives a correct signal. The number of correct signals is binomially distributed with parameters N and r. Note that $r > \frac{1}{2}$ holds as the likelihood ratio is weekly decreasing and not everywhere constant.

For each N, we now construct a strategy profile σ_N. Fix a setup a_N with voting rule (K, p). At least one of the two alternatives needs at least half of the votes to be implemented with positive probability. First, consider the case that this C needs at least $N/2$ votes, i.e., $K \geq N/2$ holds. Define the strategy profile σ_N by

$$\begin{aligned}
\sigma_N(s,\emptyset) &= \begin{cases} W, & \text{for } \ s \leq z \\ C, & \text{for } \ s > z \end{cases} \\
\sigma_N(s,h) &= \quad A, \qquad \text{for all } h \neq \emptyset \text{ with } n_C < N/2 \\
\sigma_N(s,h) &= \quad C, \qquad \text{for all } h \neq \emptyset \text{ with } n_C \geq N/2.
\end{aligned}$$

The outcome of σ_N is C if and only if at least $N/2$ voters receive a signal $s \in (z, 1]$.

Now, consider the second case that A needs at least $N/2$ votes to be implemented with positive probability. Analogously, we construct σ_N by

$$\begin{aligned}
\sigma_N(s,\emptyset) &= \begin{cases} A, & \text{for } \ s \leq z \\ W, & \text{for } \ s > z \end{cases} \\
\sigma_N(s,h) &= \quad A, \qquad \text{for all } h \neq \emptyset \text{ with } n_A \geq N/2 \\
\sigma_N(s,h) &= \quad C, \qquad \text{for all } h \neq \emptyset \text{ with } n_A < N/2.
\end{aligned}$$

Again, the outcome of σ_N is C if and only if at least $N/2$ voters receive a signal $s \in (z, 1]$. For both cases and for both states, the probability of a wrong decision is bounded above by the probability that a binomially distributed random variable $X_{(N,r)}$ with parameters N and r takes a value less or equal to $N/2$ (i.e., at least half of the voters receive the wrong signal).

By the weak law of large numbers, the realized vote share of C-voters in period one converges to the expected vote share $1 - F(z|\omega_N)$ in probability, which implies that the correct outcome is implemented with probability approaching 1. Thus, we have constructed a sequence $(a_N)_{N\in\mathbb{N}}$ of strategy profiles such that, regardless of the sequence of voting rules along the setups, the probability of an incorrect choice converges to 0 as N converges to infinity. □

Proof of Lemma 4. Consider our construction for the proof of Theorem 2. The probability of a wrong decision is bounded above by the probability that a binomially distributed random variable $X_{(N,r)}$ with parameters N and r takes a value less or equal $N/2$. By Chebyshev's inequality this probability is at most

$$\begin{aligned}
&P\left(X_{(N,r)} \leq \frac{N}{2}\right) \\
&\leq P\left(|X_{(N,r)} - rN| \geq N\Big(r - \frac{1}{2}\Big)\right) \\
&\leq \frac{r(1-r)N}{N^2\big(r-\frac{1}{2}\big)^2} \\
&= \frac{r(1-r)}{\big(r-\frac{1}{2}\big)^2} \cdot \frac{1}{N} \\
&= \mathcal{O}(N^{-1}).
\end{aligned}$$

Thus, we have constructed a bounding sequence that converges to zero at rate N^{-1}. Therefore, for the probability of a wrong decision under the strategy profiles (σ_N), the rate of convergence is at least N^{-1}. As the probability of a wrong decision is even smaller in a welfare-optimal equilibrium, this constitutes a bound for the rate of convergence of the probability of a correct decision for the sequence of welfare-optimal monotone equilibria. □

A.4 Proofs for Section 6

Proof of Theorem 3. Consider the welfare-optimal equilibrium of the simultaneous voting game. Feddersen and Pesendorfer (1996) show that this equilibrium follows a cutoff rule. The probability of a single voter abstaining is non-zero. Hence, for a fixed N, there is a positive probability that all voters abstain. For the sequential voting game, construct the strategy profile σ as follows. In the first period, the strategy is given by the strategy profile of the simultaneous voting game, except that agents wait instead of abstaining. In period two, all agents vote for the outcome that gained a simple majority in period one. If the result of the first period results is a tie, all agents who waited then vote for each alternative with equal probability. This strategy profile is outcome-equivalent to the welfare-optimal equilibrium of the simultaneous voting game.

Now, change the voting strategies such that at the history $h = (0, 0)$ where every agent waited, the welfare-optimal cutoff strategy of the induced game G_h is played. This event occurs with positive probability, and the welfare-optimal equilibrium of the induced game in period two yields a strictly higher welfare than a coin flip. Since this strictly increases the probability of the correct decision, there exists a strategy profile of the two-period model with strictly higher welfare than all equilibria of the simultaneous voting model. By McLennan (1998), there also exists an equilibrium with strictly higher welfare. □

A.5 Proofs for Section 7

Proof of Theorem 4. Consider a welfare-optimal equilibrium of the two-period voting game with a fixed voting sequence. The outcome of this equilibrium can be replicated by an asymmetric strategy profile σ of the two-period voting game with endogenous timing. If all voters vote in the same voting period, then the same argument as in Corollary 1 implies the existence of a profitable deviation. Therefore, we consider the situation that there is at least one voter in each period.

Fix a single voter i who votes in period one. Let $\varepsilon > 0$ be sufficiently small and define a deviation for voter i as follows: For a signal s with $\sigma(s) = A$ and $F(s|I),\ F(s,G) \leq \varepsilon$, voter i waits in period one and votes for A in period two instead. If the other voters in period two observe a history h where this event occurred, they play the welfare-optimal equilibrium of the induced simultaneous voting game. By the assumptions (MLRP$_<$) and (ULR), the induced prior at h is different compared to the induced prior where i votes early. Thus, the equilibrium of the induced simultaneous voting game yields a strict welfare gain.

Note that the existence proof for Theorem 1 for a welfare-optimal symmetric equilibrium also shows the existence of a welfare-optimal asymmetric equilibrium as the number of voters is finite, and the space of all asymmetric strategy profiles is therefore also compact. □

References

Austen-Smith, David and Jeffrey S. Banks. 1996. Information Aggregation, Rationality, and the Condorcet Jury Theorem. *American Political Science Review* 90.1: 34–45.

Battaglini, Marco. 2005. Sequential voting with abstention. *Games and Economic Behavior* 51.2: 445–463.

Callander, Steven J. 2002. Voting and electoral competition. *PhD thesis.* California Institute of Technology.

CNN Politics. 2021. *These are the members calling for impeachment or the 25th Amendment to be invoked.* (accessed September 9, 2021). URL: `https://edition.cnn.com/2021/01/07/politics/list-lawmakers-trump-removal-congress/index.html`.

Condorcet, Marquis de. 1785. Essai sur l'application de l'analyse a la probilite des decisions rendues a la pluralite des voix. *Paris: De l'imprimerie royale.*

Dekel, Eddie and Michele Piccione. 2000. Sequential voting procedures in symmetric binary elections. *Journal of Political Economy* 108.1: 34–55.

Dekel, Eddie and Michele Piccione. 2014. The strategic dis/advantage of voting early. *American Economic Journal: Microeconomics* 6.4: 162–79.

Duggan, John and César Martinelli. 2001. A Bayesian model of voting in juries. *Games and Economic Behavior* 37.2: 259–294.

Eyster, Erik and Matthew Rabin. 2005. Cursed equilibrium. *Econometrica* 73.5: 1623–1672.

Feddersen, Timothy and Wolfgang Pesendorfer. 1996. The swing voter's curse. *American Economic Review* 86.3: 408–424.

Feddersen, Timothy and Wolfgang Pesendorfer. 1998. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review* 92.1: 23–35.

Fey, Mark. 1998. Informational Cascades and Sequential Voting. *Working paper.*

Keefer, Philip and Razvan Vlaicu. 2007. Democracy, Credibility, and Clientelism. *Journal of Law, Economics, and Organization* 24.2: 371–406.

McLennan, Andrew. 1998. Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents. *American Political Science Review* 92.2: 413–418.

Piketty, Thomas. 2000. Voting as Communicating. *Review of Economic Studies* 67.1: 169–191.

Schmieter, Finn S. 2019. Voting with Endogenous Timing. *Master's Thesis.*

Zurcher, Anthony. 2021. *Trump impeachment: Why convicting him just got a lot harder.* BBC News. (accessed September 9, 2021). URL: `https://www.bbc.com/news/world-us-canada-55820610`.