$n \in [6, 12]$ Angry Men: The Importance of Endogenizing Jury Size When Comparing Voting Rules

Michael Thaler

Harvard University
150,000 criminal jury trials each year in the US:

- Many jury trials end in false convictions or false acquittals
  - 13% of verdicts inaccurate (Spencer, 2007)
- Many juries hang, leading to a mistrial
  - 6.2% of state juries hang (Hannaford et al., 2002)
- Suggests substantial room for improvement

**Goal:** Design voting rules to increase accuracy, decrease hung juries
Space of voting rules

Trial in which jurors vote to acquit or convict a defendant and outcome either $A$, $C$, or hang:

- **Size of jury:** $n > 0$
- **Symmetric voting rules:** $d \in \left[0, \frac{n-1}{2}\right]$
  - Reach verdict iff at least $n - d$ agree
  - Else, hung jury
  - Symmetric used in most US jury trials
- **Unanimity:** $d = 0$
- (Majority rule: $n$ odd, $d = \frac{n-1}{2}$)
Pre-1970: US norm that criminal trials have $n = 12, d = 0$.\(^1\)

In 1970s, Supreme Court issues rulings:

- **Williams v. Florida (1970):** $n \in [6, 12], d = 0$ allowed.
- **Apodaca v. Oregon (1972)** and **Johnson v. Louisiana (1972):** $n = 12, d \leq 3$ allowed.
- **Burch v. Louisiana (1979):** If $n = 6$, then only $d = 0$ allowed.
- **Ballew v. Georgia (1978):** $n < 6$ *not* allowed.
  - Justice Lewis Powell, writing for the majority: “Though the line between five- and six-member juries is difficult to justify, a line has to be drawn somewhere.”

\(^1\)Often these jurors were all men, and occasionally they were angry.
Space of voting rules in the US

Figure: Plot of $(n, d)$ used anywhere in the US as of 2018. Horizontally-striped squares used only in criminal cases; vertically-striped squares only in misdemeanor or civil cases; both-striped used for both.
How to design jury voting rules?

- Desiderata:
  - Reach some verdict
  - Acquit innocent defendants
  - Convict guilty defendants

- Questions:
  - What should \( n \) be?
  - What should \( d \) be?

This paper: Important to consider both questions simultaneously
This paper

Accuracy \equiv P(\text{Correct verdict} \mid \text{Some verdict})
Efficiency \equiv P(\text{Some verdict})

Effect of size $n$ and allowed dissenters $d$:

- **Fix $n$, vary $d$:** Accuracy $\uparrow \iff$ Efficiency $\downarrow$
- **Vary $n$, fix $d$:** Accuracy $\uparrow \iff$ Efficiency $\downarrow$
- **Vary $n$, vary $d$:** Accuracy $\uparrow$ and Efficiency $\uparrow$ possible

Unanimity is suboptimal: $(n + 2, 1)$ more accurate, efficient than $(n, 0)$

$n \in [6, 12]$ Angry Men

Michael Thaler

• Varying $n$ and $d$ has been studied calibrationally: Guerra, Luppi, Parisi (2018) vary $n$, $d$, standard of proof; King and Nesbit (2009) look at cost and asymmetric errors.

• Classic political science literature on juries: Kalven and Ziesel’s *The American Jury* (1966); Hans et al. (2003); Grofman critiques the Supreme Court on its arbitrary rulings (1976, 1980); Thomas and Pollack (1992).

A jury voting game (from Coughlan, 2000)

- Defendant on trial. State of the world: Guilty (G) or Innocent (I). Prior $P(G) = \pi; P(I) = 1 - \pi$
- Jury composed of jurors $N = \{1, \ldots, n\}$.
- Conditionally independent signals $i, g$ with accuracy $p \in (1/2, 1)$:
  - $P(i|I) = P(g|G) = p$
  - $P(i|G) = P(g|I) = 1 - p$
- Symmetric voting rule $(n, d)$:
  - If # convicting jurors $\geq n - d$, then convict (C)
  - If # acquitting jurors $\geq n - d$, then acquit (A)
  - Else, hung jury and mistrial (M)
- Timing: jurors receive private signals in period 0, simultaneously vote to convict or acquit in period 1
  - If C or A, true state revealed in period 2
  - If M, re-try with new jurors, drawn iid, who do not know previous vote
Juror utilities

- Utility of juror $j$ from outcome $Q$ given true state $S$ is $u_j(Q|S)$:
  - $u_j(C|G) = u_j(A|I) = 0$;
  - $u_j(C|I) = -q_j$ and $u_j(A|G) = -(1 - q_j)$, $q_j \in (0, 1)$ for all $j$;
  - $u_j(M|I) = -\mu_I \cdot q_j$ and $u_j(M|G) = -\mu_G \cdot (1 - q_j)$.

- Set $\mu_I$ ($\mu_G$) to be probability that future jury convicts given innocent (acquits given guilty).
Informative and strategic voting

- **Informative** voting:
  - $j$ votes to convict if signal $g$
  - $j$ votes to acquit if signal $i$

- **Strategic** voting:
  - Jurors play a Nash equilibrium in weakly undominated strategies
  - I.e. max expected utility *conditional on being a pivotal voter*
Informative equilibrium

Coughlan (2000): When exists informative voting equilibrium

- In any such equilibrium, $\mu_I = \mu_G \equiv \mu$
- Informative equilibrium if **Assumption A1:**

  \[
  \text{logit } q_j \in (\text{logit } \pi - \text{logit } p, \text{logit } \pi + \text{logit } p) \text{ for all } j
  \]

  - I.e. if preferences are not extreme compared to signals
  - Condition does not depend on $n$ or $d$
- Note: informative also equilibrium if jurors fully cursed, for any parameters (Eyster and Rabin, 2005)

Rest of this paper: Assume A1 holds; **informative voting equilibria**
Comparisons across $n$ and $d$

$p_*(p, n, d) \equiv \text{Probability of reaching the correct verdict, conditional on reaching a verdict}$

$p_M(p, n, d) \equiv \text{Probability of a hung jury and mistrial}$

In informative equilibria:

- $p_*(p, n, d|G) = p_*(p, n, d|I)$ and
- $p_M(p, n, d|G) = p_M(p, n, d|I)$
Partial ordering of voting rules

- A triple \((p, n, d)\) is **weakly better than** \((p, n', d')\) if
  \[ p_*(p, n, d) \geq p_*(p, n', d') \quad \text{and} \quad p_M(p, n, d) \leq p_M(p, n', d'). \]

- It is **strictly better** if \( p_*(p, n, d) > p_*(p, n', d') \) and
  \[ p_M(p, n, d) \leq p_M(p, n', d'). \]

- \((n, d)\) **weakly (strictly) dominates** \((n', d')\) if for all \( p \in (1/2, 1) \),
  \((p, n, d)\) is weakly (strictly) better than \((p, n', d')\).
Main result: The suboptimality of unanimity

Theorem

\((n + 2, 1)\) strictly dominates \((n, 0)\) for all \(n\).

Corollary

For all \(n\), every juror’s ex ante expected utility is strictly higher for \((n + 2, 1)\) than \((n, 0)\).
Main results

The suboptimality of unanimity

**Theorem**

\((n + 2, 1)\) strictly dominates \((n, 0)\) for all \(n\).

Intuition for why accuracy higher:

- Convict in \((n + 2, 1)\) if vote is \(n + 1\) to 1 or \(n + 2\) to 0
  - If \(n + 1\) to 1, same net convict signals as if \(n\) to 0, so likelihood of guilty also the same
  - If \(n + 2\) to 0, strictly more likely guilty than if \(n\) to 0

Intuition for why mistrials (weakly) lower:

- Given first \(n\) votes, more likely to reach verdict with two new votes:
  - More likely to go from \(n\) to 0 \(\rightarrow\) \(n\) to 2
  - Than to go from \(n - 1\) to 1 \(\rightarrow\) \(n + 1\) to 1

\(n \in [6, 12]\) Angry Men  Michael Thaler
Main results

No strict comparisons when either $n$ or $d$ is fixed

Forms of the previous results can only occur when both $n$ and $d$ are varied simultaneously:

Proposition

For any $p \in (1/2, 1)$, $(p, n, d)$ is not strictly better than $(p, n', d)$ or $(p, n, d')$. 
Cost for additional jurors

What if we explicitly penalize juries for their size?

- Suppose every hung jury leads to a retrial, and consider a cost function \( c(\# \text{ trials}, \# \text{ jurors}) = C_1 \#\text{trials} + C_2 \#\text{jurors} \).

**Proposition**

*For all \( C_1, C_2 > 0, n \geq 5, p \leq 1 - \frac{1}{1.77n} \), \((n + 2, 1) \) has lower expected cost than \((n, 0)\).*

- Tradeoff: more jurors per jury vs. larger \( p_M \) (so more juries in expectation)
- Second effect dominates if signals not too strong, juries not too small
Why model communication

- Loose estimates from motivation:
  - $p_M(p, 12, 0) = 6.2\%$
  - $p_*(p, 12, 0) = 87\%$

- No singular value of $p$ in this model can fit both estimates:
  - Low $p$: modest $p_*$ but unrealistically high $p_M$
  - High $p$: modest $p_M$ but unrealistically high $p_*$. 
Extension: Some communication

Idea: some juries communicate by revealing signals, others do not

- $\gamma$ fraction communicate, $1 - \gamma$ do not.
  - Assumption: $\gamma$ independent of $p$, $n$, $d$.
  - Retrial $\rightarrow$ same jury communication type

- **Assumption A2:**
  - $\logit q_j \in (\logit \pi - \logit p, \logit \pi)$ or
  - $\logit q_j \in (\logit \pi, \logit \pi + \logit p)$.

- Under A2: informative equilibrium exists for both jury types
Main results extend to communication model with informative equilibria:

**Proposition**

For all $\gamma$, $(n + 2, 1)$ strictly dominates $(n, 0)$ for all $n$.

**Corollary**

For all $n$ and $\gamma$, every juror’s ex ante expected utility is strictly higher for $(n + 2, 1)$ than $(n, 0)$.
Calibration and cost

Rough calibration of parameters:

- Match accuracy / mistrial rates: $p = .6615$, $\gamma = .9378$
- Trial length: 5 days, 4 hours (NCSC, 2007)
- Lost wages: $1,099$ per juror (BLS: OES, 2018)
- Trial cost: $81,958$ (Pittsburgh Post-Gazette, 1983; adjusted)

(Note: Data only for states where $n = 12$, $d = 0$)
Cost saved by using 1-dissenter juries

Compare expected cost of

- $(n, 0)$ jury and
- A mix between $(n + 2, 1)$ and $(n - 2, 1)$ juries
  - Mix so that average accuracy the same

Cost lower for 1-dissenter juries for policy-relevant sizes:

- $n = 6$ : $1,712$ saved per case
- $n = 9$ : $715$ saved per case
- $n = 12$ : $165$ saved per case
Important to endogenize jury sizes when comparing voting rules:

• Can simultaneously reduce hung jury rates, increase expected accuracy.

• Can decrease social cost with no loss to accuracy.

• Unanimity with $n$ jurors often dominated by a jury with $n + 2$ jurors that allows for one dissenter.

• Possible further work: juror heterogeneity, jury selection, non-Bayesian updating