Experimental Tests of Rational Inattention

Mark Dean and Nathaniel Neligh, Journal of Political Economy, forthcoming

presented by **Yonggyun (YG) Kim**

October 20, 2023

Introduction

Motivation

- Economic actors often fail to use all relevant information when making choices
  - Failure to notice whether or not sales tax is included in stated prices (Chetty et al, 2009); buyers of second-hand cars focusing their attention on the leftmost digit of the odometer (Lacetera et al, 2012); purchasers limit their attention to a relatively small number of websites (Santos et al, 2012)
- Several theory to explain informational limits and choice mistakes:
  - Random Utility Model: utility is randomly drawn
  - Signal Detection Theory: people receive noisy signal
  - *Rational Inattention*
- Two crucial assumptions of rational inattention
  - choice is optimal conditional on the information received
  - the DM choose what information to gather in order to maximize the utility of subsequent choice, net of costs

Motivation

- Economic actors often fail to use all relevant information when making choices
  - Failure to notice whether or not sales tax is included in stated prices (Chetty et al, 2009); buyers of second-hand cars focusing their attention on the leftmost digit of the odometer (Lacetera et al, 2012); purchasers limit their attention to a relatively small number of websites (Santos et al, 2012)
- Several theory to explain informational limits and choice mistakes:
  - Random Utility Model: utility is randomly drawn
  - Signal Detection Theory: people receive noisy signal
  - *Rational Inattention*
- Two crucial assumptions of rational inattention
  - choice is optimal conditional on the information received
  - the DM choose what information to gather in order to maximize the utility of subsequent choice, net of costs

Motivation

- Economic actors often fail to use all relevant information when making choices
  - Failure to notice whether or not sales tax is included in stated prices (Chetty et al, 2009); buyers of second-hand cars focusing their attention on the leftmost digit of the odometer (Lacetera et al, 2012); purchasers limit their attention to a relatively small number of websites (Santos et al, 2012)
- Several theory to explain informational limits and choice mistakes:
  - Random Utility Model: utility is randomly drawn
  - Signal Detection Theory: people receive noisy signal
  - *Rational Inattention*
- Two crucial assumptions of rational inattention
  - choice is optimal conditional on the information received
  - the DM choose what information to gather in order to maximize the utility of subsequent choice, net of costs

Preview of the Setting

- Examine the empirical validity of the rational inattention model

- Simple information acquisition task
  - State: Number of balls on the screen
  - DM chooses an action and the payoff depends on the state and the action
  - No time limit or extrinsic cost of information

- Various Experiments
  1. Experiment 1.1: varies the set of available options—testing monotonicity
  2. Experiment 1.2: changes the incentives for making the correct choice—testing NIAS, NIAC / ILR
  3. Experiment 1.3: changes prior beliefs—testing LIP

Preview of the Setting

- Examine the empirical validity of the rational inattention model
- Simple information acquisition task
  - State: Number of balls on the screen
  - DM chooses an action and the payoff depends on the state and the action
  - No time limit or extrinsic cost of information
- Various Experiments
  1. Experiment 1.1: varies the set of available options—testing monotonicity
  2. Experiment 1.2: changes the incentives for making the correct choice—testing NIAS, NIAC / ILR
  3. Experiment 1.3: changes prior beliefs—testing LIP

Preview of the Setting

- Examine the empirical validity of the rational inattention model
- Simple information acquisition task
  - State: Number of balls on the screen
  - DM chooses an action and the payoff depends on the state and the action
  - No time limit or extrinsic cost of information
- Various Experiments
  1. Experiment 1.1: varies the set of available options—testing monotonicity
  2. Experiment 1.2: changes the incentives for making the correct choice—testing NIAS, NIAC / ILR
  3. Experiment 1.3: changes prior beliefs—testing LIP

Preview of the Results

- Adding the new alternatives to a choice set can *increase* the likelihood of existing alternatives being chosen
    - This can be explained by RI theory, but not by RUM or SDT
- No Improving Action Switches (NIAS) and No Improving Attention Cycle (NIAC) conditions hold
    - Caplin and Dean (2015) shows that subjects behavior is consistent with the general model of rational inattention iff NIAS and NIAC hold
- Inconsistent with the Invariant Likelihood Ratio (ILR) property
    - ILR property is predicted by Shannon's entropy model
- Mixed evidence in support of Locally Invariant Posteriors (LIP) condition
    - LIP condition is predicted by uniformly posterior-separable (UPS) cost functions, which is generalized version of the entropy cost
    - 5 out of 6 tests support this prediction, but a joint test that all conditions hold simultaneously is rejected

Preview of the Results

- Adding the new alternatives to a choice set can *increase* the likelihood of existing alternatives being chosen
  - This can be explained by RI theory, but not by RUM or SDT
- No Improving Action Switches (NIAS) and No Improving Attention Cycle (NIAC) conditions hold
  - Caplin and Dean (2015) shows that subjects behavior is consistent with the general model of rational inattention iff NIAS and NIAC hold
- Inconsistent with the Invariant Likelihood Ratio (ILR) property
  - ILR property is predicted by Shannon's entropy model
- Mixed evidence in support of Locally Invariant Posteriors (LIP) condition
  - LIP condition is predicted by uniformly posterior-separable (UPS) cost functions, which is generalized version of the entropy cost
  - 5 out of 6 tests support this prediction, but a joint test that all conditions hold simultaneously is rejected

Preview of the Results

- Adding the new alternatives to a choice set can *increase* the likelihood of existing alternatives being chosen
  - This can be explained by RI theory, but not by RUM or SDT
- No Improving Action Switches (NIAS) and No Improving Attention Cycle (NIAC) conditions hold
  - Caplin and Dean (2015) shows that subjects behavior is consistent with the general model of rational inattention iff NIAS and NIAC hold
- Inconsistent with the Invariant Likelihood Ratio (ILR) property
  - ILR property is predicted by Shannon's entropy model
- Mixed evidence in support of Locally Invariant Posteriors (LIP) condition
  - LIP condition is predicted by uniformly posterior-separable (UPS) cost functions, which is generalized version of the entropy cost
  - 5 out of 6 tests support this prediction, but a joint test that all conditions hold simultaneously is rejected

Preview of the Results

- Adding the new alternatives to a choice set can *increase* the likelihood of existing alternatives being chosen
  - This can be explained by RI theory, but not by RUM or SDT
- No Improving Action Switches (NIAS) and No Improving Attention Cycle (NIAC) conditions hold
  - Caplin and Dean (2015) shows that subjects behavior is consistent with the general model of rational inattention iff NIAS and NIAC hold
- Inconsistent with the Invariant Likelihood Ratio (ILR) property
  - ILR property is predicted by Shannon's entropy model
- Mixed evidence in support of Locally Invariant Posteriors (LIP) condition
  - LIP condition is predicted by uniformly posterior-separable (UPS) cost functions, which is generalized version of the entropy cost
  - 5 out of 6 tests support this prediction, but a joint test that all conditions hold simultaneously is rejected

Background

Theory: Set-Up and Data

Decision Problem

- Single decision maker (DM)
- Ω: the finite set of states ($\omega \in \Omega$ is a generic state)
- $\mu \in \Delta(\Omega)$: the finite set of states
- A: the set of available actions ($a \in A$ is a generic action)
- $u(a, \omega)$: the DM's utility of action a in state ω
- A decision problem is defined by (μ, A)—both of which we assume can be chosen by the experimenter.

Theory: Set-Up and Data

Data

- The data observed from a particular decision problem generates a *state dependent stochastic choice (SDSC)* function
- $P_{(\mu,A)}$: the SDSC function associated with (μ, A)
  - $P_{(\mu,A)}(a|\omega)$: the probability that action $a \in A$ was chosen in state $\omega \in \Omega$
- SDSC function also implies a set of 'revealed posteriors' via Bayes' rule:

$$\gamma^a(\omega) \equiv \frac{\mu(\omega) \cdot P_{(\mu,A)}(a|\omega)}{\sum_{\nu\in\Omega} \mu(\nu) \cdot P_{(\mu,A)}(a|\nu)} \qquad (1)$$

Theory: Set-Up and Data

Data

- The data observed from a particular decision problem generates a *state dependent stochastic choice (SDSC)* function

- $P_{(\mu,A)}$: the SDSC function associated with (μ, A)
  - $P_{(\mu,A)}(a|\omega)$: the probability that action $a \in A$ was chosen in state $\omega \in \Omega$

- SDSC function also implies a set of 'revealed posteriors' via Bayes' rule:

$$\gamma^a(\omega) \equiv \frac{\mu(\omega) \cdot P_{(\mu,A)}(a|\omega)}{\sum_{\nu \in \Omega} \mu(\nu) \cdot P_{(\mu,A)}(a|\nu)} \tag{1}$$

Theory: Rational Inattention Model

Information Structure

- DM chooses an *information structure* prior to choosing an action
- Assume that the subject's choice of information structure is not observed, and so has to be inferred from their choice data
- For simplicity, we consider each signal as a posterior belief $\gamma \in \Gamma$
- $\pi : \Omega \to \Delta(\Gamma)$: information structure
  - $\pi(\gamma|\omega)$: the probability of signal γ given state ω
  - $\gamma(\omega)$: the probability of state ω conditional on receiving signal γ

Theory: Rational Inattention Model

Expected Payoffs

- G: the gross payoff of using an info. structure π in a decision problem (μ, A)

$$G(\mu, A, \pi) \equiv \sum_{\gamma \in \Gamma(\pi)} \left[\sum_{\omega \in \Omega} \mu(\omega)\pi(\gamma|\omega)\right] \cdot \left[\max_{a \in A} \sum_{\omega \in \Omega} \gamma(\omega) u(a, \omega)\right] \tag{2}$$

- $K(\mu, \pi)$: the cost of information structure π given prior μ
- DM's objective is to maximize

$$G(\mu, A, \pi) - K(\mu, \pi) \tag{3}$$

- G is observable, but K is not

Experimental Design 1

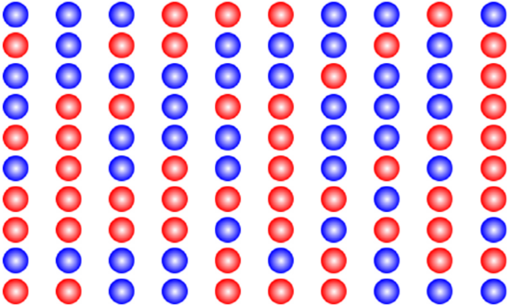
Experimental Design: Setup

- **State**: 100 balls shown on a screen, some are red, some are blue
- Prior to seeing the screen, subjects are informed of the probability distribution over such states
- Having seen the screen, they choose from a number of different actions whose payoffs are state dependent

Example Question

Remember:

- With 50% probability there will be 49 red dots
- With 50% probability there will be 51 red dots



Please select from the following options:

	Option	Pay if there are 49 red dots	Pay if there are 51 red dots
○	A	10	0
◉	B	0	10
○	C	5	5

Experiment 1.1: Testing for Responsive Attention

Theory: Other Models of Stochastic Choice

1. **Random Utility Model** (RUM) : choices are determined by the maximization of a utility function drawn from some distribution that does not depend on the decision problem
2. **Signal Detection Theory** (SDT) : people receive a noisy signal about the state of the world, then choose actions optimally given their subsequent beliefs

Definition 1

A SDSC function satisfies *Monotonicity* if, for every $\mu \in \Delta(\Omega)$, $A \subset B$, $\omega \in \Omega$ and $a \in A$, $P_{(\mu,A)}(a|\omega) \geq P_{(\mu,B)}(a|\omega)$

- Monotonicity is a necessary property of data generated by RUM and SDT
- However, Monotonicity is *not* implied by rational inattention models
  - Matejka and McKay [2015]: the introduction of a new act can increase the incentives to acquire information, which may in turn lead the DM to learn that an existing act was of high value.

1. **Random Utility Model** (RUM)
2. **Signal Detection Theory** (SDT)

Definition 1

A SDSC function satisfies *Monotonicity* if, for every $\mu \in \Delta(\Omega)$, $A \subset B$, $\omega \in \Omega$ and $a \in A$, $P_{(\mu,A)}(a|\omega) \geq P_{(\mu,B)}(a|\omega)$

- Monotonicity is a necessary property of data generated by RUM and SDT
- However, Monotonicity is *not* implied by rational inattention models
    - Matejka and McKay [2015]: the introduction of a new act can increase the incentives to acquire information, which may in turn lead the DM to learn that an existing act was of high value.

Theory: Other Models of Stochastic Choice

1. **Random Utility Model** (RUM)
2. **Signal Detection Theory** (SDT)

Definition 1

A SDSC function satisfies *Monotonicity* if, for every $\mu \in \Delta(\Omega)$, $A \subset B$, $\omega \in \Omega$ and $a \in A$, $P_{(\mu,A)}(a|\omega) \geq P_{(\mu,B)}(a|\omega)$

- Monotonicity is a necessary property of data generated by RUM and SDT
- However, Monotonicity is *not* implied by rational inattention models
    - Matejka and McKay [2015]: the introduction of a new act can increase the incentives to acquire information, which may in turn lead the DM to learn that an existing act was of high value.

Theory: Other Models of Stochastic Choice

1. **Random Utility Model** (RUM)
2. **Signal Detection Theory** (SDT)

Definition 1

A SDSC function satisfies *Monotonicity* if, for every $\mu \in \Delta(\Omega)$, $A \subset B$, $\omega \in \Omega$ and $a \in A$, $P_{(\mu,A)}(a|\omega) \geq P_{(\mu,B)}(a|\omega)$

- Monotonicity is a necessary property of data generated by RUM and SDT
- However, Monotonicity is *not* implied by rational inattention models
  - Matejka and McKay [2015]: the introduction of a new act can increase the incentives to acquire information, which may in turn lead the DM to learn that an existing act was of high value.

Experiment 1.1: Testing for Responsive Attention

- Two equally likely states: state 1 (49 red balls); state 2 (51 red balls)
- Payment: probability points with a prize of \$20
- Each subject faced 75 repetitions of two decision problems

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

- 4 treatments with $b_1 < 50 < b_2$: (40,55), (40,52), (30,55), (30,52)

Experiment 1.1: Testing for Responsive Attention

- **Monotonicity**: $A_1 = \{a, b\} \subset A_2 = \{a, b, c\} \quad \Rightarrow \quad P_{(\mu,A_1)}(b|\omega) \geq P_{(\mu,A_2)}(b|\omega)$
- SDSC data violates Monotonicity if $P_{(\mu,A_1)}(b|2) < P_{(\mu,A_2)}(b|2)$, which we show in the following slide
- Then, RUM or SDT cannot explain this data

Experiment 1.1: Testing for Responsive Attention

Table 2: Results of Experiment 1.1[44]

		$P(b\|1)$			$P(b\|2)$			
Treat	N	$\{a,b\}$	$\{a,b,c\}$	Prob	$\{a,b\}$	$\{a,b,c\}$	Prob	% Subjects
1	7	2.9	6.8	0.52	50.6	59.8	0.54	29
2	7	5.7	14.7	0.29	21.1	63.1	0.05	57
3	7	9.5	5.0	0.35	21.4	46.6	0.06	43
4	7	1.1	0.8	0.76	19.9	51.7	0.02	57
Total	28	4.8	6.6	0.52	28.3	55.6	<0.01	46

- Treatments: (b_1, b_2) are (40,55), (40,52), (30,55), (30,52)
- Col 3: prob. associated with the null hypothesis that prob. in column 1 and 2 are equal.
- % subjects: the fraction of choosing b significantly more in state 2 when c is available

Experiment 1.1: Testing for Responsive Attention

- How can Rational Inattention model explain this result?
  - In DP1, the incentive for gathering information is low, and the subject can simply choose *a* which guarantees 50 points
  - However, in DP2, with the addition of the option c, the subject may want to identify the true state with a high degree of accuracy
  - By acquiring info, half the time they will determine that the state is in fact 2, in which case b is the best option

$\Rightarrow$ there is potentially a ‘spillover’ effect of adding c to the choice set which is to increase the probability of selecting b

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

Experiment 1.1: Testing for Responsive Attention

- How can Rational Inattention model explain this result?
  - In DP1, the incentive for gathering information is low, and the subject can simply choose *a* which guarantees 50 points
  - However, in DP2, with the addition of the option c, the subject may want to identify the true state with a high degree of accuracy
  - By acquiring info, half the time they will determine that the state is in fact 2, in which case b is the best option

  $\Rightarrow$ there is potentially a 'spillover' effect of adding c to the choice set which is to increase the probability of selecting b

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

Experiment 1.1: Testing for Responsive Attention

- How can Rational Inattention model explain this result?
  - In DP1, the incentive for gathering information is low, and the subject can simply choose *a* which guarantees 50 points
  - However, in DP2, with the addition of the option c, the subject may want to identify the true state with a high degree of accuracy
  - By acquiring info, half the time they will determine that the state is in fact 2, in which case b is the best option

  $\Rightarrow$ there is potentially a 'spillover' effect of adding c to the choice set which is to increase the probability of selecting b

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

Experiment 1.1: Testing for Responsive Attention

- How can Rational Inattention model explain this result?
  - In DP1, the incentive for gathering information is low, and the subject can simply choose *a* which guarantees 50 points
  - However, in DP2, with the addition of the option c, the subject may want to identify the true state with a high degree of accuracy
  - By acquiring info, half the time they will determine that the state is in fact 2, in which case b is the best option

⇒ there is potentially a 'spillover' effect of adding c to the choice set which is to increase the probability of selecting b

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

Experiment 1.1: Testing for Responsive Attention

- How can Rational Inattention model explain this result?
  - In DP1, the incentive for gathering information is low, and the subject can simply choose *a* which guarantees 50 points
  - However, in DP2, with the addition of the option c, the subject may want to identify the true state with a high degree of accuracy
  - By acquiring info, half the time they will determine that the state is in fact 2, in which case b is the best option

$\Rightarrow$ there is potentially a ‘spillover’ effect of adding c to the choice set which is to increase the probability of selecting b

	Payoffs					
DP	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$	$U(c,1)$	$U(c,2)$
1	50	50	b_1	b_2	n/a	n/a
2	50	50	b_1	b_2	100	0

Experiment 1.2: Changing Incentives

Theory: Rational Inattention Model Conditions

- Caplin and Dean [2015] provide necessary and sufficient conditions on SDSC data such that there exists some cost function which rationalizes the general model
  - Roughly speaking, SDSC data has a *costly information acquisition representation* if the data can be explained by the maximization of $G(\mu, A, \pi) - K(\mu, \pi)$
  - *No Improving Action Switches* (NIAS) ensures that choices are consistent with efficient use of whatever information the DM has
  - *No Improving Attention Cycles* (NIAC) ensures that choices of information structure itself is rationalizable according to some underlying cost function
- **Theorem** [Caplin, Dean (2015)]: SDSC data has a costly information acquisition representation iff it satisfies NIAS and NIAC.

Theory: Rational Inattention Model Conditions

- Caplin and Dean [2015] provide necessary and sufficient conditions on SDSC data such that there exists some cost function which rationalizes the general model
  - Roughly speaking, SDSC data has a *costly information acquisition representation* if the data can be explained by the maximization of $G(\mu, A, \pi) - K(\mu, \pi)$
  - *No Improving Action Switches* (NIAS) ensures that choices are consistent with efficient use of whatever information the DM has
  - *No Improving Attention Cycles* (NIAC) ensures that choices of information structure itself is rationalizable according to some underlying cost function

- **Theorem** [Caplin, Dean (2015)]: SDSC data has a costly information acquisition representation iff it satisfies NIAS and NIAC.

Theory: Rational Inattention Model Conditions

- Caplin and Dean [2015] provide necessary and sufficient conditions on SDSC data such that there exists some cost function which rationalizes the general model
  - Roughly speaking, SDSC data has a *costly information acquisition representation* if the data can be explained by the maximization of $G(\mu, A, \pi) - K(\mu, \pi)$
  - *No Improving Action Switches* (NIAS) ensures that choices are consistent with efficient use of whatever information the DM has
  - *No Improving Attention Cycles* (NIAC) ensures that choices of information structure itself is rationalizable according to some underlying cost function

- **Theorem** [Caplin, Dean (2015)]: SDSC data has a costly information acquisition representation iff it satisfies NIAS and NIAC.

Experiment 1.2: Changing Incentives

- Two equally likely states: state 1 (49 red balls); state 2 (51 red balls)
- Two actions:

	$u(\cdot\|1)$	$u(\cdot\|2)$
a	x	0
b	0	x

- DPs 3–6: x is 5, 40, 75, 90
- **NIAS**: the subject must be more likely to choose the action a in state 1
- **NIAC**: the subject becomes no less accurate as incentives increase, i.e., $P(a|1) + P(b|1)$ increases as x increases

Experiment 1.2: Changing Incentives

- Two equally likely states: state 1 (49 red balls); state 2 (51 red balls)
- Two actions:

	$u(\cdot\|1)$	$u(\cdot\|2)$
a	x	0
b	0	x

- DPs 3–6: x is 5, 40, 75, 90
- **NIAS**: the subject must be more likely to choose the action a in state 1
- **NIAC**: the subject becomes no less accurate as incentives increase, i.e., $P(a|1) + P(b|1)$ increases as x increases

Experiment 1.2: Testing NIAS and NIAC

- The aggregate data supports both NIAS and NIAC:

Table 3: Data from Experiment 1.2[45]

DP	$P_j(a\|1)$	$P_j(a\|2)$	Prob	$P_j(a\|1) + P_j(b\|2)$
3	0.74	0.40	0.00	0.67
4	0.76	0.34	0.00	0.71
5	0.78	0.33	0.00	0.72
6	0.78	0.28	0.00	0.75

Experiment 1.2: Testing NIAS and NIAC

- Experiment 1.2 Individual level:
  - 81% show no significance violations of either condition
  - 17% violate NIAC only
  - 2% violate NIAS only
  - None violates both conditions

  $\Rightarrow$ most of subjects do not have significant violations of the NIAS and NIAC conditions and therefore act as if they maximize payoffs net of some underlying attention cost function

- NIAS and NIAC in Experiment 1.1 can also be tested (though much complicated)
  - 9 out of 196 NIAS tests (5%) significantly violates
  - 2out of 28 NIAC tests (7%) significantly violates

Experiment 1.2: Testing NIAS and NIAC

- Experiment 1.2 Individual level:
  - 81% show no significance violations of either condition
  - 17% violate NIAC only
  - 2% violate NIAS only
  - None violates both conditions

  $\Rightarrow$ most of subjects do not have significant violations of the NIAS and NIAC conditions and therefore act as if they maximize payoffs net of some underlying attention cost function

- NIAS and NIAC in Experiment 1.1 can also be tested (though much complicated)
  - 9 out of 196 NIAS tests (5%) significantly violates
  - 2out of 28 NIAC tests (7%) significantly violates

Theory: Shannon Entropy Cost and Generalization

- The previous results suggest that the SDSC data has a costly information acquisition representation, but *what types of information cost* is used?

- Shannon Entropy cost:

$$K_s(\mu,\pi) = \kappa \cdot \left\{ \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma) \cdot [-H[\gamma]] - [-H[\mu]] \right\} \tag{4}$$

  where $\pi(\gamma)$ is the unconditional probability of signal γ and $H(\gamma) \equiv \sum_{\omega \in \Omega} -\gamma(\omega) \ln \gamma(\omega)$

- Uniformly posterior-separable cost: same as (4) but $-H$ is an arbitrary convex function

Theory: Shannon Entropy Cost and Generalization

- The previous results suggest that the SDSC data has a costly information acquisition representation, but *what types of information cost* is used?

- Shannon Entropy cost:

$$K_s(\mu, \pi) = \kappa \cdot \left\{ \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma) \cdot [-H[\gamma]] - [-H[\mu]] \right\} \tag{4}$$

  where $\pi(\gamma)$ is the unconditional probability of signal γ and $H(\gamma) \equiv \sum_{\omega \in \Omega} -\gamma(\omega) \ln \gamma(\omega)$

- Uniformly posterior-separable cost: same as (4) but $-H$ is an arbitrary convex function

Theory: Shannon Entropy Cost and Generalization

- The previous results suggest that the SDSC data has a costly information acquisition representation, but *what types of information cost* is used?

- Shannon Entropy cost:

$$K_s(\mu, \pi) = \kappa \cdot \left\{ \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma) \cdot [-H[\gamma]] - [-H[\mu]] \right\} \tag{4}$$

  where $\pi(\gamma)$ is the unconditional probability of signal γ and $H(\gamma) \equiv \sum_{\omega \in \Omega} -\gamma(\omega) \ln \gamma(\omega)$

- Uniformly posterior-separable cost: same as (4) but $-H$ is an arbitrary convex function

Theory: Shannon Entropy Cost and Invariant Likelihood Ratio (ILR) property

- Under the Shannon Entropy cost function, the invariant likelihood ration (ILR) property holds:
$$\frac{\gamma^a(\omega)}{\gamma^b(\omega)} = \frac{\exp(u(a,\omega)/\kappa)}{\exp(u(b,\omega)/\kappa)} \tag{5}$$
where γ^i is the posterior belief given the choice i

- Under Experiment 1.2, it is equivalent to
$$\begin{aligned}\kappa = & \frac{5}{\ln(\gamma_3^a(1)) - \ln(\gamma_3^b(1))} = \frac{40}{\ln(\gamma_4^a(1)) - \ln(\gamma_4^b(1))} \\ = & \frac{70}{\ln(\gamma_5^a(1)) - \ln(\gamma_5^b(1))} = \frac{95}{\ln(\gamma_6^a(1)) - \ln(\gamma_6^b(1))}\end{aligned}$$
where $\gamma_j^a(1)$ is the posterior probability of state 1 in DP j

Theory: Shannon Entropy Cost and Invariant Likelihood Ratio (ILR) property

- Under the Shannon Entropy cost function, the invariant likelihood ration (ILR) property holds:

$$\frac{\gamma^a(\omega)}{\gamma^b(\omega)} = \frac{\exp(u(a,\omega)/\kappa)}{\exp(u(b,\omega)/\kappa)} \tag{5}$$

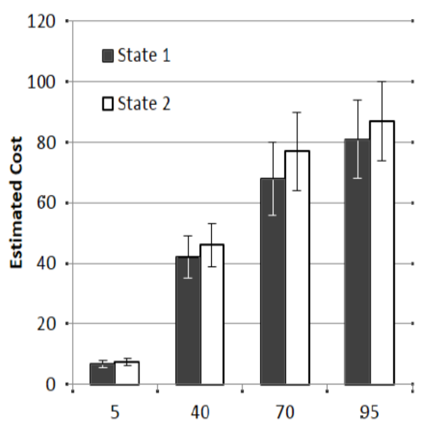
where γ^i is the posterior belief given the choice i

- Under Experiment 1.2, it is equivalent to

$$\begin{aligned}\kappa =& \frac{5}{\ln(\gamma_3^a(1)) - \ln(\gamma_3^b(1))} = \frac{40}{\ln(\gamma_4^a(1)) - \ln(\gamma_4^b(1))}\\ =& \frac{70}{\ln(\gamma_5^a(1)) - \ln(\gamma_5^b(1))} = \frac{95}{\ln(\gamma_6^a(1)) - \ln(\gamma_6^b(1))}\end{aligned}$$

where $\gamma_j^a(1)$ is the posterior probability of state 1 in DP j

Experiment 1.2: Testing ILR property



- This figure shows the extimated cost parameter κ from each decision problem
- The Shannon model predicts that these should be equal, but it is not the case.

Experiment 1.3: Changing Priors

Theory: UPS cost and LIP condition

- Uniformly posterior-separable cost:

$$K_s(\mu, \pi) = \kappa \cdot \left\{ \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma)[-H[\gamma]] - [-H[\mu]] \right\} \tag{6}$$

  where $\pi(\gamma)$ is the unconditional probability of signal γ and $-H$ is an arbitrary convex function

- Locally Invariant Posteriors (LIP) condition
  - Local changes in prior beliefs do not lead to changes in optimal posterior beliefs.

Theory: UPS cost and LIP condition

- Uniformly posterior-separable cost:

$$K_s(\mu, \pi) = \kappa \cdot \left\{ \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma)[-H[\gamma]] - [-H[\mu]] \right\} \tag{6}$$

  where $\pi(\gamma)$ is the unconditional probability of signal γ and $-H$ is an arbitrary convex function

- Locally Invariant Posteriors (LIP) condition
  - Local changes in prior beliefs do not lead to changes in optimal posterior beliefs.

Experiment 1.3: Changing Priors

- Two states: state 1 (47 red balls); state 2 (53 red balls)[1]
- Two actions:

	$u(\cdot\|1)$	$u(\cdot\|2)$
a	10	0
b	0	10

- DPs 7–10: $\Pr(s = 1)$ is .5, .6, .75, .85
- As the prior probability of state 1 increases, there are two possibilities:
    1. if the prior remains inside the convex hull, the subject must use the same posterior
    2. if the prior moves outside the convex hull, the subject should learn nothing

[1] Consider easier setting to ensure that more subjects collected some information

Experiment 1.3: Changing Priors

- Two states: state 1 (47 red balls); state 2 (53 red balls)[1]

- Two actions:

	$u(\cdot\|1)$	$u(\cdot\|2)$
a	10	0
b	0	10

- DPs 7–10: $\Pr(s = 1)$ is .5, .6, .75, .85

- As the prior probability of state 1 increases, there are two possibilities:
  1. if the prior remains inside the convex hull, the subject must use the same posterior
  2. if the prior moves outside the convex hull, the subject should learn nothing

[1] Consider easier setting to ensure that more subjects collected some information

Experiment 1.3: Changing Priors

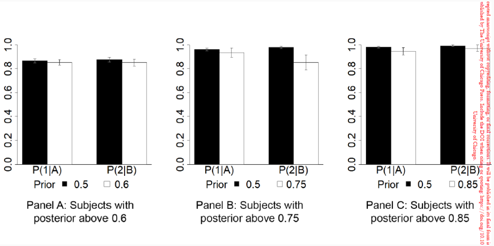
- LIP Test 1: subjects with $\gamma_7^a(1) < \mu_i(1)$ should exclusively choose action *a*, while those with $\gamma_7^a(1) > \mu_i(1)$ should choose both *a* and *b*
  - $\gamma_7^a(1)$: the posterior that the state is 1 revealed in DP 7 (prior was .5)
  - $\mu_i(1)$: the prior belief that the state is 1 in DP *i*

	$\mu(1)$		
	DP8	DP9	DP10
	0.6	0.75	0.85
$\gamma_7^a(1) < \mu_i(1)$	33%	46%	41%
$\gamma_7^a(1) \geq \mu_i(1)$	3%	10%	14%

Testing 'No Learning' prediction: Fraction of subjects who never choose *b*

Experiment 1.3: Changing Priors

- LIP Test 2: subjects who are predicted to be gathering information should use the same posteriors as they did in DP 7
- The following figure shows that data is relatively well described by LIP prediction
  - Of the six comparisons, only one shows a significant difference at the 10% level
  - However, a test of the joint hypothesis that all six conditions hold simultaneously is rejected at the 5% level



Testing 'Invariant Posterior' prediction

Conclusion

Next Step

- In this paper, the authors provide extensive experiments testing the foundations of rational inattention model
- There is an emerging literature in applied theory using the rational inattention model
- Given that this is the seminal experimental paper in the rational inattention literature, I believe that there will be a series of experimental works to be done

Experimental Design 2

Experimental Design: Setup

- **State**: number of correct simple equations

There is a 50% chance of 4 correct equations.
There is a 50% chance of 3 correct equations.

42+19=51	38+6=44	38+39=86
9+8=8	18+2=20	
41+37=78	28+15=50	