

Strategic Reasoning in Jury Decisions

an experimental study

Can Çelebi · Stefan P. Penczynski

Vienna Center for Experimental Economics

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n
Maimonides (12th c.)	1000

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n
Maimonides (12th c.)	1000
Fortescue (15th c.)	20

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n
Maimonides (12th c.)	1000
Fortescue (15th c.)	20
Blackstone (18th c.)	10

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	
Fortescue (15th c.)	20	
Blackstone (18th c.)	10	

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	
Blackstone (18th c.)	10	

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	91%

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	91%

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	91%
	1	50%

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	91%
	1	50%
	1/10	9%

Convicting the Innocent

“better that n guilty escape than one innocent suffer”

	n	threshold of reasonable doubt
Maimonides (12th c.)	1000	99.9%
Fortescue (15th c.)	20	95%
Blackstone (18th c.)	10	91%
	1	50%
	1/10	9%

death penalty at least 4.1% of US death sentences are estimated to be wrongful (Gross et al., 2014)

Voting

basic setup

states $\theta \in \{G, I\}$, equally likely

Voting

basic setup

states $\theta \in \{G, I\}$, equally likely

signals each juror sees $s_i \in \{g, i\}$ with $\Pr(s_i = \theta \mid \theta) = p > \frac{1}{2}$
(informative)

Voting

basic setup

- states $\theta \in \{G, I\}$, equally likely
- signals each juror sees $s_i \in \{g, i\}$ with $\Pr(s_i = \theta \mid \theta) = p > \frac{1}{2}$
(informative)
- payoffs convicting innocent costs q , acquitting guilty costs $1 - q$
(same for all)

Voting

basic setup

- states $\theta \in \{G, I\}$, equally likely
- signals each juror sees $s_i \in \{g, i\}$ with $\Pr(s_i = \theta \mid \theta) = p > \frac{1}{2}$
(informative)
- payoffs convicting innocent costs q , acquitting guilty costs $1 - q$
(same for all)
- majority convict iff $\geq \lfloor n/2 \rfloor + 1$ jurors vote **guilty**

Voting

basic setup

- states $\theta \in \{G, I\}$, equally likely
- signals each juror sees $s_i \in \{g, i\}$ with $\Pr(s_i = \theta \mid \theta) = p > \frac{1}{2}$
(informative)
- payoffs convicting innocent costs q , acquitting guilty costs $1 - q$
(same for all)
- majority convict iff $\geq \lfloor n/2 \rfloor + 1$ jurors vote **guilty**
- unanimity convict iff **all** n jurors vote **guilty**

Convicting the Innocent

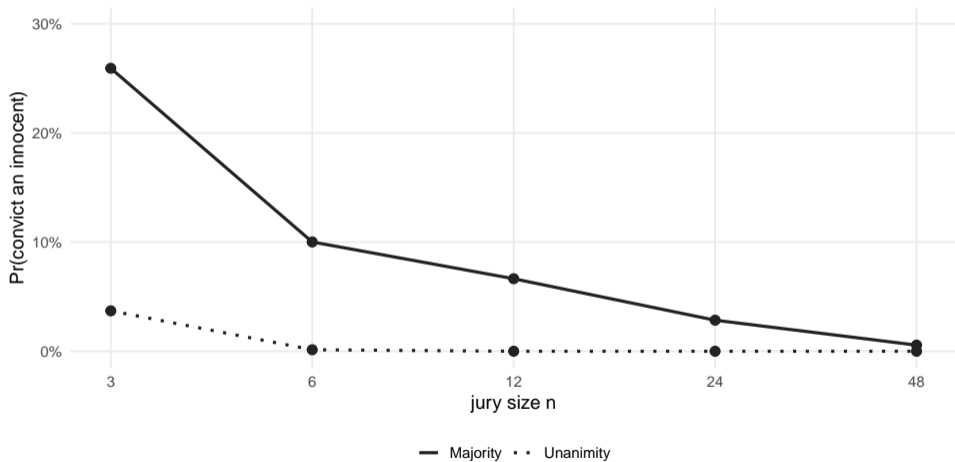
sincere voting · unanimity vs majority

Convicting the Innocent

sincere voting · unanimity vs majority · $p = 2/3$ · $q = 1/2$

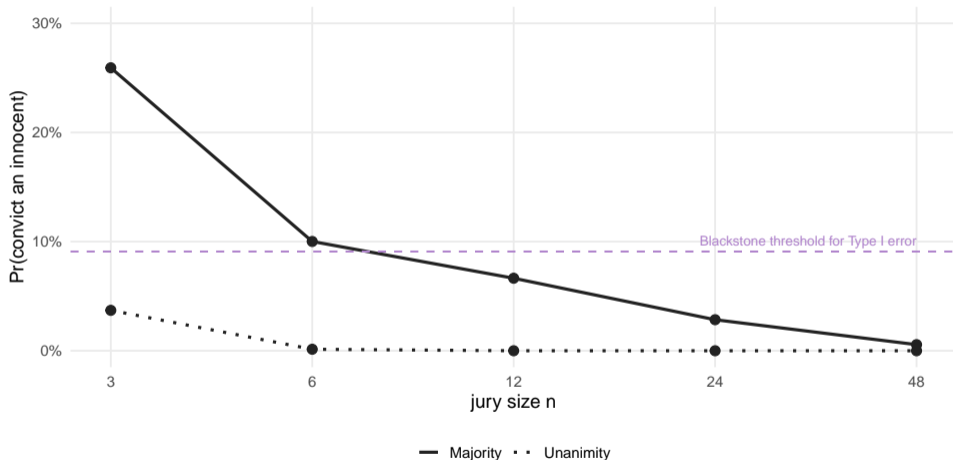
Convicting the Innocent

sincere voting · *unanimity vs majority* · $p = 2/3$ · $q = 1/2$



Convicting the Innocent

sincere voting · *unanimity vs majority* · $p = 2/3$ · $q = 1/2$



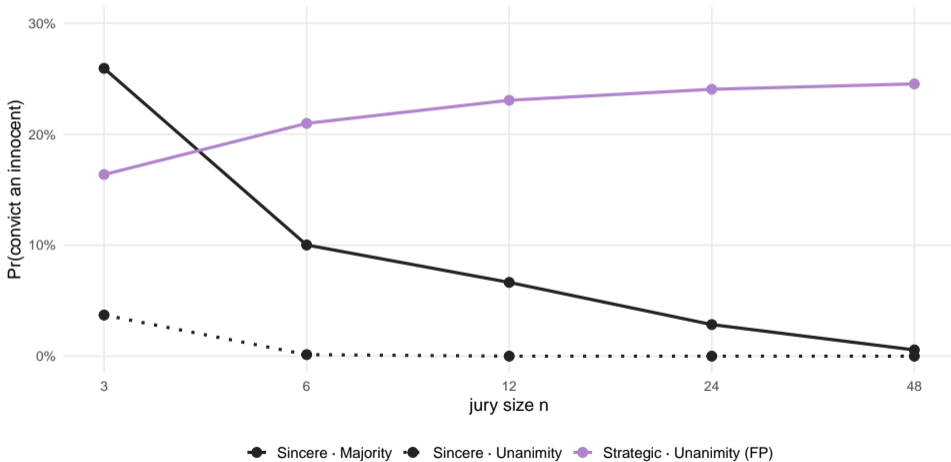
↗ Type I / Type II / Total error rates

Convicting the Innocent

Feddersen & Pesendorfer 1998 · *strategic voting*

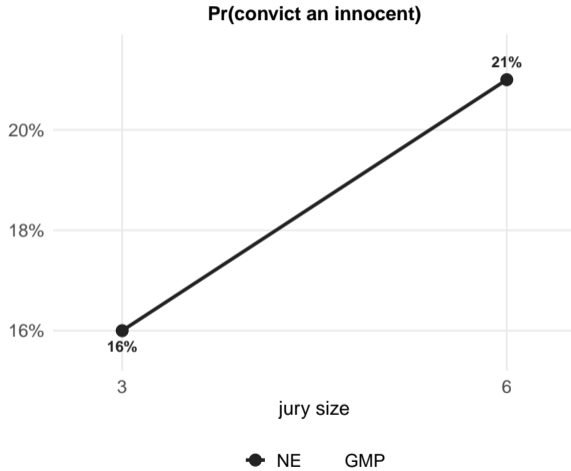
Convicting the Innocent

Feddersen & Pesendorfer 1998 · *strategic voting* · $p = 2/3$ · $q = 1/2$



Convicting the Innocent

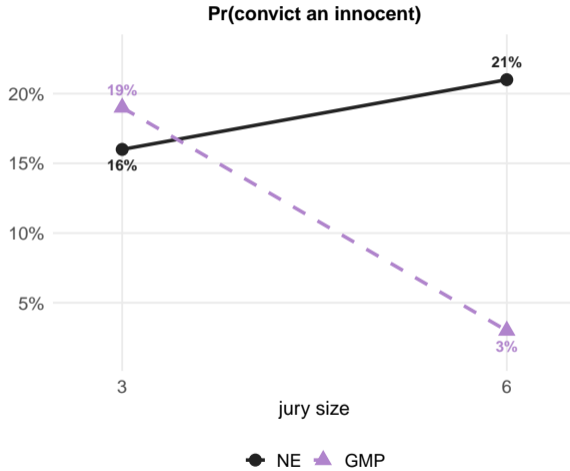
Guarnaschelli, McKelvey & Palfrey 2000 (GMP)



↗ Pr(vote G | innocent signal) ↗ error models

Convicting the Innocent

Guarnaschelli, McKelvey & Palfrey 2000 (GMP)



↗ Pr(vote G | innocent signal) ↗ error models

This Study

motivation discrepancy in the literature between theory and empirical data

This Study

motivation discrepancy in the literature between theory and empirical data

what we do (i) apply the level- k model to jury voting for the first time

(ii) elicit subjects' reasoning via a communication protocol and use the text to identify each subject's level of thinking

This Study

- motivation discrepancy in the literature between theory and empirical data
- what we do (i) apply the level- k model to jury voting for the first time
- (ii) elicit subjects' reasoning via a communication protocol and use the text to identify each subject's level of thinking
- conjecture as group size \uparrow , task complexity \uparrow
- \Rightarrow average reasoning level \downarrow
- \Rightarrow strategic voting \downarrow

Strategic voting

naive voter

prior $\Pr(G) = \Pr(I) = 1/2$

Strategic voting

naive voter

prior $\Pr(G) = \Pr(I) = 1/2$

signal $\Pr(g | G) = \Pr(i | I) = 2/3$

Strategic voting

naive voter

prior $\Pr(G) = \Pr(I) = 1/2$

signal $\Pr(g | G) = \Pr(i | I) = 2/3$

*assume the juror's threshold of reasonable doubt is 50% (if $\Pr(G | \text{evidence}) > .5$, votes **guilty**)*

Strategic voting

naive voter

prior $\Pr(G) = \Pr(I) = 1/2$

signal $\Pr(g | G) = \Pr(i | I) = 2/3$

*assume the juror's threshold of reasonable doubt is 50% (if $\Pr(G | \text{evidence}) > .5$, votes **guilty**)*

*assume the juror received an **innocent** signal*

Strategic voting

naive voter

prior $\Pr(G) = \Pr(I) = 1/2$

signal $\Pr(g | G) = \Pr(i | I) = 2/3$

*assume the juror's threshold of reasonable doubt is 50% (if $\Pr(G | \text{evidence}) > .5$, votes **guilty**)*

*assume the juror received an **innocent** signal*

$$\Pr(G | i) = \mathbf{33\%}$$

Strategic voting

$Pr(G \mid \text{pivotal})$, before your own signal

assume there are n jurors

Strategic voting

$Pr(G \mid \text{pivotal})$, before your own signal

assume there are n jurors

assume the other jurors vote sincerely

Strategic voting

$\Pr(G \mid \text{pivotal})$, before your own signal

assume there are n jurors

assume the other jurors vote sincerely

$\Pr(G \mid \text{pivotal})$

$n - 1$

majority

unanimity

Strategic voting

$\Pr(G \mid \text{pivotal})$, before your own signal

assume there are n jurors

assume the other jurors vote sincerely

	$\Pr(G \mid \text{pivotal})$	
$n - 1$	majority	unanimity
2	50%	80%

Strategic voting

$\Pr(G \mid \text{pivotal})$, before your own signal

assume there are n jurors

assume the other jurors vote sincerely

	$\Pr(G \mid \text{pivotal})$	
$n - 1$	majority	unanimity
2	50%	80%
6	50%	98%

Strategic voting

$Pr(G \mid \text{pivotal}, i)$

$n - 1$	majority	unanimity
2	33%	67%

Strategic voting

$Pr(G \mid \text{pivotal}, i)$

$n - 1$	majority	unanimity
2	33%	67%
6	33%	97%

Voting game

setup



blue urn



red urn



jurors

urn picked \sim state of the world (unknown to jurors)

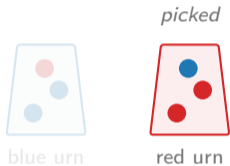
ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$

Voting game

setup



urn picked \sim state of the world (unknown to jurors)

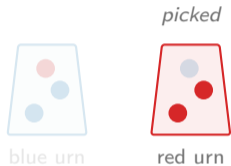
ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$

Voting game

setup



urn picked \sim state of the world (unknown to jurors)

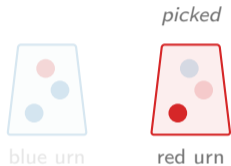
ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling

Voting game

setup



urn picked \sim state of the world (unknown to jurors)

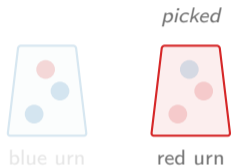
ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling

Voting game

setup



urn picked \sim state of the world (unknown to jurors)

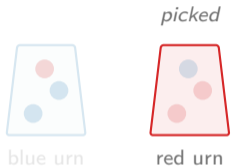
ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling

Voting game

setup



correct guess €2 · wrong guess 20¢

urn picked \sim state of the world (unknown to jurors)

ball drawn \sim private signal

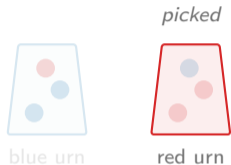
red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling ·

$q = 1/2$

Voting game

setup



urn picked \sim state of the world (unknown to jurors)

ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling ·

$q = 1/2$

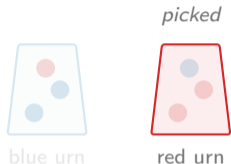


correct guess €2 · wrong guess 20¢

unanimity rule · no abstention

Voting game

setup



correct guess €2 · wrong guess 20¢
unanimity rule · no abstention

urn picked \sim state of the world (unknown to jurors)

ball drawn \sim private signal

red = guilty · blue = innocent

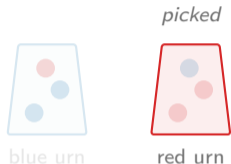
$p = 2/3$ · without replacement sampling ·

$q = 1/2$

teams 2 subjects play as a single juror

Voting game

setup



correct guess €2 · wrong guess 20¢
unanimity rule · no abstention

urn picked \sim state of the world (unknown to jurors)

ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling ·

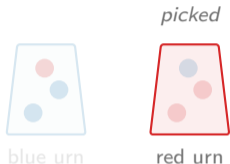
$q = 1/2$

teams 2 subjects play as a single juror

n 3 jurors \Rightarrow 6 subjects

Voting game

setup



correct guess €2 · wrong guess 20¢
unanimity rule · no abstention

urn picked \sim state of the world (unknown to jurors)

ball drawn \sim private signal

red = guilty · blue = innocent

$p = 2/3$ · without replacement sampling ·

$q = 1/2$

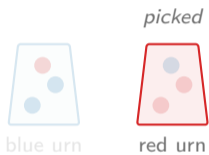
teams 2 subjects play as a single juror

n 3 jurors \Rightarrow 6 subjects

\nearrow communication protocol to elicit reasoning

Voting game

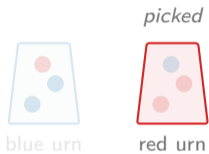
levels of reasoning



- **Level 0** random

Voting game

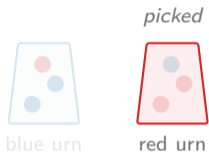
levels of reasoning



- **Level 0** random
"Turtles like red."

Voting game

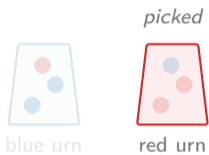
levels of reasoning



- **Level 0** random
"Turtles like red."
- **Level 1** sincere

Voting game

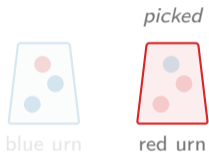
levels of reasoning



- **Level 0** random
"Turtles like red."
- **Level 1** sincere
"the urn is most likely the colour I drew"

Voting game

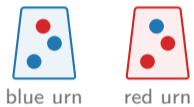
levels of reasoning



- **Level 0** random
"Turtles like red."
- **Level 1** sincere
"the urn is most likely the colour I drew"
- **Level 2** strategic

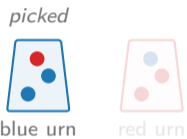
Voting game

without replacement, $n = 3$ · level-2 juror



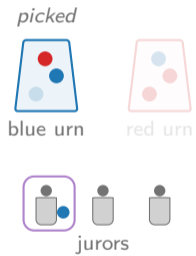
Voting game

without replacement, $n = 3$ · level-2 juror



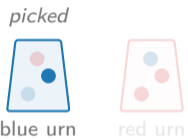
Voting game

without replacement, $n = 3$ · level-2 juror



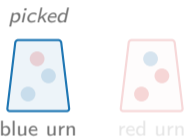
Voting game

without replacement, $n = 3$ · level-2 juror



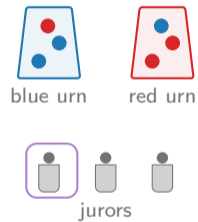
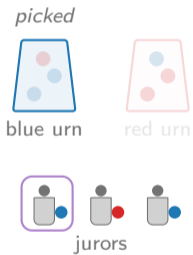
Voting game

without replacement, $n = 3$ · level-2 juror



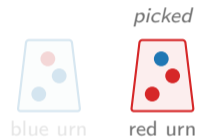
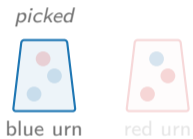
Voting game

without replacement, $n = 3$ · level-2 juror



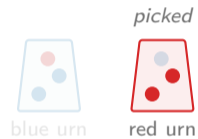
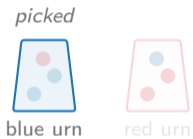
Voting game

without replacement, $n = 3$ · level-2 juror



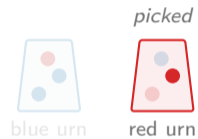
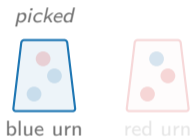
Voting game

without replacement, $n = 3$ · level-2 juror



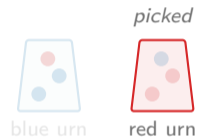
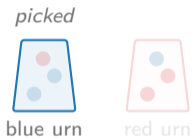
Voting game

without replacement, $n = 3$ · level-2 juror



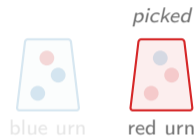
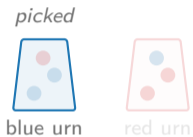
Voting game

without replacement, $n = 3$ · level-2 juror



Voting game

without replacement, $n = 3$ · level-2 juror



"I'd take red, because if anyone else takes blue it'll be blue anyway"

Voting game

with replacement, $n = 3$

Voting game

with replacement, $n = 3$



blue urn



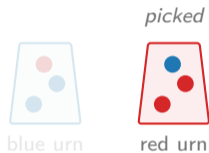
red urn



jurors

Voting game

with replacement, $n = 3$



cases

$\text{Pr}(\text{case} \mid \text{R})$

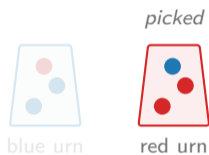


4/9



Voting game

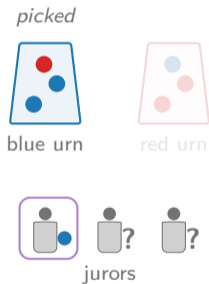
with replacement, $n = 3$



cases	$\text{Pr}(\text{case} \mid \mathbf{R})$
●●	$\frac{4}{9}$
●●	$\frac{4}{9}$
●●	$\frac{1}{9}$

Voting game

with replacement, $n = 3$



cases

$\Pr(\text{case} \mid \text{R})$

$\Pr(\text{case} \mid \text{B})$



4/9

1/9



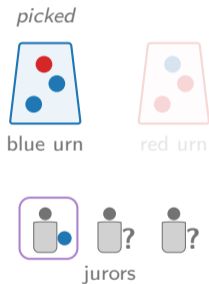
4/9



1/9

Voting game

with replacement, $n = 3$



cases

$\Pr(\text{case} \mid \text{R})$

$\Pr(\text{case} \mid \text{B})$



4/9

1/9



4/9

4/9

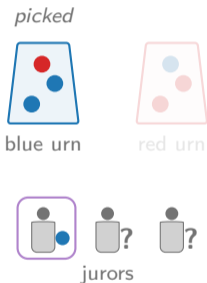


1/9

4/9

Voting game

with replacement, $n = 3$



cases	$\Pr(\text{case} \mid \text{R})$	$\Pr(\text{case} \mid \text{B})$
●●	4/9	1/9
●●	4/9	4/9
●●	1/9	4/9

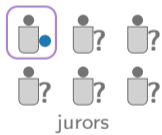
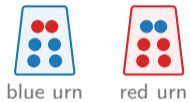
more cases to consider \Rightarrow more complex task

Voting game

with replacement, $n = 6$

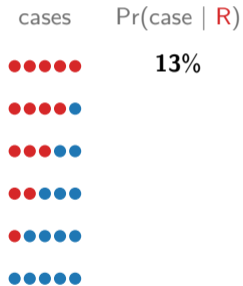
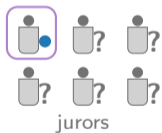
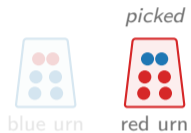
Voting game

with replacement, $n = 6$



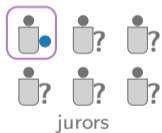
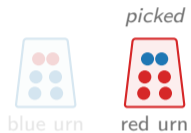
Voting game

with replacement, $n = 6$



Voting game

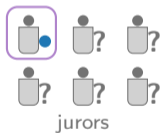
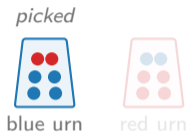
with replacement, $n = 6$



cases	Pr(case R)
●●●●●●	13%
●●●●●●	33%
●●●●●●	33%
●●●●●●	16%
●●●●●●	4%
●●●●●●	0.4%

Voting game

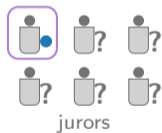
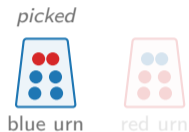
with replacement, $n = 6$



cases	Pr(case R)	Pr(case B)
●●●●●●	13%	0.4%
●●●●●●	33%	
●●●●●●	33%	
●●●●●●	16%	
●●●●●●	4%	
●●●●●●	0.4%	

Voting game

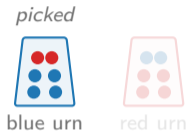
with replacement, $n = 6$



cases	Pr(case R)	Pr(case B)
●●●●●●	13%	0.4%
●●●●●●	33%	4%
●●●●●●	33%	16%
●●●●●●	16%	33%
●●●●●●	4%	33%
●●●●●●	0.4%	13%

Voting game

with replacement, $n = 6$

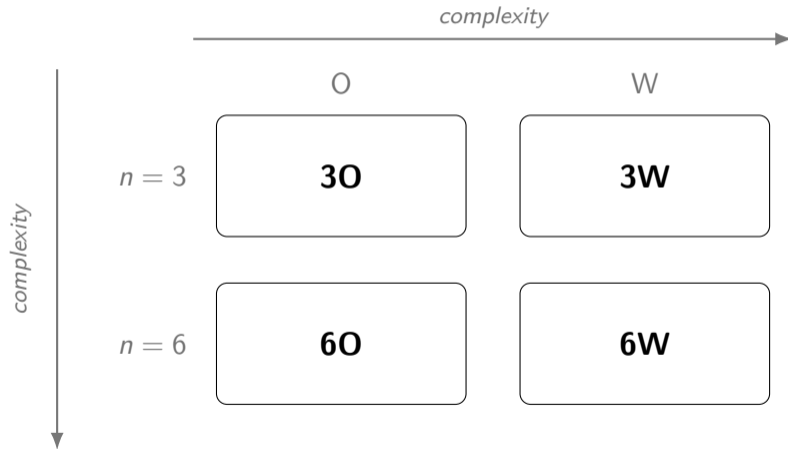


cases	Pr(case R)	Pr(case B)
●●●●●●	13%	0.4%
●●●●●●	33%	4%
●●●●●●	33%	16%
●●●●●●	16%	33%
●●●●●●	4%	33%
●●●●●●	0.4%	13%

more cases to consider \Rightarrow *more complex task*

Treatments

$2 \times 2: \{3, 6\} \times \{O, W\}$



Level- k predictions

$\sigma_k(s)$ probability of voting r after signal s , by level and treatment

	O		W	
	$n = 3$	$n = 6$	$n = 3$	$n = 6$
s	b	b	b	b
$k = 1$	0	0	0	0
$k = 2$	1	1	1	1

Level- k predictions

$\sigma_k(s)$ probability of voting r after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
s	r	b	r	b	r	b	r	b
$k = 1$	1	0	1	0	1	0	1	0
$k = 2$	1	1	1	1	1	1	1	1

Level- k predictions

$\sigma_k(s)$ probability of voting r after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
s	r	b	r	b	r	b	r	b
$k = 1$	1	0	1	0	1	0	1	0
$k = 2$	1	1	1	1	1	1	1	1

identification of levels via the elicited reasoning texts of the subjects

Hypotheses

theory

H1: level-1 players vote sincerely; level-2 players vote strategically.

Hypotheses

theory

H1: level-1 players vote sincerely; level-2 players vote strategically.

H2: Given a level of reasoning, the behavior will not differ by treatment.

Hypotheses

complexity

H3: \uparrow group size \Rightarrow \downarrow average depth of reasoning

Hypotheses

complexity

H3: \uparrow group size \Rightarrow \downarrow average depth of reasoning

H4: $O \rightarrow W \Rightarrow \downarrow$ average depth of reasoning

Experimental design

- **within-subject** design

↗ number of observations by treatment

Experimental design

- **within-subject** design
- 4 parts per experiment, one for each treatment

↗ number of observations by treatment

Experimental design

- **within-subject** design
- 4 parts per experiment, one for each treatment
- 2 rounds per part; jury decision and payoff revealed after each round

↗ number of observations by treatment

Procedures

- 6 sessions · 96 subjects · mLab, University of Mannheim

↗ classification rate by treatment

Procedures

- 6 sessions · 96 subjects · mLab, University of Mannheim
- ↗ communication protocol is used to elicit subjects' reasoning

↗ classification rate by treatment

Procedures

- 6 sessions · 96 subjects · mLab, University of Mannheim
- ↗ communication protocol is used to elicit subjects' reasoning
- two RAs classify each message into a level of reasoning

↗ classification rate by treatment

Procedures

- 6 sessions · 96 subjects · mLab, University of Mannheim
- ↗ communication protocol is used to elicit subjects' reasoning
- two RAs classify each message into a level of reasoning

message sent		no message
classified	unclassified	
525	4	215

out of 744 observations

↗ classification rate by treatment

Procedures

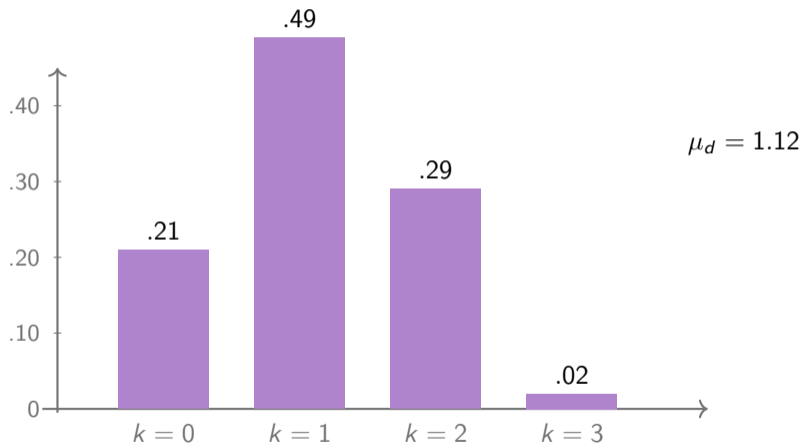
- 6 sessions · 96 subjects · mLab, University of Mannheim
- ↗ communication protocol is used to elicit subjects' reasoning
- two RAs classify each message into a level of reasoning

	message sent		no message
	classified	unclassified	
	525	4	215
matched		493	

out of 744 observations

↗ classification rate by treatment

Aggregate level- k distribution



$\sigma_k(s)$

probability of voting *red* after signal s , by level and treatment

	O		W	
	$n = 3$	$n = 6$	$n = 3$	$n = 6$
	$\sigma(b)$	$\sigma(b)$	$\sigma(b)$	$\sigma(b)$
$k = 1$	0	0	0	0
$k = 2$	1	1	1	1

$\sigma_k(s)$

probability of voting *red* after signal s , by level and treatment

	O		W	
	$n = 3$	$n = 6$	$n = 3$	$n = 6$
	$\sigma(b)$	$\sigma(b)$	$\sigma(b)$	$\sigma(b)$
$k = 1$.03	.09	.09	.04
$k = 2$.92	.91	.92	1.00

$\sigma_k(s)$

probability of voting *red* after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$
$k = 1$.03	1	.09	1	.09	1	.04	1
$k = 2$.92	1	.91	1	.92	1	1.00	1

$\sigma_k(s)$

probability of voting *red* after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$
$k = 1$.03	.97	.09	.97	.09	.97	.04	.97
$k = 2$.92	1.00	.91	1.00	.92	1.00	1.00	1.00

 full table

$\sigma_k(s)$

probability of voting *red* after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$
$k = 1$.03	.97	.09	.97	.09	.97	.04	.97
$k = 2$.92	1.00	.91	1.00	.92	1.00	1.00	1.00

- result voting behaviour aligns with level- k theory predictions

 full table

$$\sigma_k(s)$$

probability of voting *red* after signal s , by level and treatment

	O				W			
	$n = 3$		$n = 6$		$n = 3$		$n = 6$	
	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$
$k = 1$.03	.97	.09	.97	.09	.97	.04	.97
$k = 2$.92	1.00	.91	1.00	.92	1.00	1.00	1.00

- result voting behaviour aligns with level- k theory predictions
- result for a given level of reasoning, voting behaviour does not significantly differ across treatments

 full table

$\sigma(b)$ — Nash prediction vs observed

		O	W
	NE	ÇP	NE ÇP
$n = 3$.50		.32
$n = 6$.71		.66

$\sigma(b)$ — Nash prediction vs observed

		O	W	
	NE	ÇP	NE	ÇP
$n = 3$.50	.46	.32	.32
$n = 6$.71	.55	.66	.36

$\sigma(b)$ — Nash prediction vs observed

		O		W
	NE	ÇP	NE	ÇP
$n = 3$.50	.46	.32	.32
$n = 6$.71	.55	.66	.36

- under either sampling method, strategic voting is not significantly different between $n = 3$ and $n = 6$

$\sigma(b)$ — Nash prediction vs observed

		O	W	
	NE	ÇP	NE	ÇP
$n = 3$.50	.46	.32	.32
$n = 6$.71	.55	.66	.36

- under either sampling method, strategic voting is not significantly different between $n = 3$ and $n = 6$
- under either jury size, strategic voting in W is significantly lower than in O

Average level of strategic thinking

μ_d — average level of strategic thinking

	O	W
$n = 3$	1.30	1.16
$n = 6$	1.14	0.85

Average level of strategic thinking

μ_d — average level of strategic thinking

	O	W
$n = 3$	1.30	1.16
$n = 6$	1.14	0.85

$\tilde{\sigma}_2(b)$ — effective strategic voters

	O	W
$n = 3$.35	.24
$n = 6$.34	.18

Overestimating strategic voting

		O		W
	$\sigma(b)$	$\tilde{\sigma}_2(b)$	$\sigma(b)$	$\tilde{\sigma}_2(b)$
$n = 3$.46	.35	.32	.24
$n = 6$.55	.34	.36	.18

Change in level distribution

group size

	$n = 3$	$n = 6$
$k = 0$.14	.28
$k = 1$.52	.45
$k = 2$.32	.27
μ_d	1.23	1.00

↑ **level 0** · ↓ *level 1* · ↓ *level 2*

result larger juries have a significantly lower average level of reasoning

Change in level distribution

sampling

	O	W
$k = 0$.19	.23
$k = 1$.43	.55
$k = 2$.36	.22
μ_d	1.22	1.01

↑ *level 0* · ↑ *level 1* · ↓ *level 2*

result sampling without replacement results in a significantly lower average level of reasoning

$\Pr(\mathbf{R} \mid \mathbf{B})$ — convict an innocent

		O		W		
	NE	ÇP	NE	ÇP	GMP	
$n = 3$.25	.19	.16	.16	.19	
$n = 6$.25	.08	.21	.02	.03	

↗ $\Pr(\mathbf{B} \mid \mathbf{R})$ — acquit a guilty

Conclusion

motivation discrepancy in the literature between theory and empirical data

↗ total error rate by treatment

Conclusion

motivation discrepancy in the literature between theory and empirical data

what we did (i) apply the level- k model to jury voting for the first time

(ii) elicit subjects' reasoning via a communication protocol and use the text to identify each subject's level of thinking

↗ total error rate by treatment

Conclusion

motivation discrepancy in the literature between theory and empirical data

what we did (i) apply the level- k model to jury voting for the first time

(ii) elicit subjects' reasoning via a communication protocol and use the text to identify each subject's level of thinking

result contrary to NE predictions, as group size \uparrow strategic voting \downarrow ·
driven by the rise in task complexity

 total error rate by treatment

Thank you

Appendix

Appendix — index background & methods (1/2)

↗ Type I / Type II / Total

sincere voting, $p = 2/3$

↗ Pr(correct verdict)

majority $\rightarrow 1$, unanimity $\rightarrow 1/2$

↗ Pr(guilty | convicted)

FP 1998, eq. 1

↗ Pr(innocent | acquitted)

the dual

↗ Number of observations by treatment

744 obs across the 2×2 design

↗ Communication protocol

how subject reasoning is elicited

↗ Classified messages by treatment

ratio classified per treatment

↗ models & results \rightarrow

Appendix — index models & results (2/2)

↗ Pr(vote G | innocent signal)

↗ NNM & QRE error models

↗ QRE equation

↗ Cognitive Hierarchy predictions

↗ CH convergence · $n = 3, O$

↗ CH convergence · $n = 3, W$

↗ Model fit by treatment

↗ Error rates by model

↗ $\sigma_k(s)$ full table

↗ $d(k)$ by treatment

↗ $d(k) | b$ by treatment

↗ $d(k) | r$ by treatment

GMP 2000 σ plot

GMP 2000 parameter estimates

logit best-response

τ -thresholds for L2 / L3

R_k trajectories as τ grows

R_k trajectories as τ grows

NNM / QRE / CH / Lk-text, $-LL$

Pr(R | B), Pr(B | R), total per model

all four levels, observed

aggregate level distribution

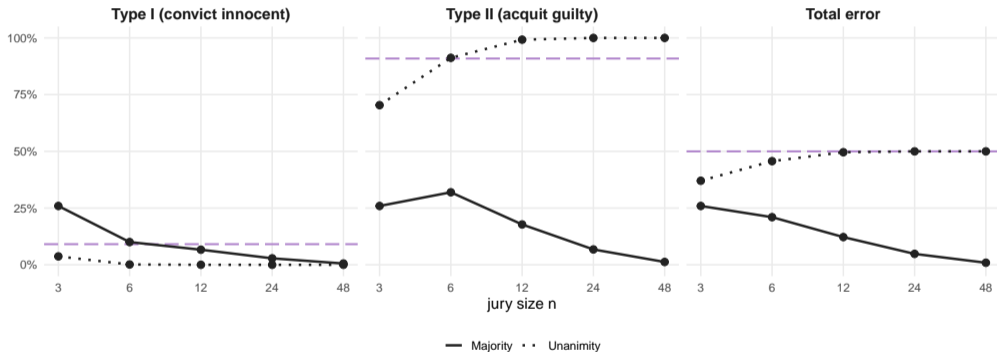
level distribution given blue signal

level distribution given red signal

← ↗ background & methods

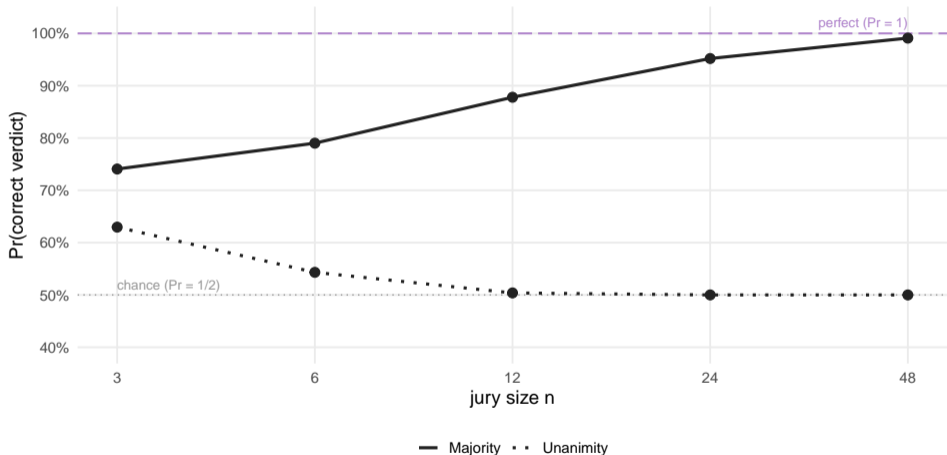
Convicting the Innocent — full picture

sincere voting, $p = 2/3$ · Type I, Type II, Total



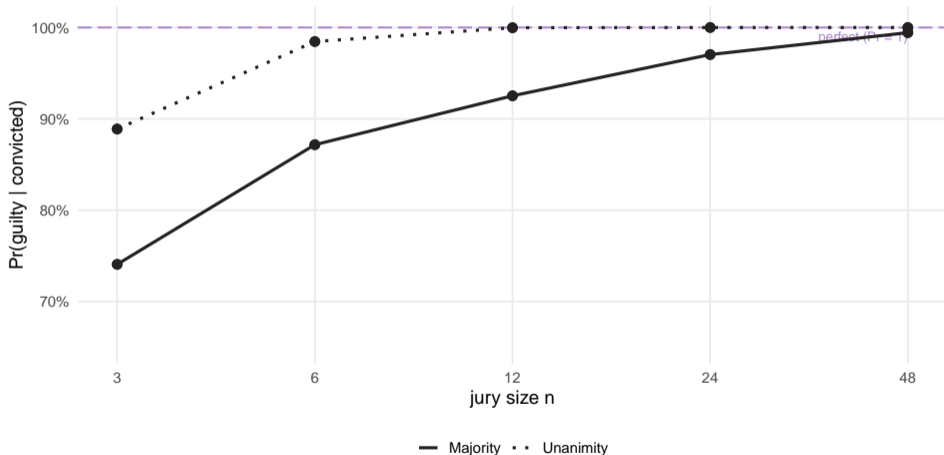
Pr(correct verdict)

sincere voting, $p = 2/3$ · majority $\rightarrow 1$, unanimity $\rightarrow 1/2$



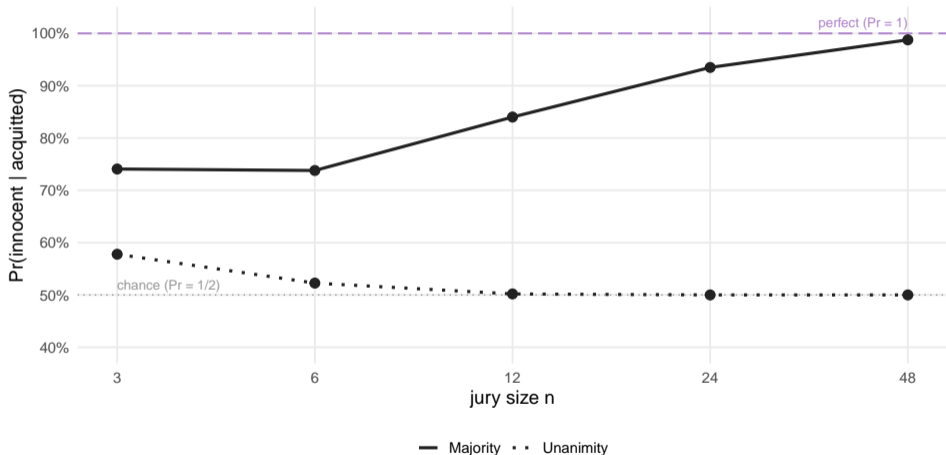
Pr(guilty | convicted)

Feddersen & Pesendorfer 1998, eq. 1 · both rules $\rightarrow 1$



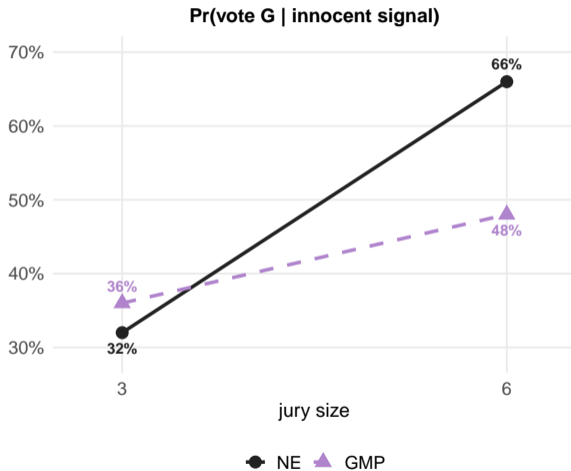
Pr(innocent | acquitted)

the dual: majority $\rightarrow 1$, unanimity $\rightarrow 1/2$ (acquittals uninformative)



Pr(vote guilty | innocent signal)

Guarnaschelli, McKelvey & Palfrey 2000 (GMP)



Error models on top of Nash

GMP 2000 — NNM and QRE

		$n = 3$	$n = 6$
	description		
NNM (γ)	best-respond with prob. γ , else votes randomly	0.90	0.76
QRE (λ)	logit response: action probability $\propto \exp(\lambda \cdot \text{expected utility})$	9.2	12.0

QRE response equation

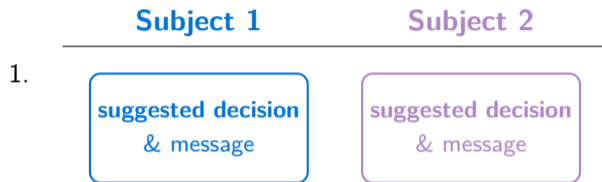
logit best-response with sensitivity λ

Each juror picks an action with probability proportional to $\exp(\lambda \cdot \text{expected utility})$.

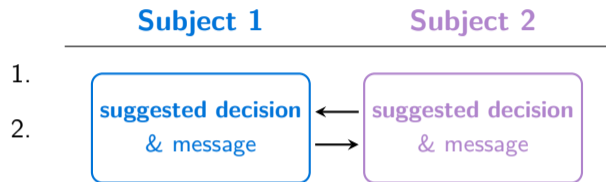
$$\Pr(\text{vote } G \mid s) = \frac{\exp(\lambda \cdot u_G(s))}{\exp(\lambda \cdot u_G(s)) + \exp(\lambda \cdot u_I(s))}$$

$\lambda = 0$: uniform random over actions · $\lambda \rightarrow \infty$: pure best-response (Nash)

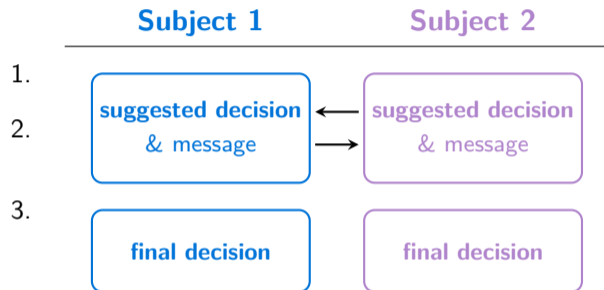
Communication protocol



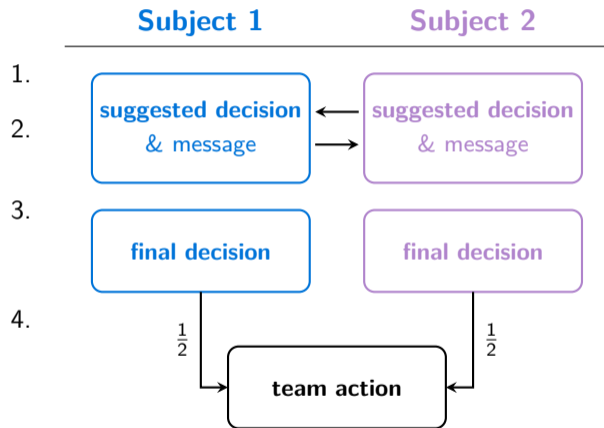
Communication protocol



Communication protocol



Communication protocol



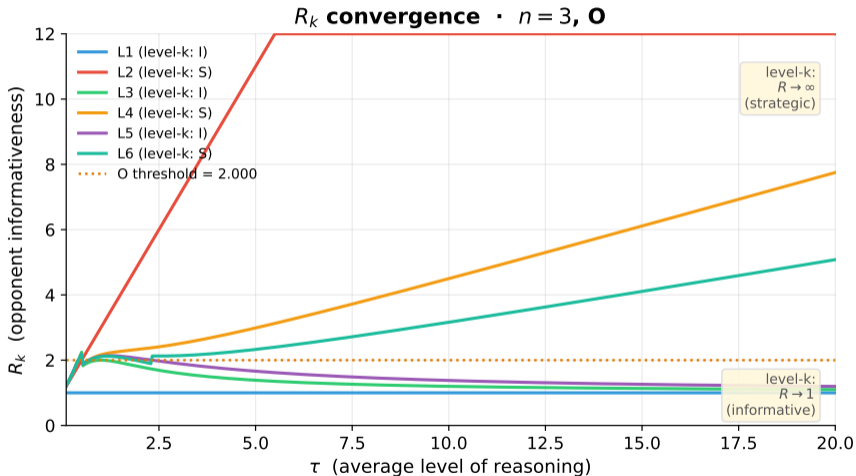
↗ index ↔ voting game ↔ procedures

Number of observations by treatment

	O	W	Σ
$n = 3$	192	192	384
$n = 6$	192	168	360
Σ	384	360	744

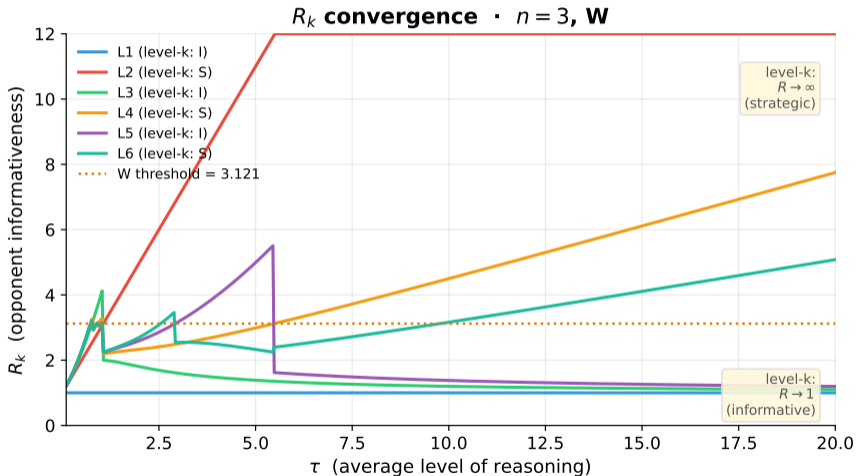
CH convergence to level- k · $n = 3, \mathbf{O}$

R_k trajectories as τ grows · even levels \rightarrow strategic, odd levels \rightarrow informative



CH convergence to level- k · $n = 3$, W

R_k trajectories as τ grows · same alternating cycle, higher threshold under W



Model fit by treatment

	NNM $\hat{\gamma}$	QRE $\hat{\lambda}$	CH $\hat{\tau} = \mu_d^{\text{CH}}$	Lk-text μ_d^{text}
3O	.84 64.68	22.6 64.47	1.84 64.42	1.30 64.54
6O	.82 59.52	40.3 62.85	1.64 58.01	1.14 58.96
3W	.94 46.05	24.3 46.93	1.81 52.13	1.16 47.05
6W	.74 59.23	36.4 63.61	1.13 54.61	0.85 53.36
pooled	.84 231.74	36.4 259.43	1.76 228.02	1.12 227.05

Error rates by model

$Pr(R|B)$ — convict an innocent implied from each model's $\sigma(r), \sigma(b)$

	NE	ÇP	NNM	QRE	CH	Lk-text
30	.25	.19	.23	.21	.19	.17
60	.25	.08	.17	.17	.09	.04
3W	.16	.16	.16	.21	.15	.15
6W	.08	.02	.12	.02	.04	.02

Error rates by model

$Pr(B|R)$ — *acquit a guilty* implied from each model's $\sigma(r), \sigma(b)$

	NE	ÇP	NNM	QRE	CH	Lk-text
30	.50	.61	.58	.60	.61	.63
60	.50	.78	.69	.61	.77	.86
3W	.57	.57	.57	.52	.58	.61
6W	.92	.86	.77	.95	.86	.91

Error rates by model

$total = Pr(R|B) + Pr(B|R)$ implied from each model's $\sigma(r), \sigma(b)$

	NE	ÇP	NNM	QRE	CH	Lk-text
30	.75	.80	.81	.81	.81	.80
60	.75	.86	.86	.78	.87	.90
3W	.73	.73	.73	.73	.73	.76
6W	1.00	.88	.88	.98	.91	.93

Cognitive Hierarchy — thresholds

Level 1 informative in every treatment.

Level 2 strategic iff $\tau \leq 1.06$); empirical $\tau \in [1, 1.5]$ clears all four \Rightarrow strategic everywhere.

Level 3 informative for $n = 3$; strategic for $n = 6$ over a wide τ range.

Classified messages by treatment

ratio of observations with classified messages

	O	W	Σ
$n = 3$.69	.65	.67
$n = 6$.65	.66	.66
Σ	.67	.66	.66

stable across treatments

$\Pr(\mathbf{B} \mid \mathbf{R})$ — acquit a guilty

		O		W		
	NE	ÇP	NE	ÇP	GMP	
$n = 3$.50	.61	.54	.57	.53	
$n = 6$.50	.78	.52	.86	.73	

$\sigma_k(s)$ — full table

observed, all four levels

	O				W			
	<i>n</i> = 3		<i>n</i> = 6		<i>n</i> = 3		<i>n</i> = 6	
	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$	$\sigma(b)$	$\sigma(r)$
<i>k</i> = 0	.75	.64	.71	.73	.40	.86	.38	.75
<i>k</i> = 1	.03	.97	.09	.97	.09	.97	.04	.97
<i>k</i> = 2	.92	1.00	.91	1.00	.92	1.00	1.00	1.00
<i>k</i> = 3	.00	—	.00	—	.00	—	—	—

k = 3 has few or no observations in several cells

Level- k distribution by treatment

aggregate

		O		W	
$d(k)$	$n = 3$	$n = 6$	$n = 3$	$n = 6$	
$k = 0$.14	.23	.14	.32	
$k = 1$.46	.40	.58	.51	
$k = 2$.36	.36	.26	.17	
$k = 3$.04	.01	.02	.00	
μ_d	1.30	1.14	1.16	0.85	

$\nearrow d(k) | b$ $\nearrow d(k) | r$ \nearrow index

Level- k distribution by treatment

given a *blue* signal

		O		W	
$d(k)$	$n = 3$	$n = 6$	$n = 3$	$n = 6$	
$k = 0$.12	.24	.17	.32	
$k = 1$.43	.36	.58	.46	
$k = 2$.38	.38	.22	.22	
$k = 3$.07	.02	.03	—	
μ_d	1.41	1.17	1.12	0.90	

↗ aggregate ↗ $d(k) | r$ ↗ index

Level- k distribution by treatment

given a *red* signal

		O		W	
$d(k)$	$n = 3$	$n = 6$	$n = 3$	$n = 6$	
$k = 0$.17	.22	.11	.33	
$k = 1$.48	.43	.59	.54	
$k = 2$.34	.34	.30	.13	
$k = 3$	—	—	—	—	
μ_d	1.17	1.12	1.19	0.80	

↗ aggregate ↗ $d(k) | b$ ↗ index

Total error rate $\cdot \Pr(\mathbf{R} | \mathbf{B}) + \Pr(\mathbf{B} | \mathbf{R})$

		O		W		
		NE	ÇP	NE	ÇP	GMP
$n = 3$.75	.80	.70	.73	.72	
$n = 6$.75	.86	.73	.88	.76	

Total error rate · $\Pr(\mathbf{R} | \mathbf{B}) + \Pr(\mathbf{B} | \mathbf{R})$

	O		W		
	NE	ÇP	NE	ÇP	GMP
$n = 3$.75	.80	.70	.73	.72
$n = 6$.75	.86	.73	.88	.76

- observed totals are close to NE · the jury is not catastrophically off

Total error rate · $\Pr(\mathbf{R} | \mathbf{B}) + \Pr(\mathbf{B} | \mathbf{R})$

		O		W		
	NE	ÇP	NE	ÇP	GMP	
$n = 3$.75	.80	.70	.73	.72	
$n = 6$.75	.86	.73	.88	.76	

- observed totals are close to NE · the jury is not catastrophically off
- a jury of all-L2 strategists \Rightarrow everyone votes red \Rightarrow uninformative

Total error rate · $\Pr(\mathbf{R} | \mathbf{B}) + \Pr(\mathbf{B} | \mathbf{R})$

	O		W		
	NE	ÇP	NE	ÇP	GMP
$n = 3$.75	.80	.70	.73	.72
$n = 6$.75	.86	.73	.88	.76

- observed totals are close to NE · the jury is not catastrophically off
- a jury of all-L2 strategists \Rightarrow everyone votes red \Rightarrow uninformative
- *plurality of reasoning types* is what keeps the jury informative

Total error rate · $\Pr(\mathbf{R} | \mathbf{B}) + \Pr(\mathbf{B} | \mathbf{R})$

	O		W		
	NE	ÇP	NE	ÇP	GMP
$n = 3$.75	.80	.70	.73	.72
$n = 6$.75	.86	.73	.88	.76

- observed totals are close to NE · the jury is not catastrophically off
- a jury of all-L2 strategists \Rightarrow everyone votes red \Rightarrow uninformative
- *plurality of reasoning types* is what keeps the jury informative
- Oregon 1934: white-supremacist legislatures cut unanimity by branding jurors uneducated or unfit

Conjecture — chain of reasoning

us level- k and Cognitive Hierarchy models
heterogeneous beliefs · different levels of thinking

conjecture

larger jury \sim \uparrow task complexity



\downarrow average reasoning level (Alaoui & Penta 2016)



\downarrow strategic voting & \downarrow Pr(convict innocent)