Rational Inattention Lecture 1

Mark Dean

Behavioral Economics G6943
Fall 2017

The Story So Far.....

- (Hopefully) convinced you that attention costs are important
- Introduced the concept of consideration sets
  - Along with sequential search and satisficing
- Showed that the model did a reasonable job in some circumstances
- But, there is something restrictive about consideration sets
  - Items are either in the consideration set and fully understood
  - Or outside the consideration set, and nothing is learned
- Seems like a good model for choice over a large number of simple alternatives
- Not for a small number of complex alternatives

A Non-Satisficing Situation

Retirement Calculators & Tools

Our calculators & tools will help you take the guesswork out of saving for retirement and assist in building an income strategy to meet your needs.

Questions?

800-343-3548

Chat with a representative.

Find an Investor Center.

Take the guesswork out of retirement savings

Saving for retirement

Get quick answers on how to get your savings started and how much you may need to save.

Living on retirement income

Learn how to make your income last throughout your retirement years.

Investing with an IRA

Learn how to choose, contribute to, and take distributions from an individual retirement account (IRA).

All retirement calculators & tools

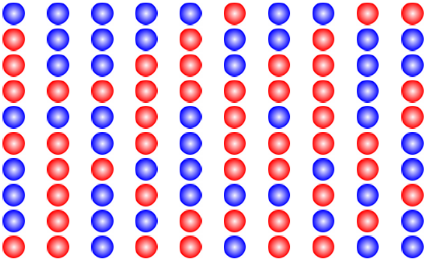
Name	Description
Roth Conversion Evaluator	Estimate the potential benefits of converting to a Roth IRA, explore how much to convert and how to pay for it.
IRA Contribution Calculator	Estimate how much you may be able to contribute to your IRA this year and how much might be tax-deductible.
Minimum Required Distribution (MRD) Calculator	Determine your required retirement account withdrawals after age 70½.
myPlan® Monitor	Get suggested action steps to help your investment plan stay on track based on your target asset mix and goals.
Guaranteed Income Estimator	Get an estimate for guaranteed income payments you can receive through a fixed income annuity (guarantees are subject to the claims-paying ability of the issuing insurance company).
Roth vs. Traditional IRA Evaluator	Answer a few questions to find out which type of IRA is right for you.
Retirement Income Planner	If you're within five years of retirement or already retired, use this tool to create a comprehensive plan to help make your money last by projecting out your expenses against potential income sources in different kinds of market conditions.
Savings Planner	Take a few minutes to tell us your priorities and we'll help you create a plan to save and pay down debt.
Portfolio Review	Identify a target asset mix that aligns with your goals. Then easily implement an investment strategy that will help you stay on track.
myPlan Snapshot®	Answer five easy questions for a quick estimate of how much you may need to save for retirement compared with your projected savings.
Fidelity Income Strategy Evaluator®	Find a right mix of income-producing investments to help meet your needs in retirement.
Retirement Quick Check	If you're more than five years from retirement, use this tool to estimate how much you'll need, compare it to what you're on track to have, and identify changes you can make today to address any shortfall.

Stay Connected

Locate an Investor Center by ZIP Code

Search

A Non-Satisficing Situation



Act	Payoff 47 red dots	Payoff 53 red dots
a	20	0
b	0	10

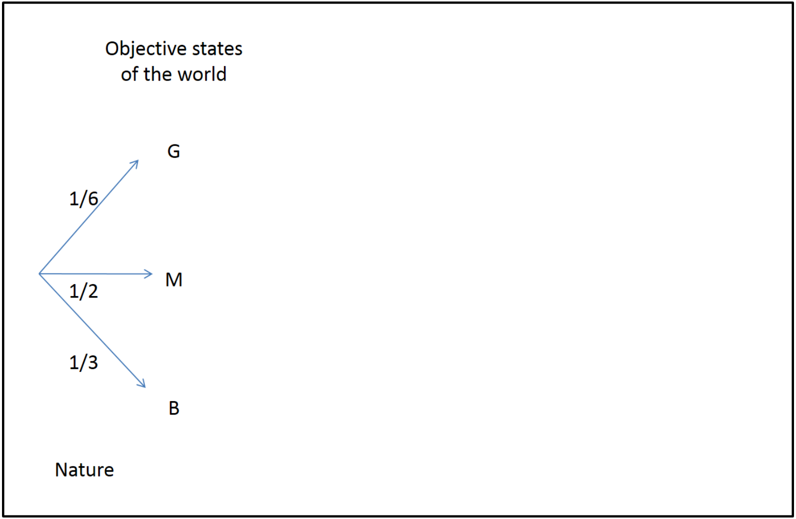
Set Up

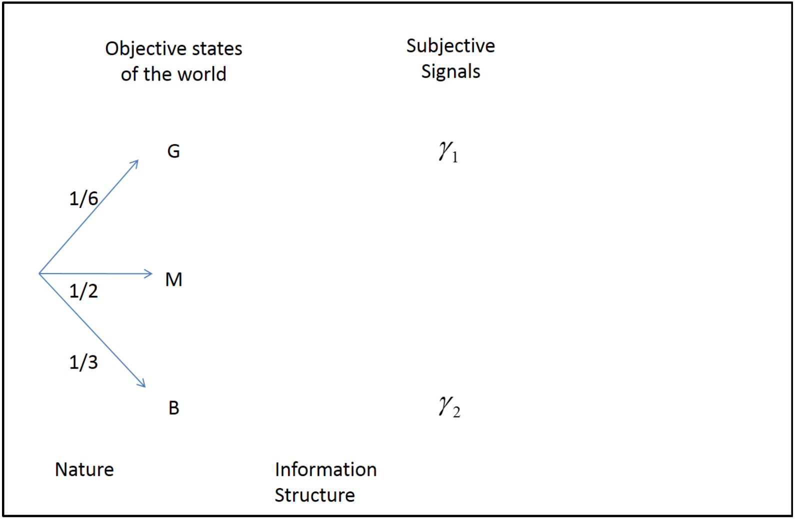
- Objective states of the world
  - e.g. Demand could be 'good', 'medium' or 'bad'
- Decision maker chooses an action
  - e.g. Set price to be high, medium, or low
- Gross payoff depends on action and state
  - e.g. Quantity sold depends on price and demand
- Decision maker get to learn something about the state before choosing action
  - e.g. Could do market research, focus groups, etc.
- Can choose what to learn conditional on the problem

The Choice Problem

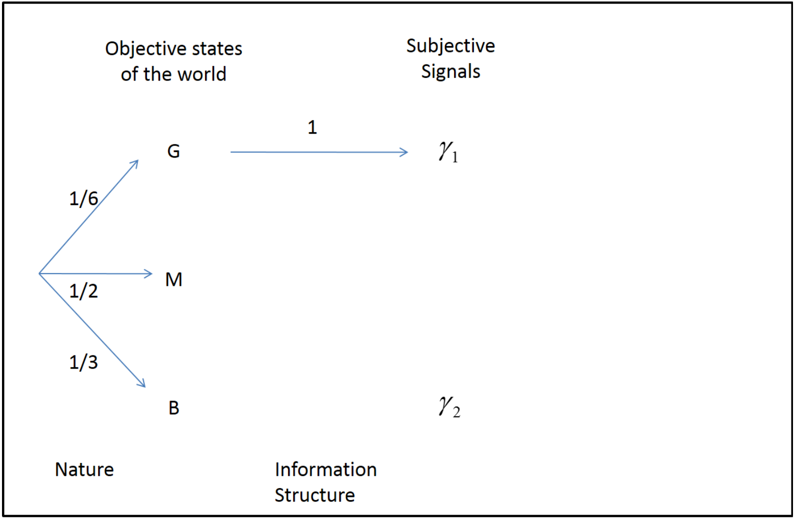
- The specifics of the process of information acquisition may be very complex
- We model the choice of information in an *abstract* way
- The decision maker chooses an *information structure*
    - Set of signals to receive
    - Probability of receiving each signal in each state of the world
- Choose action conditional on signal received
- Value of strategy given by
    - Expected value of actions taken given posterior beliefs
    - Minus cost of information
- Notice that this is an optimizing model with additional constraints
    - Subjects respond to costs and incentives
    - At least an interesting benchmark

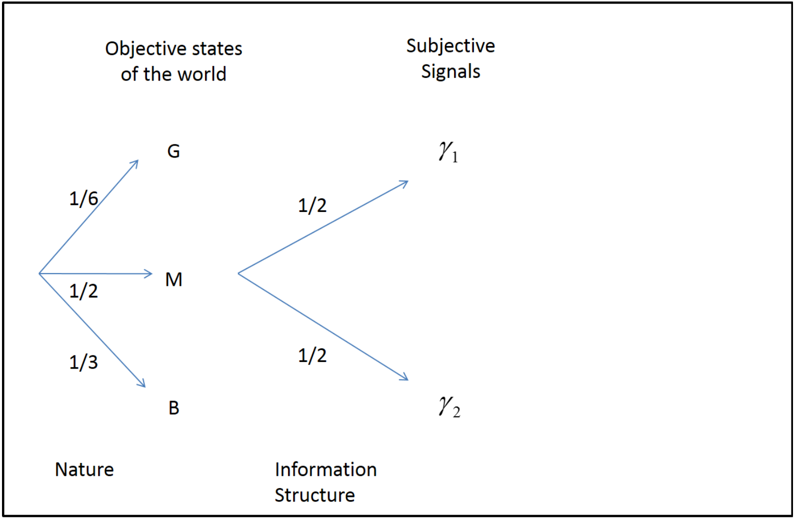
The Choice Problem



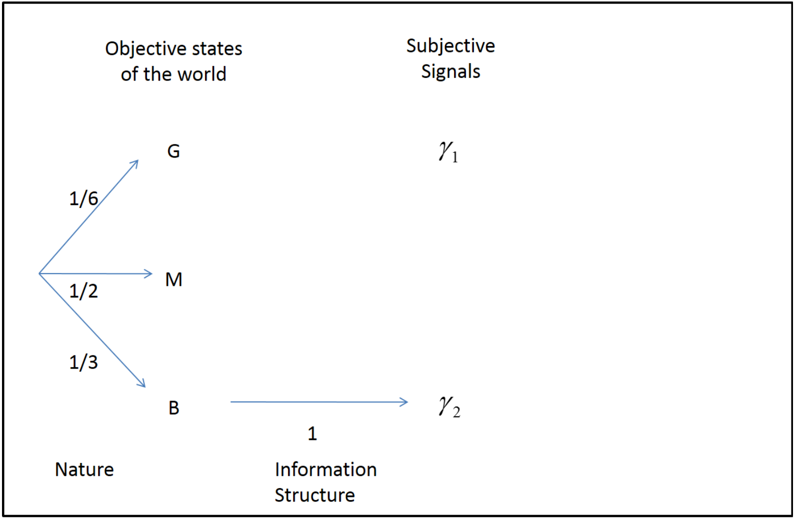

Objective states of the world
Subjective Signals
G
γ_1
1/6
M
1/2
1/3
B
γ_2
Nature
Information Structure

The Choice Problem

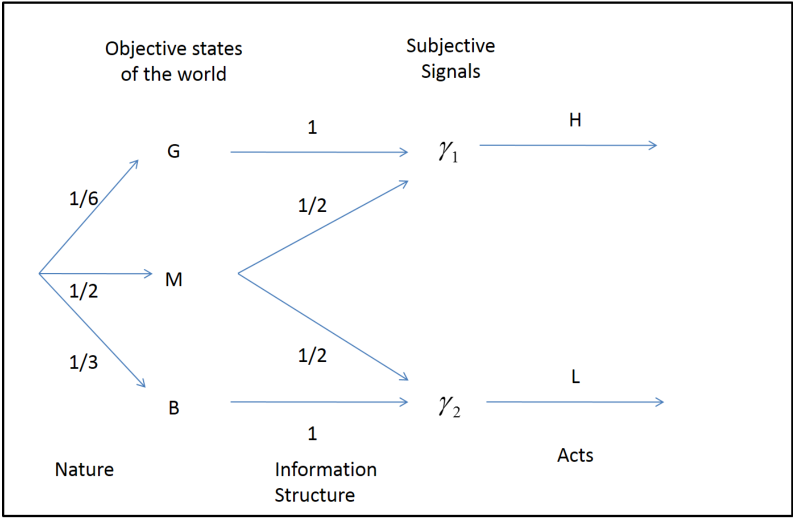


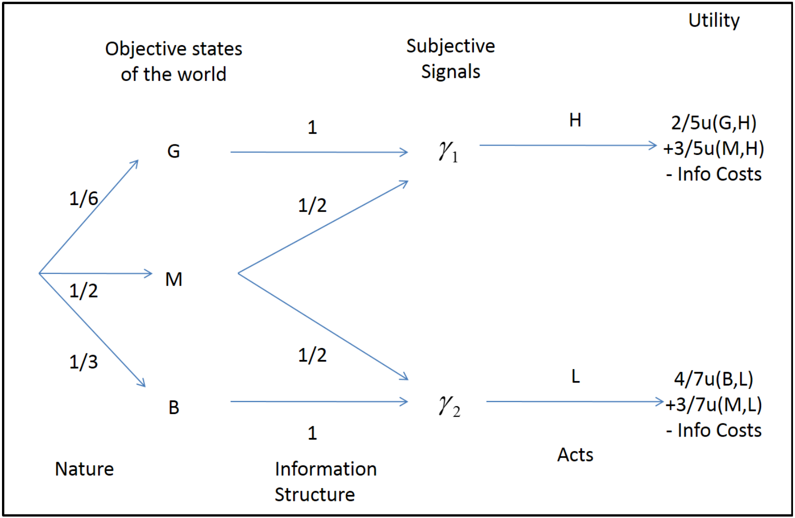

Objective states of the world
Subjective Signals
G
M
B
γ_1
γ_2
1/6
1/2
1/3
1/2
1/2
Nature
Information Structure

The Choice Problem



The Choice Problem




Utility
Objective states of the world
Subjective Signals
1
H
2/5u(G,H)
+3/5u(M,H)
- Info Costs
G
γ_1
1/6
1/2
M
1/2
1/2
1/3
B
γ_2
L
4/7u(B,L)
+3/7u(M,L)
- Info Costs
1
Acts
Nature
Information Structure

- Note that most 'real world' information gathering activities can be thought of in terms of as generating information structures
- E.g., say that you have developed a new economics class
- There are two possible states of the world
  - Class is good - $\frac{2}{3}$ of people like it on average
  - Class is bad - $\frac{1}{3}$ of people like it on average
- Each is equally likely
- Release a survey in which all 12 members of the class report if they like the class or not
- This generates an information structure
  - 13 signals: 0,1,2..... people say they like the class
  - Probability of each signal given each state of the world can be calculated

Set Up

- Ω: Objective states of the world (finite)
  - with prior probabilities μ
- a : An action - utility depends on the state
  - $U(a, \omega)$ utility of action a in state ω
  - $\mathcal{A}$: Set of actions:
- $A \subset \mathcal{A}$: Decision problem (finite)

- For each decision problem
  1. Choose information structure (π)
     - Defined by:
       - Set of signals: $\Gamma(\pi)$
       - Probability of receiving each signal γ from each state ω : $\pi(\gamma|\omega)$
  2. Choose action conditional on signal received (C)
     - $C(\gamma)$ probability distribution over actions given signal γ
- In order to maximize
  - Expected value of actions taken given posterior beliefs
  - Minus cost of information K

$$\sum_{\Omega} \mu(\omega) \sum_{\gamma \in \Gamma(\pi)} \pi(\gamma|\omega) \left(\sum_{a \in A} C(a|\gamma) U(a(\omega)) \right) - K(\mu, \pi)$$

The Value of An Information Structure

- What is the value of an information structure?
- In the end you will have to choose an *action*
  - Defined by the outcome it gives in each state of the world
- Assume in previous example, could choose three actions
  - set price H, A or L
- The following table could describe the profits each price gives at each demand level

	Price		
State	H	A	L
G	10	3	1
M	1	2	1
B	-10	-3	-1

The Value of An Information Structure

- What would you choose if you gathered no information?
  - i.e. if you had your prior beliefs

$$\mu(G) = \frac{1}{6}, \mu(M) = \frac{1}{2}, \mu(B) = \frac{1}{3}$$

- Calculate the expected utility for each act

$$\begin{aligned}
\frac{1}{6}u(H,G) + \frac{1}{2}u(H.M) + \frac{1}{3}u(H,B) &= \frac{-7}{6} \\
\frac{1}{6}u(A,G) + \frac{1}{2}u(A,M) + \frac{1}{3}u(A,B) &= \frac{1}{2} \\
\frac{1}{6}u(L,G) + \frac{1}{2}u(L,M) + \frac{1}{3}u(L,B) &= \frac{1}{3}
\end{aligned}$$

- Choose A
- Get utility $\frac{1}{2}$

The Value of An Information Structure

- What would you choose upon receiving signal R?
- Depends on beliefs conditional on receiving that signal
- Can calculate this using Bayes Rule

$$
\begin{aligned}
P(G|R) &= \frac{P(G \cap R)}{P(R)} \\
&= \frac{\mu(G)\pi(R|G)}{\mu(G)\pi(R|G) + \mu(M)\pi(R|M) + \mu(B)\pi(R|B)} \\
&= \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{4} + 0} = \frac{2}{5}
\end{aligned}
$$

The Value of An Information Structure

- We can therefore calculate posterior beliefs conditional on signal R

$$\begin{aligned} P(G|R) &= \frac{2}{5} = \gamma^R(G) \\ P(M|R) &= \frac{3}{5} = \gamma^R(M) \\ P(B|R) &= 0 = \gamma^R(B) \end{aligned}$$

- Where we use $\gamma^R(\omega)$ to mean the probability that the state of the world is ω given signal R

- And calculate the value of choosing each act given these beliefs

$$\begin{aligned}
\frac{2}{5}u(H,G)+\frac{3}{5}u(H,M) &= \frac{23}{5}\\
\frac{2}{5}u(A,G)+\frac{3}{5}u(A,M) &= \frac{12}{5}\\
\frac{2}{5}u(L,G)+\frac{3}{5}u(L,M) &= \frac{2}{5}
\end{aligned}$$

The Value of An Information Structure

- If received signal R, would choose H and receive $\frac{23}{5}$
- By similar process, can calculate that if received signal S
  - Choose L and receive $-\frac{1}{7}$
- Can calculate the value of the information structure as

$$\begin{aligned} P(R)\frac{23}{5} + P(S)\frac{-1}{7} &= \\ \frac{5}{12}\frac{23}{5} + \frac{7}{12}\frac{-1}{7} &= \frac{11}{6} \end{aligned}$$

- How much would you pay for this information structure?

The Value of An Information Structure

- Value of this information structure is $\frac{11}{6}$
- Value of being uninformed is $\frac{1}{2}$
- Would prefer this information structure to being uninformed if cost is below $\frac{8}{6}$
- Note that the value of an information structure depends on the acts available

$$
\begin{aligned}
G(\pi, A) &= \sum_{\gamma \in \Gamma(\pi)} P(\gamma) g(\gamma, A) \\
g(\gamma, A) &= \max_{a \in A} \sum_{\omega \in \Omega} \gamma(\omega) u(a, \omega)
\end{aligned}
$$

- $g(\gamma, A)$ value of receiving signal γ if available actions are A
  - Highest utility achievable given the resulting posterior beliefs

- Easy to calculate the *value* of an information structure

$$
\begin{aligned}
& G(A,\pi) \\
= \quad & \max_{C:\Gamma(\pi)\rightarrow\Delta(A)} \sum_{\Omega} \mu(\omega) \sum_{\gamma\in\Gamma(\pi)} \pi(\gamma|\omega) \left(\sum_{a\in A} C(a|\gamma) U(a,\omega) \right)
\end{aligned}
$$

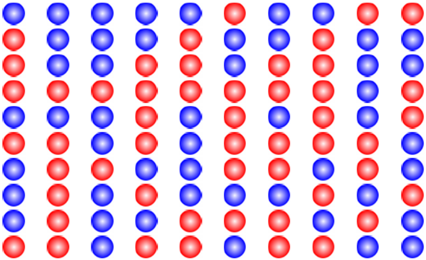
- But what is the correct information processing technology?
  - Choose variance of normal signal (e.g. Verrecchia 1982)?
  - Shannon mutual information costs (e.g. Sims 1998)?
  - Choose from set of available partitions (e.g. Ellis 2012)?
  - Sequential search (e.g. McCall 1970)?
- As usual, have two possible approaches
  1. Make further assumptions
  2. Ask if there is *any* cost function that can explain the data
- Today we take approach 2
- Next week we will follow approach 1

A Caveat

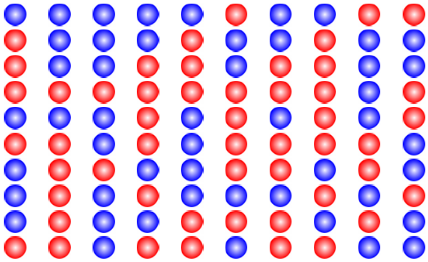
- We **will** assume throughout that costs are additively separable from utilities
- Is this assumption restrictive?
- Yes - see Chambers, Christopher P., Ce Liu, and John Rehbeck. "Nonseparable Costly Information Acquisition and Revealed Preference."
- Can you think of cases in which non-separability might be an important feature?

Data

- Let D be a collection of decision problems
- What could we observe?
- Standard choice data
  - $C(A)$: what is chosen from A
- Stochastic choice data
  - $P_A(a)$: probability of choosing alternative a
- **State dependent stochastic choice data** P_A
  - $P_A(a|\omega)$ probability of choosing action a conditional on state ω
- Also assume we observe:
  - Prior probabilities μ
  - Utilities U
- Do **not** observe
  - Information structures π_A
  - Subjective signals γ
  - Information costs K



- Subjects presented with 100 balls
- State is determined by the number of red balls
- Prior distribution of red balls known to subject



Action	Payoff 49 red balls	Payoff 51 red balls
a	10	0
b	0	10

- No time limit: trade off between effort and financial rewards

An Experimental Example

- Data: State dependant stochastic choice
  - Probability of choosing each action in each objective state of the world

Action	**State = 49 red balls**	**State = 51 red balls**
Prob choose a	$P(a\|49)$	$P(a\|51)$
Prob choose b	$P(b\|49)$	$P(b\|51)$

- Observe subject making same choice 50 times
- Can use this to estimate P_A
  - But we will not be able to observe P_A perfectly
  - Will only be able to make probabilistic statements
- Can collect this type of data in the lab
  - What about outside?

- What type of stochastic choice data $\{D, P\}$ is consistent with optimal information acquisition?
- i.e. there exists a cost function K
- For each decision problem $A \in D$ an information structure π_A and choice function C_A s.t.
    - C_A is optimal for each γ
    - π_A is optimal given K
    - C_A and π_A are consistent with P_A

$$P_A(a|\omega) = \sum_{\gamma \in \Gamma(\pi_A)} \pi_A(\gamma|\omega) C_A(a|\gamma).$$

- What 'mistakes' are consistent with optimal behavior in the face of information costs?

Notes

- This approach is very flexible
  - No in principle restriction on information structures
  - No restrictions on costs
- Nests other models of information acquisition
  - e.g. Shannon Mutual Information set costs to

$$K(\pi) = \lambda E\left(\log \frac{\mu(\omega)\pi(\gamma|\omega)}{\mu(\omega)\pi(\gamma)}\right)$$

- Can mimic a hard constraints
  - e.g. a model in which subjects choose the variance of a normal signal, set the cost of all other information structures to ∞

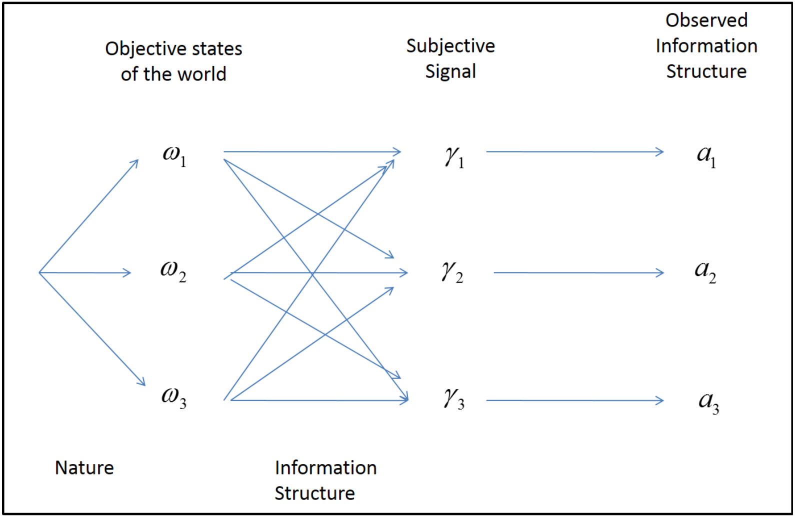
Observing Information Structures

- Key observation: State dependent stochastic choice data tells us a lot about the information structure a decision maker has used
- Assume that decision maker is 'well behaved'
  - Chooses each action in response to at most one signal
  - No mixed strategies - one action per signal
- Information structure can be observed directly from state dependent stochastic choice
  - For each chosen action a there is an associated signal $\bar{\gamma}^a$
  - Probability of signal $\bar{\gamma}^a$ in state ω is the same as the probability of choosing a in ω

$$\bar{\pi}(\bar{\gamma}^a|\omega) = P(a|\omega)$$

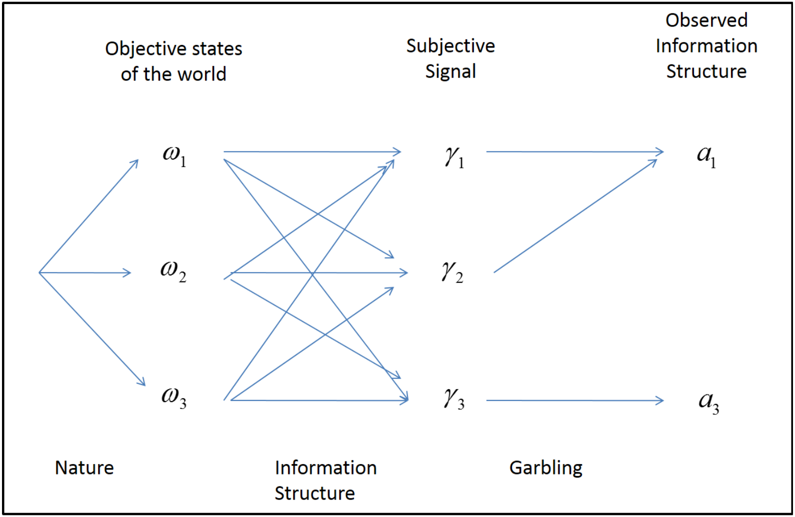
- Call $\bar{\pi}$ the 'revealed information structure'

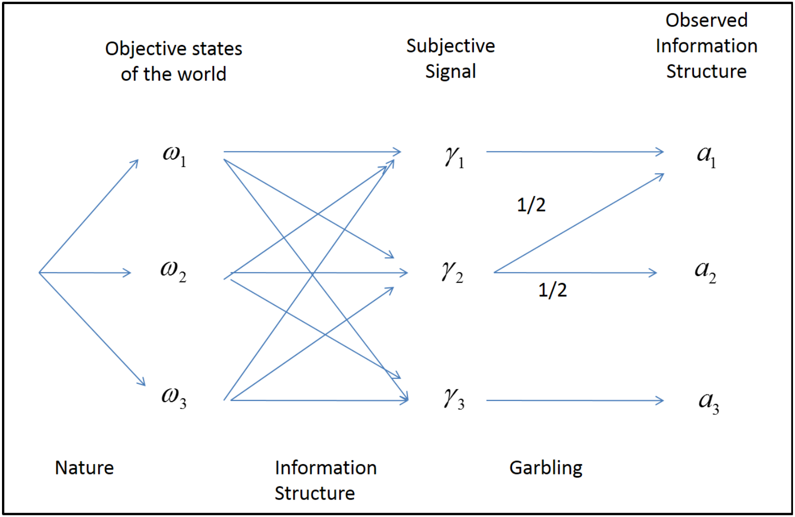
Recovering Attention Strategy



- What if decision maker is not well behaved?
  - Chooses some act in more than one subjective state
  - Mixed strategies - more than one act in an subjective state

Same Act in Different States




Objective states
of the world
Subjective
Signal
Observed
Information
Structure
ω_1
ω_2
ω_3
γ_1
γ_2
γ_3
a_1
a_2
a_3
1/2
1/2
Nature
Information
Structure
Garbling

Observing Information Structures

- Can still recover revealed information structure $\bar{\pi}$
- Not necessarily the same as true information structure π
- But will be a garbling of the true information structure
  - i.e. π is statistically sufficient for $\bar{\pi}$
- There exists a stochastic $|\Gamma(\pi)| \times |\Gamma(\bar{\pi})|$ matrix B such that if we
  - Apply π
  - For each state γ^i move to state $\bar{\gamma}^j$ with probability B^{ij}
  - We obtain $\bar{\pi}$
- i.e.

$$\begin{aligned} \sum_j B^{ij} &= 1 \; \forall \; j \\ \bar{\pi}(\bar{\gamma}^j|\omega) &= \sum_i B^{ij} \pi(\gamma^i|\omega) \; \forall \; j \end{aligned}$$

- Intuition: SDSC data cannot be more informative than the signal that created it

- Recall $G(A, \pi)$ is the *gross value* of using information structure π in decision problem A

$$
\begin{aligned}
& G(A, \pi) \\
= \quad & \max_{C:\Gamma(\pi)\rightarrow\Delta(A)} \sum_{\Omega} \mu(\omega) \sum_{\gamma\in\Gamma(\pi)} \pi(\gamma|\omega) \left(\sum_{a\in A} C(a|\gamma) U(a(\omega)) \right)
\end{aligned}
$$

- An information structure π is sufficient for information structure π' if and only if

$$G(A, \pi) \geq G(A, \pi') \ \forall \ A$$

Observing Information Structures

- $\bar{\pi}$ may not be the agent's true information structure
    - But the true information structure π must be sufficient for $\bar{\pi}$
    - π will be at least as valuable as $\bar{\pi}$ in any decision problem
- Turns out that this is all we need

Characterizing Rational Inattention

- Choice of act optimal given attentional strategy
- Choice of attention strategy optimal

Characterizing Rational Inattention

- Choice of act optimal given attentional strategy
- Choice of attention strategy optimal

Action	Payoff 49 red balls	Payoff 51 red balls
$\mathbf{a}^1$	20	0
$\mathbf{b}^1$	0	10

Prior: $\{0.5, 0.5\}$

Action	State = 49 red balls	State = 51 red balls
Prob choose a	$\frac{1}{2}$	$\frac{1}{3}$
Prob choose b	$\frac{1}{2}$	$\frac{2}{3}$

- Posterior probability of 49 red balls when action b was chosen

$$\begin{aligned}\Pr(\omega &= 49|b \text{ chosen}) = \frac{\Pr(\omega = 49, b \text{ chosen})}{\Pr(b \text{ chosen})} \\ &= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{2}{6}} = \frac{3}{7}\end{aligned}$$

- But for this posterior

$$\begin{aligned}\frac{3}{7}U(a(49)) + \frac{4}{7}U(a(51)) &= \frac{3}{7}20 + \frac{4}{7}0 = 8.6 \\ \frac{3}{7}U(b(49)) + \frac{4}{7}U(b(51)) &= \frac{3}{7}0 + \frac{4}{7}10 = 5.7\end{aligned}$$

- To avoid such cases requires

$$a \in \arg\max_{a \in A} \sum_{\Omega} \Pr(\omega|a) U(a(\omega))$$

- Which implies

Condition 1 (No Improving Action Switches) For every chosen action a

$$\sum \mu(\omega) P_A(a|\omega) \left[u(a(\omega)) - u(b(\omega))\right] \geq 0.$$

for all $b \in A$

- If $\bar{\pi}$ not true information structure, condition still holds
  - a optimal at all posteriors in which it is chosen
  - Must also be optimal at convex combination of these posteriors

- Choice of act optimal given attentional strategy
- Choice of attention strategy optimal

Action	Payoff 49 red balls	Payoff 51 red balls
$\mathbf{a}^1$	10	0
$\mathbf{b}^1$	0	10

Prior: $\{0.5, 0.5\}$

Action	State = 49 red balls	State = 51 red balls
Prob choose a	$\frac{3}{4}$	$\frac{1}{4}$
Prob choose b	$\frac{1}{4}$	$\frac{3}{4}$

Action	**Payoff 49 red balls**	**Payoff 51 red balls**
$\mathbf{a}^2$	20	0
$\mathbf{b}^2$	0	20

Prior: $\{0.5, 0.5\}$

Action	**State = 49 red balls**	**State = 51 red balls**
Prob choose a	$\frac{2}{3}$	$\frac{1}{3}$
Prob choose b	$\frac{1}{3}$	$\frac{2}{3}$

Optimal Choice of Attention Strategy

- $G(A, \pi)$ is the gross value of using information structure π in decision problem A

G	$\bar{\pi}^1$	$\bar{\pi}^2$
$\{a^1, b^1\}$	$7\frac{1}{2}$	$6\frac{2}{3}$
$\{a^2, b^2\}$	15	$13\frac{1}{3}$

- Cost function must satisfy

$$
\begin{aligned}
G(\{a^1, b^1\}, \pi^1) - K(\pi^1) &\geq G(\{a^1, b^1\}, \pi^2) - K(\pi^2) \\
G(\{a^2, b^2\}, \pi^2) - K(\pi^2) &\geq G(\{a^2, b^2\}, \pi^1) - K(\pi^1)
\end{aligned}
$$

- Which implies

$$
\frac{5}{6} = G(\{a^1, b^1\}, \pi^1) - G(\{a^1, b^1\}, \pi^2) \geq
$$

$$
K(\pi^1) - K(\pi^2) \geq
$$

$$
G(\{a^2, b^2\}, \pi^1) - G(\{a^2, b^2\}, \pi^2) = 1\frac{2}{3}
$$

Optimal Choice of Attention Strategy

- Surplus must be maximized by correct assignments

$$G(\{a^1, b^1\}, \pi^1) + G(\{a^2, b^2\}, \pi^2)$$
$$\geq G(\{a^1, b^1\}, \pi^2) + G(\{a^2, b^2\}, \pi^1)$$

- What if $\bar{\pi} \neq \pi$?
- We know that revealed and true information structure must give same value in DP it was observed

$$G(A^i, \bar{\pi}^i) = G(A^i, \pi^i)$$

- Also, as π weakly Blackwell dominates $\bar{\pi}$

$$G(A^i, \bar{\pi}^j) \leq G(A^i, \pi^j)$$

Optimal Choice of Attention Strategy

- Surplus must be maximized by correct assignments

$$G(\{a^1, b^1\}, \bar{\pi}^1) + G(\{a^2, b^2\}, \bar{\pi}^2)$$
$$\geq G(\{a^1, b^1\}, \pi^2) + G(\{a^2, b^2\}, \pi^1)$$

- What if $\bar{\pi} \neq \pi$?
- We know that revealed and true information structure must give same value in DP it was observed

$$G(A^i, \bar{\pi}^i) = G(A^i, \pi^i)$$

- Also, as π weakly Blackwell dominates $\bar{\pi}$

$$G(A^i, \bar{\pi}^j) \leq G(A^i, \pi^j)$$

Optimal Choice of Attention Strategy

- Surplus must be maximized by correct assignments

$$G(\{a^1, b^1\}, \bar{\pi}^1) + G(\{a^2, b^2\}, \bar{\pi}^2)$$
$$\geq G(\{a^1, b^1\}, \bar{\pi}^2) + G(\{a^2, b^2\}, \bar{\pi}^1)$$

- What if $\bar{\pi} \neq \pi$?
- We know that revealed and true information structure must give same value in DP it was observed

$$G(A^i, \bar{\pi}^i) = G(A^i, \pi^i)$$

- Also, as π weakly Blackwell dominates $\bar{\pi}$

$$G(A^i, \bar{\pi}^j) \leq G(A^i, \pi^j)$$

- To guarantee the existence of a cost function requires a stronger condition

Condition 2 (No Improving Attention Cycles) For an observed sequence of decision problems $A^1 ... A^K$ and associated revealed information structures $\bar{\pi}^1 ... \bar{\pi}^K$

$$\begin{aligned} & G(A^1, \bar{\pi}^1) - G(A^1, \bar{\pi}^2) \\ & + G(A^2, \bar{\pi}^2) - G(A^2, \bar{\pi}^3) \\ & + ... \\ & + G(A^K, \bar{\pi}^K) - G(A^K, \bar{\pi}^1) \\ \geq & \ 0 \end{aligned}$$

- Note that this condition relies only on observable objects

Theorem

For any data set $\{D, P\}$ the following two statements are equivalent

1. $\{D, P\}$ *satisfy NIAS and NIAC*
2. *There exists a* $K : \Pi \to \mathbb{R}$, $\{\pi^A\}_{A \in D}$ *and* $\{C^A\}_{A \in D}$ *such that* π^A *and* $C^A : \Gamma(\pi^A) \to A$ *are optimal and generate* P^A *for every* $A \in D$

Proof.

$2 \to 1$ *Trivial*

$1 \to 2$ *Rochet [1987] (literature on implementation)* $\square$

Proof

- This problem is familiar from the implementation literature
- Say there were a set of environments $X_1....X_N$ and actions $B_1....B_M$ such that the utility of each environment and each state is given by
$$u(X_i, B_j)$$
- Say we want to implement a mechanism such that action $Y(X_i)$ is taken at in each environment.
- We need to find a taxation scheme $\tau : B_1....B_M \rightarrow \mathbb{R}$ such that
$$\begin{aligned} u(X_i, Y(X_i)) - \tau(Y(X_i)) &\geq u(X_i, B) - \tau(B) \\ &\quad \forall\ B_1....B_M \end{aligned}$$
- This is the same as our problem.

Proof

- Our problem is equivalent to finding $\theta : D \to \mathbb{R}$, such that, for all A_i, $A_j \in D$

$$G(A_i, \pi^i) - \theta(A_i) \geq G(A_i, \pi^j) - \theta(A_j)$$

  - Just define $K(\pi) = \theta(A_i)$ if $\pi = \pi^i$ for some i, or $= \infty$ otherwise

- We can apply a proof from Rockerfellar [1970] to show that NIAC gives us this condition

Proof

- Pick some arbitrary A_0 and define

$$T(A) = \sup_{\text{all chains } A_0 \text{ to } A = A_M} \sum_{n=0}^{M-1} G(A_{i+1}, \pi^i) - G(A_i, \pi^i)$$

- NIAC implies that $T(A_0) = 0$
- Also note that

$$T(A_0) \geq T(A_i) + G(A_0, \pi^i) - G(A_i, \pi^i)$$

- So $T(A_i)$ is bounded

Proof

- Furthermore, for any A_i A_j we have

$$T(A_i) \geq T(A_j) + G(A_i, \pi^j) - G(A_j, \pi^j)$$

- So, setting $\theta(A_j) = G(A_j, \pi^j) - T(A_j)$, we get

$$G(A_i, \pi^i) - \theta(A_i) \geq G(A_i, \pi^j) - \theta(A_j)$$

- So far we have been completely agnostic about the cost function
- Perhaps we want to impose some more structure
  - e.g. information structure that are more (Blackwell) Informative are (weakly) more expensive
- Turns out we get this 'for free'
- Say we observe π^A in A and π^B in B such that π^A is sufficient for π^B
- It must be the case that

$$\begin{aligned} G(B,\pi^B) - K(\pi^B) &\geq G(B,\pi^A) - K(\pi^A) \Rightarrow \\ K(\pi^A) - K(\pi^B) &\geq G(B,\pi^A) - G(B,\pi^B) \end{aligned}$$

- But by Blackwell's theorem

$$G(B,\pi^A) \geq G(B,\pi^B)$$

Restrictions on the Cost Function

- Any behavior that can be rationalized can be rationalized with a cost function that
  - Is weakly monotonic with respect to Blackwell?
  - Allows mixing
  - Positive with free inattention
- Reminiscent of Afriat's theorem
- Can also extend to 'sequential rational inattention'

- Say $\bar{\pi}^A$ is the revealed attn. strategy in decision problem A.
- Assuming weak monotonicity, it must be that

$$K(\bar{\pi}^A) - K(\pi) \leq G(A, \bar{\pi}^A) - G(A, \pi)$$

- If $\bar{\pi}^B$ is used in decision problem B then we can bound relative costs

$$G(B, \bar{\pi}^A) - G(B, \bar{\pi}^B) \leq K(\bar{\pi}^A) - K(\bar{\pi}^B) \leq G(A, \bar{\pi}^A) - G(A, \bar{\pi}^B)$$

- Tighter bounds can be obtained using chains of observations

$$\begin{aligned} & \max_{\{A^1 \ldots A^n \in D | A^1 = B, A^n = A\}} \sum \left[G(A^i, \bar{\pi}^{A^i}) - G(A^i, \bar{\pi}^{A^{i+1}}) \right] \\ \leq \quad & K(\bar{\pi}^A) - K(\bar{\pi}^B) \\ \leq \quad & \min_{\{A^1 \ldots A^n \in D | A^1 = A, A^n = B\}} \sum \left[G(A^i, \bar{\pi}^{A^i}) - G(A^i, \bar{\pi}^{A^{i+1}}) \right] \end{aligned}$$

What If Utility and Priors Are Unobservable?

- Can add 'there exists' to the statement of the NIAS and NIAC conditions
- Data has an optimal costly attention representation if **there exists** $\mu \in \Delta(\Omega)$ and $U : X \to \mathbb{R}$ such that
  - NIAS is satisfied
  - NIAC is satisfied
- If μ is known but U is unknown, conditions are linear and (relatively) easy to check
- If μ and U are unknown, conditions are harder to check
  - Still not vacuous
- Alternatively, can enrich data so that these objects can be recovered

Rational Inattention vs Random Utility

- Alternative model of random choice: Random Utility
  1. Agent receives some information about the state of the world
  2. Draws a utility function from some set
  3. Chooses in order to maximize utility given information
- Key differences between Random Utility and Rational Inattention
  1. Random Utility allows for multiple utility functions
  2. Rational Inattention allows attention to vary with choice problem
- How can we differentiate between the two?

Monotonicity

- Random Utility implies monotonicity
  - In fact, fully characterized by Block Marschak monotonicity
- For any two decision problems $\{A, A \cup b\}$, $a \in A$ and $b \notin A$

$$P_A(a|\omega) \geq P_{A \cup b}(a|\omega)$$

- Rational Inattention can lead to violations of monotonicity (Ergin, Matejka and McKay)

Act	**Payoff 49 red dots**	**Payoff 51 red dots**
a	23	23
b	20	25
c	40	0

- Adding act c to $\{a, b\}$ can increase the probability of choosing b in state 51

- Introduce an experimental interface that can be used to collect state dependent stochastic choice data
- Use it to perform some basic tests
  - Whether subjects actively adjust their attention
  - Whether they do so optimally
  - Measure attention costs
- Rule out alternative models with fixed attention
  - Signal Detection Theory
  - Random Utility Models
- Experimental note: subjects paid in probability points to keep utility 'linear'

An Aside: Testing Axioms with Stochastic Data

- Much of the following is going to come down to testing axioms of the following form

$$P(a|1) \geq P(a|2)$$

- These are conditions on the **population** probabilities
- We don't observe these, instead we observe **sample** estimates $\bar{P}(a|1)$ and $\bar{P}(a|2)$
- What to do?

An Aside: Testing Axioms with Stochastic Data

- We can make **statistical statements** about the validity of the axioms
- But there are two was to do this
  1. Can we reject a **violation** of the axiom
     - i.e., is it the case that $\bar{P}(a|1) > \bar{P}(a|2)$ and we can reject the hypothesis that $P(a|1) = P(a|2)$ at (say) the 5% level
  2. Can we find a **significant** violation of the axiom
     - i.e. is it the case that $\bar{P}(a|1) < \bar{P}(a|2)$ and we can reject the hypothesis that $P(a|1) = P(a|2)$ at (say) the 5% level
- (1) Is clearly a much tougher test that (2)
- If we have low power we will never be able to do (1)

Experimental Results

- Experiment 1: Extensive Margin
- Experiment 2: Spillovers
- Experiment 3: Intensive Margin

Experiment 2				
Decision Problem	Payoffs			
	$U(a,1)$	$U(a,2)$	$U(b,1)$	$U(b,2)$
1	5	0	0	5
2	40	0	0	40
3	70	0	0	70
4	95	0	0	95

- Two equally likely states
- Two acts (a and b)
- Symmetric change in the value of making correct choice
- 46 subjects

Testing NIAC and NIAS

- In the symmetric 2x2 case, NIAS and NIAS have specific forms
- NIAS:

$$P_A(a|\omega_1) \geq \max\left\{\alpha P_A(a|\omega_2), \alpha P_A(a|\omega_2) + \beta\right\}, \qquad (1)$$

where

$$\begin{aligned}\alpha &= \frac{u(b(\omega_2)) - u(a(\omega_2))}{u(a(\omega_1)) - u(b(\omega_1))} \\ \beta &= \frac{u(a(\omega_1)) + u(a(\omega_2)) - u(b(\omega_1)) - u(b(\omega_2))}{(a(\omega_1)) - u(b(\omega_1))}\end{aligned}$$

- In this case boils down to

$$P(a|\omega_1) \geq P(a|\omega_2)$$

- NIAC:

$$
\begin{aligned}
& \Delta P(a|\omega_1)\left(\Delta\left(u(a(\omega_1)) - u(b(\omega_1))\right)\right) + && (2)\\
& \Delta P(b|\omega_2)\left(\Delta\left(u(b(\omega_2)) - u(a(\omega_2))\right)\right) && (3)\\
\geq\ & 0 && (4)
\end{aligned}
$$

- In this case boils down to

$$
\begin{aligned}
& P_1(a|\omega_1) + P_1(b|\omega_2)\\
\leq\ & P_2(a|\omega_1) + P_2(b|\omega_2)\\
\leq\ & P_3(a|\omega_1) + P_3(b|\omega_2)\\
\leq\ & P_4(a|\omega_1) + P_4(b|\omega_2)
\end{aligned}
$$

Do People Optimally Adjust Attention?

- Alternative model: Choose optimally conditional on fixed signal
  - e.g. Signal detection theory [Green and Swets 1966]
- In general, choices can vary with incentives
  - Changes optimal choice in posterior state
- But not in this case
  - Optimal to choose *a* if $\gamma_1 > 0.5$, regardless of prize
- Change in choice between decision problems rules out Signal Detection Theory
  - Also rational inattention with fixed entropy

Testing NIAS: Experiment 1

- NIAS test: For each decision problem

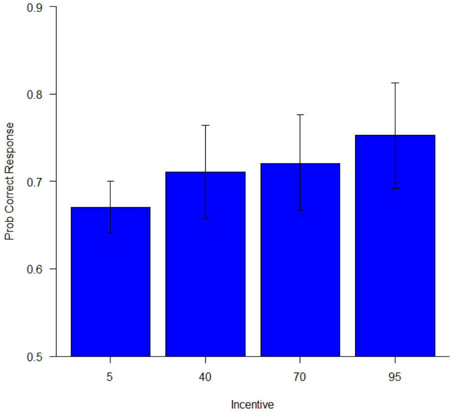
$$P(a|1) \geq P(a|2)$$

- From the aggregate data

Table 2: NIAS Test

DP	$P_j(a|1)$	$P_j(a|2)$	Prob
1	0.74	0.40	0.000
2	0.76	0.34	0.000
3	0.78	0.34	0.000
4	0.78	0.27	0.000

Testing NIAC: Experiment 1

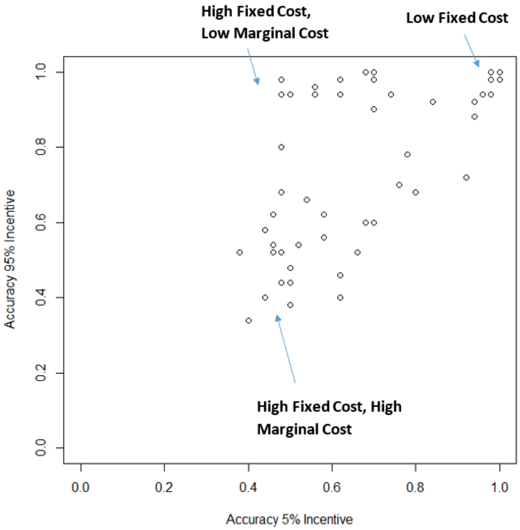


NIAC And NIAS: Individual Level

Violate	%
NIAS Only	2
NIAC Only	17
Both	0
Neither	81

- Counting only statistically significant violations

Recovering Costs - Individual Level



Experiment 2: Spillovers

Table 1: Experiment 1						
	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

Table 2: Treatments for Exp. 1		
Treatment	Payoffs	
	b_1	b_2
1	40	55
2	40	52
3	30	55
4	30	52

Experiment 2: Spillover

Table 8: Results of Experiment 1

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

Experiment 3: Intensive Margin

Experiment 3								
	Payoffs							
Decision Problem	U_1^a	U_2^a	U_3^a	U_4^a	U_1^b	U_2^b	U_3^b	U_4^b
9	1	0	10	0	0	1	0	10
10	10	0	1	0	0	10	0	1
11	1	0	1	0	0	1	0	1
12	10	0	10	0	0	10	0	10

- 4 states of the world: 29, 31, 69, 71 red balls
- Change which states it is important to differentiate between

Testing NIAC: Experiment 3

Experiment 3								
	Payoffs							
Decision Problem	U_1^a	U_2^a	U_3^a	U_4^a	U_1^b	U_2^b	U_3^b	U_4^b
9	1	0	10	0	0	1	0	10
10	10	0	1	0	0	10	0	1

- Comparing DP 9 and 10
  - DP9: important to differentiate between states 3 and 4
  - DP10: important to differentiate between states 1 and 2

$$\begin{aligned} & P_{10}(a|\omega_1) + P_{10}(b|\omega_2) + P_9(a|\omega_3) + P_9(b|\omega_4) \\ \geq \quad & P_9(a|\omega_1) + P_9(b|\omega_2) + P_{10}(a|\omega_3) + P_{10}(b|\omega_4) \end{aligned}$$

  - Average LHS: 73%, Average RHS: 65% (24 subjects)
- Overall 79% of subjects in line of NIAC

Summary

- These are clearly extremely simple experimental tests
- A lot more work to be done
  - identifying where people are optimal and where they are not
  - identifying types of mistakes that they are making
  - measuring costs.

Other Approaches

- There are lots of other papers testing the rational inattention hypothesis for specific cost functions:
  - Shannon mutual information (e.g. Sims 2003)
  - Shannon capacity (e.g. Woodford 2012)
  - Choice of optimal partitions (Ellis 2012)
  - All or nothing (Reis 2006)
- We will talk (in particular) about mutual information next week.

de Oliveira et al [2017]

- One other paper considers optimal information acquisition without making any assumption about the cost functions
- Rather than state dependant stochastic choice data, uses preferences over menus
  - i.e would you prefer to make a choice for menu A or menu B
- Timeline is as follows
  - Choose between menu
  - State resolves itself
  - Choose what information processing to do
  - Choose an alternative based on signal

de Oliveira et al [2017]

- Two key conditions for rational inattention

1. Preference for Flexibility
   - $A \cup \{a\} \succeq A$
   - Always prefer to have more options
   - Note relation to 'too much choice'
2. Preference for Early Resolution of Uncertainty
   - Define $\frac{1}{2}$ mixture of A and B as

$$\left\{c = \frac{1}{2}a + \frac{1}{2}b | a \in A, b \in B\right\}$$

   - Choosing from $\frac{1}{2}A + \frac{1}{2}B$ is like choosing from A, choosing from B then flipping a coin to see which choice you get
   - This is costly from an informational standpoint

$$\begin{aligned} A &\sim B \Rightarrow \\ A &\succeq \frac{1}{2}A + \frac{1}{2}B \end{aligned}$$

Summary

- Introduced 'Rational Inattention'
  - A class of models in which it is costly to learn about the state of the world
- Introduced 'State Dependent Stochastic Choice data'
  - A handy data set for testing models of rational inattention
- Introduced an experimental method for collecting SDSC
- Very early stage in the research program, lots of open questions
  - Many other experiments to be run
  - SDSC data in the wild
  - Link between menu choice and stochastic choice