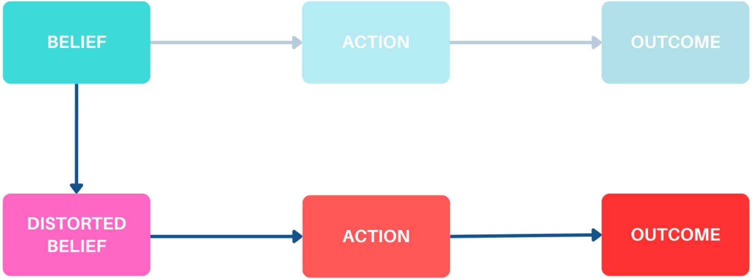
Coherent Distorted Beliefs

C. Chambers Y. Masatlioglu C. Raymond

McGill
Sep, 2024

Traditional View

Motivated Belief



Models of Distorted Beliefs

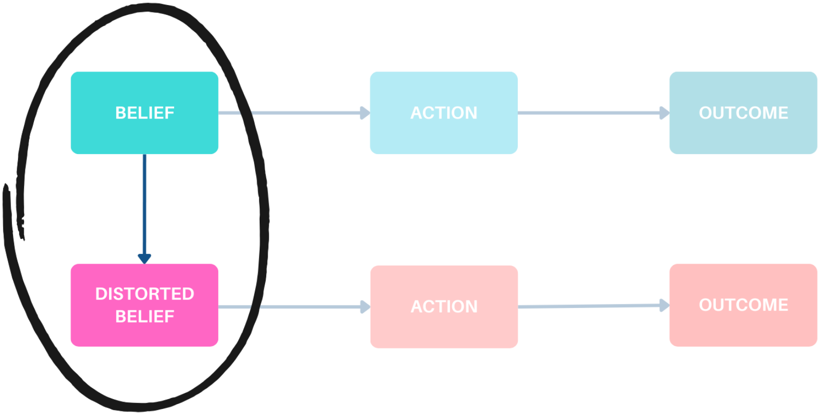
Motivated reasoning about states: Benabou and Tirole (2002), Brunnermeier and Parker (2005), Bracha and Brown (2012), Caplin and Leahy (2019), Mobius et al (2022)

Probabilistic Biases: Grether (1980) Rabin (2002), Benjamin et al (2016), Bodoh-Creed et al (2018), Benjamin (2019), Noor and Payro (2022)

Ambiguity aversion with "distorted beliefs" (e.g., variational preferences): Maccheroni et al (2006), Strzalecki (2011), Cerreia-Vioglio et al (2012)

Nonexpected utility: Kahneman and Tversky (1979), Quiggin (1980), Chew (1983), Tversky and Kahneman (1992)

Belief Distortion



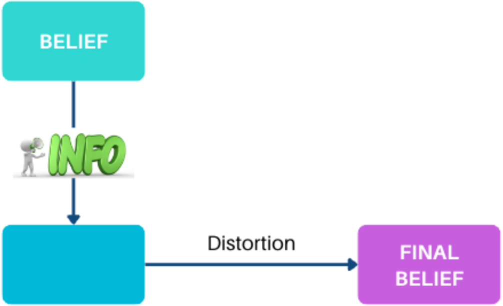
Belief Distortion

$$p \xRightarrow{\quad\quad\phi\quad\quad} q$$

Timing of Information

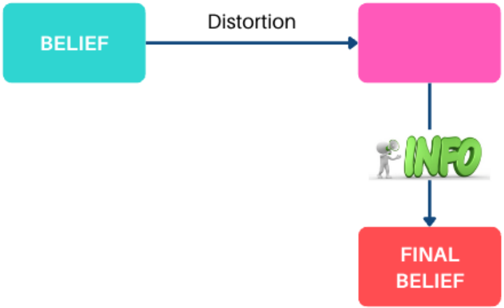
Timing of Information

Case I: Before distortion



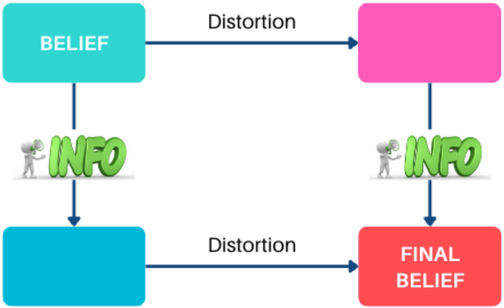
Timing of Information

Case II: After distortion



Coherent Distortion

Distortion invariant to the timing of information



Coherent Distortion

Say ϕ is *coherent* when for every $E \subseteq \Omega$, $\phi(p)(\cdot|E) = \phi(p(\cdot|E))$.

$$\begin{tikzcd}
p \arrow[r, Rightarrow, "\phi"] \arrow[d, "E"'] & q \arrow[d, "E"'] \\
p(\cdot|E) \arrow[r, Rightarrow, "\phi"] & q(\cdot|E)
\end{tikzcd}$$

This paper

Say ϕ is *coherent* when for every $E \subseteq \Omega$, $\phi(p)(\cdot|E) = \phi(p(\cdot|E))$.

$$\begin{tikzcd}
p \arrow[r, Rightarrow, "\phi"] \arrow[d, "E"'] & q \arrow[d, "E"'] \\
p(\cdot|E) \arrow[r, Rightarrow, "\phi"] & q(\cdot|E)
\end{tikzcd}$$

Main Question: What is the class of ϕ that are coherent?

"Subjective Bayesianism"

This paper

Say ϕ is *coherent* when for every $E \subseteq \Omega$, $\phi(p)(\cdot|E) = \phi(p(\cdot|E))$.

$$\begin{tikzcd} p \arrow[r, Rightarrow, "\phi"] \arrow[d, "E"'] & q \arrow[d, "E"'] \\ p(\cdot|E) \arrow[r, Rightarrow, "\phi"] & q(\cdot|E) \end{tikzcd}$$

Main Question: What is the class of ϕ that are coherent?

"Subjective Bayesianism"

Coherent distortions

What motivates coherency?

1. "Subjective Bayesianism:" Our agent acts as if they have a (distorted) prior and then update beliefs with the receipt of information in accordance with Bayes' rule.

2. Distortions are introspection proof: if our agent thinks about what they would do if they potentially learned information, hypothetical updated likelihoods do not contradict current beliefs

3. Final beliefs robust to timing of information versus distortion: don't need to know if our agent distorts and then accesses private information, or the opposite

4. Under the assumption that our agent is an expected utility maximizer conditional on beliefs, our agent will be immune to subjective Dutch books.

Order effects may be real and appear in data. But, *unless we are interested in these types of framing issues*, we may want to avoid this dependence. "First Rationality Check"

Notation

1. Ω: a finite set of *states of the world* with $|\Omega| \geq 3$
2. Information is in the form of $E \subseteq \Omega$ with $p(E) > 0$
3. A *distortion function* $\phi : \Delta(\Omega) \to \Delta(\Omega)$
4. *Positivity*: $\phi(p)(\omega) > 0$ iff $p(\omega) > 0$
5. $p(\cdot|E)$ is the Bayesian update of p after receiving E

Coherence

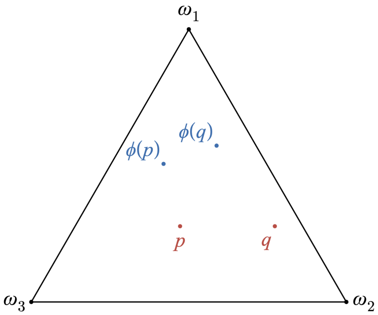
Say ϕ is *coherent* if for every $p \in \Delta(\Omega)$ and every $E \subseteq \Omega$ for which $p(E) > 0$,

$$\phi(p)(\cdot|E) = \phi(p(\cdot|E))$$

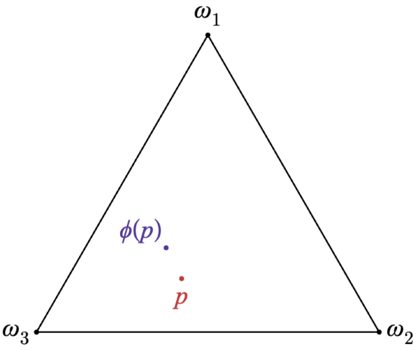
$$\begin{tikzcd}
p \arrow[r, Rightarrow, "\phi"] \arrow[d, "E"'] & q \arrow[d, "E"'] \\
p(\cdot|E) \arrow[r, Rightarrow, "\phi"] & q(\cdot|E)
\end{tikzcd}$$

Main Question: What is the class of ϕ that are coherent?

Illustration

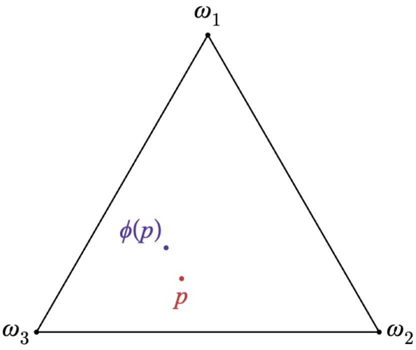


Illustration



Assume $E = \{\omega_1, \omega_2\}$

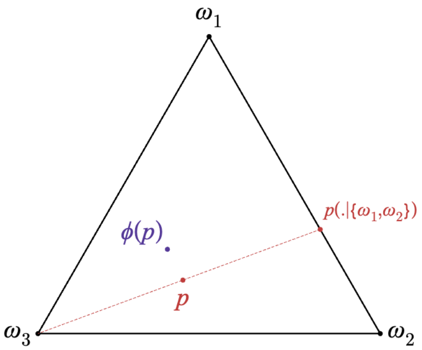
Illustration



Assume $E = \{\omega_1, \omega_2\}$

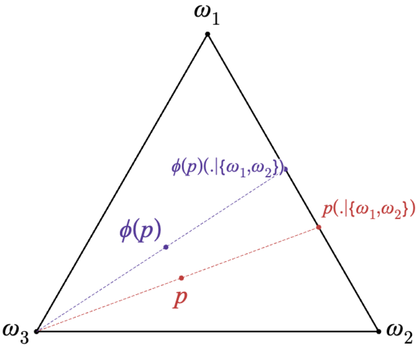
Illustration

Assume $E = \{\omega_1, \omega_2\}$



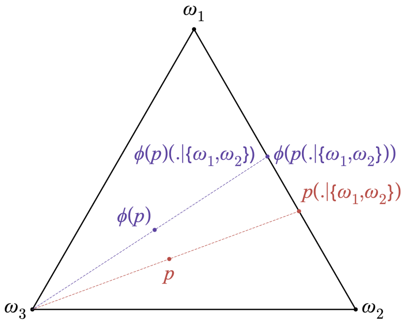
Illustration

Assume $E = \{\omega_1, \omega_2\}$



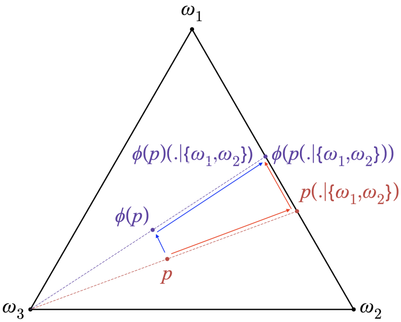
Illustration

Assume $E = \{\omega_1, \omega_2\}$



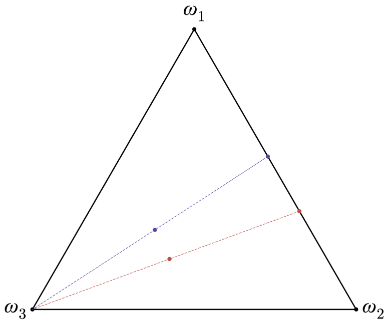
Illustration

Assume $E = \{\omega_1, \omega_2\}$



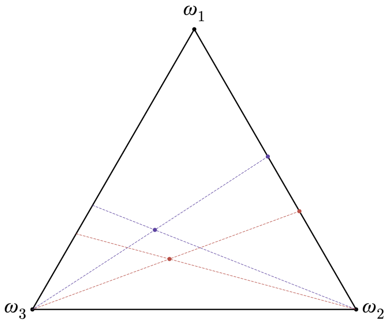
Illustration

Assume $E = \{\omega_1, \omega_2\}$



Illustration

Assume $E = \{\omega_1, \omega_2\}$



Result

Theorem

Let $|\Omega| \geq 3$. A positive and continuous distortion function is coherent iff there is $\alpha > 0$ and for each $\omega \in \Omega$, $\psi(\omega) > 0$ so that

$$\phi(p)(\omega) = \frac{\psi(\omega)p(\omega)^\alpha}{\sum_{\omega'} \psi(\omega')p(\omega')^\alpha}$$

Similar to Aczél and Saaty (1983); Genest, Weerahandi and Zidek (1984)

Theorem

Let $|\Omega| \geq 3$. A positive and continuous distortion function is coherent iff there is $\alpha > 0$ and for each $\omega \in \Omega$, $\psi(\omega) > 0$ so that

$$\phi(p)(\omega) = \frac{\psi(\omega)p(\omega)^\alpha}{\sum_{\omega'} \psi(\omega')p(\omega')^\alpha}$$

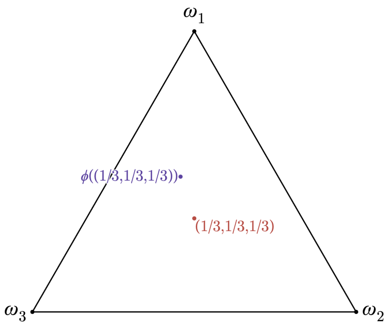
Similar to Aczél and Saaty (1983); Genest, Weerahandi and Zidek (1984)

Identification

How to identify weights?

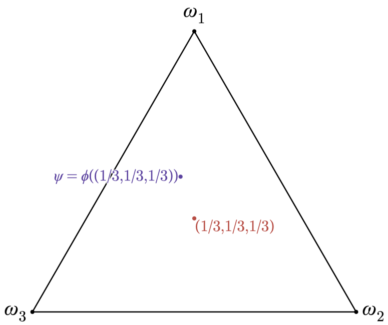
Identification

How to identify weights?



Identification

How to identify weights?

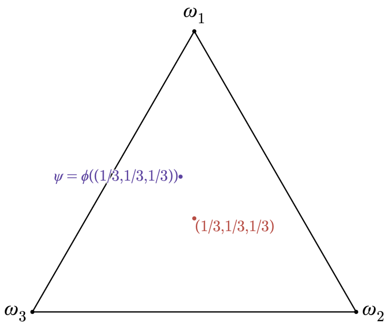


Identification

How to identify α?

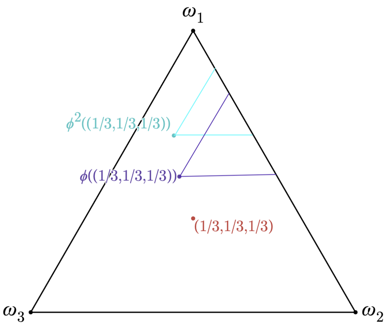
Identification

How to identify α?



Identification

How to identify α?



$$\alpha = \frac{\ln \phi^2(p_u)(\omega_i) - \ln \phi^2(p_u)(\omega_j)}{\ln \phi(p_u)(\omega_i) - \ln \phi(p_u)(\omega_j)} - 1$$

p_u uniform prior

$\alpha = 1$

Corollary

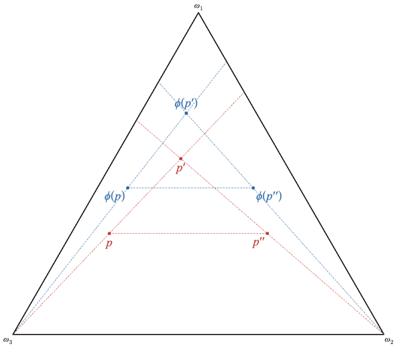
Let ϕ have a power-weighted distorted belief representation. Then $\alpha = 1$ if and only if

$$\frac{\frac{\phi(p)(\omega_i)}{\phi(p)(\omega_j)}}{\frac{\phi(q)(\omega_i)}{\phi(q)(\omega_j)}} = \frac{\frac{p(\omega_i)}{p(\omega_j)}}{\frac{q(\omega_i)}{q(\omega_j)}}$$

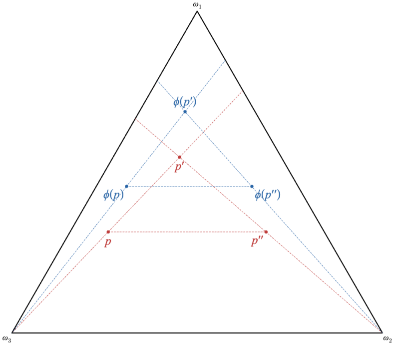
for any p, q and ω_i, ω_j for which $\min\{p(\omega_j), q(\omega_i)\} > 0$.

$\alpha = 1$

- $\Pi = \{\{\omega_1\}, \{\omega_2, \omega_3\}\}$
- This agent does not distinguish between ω_2 and ω_3
- For all $E \in \Pi$, $p(\{\omega_2, \omega_3\}) = p'(\{\omega_2, \omega_3\})$, then for all $E \in \Pi$, $\phi(p)(\{\omega_2, \omega_3\}) = \phi(p')(\{\omega_2, \omega_3\})$



$\alpha = 1$



Corollary

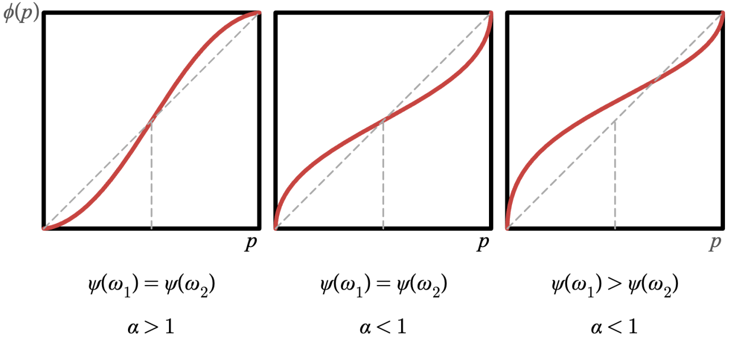
Let Π be a nontrivial partition of Ω. Then ϕ satisfies positivity, continuity, coherency and Π-marginality if and only if $\alpha = 1$ and the map ψ is Π-measurable.

Two States

- Consider only two states
- $p(\omega_1) = p$ and $p(\omega_2) = 1 - p$
- Consider $\phi(p)$

Two States

- Consider only two states
- $p(\omega_1) = p$ and $p(\omega_2) = 1 - p$
- Consider $\phi(p)$



What kind of models generate coherent distortions?

Consider "Λ-KL motivated beliefs problem:"

$$\max_{q \in \Delta(\Omega)} u \cdot p - \frac{1}{K}[\sum_{\omega} q(\omega)[\ln(q(\omega)) - \Lambda \ln(p(\omega))]]$$

$u \cdot q$ captures the anticipated expected utility of the DM, conditional on the distorted beliefs.

$\frac{1}{K}[\sum_{\omega} q(\omega)[\ln(q(\omega)) - \Lambda \ln(p(\omega))]]$ is the cost of distorting beliefs to q, given objective beliefs p.
$K > 0$ is weight on costs, > 0 changes structure of costs, $\Lambda = 1$ is Kullback-Liebler (KL) divergence betweeen p and q.

Taking stock

Proposition

The solution to a Λ-KL motivated beliefs problem with utility u is that $q = \phi(p)$ where ϕ is a weighted power distortion function. Moreover, for a given weighted power distortion function ϕ, we can find a set of utilities u and a K and Λ such that $\alpha = \Lambda$ and $\psi(\omega) = e^{Ku(\omega)}$.

Caplin and Leahy (2019), discuss $\Lambda = 1$; Mayraz (2011) considers distortions where $\alpha = 1$ and $\psi(\omega) = e^{Ku(\omega)}$.

Similar exercise for Blackwell distortions (again consistent with Caplin and Leahy, 2019)

Taking stock

Consider lottery space over outcomes Outcome space X, with elements x_i.

Suppose individuals choose lottery p in order to maximize expected utility given beliefs $\phi(p)$: denote induced preferences as $\succ_{EU,\phi}$

Proposition

A distorted belief ϕ is positive, coherent, and continuous, and $\succ_{EU,\phi}$ is continuous if and only if $\alpha = 1$ and $\succ_{EU,\phi}$ has a weighted utility representation (Chew, 1983).

Choose lotteries using the function

$$\sum_i \frac{\psi(x_i)p(x_i)}{\sum_i \psi(x_i)p(x_i)} u(x_i)$$

where $\alpha = 1$.

Distorting Blackwell Experiments

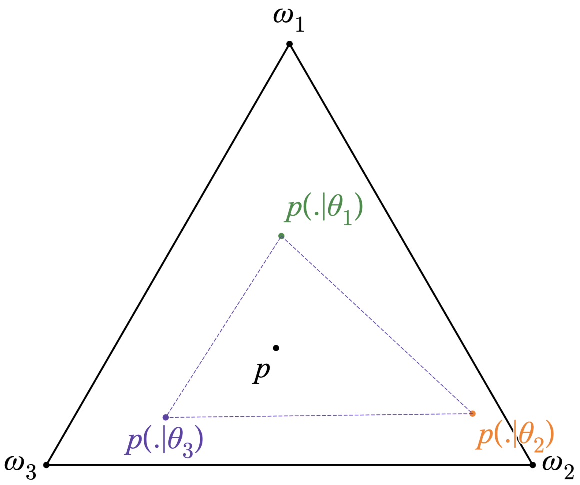
Distorting Blackwell Experiments

- So far, information about states
- Consider Blackwell Experiments

Ω	Signals			
	θ_1	θ_2	θ_3	
ω_1	$\sigma_{\omega_1}(\theta_1)$	$\sigma_{\omega_1}(\theta_2)$	$\sigma_{\omega_1}(\theta_3)$	1
ω_2	$\sigma_{\omega_2}(\theta_1)$	$\sigma_{\omega_2}(\theta_2)$	$\sigma_{\omega_2}(\theta_3)$	1
ω_3	$\sigma_{\omega_3}(\theta_1)$	$\sigma_{\omega_3}(\theta_2)$	$\sigma_{\omega_3}(\theta_3)$	1

- Blackwell experiment $\sigma : \Omega \to \Delta(\Theta)$, prior p

Illustration



Distorting Blackwell Experiments

- Some models assume that priors are not distorted, only the rows of the Blackwell experiment: Benabou and Tirole (2002), Gottlieb (2014), Caplin and Leahy (2019)

Ω	Signals			
	θ_1	θ_2	θ_3	
ω_1	$\sigma_{\omega_1}(\theta_1)$	$\sigma_{\omega_1}(\theta_2)$	$\sigma_{\omega_1}(\theta_3)$	1
ω_2	$\sigma_{\omega_2}(\theta_1)$	$\sigma_{\omega_2}(\theta_2)$	$\sigma_{\omega_2}(\theta_3)$	1
ω_3	$\sigma_{\omega_3}(\theta_1)$	$\sigma_{\omega_3}(\theta_2)$	$\sigma_{\omega_3}(\theta_3)$	1

$$\xrightarrow{\{g_\omega\}_{\omega\in\Omega}}$$

Ω	Signals			
	θ_1	θ_2	θ_3	
ω_1	$g_{\omega_1}(\sigma_{\omega_1})(\theta_1)$	$g_{\omega_1}(\sigma_{\omega_1})(\theta_2)$	$g_{\omega_1}(\sigma_{\omega_1})(\theta_3)$	1
ω_2	$g_{\omega_2}(\sigma_{\omega_1})(\theta_1)$	$g_{\omega_2}(\sigma_{\omega_1})(\theta_2)$	$g_{\omega_2}(\sigma_{\omega_1})(\theta_3)$	1
ω_3	$g_{\omega_3}(\sigma_{\omega_1})(\theta_1)$	$g_{\omega_3}(\sigma_{\omega_1})(\theta_2)$	$g_{\omega_3}(\sigma_{\omega_1})(\theta_3)$	1

- State dependent noise distortion: $g_\omega : \Delta(\Theta) \to \Delta(\Theta)$
- Coherence only for updating on signals: $\{g_\omega\}$ is *Blackwell signal-coherent* if for every $\omega \in \Omega$ and every $S \subseteq \Theta$, $g_\omega(\sigma_\omega(\theta|S)) = g_\omega(\sigma_\omega)(\theta|S)$

A Second (Same?) Result

Corollary

Suppose that $|\Theta| \geq 3$. Given Blackwell distortion functions $\{g_\omega\}$ where each is positive, continuous the following are equivalent:

1. *$\{g_\omega\}$ is Blackwell signal-coherent*
2. *for each state ω there exists a $\psi_\omega : \Theta \to \mathbb{R}_{++}$ and $\alpha_\omega > 0$ such that*

$$g_\omega(\sigma_\omega)(\theta) = \frac{\psi_\omega(\theta)\sigma_\omega(\theta)^{\alpha_\omega}}{\sum_{\theta' \in \Theta} \psi_\omega(\theta')\sigma_\omega(\theta')^{\alpha_\omega}}$$

Gretherian updating

Grether's framework

Grether (1980) style models: separate distortions for prior and signals. Used widely in economics (Benjamin, 2019)

Distinguish between states Ω (with subset E) and signals Θ (with subset S).

Blackwell experiment $\sigma : \Omega \to \Delta(\Theta)$, prior p.

$$\phi(p, \sigma|\theta)(\omega) = \frac{p(\omega)^{\alpha}\sigma(\theta|\omega)^{\beta}}{\sum_{\omega'} p(\omega')^{\alpha}\sigma(\theta|\omega')^{\beta}}.$$

Grether's framework

$$\phi(p,\sigma|\theta)(\omega) = \frac{p(\omega)^{\alpha}\sigma(\theta|\omega)^{\beta}}{\sum_{\omega'} p(\omega')^{\alpha}\sigma(\theta|\omega')^{\beta}}.$$

Two parameter model allows to empirically study weight given to beliefs vs. signals. Captures a large literature in psychology and economics on "probabilistic biases"

E.g., sample-size neglect (Kahneman and Tverksy, 1972; Benjamin et al 2016); base rate neglect (Bodoh Creed, 2010; Benjamin et al, 2019)

Grether's framework

Generalize Grether's formulation:

- f distorts beliefs about states ($f : \Delta(\Omega) \to \Delta(\Omega)$)
- g_ω distorts state-dependent signal distortion ($g_\omega : \Delta(\Theta) \to \mathbb{R}_+^\Theta$)

Goal is to ask what joint distortions satisfy coherency.

Gretherian update

Bayesian Update

$$\mathcal{B}_\sigma(p,\theta)(\omega) = \frac{p(\omega)\sigma(\theta|\omega)}{\sum_{\omega'} p(\omega')\sigma(\theta|\omega')}$$

Gretherian update

$$\mathcal{B}_{g\circ\sigma}(f(p),\theta)(\omega) = \frac{f(p)(\omega)g_\omega(\sigma(\cdot|\omega))(\theta)}{\sum_{\omega'} f(p)(\omega')g_{\omega'}(\sigma(\cdot|\omega')(\theta))}.$$

Gretherian-coherent

Extend definition of coherency and ask when two operations are equivalent: (i) distort prior, and update with distorted Blackwell matrix, (ii) update with true prior and Blackwell matrix then distort posterior

Gretherian-coherent if

$$\mathcal{B}_{g \circ \sigma}(f(p), \theta)(\omega) = f\left(\mathcal{B}_{\sigma}(p, \theta)\right)(\omega)$$

Operation (i) uses two distortions, but (ii) uses only one

Result

Theorem

Suppose that $|\Omega| \geq 3$ *and* $|\Theta| \geq 2$. *Suppose* f *and* g_ω *for all* $\omega \in \Omega$ *are given. Then* f *and* g_ω *are positive and continuous and the pair* (f, g) *is Gretherian-coherent iff there exists* $\psi(\omega) > 0$ *for all* ω, $\gamma(\theta) > 0$ *for all* θ, *and* $\alpha > 0$ *for which*

$$f(p)(\omega) = \frac{\psi(\omega)p(\omega)^\alpha}{\sum_{\omega'} \psi(\omega')p(\omega')^\alpha} \quad \text{and} \quad g_\omega(\sigma(\cdot|\omega))(\theta) = \gamma(\theta)(\sigma(\theta|\omega))^\alpha \quad \text{for all } \omega$$

Grether's rule is coherent only when exponents for states and signals are the same

Why allow distorted probabilities to not sum to 1?

If $\sum_\theta g_\omega(\sigma(\cdot|\omega))(\theta) = 1$ for each ω and σ, then $\alpha = 1$ (see Genest, McConway and Schervish, 1986)

Distorting Joint Probabilities

Distorting Joint Probabilities

- Distinguish between states Ω and signals Θ
- Space of uncertainty is $\Omega \times \Theta$. Green and Stokey, "two representations" (1978) published (2022)
- A new distortion function $\phi : \Delta(\Omega \times \Theta) \to \Delta(\Omega \times \Theta)$

Ω	Signals			Prior
	θ_1	θ_2	θ_3	
ω_1	p_{11}	p_{12}	p_{13}	$p(w_1)$
ω_2	p_{21}	p_{22}	p_{23}	$p(w_2)$
ω_3	p_{31}	p_{32}	p_{33}	$p(w_3)$
	$p(\theta_1)$	$p(\theta_2)$	$p(\theta_3)$	1

$\xrightarrow{\phi}$

Ω	Signals			Prior
	θ_1	θ_2	θ_3	
ω_1	q_{11}	q_{12}	q_{13}	$q(w_1)$
ω_2	q_{21}	q_{22}	q_{23}	$q(w_2)$
ω_3	q_{31}	q_{32}	q_{33}	$q(w_3)$
	$q(\theta_1)$	$q(\theta_2)$	$q(\theta_3)$	1

Marginality

- Marginality: $p|_\Omega = p'|_\Omega$ implies $\phi(p)|_\Omega = \phi(p')|_\Omega$.
- Marginality rules out confounding of state probability distortions with signal probability distortions.
- Only need to know the marginal on states to know the distorted marginal on states

Two Coherence Condition

And Bayesian updating on signals, rather than states:

Weak signal coherence: For all $\theta \in \Theta$, $\phi(p)(\cdot|\{\theta\}) = \phi(p(\cdot|\{\theta\}))$

Strong signal coherence: For all $E \subseteq \Theta$, $\phi(p)(\cdot|E) = \phi(p(\cdot|E))$.

Result

Theorem

Suppose that $|\Omega| \geq 3$ *and* $|\Theta| \geq 2$. *Suppose further that* $\phi : \Delta(\Omega \times \Theta) \to \Delta(\Omega \times \Theta)$ *satisfies positivity, continuity, weak signal coherence, and marginality. Then for each* $\omega \in \Omega$, *there is* $\psi(\omega) > 0$ *for which for all* $\theta \in \Theta$ *we have*

$$\phi(p)(\omega, \theta|\theta) = \frac{\psi(\omega)p(\omega, \theta)}{\sum_{\omega'} \psi(\omega')p(\omega', \theta)}$$

when $p(\theta) > 0$ *and*

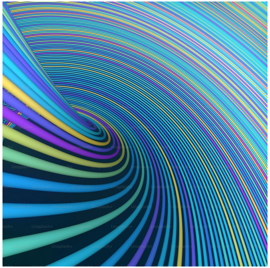
$$\phi(p)(\omega) = \frac{\psi(\omega)p(\omega)}{\sum_{\omega'} \psi(\omega')p(\omega')}.$$

Further, under either (i) $|\Omega| \geq |\Theta|$ or (ii) $|\Theta| \geq 3$ and f is strongly signal coherent, we have

$$\phi(p)(\omega, \theta) = \frac{\psi(\omega)p(\omega, \theta)}{\sum_{\omega', \theta'} \psi(\omega')p(\omega', \theta')}.$$

Conclusion

- Coherency is a normatively appealing property
- Can be tested empirically
- Imposes strong restrictions on functional forms
- Allows for functional forms used in the literature



Iterates converge

Theorem

For any $\alpha > 0$ and ψ, and all $p \in \Delta(\Omega)$, $\phi^*(p) = \lim_n \phi^n(p)$ exists, and has the following form:

1. *If $0 < \alpha < 1$, then for all ω for which $p(\omega) > 0$, $\phi^*(p)(\omega) \propto \psi(\omega)^{\frac{1}{1-\alpha}}$. Otherwise, $\phi^*(p)(\omega) = 0$.*
2. *If $\alpha = 1$, then for all ω for which $\psi(\omega)$ is maximal amongst the set of ω for which $p(\omega) > 0$, $\phi^*(p)(\omega) \propto p(\omega)$, otherwise, $\phi^*(p)(\omega) = 0$.*
3. *If $1 \leq \alpha$, then for all ω for which $p(\omega)^{\alpha-1}\psi(\omega)$ is maximal, we have $\phi^*(p)(\omega) \propto \psi(\omega)^{\frac{1}{1-\alpha}}$. Otherwise, $\phi^*(p)(\omega) = 0$.*