Confidence in Inference

En Hua Hu

Toronto

Motivation

- Widespread reliance on samples of signals in decision-making.
  - Google map review of restaurants.
  - Pharmaceutical lab tests on drugs.

Motivation

- Widespread reliance on samples of signals in decision-making.
  - Google map review of restaurants.
  - Pharmaceutical lab tests on drugs.
- Extensive economic analyses - focused on the epistemic aspect.
  - Experimental and empirical evidence on belief updating.
  - Theoretical works on updating rules and belief dynamics.

Motivation

- **This paper**: Focus on empirical content - choice between objects associated with samples of signals.

Motivation

- **This paper**: Focus on empirical content - choice between objects associated with samples of signals.
- **Question**: Do our economic models account well for actual choices?

Motivation

- **This paper**: Focus on empirical content - choice between objects associated with samples of signals.
- **Question**: Do our economic models account well for actual choices?
- **Short Answer**: Not at all.

Results

1. **Theory**: characterize the empirical content of updating models.
   - Result: all models imply a behavioral *separability* axiom.

2. Separability is violated in a **thought experiment**.
   - Intuition: confidence in correctly interpreting samples matters for choice.

3. **Experiment**: test separability + novel confidence elicitation.
   - Collect *choice* data: 95% of subjects violate separability.
   - Evidence of the role of confidence in choice.

Thought Experiment

- Consider a venture capitalist choosing between two projects:
  - Project A: **4** experts predict success and **1** expert predicts failure.
  - Project B: **1** expert predicts success and **0** expert predict failure.

Thought Experiment

- Consider a venture capitalist choosing between two projects:
  - Project A: **4** experts predict success and **1** expert predicts failure.
  - Project B: **1** expert predicts success and **0** expert predict failure.
  - Common choice is $\begin{bmatrix} 4 \\ 1 \end{bmatrix} \succ \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with low confidence.

Thought Experiment

- Consider a venture capitalist choosing between two projects:
  - Project A: **4** experts predict success and **1** expert predicts failure.
  - Project B: **1** expert predicts success and **0** expert predict failure.
  - Common choice is $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix}$ with low confidence.
- **Separability**: adding identical samples preserves preference.
  - $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} \Rightarrow \forall x, y, \begin{bmatrix}4\\1\end{bmatrix} + \begin{bmatrix}x\\y\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}x\\y\end{bmatrix}$

Thought Experiment - Separability

Theorem

The following are equivalent.

1. $\succeq$ *is complete, transitive, continuous, and separable.*
2. $\succeq$ *is the choice of a DM who perceives signals as iid and updates monotonically with respect to the likelihood ratio of samples.*

Takeaway: The choice behavior of any conventional model must be separable.

Thought Experiment

- Suppose now the venture capitalist received the following predictions:
  - Project A: **40** experts predict success and **10** experts predict failure.
  - Project B: **10** experts predict success and **0** expert predict failure.

Thought Experiment

- Suppose now the venture capitalist received the following predictions:
  - Project A: **40** experts predict success and **10** experts predict failure.
  - Project B: **10** experts predict success and **0** expert predict failure.
- Common choice is $\begin{bmatrix} 40 \\ 10 \end{bmatrix} \prec \begin{bmatrix} 10 \\ 0 \end{bmatrix}$ with higher confidence than earlier.

Thought Experiment

- Suppose now the venture capitalist received the following predictions:
  - Project A: **40** experts predict success and **10** experts predict failure.
  - Project B: **10** experts predict success and **0** expert predict failure.
- Common choice is $\begin{bmatrix} 40 \\ 10 \end{bmatrix} \prec \begin{bmatrix} 10 \\ 0 \end{bmatrix}$ with higher confidence than earlier.
- $\begin{bmatrix} 4 \\ 1 \end{bmatrix} \succ \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 40 \\ 10 \end{bmatrix} \prec \begin{bmatrix} 10 \\ 0 \end{bmatrix}$ is ruled out by separability.
- $\begin{bmatrix} 4 \\ 1 \end{bmatrix} \succ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \end{bmatrix} \succ \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

Thought Experiment

- Suppose now the venture capitalist received the following predictions:
  - Project A: **40** experts predict success and **10** experts predict failure.
  - Project B: **10** experts predict success and **0** expert predict failure.
- Common choice is $\begin{bmatrix}40\\10\end{bmatrix} \prec \begin{bmatrix}10\\0\end{bmatrix}$ with higher confidence than earlier.
- $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}40\\10\end{bmatrix} \prec \begin{bmatrix}10\\0\end{bmatrix}$ is ruled out by separability.
- $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} \Rightarrow \begin{bmatrix}4\\1\end{bmatrix} + \begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix}$

Thought Experiment

- Suppose now the venture capitalist received the following predictions:
  - Project A: **40** experts predict success and **10** experts predict failure.
  - Project B: **10** experts predict success and **0** expert predict failure.
- Common choice is $\begin{bmatrix}40\\10\end{bmatrix} \prec \begin{bmatrix}10\\0\end{bmatrix}$ with higher confidence than earlier.
- $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}40\\10\end{bmatrix} \prec \begin{bmatrix}10\\0\end{bmatrix}$ is ruled out by separability.
- $\begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} \Rightarrow \begin{bmatrix}4\\1\end{bmatrix} + \begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}4\\1\end{bmatrix} \succ \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix}$

Intuition & Features of Choice

- Intuition:
  - Confidence in signal interpretation **changes** as sample sizes change.
  - Separability: interpretation of each signal is **identical** and **independent** of the observed sample.

Intuition & Features of Choice

- Intuition:
  - Confidence in signal interpretation **changes** as sample sizes change.
  - Separability: interpretation of each signal is **identical** and **independent** of the observed sample.
- Two complementary features of the choice process:
  - Choice was purely based on sample size and proportion of success.
  - As sample sizes grow, more confident in choice and more willing to neglect the sample size.

Experiment Design

- Collect indifference curves via revealed preference.
  - **Separability $\Leftrightarrow$ Indifference curves are parallel straight lines.**
  - 3 between-subject treatments: two iid info structures and one structure with informativeness uncertainty.
- Novel incentive-compatible confidence elicitation.
- Pre-registered: analysis on the full sample and a sub-sample who satisfy a weak monotonicity condition (more coherent and confident).

Experimental Design

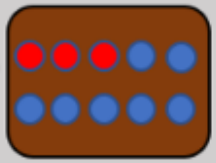
Experiment Set-Up

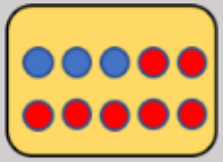
1) There are 200 boxes, half of them are golden, the other half are wooden.

2) Each box contains 10 balls. The color of these balls depends on whether the box is golden or wooden.

3)

Wooden Box	Golden Box
7 blue balls 3 red balls	3 blue balls 7 red balls





4) You do not know which boxes are wooden or golden.

5) You will choose between boxes after seeing draws from them. The draws are made with replacement, meaning after each draw, we put the ball back.

Experimental Design

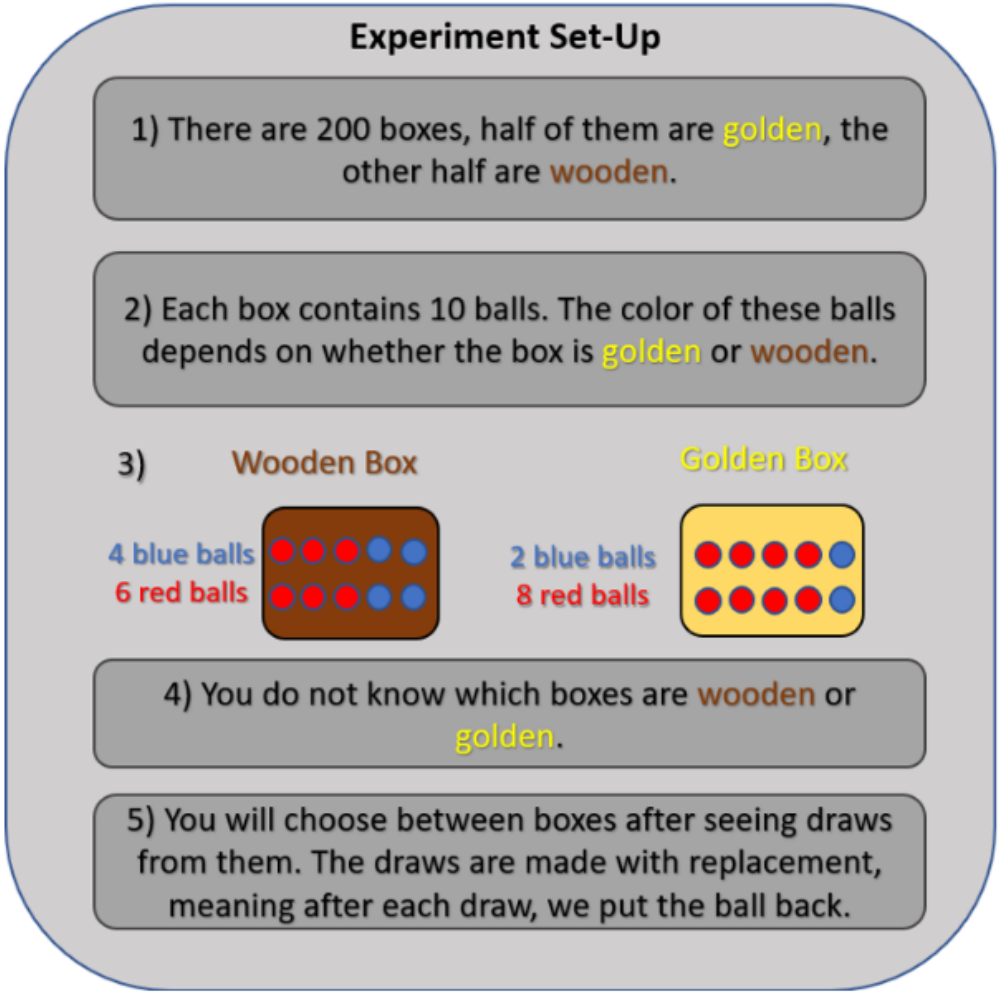


Figure: Asymmetric Treatment

Experimental Design

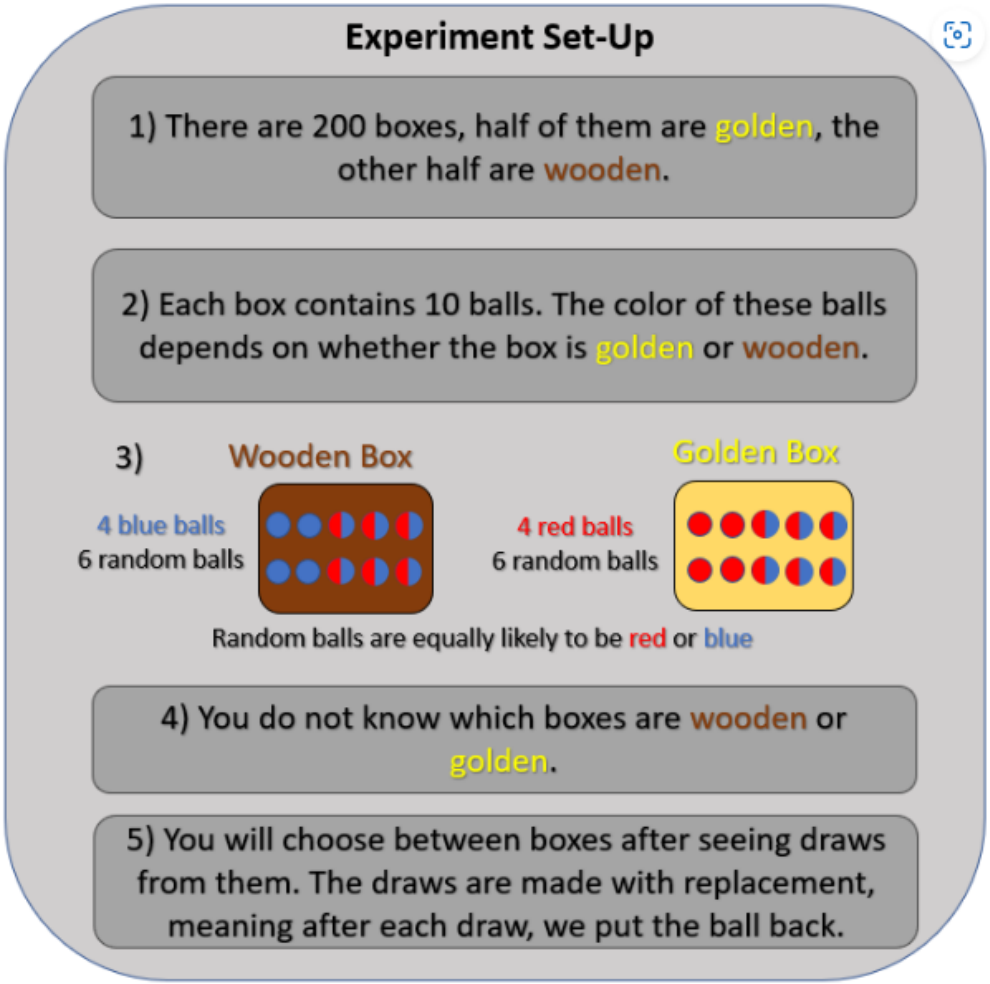


Figure: Correlated Treatment

Experimental Design

Choose Box A		Choose Box B
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **0 out of 10 (0%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **1 out of 10 (10%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **2 out of 10 (20%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **3 out of 10 (30%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **4 out of 10 (40%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **5 out of 10 (50%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **6 out of 10 (60%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **7 out of 10 (70%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **8 out of 10 (80%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **9 out of 10 (90%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ○	Box B - if **10 out of 10 (100%)** balls drawn are red.

Experimental Design

Choose Box A		Choose Box B
Box A - **1 out of 4 (25%)** balls drawn were red.	◉ ○	Box B - if **0 out of 10 (0%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	◉ ○	Box B - if **1 out of 10 (10%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	◉ ○	Box B - if **2 out of 10 (20%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **3 out of 10 (30%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **4 out of 10 (40%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **5 out of 10 (50%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **6 out of 10 (60%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **7 out of 10 (70%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **8 out of 10 (80%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **9 out of 10 (90%)** balls drawn are red.
Box A - **1 out of 4 (25%)** balls drawn were red.	○ ◉	Box B - if **10 out of 10 (100%)** balls drawn are red.

Indifference Curves

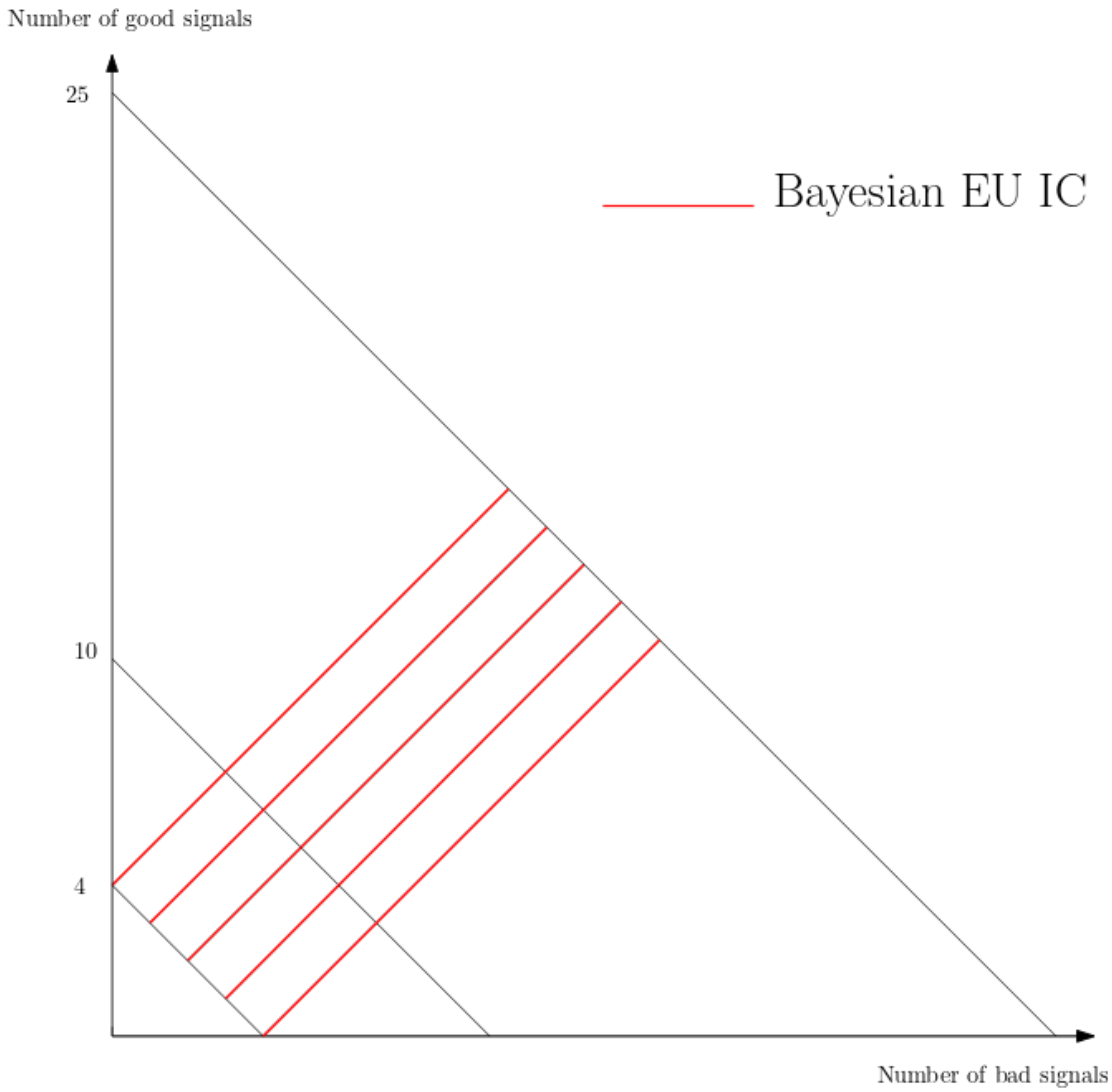


Figure: Bayesian EU IC, symmetric treatment

Indifference Curves

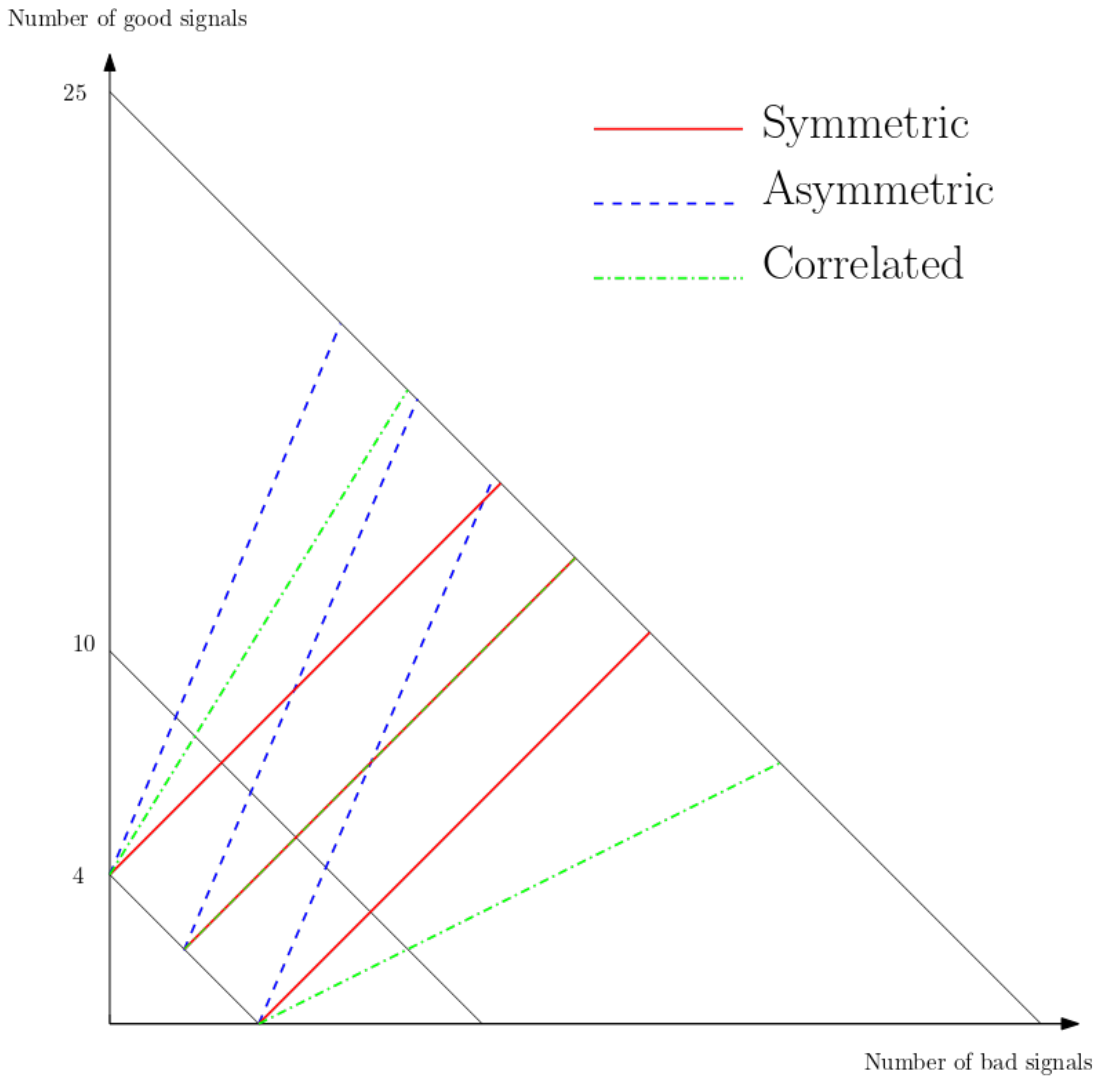


Figure: Bayesian EU IC, all treatment

Indifference Curves

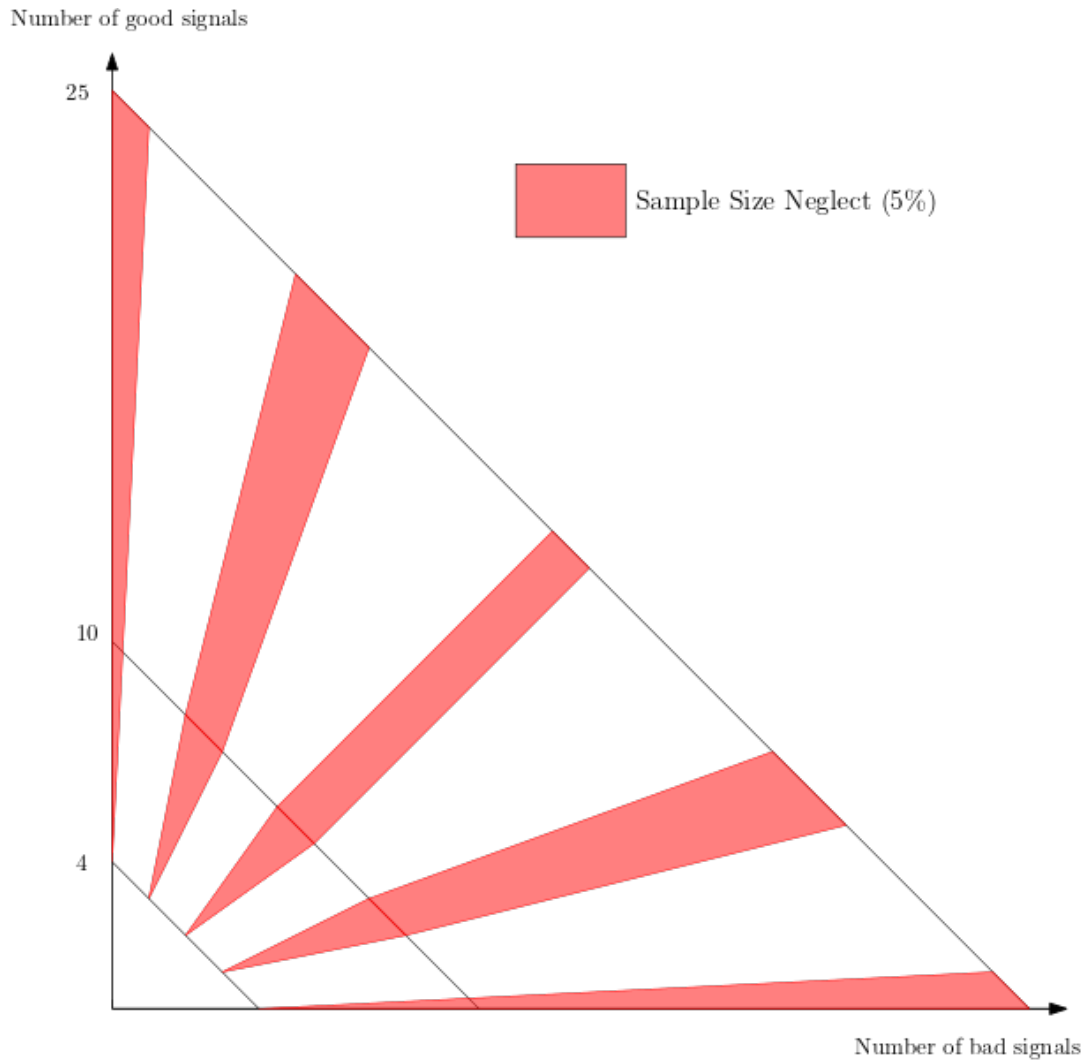


Figure: Sample Size Neglect

Indifference Curves

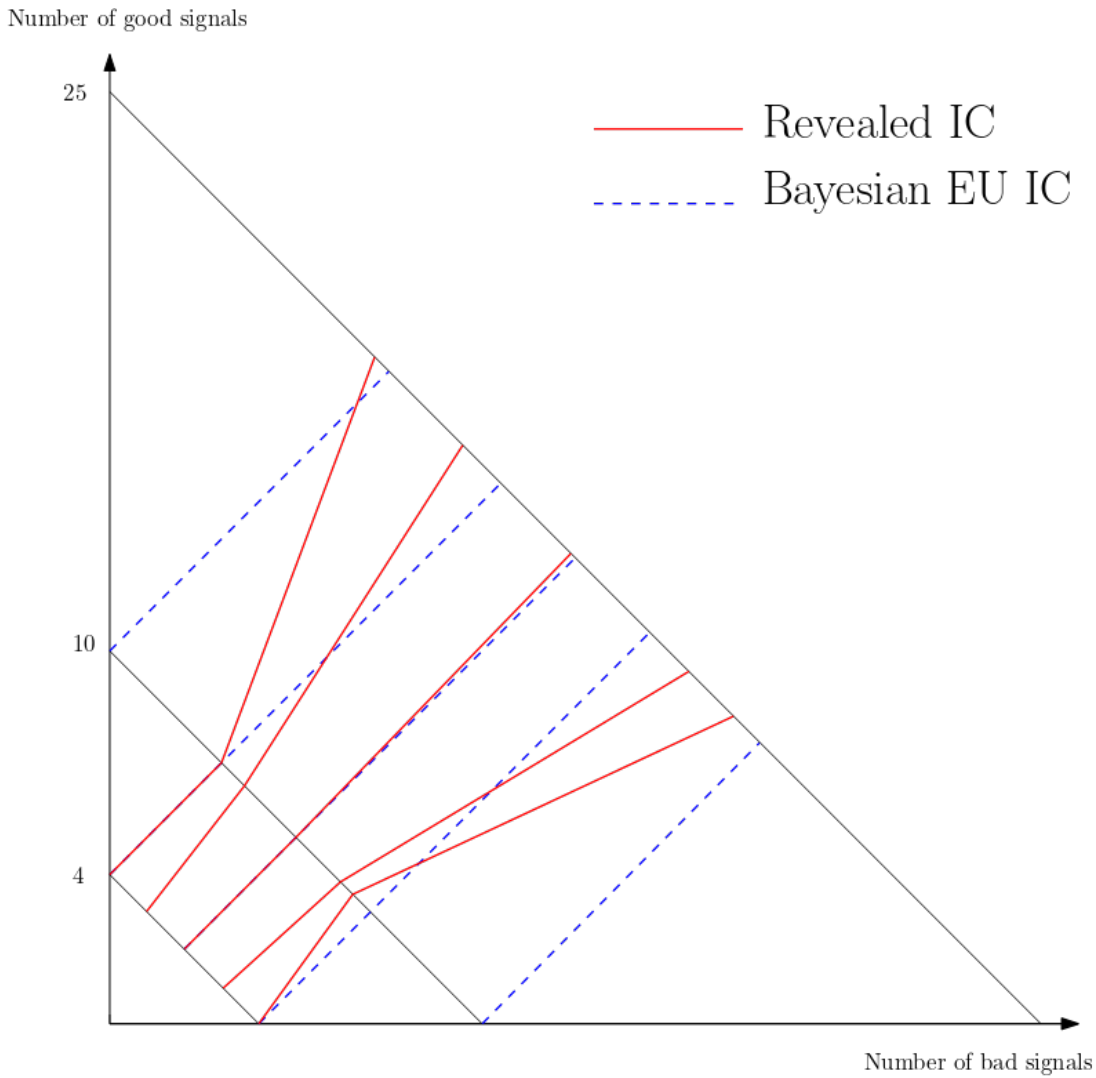


Figure: Full Sample IC, symmetric treatment

Indifference Curves

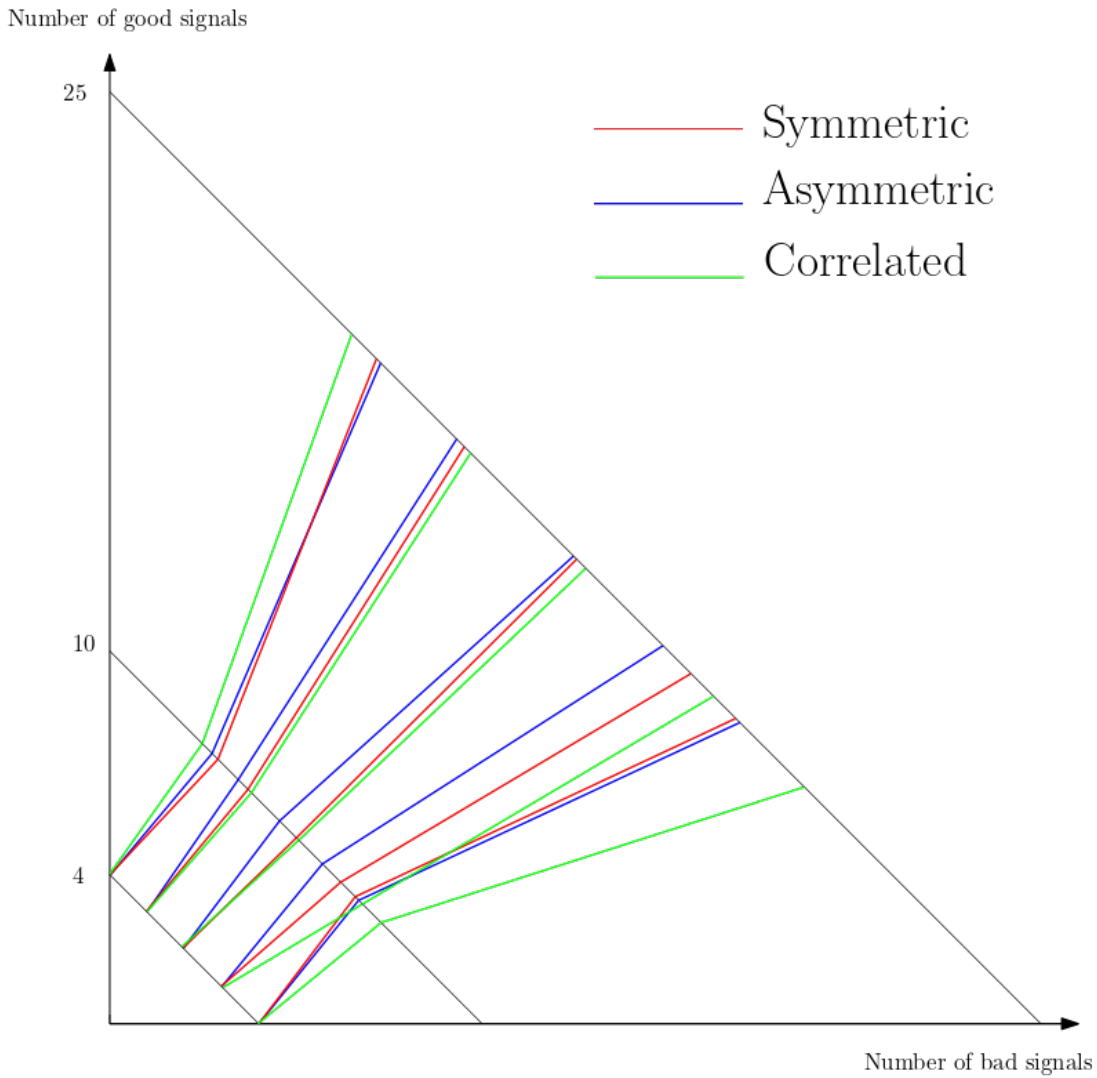


Figure: All treatment - Full Sample

Indifference Curves

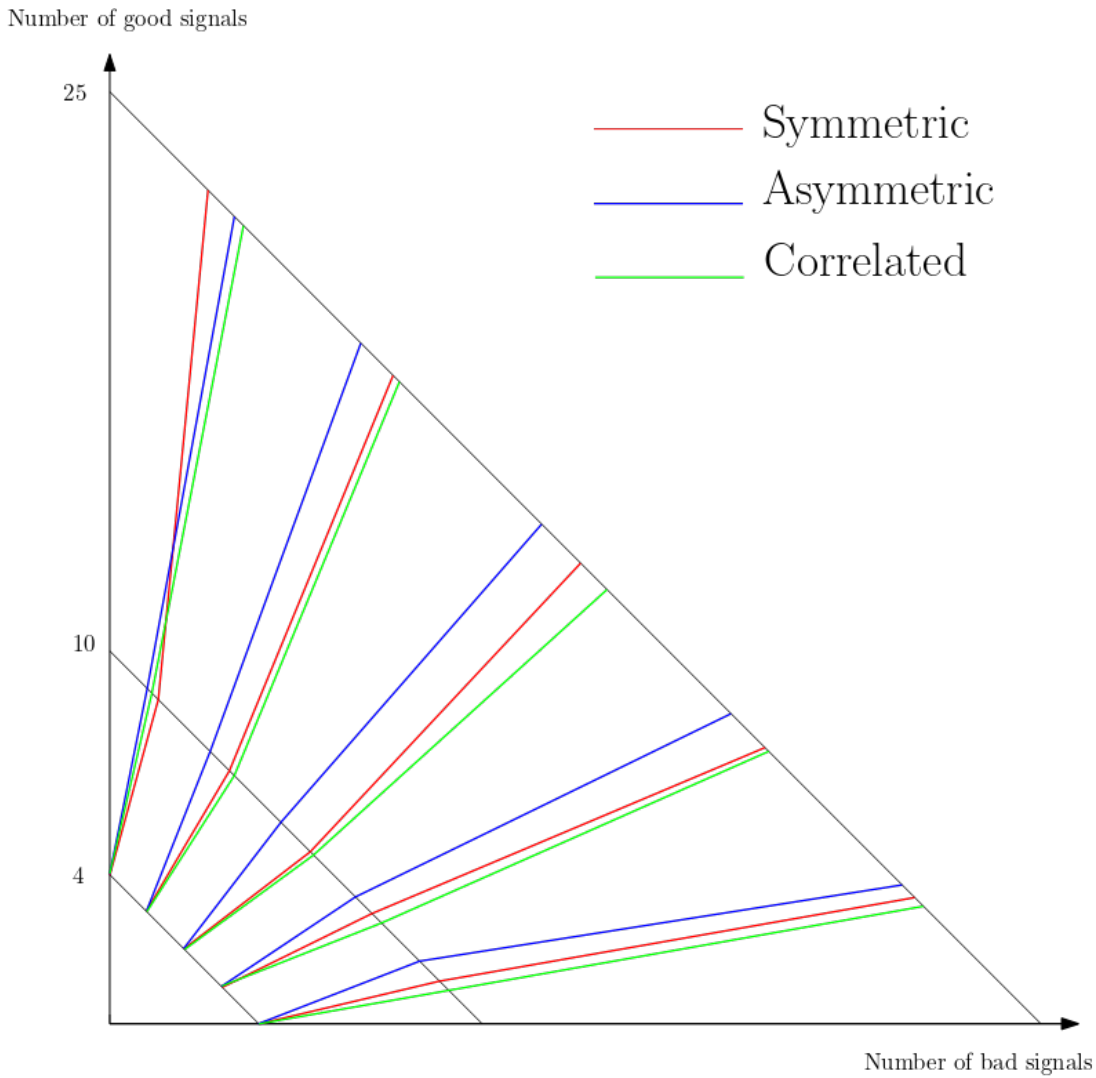


Figure: All treatment - Sub-Sample

Confidence and Sample Size Neglect

- Thought experiment suggests sample size neglect is related to confidence.
- I measure lack of confidence as WTP to learn the correct choice.
  - Well correlated with unincentivized measure.
- Sample size neglect is highly positively correlated with confidence.

Confidence and Sample Size Neglect

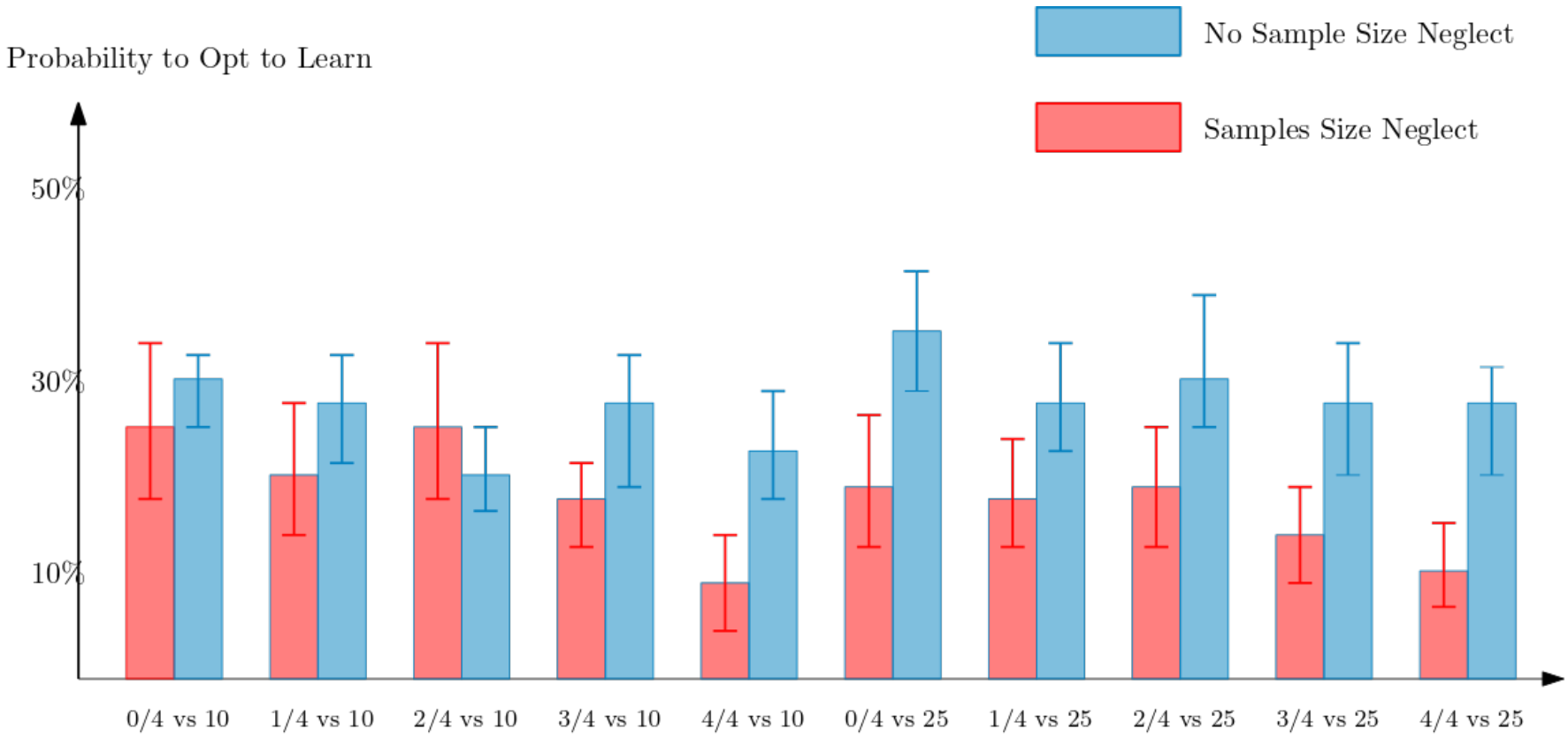


Figure: Full Sample

Confidence and Sample Size Neglect

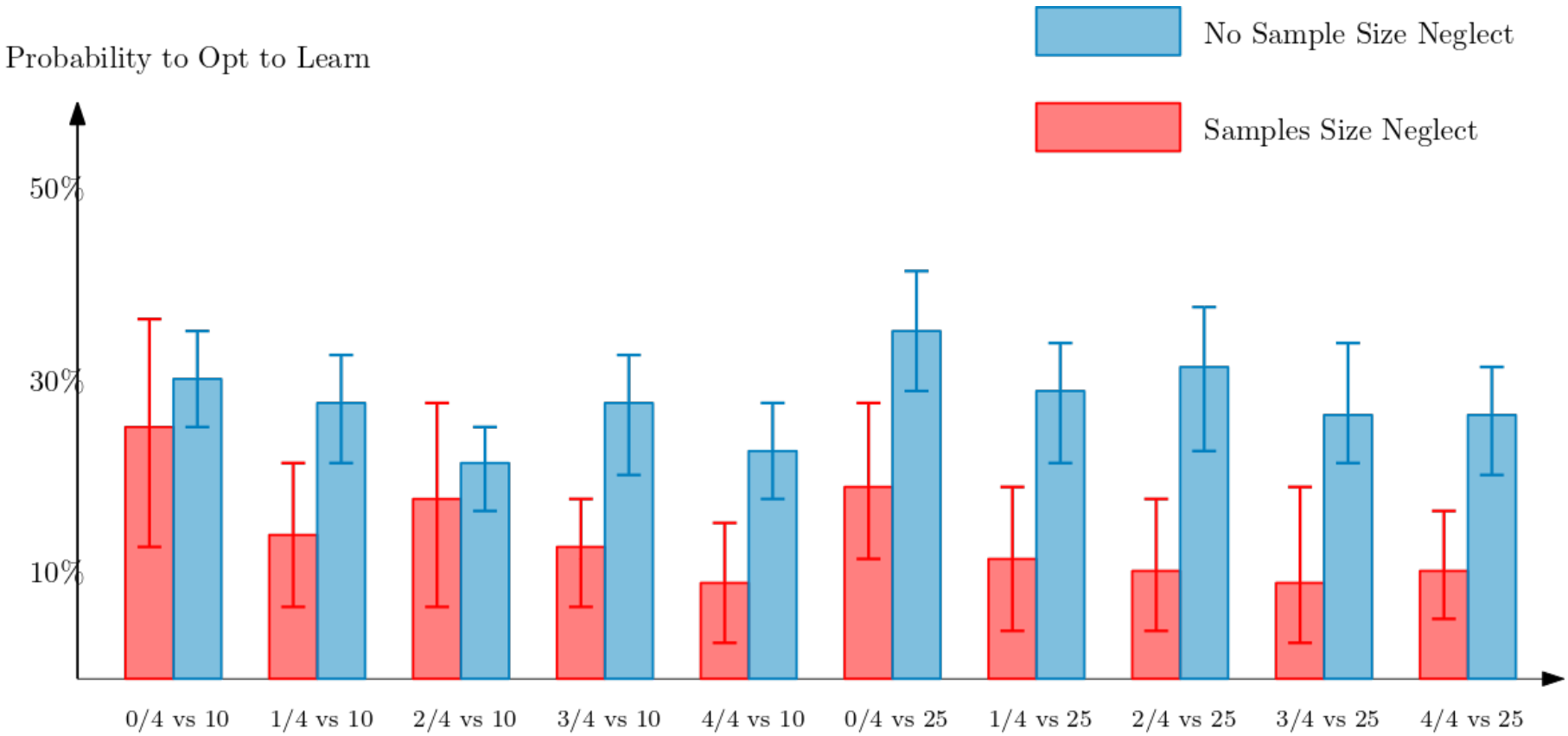


Figure: Sub-Sample

(Limited) Literature

1. Inference from Samples of Signals
   - Previous works: focus on beliefs + do not consider confidence + correlation neglect
     - Grether (1980), Griffin&Tversky (1992), Enke&Zimmermann (2019), Mobius et.al (2022)
   - This paper: focus on choices + show confidence matters + independence neglect

(Limited) Literature

1. Inference from Samples of Signals

2. Incentive-Compatible Confidence Elicitation
   - Previous works: lack incentive-compatibility - only incentive compatible under strong assumptions - too complex - require specific environment.
     - Coffman (2014), Karni (2018, 2020), Chambers&Lambert (2021), Enke&Graeber (2023), Nielsen&Rigotti (2023), Halevy et.al (2023)
   - This paper: incentive-compatible + simple to understand and implement + concrete interpretation in terms of economic fundamentals.

(Limited) Literature

1. Inference from Samples of Signals
2. Incentive-Compatible Confidence Elicitation
3. Uncertainty about Informativeness of Signals
   - Ambiguity: Epstein&Halevy (2023), Liang (2023)
     - Ambiguous information + Ambiguity attitude + Updating rule $\Rightarrow$ Behaviors.
   - Imprecise Cognition: Enke&Graeber (2023), Khaw, Li&Woodford (2021)
     - Uncertain + Anchoring towards mental default $\Rightarrow$ Behaviors.
   - This paper: Separability violation $\Rightarrow$ uncertain about informativeness $\Rightarrow$ learn from samples + confidence in samples.

(Limited) Literature

1. Inference from Samples of Signals
2. Incentive-Compatible Confidence Elicitation
3. Uncertainty about Informativeness of Signals

Conclusion

- Prevailing models of inference satisfy a behavioral separability axiom and leave no room for the role of confidence in inference.
- These models are tested and rejected.
- Confidence is highly positively correlated with sample size neglect.
- Not shown today, happy to talk about:
  - Individual analysis: 95% of subject violate separability.
  - Methodological contribution in confidence elicitation.
  - Behaviors are rationalized by a model of signal uncertainty.

Thank You

Thank You for Your Time.

Experimental Results

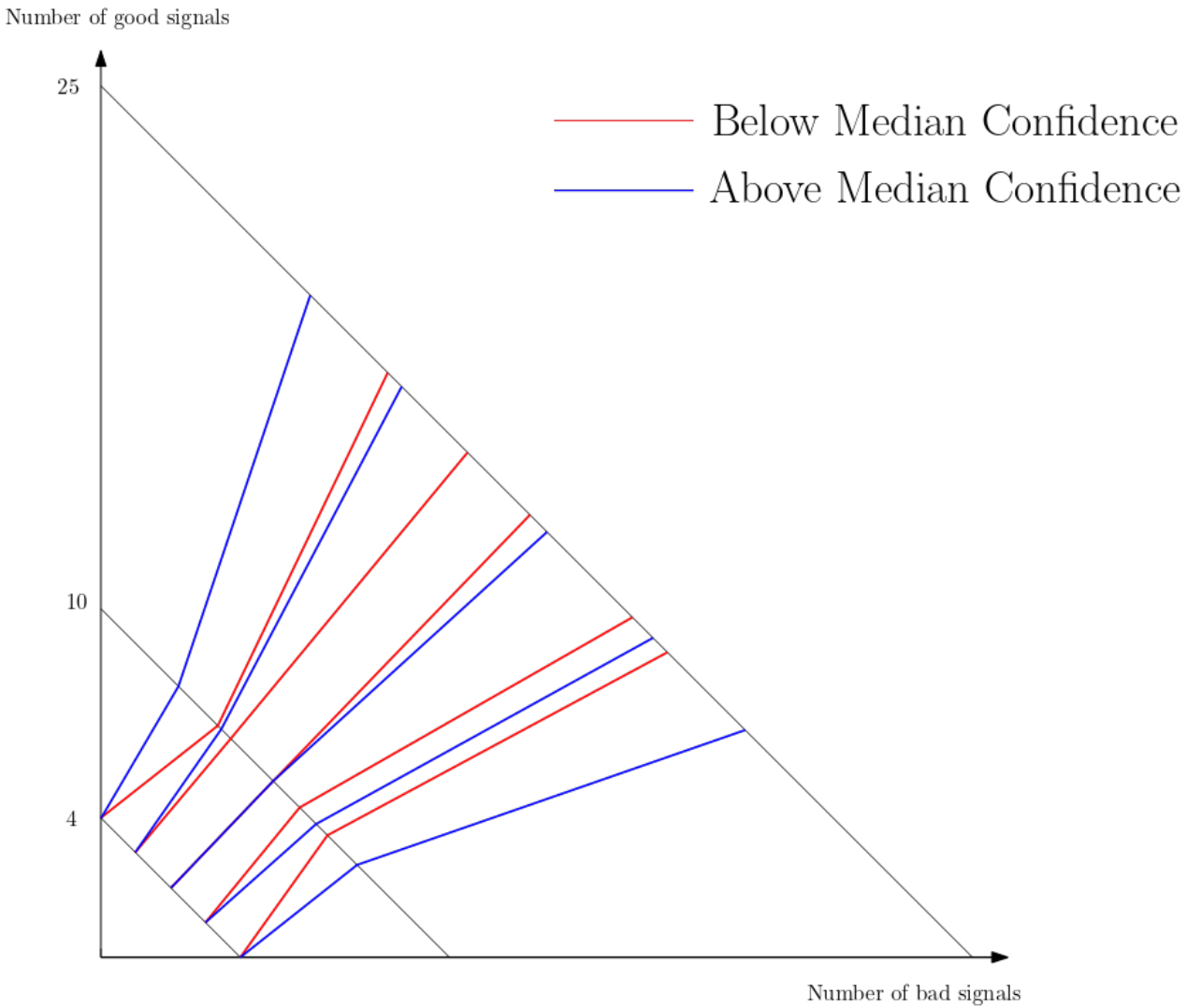
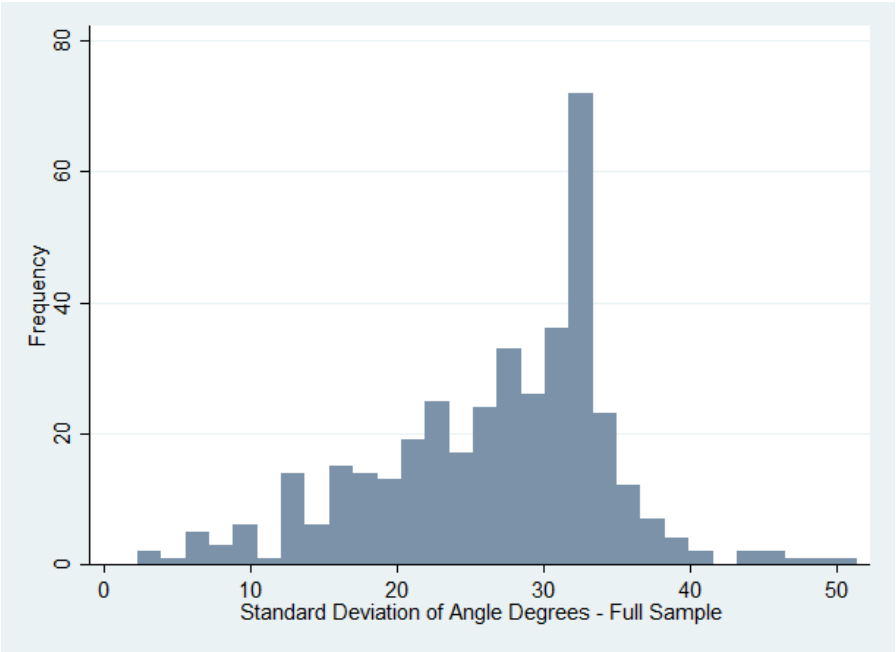
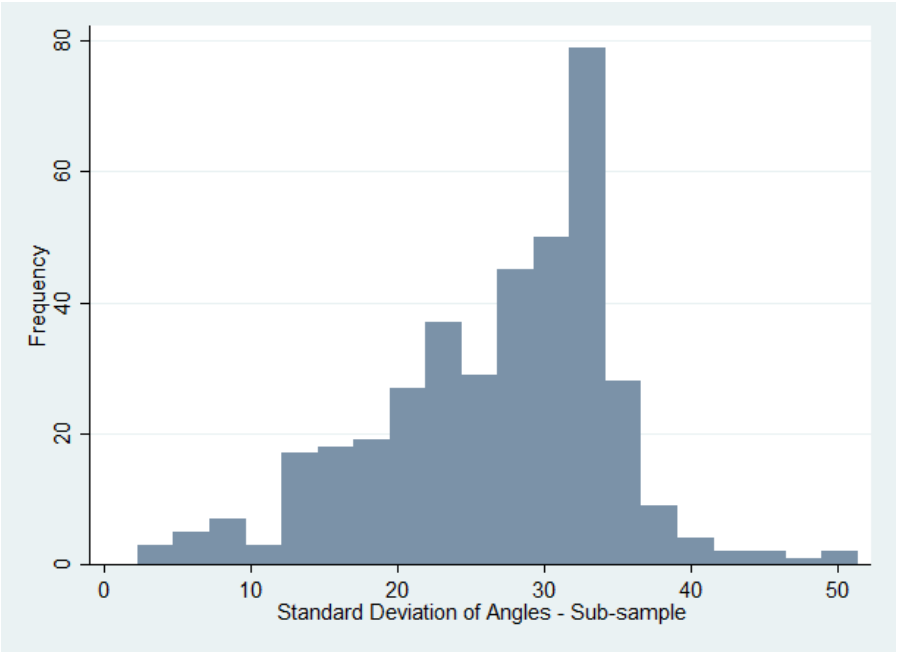


Figure: IC by Confidence

Classifying Subjects Through Standard Deviation of Angles

- Full and sub-sample: 95% of subjects have $std > 10$





Back

Belief Updating ICs

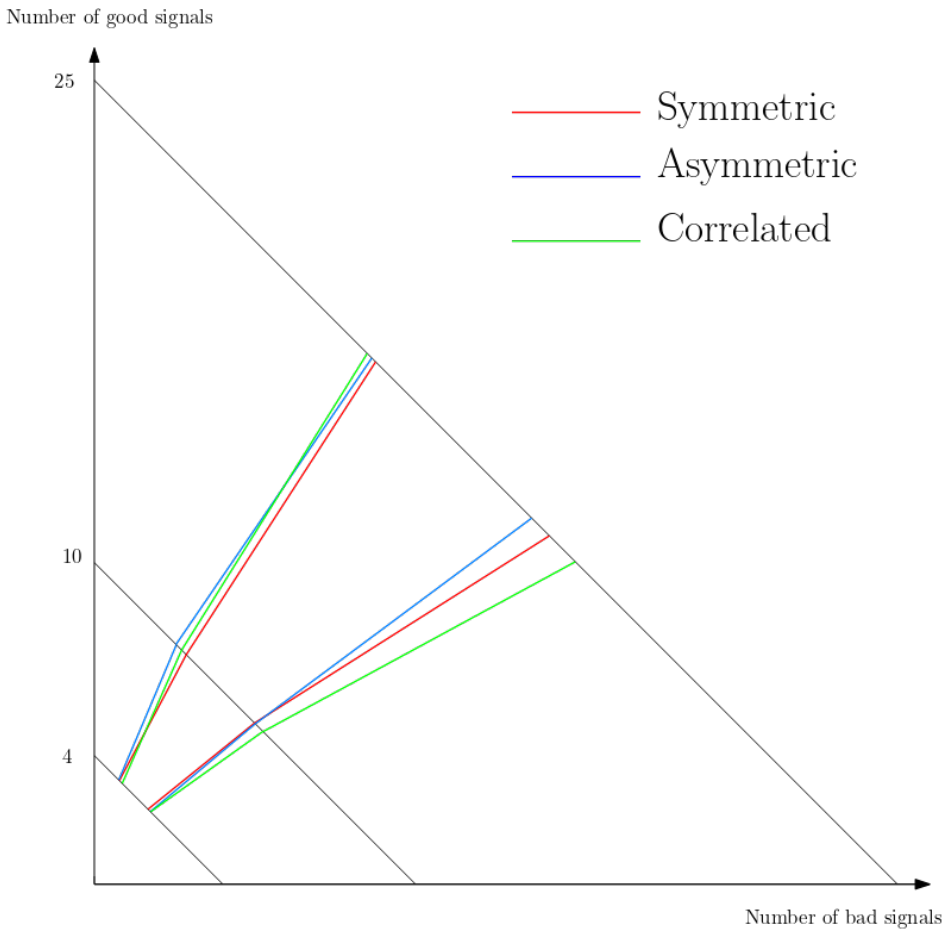


Figure: Full Sample

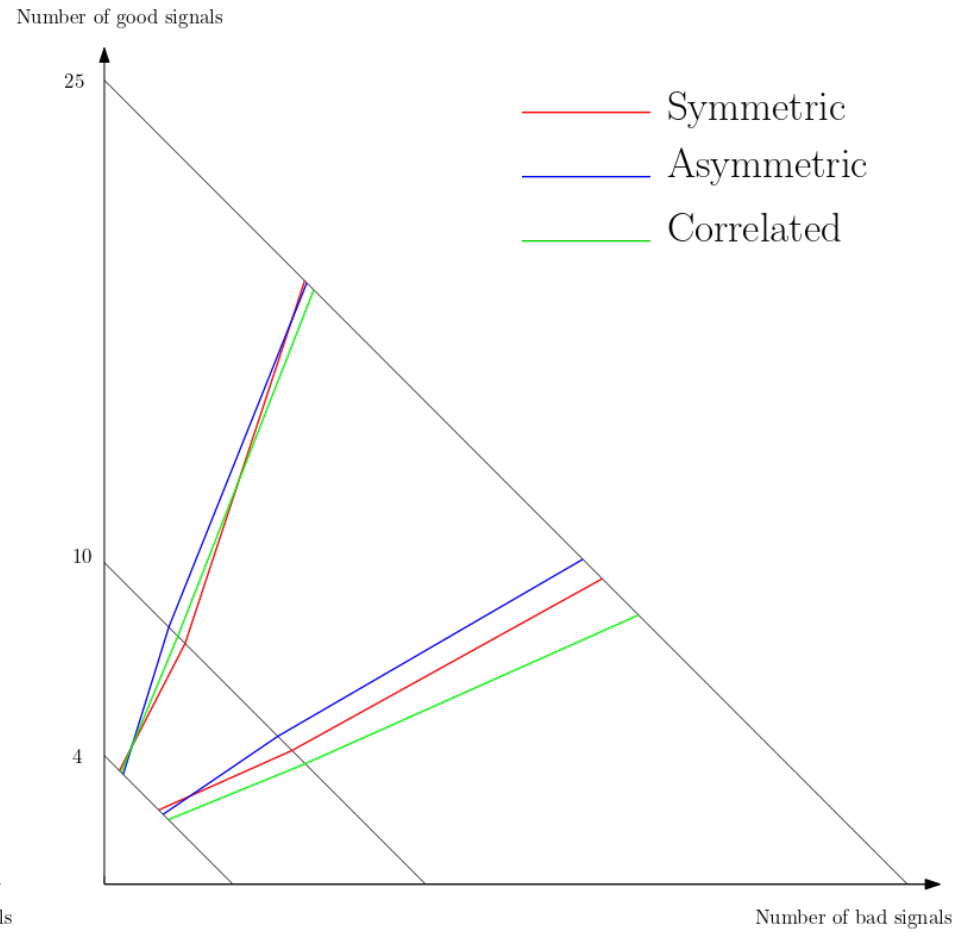


Figure: Sub-Sample