Foundations for Optimal Inattention

Andrew Ellis

London School of Economics

March 2014

Motivation

- Vast amount of information is freely available
- Individuals appear not to pay attention to all of it

Inattention can imply:

1. Delayed response to shocks (Sims (1998, 2003))
2. Sticky-prices (Mackowiak and Wiederholt (2009))
3. Under-diversification (Van Nieuweburgh and Veldkamp (2010))
4. Sticky investment (Woodford (2008))
5. Coordination failure (Hellwig and Veldkamp (2009))
6. Specialization (Dow (1991))
7. Price discrimination (Rubinstein (1993))
8. Extreme price swings (Gul, Pesendorfer and Strazlecki (2011))

This Paper

- Studies an agent who may **exhibit inattention** to information
  - **Coarser, subjective information** rather than objective info.
  - Behavior not well understood
    - ⋆ Cost of attention, subjective information and underlying preference are all **unobservable**
- Provides **behavioral foundations** for the **optimal inattention** model based on **conditional choices**
  - Subjective information **maximizes EU** net of cost
  - Actions **maximize EU** given subjective information
  - Agent **considers every feasible action**
    - ⋆ Explicitly rule out inattention to **the available alternatives** (e.g. Masatlioglu, Nakajima and Ozbay (2012))

Contributions

- Clarify and justify the assumptions made by previous work
- What can be identified from choices?
  - Which choices reveal optimal inattention?
  - How can one identify tastes, subjective information and attention cost?
- What does this theory imply about choices?
  - Conditional choices are **observable** but violate almost all of the properties satisfied by a fully attentive EU DM
    - Weak Axiom of Revealed Preference (WARP), Independence, Continuity and Consequentialism are all violated
    - Conditions on different information when facing different problems

Example

- Benevolent doctor treats patients suffering from a given disease
- Three drugs (Generic (g), Merck (m) and Pfizer (f)) treat it
  - One of the three will be strictly more effective than the others
  - Which works best for a given patient is initially unknown, and the doctor can, in principle, determine it
- Modeled as a three state decision problem

$$\Omega = \{generic, merck, pfizer\}$$

  - State indicates which drug is the most effective
- Objective information is represented by the partition

$$P = \{\{generic\}, \{merck\}, \{pfizer\}\}$$

Example

- Two patients are identical except for their insurance plans
  - One's plan covers all three drugs
  - Other's plan does not cover Pfizer's drug
- Prescribing a drug is choosing an act
  - Acts give state-contingent outcomes (patient's health)
- Each patient corresponds to a choice problem
  - First patient is $\{g, m, f\}$
  - Second patient is $\{g, m\}$

	generic	*merck*	*pfizer*
$c(\{g, m, f\} \vert \cdot)$	$\{m\}$	$\{m\}$	$\{f\}$
$c(\{g, m\} \vert \cdot)$	$\{g\}$	$\{m\}$	$\{m\}$

Example

	generic	*merck*	*pfizer*
$c(\{g, m, f\} \mid \cdot)$	$\{m\}$	$\{m\}$	$\{f\}$
$c(\{g, m\} \mid \cdot)$	$\{g\}$	$\{m\}$	$\{m\}$

- Which choices reveal she **does not process** all information?
  - If she processes all information, then in a given state, each choice maximizes the same conditional preference relation
  - So her choices satisfy WARP
- Choices conditional on the state "*generic*" violate WARP
  - Implies that the consumer **does not** pay attention to P

Example

	generic	*merck*	*pfizer*
$c(\{g,m,f\}\|\cdot)$	$\{m\}$	$\{m\}$	$\{f\}$
$c(\{g,m\}\|\cdot)$	$\{g\}$	$\{m\}$	$\{m\}$

- To what **does** she pay attention?

Observation

If $c(B|\omega) \neq c(B|\omega')$, then ω and ω' must be in different cells of her subjective information when facing B

- Asks "Is Pfizer's drug the most effective?" when faces $\{g,m,f\}$
- Asks "Is generic drug the most effective?" when faces $\{g,m\}$

Aside: Observability

This data is the natural generalization of the domain considered by the papers studying implications of inattention

How can one observe these conditional choices?

1. Laboratory
2. Any setting in which information and state are iid
    - Doctor treats "many" patients who are *a priori* identical except for their insurance plans
    - Modeler learns state by observing reaction to treatment

Outline

1. Model
2. Foundations
3. Representation Theorem
4. Identification

Formal Setup

- **Anscombe-Aumann setting** with information
  - Objective state space: Ω
  - Objective information: P (finite partition)
  - Acts: mappings from Ω to (lotteries over) consequences
- **Data:** DM's conditional choice correspondence
  - Choice from **each** feasible set of acts in **every** state
  - She chooses $c(B|\omega)$ from the problem B in the state ω
    - ⋆ $c(B|\cdot)$ must be P-measurable

The Model

$c(\cdot)$ has **an optimal inattention representation** if there exists a $(u, \pi, \gamma, \hat{P})$ so that for every problem B,

$$\hat{P}(B) \in \arg\max_{Q} [\sum_{E \in Q} \pi(E) \max_{f \in B} \int u \circ f d\pi(\cdot|E) - \gamma(Q)]$$

and for every problem B and state ω,

$$c(B|\omega) = \arg\max_{f \in B} \int u \circ f d\pi(\cdot|\hat{P}(B)(\omega))$$

The Model

Where:

- u is an affine, continuous, and unbounded **utility index**
- π is a **prior** probability measure on Ω
  - Assigns positive probability to each cell of P
- γ is the **attention cost function**
  - maps each partition to a cost between 0 and ∞
  - $\gamma(\{\Omega\}) = 0$
  - $Q \gg R \implies \gamma(Q) \geq \gamma(R)$
- $\hat{P}$ is an **attention rule** mapping problems to subjective information
  - DM pays attention to $\hat{P}(B)$ when facing B

The Model: Special Cases

- A **fully attentive DM** processes all information
  - $c(\cdot)$ has a full attention representation if

$$c(B|\omega) = \arg\max_{f \in B} \int u \circ f d\pi(\cdot|P(\omega))$$

  - Standard Dynamic SEU model
  - $\gamma(P) = 0$ and $\hat{P}(B) = P$ for all B
- A **DM with fixed attention** processes the same information, regardless of the problem faced
  - $c(\cdot)$ has a fixed attention representation if

$$c(B|\omega) = \arg\max_{f \in B} \int u \circ f d\pi(\cdot|Q(\omega))$$

    for some Q that is coarser than P
  - $\gamma(Q) = 0$ and $\gamma(R) < \infty$ only if $Q \gg R$
  - $\hat{P}(B) = Q$ for all B

The Model: Special Cases

$c(\cdot)$ has **a constrained attention representation** if there exists a $(u, \pi, \mathbb{P}^*, \hat{P})$ so that for every problem B,

$$\hat{P}(B) \in \arg\max_{Q \in \mathbb{P}^*} \sum_{E \in Q} \pi(E)[\max_{f \in B} \int u \circ f d\pi(\cdot|E)],$$

where $\mathbb{P}^*$ is an **attention constraint**

and for every problem B and state ω,

$$c(B|\omega) = \arg\max_{f \in B} \int u \circ f d\pi(\cdot|\hat{P}(B)(\omega))$$

- Special case where $range(\gamma(\cdot)) = \{0, \infty\}$

Aside: Alternative Settings

- All conditional choices from a fixed problem, e.g. Sims (2003)
  - **Cannot identify** tastes, prior or constraint
  - Admits testable implications **under additional assumptions**
- Choice from each problem, in a fixed state (Van Zandt, 1996)
  - **No testable implications**
  - Does not model underlying uncertainty
- Preference over menus (de Oliveira-Denti-Mihm-Ozbek, 2013)
  - Implied properties for choice of actions unclear
- Stochastic conditional choice (Caplin and Dean, 2013)
  - Focus on testability rather than identification and behavior
  - Signals rather than partitions – equivalent if Savage-style state space where information is part of state

Outline

1. Model
2. Foundations
    1. Independence of Never Relevant Acts
    2. Attention Constrained Independence
    3. Monotonicity
    4. Subjective Consequentialism
    5. Continuity
    6. Unbounded
3. Representation Theorem
4. Identification

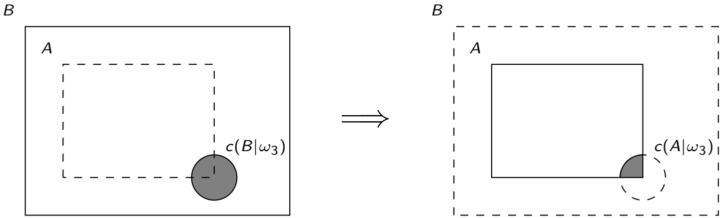
Weak Axiom of Revealed Preference

- WARP, aka Independence of **Irrelevant** Acts, requires that

$$[A \subset B \,\&\, c(B|\omega) \cap A \neq \emptyset] \implies c(A|\omega) = c(B|\omega) \cap A$$

  - Only considers choices in state ω

- Graphically (with $\Omega = \{\omega_1, \omega_2, \omega_3\}$):



- Violates WARP only if she pays attention to different information

- Independence of **Never** Relevant Acts gives one situation where an optimally inattentive DM **does not violate** WARP

Axiom 1: Independence of Never Relevant Acts (INRA)

For any $A \subset B$, if $c(B|\omega') \cap A \neq \emptyset$ for **every** state ω', then $c(A|\omega) = c(B|\omega) \cap A$ for any ω

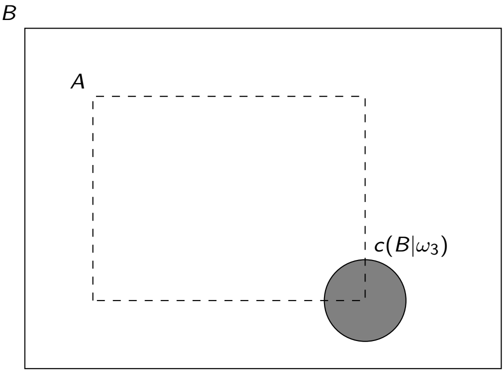
- If two patients differ only in that one's plan drops the drug h but the doctor **never** prescribes h to the patient with better insurance, then she prescribes the same drugs to both

INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:

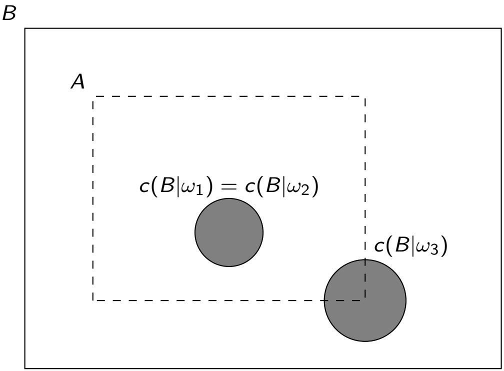
INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



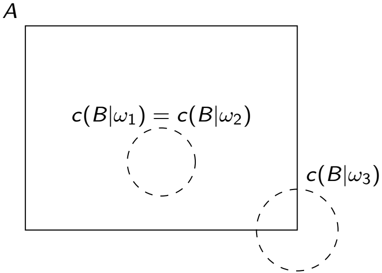
INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



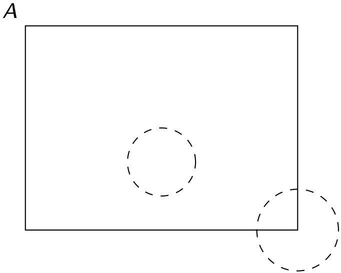
INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



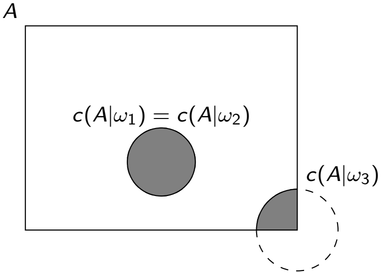
INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



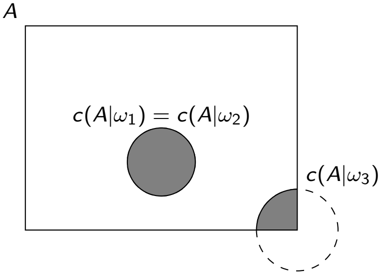
INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



INRA

Graphically, with $\Omega = \{\omega_1, \omega_2, \omega_3\}$:



Independence of Never Relevant Acts

- Fix a problem B and an act f so that $\{f\} \neq c(B|\omega')$ for all ω'
  - f is "never relevant"
- Set $A = B \setminus \{f\}$, noting that $A \cap c(B|\omega') \neq \emptyset$ for all ω'
- Suppose her subjective information is Q when facing B
  - Conditional on any cell of Q, an act in A at least as good as f
  - Benefit from Q when facing A **is the same as** when facing B
  - Q **optimal** when facing $B \implies Q$ **still optimal** when facing A
- The DM should **pay attention to the same** information when facing B as when facing A
- So her choices from A and B **should not** violate WARP

INRA: Example and Counterexample

The first doctor chose:

	generic	*merck*	*pfizer*
$c(\{g, m, f\}\|\cdot)$	$\{m\}$	$\{m\}$	$\{f\}$
$c(\{g, m\}\|\cdot)$	$\{g\}$	$\{m\}$	$\{m\}$

A second doctor chooses:

	generic	*merck*	*pfizer*
$c'(\{g, m, f\}\|\cdot)$	$\{m\}$	$\{m\}$	$\{m\}$
$c'(\{g, m\}\|\cdot)$	$\{g\}$	$\{m\}$	$\{m\}$

- First satisfies INRA (but not WARP)
- Second violates INRA and cannot have optimal inattention repn.

Notation

For any problems A, B and any $\alpha \in [0,1]$,
$\alpha A + (1-\alpha)B = \{\alpha f + (1-\alpha)g : f \in A, g \in B\}$

- Suppose $f \in c(B|\omega)$, $g \in c(C|\omega)$, $h \in c(D|\omega)$, and $\alpha, \beta \in (0,1)$
- **Independence** implies

$$\begin{aligned} & (1-\alpha)f + \alpha g \in c((1-\alpha)B + \alpha C|\omega) \\ \iff \quad & (1-\beta)f + \beta h \in c((1-\beta)B + \beta D|\omega) \end{aligned}$$

- If attends to same info. when facing B, C, D, $(1-\alpha)B + \alpha C$ and $(1-\beta)B + \beta D$, then an inattentive DM does not violate Independence

- Attention Constrained Independence gives one situation where the DM **does not violate** Independence

Axiom 2: Attention Constrained Independence (ACI)

For every $\alpha \in [0,1]$, $\omega \in \Omega$, problem B and acts f, h, h':

$$(1-\alpha)f + \alpha h \in c((1-\alpha)B + \alpha\{h\}|\omega) \iff (1-\alpha)f + \alpha h' \in c((1-\alpha)B + \alpha\{h'\}|\omega)$$

- If there is a (state-independent) probability α that the patient will take drug h regardless of what the doctor prescribes, then her choices are **unaffected** by the identity of h
- Independence holds when $\alpha = \beta$ and C, D are singletons
- Choice MAY be affected by magnitude of α
  - Choose from B but get h (h') with probability α
  - α "small" implies likely to get choice
  - α "large" implies unlikely to get choice

Attention Constrained Independence

- Fix problems B, $\{h\}$, and $\{h'\}$
- $\{h\}$, $\{h'\}$ are singletons $\implies$ DM's choice is the same no matter what her subjective information is
  - Difference between the benefits of any two partitions **is the same** for $\alpha B + (1-\alpha)\{h\}$ and $\alpha B + (1-\alpha)\{h'\}$
- Q **optimal** when facing $\alpha B + (1-\alpha)\{h\}$ if and only if
  Q **optimal** when facing $\alpha B + (1-\alpha)\{h'\}$
- The DM **pays attention to the same** information when facing $\alpha B + (1-\alpha)\{h\}$ as when facing $\alpha B + (1-\alpha)\{h'\}$
- So her choices from $\alpha B + (1-\alpha)\{h\}$ and $\alpha B + (1-\alpha)\{h'\}$ **do not** violate Independence

For two constant acts x and y, say that x is ***revealed* (**resp. ***strictly*) *preferred*** to y if there exists a state ω so that $x \in c(\{x, y\}|\omega)$ (resp. and $y \notin c(\{x, y\}|\omega)$)

Axiom 3: Monotonicity

If $f, g \in A$ and $f(\omega')$ is **revealed preferred** to $g(\omega')$ for every ω', then $g \in c(A|\omega) \implies f \in c(A|\omega)$
Moreover, if $f(\omega')$ is revealed **strictly** preferred to $g(\omega')$ for every $\omega' \in P(\omega)$, then $g \notin c(A|\omega)$

- Restatement of Anscombe-Aumann monotonicity

- Tastes are **state independent**
- DM **never chooses** an act that **always** yields a worse outcome

Axiom 4: Subjective Consequentialism

For each choice problem B and state ω:
If "the only states in which f and g differ must be in a different cell of her subjective information than ω,"
then whenever $f, g \in B$

$$f \in c(B|\omega) \iff g \in c(B|\omega)$$

- The DM's choice between any two acts is unaffected by their consequences in states that she knows did not occur
- Recall: $c(B|\omega) \neq c(B|\omega')$ only if ω and ω' in different cells of subjective information

Axiom 4: Subjective Consequentialism

For each choice problem B and state ω:
If $f(\omega) = g(\omega)$ and
for all $\omega' \neq \omega$ either $f(\omega') = g(\omega')$ or $c(B|\omega') \neq c(B|\omega)$,
then whenever $f, g \in B$

$$f \in c(B|\omega) \iff g \in c(B|\omega)$$

- The DM's choice between any two acts is unaffected by their consequences in states that she knows did not occur
- Recall: $c(B|\omega) \neq c(B|\omega')$ only if ω and ω' in different cells of subjective information

Axiom 5: Continuity

Formal Statement

- Technical continuity condition required
- Ensures the continuity of the underlying preference relation
- Complication: the DM's choices from different problems may be conditioned on different information so her choices may appear discontinuous to the modeler
- It is implied by the combination of WARP and upper hemi-continuity

Last axiom guarantees range of $u(\cdot)$ is all of $\mathbb{R}$

X is set of lotteries

Axiom 6: Unbounded

There exist $x, y \in X$ so that $\{x\} = c(\{x, y\}|\omega)$, and for every $\beta \in [0, 1]$, there exists a $z \in X$ so that

$$\{\beta z + (1-\beta)y\} = c(\{\beta z + (1-\beta)y, x\}|\omega)$$

and a $z' \in X$ so that

$$\{y\} = c(\{\beta z' + (1-\beta)x, y\}|\omega)$$

- Standard, technical axiom
- Technical remark: needed for sufficiency proof, not just identification

Outline

Representation Theorem: Sufficiency of Axioms

Theorem

If $c(\cdot)$ satisfies INRA, ACI, Monotonicity, Subjective Consequentialism, Continuity, and Unbounded, then $c(\cdot)$ has an optimal inattention representation.

- The six axioms are sufficient for the DM to behave **as if** she has optimal inattention

Skip To Necessity

Proof Idea

- Choice of acts very poorly behaved
- Consider choice on another domain: plans
- A plan is a mapping from states to acts
  - In example, doctor chooses the plan "pick g in state '*generic*', otherwise pick m" from $\{g, m\}$ and chooses the plan "pick f in state '*pfizer*', otherwise pick m" from $\{g, m, f\}$
- INRA guarantees that her choice over plan maximizes a preference relation $\succeq$
- Other axioms ensure $\succeq$ well-behaved and has desired represention
- Show choosing plan equivalent to choosing info.

Representation Theorem: Necessity of Axioms

- Necessity of axioms complicated because I have not restricted attention to "regular" tie-breaking rules
  - May be more than one optimal partition for a given problem
  - Axioms **necessary given some conditions** on tie-breaking
  - The axioms are **generically necessary** for any tie-break rule
- Since tie-breaking is of secondary interest, my axioms capture the behavioral content of optimal inattention

Representation Theorem: Necessity of Axioms

Theorem

If $c(\cdot)$ has an optimal inattention representation, then $c(\cdot)$ satisfies Monotonicity, Subjective Consequentialism, Continuity and Unbounded.

*Moreover, there exists $c'(\cdot)$ satisfying **all six axioms** and there exists **an open, dense** subset K of choice problems so that:*

1. *$c(\cdot)$ and $c'(\cdot)$ have optimal inattention representations parametrized by $(u(\cdot), \pi(\cdot), \gamma(\cdot), \hat{P}(\cdot))$ and $(u(\cdot), \pi(\cdot), \gamma(\cdot), \hat{Q}(\cdot))$*
2. *$c(B|\omega) = c'(B|\omega)$ **for every** $\omega \in \Omega$ and $B \in K$.*

- $c(\cdot)$ and $c'(\cdot)$ differ only because of tie-breaking
- Tie-breaking is **unnecessary** for **"most" problems**

Special Cases

Axiom: Consequentialism

For any ω, if $f, g \in B$ and $f(\omega') = g(\omega')$ for all $\omega' \in P(\omega)$, then

$$f \in c(B|\omega) \iff g \in c(B|\omega).$$

- Standard property of models of choice under uncertainty
- Guarantees that DM respects the objective information partition
- Implies Subjective Consequentialism

Special Cases

Corollary

$c(\cdot)$ satisfies ***Consequentialism*** *as well as INRA, ACI, Monotonicity, Continuity, and Unbounded if and only if $c(\cdot)$ has a full attention representation.*

- Optimally inattentive DM respects objective information partition if and only if she processes all available information
- Novel characterization of dynamic SEU model
- Only behavioral difference between optimally inattentive and fully attentive is latter must satisfy Consequentialism

Special Cases

Axiom: Independence

$f \in c(A|\omega) \& g \in c(B|\omega) \iff \alpha f + (1-\alpha)g \in c(\alpha A + (1-\alpha)B|\omega)$

Corollary

The following are equivalent:
*(i) $c(\cdot)$ satisfies **WARP** as well as ACI, Subjective Consequentialism, Monotonicity, Continuity, and Unbounded*
*(ii) $c(\cdot)$ satisfies **Independence** as well as INRA, Subjective Consequentialism, Monotonicity, Continuity, and Unbounded*
(iii) $c(\cdot)$ has a fixed attention representation

- WARP and Independence are equivalent for an optimally inattentive DM
- DM violates WARP only if $\hat{P}(B) \neq \hat{P}(A)$
  - If she never violates WARP, then $\hat{P}(B) = \hat{P}(B')$ for all B, B'

Axiom 2*: Strong ACI

For every $\alpha, \beta \in (0,1]$, $\omega \in \Omega$, problem B and act h, h':

$$(1-\alpha)f + \alpha h \in c((1-\alpha)B + \alpha\{h\}|\omega) \iff$$
$$(1-\beta)f + \beta h' \in c((1-\beta)B + \beta\{h'\}|\omega)$$

- Similar interpretation to ACI
- Requires choice unaffected by **magnitude of** α as well as identity of h

- Choose from B but get h (h') with probability α (β)
  - Suppose $\alpha < \beta$
  - α "small" implies likely to get choice
  - β "large" implies unlikely to get choice
  - Pays attention to same information in both cases

Constrained Attention Representation

Corollary

*If $c(\cdot)$ satisfies INRA, **Strong** ACI, Monotonicity, Subjective Consequentialism, Continuity, and Unbounded, then $c(\cdot)$ has a **constrained attention** representation.*

*If $c(\cdot)$ has a **constrained attention** representation, then there exists $c'(\cdot)$ satisfying **the above six axioms** and **an open, dense** subset K of problems so that:*

1. *$c(\cdot)$ and $c'(\cdot)$ have optimal inattention representations parametrized by $(u(\cdot), \pi(\cdot), \mathbb{P}^*, \hat{P}(\cdot))$ and $(u(\cdot), \pi(\cdot), \mathbb{P}^*, \hat{Q}(\cdot))$*
2. *$c(B|\omega) = c'(B|\omega)$ **for every** $\omega \in \Omega$ and $B \in K$.*

Outline

1. Model
2. Foundations
3. Representation Theorem
4. Identification
    1. Uniqueness result
    2. Comparative behavior

Identification

Theorem

If c has an optimal inattention representation, then:

- *The support of γ is **unique***
- *There is a **unique, canonical** (coarsest) $\hat{P}$*
- *u is **unique** up to a positive affine transformation*
- *If the relative likelihood of all events is decision-relevant:*
  - *π is unique* details
  - *γ is unique up to the same affine transformation as u*

- Reason for non-uniqueness: only *ex post* choice observed
- If **either**
  - *ex ante* choice is also observed **or**
  - cost function is "rich enough"

then π and γ are uniquely identified

Identification: Prior

Example: Suppose $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $P = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$.

1. If $\gamma(P) = 0$, then **any** prior represents her choices
   - No choices reveal tradeoffs between any two states
2. If $\gamma(Q) > 0$ for all $Q \neq \{\Omega\}$, then π is **uniquely identified**
   - Optimal to pay attention to $\{\Omega\}$ with "small stakes" bets
3. In general:
   - If $\gamma(P) > \gamma(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$, then when she pays attention to the latter, her choices reveal tradeoff between ω_1 and ω_2
   - If $\gamma(P) > \gamma(\{\{\omega_1, \omega_3\}, \{\omega_2\}\})$, then when she pays attention to the latter, her choices reveal tradeoff between ω_1 and ω_3

Identification: How Is γ Identified?

- Suppose DM only pays attention to either Q or $\{\Omega\}$
- Construct a problem that has
  - "bets" on each cell of Q at stakes x and y
  - constant act giving y for sure
- If x s.t. $u(x) < \gamma(Q)$, then optimal to pay attention to $\{\Omega\}$
- If x s.t. $u(x) > \gamma(Q)$, then optimal to pay attention to Q
- $\gamma(Q)$ is smallest $u(x)$ so that DM pays attention to Q
- Generalizes but must choose bets and partitions carefully

Comparative Behavior

- What is behavioral interpretation of γ?

Definition

$c_1(\cdot)$ **pays more attention than** $c_2(\cdot)$ if for all* B and $\omega, \omega' \in \Omega$

$$c_2(B|\omega) \neq c_2(B|\omega') \implies c_1(B|\omega) \neq c_1(B|\omega')$$

- Whenever DM2 distinguishes ω and ω', so does DM1
  - Implies that $\hat{P}_1(B) \gg \hat{P}_2(B)$ for all B
  - If same ex ante preference, then DM1 gets at least as high expected benefit from information

- Conjecture: $\gamma_1(Q) \leq \gamma_2(Q)$ for all Q if and only if $c_1(\cdot)$ pays more attention than $c_2(\cdot)$

Comparative Behavior

- This conjecture is false
- Suppose $\Omega = \{1, 2, 3\}$ & $\pi(i) = \frac{1}{3}$ for all $i \in \Omega$
  - $u \circ f = (24, 0, 0)$, $u \circ g = (0, 12, 12)$, and $u \circ h = (0, 0, 21)$
  - Let $Q = \{\{1\}, \{2, 3\}\}$ and $R = \{\{1\}, \{2\}, \{3\}\}$
    - ⋆ $\gamma_1(Q) = 4$ and $\gamma_1(R) = 8$
    - ⋆ $\gamma_2(Q) = 7$ and $\gamma_2(R) = 9$
  - Gain from paying attention to x given $\{f, g, h\}$ is $G(x)$
    - ⋆ $G(\{\Omega\}) = 8$, $G(Q) = 16$, $G(R) = 19$
    - ⋆ $G(Q) - \gamma_1(Q) = 12$ and $G(R) - \gamma_1(R) = 11$
    - ⋆ $G(Q) - \gamma_2(Q) = 9$ and $G(R) - \gamma_2(R) = 10$
  - Conclude $\hat{P}_1(\{f, g, h\}) = Q$ and $\hat{P}_2(\{f, g, h\}) = R$

Comparative Behavior

Proposition

Suppose that:

1. $(u, \pi, \gamma_1, \hat{P}_1)$ *represents* $c_1(\cdot)$ *and* $(u, \pi, \gamma_2, \hat{P}_2)$ *represents* $c_2(\cdot)$,
2. $supp(\gamma_1) = supp(\gamma_2)$,
3. *and* $[Q, R \in supp(\gamma_2) \implies Q \gg R \text{ or } R \gg Q]$.

Then:

- c_1 *pays more attention* c_2 *if and only if* $Q \gg R \implies \gamma_1(Q) - \gamma_1(R) \leq \gamma_2(Q) - \gamma_2(R)$

- Marginal, not total, cost impacts ex-post choice
- Implies $\gamma_1(Q) \leq \gamma_2(Q)$ for all Q
- Also requires decreased marginal cost

Discussion and Interpretation

1. Attention is a positive, not a negative (INRA)
   - DM chooses her subjective information because she **wants** to choose acts, not because she **wants to avoid** choosing acts
2. Behavioral distinction between costly and constrained attention is ACI versus strong ACI
3. "Stickiness" is a defining feature of optimal inattention
   - Sims's seminal 1998 paper title: "Stickiness"
   - Identified with violations of consequentialism
4. Dynamic Optimal Inattention
   - Consequentialism violated
   - Dynamic consistency undefined

Thank you

Axiom 5: Continuity (i)

If for every n

$$f_n \in c(B_n|\omega) \,\&\, \forall\omega'[c(B|\omega') \neq c(B|\omega) \implies c(B_n|\omega') \neq c(B_n|\omega)],$$

then $B_n \to B$ and $f_n \to f$ imply that $f \in c(B|\omega)$.
(ii) For any $x, y \in X$ and $\omega \in \Omega$, if $\{y\} \; \overline{IS} \; \{x\}$ and $x \in c(\{x, y\}|\omega)$, then $y \in c(\{x, y\}|\omega)$ as well

- The acts in problem B_1 are **indirectly selected** over the acts in problem B_n ($B_1 \; IS \; B_n$) if there is a finite sequence of problems $B_2, ..., B_{n-1}$ so that the DM chooses an act from B_{i-1} in **every** state of the world when the available acts are B_i
- Let $\overline{IS}$ be the sequential closure of IS

Return

Identification

- Define $BEg = \{fEg : f \in B\}$ for any problems B and act g
  - All acts in BEg equal g on E^c

Definition

The likelihood of E is **not decision-relevant** if for any* B and g:
(i) $\forall \omega \in E : \; f \in c(B|\omega) \iff fEg \in c(BEg|\omega)$, and
(ii) $\forall \omega' \in E^c : \; f' \in c(B|\omega') \iff f'E^c g \in c(BE^c g|\omega')$.

- **Strong condition**
- Can "split up" **any** problem without changing choices

Theorem

If the likelihood of all events is decision-relevant, then π and γ are unique as claimed.* Return

Fix any B, $f, g \in B$ and ω.

Suppose for all ω' either $f(\omega') = g(\omega')$ or $c(B|\omega') \neq c(B|\omega)$

- Consider any $\omega'' \in P(\omega)$
- By measurability, $c(B|\omega'') = c(B|\omega)$
- Since either $f(\omega'') = g(\omega'')$ or $c(B|\omega'') \neq c(B|\omega)$, $f(\omega'') = g(\omega'')$
- Since ω'' arbitary, $f(\omega'') = g(\omega'') \forall \omega'' \in P(\omega)$
- Consequentialism implies that

$$f \in c(B|\omega) \iff g \in c(B|\omega).$$

Return