Rules and Commitment in Communication

Guillaume Frechette Alessandro Lizzeri Jacopo Perego

September 2018

Overview

We study strategic information transmission with **(partial) commitment**.

Recently, a lot of interest in the implications of commitment in communication:

- Design incentives via information — Kamenica and Gentzkow (2011); Bergemann and Morris (2016, 2018); Mathevet, Perego, and Taneva (2017)
- Political campaigns — Alonso and Camara (2016); Chan, Gupta, Li and Wang. (2017)
- Rating systems — Duffie, Dworczak, and Zhu (2017)
- Stress tests for the banking system — Inostroza and Pavan (2017)

Overview

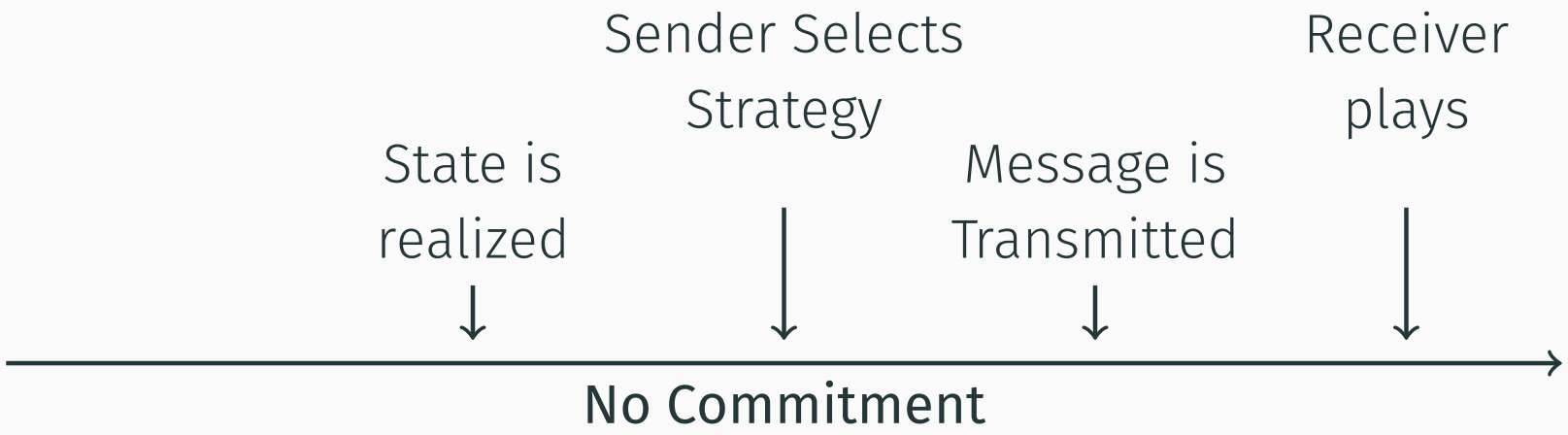
Despite all this attention, the baseline model mostly lacks empirical corroboration:

- Information can be notoriously hard to “measure” in the field.
- Identifying variations in commitment can also be quite problematic.

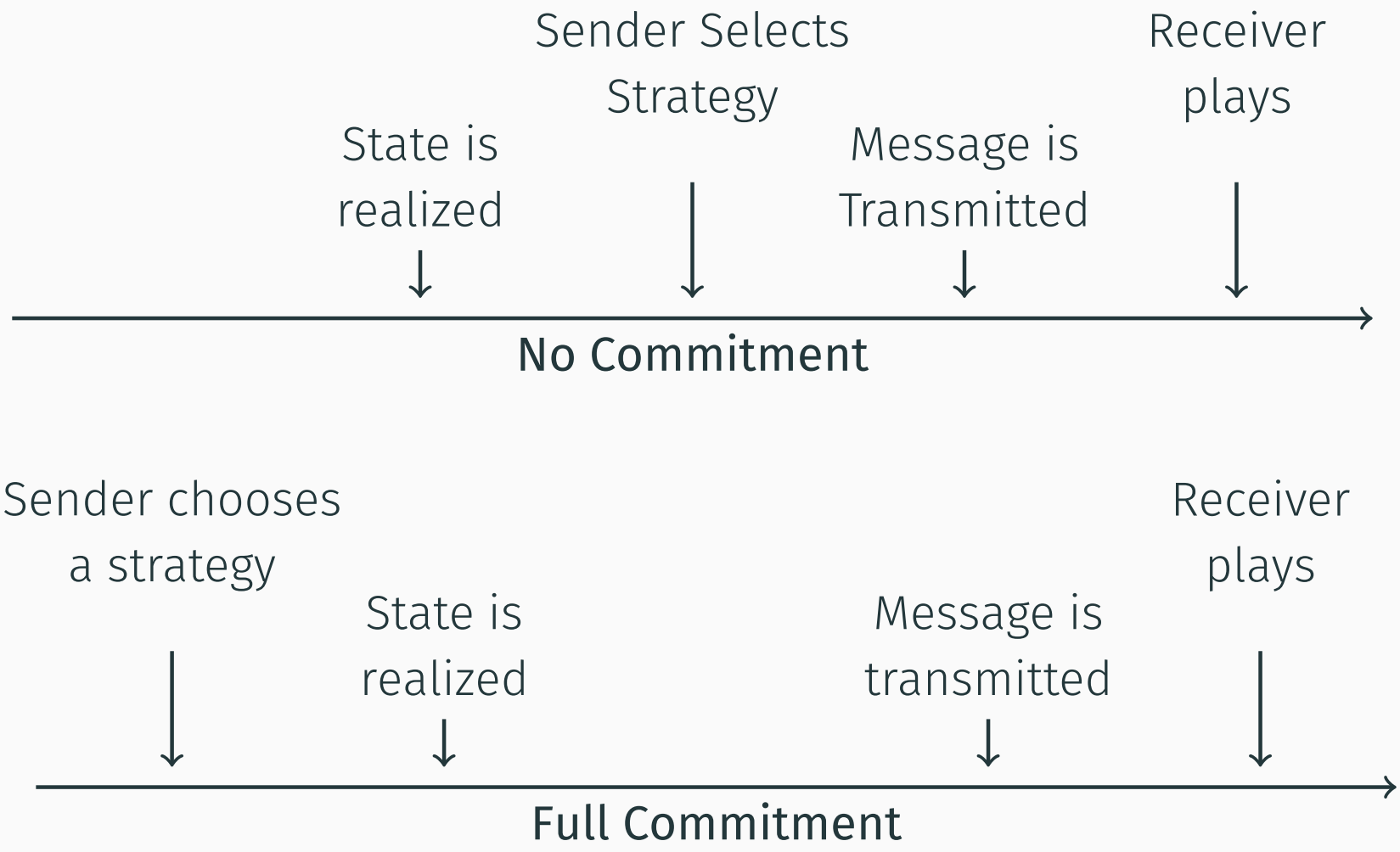
This paper:

Introduce a **novel framework** that allows for a comprehensive experimental assessment of the role of **commitment** in communication.

What is Commitment



What is Commitment



Special Cases

- **Cheap Talk**. Crawford and Sobel (1982)
  - No commitment and unverifiable information.
- **Disclosure**. Grossman (1981), Milgrom (1981), Okuno-Fujiwara et al (1990)
  - No commitment and verifiable Information.
- **Bayesian Persuasion**. Kamenica and Gentzkow (2011)
  - Full commitment and Unverifiable information.

Variations around a common basic structure, different predictions.

Literature in Lab and Field

- **Cheap talk experiments**: Dickhaut, McCabe, and Mukherji (1995); Blume, De Jong, Kim, and Sprinkle (1998); Cai and Wang (2006); Sanchez-Pages and Vorsatz (2007); Wang, Spezio, Camerer (2010)

- **Disclosure experiments**: Forsythe, Isaac, and Palfrey (1989); King and Wallin (1991); Dickhaut, Ledyard, Mukherji, and Sapra (2003); Forsythe, Lundholm, and Rietz (1999); Benndorf, K?bler, and Normann (2015); Hagenbach and Perez-Richet (2015); Jin, Luca, and Martin (2016)

- **Disclosure field:** Mathios (2000); Jin and Leslie (2003); Dranove and Jin (2010)

What We Do

We introduce a theoretical framework nesting the existing models of communication under the same umbrella.

- This framework generates **comparative statics** that we use to explicitly test the role of commitment in communication.
- Moreover, it gives us a broader perspective on the communication literature.
    1. A critical assessment of *over-communication* in cheap talk.
    2. A novel policy role for **rules** governing communication.

Main Results

- Data match **qualitative predictions** of the model providing validation of the strategic rational behind the role of commitment commitment in communication.
- There are also interesting **quantitative deviations**.
- Most notably, we find that **rules matter**, in ways that are entirely unpredicted by the theory. (policy implications)

benchmark treatment

Setting the Stage

We start with a **simple** implementation of a persuasion game.

- Goal is to characterize a number of **basic facts** about senders and receivers' behavior.

Moreover, this benchmark treatment allows us to

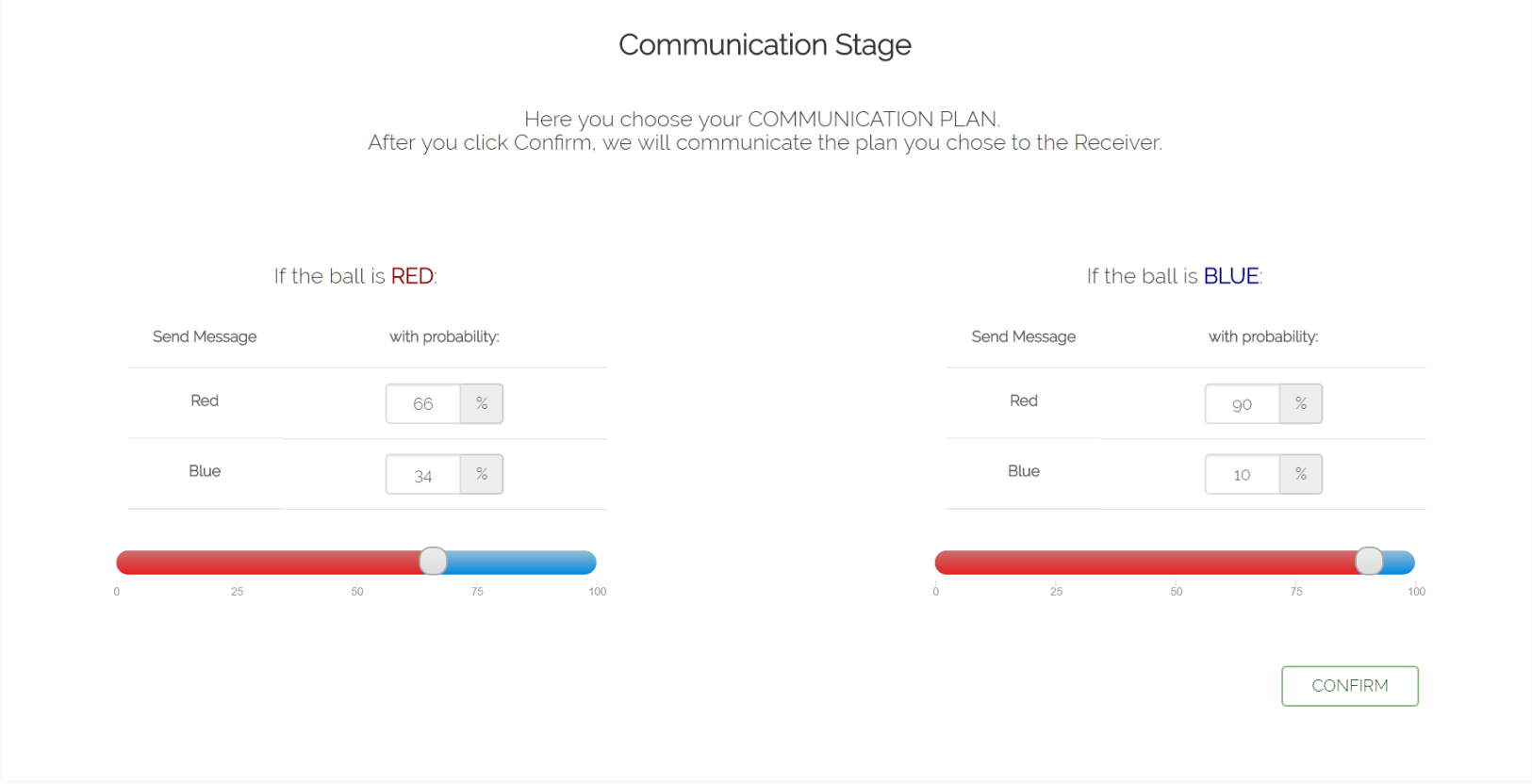
- Introduce main ingredients of the design
- Familiarize with the main strategic tensions

The Base Game

A Sender-Receiver binary guessing game.

- Binary payoff-state $\theta \in \{B, R\}$, prior belief $\mu_0(R) = \frac{1}{3}$.
- Sender sends messages (b or r) to Receiver as a function of the state.
- Preferences are **misaligned**:
  - Receiver's payoff is 1 if she guesses the state correctly.
  - Sender's payoff is 1 if Receiver guesses red.

Sender communicates **under commitment**.

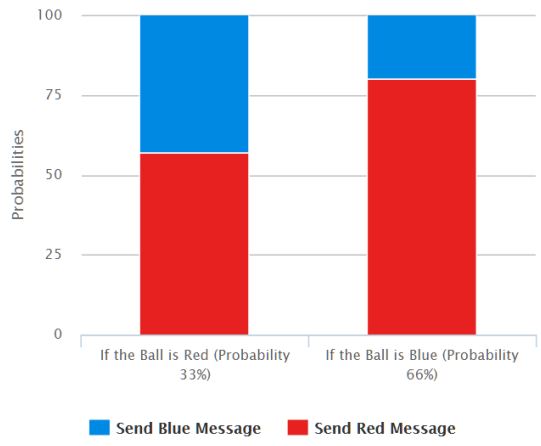


Receiver responds by **guessing the state**.

Guessing Stage

The message you will receive will come from the plan below:

COMMUNICATION PLAN:



Choose your **GUESSING PLAN**:

If I Receive Message...	...my guess will be:	
The Ball is Red	RED	BLUE
The Ball is Blue	RED	BLUE

Main Strategic Tension

- Sender wants to persuade Receiver to guess red.
- Absent communication, Receiver should always guess $blue$.

Main Strategic Tension

- Sender wants to persuade Receiver to guess red.
- Absent communication, Receiver should always guess $blue$.
- With communication, the equilibrium predicts that

		message	
		r	*b*
State	R	100%	0
	B	50%	50%

- Sender exploits commitment to **strategically lie** when state is B, but just enough to maintain credibility.
- Commitment assumption is **key**.

Potential Challenges

Simple in theory. In practice, it is more subtle.

- Sender's task is non-trivial: she needs to optimally design an information policy.
- **Bayesian paradigm**: Receiver needs to be Bayesian, Sender needs to be Bayesian and believe Receiver is Bayesian, etc.
- Understanding **commitment power**: how to use it, how to react to it.

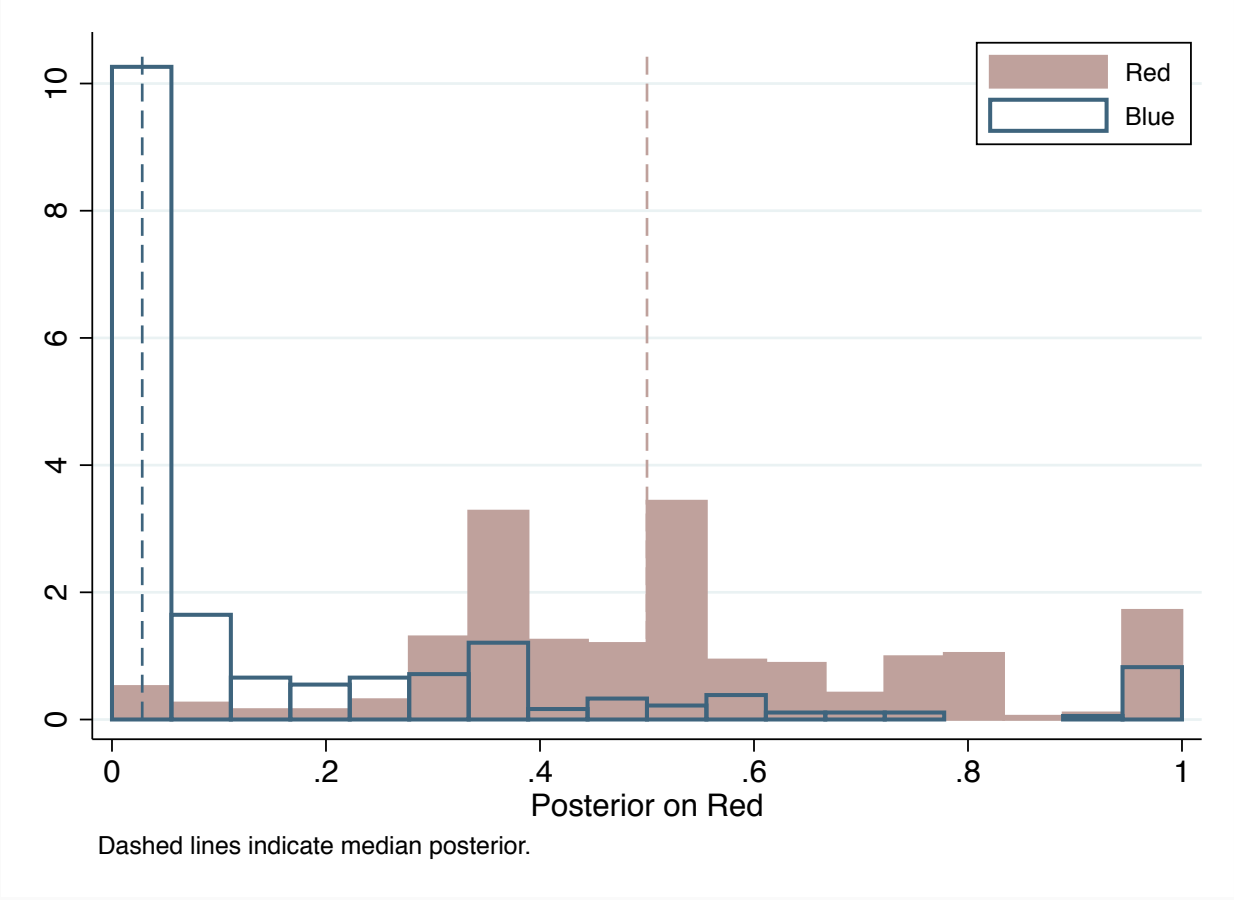
preliminary results

Results: Senders

Fact 1.

Senders induce posteriors that on average are in line with theoretical predictions.

Moreover, they communicate by employing a natural language.



Results: Senders

Fact 2. Most senders do engage in communication strategies that are somewhat sophisticated.

We use k-means clustering algorithm to create **behavioral types**.

Results: Senders

Fact 2. Most senders do engage in communication strategies that are somewhat sophisticated.

We use k-means clustering algorithm to create **behavioral types**.

This categorization is quite clean: 88% of data fall in three meaningful types:

- Under-communication – 24%
- Equilibrium-like – 35%
- Over-communication – 29%

Results: Senders

Fact 2. Most senders do engage in communication strategies that are somewhat sophisticated.

We use k-means clustering algorithm to create **behavioral types**.

This categorization is quite clean: 88% of data fall in three meaningful types:

- Under-communication – 24%
- Equilibrium-like – 35%
- Over-communication – 29%

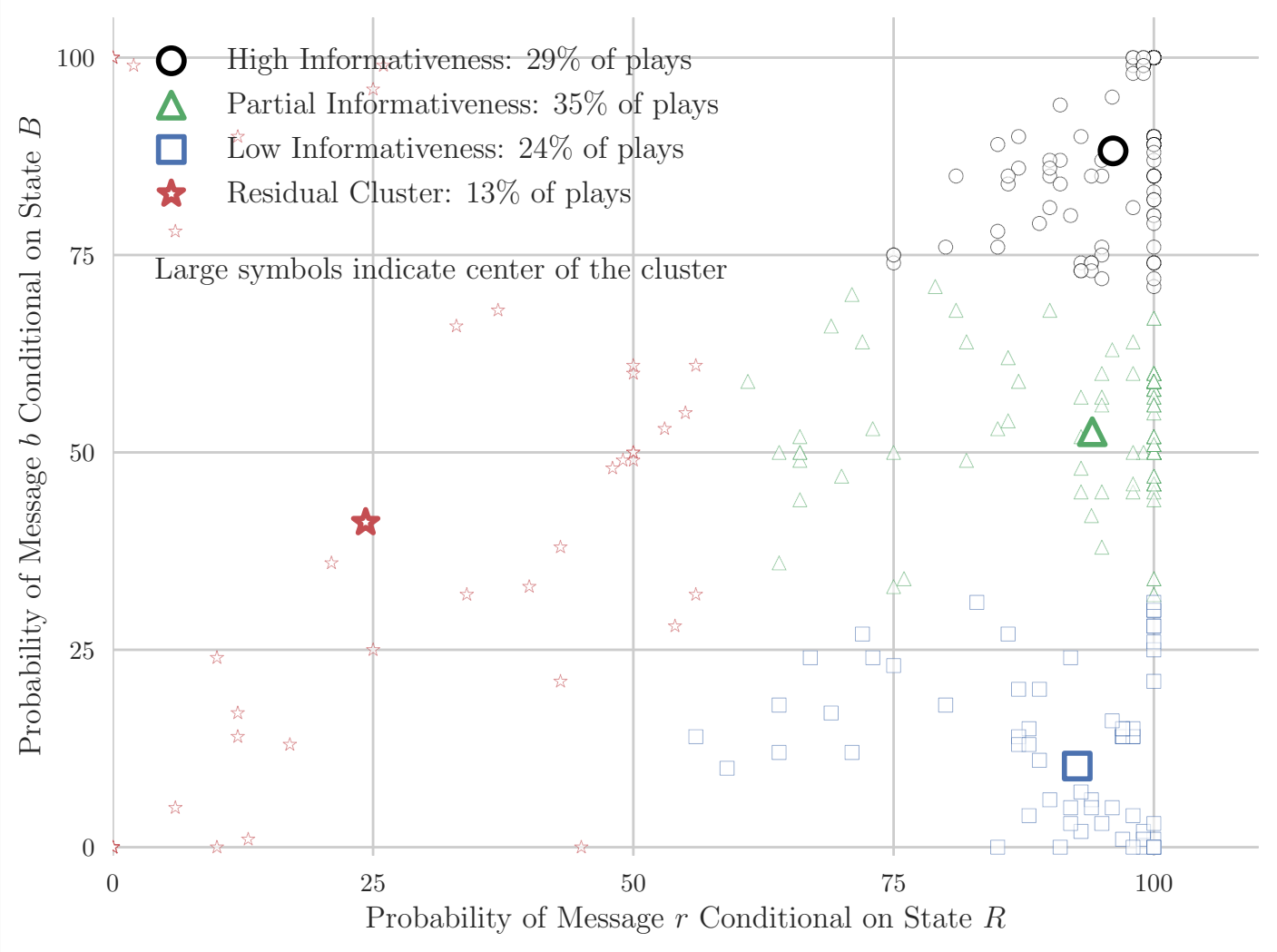
Fact 3. Do so persistently.

Median sender plays a strategy in the same cluster nine times out of ten.

Results: Senders

benchmark treatment

Figure 1: Sender's Strategy Grouped in Clusters



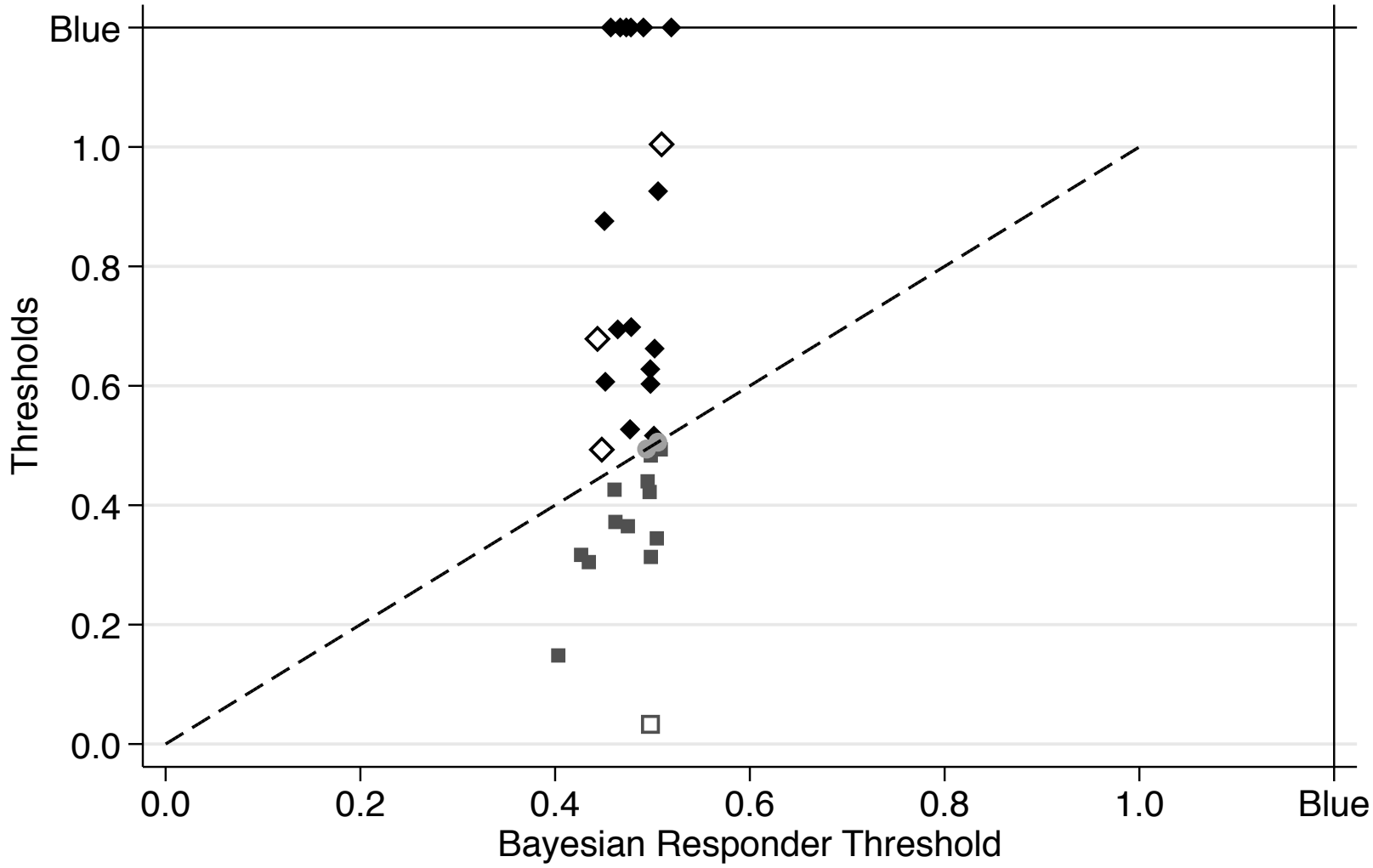
Results: Receivers

Fact 4.
Receivers' behavior:

- Bayesian paradigm does a poor job at predicting receivers behavior.
- Yet, they do react to information, and their behavior can be explained by threshold strategies.
- Large heterogeneity: they differ in the amount of information they require to be successfully persuaded.

Results: Receivers

benchmark treatment



Diamonds for subjects harder to convince than a Bayesian, square for subjects easier to convince than a Bayesian, circle otherwis
Hollow marker for subjects whose threshold is accurate in less than 80% of cases.
Blue indicates a subject who's choices are best explained by the rule always guess blue.
Values are jittered slightly to make multiple overlapping thresholds distinguishable.

Results: Takeaways

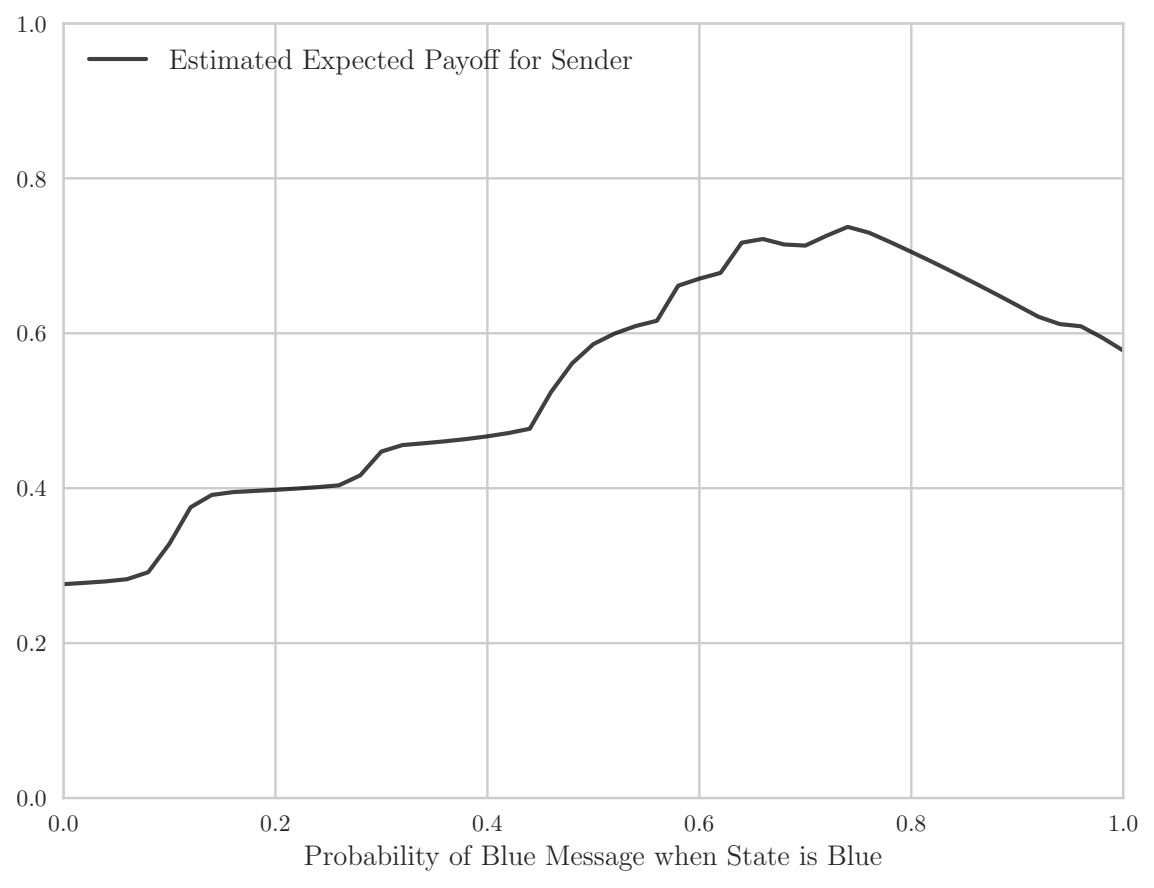
Two **qualitative** takeaways:

1. Despite all these deviations, what about the main qualitative prediction of the model, **strategic lying**?

Results: Takeaways

Two **qualitative** takeaways:

1. Despite all these deviations, what about the main qualitative prediction of the model, **strategic lying**?



Results: Takeaways

Two **qualitative** takeaways:

2. A critical assessment of existing experimental evidence on communication.

   A central finding in the literature on experimental cheap talk is *over-communication*.

   "Players tend to be truthful as a sender and credulous as a receiver." (Blume et al., 2017)

Results: Takeaways

Two **qualitative** takeaways:

2. A critical assessment of existing experimental evidence on communication.

   A central finding in the literature on experimental cheap talk is *over-communication*.

   "Players tend to be truthful as a sender and credulous as a receiver." (Blume et al., 2017)

   Some of our findings contrast this conclusion.
   - Increasing commitment.
   - Behavior in revision stage.
   - Analysis in U100S: communication types.
   - Comparative statics U vs V.

general framework

Commitment in Communication

Two main limitations in the previous analysis

- It does not quite address the *defining* assumption in persuasion: the role commitment.
- We did not provide an actual "test" for how commitment affects subjects behavior.

Simply, we cannot address these more substantive questions in the context our benchmark model.

A Richer Model

To do so, we need to enrich our model and we do so in **two directions**:

1. Allow for **partial** commitment.
2. Introduce **verifiable** information (what we call "**rules**")

A Richer Model

To do so, we need to enrich our model and we do so in **two directions**:

1. Allow for **partial** commitment.
2. Introduce **verifiable** information (what we call "**rules**")

The value of this richer model is twofold.

- The interaction between commitment and rules generates sharp **comparative statics** that we use to test persuasion.
- **Umbrella framework**. A broader perspective on communication games.

1. What is Partial Commitment?

1. **Commitment Stage**
   - **Sender** chooses an information structure.

1. What is Partial Commitment?

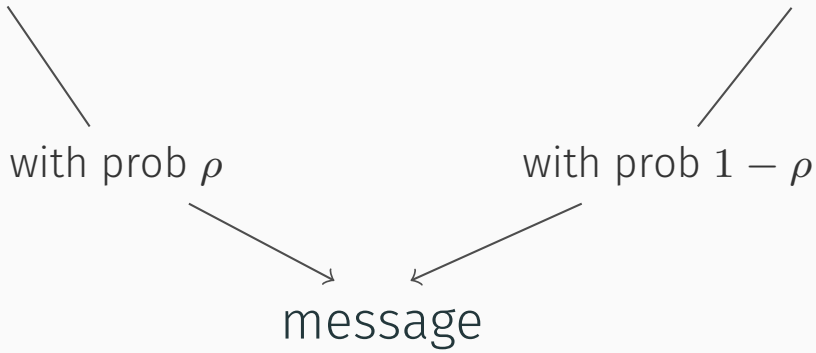
1. Commitment Stage
   - **Sender** chooses an information structure.

2. Revision Stage
   - **Sender** learns the state.
   - She can revise her message.

1. What is Partial Commitment?

1. **Commitment Stage**
   - **Sender** chooses an information structure.

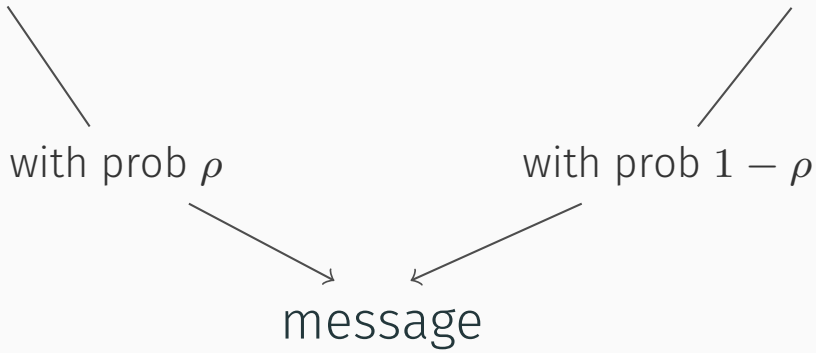
2. **Revision Stage**
   - **Sender** learns the state.
   - She can revise her message.



1. What is Partial Commitment?

1. **Commitment Stage**
   - **Sender** chooses an information structure.

2. **Revision Stage**
   - **Sender** learns the state.
   - She can revise her message.



3. **Guessing Stage**
   - **Receiver** observes message and guesses the state.

2. What are Rules?

Enrich the message space: r, b and n.

Then, we consider two extreme situations:

- **Unverifiable** information. As in *cheap talk* models

  Sender can send whichever message she wants.

- **Verifiable** information. As in models of *disclosure*

  If state is R, Sender can't send message b message.

  If state is B, Sender can't send message r message.

Interacting Rules and Commitment

Theorem

Under *unverifiable* information, increasing commitment **increases** informativeness.

Under *verifiable* information, increasing commitment **decreases** informativeness.

Theorem

Under *unverifiable* information, increasing commitment **increases** informativeness.

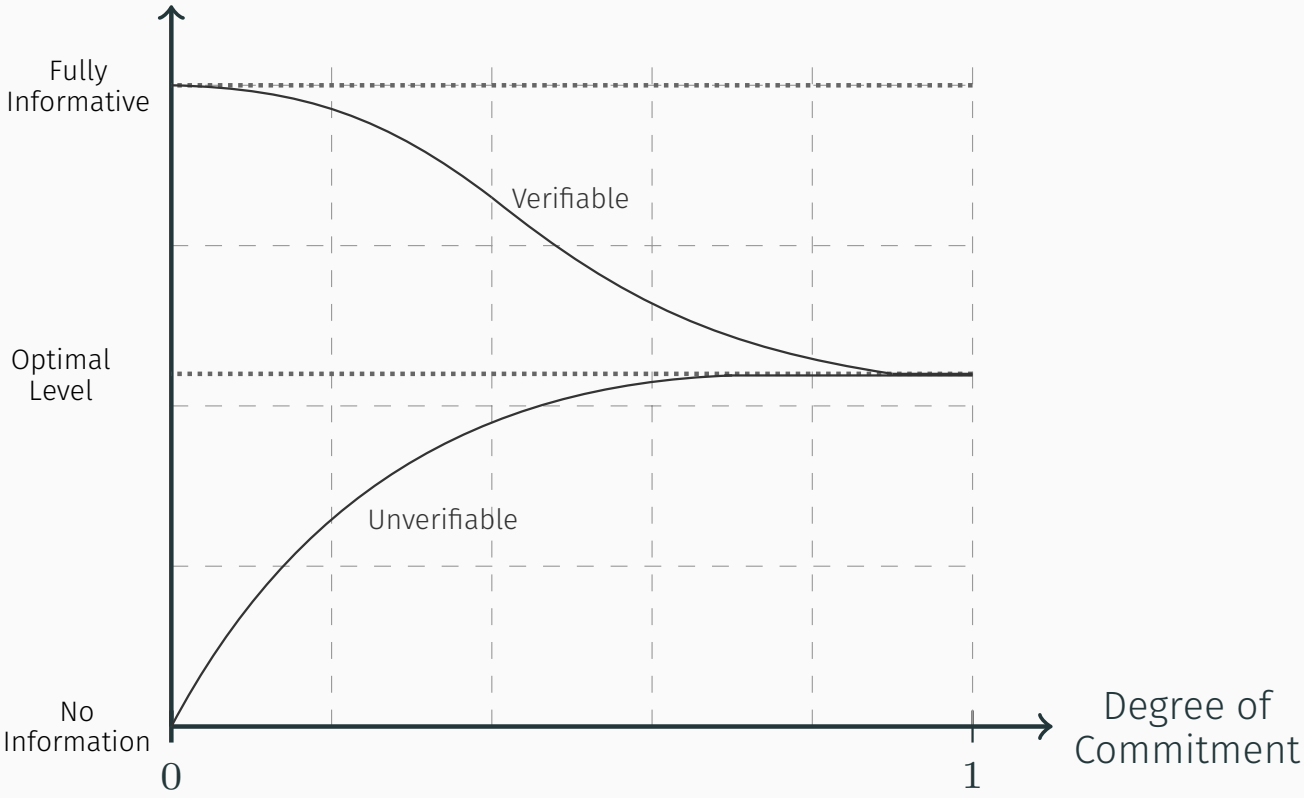
Under *verifiable* information, increasing commitment **decreases** informativeness.

Remark

At $\rho = 1$ (full commitment), equilibrium informativeness is rule-independent.

Therefore, rules are predicted to have **no welfare effect** under full commitment.

Interacting Rules and Commitment



Main Strategic Tension (Revisited)

Sender's **equilibrium behavior** in two extreme cases:

U100

State	messages: *r*	*b*	*n*
R	100%	0	0
B	50%	50%	0

V100

State	messages: *r*	*b*	*n*
R	0	0	100%
B	0	50%	50%

Intuition:

- **U100**. Lie as much as you can, but preserve incentives.
- **V100**. Conceal the good news: "No news, good news."

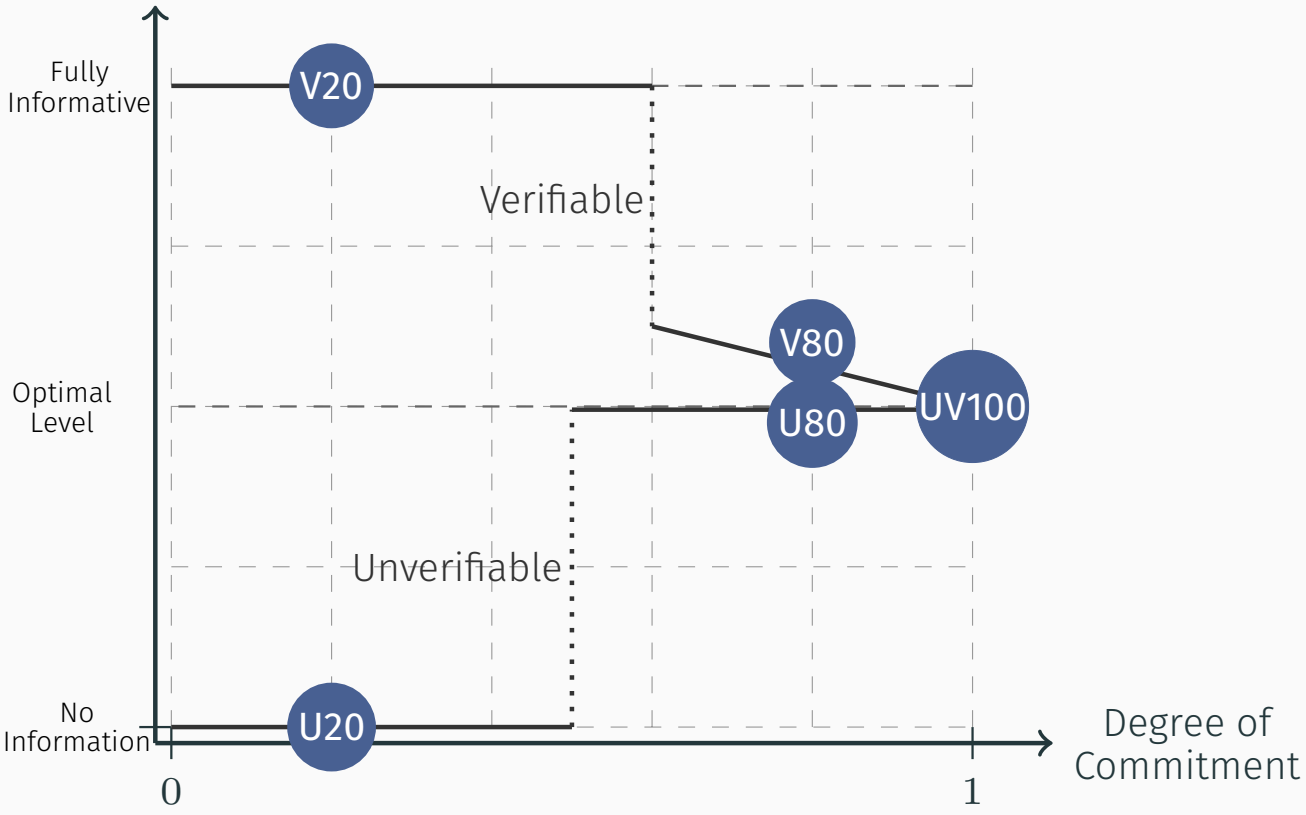
Treatments

Our Treatments (2x3 factorial design):

Rules: Verifiable (V) *vs* Unverifiable (U).
Commitment: $\rho \in \{20, 80, 100\}$.

	Commitment		
Rules	V20	V80	V100
	U20	U80	U100

Treatments


Fully
Informative
Optimal
Level
No
Information
V20
Verifiable
V80
U80
UV100
Unverifiable
U20
0
1
Degree of
Commitment

Equilibrium Predictions

	Sender								Receiver	
	Commitment				*Revision*				*Guessing*	
Treat.	Ball	Message			Ball	Message			Mes.	Guess
		red	blue	no		red	blue	no		
V20	R	1		0	R	1		0	red	*red*
	B		x	$1-x$	B		x	$1-x$	blue	*blue*
									no	*blue*
V80	R	0		1	R	1		0	red	*red*
	B		$\frac{3}{4}$	$\frac{1}{4}$	B		0	1	blue	*blue*
									no	*red*
V100	R	0		1					red	*red*
	B		$\frac{1}{2}$	$\frac{1}{2}$					blue	*blue*
									no	*red*
U20	R	x	y	$1-x-y$	R	1	0	0	red	*blue*
	B	x	y	$1-x-y$	B	1	0	0	blue	*blue*
									no	*blue*
U80	R	1	0	0	R	1	0	0	red	*red*
	B	$\frac{3}{8}$	$\frac{5}{8}$	0	B	1	0	0	blue	*blue*
									no	*blue*
U100	R	1	0	0					red	*red*
	B	$\frac{1}{2}$	$\frac{1}{2}$	0					blue	*blue*
									no	*blue*

Experimental Details

Implementation:

- Two unpaid practice rounds.
- 25 periods in **fixed roles** (last 10 for data analysis).
- Random rematching between periods.

General Information:

- Six treatments, four sessions per treatment.
- 384 subjects (≈ 16 per session; between 12 and 24).
- Average earnings: \$24 (including \$10 show up fee).
- Average duration: 100 minutes.

main results

Testing Commitment: Senders

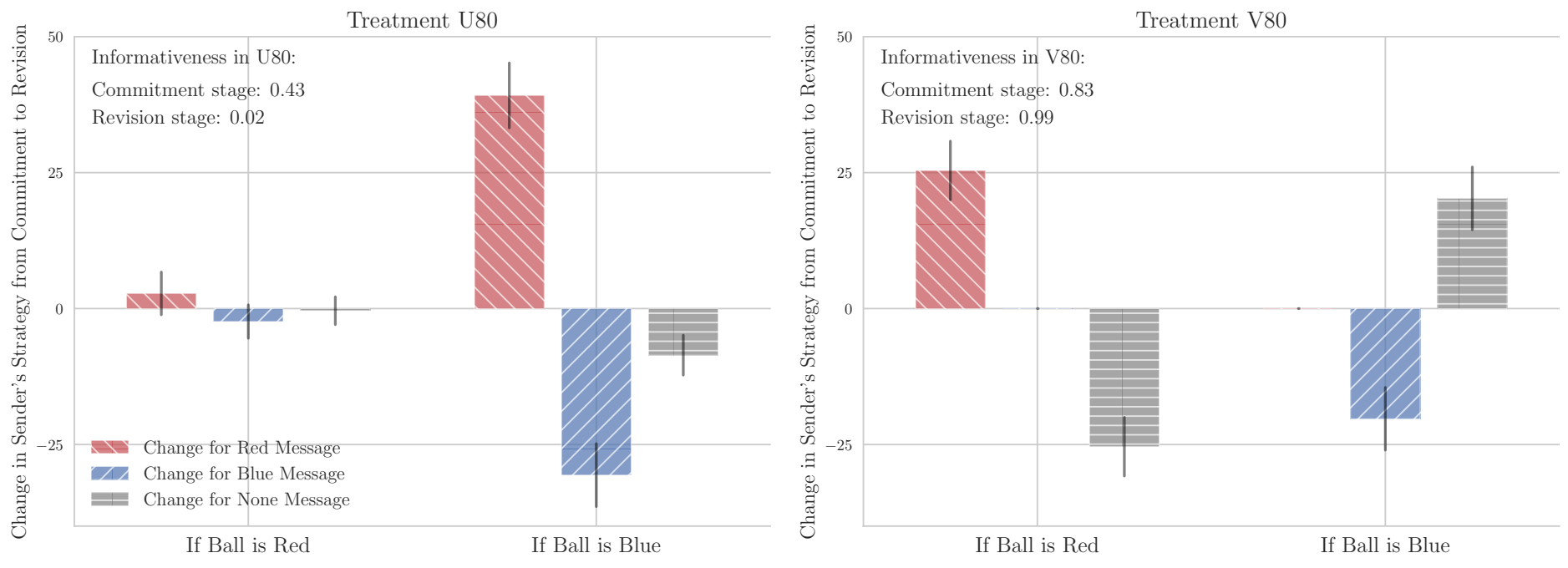
Do Senders *understand* and *use* commitment power?

Identification Strategy:

- Focus on treatments of **partial commitment**
- Compare Senders' behavior between *Commitment* Stage and *Revision* Stage.

Testing Commitment: Senders

Figure 2: Sender's Strategy: Commitment vs. Revision, $\rho = 0.8$



Testing Commitment: Senders

We draw two main conclusions:

- Senders do react to commitment.
- Moreover, as predicted by the theory, the change in behavior between commitment and revision stage are in opposite directions with verifiable vs. unverifiable messages.

Testing Commitment: Receivers

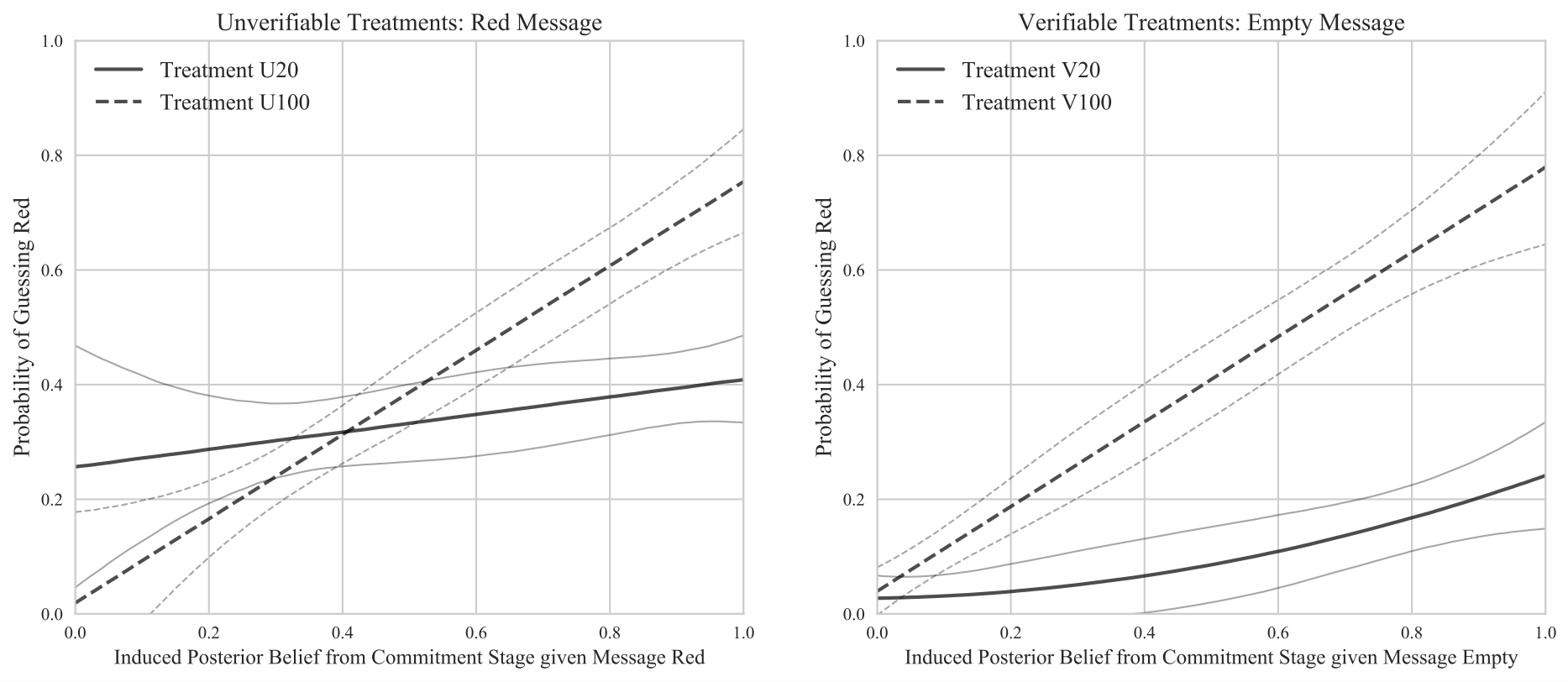
Do Receivers understand the implications of Senders' commitment power?

Identification Strategy:

- Isolate messages with the same "Commitment-Stage" informational content.
- Compare Receivers' guesses in U20 vs U100 (V20 vs V100) conditional on those messages.

Testing Commitment: Sender

Figure 3: Receiver's Response to Persuasive Messages: $\rho = 0.2$ vs. $\rho = 1$



Testing Commitment: Receivers

We draw two main conclusions:

- Receivers do react to commitment power.
- Moreover, they do so in ways predicted by the theory.

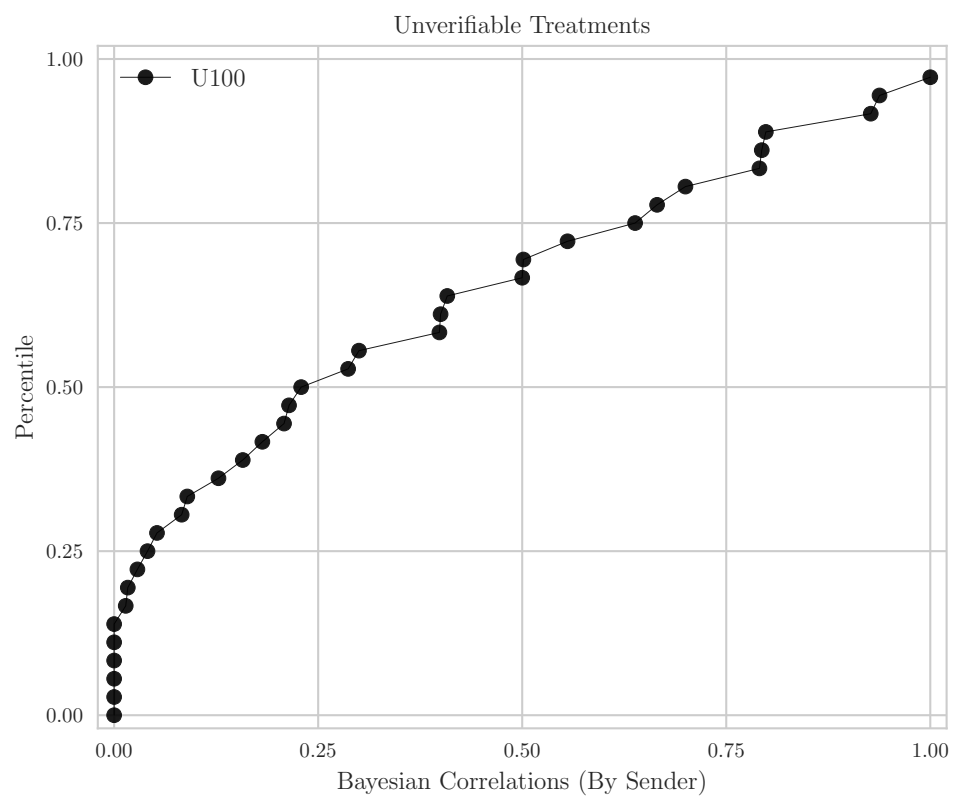
Main Comparative Statics

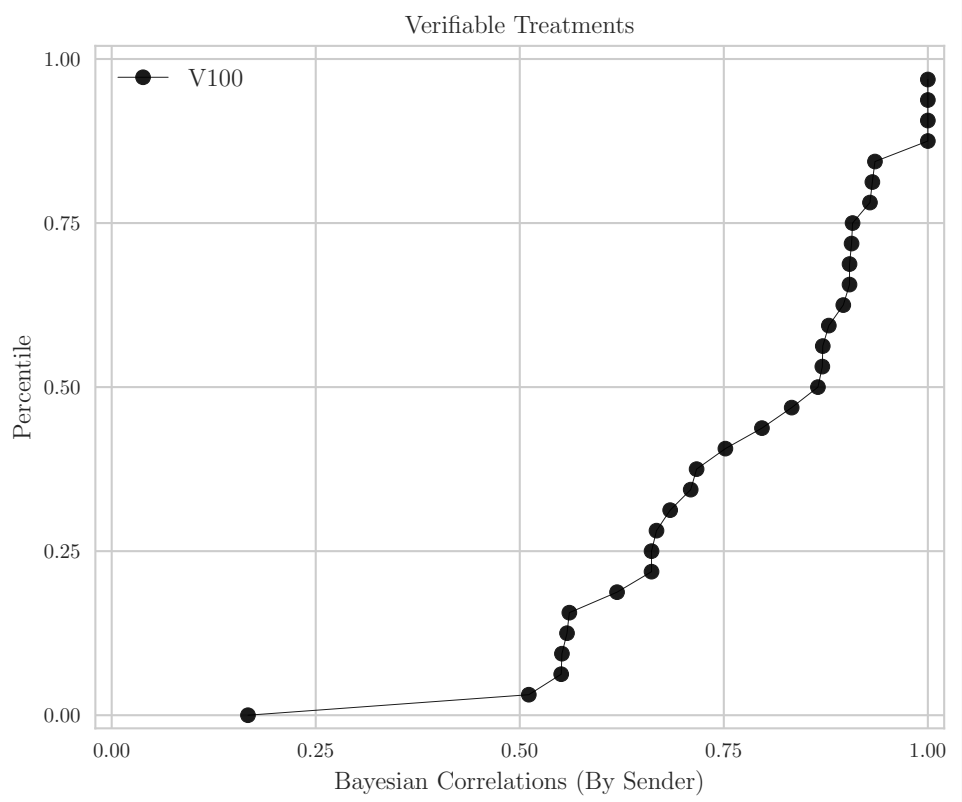
Next, we test our main comparative statics.

Recall our main result:

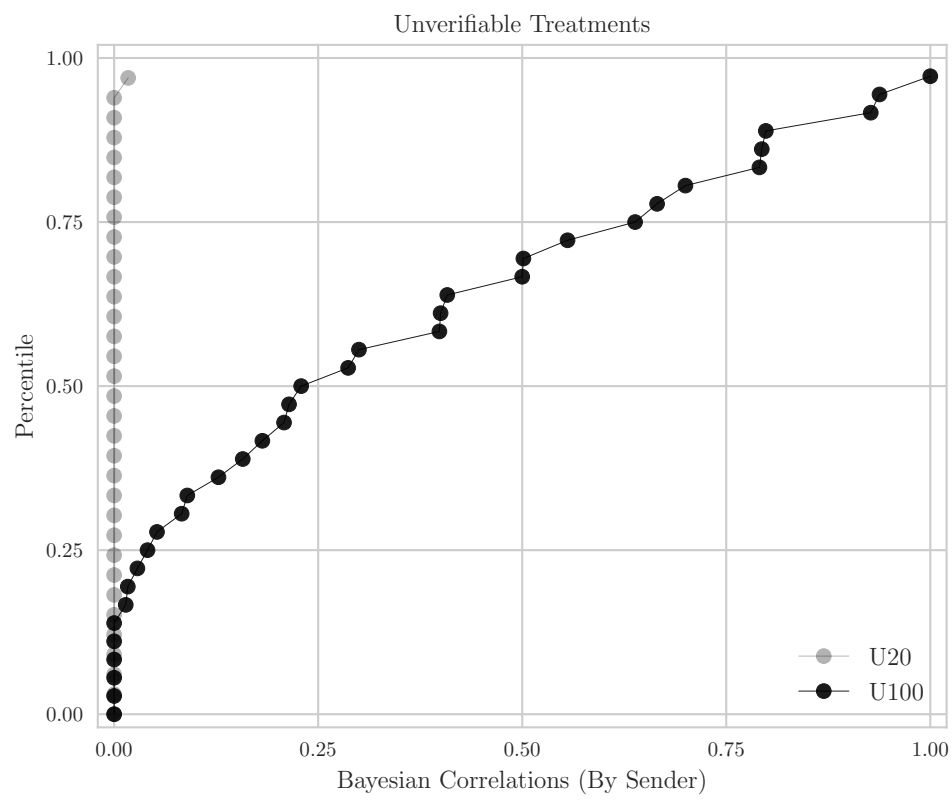
- Under *unverifiable* information, increasing commitment **increases** informativeness.
- Under *verifiable* information, increasing commitment **decreases** informativeness.

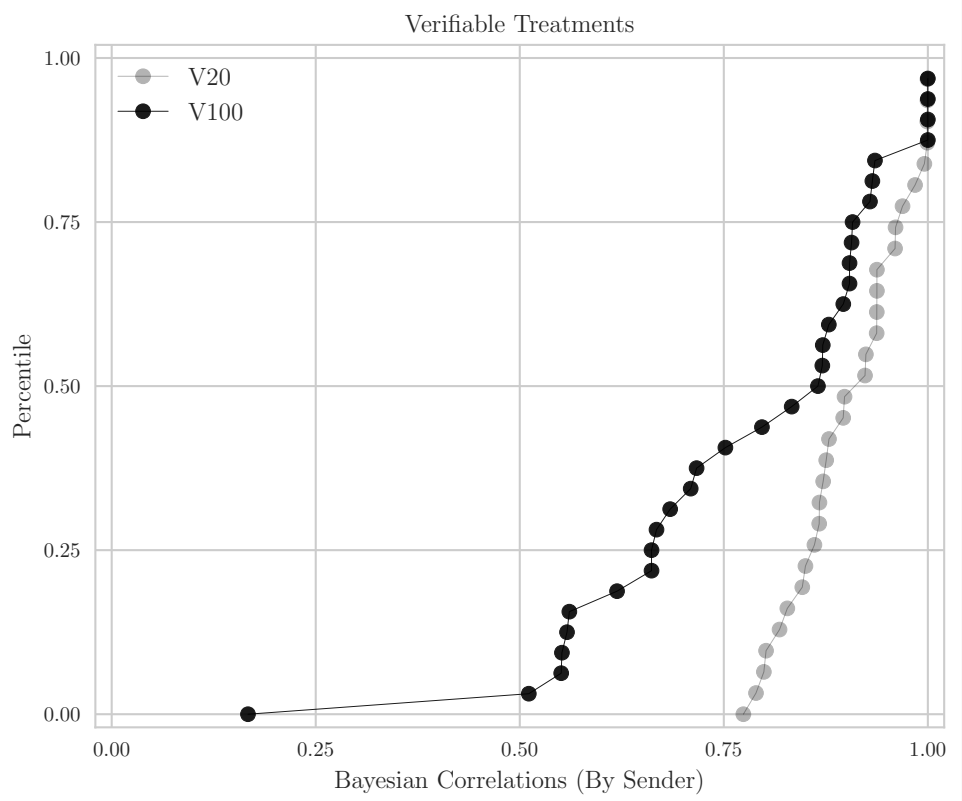
Main Comparative Statics





Main Comparative Statics





Main Comparative Statics

We observe a first-order stochastic

- *increase* in the distribution of informativeness when information is unverifiable.
- *decrease* in the distribution of informativeness when information is verifiable.

Overall, the model does a pretty good job at matching these qualitative predictions.

Increasing commitment:

Unverifiable: Truthfully revealing bad news.

Verifiable: Strategically concealing good news.

The Role of Rules

So far rules are purely instrumental.

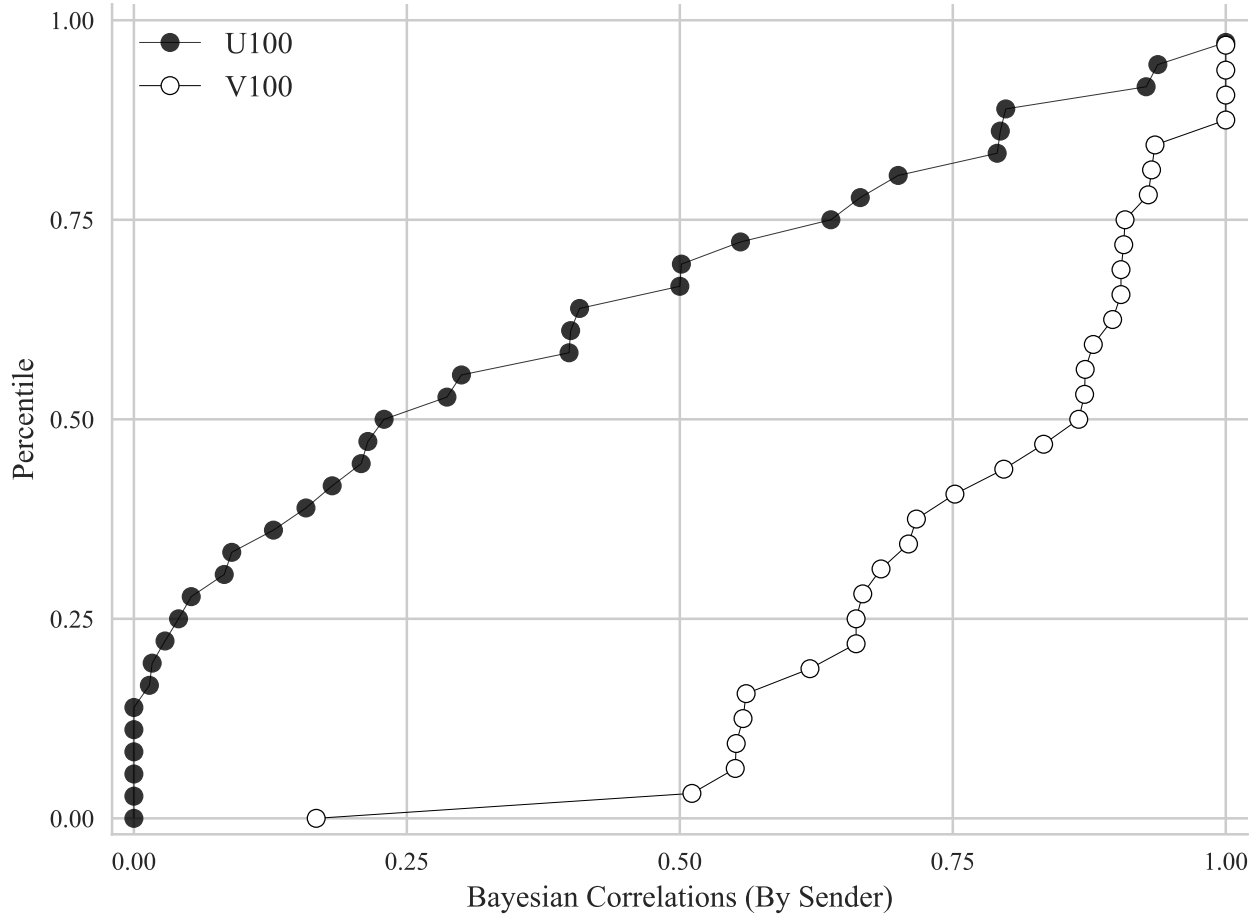
A concluding aside on rules in communication.

Remark

At $\rho = 1$ (full commitment), equilibrium informativeness is rule-independent.

Therefore, rules are predicted to have no welfare effect under full commitment.

And yet... Rules matter disproportionately (when they shouldn't):



Role of Rules

Why?

U100

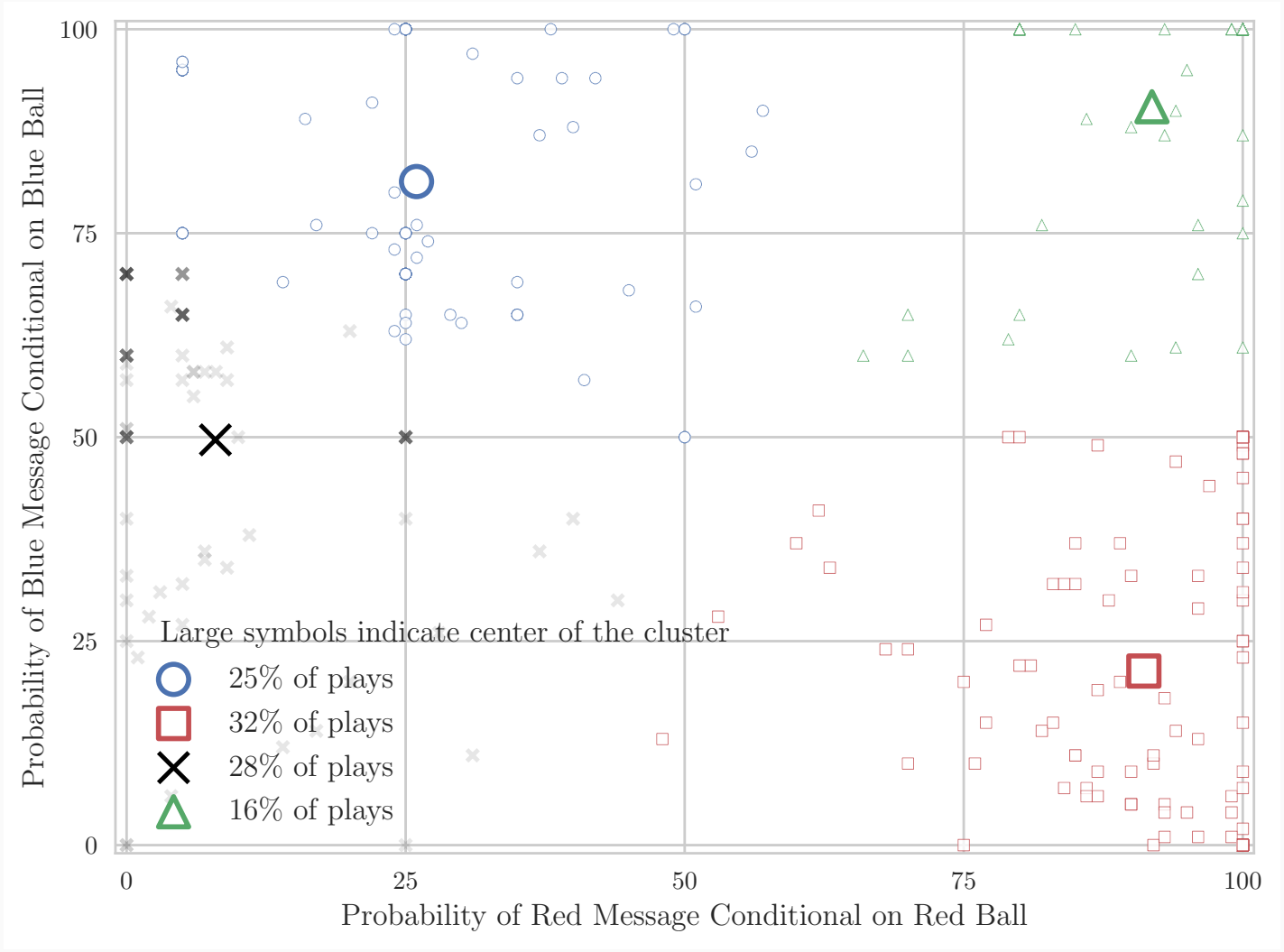
		messages		
		r	*b*	*n*
Ball	R	100%	0	0
	B	50%	50%	0

V100

		messages		
		r	*b*	*n*
Ball	R	0	0	100%
	B	0	50%	50%

The Role of Rules

Figure 4: Sender's Strategy Grouped in Clusters



The Role of Rules

In the data, we observe a combination of:

- Senders' failure to fully embrace a strategy in which good news are concealed.
- Receivers' are overly (i.e. more than Bayesian) skeptical about silence.

The art of concealing information is not a simple one.

An increased policy role for the introduction of rules.

A broader perspective

- A key finding in the literature on experimental cheap talk is *over-communication*.
- "Players tend to be truthful as a sender and credulous as a receiver." (Blume et al., 2017)
- Some of our findings contrast this conclusion.
  - Increasing commitment.
  - Behavior in revision stage.
  - Analysis in U100S: communication types.
  - Comparative statics U vs V.

conclusion

Conclusion

- The first experimental assessment of Bayesian Persuasion.
- We designed tests via a novel framework that nests existing communication models under the same umbrella.
- Subjects react to commitment and match the **qualitative predictions** of our model.
- Substantial heterogeneity and existence of persistent **communication types**.
- An **over-reaction** to *rules* and its policy-implications.

Correlations by Treatment

Theory:

	Commitment (ρ)		
	20%	80%	100%
Verifiable	1	0.57	0.50
Unverifiable	0	0.50	0.50

Data:

	Commitment (ρ)				
	20%		80%		100%
Verifiable	0.80	$\approx$	0.78	$>$	0.67
	$\vee$		$\vee$		
Unverifiable	0.09	$<$	0.21	$\approx$	0.21

Correlations by Treatment

Theory:

	Commitment (ρ)		
	20%	80%	100%
Verifiable	1	0.57	0.50
Unverifiable	0	0.50	0.50

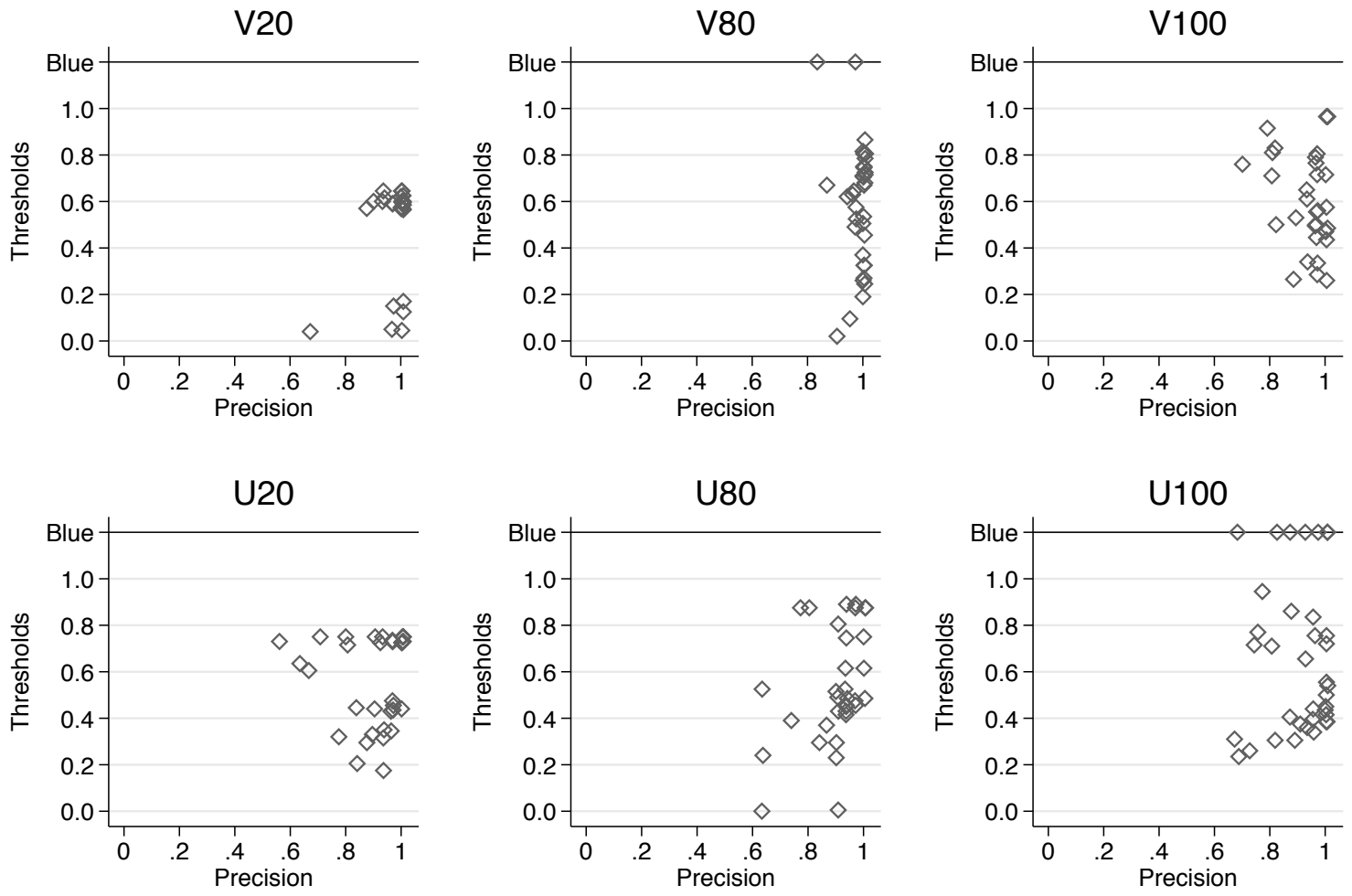
Data:

	Commitment (ρ)				
	20%		80%		100%
Verifiable	0.80	$\approx$	0.78	$>$	0.67
	$\vee$		$\vee$		
Unverifiable	0.09	$<$	0.21	$\approx$	0.21

Data + Bayesian Rec:

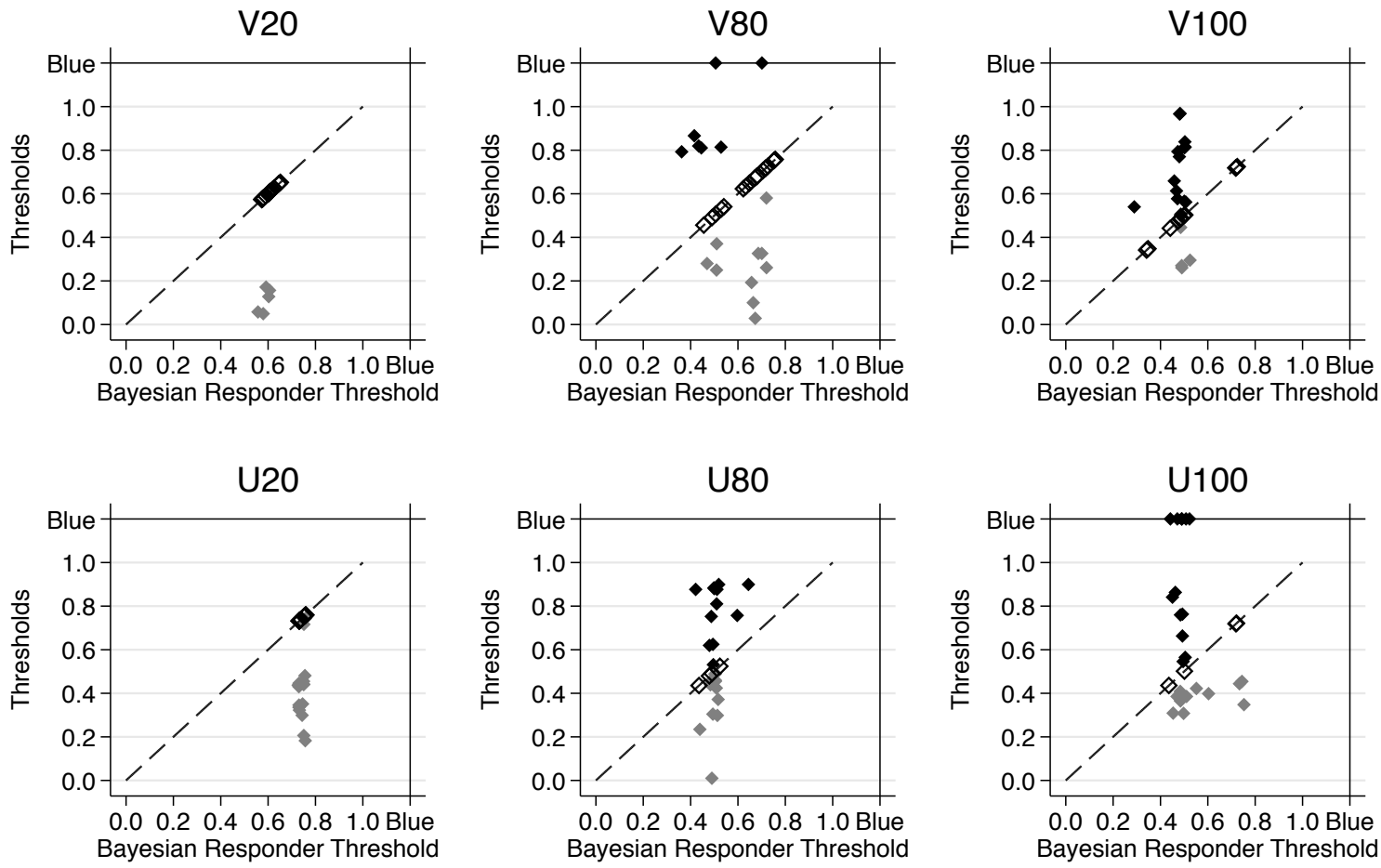
	Commitment (ρ)				
	20%		80%		100%
Verifiable	0.89	$\approx$	0.85	$>$	0.78
	$\vee$		$\vee$		
Unverifiable	0.00	$<$	0.33	$\approx$	0.34

Estimated Thresholds and Precisions



Blue indicates a subject who's choices are best explained by the rule always guess blue.
Values are jittered slightly to make multiple overlapping thresholds distinguishable.

Thresholds: Receivers vs Bayesian



Black for subjects harder to convince than a Bayesian, gray for subjects easier to convince than a Bayesian.
Blue indicates a subject who's choices are best explained by the rule always guess blue.
Values are jittered slightly to make multiple overlapping thresholds distinguishable.
Only includes responders for whom the best fitting threshold accounts for at least 80% of choices.

Enlarging Message Space