Lecture 7

Noisy information models and rational inattention: Foundations

Chen Lian[1]

[1]UC Berkeley and NBER

March 20, 2025

Outline

Overview: Lecture 1

- **Noisy-information models:** a basic model to capture imperfect knowledge
  - accommodate rational and behavioral interpretations
- **Rational inattention (Sims):** a behavioral foundation of the "noise"
  - limited cognitive capacity
- **Sparsity (Gabaix):** a close cousin of rational inattention
- **Imprecise Perceptual Judgments (Woodford):** an alternative foundation of noise
- **Foundations** for many applications (survey evidence, GE dampening, sentiments
  - today: narrow thinking (Lian, 2021)

Formulation

- θ : an exogenous aggregate fundamental. For example:
  - TFP in RBC
  - monetary policy in NK
- Each household/firm i's knowledge about θ is captured by a noisy signal

$$s_i = \theta + \varepsilon_i$$

Expectations under Noisy Signals

Gaussian case:

$$\theta \sim \mathscr{N}\left(\mu, \sigma^2\right) \quad \text{and} \quad \varepsilon_i \sim \mathscr{N}\left(0, \sigma_\varepsilon^2\right),$$

where ε_i is i.i.d. and independent of x.

$$\begin{aligned} E_i[\theta] = E[\theta|s_i] &= \lambda s_i + (1-\lambda)\mu \\ &= \lambda(\theta + \varepsilon_i) + (1-\lambda)\mu, \end{aligned}$$

where

$$\lambda = \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2} \in [0,1].$$

Key prediction: belief about θ **under-reacts** to innovation in θ

"Heuristics" to think about conditional expectation:

- Run a predictive regression of θ on s_i
- Similar if you have multiple Gaussian signals

Interpretation of Noisy Signals and Macro Implications

First-generation ("70s"):

- Disperse information about the aggregate fundamental

Lucas (1972) island model: an "impressionism" version

- Disperse information about money supply m_t on each island

$$p_{i,t} = E_{i,t}[m_t] \quad \text{and} \quad p_t = \lambda m_t,$$

where I log-linearize around the steady state so $\mu_m = 0$.

- Monetary policy has real effects

$$y_t = m_t - p_t = (1 - \lambda) m_t$$

- THE "micro-foundation" of the Phillips curve in 70s

Interpretation of Noisy Signals and Macro Implications

The "80s" critique:

- Public available statistics reveals the aggregate fundamental
- The sticky-price/NK paradigm becomes dominant

The "00s" behavioral revival:

- Impossible for a human being to process all available info
- Noise captures cognitive limitation/rational inattention

What do Noisy Signals Imply?

Imperfect knowledge about the **fundamental**: ("first-order uncertainty")

$$\int E_i[\theta]\, di = \lambda\theta + (1-\lambda)\mu$$

Imperfect knowledge about others' **information/signals**: ("higher-order uncertainty")

$$\int\int E_i[s_j]\, djdi = \lambda\bar{s} + (1-\lambda)\mu,$$

where $\bar{s} = \int s_i di = \theta$.

- When **noises are independent across agents**

Imperfect knowledge about others' **actions**:

$$\int\int E_i[x_j]\, djdi = \int E_i[\bar{x}]\, di = \lambda\bar{x} + (1-\lambda)\mu_{\bar{x}},$$

where $\bar{x} = \int x_i di$ is the aggregate action.

- Capture **frictions in strategic interactions** (key later)

Pause for Questions

Outline

Overview

Rational Inattention:

- Impossible for a human being to process all available info
- Limited cognitive capacity and optimally allocate attention
- Trade-off between
  - cost of attention
  - benefit of attention (better decisions)
- "Micro-foundation" of the noisy information approach

$$s_i = \theta + \varepsilon_i,$$

where the variance of $\varepsilon_i \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$ is endogeneized

Entropy

- Sims (03): Cost of attention determined by quantity of information
- Entropy: measure of **the quantity of information** in information theory
- How much information is required to describe θ with probability density function $p(\theta)$?

$$H(\theta) = -\mathbb{E}\left[\log_2(p(\theta))\right],$$

  - Discrete: $H(\theta) = -\sum_{\theta \in \Theta} \log_2(p(\theta))\, p(\theta)$
  - Continuous: $H(\theta) = -\int_{\theta \in \Theta} \log_2(p(\theta))\, p(\theta)\, d\theta$
  - Note $\log_2$ is concave. If θ is more "dispersed," $H(\theta)$ is larger.

Entropy: Examples

Bernoulli:

- Flipping a fair coin. One bit of information.

$$H(\theta) = -\left(\frac{1}{2}\log_2\left(\frac{1}{2}\right) + \frac{1}{2}\log_2\left(\frac{1}{2}\right)\right) = 1$$

- Flipping a fair coin n times and θ captures the # of heads.

$$H(\theta) = n$$

Uniform: Suppose $\theta \in U[0, a]$

$$H(\theta) = -\int_0^a \log_2\left(\frac{1}{a}\right)\frac{1}{a}d\theta = \log_2 a$$

Entropy: Examples

Normal Variables:

- If $\theta \sim \mathcal{N}\left(\mu, \sigma^2\right)$

$$H(\theta) = \frac{1}{2}\log_2\left(2\pi e \sigma^2\right)$$

- If θ is an $n \times 1$ vector of normal variables $\theta \sim \mathcal{N}\left(\mu, \Sigma\right)$

$$H(\theta) = \frac{1}{2}\log_2\left((2\pi e)^n \det\left(\sum\right)\right)$$

Mutual Information

- **Measure:** How informative is a signal about the underlying random variable?
- **Mutual Information:** How much does observing one random variable reduce the entropy of the another?

$$I(\theta, s) = H(\theta) - H(\theta|s)$$

- **Conditional entropy:** The remaining amount of info in θ given that the value of s

$$H(\theta|s) = H(\theta, s) - H(s)$$

- **Symmetry** of mutual information:

$$I(\theta, s) = I(s, \theta) = H(\theta) + H(s) - H(\theta, s)$$

Mutual Information: Examples

Normal fundamentals and signals:

- Fundamental $\theta \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- Signal $s = \theta + \varepsilon$, where $\varepsilon \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$ and independent of θ

$$
\begin{aligned}
I(\theta, s) &= H(\theta) - H(\theta|s) \\
&= \frac{1}{2}\log_2\left(2\pi e \sigma_\theta^2\right) - \frac{1}{2}\log_2\left(2\pi e \sigma_{\theta|s}^2\right) \\
&= \frac{1}{2}\log_2\left(1 + \frac{\sigma_\theta^2}{\sigma_\varepsilon^2}\right)
\end{aligned}
$$

Pause for Questions

Rational Inattention: the Entropy Approach

Rational Inattention: the Entropy Approach (Sims, 2003)

- **Bound on information flow** measured based on mutual information

$$I(\theta, s) \leq \kappa$$

- Equivalently, **cognitive costs of attention** proportional to mutual information

$$\eta I(\theta, s)$$

A Simple Rational Inattention Problem

Payoff: A one-dimensional tracking problem.

$$-\frac{\omega}{2}\mathbb{E}\left[(a_i-\theta)^2\right]-\eta I(\theta,s_i),$$

where $\theta \sim \mathcal{N}(\mu,\sigma^2)$.

Attention to θ can be captured by a Gaussian signal s_i about θ

$$s_i=\theta+\varepsilon_i,$$

where $\varepsilon_i \sim \mathcal{N}(0,\sigma_\varepsilon^2)$ and independent of θ.

- The Gaussian form of the signal can be proved as optimal
- Same formulation as noisy-information, but σ_ε^2 is endogenously chosen

An Equivalent Formulation

Given signal s_i, the **optimal action** is

$$a_i = E[\theta|s_i] = \lambda s_i + (1-\lambda)\mu,$$

where $\lambda = \frac{\sigma^2}{\sigma^2+\sigma_\varepsilon^2}$ increases with the precision of the signal.

Cognitive cost:

$$I(\theta, s_i) = \frac{1}{2}\log_2\left(\frac{1}{1-\lambda}\right)$$

Expected tracking error:

$$\begin{aligned}\mathbb{E}\left[(a_i-\theta)^2\right] &= \mathbb{E}\left[(\lambda(\theta+\varepsilon_i)+(1-\lambda)\mu-\theta)^2\right] \\ &= \frac{\sigma^2\sigma_\varepsilon^2}{\sigma^2+\sigma_\varepsilon^2} = (1-\lambda)\sigma^2\end{aligned}$$

An Equivalent Formulation

$$\min_{\lambda} \frac{\omega}{2}\sigma^2(1-\lambda) + \frac{\eta}{2}\log_2\left(\frac{1}{1-\lambda}\right)$$

FOC:

$$\frac{\omega}{2}\sigma^2 = \frac{\eta}{2\ln 2}\frac{1}{1-\lambda^*}$$

Optimal attention:

$$\lambda^* = \begin{cases} 1 - \frac{\eta}{\ln 2}\frac{1}{\omega\sigma^2} & \omega\sigma^2 \geq \frac{\eta}{\ln 2} \\ 0 & \omega\sigma^2 < \frac{\eta}{\ln 2} \end{cases}$$

Under-appreciated point: "entropy" cost can generate **zero-attention corner solution**

- If benefits of tracking $\omega\sigma^2$ is too small/cost η of attention is too large

Pause for Questions

Optimal Price Setting under Rational Inattention

Expected loss: (Mackowiak & Wiederholt, 09)

$$-\frac{|\pi_{pp}|}{2}\mathbb{E}\left[\left(p_{i,t}-p^{*}_{i,t}\right)^{2}\right],$$

where $p^{*}_{i,t}=\Delta_t-\frac{\pi_{pz}}{\pi_{pp}}z_{i,t}$.

Information choice:

- One signal about the aggregate component $\Delta_t\sim\mathcal{N}\left(0,\sigma^2_{\Delta}\right)$

$$s_{1,i,t}=\Delta_t+\varepsilon_{i,t},$$

  where $\varepsilon_{i,t}\sim\mathcal{N}\left(0,\sigma^2_{\varepsilon}\right)$.

- One signal about the idiosyncratic component $z_{i,t}\sim\mathcal{N}\left(0,\sigma^2_{z}\right)$

$$s_{2,i,t}=z_{i,t}+v_{i,t},$$

  where $v_{i,t}\sim\mathcal{N}\left(0,\sigma^2_{v}\right)$.

- Implicit restriction: cannot have a signal directly about $p^{*}_{i,t}$
  - otherwise optimal to have $s^{*}_{i,t}=p^{*}_{i,t}+\varepsilon^{*}_{i,t}$

Optimal Price Setting under Rational Inattention

Attention allocation:

$$\underbrace{\frac{1}{2}\log_2\left(1+\frac{\sigma_\Delta^2}{\sigma_\varepsilon^2}\right)}_{\kappa_\Delta}+\underbrace{\frac{1}{2}\log_2\left(1+\frac{\sigma_z^2}{\sigma_v^2}\right)}_{\kappa_z}\leq\kappa.$$

Expected loss:

$$-\frac{|\pi_{pp}|}{2}\left[2^{-2\kappa_\Delta}\sigma_\Delta^2+\left(\frac{\pi_{pz}}{\pi_{pp}}\right)^2 2^{-2\kappa_z}\sigma_z^2\right]$$

Optimal Price Setting under Rational Inattention

- **Optimal Allocation of Attention:**

$$\kappa_\Delta^* = \begin{cases} \kappa & \text{if } \sigma_\Delta^2 / \left(\frac{\pi_{pz}}{\pi_{pp}}\sigma_z\right)^2 \geq 2^{2\kappa} \\ \frac{1}{2}\kappa + \frac{1}{4}\log_2\left(\sigma_\Delta^2 / \left(\frac{\pi_{pz}}{\pi_{pp}}\sigma_z\right)^2\right) & \text{if } \sigma_\Delta^2 / \left(\frac{\pi_{pz}}{\pi_{pp}}\sigma_z\right)^2 \in \left[2^{-2\kappa}, 2^{2\kappa}\right] \\ 0 & \text{if } \sigma_\Delta^2 / \left(\frac{\pi_{pz}}{\pi_{pp}}\sigma_z\right)^2 \leq 2^{-2\kappa} \end{cases}$$

- See paper for the detailed GE equilibrium
- Take-home lesson:
  - Variances of idiosyncratic shocks are larger so endogenous inattentive to the aggregate nominal shock
  - Micro-foundation of the Phillips curve
  - Monetary shocks have real effects

Rational Inattention: Beyond Shannon

Implicit assumptions embedded in the Shannon cost:

- Each pair of states is equally difficult to distinguish
- 2 vs 2.00001; 2 vs 1000;

Alternatives to the Shannon cost:

- Hébert & Woodford (21, AER): "Neighborhood-Based Information Costs"
  - certain pairs of states are easy to distinguish, whereas others are difficult to distinguish
- Pomatto, Strack & Tamuz (23, AER). "Log-likelihood Ratio Cost"
  - axiom 1: the cost of generating two independent signals is the sum of their costs
  - axiom 2: generating a signal with probability half costs half its original cost

Rational Inattention: Beyond Shannon

Uniformly posterior separable (UPS) cost:

$$C(\mu) = \mathbb{E}_{q\sim\mu}\left[H(q)\right] - H(p),$$

- Cost of acquiring information depends only on distribution of posterior beliefs q and the function H does not depend on prior.
- Nests Shannon, Hébert & Woodford, Pomatto, Strack & Tamuz
- Denti (22, AER); Caplin, Dean & Leahy (22, JPE): axioms for the UPS cost and the entropy cost
- Dean & Neligh (23, JPE): experimental evidence: some support of UPS but not entropy
- Hebert and Woodford (23, JET), Morris and Strack (2019), and Bloedel and Zhong (2021) provide conditions under which the ex ante cost of a sequential sampling procedure is posterior separable

Rational Inattention: General Problems based on the Shannon Cost

Fully general cost:

- Caplin & Dean (15, AER): Axioms for whether a dataset is consistent with a model of information acquisition that puts no restrictions on the cost of information.

Static multidimensional rational Inattention:

- Koszegi and Matejka (2020, QJE)

Dynamic rational inattention and learning:

- Maćkowiak, Matějka, & Wiederholt (18, JET); Afrouzi & Yang (21, with a nice toolbox!); Miao, Wu, & Young (22, ECMA)

Beyond Linear-Quadratic-Gaussian:

- Matejka (2016); Jung, Kim, Matejka, Sims (19, Restud): leading to discrete actions

Rational Inattention: Applications

- Matějka & McKay (15, AER): rational inattention as a foundation to discrete choice models
- Maćkowiak & Wiederholt (15, Restud): business cycle dynamics under rational inattention.
- Nieuwerburgh & Veldkamp (09, JF); Nieuwerburgh, & Veldkamp (10, Restud); Kacperczyk, Nieuwerburgh & Veldkamp (16, ECMA): portfolio choices

Pause for Questions

Outline

Overview: Gabaix (2014)

Sparsity:

- A vector $m \in \mathbb{R}^{1,000,000}$ if it has a few non-zero elements

Gabaix (2014): A model of inattention that

- Handles multi-dimensional problems in a tractable way
- Generate the "sparsity" pattern

An Unconstrained Sparsity Problem

- **Objective** (quadratic or quadratic approximation)

$$u\left(a,\vec{\theta}\right) \quad \text{with} \quad \vec{\theta}=(\theta_1,\cdots,\theta_N)$$

- Rational, non-sparse, **optimal action**

$$a^r=\sum_{i=1}^{N} a_{\theta_i}\theta_i$$

- **Attention:** $m_i\in[0,1]$ (normalize $\theta_i^d=0$)

$$\theta_i^s\left(\vec{m}\right)=m_i\theta_i$$

  - no noise & deterministic inattention/belief

- **Sparse action** (optimal action given sparsely perceived fundamental):

$$a^s\left(\vec{m}\right)=\sum_{i=1}^{N} a_{\theta_i}\theta_i^s$$

An Unconstrained Sparsity Problem

- **Optimal attention allocation:**

$$\max_{\vec{m}} u\left(a^s\left(\vec{m}\right),\vec{\theta}\right)-\mathscr{C}\left(\vec{m}\right)$$

- **Loss:** $L\left(\vec{m}\right)\equiv\mathbb{E}\left[u\left(a^s\left(\vec{m}\right),\vec{\theta}\right)-u\left(a^r,\vec{\theta}\right)\right]$

$$L\left(\vec{m}\right)=\frac{1}{2}\sum_{i,j\in\{1,\cdots,N\}}\left(1-m_i\right)\Lambda_{ij}\left(1-m_j\right),$$

where $\Lambda_{ij}=-Cov\left(\theta_i,\theta_j\right)a_{\theta_i}u_{aa}a_{\theta_j}$ and $a_{\theta_i}=-u_{aa}^{-1}u_{a\theta_i}$,

- **Cost of attention:**

$$\mathscr{C}\left(\vec{m}\right)=\kappa\sum_{i=1}^{N}m_i^{\alpha}$$

- **Equivalent problem:**

$$\min_{\vec{m}}\frac{1}{2}\sum_{i,j\in\{1,\cdots,N\}}\left(1-m_i\right)\Lambda_{ij}\left(1-m_j\right)+\kappa\sum_{i=1}^{N}m_i^{\alpha}$$

The One-Dimensional Problem

One-dimensional problem: $u(a,\theta) = -\frac{\omega}{2}(a_i - \theta)^2$

$$\min_m \frac{\omega}{2}(1-m)^2 \sigma_\theta^2 + \kappa m^\alpha$$

Optimal attention:

$$m^* = \mathscr{A}_\alpha\left(\frac{\omega\sigma_\theta^2}{\kappa}\right),$$

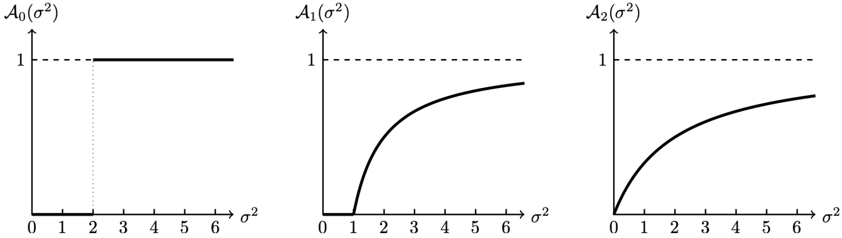
where $\mathscr{A}_\alpha$ is the "attention function."

Compared to the **entropy-based problem:**

$$\min_\lambda \frac{\omega}{2}\sigma^2(1-\lambda) + \frac{\eta}{2}\log_2\left(\frac{1}{1-\lambda}\right)$$

- Essentially just a change of the functional form

The Attention Function and Sparsity ("$\alpha \leq 1$")



Differences between Sparsity and Rational Inattention

Deterministic vs Stochastic

- Sparsity: Decisions are deterministic
- Rational inattention: Decisions are stochastic
- Huge literature in decision theory on stochastic decisions
- Not that relevant for macro (law of large numbers)

Simpler for multi-dimensional problems

- Independently distributed fundamentals: equally simple
- Correlated fundamentals: sparsity simpler
  - cost function separate: $\mathscr{C}(\vec{m}) = \kappa \sum_{i=1}^{N} m_i^{\alpha}$

Sparse Optimization with Constraints

$$\max u(\vec{c}) \quad \text{s.t.} \quad \vec{p} \cdot \vec{c} \leq w$$

- **Question:** how to maintain the budget with behavioral mistakes?
- **Solution:** MRS optimality holds with perceived prices
$$\frac{\partial u}{\partial c_i} / \frac{\partial u}{\partial c_j} = p_i^s / p_j^s$$
- **Sparse decision:**
$$u'(\vec{c}^s(\lambda)) = \lambda \vec{p}^s \quad \text{and} \quad p \cdot \vec{c}^s(\lambda) = w$$
- Optimal attention part same
- Implications: asymmetric Slutsky Matrix

Sparse Optimization with Constraints

Alternative: **the "residual" decision approach**

$$u(\vec{c}) + v(y) \quad \text{s.t.} \quad \vec{p} \cdot \vec{c} + y \leq w,$$

where $v(y)$ is the "continuation" value.

- $\vec{c}$ made "behaviorally and y absorbs the changes

Example (Sims, 03): value from borrowing and saving.

Open questions: in practice, agents make decisions sequentially

- May have "hints" if "closer" to the constraint
- "Foundations" of those reduced-form approaches

Pause for Questions

Outline

Overview: Mankiw & Reis (2002)

Concept of **sticky information**:

- Pricing decisions are not always based on current information
- Only a fraction of firm obtains new info each period
- Firms can always change their prices (no sticky "prices")
- Key prediction: inertia to the inflation process

Comparison with the **sticky price** (Calvo) model:

- Though price is sticky, inflation is forward looking and "jumpy"

$$\pi_t = \kappa y_t + \beta E_t [\pi_{t+1}]$$

- Trouble explaining why inflation is so persistent
- Trouble explaining why shocks to monetary policy have a delayed and gradual effect on inflation ("hump-shape" response)

Differences between sticky-info and noisy-info more nuanced

- Especially in terms of macro implications

Sticky Information

- In each period
  - a fraction λ of firm obtains new info and re-optimizes
  - the other firms continue to set prices based on outdated info
- Firm sets its price every period
  - can change prices even based on outdated info
  - e.g., anticipated news & steady-state inflation
- Connection with noisy-info/rational-inattention
  - a fraction λ of firm is "attentive" this period
  - the rest $1-\lambda$ of firm is "inattentive" this period
  - perfectly precise signals about the current shock if "attentive"
  - infinitely imprecise signals if "inattentive"

Pricing Decisions under Sticky-info

- "Desired"/"Target" price:
$$p_t^* = p_t + \alpha y_t,$$
where p_t is nominal price and y_t is output gap.

- A firm that last updated its information k periods ago sets the price
$$p_t^k = E_{t-k}\left[p_t^*\right]$$

- The aggregate price level
$$p_t = \sum_{k=0}^{+\infty} \lambda \left(1-\lambda\right)^k p_t^k$$

Sticky Information PC

- From the previous slides

$$p_t = \sum_{k=0}^{+\infty} \lambda (1-\lambda)^k E_{t-k} [p_t + \alpha y_t]$$

- Take a first difference and collect terms

$$\pi_t = \frac{\alpha\lambda}{1-\lambda} y_t + \lambda \sum_{j=0}^{+\infty} (1-\lambda)^j E_{t-1-j} [\pi_t + \alpha \Delta y_t],$$

  where $\Delta y_t = y_t - y_{t-1}$
  - Past expectations about current conditions
- Compared to NKPC

$$\pi_t = \frac{\alpha\lambda^2}{1-\lambda} y_t + E_t [\pi_{t+1}],$$

  where λ is prob. of adjustment and the discount rate $\beta = 1$.
  - Current expectations about future conditions

A Sudden Permanent Drop in Agg. Demand $m_t = p_t + y_t$

- Key feature of sticky info: **inertia/hump-shape** in π_t response

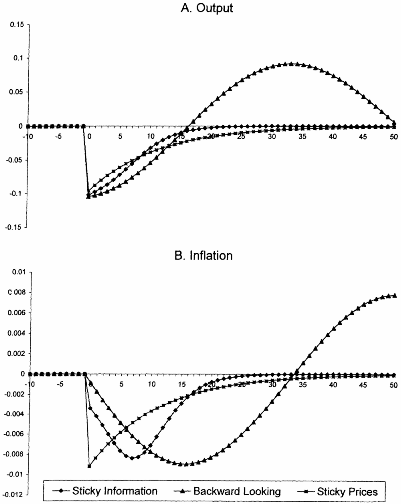


FIGURE I
Dynamic Paths after a 10 Percent Fall in the Level of Aggregate Demand at Time 0

Micro-foundation of Sticky Information

Reis (2006): Inattentive Producers

- **Fixed cost** to acquire information
- Difference from RI:
  - cost does not depend on the "distance" between prior and posterior
- The producer optimally only **updates her info sporadically**
  - inattentive in between

Alvarez et al. (2016): Monetary Shocks in Models with Inattentive Producers

- A general version of Reis (2006) beautifully solved

Matejka (2016): Rationally Inattentive Seller: Sales and Discrete Pricing

- Entropy approach but beyond Gaussian
- Prices tend to remain constant for a period of time and then jump back and forth between a few given values

Sticky Information v.s. Noisy Information

Differences more nuanced, especially in terms of macro implications

First moments:

- Both: **average expectations under-reacts** to aggregate shock
  - initial underreaction
    $$\frac{d\bar{E}_0[m_0]}{dm_0} = \lambda$$
  - slow learning (soon)
- Both: generate inertia in inflation dynamics

Second moments/disagreement: (Coibion & Gorodnichenko, 2012)

- Sticky info: more disagreement after a larger realization of shocks
- Noisy info: in a stationary environment
  - disagreement independent of the realization of shocks

Pause for Questions

Outline

Overview

A conceptual issue: what is the "noise" in rational inattention?

- Either attentive (the true value) or inattentive (the prior)
- What is the interpretation of the in-between (noise)?

Woodford (20): Modeling Imprecision in Perception, Valuation and Choice

- Imprecise perceptual judgments in psychophysics

Psychophysics

Measure and model the relationship between

- Objective physical properties of environment
- The way these are subjectively perceived

Finding: the **stochasticity** of subjective judgments

- People unable to make completely accurate comparisons
  - which of two weights is heavier, which of two lights is brighter, etc
- But probability of a given response increases monotonically with increases in its relative magnitude
  - psychometric function

Imprecise Perceptual Judgments

- Internal representation of a stimuli x summarized by

$$s \sim \mathcal{N}\left(m(x), v^2\right)$$

  - "noisy" signal, but allowing mental coding to be in a "different space"

- As a result,

$$\text{Prob}[\text{''}x_2 \;\; \text{greater} \;\; x_1\text{''}|x_1, x_2] = \Phi\left(\frac{m(x_2) - m(x_1)}{\sqrt{2}v}\right)$$

- "Weber's Law"
  - discrimination threshold increases in proportion to the **relative difference** between two stimuli
  - $m(x)$ is the logarithm of x

Encoding and Decoding as Distinct Processes

Encoding: from the stimuli to the international representation

$$s \sim \mathcal{N}\left(m(x), v^2\right)$$

Decoding: inference based on the international representation s, $E\left[x|s\right]$

- Similar to inference based on a "noisy" signal

Why distinct process?

- Internal representation measurable with development of neuro-science
- Encoding/Decoding process can incorporate other cognitive features
  - Efficient coding theory (encoding)
  - Selective retrieval from memory (decoding)

Application: Small-Stakes Risk Aversion

- Empirical studies documenting **small-stakes risk aversion**
- Rabin (00): based on expected utility theory, small-stakes risk aversion
  - leads to extraordinary risk aversion with respect to large gambles
- Standard explanation: loss aversion (Kahneman & Tversky, 79)
- Khaw, Li, & Woodford (21, Restud). Cognitive imprecision and small-stakes risk aversion.

Cognitive Imprecision and Small-stakes Risk Aversion

- A choice between
  - money $C > 0$ with certainty
  - a gamble that pays X with probability 1/2, but has probability 1/2 of paying nothing
  - C and X vary across trials
- Internal representation of C and X

$$s = (s_C, s_X)$$

Cognitive Imprecision and Small-stakes Risk Aversion

- If risk neutral, accept the game if and only if

$$E[X|s] > 2E[C|s]$$

- Probability of acceptance:

$$\Phi\left(\frac{\log(X/C) - \beta^{-1}\log 2}{\sqrt{2}v}\right),$$

where β increases with the precision of the signal.

- "Indifference" point:

$$X^{indiff}/C = 2^{1/\beta} > 2 \Longrightarrow \text{Risk Aversion}$$

- Caveat: coding in the "log" space important
- Additional evidence: measured risk aversion
  - correlated with degree of randomness in responses
  - increases with cognitive loads of subjects in laboratory experiments

Cognitive Imprecision and Small-stakes Risk Aversion

- If risk neutral, accept the game if and only if

$$E\left[X|s\right] > 2E\left[C|s\right]$$

- Probability of acceptance:

$$\Phi\left(\frac{\log\left(X/C\right) - \beta^{-1}\log 2}{\sqrt{2}v}\right),$$

  where β increases with the precision of the signal.

- "Indifference" point:

$$X^{indiff}/C = 2^{1/\beta} > 2 \Longrightarrow \text{Risk Aversion}$$

- Caveat: coding in the "log" space important

- Additional evidence: measured risk aversion
  - correlated with degree of randomness in responses
  - increases with cognitive loads of subjects in laboratory experiments

Outline

A New Approach to Narrow Bracketing

Economic decisions are made disjointly ("**narrow bracketing**")

- "We tend to make decisions as problems arise, even when we are specifically instructed to consider them jointly."

— *Kahneman (2011)*

Existing model of narrow bracketing

- Directly **imposing each decision is made in isolation**

This paper: **narrow thinking** approach to narrow bracketing (Lian, 2021, Restud)

- Application: a "smooth" model of **mental accounting**
- Connect mental accounting with narrow bracketing

Narrow Thinking (Lian, 2021)

Definition: Diff. decisions based on **diff. & non-nested information**

- When buying food, know food price, but not gasoline price & v.v.
- Psy foundation: bounded recall & "what you see is all there is"

Rep.: **Incomplete info**, common interest, game among multiple selves

- Each decision is made with imperfect perception of other decisions

Essence: Capture **difficulty in coordinating multiple decisions**

- As if each decision is made caring less about other decisions
- A smooth model of narrow bracketing

Narrow Thinking $\Longrightarrow$ Narrow Bracketing

A simple consumer theory example:

$$u(x_1, x_2, \vec{p}) = v(x_1, x_2) + w - p_1 x_1 - p_2 x_2$$

Narrow thinker: self i knows p_i & receives a noisy signal about other p_{-i}

A smooth model of **narrow bracketing**

$$\frac{\partial x_i^{\text{Narrow}}}{\partial p_i} = \omega_i \frac{\partial x_i^{\text{Neglect}}}{\partial p_i} + (1 - \omega_i) \frac{\partial x_i^{\text{Standard}}}{\partial p_i}$$

- $x_i^{\text{Neglect}}(\vec{p})$: completely neglecting the other decision

Intuition.

- x_{-i} not as responsive to p_i.
- Indirect effect through x_{-i} dampened.
- Effectively cares less about the other decision

Application: a Smooth Model of Mental Accounting

Application: **mental accounting**

- Separable, non quasi-linear utility. Interaction from the budget.

$$\sum_{i=1}^{N} v_i(x_i) + h\left(w - \sum p_i x_i\right)$$

- A **smooth** model of mental accounting

$$\frac{\partial x_i^{\text{Narrow}}}{\partial p_i} = \omega_i \frac{\partial x_i^{\text{Explicit}}}{\partial p_i} + (1 - \omega_i) \frac{\partial x_i^{\text{Standard}}}{\partial p_i} \quad \forall i,$$

- $x_i^{\text{Explicit}}(\vec{p})$: **explicit budget**, e.g. allocates \$100 to food
- **Intuition**: for $x_i^{\text{Explicit}}(\vec{p})$, each decision can be made in isolation

Narrow bracketing/mental accounting, by the same underlying friction

New predictions about what drives the **degree** of mental accounting

- Spending shares
- Cognitive limitations

Related Literature

Difference from **rational inattention/sparsity**:

- Imperfect knowledge about the fundamental (e.g. prices)
- **Same information** for **all** decisions
  - Gabaix (2014); Koszegi & Matejka (2020)
  - When buy food/gasoline, always knows food but not gasoline price
- **Fully consider other decisions' influence** on current decision

Pause for Questions