Rational Inattention

Mark Dean

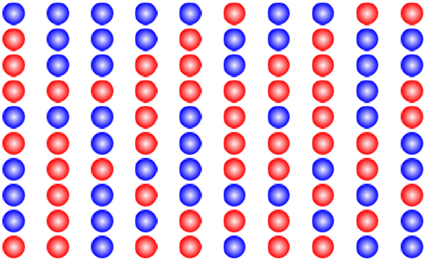
Behavioral Economics Spring 2017

The Story So Far.....

- (Hopefully) convinced you that attention costs are important
- Introduced the 'satisficing' model of search and choice
- But, this model seems quite restrictive:
  - Sequential Search
  - 'All or nothing' understanding of alternatives
- 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

- You are deciding whether or not to buy a used car
- The car might be high quality
  - in which case you want to buy it
- Or of low quality
  - in which case you don't
- The more attention you pay to the problem, the better information you will get about the quality of the car
- But this is not really a situation of satisficing....



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

- An alternative model of information gathering
- The world can be in one of a number of different ‘states’
    - 47 or 53 balls on a screen
    - Demand for your product can be high or low
    - Quality of a used car can be good or bad
    - A firm could be profitable or not
- Initially have some beliefs about the likelihood of different states of the world
    - This is your ‘prior’

- By exerting effort, we can learn more about the 'state'
  - Count some of the balls
  - Run a customer survey
  - Ask a mechanic to look at the car
  - Read some stock market reports
- The more inforrmation you gather, the better choices you will subsequently make
  - Less likely to buy a bad car
  - Invest in a bad stock
  - Price your product badly
- But this learning comes with costs
  - Time, Cognitive effort, Money, etc

Rational Inattention

- Key decision

1. How much information to gather?
   - Better information $\Rightarrow$ Better choice
   - But at more cost
2. What **type** of information to gather?
   - Want to gather information that is **relevant** to your choice

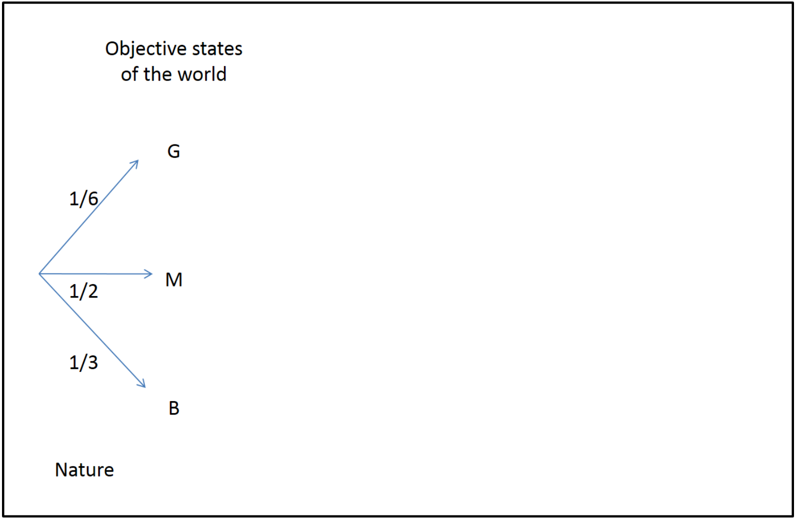
- This is the model of **rational inattention**
- Heavly used in economics
  - Consumption/savings
  - Portfolio choice
  - Pricing of firms

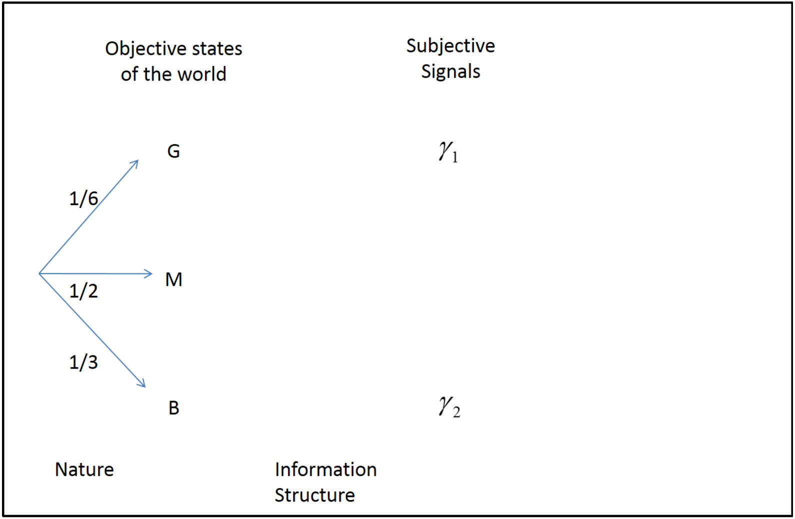
The Choice Problem

- 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
- Then chooses what action to take based only on the signal.
- More informative information structures are more costly, but lead to better decisions
  - Sets up a trade off

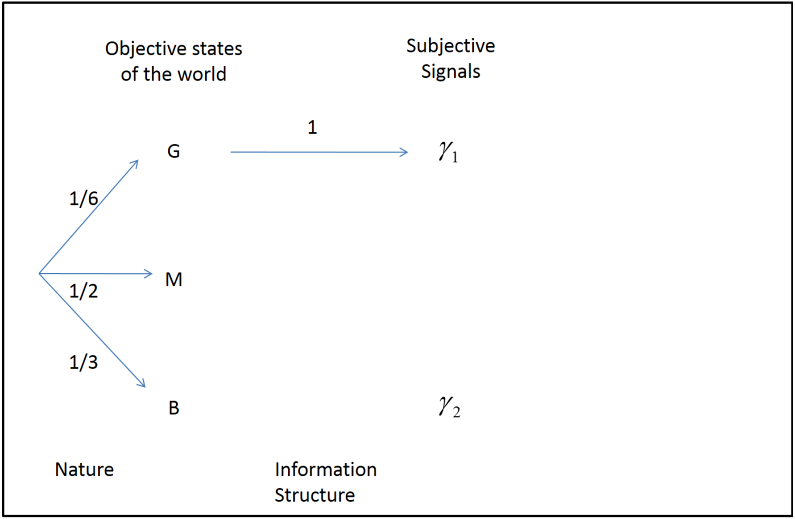
- This may seem like a really weird way of setting up the problem
- After all, who goes about choosing information structures?
- I'm going to claim that this is a good modelling tool
    - Even if you don't choose information structures **directly,** I can still think of your information gathering as generating an information structure
- Will come back to this point after I have explained what an information structure is

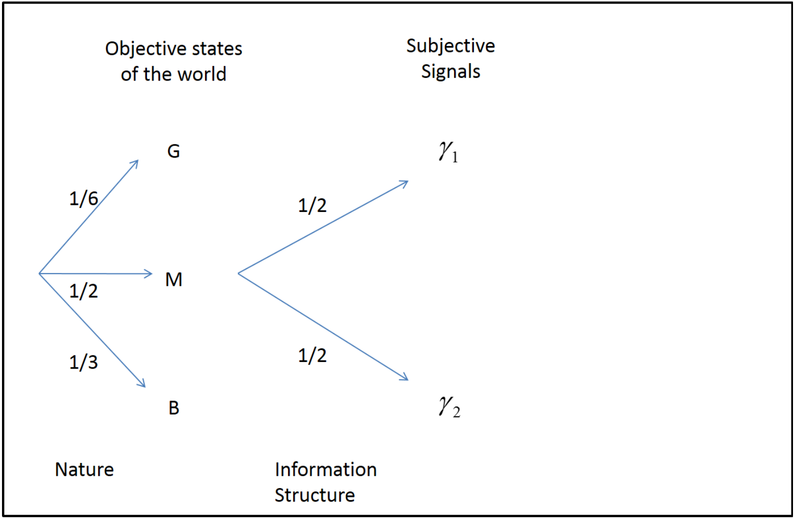
- Objective states of the world
  - e.g. Demand could be 'good', 'medium' or 'bad'
- At the end of the day, decision maker chooses an action
  - e.g. Set price to be high, average, or low
- Gross payoff depends on action and state
  - e.g. Quantity sold depends on price and demand
- Decision maker get to learn something about the state before choosing action
  - e.g. Could do market research, focus groups, etc.
  - This we model as choice of information structure

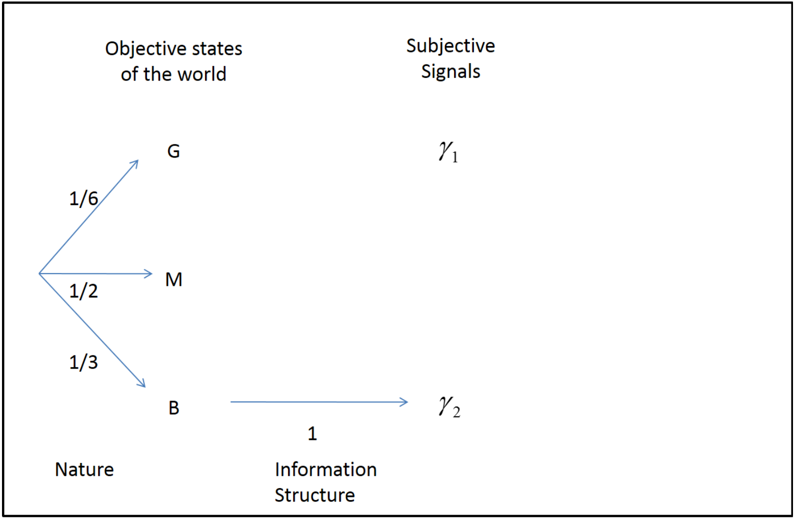

Objective states
of the world
G
1/6
M
1/2
1/3
B
Nature

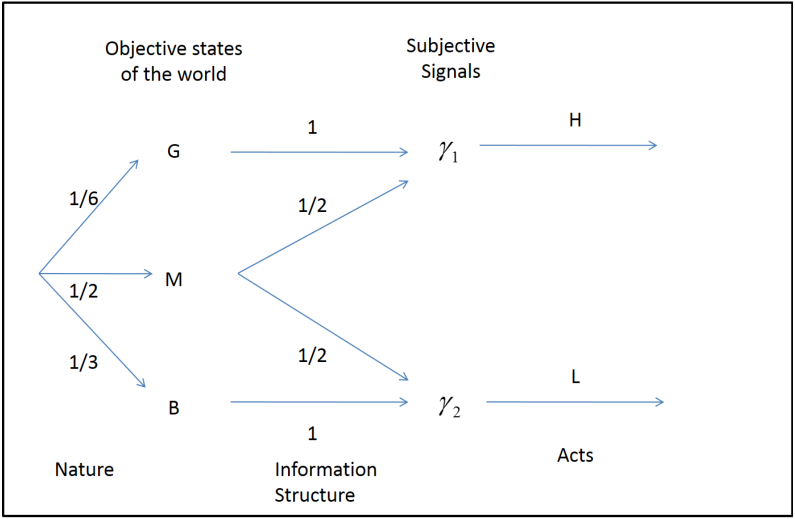

Objective states of the world
Subjective Signals
G
γ_1
1/6
M
1/2
1/3
B
γ_2
Nature
Information Structure

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


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


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

Describing an Information Structure

- $\Omega = \{\omega_1, ...\omega_M\}$: States of the world (number of balls, quality of the car, etc)
  - with prior probabilities μ
- Information structure defined by:
  - Set of signals: $\Gamma(\pi)$
  - Probability of receiving each signal γ from each state $\omega : \pi(\gamma|\omega)$
- In previous example

	Signal (Γ)	
State (Ω)	*R*	*S*
G	1	0
M	$\frac{1}{2}$	$\frac{1}{2}$
B	0	1

Information Structures as Metaphors

- 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 50 members of the class report if they like the class or not
- This generates an information structure
  - 51 signals: 0,1,2..... people say they like the class
  - Probability of each signal given each state of the world can be calculated

What Information Structure to Choose?

- Better information will lead to better choices
- But will cost more
  - Time, effort, money etc
- How to decide what information structure to choose?
- Trade off
  - Benefit of information (easy to measure)
  - Cost of information (hard to measure)
- Assume that this trade off is done *optimally*

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
- 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

- Let $u(a(\omega))$ be the utility (profit) that action a gives in state ω

The Value of An Information Structure

- What would you choose if you gathered no information?
  - i.e. if you had your prior beliefs
  - Use μ to describe the prior

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

- 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
- Luckily we can calculate this using Bayes Rule

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

- 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

The Value of An Information Structure

- 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

The Choice of Information Structure

- What information structure would you choose?
- In general, more information means better choices, and higher values
- Without further constraints, would choose to be fully informed
- To make the problem interesting and realistic, need to introduce a **cost to information** K
- The 'net value' of an information structure π in choice set A is

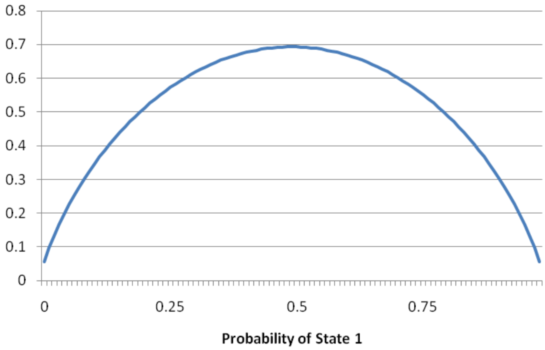
$$G(\pi, A) - K(\pi)$$

What is the cost of information?

- What form should information costs K take?
- Good question!
- Many alternatives have been considered in the literature
  - Pay for the precision of a normal signals (we will see an example of this later)
  - 'All or Nothing'
- One popular alternative is 'Shannon mutual information' (Sims 2003)
  - A way of measuring how much information is gained by using an information structure

- Shannon Entropy is a measure of how much 'missing information' there is in a probability distribution
- In other words - how much we do not know, or how much we would learn from resolving the uncertainty
- For a random variable X that takes the value x_i with probability $p(x_i)$ for $i = 1...n$, defined as

$$\begin{aligned} H(X) &= E(-\ln(p(x_i))) \\ &= -\sum_i p(x_i) \ln(p_i) \end{aligned}$$



- Can think of it as how much we learn from result of experiment
  - i.e. actually determining what x is
- Lower entropy means that you are more informed

- Related to the notion of entropy is the notion of Mutual Information

$$I(X,Y)=\sum_x\sum_y p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$$

- Measure of how much information one variable tells you about another
- Note that $I(X,Y)=0$ if X and Y are independent
- Can be rewritten as

$$\begin{aligned} & \sum_y p(y)\sum_x p(x|y)\ln p(x|y)-\sum_y p(x)\ln p(x) \\ = \; & H(X)-\sum_y P(y)H(X|y) \end{aligned}$$

- The expected reduction in entropy about variable x from observing y

Mutual Information and Information Costs

- Mutual Information measures the expected reduction in entropy from observing a signal
- We can use it as a measure of information costs

$$K(\pi, \mu) = -\kappa \left[\text{ expected entropy of signals - entropy of prior} \right]$$

$$= -\kappa \left[\sum_{\gamma \in \Gamma(\pi)} P(\gamma) \sum_{\omega \in \Omega} \gamma(\omega) \ln \gamma(\omega) - \sum_{\omega \in \Omega} \mu(\omega) \ln \mu(\omega) \right]$$

- Can be justified by information theory
    - Mutual Information related to the number of bits of information that need to be sent to achive the information structure

Working with Rational Inattention

- Now we have defined information costs, the optimization problem is well defined
- For any set of alternatives A, choose π to maximize

$$G(\pi, A) - K(\pi)$$

- What does this tell us about behavior?

A Simple Example

- Consider the case of two state and two acts

	ω_1	ω_2
a	$U(a(\omega_1))$	$U(a(\omega_2))$
b	$U(b(\omega_1))$	$U(b(\omega_2))$

- It is easy to show that decision maker will never choose more than 2 signals
  - Why?
  - After you receive a signal you will either choose a or b
  - If you use (say) 3 signals you will take the same action after 2 of them
  - But this is a waste of information!
  - Just merge those two signals

A Simple Example

- Assume $\mu(1) = \mu(2) = 0.5$
- Assume that they do choose two signals
  - γ^a, after which a is chosen
  - γ^b, after which b is chosen
- There are several ways to set up the resulting optimization problem
  - For example, choosing probabilites $\pi(\gamma|\omega)$
  - I'll show you one that can sometimes be particularly useful

Solving for Optimal Behavior

- Choose
  - $P(\gamma^a)$: Probability of signal γ^a
  - $\gamma^a(\omega_1)$: Posterior probability of state ω_1 following γ^a
  - $\gamma^b(\omega_1)$: Posterior probability of state ω_1 following γ^b
- To maximize

$$P(\gamma^a)\left[\gamma^a(\omega_1)u(a(\omega_1)) + (1-\gamma^a(\omega_1))u(a(\omega_2))\right] +$$
$$(1-P(\gamma^a))\left[\gamma^b(\omega_1)u(b(\omega_1)) + (1-\gamma^b(\omega_1))u(b(\omega_2))\right]$$

$$-\kappa\left[\begin{array}{c} P(\gamma^a)\left(\begin{array}{c}\gamma^a(\omega_1)\ln\gamma^a(\omega_1)+ \\ (1-\gamma^a(\omega_1))\ln(1-\gamma^a(\omega_1))\end{array}\right)+ \\ (1-P(\gamma^a))\left(\begin{array}{c}\gamma^b(\omega_1)\ln\gamma^b(\omega_1)+ \\ (1-\gamma^b(\omega_1))\ln(1-\gamma^b(\omega_1))\end{array}\right)\end{array}\right]$$

- subject to

$$P(\gamma^a)\gamma^a(\omega_1) + (1-P(\gamma^a))\gamma^b(\omega_1) = \mu(\omega_1)$$

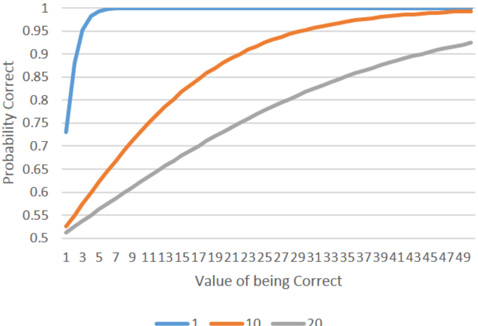
- This can be solved using standard optimization techniques
- You will show

$$
\begin{aligned}
\frac{\gamma^a(\omega_1)}{\gamma^b(\omega_1)} &= \exp\left(\frac{u(a(\omega_1)) - u(b(\omega_1))}{\kappa}\right) \\
\frac{\gamma^a(\omega_2)}{\gamma^b(\omega_2)} &= \exp\left(\frac{u(a(\omega_2)) - u(b(\omega_2))}{\kappa}\right)
\end{aligned}
$$

- Ratio of beliefs in each state depends only on the 'cost of mistakes' in that state
- Posterior beliefs do not depend on priors

- We can use these formulae to calculate how probability of correct choice changes with reward.
- Assume
  - $u(a(\omega_1)) = u(b(\omega_2)) = c$, $u(a(\omega_2)) = u(b(\omega_2)) = 0$,
- Imples that

$$\pi(\gamma^a|\omega_1) = \pi(\gamma^b|\omega_2) = \frac{\exp\left(\frac{c}{\kappa}\right)}{1 + \exp\left(\frac{c}{\kappa}\right)}$$



A More General Solution

$$P(a|\omega) = \frac{P(a) \exp^{\frac{u(a(\omega))}{\kappa}}}{\sum_{c \in A} P(c) \exp^{\frac{u(c(\omega))}{\kappa}}}$$

- Where
  - $P(a|\omega)$ is the probability of choosing a in state ω
  - $P(a)$ is the unconditional probability of choosing a
- See Matejka and McKay [2015]
- As costs go to zero, deterministically pick best option in that state
- As costs go to infinity, deterministically pick the best option ex ante
- Sometimes the model can be solved analytically
- Sometimes need a numerical solution (e.g. Blahut Arimoto)

Application: Price Setting with Rationally Inattentive Consumers

- Consider buying a car
- The price of the car is easy to observe
- But quality is difficult to observe
- How much effort do consumers put into finding out quality?
- How does this affect the prices that firms charge?
- This application comes from Martin [2017]

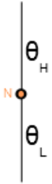
Application: Price Setting with Rationally Inattentive Consumers

- Model this as a simple game
  1. Quality of the car can be either high or low
  2. Firm decides what price to set depending on the quality
  3. Consumer observes price, then decides how much information to gather
  4. Decides whether or not to buy depending on their resulting signal
  5. Assume that consumer wants to buy low quality product at low price, but not at high price
- Key point: prices may convey information about quality
- And so may effect how much effort buyer puts into determining quality

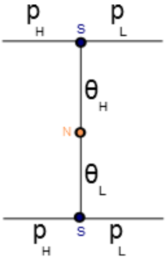
Market Setting

- One off sales encounter
  - One buyer, one seller, one product

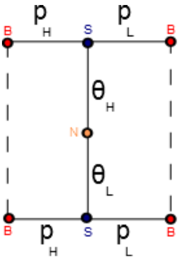
- Nature determines quality $\theta \in \{\theta_L, \theta_H\}$
  - Prior $\mu = \Pr(\omega_H)$



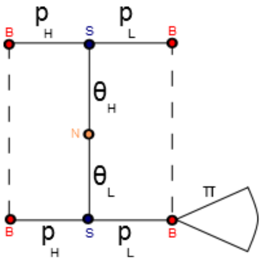
- Seller learns quality, sets price $p \in \{p_L, p_H\}$



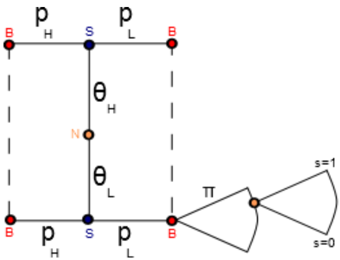
- Buyer learns p, forms interim belief μ_p (probability of high quality given price)
  - Based on prior μ and seller strategies



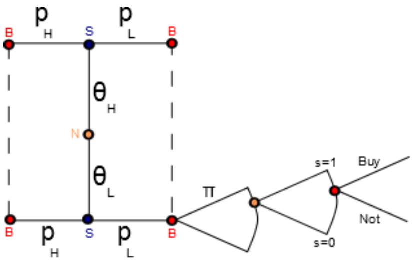
- Choose attention strategy contingent on price $\{\pi^H, \pi^L\}$
  - Costs based on Shannon mutual information



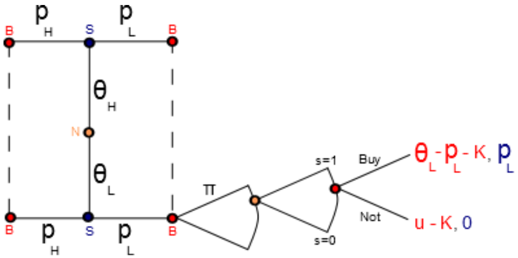
- Nature determines a signal
  - Posterior belief about product being high quality



- Decides whether to buy or not
    - Just a unit of the good



- Standard utility and profit functions (risk neutral EU)
  - $u \in \mathbb{R}_+$ is outside option, $K \in \mathbb{R}_+$ is Shannon cost



Equilibrium

- How do we make predictions in this setting?
- We need to find
  - A pricing strategy for low and high quality firms
  - An attention strategy for the consumer upon seeing low and high prices
  - A buying strategy for the consumers
- Such that
  - Firms are optimizing profits given the behavior of the customers
  - Consumers are maximizing utility given the behavior of the firms

Equilibrium

- There is **no** equilibrium in which low quality firm charges p_L and high quality firm charges p_H
- Why?
- If this were the case, the consumer would be completely inattentive with probability 1 at both prices
    - Price conveys all information
- Incentive for the low quality firm to cheat and charge the high price
- Would sell with probability 1

Equilibrium

- Always exists "Pooling low" Equilibrium
  - High quality sellers charge a *low price* with probability 1
  - Low quality sellers charge a *low price* with probability 1
  - Buyer believes that high price is a signal of low quality
- However, this is not a 'sensible' equilibrium:
  - Perverse beliefs on behalf of the buyer:
  - High price implies low quality
  - Allowed because beliefs never tested in equilibrium

Theorem

For every cost λ, there exists an equilibrium ("mimic high") where high quality sellers price high with probability 1 and low quality sellers price high with a unique probability $\eta \in [0, 1]$.

Explaining the Equilibrium

- How do rationally inattentive consumers behave?
- If prices are low, do not pay attention
- If prices are high, choose to have two signals
  - 'bad signal' - with high probability good is of low quality
  - 'good signal' - with high probability good is of high quality
- Buy item only after good signal

Explaining the Equilibrium

- Give rise to two posteriors (prob of high quality):
  - $\gamma^0_{p_H}$ (bad signal)
  - $\gamma^1_{p_H}$ (good signal)
- We showed that these optimal posterior beliefs are determined by the relative rewards of buying and not buying in each state

$$\ln\left(\frac{\gamma^1_{p_H}}{\gamma^0_{p_H}}\right) = \frac{(\theta_H - p_H) - u}{\kappa}$$

$$\ln\left(\frac{1-\gamma^1_{p_H}}{1-\gamma^0_{p_H}}\right) = \frac{(\theta_L - p_H) - u}{\kappa}$$

- Let $\mu_{p_H}(H)$ be the prior probability that the good is of high quality given that it is of high price
- Let $d_{p_H}^{\theta_L}$ be the probability of buying a good if it is actually low quality if the price is high:
  - i.e $\pi_{p_H}(\gamma_{p_H}^1|\theta_L)$
- Using Bayes rule, we (you!) can show:

$$d_{p_H}^{\theta_L} = \frac{\left(\frac{1-\gamma_{p_H}^1}{\gamma_{p_H}^1-\gamma_{p_H}^0}\right)\left(\mu_{p_H}(H) - \gamma_{p_H}^0\right)}{\left(1-\mu_{p_H}(H)\right)}$$

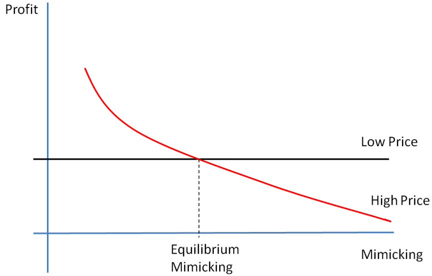
- Conditional demand is
  - Strictly increasing in interim beliefs μ_{p_H}
  - So strictly decreasing in ‘mimicking’ η

- What about firm behavior?
- If the low quality firm sometimes prices high and sometimes prices low, we need them to be **indifferent** between the two

$$d_{p_H}^{\theta_L} \times p_H = p_L \Rightarrow d_{p_H}^{\theta_L} = \frac{p_L}{p_H}$$

- As low quality firms become more likely to mimic, it decreases the probability that the low quality car will be bought
- And so reducs the value of setting the high price

Firm Behavior



- What is the unique value of η when $\eta \in (0,1)$?

$$\eta = \frac{\lambda}{1-\lambda} \frac{\left(1-\gamma^0_{p_H}\right)\left(1-\gamma^1_{p_H}\right)}{\gamma^0_{p_H}\left(1-\gamma^1_{p_H}\right) + \frac{p_L}{p_H}\left(\gamma^1_{p_H} - \gamma^0_{p_H}\right)}$$

- We can use a model of rational inattention to solve form
  - Consumer demand
  - Firm pricing strategies
- Can use the model to make predictions about how these change with parameters of the model
  - E.g as $\kappa \to 0$, $\eta \to 0$

Discrimination [Bartos et al 2016]

- A second recent application of the rational inattention model has been to study discrimination
- Imagine you are a firm looking to recruit someone for a job
- You see the name of the applicant at the top of the CV
- This gives you a clue to which 'group' an applicant belongs to
    - e.g. British vs American
- You have some prior belief about the abilities of these groups
    - e.g. British people are better than Americans
- Do you spend more time looking at the CVs of Brits or Americans?

A Formal Version of the Model

- You are considering an applicant for a position
  - Hiring for a job
  - Looking for someone to rent your flat
- An applicant is of quality q, which you do not observe
- If you hire the applicant you get payoff q
- Otherwise you get 0

Information

- Initially you get to observe which group the applicant comes from
  - Brits (B) or Americans (A)
- Your prior beliefs depend on this group
- If the persion is British you believe

$$q \sim N(q_B, \sigma^2)$$

- American

$$q \sim N(q_A, \sigma^2)$$

  with $q_B < q_A$
- This is your 'bias'

- Before deciding whether to hire the applicant you receive a normal signal
$$y = q + \varepsilon$$
Where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$
- You get to choose the **precision** of the signal
  - i.e. get to choose σ_ε^2
- Pay a cost based on the precision of the signal
  - $M(\sigma_\varepsilon^2)$
- Note, it doesn't have to be the case that costs are equal to Shannon
  - Only assume that lower variance gives higher costs

- What are the benefits of information?
- What do you believe after seeing signal if variance is σ_ε^2?

$$q' = \alpha y + (1 - \alpha) q_G$$

Where q_G is the beliefs given the group (i.e. q_B or q_A)

$$\alpha = \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}$$

- As signal gets more precise (i.e σ_ε^2 falls) then
  - More weight is put on the signal
  - Less weight put on the bias
- If information was free then bias wouldn't matter

- If you got signal y, what would you choose?
- If
$$q' = \alpha y + (1 - \alpha) q_G > 0$$
- Will hire the person
- Otherwise will not

- Value of the information structure is the value of the choice for each y

$$\max\left\{\alpha y + (1-\alpha)q_G, 0\right\}$$

- Integrated over all possible values of y

$$G(\sigma_\varepsilon^2) = \int_{-\frac{(1-\alpha)}{\alpha}q_G}^{\infty} \alpha y + (1-\alpha)q_G \, dy$$

- So the optimal strategy is to

1. Choose the precision of the signal σ_ε^2 to maximize

$$G(\sigma_\varepsilon^2) - M(\sigma_\varepsilon^2)$$

2. Hire the worker if and only if

$$\alpha y + (1-\alpha) q_G > 0$$

or

$$\varepsilon > q + \frac{(1+\alpha)}{\alpha} q_G$$

- What type of question can we answer with this model?

1. Do Brits or Americans recieve more attention
2. Does 'Rational Inattention' help or hurt the group that descriminated against?
    - i.e. would Americans do better or worse if σ_ε^2 had to be the same for both groups?

Cherry Picking or Lemon Dropping

- It turns out the answer depends on whether we are in a 'Cherry Picking' or 'Lemon Dropping' market
- Cherry Picking: would not hire the 'average' candidate from either group
  - i.e. $q_B < q_A < 0$
  - Only candidates for which good signals are received are hired
  - e.g. hiring for a job
- Lemon Dropping: would hire the 'average' candidate from either group
  - i.e. $0 < q_B < q_A$
  - Only candidates for which bad signals are recieved are not hired
  - e.g. looking for people to rent an apartment

Theorem
In Cherry Picking markets, the 'worse' group gets less attention, and rational attention hurts the 'worse' group

Theorem
In Lemon Dropping markets, the 'worse' group gets more attention, and rational attention hurts the 'worse' group

- 'Hurts' in this case means relative to a situation in which the 'worse' group had to be given the same attention as the 'better' group
- Minorites get screwed either way!

Theorem

- Intuition:

1. Attention is more valuable to the hirer the further away a group is from the threshold on average
   - If you are far away from the threshold, less likely information will make a difference to my choice
   - In the cherry picking market the ‘worse’ group is further away from the threshold, and so get less attention
   - In the lemon dropping market the worse group is closer to the threshold and gets more attention
2. Attention is more likely to get you hired in the cherry picking market, less likely to get you hired in the lemon dropping market
   - In the first case only hired if there is high quality evidence that you are good
   - In the latter case hired unless there is high quality evidence that you are bad

- Market 1: Lemon Dropping - Housing Applications
- Market 2: Cherry Picking - Job Applications
- Experiment run in Czech Republic
- In each case used dummy applicants with different 'types' of name
    - White
    - Asian
    - Roma

TABLE 1—CZECH RENTAL HOUSING MARKET: INVITATION RATES AND INFORMATION ACQUISITION BY ETHNICITY, COMPARISON OF MEANS

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	Percentage point difference: W − E, (*p*-value) (3)	Asian minority name (A) (4)	Percentage point difference: W − A, (*p*-value) (5)	Roma minority name (R) (6)	Percentage point difference: W − R, (*p*-value) (7)	Percentage point difference: R − A, (*p*-value) (8)
Panel A. Invitation for a flat visit								
No Information Treatment ($n = 451$)	0.78	0.41	37 (0.00)	0.39	39 (0.00)	0.43	36 (0.00)	3 (0.57)
Monitored Information Treatment ($n = 762$)	0.72	0.49	23 (0.00)	0.49	23 (0.00)	0.49	23 (0.00)	0 (0.92)
Monitored Information Treatment[a] ($n = 293$)	0.84	0.66	18 (0.00)	0.71	13 (0.00)	0.62	21 (0.00)	−9 (0.20)
Monitored Information Treatment[b] ($n = 469$)	0.66	0.37	29 (0.00)	0.35	31 (0.00)	0.39	27 (0.00)	4 (0.51)
Treatment with additional text in the e-mail ($n = 587$)	0.78	0.52	26 (0.00)	0.49	29 (0.00)	0.55	23 (0.00)	5 (0.29)
Panel B. Information acquisition in the Monitored Information Treatment								
Opening applicant's personal website	0.33	0.41	−8 (0.03)	0.38	−5 (0.24)	0.44	−11 (0.01)	6 (0.15)
Number of pieces of information acquired	1.29	1.75	−0.46 (0.01)	1.61	−0.32 (0.09)	1.88	−0.59 (0.00)	0.27 (0.17)
At least one piece of information acquired	0.30	0.40	−10 (0.01)	0.37	−7 (0.12)	0.44	−13 (0.00)	7 (0.12)
All pieces of information acquired	0.19	0.26	−8 (0.02)	0.24	−6 (0.12)	0.28	−10 (0.01)	4 (0.33)
Number of pieces of information acquired[a]	3.91	4.24	−0.33 (0.06)	4.23	−0.32 (0.15)	4.25	−0.34 (0.09)	0.02 (0.90)
At least one piece of information acquired[a]	0.92	0.98	−6 (0.02)	0.97	−5 (0.15)	0.98	−7 (0.03)	2 (0.47)
All pieces of information acquired[a]	0.56	0.64	−7 (0.23)	0.64	−8 (0.30)	0.64	−7 (0.30)	−0 (0.96)

TABLE 4—CZECH LABOR MARKET: INVITATION RATES AND INFORMATION ACQUISITION BY ETHNICITY, COMPARISON OF MEANS

	White majority name (W) (1)	Pooled Asian and Roma minority name (E) (2)	Percentage point difference: W − E, (p-value) (3)	Asian minority name (A) (4)	Percentage point difference: W − A, (p-value) (5)	Roma minority name (R) (6)	Percentage point difference: W − R, (p-value) (7)	Percentage point difference: R − A, (p-value) (8)
Panel A. Employer's response								
Callback	0.43	0.20	23 (0.00)	0.17	26 (0.00)	0.25	18 (0.01)	8 (0.22)
Invitation for a job interview	0.14	0.06	8 (0.03)	0.05	9 (0.03)	0.08	6 (0.18)	3 (0.46)
Invitation for a job interview[a]	0.19	0.09	10 (0.06)	0.09	10 (0.12)	0.10	9 (0.16)	1 (0.83)
Panel B. Information acquisition								
Opening applicant's resume	0.63	0.56	7 (0.22)	0.47	16 (0.03)	0.66	−3 (0.69)	19 (0.01)
Acquiring more information about qualification[a]	0.16	0.10	6 (0.27)	0.06	10 (0.12)	0.14	2 (0.73)	8 (0.24)
Acquiring more information about other characteristics[a]	0.18	0.18	0 (0.92)	0.19	−1 (0.85)	0.18	0 (0.99)	1 (0.85)

Other Applications

- Consumption and Savings [Sims 2003]
  - Standard permanent income hypothesis: consumption responds immediately and fully to changes in income
  - Rational Inattention: consumption responses occur gradually over time
  - Fits stylized facts in the macro literature
- Discrete Pricing [Matejka 2010]
  - Standard model: Firms prices should respond continuously to cost shocks
  - Rational Inattention: Firms will 'jump' between a small number of discrete prices
  - In line with observed date
- Home Bias [Van Nieuwerburgh and Veldkamp 2009]
  - Standard model: investors should diversify portfolio internationally
  - Rational Inattention: investors should specialize in assets they know more about
  - Leads to 'Home Bias' in investment

Summary

- Rational Inattention provides a way of modelling how people choose to learn about the state of the world
  - Applicable in cases in which satisficing is not appropriate
- Assumes people choose information to maximize value net of costs
  - Value depends on the choices to be made
  - Costs generally based on Shannon Entropy
- We can make predictions about learning and choice based on the rewards available in the environment
- Can be used to address a number of ‘puzzles’