# *A Conversational War of Attrition*[*]

Katalin Bogzar — Michigan
Moritz Meyer-ter-Vehn[†] — UCLA
Lones Smith[‡] — Michigan

September 25, 2009

(in progress)

**Abstract**

We introduce and explore a simple model of costly dynamic deliberation. Two partially informed individuals with common interests sequentially exchange coarse messages before arriving at a common verdict. In this setting, longer "conversations" are better, and the longest one can potentially last any number of periods. Except when forced by a cloture rule, bounded equilibria are less robust predictions as off-equilibrium beliefs must be interpreted as mistakes. We also find that as a conversation transpires, it becomes increasingly likely that the case at hand is moot. So, surprisingly, information exchange correlates negatively with the value of the exchange process.

[*]We have benefited from comments at SWET 08, Penn State, UC Davis, SED 08, Games 08, ESEM 08, CalTech, and UBC
[†]Moritz thanks the NSF for financial support of this research.
[‡]Lones thanks the NSF for financial support of this research.

# 1 Introduction

> It's not easy to raise my hand and send a boy off to die without talking about it first... We're talking about somebody's life here. We can't decide in five minutes. Supposin' we're wrong.
>
> Juror #8 (Henry Fonda), *Twelve Angry Men*

Many important decisions are made by groups. Academic departments vote on whether to recommend tenure for assistant professors. Juries must weigh in on the guilt of the accused. FDA panels opt whether to approve new drugs. The Federal Reserve Board regularly decides whether to adjust the interest rate. Executive committees at firms choose whether to proceed with new technologies or pursue new product lines.

Group decisions must balance the costs of prolonged discussion against the gains of a possibly better informed decision. Organizational cultures vary in their responses to this essential conflict. Debate may be typically vibrant and long-winded in one group, but curt and quick in another. Rules may actively encourage debate, or stifle it. Motions of cloture may be allowed, cutting off further discussions. Or open discussion might be actively fostered, with mandated open-ended "question periods".

Some organizations systematically make better decisions than others. One department may tenure poorly and sink into obscurity, while another promotes wisely and prospers. Some juries are governed by passion and haste, while in others, more nuanced debate prevails. And group decisions can be tragically stupid and foreseen — like the FDA's 1999 approval of Vioxx, or Coca Cola's 1985 ill-fated introduction of "New Coke".

This paper explores the common assertions about the "wisdom of the majority" — the claim that groups make smarter decisions than individuals. Implicit in this argument is that the individuals' information is adequately aggregated for the decision. Absent a conflict of interest this aggregation poses no problem if the group members can just "put all their pertinent information on the table". This in turn seems unproblematic for hard facts that can be described in a common terminology such as, say, the forensic doctor's estimate of the time of death in a murder trial. In contrast, much of the pertinent information to determine the guilt of a defendant "beyond a reasonable doubt" may consist of subjective assessments, say, on the credibility of a testimony, or on information that is difficult to express because of differing backgrounds or terminologies. As it is of course not feasible to share this information completely, a natural way to aggregate it partially is a dynamic arguing process.

It is natural to ask why decisions take as long as they do. One possible response is that group members fundamentally disagree on the payoffs of various actions. This is no doubt often true, and possibly the cause of many a hung jury or indecisive committee. We explore an orthogonal reason for delay in joint decisions. For it is quite clear to anyone who has sought to formulate his information that communication is easier said than done. Contrary to toy Bayesian models of learning, information is often not so easily quantified. Gut feelings and instincts are often critical. We venture that a boundedly rational world with individuals unable to clearly state the strength or precision of their information. Whenever asked about a binary choice, they can tell which they think best.

While individuals share identical preferences, an *ex ante* conflict emerges if they have conflicting information. This paper introduces a highly stylized model of this arguing process and asks how much information will be aggregated in equilibrium when the group members face explicit delay costs. We focus for simplicity on the smallest "jury" of two who can only argue for one of two possible verdicts, Acquit or Convict, in any period. The arguing process stops, and a verdict is taken as soon as the two jurors argue for the same verdict. One can alternatively interpret the game form as a dynamic voting procedure with a unanimity rule in the sense that the voting process is repeated until the vote is unanimous. The jurors can only express which verdict they favor but cannot explain why. Equivalently, and maybe more plausibly, the jurors can express more sophisticated arguments they are unable to understand each other's arguments beyond a statement which verdict the other favors.

In the resulting "conversational war of attrition", the informativeness of communication is endogenous to the delay cost. The game turns out to have multiple equilibria and so how much information will be aggregated depends on the choice of an equilibrium. We show that equilibria are of two kinds. In a *communicative* equilibrium any type of either juror interprets persistence of her colleague as increasingly strong evidence for the verdict favored by the colleague and will give in accordingly. In an *insistent* equilibrium communication breaks down at some point. All remaining types of one juror will give in at that point and her colleague will insist on her verdict forever, interpreting potential deviations from her opponent's strategy as mistakes rather than signals of informational strength. Thus, insistent equilibria aggregate less information than communicative equilibria.

Our first main result (Proposition 1) is that only communicative equilibria satisfy forward induction and that the slowest equilibrium is communicative and Pareto dominant. Thus, the jurors should argue as long as possible, and they will do so if they are willing to always interpret each others' utterings as signals rather than uninformative mistakes.

Our other major result finds a puzzling character of the long conversations. We investigate for any fixed equilibrium how the length of the conversation correlates with the joint information of the jurors. Proposition 2 shows that as the conversation transpires the probability increases that the jurors' joint information is moot and does not clearly support either one of the two verdicts. Thus, ex-post, investment into information aggregation is negatively correlated with the value of the aggregated information. In other words it exhibits diminishing total returns, rather than just diminishing marginal returns.

If the jurors anticipated a conflict of interest they would partially conceal their private information to nudge the verdict in their favor. This observation goes back to Crawford and Sobel (1982) and has been analyzed in a dynamic cheap talk setting by Aumann and Hart (2003) and a static committee decision problem by Li, Rosen and Suen (2001). But unlike here, communication in these models is not just "cheap" but actually free. So partially conflicting interest in these papers acts like our costly communication, and yields coarse and not fully revealing optimal communication.

Our insights relate to the "swing voter's curse" in Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997). They explored how like-minded, strategic voters take into account the extra information inferred from the condition that their vote is pivotal. This argument assumes that group members cannot communicate and infer each other's information solely from anticipated voting behavior. Our paper essentially introduces a role for communication before a committee decision, while preventing perfect communication which would render a consecutive vote unnecessary.

The communication role of repeated voting has previously been studied by Piketty 2000 who focuses on large electorates signalling information about third alternatives in the first of two voting rounds. He finds that multi-round elections can be more efficient than single round elections, which is broadly resonant with our results.

More closely related to our paper is Ettore, Li and Suen (2009a,b) studying how repeated voting allows jurors with partially conflicting interests to share more of their information than a single vote. Just like in our model juror $A$ infers strong opposing information from the insistence of her opponent $B$ and will eventually give in unless she is certain that $A$ is the better verdict. Because of the conflicting interests the costs of information aggregation are bounded away from zero even as the frictions vanish (unlike in our model) but for moderate levels of interest conflict the benefits of dynamic information aggregation still outweigh its costs, unlike in a war or attrition with purely conflicting interests or as in Gul and Lundholm (1995).

# 2 Model

**A. Conversational War of Attrition**

Two jurors alternate in arguing to $\mathcal{C}$onvict or $\mathcal{A}$cquit the defendant of some crime. The defendant can be either $\mathcal{G}$uilty or $\mathcal{I}$nnocent. Before the game each juror privately observes a signal about the $\mathcal{G}$uilt or $\mathcal{I}$nnocence of the defendant. The game ends once agreement is reached, i.e. once both jurors argue for the same verdict. We will be interested in sequential equilibria of this conversational war of attrition and refer to equilibrium paths as *conversations.*

The jurors have identical cardinal preferences: They want to take *the right decision*, i.e. convict the guilty and acquit the innocent. Additionally, they face linear costs of arguing.

**B. Information**

Initially, the jurors share a flat common prior on the state of the world $\theta = \mathcal{G}, \mathcal{I}$. Then, they receive conditionally independent signals $\lambda, \mu$ from distributions $\Phi^{\theta}(\lambda), \Gamma^{\theta}(\mu)$. Signals are private, but all else is common knowledge.

It will simplify matters to represent these signals in a non-standard fashion, as log-likelihood-ratios, and with a different reference state for each juror. More precisely, we assume that $d\Phi^{\mathcal{G}}(\lambda)$ and $d\Phi^{\mathcal{I}}(\lambda)$ (and $d\Gamma^{\mathcal{I}}(\mu)$ and $d\Gamma^{\mathcal{G}}(\mu)$) are mutually absolutely continuous, and define $\ell = \log\left(d\Phi^{\mathcal{G}}(\lambda)/d\Phi^{\mathcal{I}}(\lambda)\right)$ as a measure of how strongly signal $\lambda$ indicates $\mathcal{G}$uilt. $m = \log\left(d\Gamma^{\mathcal{I}}(\mu)/d\Gamma^{\mathcal{G}}(\mu)\right)$ measures how strongly signal $\mu$ indicates $\mathcal{I}$nnocence. These transformed *types* $(\ell, m)$ of the jurors are more tractable than the raw signals $(\lambda, \mu)$ because they enter additively in the posterior probabilities:

$$\pi(\ell, m) = \Pr(\mathcal{G}|\ell, m) = \frac{e^{\ell-m}}{e^{\ell-m}+1}.$$

To further streamline the exposition we assume that the distributions of $\ell$ and $m$ admit weakly log-concave and strictly positive pdfs $f$ and $g$ on $(-\infty, \infty)$. [1] We denote by $h(\ell, m)$ the joint distribution of $(\ell, m)$, and by $r(\ell, m) = h(\ell, m)/f(\ell)g(m)$ their unconditional correlation. Lemma 8 in the Appendix computes $r(\ell, m)$ explicitly.

[1] Both of these assumptions are primarily for notational convenience. Proposition 1 does not rely on them, and neither do most of the auxiliary results. Proposition 2 however would fail if pdfs are sufficiently log-convex.

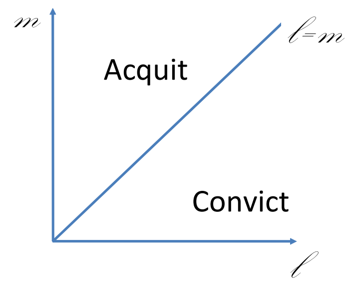


Figure 1: Correct verdict in log-likelihoodratio space

**C. Preferences**

Payoffs are the sum of waiting costs and decision costs. *Waiting costs* are explicit, equal to $\kappa \ll 1$ per argument. *Decision costs* are zero for the right decision, and mistakes are symmetrically penalized: They equal 1 of verdict $\mathcal{C}$ in state $\mathcal{I}$ and verdict $\mathcal{A}$ in state $\mathcal{G}$. Also, jurors are risk-neutral and do not discount future payoffs. Thus, for signals $(\ell, m)$ the ex-post optimal verdict entails statistical decision costs of $\min\{\pi(\ell, m), 1 - \pi(\ell, m)\}$ and the ex-post preference of $\mathcal{C}$ over $\mathcal{A}$ is given by $2\pi(\ell, m) - 1$. By virtue of the log-likelihood-ratio transformation, $\mathcal{C}$onviction is the *correct verdict* if and only if $\pi(\ell, m) = \frac{e^{\ell - m}}{e^{\ell - m}+1} \geq \frac{1}{2}$, which is the case whenever $\ell \geq m$. So with perfect information sharing, the stronger juror type should get her way.

**D. The Stopping Subgame**

After the initial argument, $\mathcal{C}$onvict or $\mathcal{A}$cquit, the game becomes a stopping game. If the initial juror argues for $\mathcal{A}$cquit, the second juror takes on the role of an accuser. He can either keep the conversation going by arguing to $\mathcal{C}$onvict, or terminate it by arguing for $\mathcal{A}$cquittal. Similarly, the first juror becomes the defender.

In the body of this paper we focus on the subgame where the initial juror voted for $\mathcal{A}$cquittal and his type indicates $\mathcal{I}$nnocence, i.e. $m > m_0$ for some cutoff $m_0$. This unusual assumption has two advantages. First, we can identify jurors with the verdict they are arguing for and will refer to them as the $\mathcal{C}$-juror and $\mathcal{A}$-juror, or just $\mathcal{C}$ (he) and $\mathcal{A}$ (she). Second, we only need to solve a stopping game, where the strategy of each juror is given by a distribution over stopping times.

In Appendix 7.1 we discuss how all of our insights carry over to the game including the initial vote. In particular, we show that a best-response of the initial juror necessarily uses a cutoff rule. It may be worth noting (probably not here, though) that there are equilibria where the initial vote is anti-monotone: Types $m > m_0$ with signals indicating $\mathcal{I}$nnocence argue to $\mathcal{C}$onvict and vice versa. In such an equilibrium, the initial juror anticipates that her colleague is a "devil's advocate" who will only agree with the initial argument if her own information is overwhelmingly supportive. More precisely, after an initial argument for $\mathcal{C}$ the second juror chooses a cutoff $\ell_1 \gg 0$ and agrees to $\mathcal{C}$onvict only if $\ell > \ell_1$ but otherwise insists on $\mathcal{A}$cquittal forever. Similarly after an initial argument for $\mathcal{A}$ she chooses a cutoff $\ell_1 \ll 0$ and agrees to $\mathcal{A}$cquit only if $\ell > \ell_1$. Anticipating this contrarian response, the initial juror will indeed optimally argue against the verdict indicated by her signal. Reversely, the contrarian behaviour is an optimal response to the reversed strategy of the initial juror.

Our analysis extends easily to these "devil's advocate" equilibria with anti-monotonic initial vote. Still, we exclude them for notational convenience by assuming that the type space of the $\mathcal{A}$-juror, conditional on reaching the subgame analyzed here, is bounded below and unbounded above, i.e. $m \in [m_0, \infty)$.

# 3 Equilibrium Analysis

In this section we will show that best responses are in monotone cutoff strategies and derive an explicit formula for equilibrium cutoff types $\ell_n, m_n$. This formula implies that equilibria satisfy a "zig-zag" property of equilibria, and that they are ordered by length. We then prove as another auxiliary result that a planner's solution to the game exists and constitutes an efficient equilibrium of the game.

## 3.1 Monotonicity

In a *monotone strategy*, stronger types $\ell'$ of the $\mathcal{C}$-juror hold out weakly longer than weaker types $\ell < \ell'$: If $\ell$ holds out until round $n$ with positive probability then $\ell'$ holds out until round $n$ with certainty.

**Lemma 1 (Monotone Best Responses)** *Every best response strategy of one juror to any strategy of her colleague is monotone.*

What makes this lemma non-trivial is that stronger type $\ell' > \ell$ is not just more convinced of $\mathcal{G}$ because of his own signal, but he also faces a different conditional distribution $r(\ell, m)g(m)$ of opponent types $m$. Still, by conditioning on the true state $\mathcal{G}, \mathcal{I}$, the proof shows that the benefit of holding out from round $n$ to round $n' > n$ satisfies single-crossing in $\ell$.

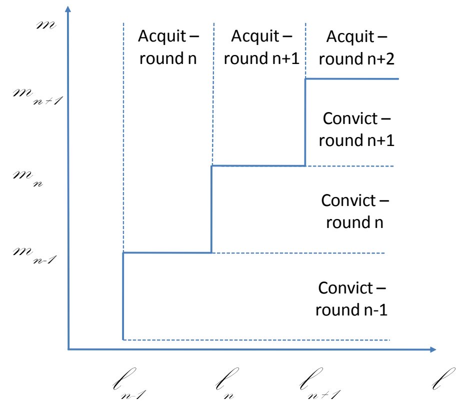


Figure 2: Equilibrium Decision Profile

The lemma yields a standard skimming-property of equilibria: Increasingly strong types quit in every round until, potentially, in some round communication terminates. We represent such strategies by increasing sequences of cutoff values $L = (\ell_n)_{n\in\mathbb{N}}$ and $M = (m_n)_{n\in\mathbb{N}}$: Type $\ell$ quits in round $n$ if $\ell \in (\ell_{n-1}, \ell_n)$, where $\ell_0 = -\infty$. The equilibrium decision profile is illustrated in figure 2.

We say that $\mathcal{C}$ *terminates the equilibrium* in round $n$, if *f*or some type profile $(\ell, m)$ the conversation reaches $\mathcal{C}$'s $n^{th}$ argument, i.e. $\ell_{n-1}, m_{n-1} < \infty$, but then $\mathcal{C}$ quits with certainty, i.e. $\ell_n = \infty$. In this case we also say that all later cutoffs $m_n, \ell_{n+1}, m_{n+1}$ are *off the equilibrium path.* The analogous definition applies for $\mathcal{A}$.

## 3.2 Equilibrium Cutoff Types

Given $\mathcal{A}$'s cutoffs $M = (m_n)_{n\in\mathbb{N}}$, the $\mathcal{C}$-juror's problem is to minimize his cost $C_n(\ell)$ by choice of the optimal stopping time $n = 1, 2, \dots, \infty$.

$$\begin{aligned}
C_n(\ell) &= \sum_{\nu=1}^{n-1} \Pr\left(\mathcal{A} \text{ quits in round } \nu\right)\left(2\nu\kappa + 1 - \pi\left(\ell, m\right)\right) \\
&\quad + \Pr\left(\mathcal{A} \text{ holds out until round } n\right)\left(\left(2n-1\right)\kappa + \pi\left(\ell, m\right)\right) \\
&= \sum_{\nu=1}^{n-1}\left(\int_{m_{\nu-1}}^{m_\nu} (2\nu\kappa + 1 - \pi(\ell, m)) r(\ell, m) g(m) dm\right) \\
&\quad + \int_{m_{n-1}}^{\infty} ((2n-1)\kappa + \pi(\ell, m)) r(\ell, m) g(m) dm
\end{aligned}$$

At the cutoff $\lambda_n$, the $\mathcal{C}$-juror is indifferent between quitting in round $n$ and holding out one more round. The indifference condition $C_n\left(\lambda_n\right) - C_{n+1}\left(\lambda_n\right) = 0$ can be written as[2]

$$\int_{m_{n-1}}^{m_n} (2\pi - 1 - \kappa) r g = \int_{m_n}^{\infty} 2\kappa r g. \tag{1}$$

The LHS represents the net benefit from enforcing $\mathcal{C}$ by holding out one more round. $2\pi - 1$ is the net improvement of the decision, which can be positive or negative depending on whether the opponent's type $m \in [m_{n-1}, m_n]$ is stronger or weaker than $\ell_n$, and $\kappa$ is the extra waiting cost to achieve this. The RHS represents the extra waiting cost $2\kappa$ if $m \geq m_n$ and $\mathcal{A}$ holds out in round $n$. The next lemma shows that equilibrium is fully characterized by this marginal analysis.

**Lemma 2 (Equilibrium Cut-offs)** *On the equilibrium path, the cutoff $\ell_n$ is either characterized by the indifference condition (1) or $\ell_n = \infty$. The analogous condition holds for $\mathcal{A}$'s cutoffs $m_n$.*

The lemma is immediate when there exist types $\ell, \ell'$ who optimally quit in rounds $n$ and $n+1$, given $\mathcal{A}$'s strategy $M$. The proof shows that such $\ell, \ell'$ exist whenever $m_{n-1}, \ell_n < \infty$.

In many proofs we will use an off-equilibrium version of (1): Let $P = C_n(\lambda_n) - C_{n+1}(\lambda_n)$ be the *propensity to hold out* one more round. We can write $P = P(m_{n-1}, \ell_n, m_n)$ because this propensity depends on $\mathcal{A}$'s strategy solely via the interval of types $[m_{n-1}, m_n]$ that are about to quit.

## 3.3 Zig-zag

To achieve indifference in equation (1), the ex-post preference for $\mathcal{C}$onviction $2\pi(\ell_n, m) - 1$ must exceed the cost of waiting one more round $\kappa$ in the most optimistic case $m = m_{n-1}$. Let $\varkappa$ be the public log-likelihood-ratio of $\mathcal{G}$ that leaves a payoff difference of $\kappa$ between $\mathcal{C}$onvict and $\mathcal{A}$cquit. So $\varkappa = \log \frac{\kappa}{1-\kappa}$, or equivalently $\kappa = \frac{e^{\varkappa}}{1+e^{\varkappa}}$.

We say that the monotone strategy profile $(L, M)$ satisfies the *zig-zag property* if all cutoff types $\ell_n, m_n$ on the equilibrium path obey the inequalities:

$$m_{n-1} + \varkappa \quad < \quad \ell_n \quad < \quad m_n - \varkappa.$$

With positive probability, increasingly strong types are quitting in alternating rounds.

**Lemma 3 (Zig-Zag)** *Any equilibrium $(L, M)$ satisfies the zig-zag property.*

[2] The omitted arguments in the integral are $\ell_n$ and the integration variable $m \in [m_{n-1}, m_n]$.

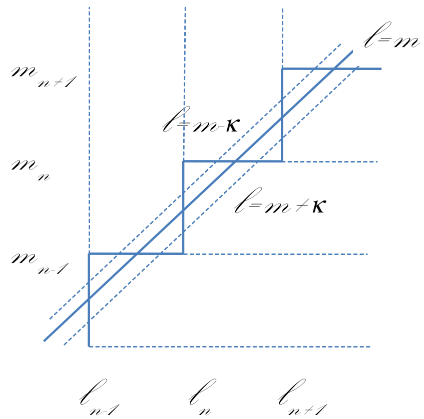


Figure 3: The equilibrium cutoff profiles $(\ell_n, m_n)$ zig-zag across a corridor around the indifference line. Thanks to the log-likelihoodratio representation, this picture is linear. The triangles are the areas in the type space where the wrong verdict is chosen and jurors' goal is to minimize the integral of $|2\pi(\ell, m) - 1|$ over these triangles.

**Proof** The first inequality, which is equivalent to $2\pi(\ell_n, m_{n-1}) - 1 > \kappa$, follows from (1) because $\pi$ is decreasing in $m$ and thus $0 < \int_{m_{n-1}}^{m_n}(2\pi - 1 - \kappa)rg < (2\pi(\ell_n, m_{n-1}) - 1 - \kappa)\int_{m_{n-1}}^{m_n} rg$. The second inequality owes to the analogous indifference condition of $m_n$.

Thus, in every equilibrium the public posterior $2\pi(\ell_n, m_n) - 1$ changes sign after every move. This property is illustrated in figure 3.

This lemma has two interesting implications. First, it shows that consecutive cutoff types $\ell_n, \ell_{n+1}$ are separated by at least $2\varkappa$. Thus, every profile of types quits in finite time. Second, the lemma highlights that the type $\ell_n$ of $\mathcal{C}$, who is indifferent between quitting and holding out in round $n$, will learn $\mathcal{A}$ is right if $\mathcal{A}$ holds out one more round. For then $m \in [m_n, \infty)$ is stronger than $\ell_n$. So $\mathcal{C}$ will then quit in round $n + 1$ not because he is tired of the conversation as in a regular war of attrition, but because he is actually convinced that his colleague was right and he was wrong.

## 3.4 Length of Equilibria

Comparing equilibria, we call $(L, M)$ *longer*, or slower, than $(L', M')$ if for all type combinations $(\ell, m)$ the decision occurs weakly later in $(L, M)$. This means $\ell_n \leq \ell'_n$ and $m_n \leq m'_n$ for all cutoffs on the equilibrium path, i.e. as long as $m'_{n-1} < \infty$ and $\ell'_n < \infty$.

**Lemma 4 (Ordered Equilibria)** *Any two equilibria are ordered by length.*

The outline of the proof is straight-forward: The indifference condition (1) pins down cutoff $m_n$ as a function of the preceding cutoffs $m_{n-1}$ and $\ell_n$. Thus, any equilibrium $(L, M)$ is already determined by the initial cutoff $\ell_1$. The log-concavity assumption then ensures that the order of the initial cutoffs $\ell_1 \leq \ell_1'$ is inherited by all later cutoffs.

### 3.5 Efficient Equilibria

Common interest games satisfy a second welfare theorem: Any planner's solution for the game is also an equilibrium, since any deviation from an efficient strategy profile raises joint costs, and thus own costs.

**Lemma 5 (Efficiency and Equilibrium)**
$(a)$ *A cost-minimizing strategy profile exists.*
$(b)$ *Any cost-minimizing strategy profile is an equilibrium.*

Part $(b)$ follows from the above argument. Part $(a)$ follows from a standard diagonal argument applied to the sequences of cutoffs $L$ and $M$.

## 4 Long Equilibria Are Better

How much information is aggregated in equilibrium depends on the choice of an equilibrium. We will show now that there are infinitely many equilibria in which communication between the jurors breaks down and less information is aggregated than would be Pareto optimal.

The easiest such equilibrium corresponds to an asymmetric outcome of a standard war of attrition. One juror, say $\mathcal{A}$, quits immediately by setting $m_1 = \infty$, anticipating that her colleague would insist on $\mathcal{C}$onvict forever. In a conversational war of attrition this insistent behavior need not start in period 0.

We call an equilibrium *insistent* if either juror terminates the equilibrium in some round, i.e. $m_n = \infty$ or $\ell_n = \infty$ for some $n \in \mathbb{N}$. This definition captures the notion that the other juror is insisting on her point, since the behavior on the equilibrium path for $(L, M)$ with, say, $m_{n-1} < \infty$ and $\ell_n = \infty$ is identical to the strategy profile where $\mathcal{A}$ insists on her verdict starting in round $n$, i.e. $m_{n-1} = m_n = \cdots$.

**Lemma 6 (Insistent Equilibria)** *For either juror and any $n \in \mathbb{N}$, there is an insistent equilibrium that is terminated by that juror in round $n$. These equilibria are unique up to changes off the equilibrium path.*

The critical step in the proof is the inference off the equilibrium path: If according to equilibrium, $\mathcal{A}$ terminates the equilibrium in round $n$ and all remaining types of $\mathcal{C}$ insist forever after, what will $\mathcal{C}$ infer if $\mathcal{A}$ holds out? Sequential equilibrium allows $\mathcal{C}$ to interpret $\mathcal{A}$'s holding out as a tremble of a weak type. This rationalizes insisting behavior by $\mathcal{C}$.

Insistent equilibria thus represent a communication failure. Strong types, say of $\mathcal{A}$, quit not because they are convinced that $\mathcal{C}$'s type is stronger, but because $\mathcal{A}$ cannot convince $\mathcal{C}$ that her own type is stronger. This could not happen if $\mathcal{C}$ was required to interpret off-equilibrium behavior of $\mathcal{A}$ as a signal of strength, rather than as a mistake. Proposition 1 formalizes this argument and shows that insistent equilbria do not satisfy the forward induction refinement, as defined in Cho (1987).

It is not essential for an insistent equilibrium that no types $\ell$ quit after some round $n$. It suffices that so few of them $[\ell_n, \ell_{n+1}]$ quit that waiting one more round is not worthwhile for $\mathcal{A}$, implying $m_n = \infty$. This communication break-down is avoided precisely when $\ell_n, m_n < \infty$ for all $n$. Therefore, we will call such an equilibrium *communicative.*

**Proposition 1 (Longer Equilibria are Better)**

($a$) *Communicative equilibria exist and are longer than any insistent equilibrium.*

($b$) *Equilibria are ex-ante Pareto ranked according to their length.*

($c$) *Communicative equilibria satisfy forward induction but insistent equilibria do not.*

Thus, maximal delay is optimal and to be expected whereas communication break-down is both unstable and unfavorable to both jurors.

**Proof** ($a$) The proof of Lemma 6 showed that every insistent equilibrium is strictly Pareto dominated. Thus an efficient equilibrium, which exists by Lemma 5, must be communicative. By Lemma 4 it is slower than any insistent equilibrium.

($b$) This follows from the proof of Lemma 6 if the shorter equilibrium is insistent. Further, Lemma 11 shows that if there are multiple communicative equilibria, then they all share the same expected total costs and are thus all ex-ante Pareto optimal.

($c$) In a communicative equilibrium every action node is reached on the equilibrium path, so forward induction has no bite. In an insistent equilibrium, say with $m_{n-1} = m_n = \cdots < \infty$ and $\ell_n = \infty$, forward induction forces $\mathcal{A}$ to infer that $\mathcal{C}$'s type is strong if $\mathcal{C}$ unexpectedly deviates by holding out in round $n$. More precisely, $\mathcal{A}$ must restrict his belief to types $\ell > m_{n-1}$. But then type $m_{n-1}$ optimally quits in round $n$, implying that this equilibrium does not satisfy forward induction. In the appendix we show that any insistent equilibrium with $\ell_n = \infty$ but $m_{n-1} < m_n$ does not satisfy forward induction either.

# 5 Conversations get Moot

We now fix an equilibrium $(L, M)$ and interpret it as an institution through which the jurors reach verdicts. We represent different cases as independent draws $(\ell, m)$. Our goal is to show that as a conversation evolves it is getting increasingly likely that the case $(\ell, m)$ is moot. Thus, when comparing conversations across different cases $(\ell, m)$ ex-post, our model predicts a negative correlation between the realized costs of information aggregation and the value of the aggregated information.

To make this idea precise we say that *conversations get moot over time* in equilibrium $(L, M)$ if, conditional on the correct verdict $\mathcal{A}$ or $\mathcal{C}$, and conditional on the conversation ending in round $n$, the probability of a moot case, i.e. $|\ell - m| < x$ for some threshold $x$, is increasing in $n$ for any $x$. In other words, the mootness of remaining cases is increasing over time in the sense of first-order-stochastic-dominance.

**Proposition 2** *In any equilibrium $(L, M)$, conversations get moot over time.*

We prove this by showing that the distribution $r(\ell, m) f(\ell) g(m)$ of $\ell - m$ conditional on $\ell \geq \ell', m = \ell'$ is decreasing in $\ell'$. Intuitively, the conditional expectation $\mathbb{E}[\ell - \ell' | \ell > \ell']$ of random variable $\ell$ exceeding threshold $\ell'$ is decreasing in the threshold, when the hazard rate of $\ell$ is increasing. This in turn is implied by assuming log-concavity of $f$.

This result is at odds with the standard result that higher investment leads to lower marginal returns but higher total returns. To reconcile these two views, consider the timing of our result. The negative correlation is established ex-post. When we consider the jurors' ex-ante problem of searching $(\ell, m)$ in the joint type space, the search exhibits diminishing marginal returns. Both the probability of terminating the conversation[3] and the expected value of the true information is decreasing in the number of rounds. As the benefit of investment into information aggregation is realized only when the information is utilized and the verdict is taken, the total return on investment ex-post is actually the marginal return to the investment ex-ante. This explains why the properties of investment into information differ from investment in, say, capital assets.

[3] not the hazard rate, but the probability as seen from the start of the conversation

Diminishing total returns are a robust feature of investment into information: Consider the dynamic information acquisition problem of a single agent. She will stop investing and take the decision early when early signals clearly indicate the optimal decision. A long spell of experimentation to the contrary indicates that the realized signals had little value. A late stopping-time can thus indicate that the agent has run out of promising experiments, i.e. she concludes that she just does not know and takes an uninformed decision. In this setting investment into information is negatively correlated with the value of the information for the decision, just as in our conversational war of attrition.

# 6 Conclusion

We have introduced a common interest model of information aggregation via coarse and costly arguing. In this model equilibrium communication break-down is possible, but unfavorable to both jurors, and not stable. In any equilibrium long conversations indicate that the case at hand is moot.

The correlation between the length of a conversation and the mootness of the aggregated information can be tested empirically. While mootness $(\ell - m)$ is not observable directly, it may be correlated with the probability that a verdict gets overturned in the future, when new information becomes available.

The arguing protocol is highly stylized in (1) allowing for only two possible arguments at every action node, and (2) entangling the arguing with the decision taking by assuming that a verdict is taken as soon as both jurors argue for the same verdict. This protocol is not optimal in, say, the class of protocols with fixed transmission costs per bit. Strong types can only signal their strength through holding out for many rounds. We now discuss that none of our main results is driven by assumption (1) or (2).

Suppose that alternatively to arguing for $\mathcal{A}$ or $\mathcal{C}$, a juror can also choose an argument from some set $X$ at each of his action nodes. The game ends when the jurors sequentially argue for the same verdict, $\mathcal{A}$ or $\mathcal{C}$. The arguments $x \in X$ have no intrinsic meaning but need to be interpreted according to equilibrium beliefs. The set of arguments $X = \{1, ..., x\}$ is finite, non-empty, and can be history-dependent $X = X(h)$.

We focus on monotone cutoff strategies: At every information set of juror 1, say, there are cutoffs $\ell_0 \leq \cdots \leq \ell_x$ such that types $\ell \in (\ell_{y-1}; \ell_y]$ play $y$, $\ell \in (-\infty; \ell_1]$ argues for $\mathcal{A}$, and $\ell \in (\ell_x; \infty)$ argues for $\mathcal{C}$. We assume that an efficient strategy profile exists, and thus is an equilibrium (Lemma 5).

**Longer Equilibria are Better**

These more general protocols still allow for insistent equilibria. In some round $n$ (after any history $h$) all remaining types $\ell$ of juror 1, say, propose to end the conversation by voting for one verdict, say $\mathcal{C}$, and juror 2 seconds the vote to end the game. Off the equilibrium path, either juror interprets deviations as trembles by an opponent type who really favors $\mathcal{C}$, and thus insists on $\mathcal{C}$. For the same reason as in the baseline model, this constitutes a sequential equilibrium, which does not satisfy forward-induction and which is Pareto-inferior to a communicative equilibrium.

**Conversations geg Moot**

Suppose that the conversation has already revealed that the jurors' types indicate different verdicts, i.e. $\ell, m > 0$,[4] and that juror 1 has just voted for some verdict, say $\mathcal{C}$. In the subsequent move the weakest remaining types $m$ of 2 will end the game by voting for $\mathcal{C}$. Conditional on $\mathcal{C}$ being the correct verdict, the remaining cases are more moot than the ones resolved by this action. In this sense, conversations get moot over time also in this more general setting.

One salient feature of the baseline model, that is not robust to this generalization, is that jurors have fixed roles during a conversation: Each non-terminal argument of the $\mathcal{C}$-juror, say, moves the public posterior towards guilt. This may help to explain why parties rarely switch sides during a conversation, first arguing one case and then its opposite. Consider the international discussion on global warming. Some countries insist it is happening, others say the evidence is not clear enough, some are getting convinced of the case made by others but the roles of the main actors do not switch over the course of the debate. The generalization of the model predicts that such switches become more common as the jurors' language becomes more sophisticated.

Another simplification in our model is the assumption of common interests. Crawford and Sobel (1982) and Li, Rosen and Suen (2001) find that conflicting interests obstruct communication and ex-post dominated verdicts are unavoidable in equilibrium. This result breaks down when we combine our technological communication frictions with their strategic frictions. It may be interesting whether the waiting costs necessarily outweigh the improvement in the decision, as is the case in Gul and Lundholm (1995), or whether repeated costly communication can enhance welfare as in Damiano, Li and Suen (2009a,b).

[4] For this more general protocol we need to assume that the pdfs $f$ and $g$ of $\ell$ and $m$ are log-concave in both tails. The assumption $\ell, m > 0$ then ensures that we can utilize log-concavity when comparing mootness.

# 7 Appendix

## 7.1 Game Including Initial Vote

The analysis in the paper assumes that the initial juror voted to $\mathcal{A}$cquit, and thereby took up the role of the defender. We now endogenize this initial argument. A strategy for juror 1 defines for each type $m$ an initial vote $\iota_m \in \{\mathcal{A}, \mathcal{C}\}$, and a stopping rule that is conditional on this initial vote. Similarly a strategy for juror 2 is given by a stopping rule contingent on the initial vote.

By Lemma 1, the strategies following an initial $\mathcal{A}$-vote are described by increasing sequences of cutoff values $(L, M)$ (and similarly by decreasing sequences of cutoffs after an initial $\mathcal{C}$-vote). We will now show that while juror 1 optimally uses a cutoff rule also in the initial round.

**Lemma 7** *Fix any strategy of juror* 2*. The best response of juror* 1 *is to pick a cutoff value* $m_0$ *and either*
($a$) *vote straight-forwardly, i.e. vote* $\mathcal{A}$ *if* $m > m_0$ *and vote* $\mathcal{C}$ *if* $m < m_0$*, or*
($b$) *vote reversely, i.e. vote* $\mathcal{A}$ *if* $m < m_0$ *and vote* $\mathcal{C}$ *if* $m > m_0$*.*

**Proof** Fix best responses of juror 1 after the intial vote. Then denote by $c(\mathcal{A}, \theta)$ the expected cost of initially voting $\mathcal{A}$ given true state $\theta \in \mathcal{G}, \mathcal{I}$ (and similarly for $\mathcal{C}$). Thus, the benefit of voting $\mathcal{A}$ rather than $\mathcal{C}$ is given by

$$\pi(m)(c(\mathcal{C}, \mathcal{G}) - c(\mathcal{A}, \mathcal{G})) + (1 - \pi(m))(c(\mathcal{C}, \mathcal{I}) - c(\mathcal{A}, \mathcal{I})).$$

The terms in the parentheses do not depend on $m$. As $\pi$ is decreasing in $m$ the whole expression is increasing in $m$ if $c(\mathcal{C}, \mathcal{I}) + c(\mathcal{A}, \mathcal{G})$ is greater than $c(\mathcal{C}, \mathcal{G}) + c(\mathcal{A}, \mathcal{I})$ and decreasing otherwise. In either case, the optimal initial voting rule is described by some cutoff point $m_0$.

Case ($a$) captures the equilibria discussed in the body of the paper where we assume that $m \in [m_0, \infty)$. Case ($b$) are the "devil's advocate" equilibria discussed in Section 2. If juror 1 expects contrarian behaviour from juror 2, voting for a verdict makes it less likely to prevail in the end. Thus $c(\mathcal{C}, \mathcal{I}) + c(\mathcal{A}, \mathcal{G})$ is less than $c(\mathcal{C}, \mathcal{G}) + c(\mathcal{A}, \mathcal{I})$ and reverse voting is optimal. The conceptually easiest "devil's advocate" equilibrium has juror 2 insist on her verdict forever. This is not the only kind of "devil's advocate" equilibrium. Rather there are equilibria in which the conversation may continue a finite number of arguments (much like the insistent and communicative equilibria after a straight-forward initial vote in section 4). These are "double devil's advocate" equilibria: The marginal type of the $\mathcal{C}$-juror actually has evidence for $\mathcal{I}$nnocence, i.e. $\lambda < 0$, but is reluctant to vote for $\mathcal{A}$cquit because he correctly anticipates that the marginal type of her colleague, who argues for $\mathcal{A}$cquittal, has evidence for $\mathcal{G}$uilt, i.e. $\mu < 0$.

The main insights of the paper carry over to the game including the initial vote.

1. While equilibria in the larger game are no longer ranked according to their length, there are still communicative and insistent equilibria and it is still true that the latter do not satisfy forward induction and are Pareto dominated.

2. Long conversations are still indicate moot informaton, at least eventually. The complication is that a juror may take up either side of the argument and the result obtains once the conversation has reached the tail of the distribution, if we make the log-concavity condition on both tails of the distribution.

## 7.2 Omitted Proofs

**Lemma 8 (Unconditional Correlation)** *Jurors' types are correlated with "local" correlation $r(\ell, m) = h(\ell, m)/f(\ell)g(m)$ given by:*

$$r(\ell, m) = \frac{2(e^\ell + e^m)}{(1 + e^\ell)(1 + e^m)} \in (0, 2)$$

**Proof** Denote by $f^\theta(\ell), g^\theta(m)$ and $h^\theta(\ell, m)$ the pdfs in state $\theta \in \{\mathcal{I}, \mathcal{G}\}$. The likelihood ratios satisfy the no-introspection condition as in (**?**) $\frac{f^\mathcal{G}(\ell)}{f^\mathcal{I}(\ell)} = e^\ell$, $\frac{g^\mathcal{G}(m)}{g^\mathcal{I}(m)} = e^{-m}$ and $\frac{h^\mathcal{G}(\ell,m)}{h^\mathcal{I}(\ell,m)} = e^{\ell-m}$. Therefore:

$$\begin{aligned}
h(\ell, m) &= \frac{1}{2}h^\mathcal{I}(\ell, m) + \frac{1}{2}h^\mathcal{G}(\ell, m) \\
&= \frac{1 + e^{\ell-m}}{2}h^\mathcal{I}(\ell, m) = \frac{1 + e^{\ell-m}}{2}f^\mathcal{I}(\ell)g^\mathcal{I}(m) \\
&= \frac{1 + e^{\ell-m}}{2}\frac{2}{1 + e^\ell}f(\ell)\frac{2}{1 + e^{-m}}g(m) \\
&= \frac{2(e^\ell + e^m)}{(1 + e^\ell)(1 + e^m)}f(\ell)g(m).
\end{aligned}$$

**Lemma 9 (Log-supermodular Correlation)** *As function of $\lambda$ and $\delta$, the correlation function $r(\lambda+\delta,\lambda)$ is strictly log-supermodular, i.e.*

$$\frac{r(\ell'+\delta,\ell')}{r(\ell+\delta,\ell)} > \frac{r(\ell'+\delta',\ell')}{r(\ell+\delta',\ell)}$$

*for all $\ell' > \ell$ and $\delta' > \delta$.*

**Proof** By Lemma 8 we have

$$r(\ell,\ell+\delta) = \frac{2e^{\ell}\left(1+e^{\delta}\right)}{\left(1+e^{\ell}\right)\left(1+e^{\ell+\delta}\right)},$$

so all but the $1+e^{\ell+\delta}$ term cancel in $\frac{r(\ell'+\delta,\ell')}{r(\ell+\delta,\ell)} > \frac{r(\ell'+\delta',\ell')}{r(\ell+\delta',\ell)}$. Thus, it suffices to show that

$$\frac{1+e^{\ell+\delta}}{1+e^{\ell'+\delta}} > \frac{1+e^{\ell+\delta'}}{1+e^{\ell'+\delta'}}$$

which follows from the fact that $e^{\ell+\delta}$ is strictly supermodular in $\ell$ and $\delta$.

**Proof of Lemma 1: Monotonicity** Fix a strategy of the $\mathcal{A}$-juror. We will show that the $\mathcal{C}$-juror's benefits from holding out satisfy single-crossing in $\ell$. Denote by $\Pi(n',n,\theta)$ the probability that the $\mathcal{A}$-juror will give in between rounds $n$ and $n'>n$ conditional on state $\theta$, and by $w(n',n,\theta)$ the expected incremental waiting cost from holding out until $n'$ rather than quitting in round $n$.

The total incremental cost from holding out until round $n'$ vs. quitting in round $n$ is thus:

$$\pi(\ell)(-\Pi(n',n,\mathcal{G})+w(n',n,\mathcal{G}))+(1-\pi(\ell))(\Pi(n',n,\mathcal{I})+w(n',n,\mathcal{I})) \qquad (*)$$

Incremental waiting costs $w$ are always positive while incremental decision costs are negative (resp. positive) when holding out until $n'$ changes the verdict to $\mathcal{C}$ and the state is $\mathcal{G}$ (resp. $\mathcal{I}$).

By the single-crossing property it suffices to show that (*) is strictly decreasing in $\ell$ if it is non-positive for some $\ell$: If type $\ell$ weakly prefers to hold out from round $n$ to $n'$, then every more extreme type $\ell' > \ell$ does so strictly.

The second term in (*) is strictly positive. So for (*) to be non-positive for some $\ell$ it needs to be the case that the first term in (*) is strictly negative: Holding out until round $n'$ can only be good for type $\ell$ if it is profitable conditional on the $\mathcal{G}$uilt of the defendant. As $\frac{d}{d\ell}\pi(\ell) > 0$, $\frac{d}{d\ell}(1-\pi(\ell)) < 0$, and $\Pi(n', n, \theta) + w(n', n, \theta)$ is independent of $\ell$, (*) is strictly decreasing in $\ell$.

**Proof of Lemma 2: Equilibrium Cutoffs** Fix $M = (m_n)$ with $m_{n-1} < \infty$ and assume that $\ell_n < \infty$. If there exist types $\ell, \ell'$ who optimally quit in rounds $n$ and $n+1$, then $\ell_n \in [\ell, \ell']$ is uniquely defined by equation (1). We now argue that such types $\ell, \ell'$ exist.

If no type $\ell$ optimally quits in round $n$, i.e. $\ell_{n-1} = \ell_n$, then any type $m$ of the $\mathcal{A}$-juror would rather quit in round $n-1$ than in round $n$: This entails the same decision but lower waiting costs. So, no type of $\mathcal{A}$ will quit in round $n$, i.e. $m_{n-1} = m_n$. But then the same logic implies $\ell_n = \ell_{n+1}$. Iterating this argument, we get $m_{n-1} = m_n = m_{n+1} = \cdots < \infty$ and $\ell_n = \ell_{n+1} = \cdots < \infty$. This strategy profile cannot constitute an equilibrium. After round $n$ any remaining type expects infinite waiting costs and will optimally quit.

If no type $\ell'$ optimally quits in round $n+1$, i.e. if $\ell_n = \ell_{n+1} < \infty$, the above argument implies $m_n = m_{n+1}$. If $m_n = m_{n+1} < \infty$ we get the prior contradiction. But if $m_n = \infty$, i.e. the $\mathcal{A}$-juror quits in round $n$ with certainty, then it is weakly optimal for all remaining types $\ell' > \ell_n$ to quit, off the equilibrium path, in round $n+1$.

**Proof of Lemma 4: Ordered Equilibria** We need to show that for all rounds $n$ we have $\ell_n < \ell'_n$ as long as $m'_{n-1} < \infty$, and $m_n < m'_n$ as long as $\ell'_n < \infty$. We proceed by induction, using as basis that $m_0 = m'_0$ and $\ell_1 < \ell'_1$.

The induction step will be that $m_{n-1} \leq m'_{n-1}$ and $\ell_n - m_{n-1} < \ell'_n - m'_{n-1}$ implies that $\ell_n < \ell'_n$ and $m_n - \ell_n < m'_n - \ell'_n$ (as long as $\ell'_n < \infty$).[5]

The first part, that $\ell_n < \ell'_n$, follows by adding the two inequalities of the induction hypothesis.

The second part is based on properties of the propensity to hold out, proven below in Lemma 10. We then define $\delta := \ell'_n - \ell_n > 0$ and compare

$$\begin{aligned} 0 &= P(m_{n-1}, \ell_n, m_n) \\ &< P(m_{n-1} + \delta, \ell_n + \delta, m_n + \delta) \\ &< P(m'_{n-1}, \ell'_n, m_n + \delta). \end{aligned}$$

[5] Of course, there is another, analogous half of the induction step: Show that $\lambda_n \leq \lambda'_n$ and $\mu_n - \lambda_n < \mu'_n - \lambda'_n$ implies that $\mu_n < \mu'_n$ and $\lambda_{n+1} - \mu_n < \lambda'_{n+1} - \mu'_n$ (as long as $\mu'_n < \infty$).

The first inequality is part (3) of Lemma 10. The second inequality follows because the induction hypothesis $\ell_n - m_{n-1} < \ell'_n - m'_{n-1}$ can be rewritten as $m'_{n-1} < m_{n-1} + \delta$, and $P$ is decreasing in its first argument by part (2) of Lemma 10.

From $P(m'_{n-1}, \ell'_n, m'_n) = 0 < P(m'_{n-1}, \ell'_n, m_n + \delta)$, part (1) of Lemma 10 implies $m_n + \delta < m'_n$, and after rearranging $m_n - \ell_n < m'_n - \ell'_n$ as desired.

**Lemma 10 (Monotone Propensity to Hold Out)**

$$P(m_{n-1}, \ell_n, m_n) = \int_{m_{n-1}}^{m_n} (2\pi - 1 - \kappa) rg - \int_{m_n}^{\infty} 2\kappa rg$$

*is*

1. *Decreasing in $m_{n-1}$ and $m_n$ as long as $m_{n-1} + \varkappa < \ell_n < m_n - \varkappa$.*

2. *Increasing in $\ell_n$ with $\lim_{\ell_n \to \infty} P(m_{n-1}, \ell_n, \infty) > 0$.*

3. *"Diagonally increasing", i.e. $P(m_{n-1}, \ell_n, m_n) < P(m_{n-1} + \delta, \ell_n + \delta, m_n + \delta)$ for $\delta > 0$.*

**Proof** (1) follows because $-(2\pi(\ell_n, m_{n-1}) - 1 - \kappa) < 0$ for $m_{n-1} + \varkappa < \ell_n$, and because $2\pi(\ell_n, m_{n-1}) - 1 - \kappa < 2\kappa$ for $\ell_n < m_n - \varkappa$. Intuitively, holding out and reaching a $\mathcal{C}$onvict verdict against types $m \in [m_{n-1}, m_n]$ of the colleague becomes less attractive as the strength of these types increases.

The first part of (2) follows by Lemma 1. The second part follows because $2\pi(\ell, m) - 1 - \kappa$ is positive for $\ell > m + \varkappa$, and the "density function" $rg$ is putting all of its weight on such $m$ as $\ell \to \infty$: If $\mathcal{A}$ is about to terminate, a sufficiently strong type of $\mathcal{C}$ will prefer to hold out.

For part (3) we need to show that

$$\begin{aligned}
& \int_{m_{n-1}}^{m_n} (2\pi(\ell_n, m) - 1 - \kappa) r(\ell_n, m) g(m)\, dm - 2\kappa \int_{m_n}^{\infty} r(\ell_n, m) g(m)\, dm \\
\leq \; & \int_{m_{n-1}}^{m_n} (2\pi(\ell_n + \delta, m + \delta) - 1 - \kappa) r(\ell_n + \delta, m + \delta) g(m + \delta)\, dm - \\
& -2\kappa \int_{m_n}^{\infty} r(\ell_n + \delta, m + \delta) g(m + \delta)\, dm
\end{aligned}$$

By the log-concavity of $g$ and Lemma 9 the density $r(\ell_n, m) g(m)$ MLRP-dominates $r(\ell_n + \delta, m + \delta) g(m + \delta)$. The integrand (net of this density!) is the same on both sides of the inequality as $\pi(\ell_n, m) = \pi(\ell_n + \delta, m + \delta)$.

If the integrand was decreasing in $m$, the lemma would be established. But while $2\pi(\ell_n, m) - 1 - \kappa$ is decreasing in $m$, the integrand jumps up at $m_n$ because $2\pi(\ell_n, m_n) - 1 - \kappa < -2\kappa$.

We abbreviate notation by setting $\phi(m) = 2\pi(\ell_n, m) - 1 - \kappa$, $\psi(m) = \frac{r(\ell_n,m)g(m)}{\int_{m_{n-1}}^{m_n} r(\ell_n,m)g(m)dm}$, $\psi^\delta(m) = \frac{r(\ell_n+\delta,m+\delta)g(m+\delta)}{\int_{m_{n-1}}^{m_n} r(\ell_n+\delta,m+\delta)g(m+\delta)dm}$. Thus, we need to show:

$$\begin{aligned}
&\int_{m_{n-1}}^{m_n} \phi(m)\psi(m)\,dm - 2\kappa\int_{m_n}^{\infty}\psi(m)\,dm \\
<& \int_{m_{n-1}}^{m_n} \phi(m)\psi^\delta(m)\,dm - 2\kappa\int_{m_n}^{\infty}\psi^\delta(m)\,dm.
\end{aligned}$$

Now first, $\psi(m)$ MLRP dominates $\psi^\delta(m)$, both are probability densities on $[m_{n-1}, m_n]$ and $\phi$ is decreasing on this interval. This implies

$$\int_{m_{n-1}}^{m_n} \phi(m)\psi(m)\,dm < \int_{m_{n-1}}^{m_n} \phi(m)\psi^\delta(m)\,dm.$$

Second, we know that

$$\begin{aligned}
\int_{m_n}^{\infty}\psi(m)\,dm &= \frac{\int_{m_n}^{\infty} r(\ell_n, m)\,g(m)\,dm}{\int_{m_{n-1}}^{m_n} r(\ell_n, m)\,g(m)\,dm} > \\
&> \frac{\int_{m_{n-1}}^{m_n} r(\ell_n+\delta, m+\delta)\,g(m+\delta)\,dm}{\int_{m_{n-1}}^{m_n} r(\ell_n+\delta, m+\delta)\,g(m+\delta)\,dm} = \int_{m_n}^{\infty}\psi^\delta(m)\,dm
\end{aligned}$$

as $r(\ell_n, m)\,g(m)$ MLRP dominates $r(\ell_n+\delta, m+\delta)\,g(m+\delta)$. This completes the proof.

**Proof of Lemma 5: Efficiency and Equilibrium** In looking for the optimal strategy profile we can restrict ourselves to monotone strategies represented by vectors of cutoff values $(L, M)$ because any best-response is in monotone strategies by Lemma 1. Rather than considering the cutoffs in log-likelihood notation $\ell_n, m_n \in (-\infty, \infty)$ it is useful to compactify the type space. Thus, consider as transformed types values of the cdfs $\alpha_n := \int_{-\infty}^{\ell_n} f, \beta_n := \int_{m_0}^{m_n} g \in [0,1]$ and vectors $a = (\alpha_n)_n, b = (\beta_n)_n$ thereof. We can restrict attention to strategy profiles $(a, b)$ with either $\lim\alpha_n = 1$ or $\lim\beta_n = 1$ (profiles that do not satisfy this property have expected total costs of $\infty$).

So consider any sequence $(a^\nu, b^\nu)_\nu$ of strategy profiles that approaches the lower cost limit $\lim_{\nu\to\infty} c_\kappa(a^\nu, b^\nu) = c(\kappa)$ [6]. We will show that there exists a subsequence $(a^{\nu_\iota}, b^{\nu_\iota})_\iota$ that converges point-wise to a cost-minimizing strategy profile $(a^*, b^*)$, i.e. with $c(a^*, b^*) = c(\kappa)$.

The proof is a typical "diagonal argument". Consider the sequence of the first cutoff points $(\alpha_1^\nu)_\nu$. As a sequence in the compact space $[0,1]$ it has a convergent subsequence which we denote by $(\alpha_1^{\nu_\iota^1})_\iota$. Similarly, considering the first $n \in \mathbb{N}$ cutoff points we can find a subsequence $(\nu_\iota^n)_\iota$ such that $(\alpha_1^{\nu_\iota^n}, \beta_1^{\nu_\iota^n}, \ldots, \alpha_n^{\nu_\iota^n}, \beta_n^{\nu_\iota^n})_\iota$ converges in $[0,1]^{2n}$. Thus the diagonal sub-sequence $(\alpha^{\nu_\iota^\iota}, \beta^{\nu_\iota^\iota})_\iota$ converges point-wise against some sequence of cutoffs $(a^*, b^*)$.

It remains to be shown that $c(a^*, b^*) = \lim_{\iota\to\infty} c_\kappa(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota}) = c(\kappa)$. We will show for any $\varepsilon$ that $c(a^*, b^*) \leq c_\kappa(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota}) + \varepsilon$ for all sufficiently large $\iota$. First note that for sufficiently large $n$ and $\iota$ we have that $\alpha_n^{\nu_\iota^\iota}$ and $\alpha_n^*$ (or $\beta_n^{\nu_\iota^\iota}$ and $\beta_n^*$) are arbitrarily close to 1 and therefore the probability of changing the verdict after round $n$ by switching from strategy profile $(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota})$ to $(a^*, b^*)$ is going to zero. Also, the loss in case of a wrong verdict is bounded by 1, and so the expected loss from switching is bounded by $\frac{\varepsilon}{2}$.

As for the cutoff types 1 through $n$, there exists $\iota$ above which these first $2n$ cutoff points of $(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota})$ are uniformly so close to those of $(a^*, b^*)$, that the expected loss of inducing the wrong verdict before round $n$ by changing strategy profile $(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota})$ to $(a^*, b^*)$ is bounded by $\frac{\varepsilon}{2}$, proving that $c_\kappa(a^*, b^*) \leq c_\kappa(a^{\nu_\iota^\iota}, b^{\nu_\iota^\iota}) + \varepsilon$.

**Proof of Lemma 6: Insistent Equilibria** We will show existence of the equilibrium terminated by $\mathcal{C}$ in round $n$.

Consider the socially optimal cutoff profile $(L, M)$ subject to $\ell_n = \infty$ and $m_{n-1} = m_n = \cdots$. $(L, M)$ is an equilibrium: There are no profitable deviation after the $n$th argument of $\mathcal{C}$ because those are off the equilibrium path. There are also no deviations before round $n$ by definition of $(L, M)$ and Lemma 5. The discussion after the statement of the lemma implies that the equilibrium is sequential.

To show that $\ell_n$ is on the equilibrium path, i.e. $m_{n-1}, \ell_n < \infty$, assume to the contrary that the equilibrium path terminates at an earlier round, say in round $n' < n$ with the move of the $\mathcal{C}$-juror, i.e. $m_{n'-1} < \ell_{n'} = \infty$. As the $n'^{th}$ argument of the $\mathcal{A}$-juror is off the equilibrium path, this strategy profile remains optimal, and thus an equilibrium, after setting $m_{n'} = \infty$. However, part (2) of Lemma 10 states that sufficiently strong types $\ell$ will hold out if all opponent types are about to quit, i.e.

[6] For a formal definition of $c_\kappa(a^\nu, b^\nu)$, compare Section **??**

$\lim_{\ell\to\infty} P(m_{n'-1},\ell,\infty) > 0$. Thus, $\ell_{n'} = \infty$ is not optimal for the $\mathcal{C}$-juror and thus not a cost-minimizing strategy profile, in contradiction to its definition. This argument implies that longer equilibria strictly Pareto dominate shorter insistent equilibria.

Uniqueness follows by Lemma 4.

**Lemma 11** *If there are multiple communicative equilibria $(L, M)$ and $(L', M')$ then for every intermediate value $\ell \in [\ell_1, \ell_1']$ there is a communicative equilibrium $(L'', M'')$ with $\ell_1'' = \ell$.*

*All communicative equilibria share the same ex-ante expected total costs.*

**Proof** Remember that the equilibria $(L, M)$ and $(L', M')$ are uniquely determined by their initial cutoff values $\ell_1$ and $\ell_1'$. Any value $\ell_1'' \in [\ell_1, \ell_1']$ defines a sequence $m_1'', \ell_2'', m_2'', \ldots$ and by Lemma 4 we have $\ell_n'' \in [\ell_n, \ell_n'] \subset \mathbb{R}$ and $m_n'' \in [m_n, m_n'] \subset \mathbb{R}$, implying that all cutoffs $\ell_n'', m_n''$ are finite and $(L'', M'')$ is an equilibrium by Lemma 2.

For the second claim, let $(L, M)(\ell)$ be the equilibrium defined by $\ell \in [\ell_1, \ell_1']$ and by $\ell_n(\ell), m_n(\ell)$ the respective cutoff values. As $\ell_n(\ell), m_n(\ell)$ are finite and the signal distributions are continuous and bounded on any compact set, $\ell_n(\ell), m_n(\ell)$ are differentiable as functions of $\ell$. We denote by $C(\ell) = C(L, M)_\ell$ the total cost (waiting cost plus decision cost) of the equilibrium $(L, M)_\ell$ and will now show that $C'(\ell) = 0$ on $[\ell_1, \ell_1']$.

For any $\varepsilon$ there exists $n \in \mathbb{N}$ such that $C(L, M)_\ell$ decreases by at most $\varepsilon^2$ if we set the cutoffs $\ell_{n'}, m_{n'}$ equal to $\infty$ for rounds $n' \geq n$. Now $C(\ell + \varepsilon) - C(\ell)$ is a sum of terms that are either of the kind $\frac{\partial}{\partial \ell_n} c(L, M)_\ell \frac{d\ell_n}{d\ell}(\ell)\,\varepsilon$ or of degree $\varepsilon^2$. As $(L, M)$ is an equilibrium $\frac{\partial}{\partial \ell_n} c(L, M)_\ell = 0$ and the first class of terms vanishes, implying that $C'(\ell) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(C(\ell + \varepsilon) - C(\ell)) = 0$.

**Proof of Proposition 1 (Longer Equilibria are Better) continued** We still need to show that an insistent equilibrium with $\ell_n = \infty$ but $m_{n-1} < m_n$ does not satisfy forward induction. Lemma 10 shows that there are types $\ell > m_{n-1}$ that optimally hold out if $\mathcal{A}$ increases her next cutoff $m_n$ sufficiently. Let $\widetilde{\ell} := \inf\{\ell | \exists m : P(m_{n-1}, \ell, m) \geq 0\}$ be the smallest such type.

Forward induction restricts $\mathcal{A}$'s beliefs in round $n$ to the interval $[\widetilde{\ell}, \infty)$. This restricts her best-responses to $m_n > \widetilde{\ell} + \varkappa$. This implies that any type of $\mathcal{C}$ with

$\ell \geq m_n - \varkappa$ optimally holds out in round $n$ and thus $\ell_n = \infty$ is not optimal. This follows by

$$\begin{aligned} 0 &\leq P(m_{n-1}, \widetilde{\ell}, \widetilde{\ell} + \varkappa) \\ &< P(m_{n-1}, m_n - \varkappa, \widetilde{\ell} + \varkappa) \\ &< P(m_{n-1}, m_n - \varkappa, m_n). \end{aligned}$$

Line (1) follows by definition of $\widetilde{\ell}$ and the fact that $P(m_{n-1}, \widetilde{\ell}, \cdot)$ is maximized by $\widetilde{\ell} + \varkappa$. Line (2) follows by the assumption that $m_n > \widetilde{\ell} + \varkappa$ and the fact that $P$ is strictly increasing in its second argument. Line three follows again by the argument that $P(m_{n-1}, m_n - \varkappa, \cdot)$ is maximized by $m_n$.

**Proof of Proposition 2** We will show that the likelihood ratio $\frac{\Pr(\ell - m = \delta, (\ell, m) \in E_n)}{\Pr(\ell - m = \delta', (\ell, m) \in E_n)}$ of more moot information $\delta \leq \delta'$ is increasing in $n$, where $E_n := \{(\ell, m) : m \in [m_{n-1}, m_n], \ell \geq m\}$ is the event that $\mathcal{C}$ is the correct verdict and the conversation terminates in round $n$.

After some rearranging, this condition translates into:

$$\frac{\int_{m_n}^{m_{n+1}} f(m+\delta) g(m) r(m+\delta, m) dm}{\int_{m_{n-1}}^{m_n} f(m+\delta) g(m) r(m+\delta, m) dm} \geq \frac{\int_{m_n}^{m_{n+1}} f(m+\delta') g(m) r(m+\delta', m) dm}{\int_{m_{n-1}}^{m_n} f(m+\delta') g(m) r(m+\delta', m) dm}.$$

The proof now follows by the log-concavity of $f$, i.e. $\frac{f(m'+\delta)}{f(m+\delta)} \geq \frac{f(m'+\delta')}{f(m+\delta')}$ for $m' \geq m$ and $\delta' \geq \delta$, and the log-concavity of $r$, i.e. $\frac{r(m'+\delta, m')}{r(m+\delta, m)} \geq \frac{r(m'+\delta', m')}{r(m+\delta', m)}$ derived in Lemma 9.

If we were to replace the integrals over $m$ in the desired inequality above with the values of the integrand at some intermediate points $m \in [m_{n-1}, m_n]$ and $m' \in [m_n, m_{n+1}]$, the $g$-terms would cancel and the inequality would obtain because of the two log-concavity conditions. For the integrals, we use an argument from (**?**) to obtain

$$\begin{aligned} \frac{\int_{m_n}^{m_{n+1}} f(m+\delta) g(m) r(m+\delta, m) dm}{\int_{m_{n-1}}^{m_n} f(m+\delta) g(m) r(m+\delta, m) dm} &= \frac{\int_{m_n}^{m_{n+1}} \frac{f(m+\delta) g(m) r(m+\delta, m)}{f(m_n+\delta) g(m_n) r(m_n+\delta, m_n)} dm}{\int_{m_{n-1}}^{m_n} \frac{f(m+\delta) g(m) r(m+\delta, m)}{f(m_n+\delta) g(m_n) r(m_n+\delta, m_n)} dm} \geq \\ \geq \frac{\int_{m_n}^{m_{n+1}} \frac{f(m+\delta') g(m) r(m+\delta', m)}{f(m_n+\delta') g(m_n) r(m_n+\delta', m_n)} dm}{\int_{m_{n-1}}^{m_n} \frac{f(m+\delta') g(m) r(m+\delta', m)}{f(m_n+\delta') g(m_n) r(m_n+\delta', m_n)} dm} &= \frac{\int_{m_n}^{m_{n+1}} f(m+\delta') g(m) r(m+\delta', m) dm}{\int_{m_{n-1}}^{m_n} f(m+\delta') g(m) r(m+\delta', m) dm}. \end{aligned}$$

# References


[1] Aumann, Robert and Sergiu Hart (2003) "Long Cheap Talk", Econometrica 71 (6), 1619-1660

[2] Austen-Smith, David and Jeffrey Banks (1996) "Information Aggregation, Rationality, and the Condorcet Jury Theorem", American Political Science Review 90 (1), 34-45

[3] Cho, In-Koo (1987) "A Refinement of Sequential Equilibrium", Econometrica 55 (6), 1367-1389

[4] Crawford, Vincent P and Sobel, Joel (1982) "Strategic Information Transmission", Econometrica 50 (6), 1431-1451

[5] Damiano, Ettore and Li, Hao and Suen, Wing (2009a) "Delay in Strategic Information Aggregation", mimeo

[6] Damiano, Ettore and Li Hao and Suen, Wing (2009b) "Optimal Deadlines for Agreements", mimeo

[7] Feddersen, Timothy and Wolfgang Pesendorfer (1998) "Voting Behavior and Information Aggregation in Elections with Private Information", Econometrica 65 (5), 1029-1058

[8] Geanakoplos, John and Polemarchakis, Heracles M. (1982) "We Can't Disagree Forever", JET 28 (1), 192-200

[9] Gul, Faruk and Lundholm, Russell (1995) "Endogenous Timing and the Clustering of Agents' Decisions", JPE 103 (5), 1039-1066

[10] Li, Hao and Rosen, Serwin and Suen, Wing (2001) "Conflicts and Common Interests in Committees", AER 91 (5), 1478-1497

[11] Milgrom, Paul and Roberts, John (1990) "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities", Econometrica 58 (6), 1255-1277

[12] Piketty, Thomas (2000) "Voting as Communicating", ReStud 67 (1), 169-191

[13] Smith, Lones and Peter Sorensen (2000) "Pathological Outcomes of Observational Learning", Econometrica 68 (2), 371-398