Model Uncertainty

Massimo Marinacci

AXA-Bocconi Chair in Risk
Department of Decision Sciences and IGIER
Università Bocconi

University of Duisburg-Essen
25 May 2016

The problem

- Uncertainty and information are twin notions
- Uncertainty is indeed a form of partial / limited knowledge about the possible realizations of a phenomenon
  - toss a die: what face will come up?
- The first order of business is to frame the problem properly
- First key breakthrough: *probabilities*
- You can assign numbers to alternatives that quantify their relative likelihoods (and manipulate them according to some rules; probability calculus)

Probability: emergence and consolidation

- 16th-17th centuries: probability and its calculus emerged with the works of Cardano, Huygens, Pascal et al.
- 18th-19th centuries: consolidation phase with the works of the Bernoullis, Gauss, Laplace et al.
- Laplace canon (1812) based on equally likely cases / alternatives: the probability of an event equals the number of "favorable" cases over their total number
- Later, the "equally likely" notion came to be viewed as an objective / physical feature (faces of a die, sides of a fair coin) until...

20th century: the Bayesian leap

- 1920s: de Finetti and Ramsey freed probability of physics and rendered “equally likely” a *subjective evaluation*
- In doing so, they could attach probabilities to any event
  - “tomorrow it will rain”
  - “left wing parties will increase their votes in the next elections”
- Such probabilities (often called subjective) quantify the decision maker *degree of belief*
- In this way, all uncertainty can be probabilized: Bayesianism

Road map

- Probabilities: a (brief) historical detour
- **Types of uncertainty: physical vs epistemic**
- Decision problems
  - toolbox
  - Savage setup
  - classical subjective expected utility
- Model uncertainty: ambiguity / robustness models
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model uncertainty resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?

Types of uncertainty

All uncertainty relevant for decision making is ultimately subjective

- To paraphrase Protagoras, in decision problems “DMs are the measure of all things”
- Yet, in applications (especially with data) it is convenient to distinguish between *physical* and *epistemic* uncertainty
  - It traces back to Cournot and Poisson around 1840
- This distinction is a pragmatic *divide et impera* approach (combining objective and subjective views often regarded as dichotomic)
- Caveat, again: relevant for decision problems with data (not for one-of-a-kind decisions / events)

Types of uncertainty: physical

- Examples of physical uncertainty: coin / dice tossing, measurement errors
- Physical uncertainty deals with *variability in data* (e.g., economic time series), because of their inherent randomness, measurement errors, omitted minor explanatory variables
- In applications, physical uncertainty characterizes data generating processes (DGP), i.e., probability models for data

Types of uncertainty: physical

- Physical uncertainty is irreducible
  - take either an urn with 50 white and 50 black balls or a fair coin, the probability of each alternative is 1/2
- There is nothing to learn and information is captured by conditioning
- Here *probability* is a *measure of randomness* / *variability*

Types of uncertainty: epistemic

- Epistemic uncertainty deals with the truth of propositions
  - "tomorrow it will rain"
  - "left wing parties will increase their votes in the next elections"
  - "the parameter that characterizes the DGP has value x"
  - "the composition of the urn is 50 white and 50 black balls"
- It is reducible through learning via Bayes' rule
  - take an urn with only black and white balls, in unknown (and so uncertain) proportion; repeated drawing enables to learn about such uncertainty and, hence, to reduce it
- Here *probability* is a *measure of degree of belief*

Road map

- Probabilities: a (brief) historical detour
- Types of uncertainty: physical vs epistemic
- **Decision problems**
  - **toolbox**
  - **Savage setup**
  - **classical subjective expected utility**
- Model uncertainty: ambiguity / robustness models
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model uncertainty resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?

Decision problems: the toolbox, I

A decision problem consists of

- a space A of actions
- a space C of material (e.g., monetary) consequences
- a space S of environment states
- a consequence function $\rho : A \times S \longrightarrow C$ that details the consequence

$$c = \rho(a, s)$$

of action a when state s obtains

Example (i): natural hazards

Public officials have to decide whether or not to evacuate an area because of a possible earthquake

- A two actions a_0 (no evacuation) and a_1 (evacuation)
- C monetary consequences (damages to infrastructures and human casualties; Mercalli-type scale)
- S possible peak ground accelerations (Richter-type scale)
- $c = \rho(a, s)$ the monetary consequence of action a when state s obtains

Example (ii): monetary policy example

ECB or the FED have to decide some target level of inflation to control the economy unemployment and inflation

- Unemployment u and inflation π outcomes are connected to shocks (w, ε) and the policy a according to

$$
\begin{aligned}
u &= \theta_0 + \theta_{1\pi}\pi + \theta_{1a}a + \theta_2 w \\
\pi &= a + \theta_3 \varepsilon
\end{aligned}
$$

$\theta = (\theta_0, \theta_{1\pi}, \theta_{1a}, \theta_2, \theta_3)$ are five structural coefficients

- $\theta_{1\pi}$ and θ_{1a} are slope responses of unemployment to actual and planned inflation (e.g., Lucas-Sargent $\theta_{1a} = -\theta_{1\pi}$; Samuelson-Solow $\theta_{1a} = 0$)
- θ_2 and θ_3 quantify shock volatilities
- θ_0 is the rate of unemployment that would (systematically) prevail without policy interventions

Example (ii): monetary policy

Here:

A the target levels of inflation

C the pairs $c = (u, \pi)$

S has structural and random components

$$s = (w, \varepsilon, \theta) \in W \times E \times \Theta = S$$

The reduced form is

$$\begin{aligned} u &= \theta_0 + (\theta_{1\pi} + \theta_{1a})\, a + \theta_{1\pi}\theta_3\varepsilon + \theta_2 w \\ \pi &= a + \theta_3\varepsilon \end{aligned}$$

and so ρ has the form

$$\rho(a, w, \varepsilon, \theta) = \begin{bmatrix} \theta_0 \\ 0 \end{bmatrix} + a \begin{bmatrix} \theta_{1\pi} + \theta_{1a} \\ 1 \end{bmatrix} + \begin{bmatrix} \theta_2 & \theta_{1\pi}\theta_3 \\ 0 & \theta_3 \end{bmatrix} \begin{bmatrix} w \\ \varepsilon \end{bmatrix}$$

Decision problems: the toolbox, II

- The quartet (A, S, C, ρ) is a *decision form under uncertainty*
- The decision maker (DM) has a preference $\succsim$ over actions
  - we write $a \succsim b$ if the DM (weakly) prefers action a to action b
- The quintet $(A, S, C, \rho, \succsim)$ is a *decision problem under uncertainty*
- DMs aim to select actions $\hat{a} \in A$ such that $\hat{a} \succsim a$ for all $a \in A$

Consequentialism

What matters about actions is not their label / name but the *consequences* that they determine when the different states obtain

- *Consequentialism*: two actions that are realization equivalent – i.e., that generate the same consequence in every state – are indifferent
- Formally,

$$\rho(a, s) = \rho(b, s) \quad \forall s \in S \Longrightarrow a \sim b$$

  or, equivalently,

$$\rho_a = \rho_b \Longrightarrow a \sim b$$

- Here $\rho_a : S \to C$ is the section of ρ at a given by $\rho_a(s) = \rho(a, s)$

Savage setup

- Identify actions that are realization equivalent
- Formally, in place of actions we consider the maps $\mathbf{a} : S \to C$ that they induce as follows:

$$\mathbf{a}(s) = \rho_a(s) \qquad \forall s \in S$$

- These maps are called *acts* — they are state contingent consequences
- $\mathbf{A}$ denotes the collection of all the acts
- We can directly consider the preference $\succsim$ on $\mathbf{A}$ by setting $\mathbf{a} \succsim \mathbf{b}$ if and only if $a \succsim b$
- The quartet $(\mathbf{A}, S, C, \succsim)$ represents the decision problem a la Savage (1954), a reduced form of problem $(A, S, C, \rho, \succsim)$

Physical uncertainty: probability models

- Because of their ex-ante structural information, DMs know that states are generated by a probability model m that belongs to a given subset M of $\Delta(S)$
- Each m describes a possible *DGP*, and so it represents *physical uncertainty* (risk)
- DMs thus posit a model space M in addition to the state space S, a central tenet of classical statistics a la Neyman-Pearson-Wald
- When the model space is based on experts' advice, its nonsingleton nature may reflect different advice

Models: a toy example

Consider an urn with 90 Red, or Green, or Yellow balls

- DMs bet on the color of a ball drawn from the urn
- State space is $S = \{R, G, Y\}$
- Without any further information, $M = \Delta(\{R, G, Y\})$
- If DMs are told that 30 balls are red, then

$$M = \{m \in \Delta(\{R, G, Y\}) : m(R) = 1/3\}$$

Models and experts: probability of heart attack

Two DMs: John and Lisa are 70 years old

- smoke
- no blood pressure problem
- total cholesterol level 310 mg/dL
- HDL-C (good cholesterol) 45 mg/dL
- systolic blood pressure 130

What's the probability of a heart attack in the next 10 years?

Models and experts: probability of heart attack

Based on their data and medical models, experts say

Experts	John's m	Lisa's m
Mayo Clinic	25%	11%
National Cholesterol Education Program	27%	21%
American Heart Association	25%	11%
Medical College of Wisconsin	53%	27%
University of Maryland Heart Center	50%	27%

Table from Gilboa and Marinacci (2013)

Models: adding a consistency condition

- Cerreia, Maccheroni, Marinacci, Montrucchio (*PNAS* 2013) take the "physical" information M as a primitive and thus enrich the standard Savage framework
- DMs know that the true model m that generates observations belongs to the posited collection M
- In terms of preferences: betting behavior must be *consistent* with datum M. Formally,

$$m(F) \geq m(E) \quad \forall m \in M \Longrightarrow cFc' \succsim cEc'$$

  where cFc' and cEc' are bets on events F and E, with $c \succ c'$
- The quintet $(\mathbf{A}, S, C, M, \succsim)$ forms a Savage *classical* decision problem
- Remark: we abstract away from model misspecification issues

Classical subjective EU

We show that a preference $\succsim$ that satisfies Savage's axioms and the consistency condition is represented by the criterion

$$V(\mathbf{a}) = \sum_{m \in M} \left(\sum_{s \in S} u(\mathbf{a}(s))\, m(s) \right) \mu(m) \tag{1}$$

That is, acts **a** and **b** are ranked as follows:

$$\mathbf{a} \succsim \mathbf{b} \iff V(\mathbf{a}) \geq V(\mathbf{b})$$

Here

- u is a von Neumann-Morgenstern utility function that captures *risk attitudes* (i.e., attitudes toward physical uncertainty)
- μ is a *subjective prior probability* that quantifies the epistemic uncertainty about models; its support is included in M
- If M is based on the advice of different experts, the prior may reflect the *different confidence* that DMs have in each of them

Classical subjective EU

We call this representation *Classical Subjective Expected Utility* because of the classical statistics tenet on which it relies

- If we set

$$U(\mathbf{a}, m) = \sum_{s \in S} u(\mathbf{a}(s))\, m(s)$$

  we can write the criterion as

$$V(\mathbf{a}) = \sum_{m \in M} U(\mathbf{a}, m)\, \mu(m)$$

- In words, the criterion considers the expected utility $U(\mathbf{a}, m)$ of each possible DGP m, and averages them out according to the prior μ

Classical subjective EU

- Each prior μ induces a *predictive probability* $\bar{\mu} \in \Delta(S)$ through reduction

$$\bar{\mu}(E) = \sum_{m \in M} m(E)\,\mu(m)$$

  In turn, the predictive probability enables to rewrite the representation as

$$V(\mathbf{a}) = U(\mathbf{a}, \bar{\mu}) = \sum_{s \in S} u(\mathbf{a}(s))\,\bar{\mu}(s)$$

- This reduced form of V is the original Savage subjective EU representation

Classical subjective EU: some special cases

- If the support of μ is a singleton $\{m\}$, DMs subjectively (and so possibly wrongly) believe that m is the true model
  The criterion thus reduces to a Savage EU criterion $U(\mathbf{a}, m)$
- If M is a singleton $\{m\}$, DMs know that m is the true model (a rational expectations tenet)
  - (i) There is no epistemic uncertainty, but only physical uncertainty (quantified by m)
  - (ii) The criterion again reduces to the EU representation $U(\mathbf{a}, m)$, but now interpreted as a *von Neumann-Morgenstern criterion*

Classical subjective EU: some special cases

- Classical subjective EU thus encompasses both the Savage and the von Neumann-Morgenstern representations
- If $M \subseteq \{\delta_s : s \in S\}$, there is no physical uncertainty, but only epistemic uncertainty (quantified by μ). By identifying s with δ_s, wlog we can write $\mu(s)$ and so the criterion takes the form

$$V(\mathbf{a}) = \sum v(\mathbf{a}(s))\,\mu(s)$$

where it is v that matters

Classical subjective EU: monetary policy example

Back to the monetary example

$$u = \theta_0 + \theta_{1\pi}\pi + \theta_{1a}a + \theta_2 w$$
$$\pi = a + \theta_3\varepsilon$$

- Distribution q of shocks (w, ε)
- θ is deterministic, fixed
- Each model m corresponds to a shock distribution q and to a possible model economy θ

Classical subjective EU: monetary policy example

Suppose:
(i) shocks distribution q is known
(ii) model economy θ is unknown

- Each model m is thus uniquely parametrized by θ, and so belief μ is directly on θ
- The monetary policy problem is then

$$\max_{\mathbf{a}\in\mathbf{A}} V(\mathbf{a}) = \max_{\mathbf{a}\in\mathbf{A}} \sum_{\theta\in\Theta} \left(\sum_{(w,\varepsilon)\in W\times E} u(\mathbf{a}(w,\varepsilon,\theta))\, q(w,\varepsilon) \right) \mu(\theta)$$

Classical subjective EU: portfolio

- Frictionless financial market with n assets
- Each with uncertain gross return r_i after one period
- $a = (a_1, ..., a_n) \in \Delta_{n-1}$ is vector of portfolio weights
- If initial wealth is 1,

$$\rho(a, s) = a \cdot s = \sum_{i=1}^{n} a_i r_i$$

  is the end-of-period wealth when $s = (r_1, ..., r_n)$ obtains
- The portfolio decision problem is

$$\max_{a \in A} V(a) = \max_{a \in \Delta_{n-1}} \sum_{m \in M} \left(\sum_{s \in S} u(a \cdot s) \, m(s) \right) \mu(m)$$

Classical subjective EU: portfolio

- Two assets: a risk free with return r_f and a risky one with uncertain return r
- State space is the set R of all possible returns of the risky asset
- If $a \in [0, 1]$ is the fraction of wealth invested in the risky asset, the portfolio problem becomes

$$\max_{a \in [0,1]} \sum_{m \in M} \left(\sum_{r \in R} u\left(\left(1 - a\right) r_f + ar\right) m\left(r\right) \right) \mu\left(m\right)$$

Classical subjective EU: portfolio

- Suppose $r - r_f = \beta x + (1 - \beta)\,\varepsilon$, with $\beta \in [0, 1]$
- x is a predictor for the excess return and ε is a shock with distribution q
- The higher β, the more predictable the excess return
- $s = (\varepsilon, \beta)$, where ε and β are its random and structural components
- Each model corresponds to a shock distribution q and to a predictability structure β

Classical subjective EU: portfolio

- If q is known, the only unknown is β:

$$\max_{a\in[0,1]} \int_{[0,1]} \left(\int_E u\left(r_f + a\left(\beta x + (1-\beta)\,\varepsilon\right)\right) dq\left(\varepsilon\right) \right) d\mu\left(\beta\right)$$

Here only predictability uncertainty

- If q and ε are both unknown:

$$\max_{a\in[0,1]} \int_{\Delta(E)\times[0,1]} \left(\int_E u\left(r_f + a\left(\beta x + (1-\beta)\,\varepsilon\right)\right) dq\left(\varepsilon\right) \right) d\mu\left(q,\beta\right)$$

Now both parametric and predictability uncertainty (Barberis, 2000)

Road map

- Probabilities: a (brief) historical detour
- Types of uncertainty: physical vs epistemic
- Decision problems
  - toolbox
  - Savage setup
  - classical subjective expected utility
- **Model uncertainty: ambiguity / robustness models**
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model uncertainty resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?

Ambiguity / Robustness: the problem

- Physical and epistemic uncertainties need to be treated differently
- The standard expected utility model does not
- Since the 1990s, a strand of economic literature has been studying *ambiguity* / *Knightian uncertainty* / *robustness*
- We consider two approaches
  - non-Bayesian (Gilboa and Schmeidler, *J. Math. Econ.* 1989; Schmeidler, *Econometrica* 1989)
  - Bayesian (Klibanoff, Marinacci, Mukerji, *Econometrica* 2005)
- Both approaches broaden the scope of traditional EU analysis
- Normative focus (no behavioral biases or "mistakes"; see Gilboa and Marinacci, 2013)

Ambiguity / Robustness: the problem

- Intuition: betting on coins is greatly affected by whether or not coins are well tested
- Models correspond to possible biases of the coin
- By symmetry (uniform reduction), heads and tails are judged to be equally likely when betting on an untested coin, never flipped before
- The same probabilistic judgement holds for a well tested coin, flipped a number of times with an approximately equal proportion of heads to tails
- The evidence behind such judgements, and so the confidence in them, is dramatically different: ceteris paribus, DMs may well prefer to bet on tested (phys. unc.) rather than on untested coins (phys. & epist. unc.)
- Experimental evidence: Ellsberg paradox

Ambiguity / Robustness: relevance

- A more robust rational behavior toward uncertainty emerges
- A more accurate / realistic account of how uncertainty affects valuation (e.g., uncertainty premia in market prices)
- Better understanding of exchange mechanics
  - a dark side of uncertainty: no-trade or small-trade results because of cumulative effects of physical and epistemic uncertainty; See the recent financial crisis
- Better calibration and quantitative exercises
  - applications in Finance, Macroeconomics, and Environmental Economics
- Better modelling of decision / policy making
  - applications in Risk Management; e.g., the otherwise elusive precautionary principle may fit within this framework

Ambiguity / Robustness: relevance

- Caveat: risk and model uncertainty can work in the same direction (magnification effects), as well as in different directions
- Magnification effects: large "uncertainty prices" with reasonable degrees of risk aversion
- Combination of sophisticated formal reasoning and empirical relevance

Ambiguity / Robustness: a Bayesian approach

- A first distinction: DMs do not have attitudes toward uncertainty per se, but rather toward physical uncertainty and toward epistemic uncertainty
- Such attitudes may differ: typically DMs are more averse to epistemic than to physical uncertainty
- Inferred from lab experiments, but in the end it is an empirical question

Bayesian approach: a tacit assumption

Suppose acts are monetary

- Classical subjective EU representation can be written as

$$
\begin{aligned}
V(\mathbf{a}) &= \sum_{m\in M} U(\mathbf{a}, m)\,\mu(m) \\
&= \sum_{m\in M} \left(u \circ u^{-1}\right)(U(\mathbf{a}, m))\,\mu(m) \\
&= \sum_{m\in M} u(c(\mathbf{a}, m))\,\mu(m)
\end{aligned}
$$

  where $c(\mathbf{a}, m)$ is the certainty equivalent

$$c(\mathbf{a}, m) = u^{-1}(U(\mathbf{a}, m))$$

  of act **a** under model m

- Recall that $U(\mathbf{a}, m) = \sum_{s\in S} u(\mathbf{a}(s))\,m(s)$

Bayesian approach: a tacit assumption

- The profile

$$\{c(\mathbf{a}, m) : m \in \operatorname{supp} \mu\}$$

is the scope of the model uncertainty that is relevant for the decision

- In particular, DMs use the decision criterion

$$V(\mathbf{a}) = \sum_{m \in M} u(c(\mathbf{a}, m)) \mu(m)$$

to address model uncertainty, while

$$U(\mathbf{a}, m) = \sum_{s \in S} u(\mathbf{a}(s)) m(s)$$

is how DMs address the physical uncertainty that each model m features

- Identical attitudes toward physical and epistemic uncertainties, both modeled by the same function u

Bayesian approach: representation

- The smooth ambiguity model generalizes the representation by distinguishing such attitudes
- Acts are ranked according to the smooth (ambiguity) criterion

$$\begin{aligned} V(\mathbf{a}) &= \sum_{m \in M} \left(v \circ u^{-1}\right)\left(U(\mathbf{a}, m)\right)\mu(m) \\ &= \sum_{m \in M} v\left(c(\mathbf{a}, m)\right)\mu(m) \end{aligned}$$

- The function $v : C \longrightarrow \mathbb{R}$ represents attitudes toward model uncertainty
- A negative attitude toward model uncertainty is modelled by a concave v, interpreted as aversion to (mean preserving) spreads in certainty equivalents $c(\mathbf{a}, m)$
- Ambiguity aversion amounts to a higher degree of aversion toward epistemic than toward physical uncertainty, i.e., a v more concave than u

Bayesian approach: representation

- Setting $\phi = v \circ u^{-1}$, the smooth criterion can be written as

$$V(\mathbf{a}) = \sum_{m \in M} \phi(U(\mathbf{a}, m))\,\mu(m)$$

- This formulation holds for any kind of acts (not just monetary)
- Ambiguity aversion corresponds to the concavity of ϕ
- If $\phi(x) = -e^{-\lambda x}$, it is a Bayesian version of the multiplier preferences of Hansen and Sargent (*AER* 2001, book 2008)
- Sources of uncertainty now matter (no longer "uncertainty is reduced to risk")

Bayesian approach: example

- Call I the tested coin and II the untested one
- Actions a_I and a_{II} are, respectively, bets of one euro on coin I and on coin II
- $S = \{H, T\} \times \{H, T\} = \{HH, HT, TH, TT\}$
- The next table summarizes the decision problem

	HH	HT	TH	TT
$\mathbf{a}_I$	1	1	0	0
$\mathbf{a}_{II}$	1	0	1	0

Bayesian approach: example

- Given the available information, it is natural to set

$$M = \left\{ m \in \Delta(S) : m(HH \cup HT) = m(TH \cup TT) = \frac{1}{2} \right\}$$

- M consists of all models that give probability $1/2$ to either outcome for the tested coin; no specific probability is, instead, assigned to the outcome of the untested coin

Bayesian approach: example

- Normalize $u(1) = 1$ and $u(0) = 0$, so that

$$V(\mathbf{a}_I) = \sum_{m \in M} \phi(m(HH \cup HT))\, d\mu(m) = \phi\left(\frac{1}{2}\right)$$

and

$$V(\mathbf{a}_{II}) = \sum_{m \in M} \phi(m(HH \cup TH))\, d\mu(m)$$

- If μ is uniform, $V(\mathbf{a}_{II}) = \int_0^1 \phi(x)\, dx$. If ϕ is strictly concave, by the Jensen inequality we then have

$$V(\mathbf{a}_{II}) = \int_0^1 \phi(x)\, dx < \phi\left(\int_0^1 x dx\right) = \phi\left(\frac{1}{2}\right) = V(\mathbf{a}_I)$$

Bayesian approach: extreme attitudes and maxmin

- Under extreme ambiguity aversion (e.g., as $\lambda \uparrow \infty$ when $\phi(x) = -e^{-\lambda x}$), the smooth ambiguity criterion in the limit reduces to the maxmin criterion

$$V(\mathbf{a}) = \min_{m \in \operatorname{supp} \mu} \sum_{s \in S} u(\mathbf{a}(s)) \, m(s)$$

- Pessimistic criterion: DMs maxminimize over all possible probability models in the support of μ
- The prior μ just selects which models in M are relevant
- Waldean version of Gilboa and Schmeidler (*J. Math. Econ.* 1989) seminal maxmin decision model

Bayesian approach: extreme attitudes and maxmin

- If $\mathrm{supp}\,\mu = M$, the prior is actually irrelevant and we get back to the Wald (1950) maxmin criterion

$$V(\mathbf{a}) = \min_{m \in M} \sum_{s \in S} u(\mathbf{a}(s))\, m(s)$$

- When M consists of all possible models, it reduces to the statewise maxmin criterion

$$V(\mathbf{a}) = \min_{s \in S} u(\mathbf{a}(s))$$

  A very pessimistic (paranoid?) criterion: probabilities, of any sort, do not play any role (Arrow-Hurwicz decision under ignorance)

- Precautionary principle

Bayesian approach: extreme attitudes and no trade

In a frictionless market a primary asset y that pays $y(s)$ if state s obtains, is traded

- Its market price is p
- Investors may trade x units of the asset (buy if $x > 0$, sell if $x < 0$, no trade if $x = 0$)
- State contingent payoff is $\mathbf{x}(s) = y(s)x - px$
- Trade occurs only if $V(\mathbf{x}) \geq V(\mathbf{0}) = 0$

Bayesian approach: extreme attitudes and no trade

- Dow and Werlang (*Econometrica* 1992): under maxmin behavior, there is no trade on asset y whenever

$$\min_{m \in \operatorname{supp} \mu} \mathrm{E}_m(y) < p < \max_{m \in \operatorname{supp} \mu} \mathrm{E}_m(y) \tag{2}$$

- High ambiguity aversion may freeze markets
- Inequality (2) requires $\operatorname{supp} \mu$ to be nonsingleton: the result requires ambiguity
- More generally: a lower trade volume on asset y corresponds to a higher ambiguity aversion (e.g., higher λ when $\phi(x) = -e^{-\lambda x}$) if (2) holds
- Bottom line: it reinforces the idea that uncertainty can be an impediment to trade

Bayesian approach: quadratic approximation

- The smooth ambiguity criterion admits a simple quadratic approximation that leads to a generalization of the classic mean-variance model (Maccheroni, Marinacci, Ruffino, *Econometrica* 2013)
- The robust mean-variance rule ranks acts **a** by

$$\mathrm{E}_{\bar{\mu}}\left(\mathbf{a}\right) - \frac{\lambda}{2}\sigma^2_{\bar{\mu}}\left(\mathbf{a}\right) - \frac{\theta}{2}\sigma^2_{\mu}\left(\mathrm{E}\left(\mathbf{a}\right)\right)$$

  where λ and θ are positive coefficients
- Here $\mathrm{E}(\mathbf{a}) : M \to \mathbb{R}$ is the random variable

$$m \mapsto \mathrm{E}_m(\mathbf{a}) = \sum_{s \in S} \mathbf{a}\left(s\right) m\left(s\right)$$

  that associates the EV of act **a** under each possible model m
- $\sigma^2_{\mu}\left(\mathrm{E}(\mathbf{a})\right)$ is its variance

Bayesian approach: quadratic approximation

- The robust mean-variance rule

$$\mathrm{E}_{\bar{\mu}}\left(\mathbf{a}\right)-\frac{\lambda}{2}\sigma_{\bar{\mu}}^{2}\left(\mathbf{a}\right)-\frac{\theta}{2}\sigma_{\mu}^{2}\left(\mathrm{E}\left(\mathbf{a}\right)\right)$$

  is determined by the three parameters λ, θ, and μ. When $\theta = 0$ we return to the usual mean-variance rule
- The taste parameters λ and θ model DMs' attitudes toward physical and epistemic uncertainty, resp.
- Higher values of these parameters correspond to stronger negative attitudes

Bayesian approach: quadratic approximation

- The information parameter μ determines the variances $\sigma_{\bar{\mu}}^2(\mathbf{a})$ and $\sigma_{\mu}^2(\mathrm{E}(\mathbf{a}))$ that measure the physical and epistemic uncertainty that DMs perceive in the evaluation of act **a**
- Higher values of these variances correspond to a DM's poorer information regarding such uncertainties
- As usual, the risk premium is

$$\frac{\lambda}{2}\sigma_{\bar{\mu}}^2(\mathbf{a})$$

- Novelty: the ambiguity premium is

$$\frac{\theta}{2}\sigma_{\mu}^2(\mathrm{E}(\mathbf{b}))$$

Ambiguity / Robustness: a non Bayesian approach

- Need to relax the requirement that a single number quantifies beliefs: the multiple (prior) probabilities model
- DMs may not have enough information to quantify their beliefs through a single probability, but need a set of them
- Expected utility is computed with respect to each probability and DMs act according to the minimum among such expected utilities

Non Bayesian approach: representation

- Epistemic uncertainty quantified by a set C of priors
- DMs use the criterion

$$\begin{aligned} V(\mathbf{a}) &= \min_{\mu \in C} \sum_{m \in M} \left(\sum_{s \in S} u(\mathbf{a}(s))\, m(s) \right) \mu(m) \\ &= \min_{\mu \in C} \sum_{s \in S} u(\mathbf{a}(s))\, \bar{\mu}(s) \qquad (3) \end{aligned}$$

- DMs consider the least among all the EU determined by each prior in C
- The predictive form (3) is the original version axiomatized by Gilboa and Schmeidler (*J. Math. Econ.* 1989)

Non Bayesian approach: comments

- This criterion is less extreme than it may appear at a first glance
- The set C incorporates
  - the attitude toward ambiguity, a taste component
  - its perception, an information component
- A smaller set C may reflect both better information – i.e., a lower perception of ambiguity – and / or a less averse uncertainty attitude
- In sum, the size of C does not reflect just information, but taste as well

Non Bayesian approach: comments

- With singletons $\mathsf{C} = \{\mu\}$ we return to the classical subjective EU criterion
- When C consists of all possible priors on M, we return to the Wald maxmin criterion

$$\min_{m \in M} \sum_{s \in S} u(\mathbf{a})\, m(s)$$

- No trade results (kinks)

Non Bayesian approach: variational model

- In the maxmin model, a prior μ is either "in" or "out" of the set C
- Maccheroni, Marinacci, Rustichini (*Econometrica* 2006): general variational representation

$$V(\mathbf{a}) = \inf_{\mu \in \Delta(M)} \left(\sum_{m \in M} \left(\sum_{s \in S} u(\mathbf{a}(s))\, m(s) \right) \mu(m) + c(\mu) \right)$$

where $c(\mu)$ is a convex function that weights each prior μ

- If c is the dichotomic function given by

$$\delta_{\mathrm{C}}(\mu) = \begin{cases} 0 & \text{if } \mu \in \mathrm{C} \\ +\infty & \text{else} \end{cases}$$

we get back to the maxmin model with set of priors C

Non Bayesian approach: multiplier model

- If c is given by the relative entropy $R(\mu||\nu)$, where ν is a reference prior, we get the multiplier model

$$V(\mathbf{a}) = \inf_{\mu \in \Delta(M)} \left(\sum_{m \in M} \left(\sum_{s \in S} u(\mathbf{a}(s))\, m(s) \right) \mu(m) + \alpha R(\mu||\nu) \right)$$

popularized by Hansen and Sargent in their studies on robustness in Macroeconomics

- Also the mean-variance model is variational, with c given by a Gini index

Road map

- Probabilities: a (brief) historical detour
- Types of uncertainty: physical vs epistemic
- Decision problems
  - toolbox
  - Savage setup
  - classical subjective expected utility
- Model uncertainty: ambiguity / robustness models
- **Issues**
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model uncertainty resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?

Optima: more prudent?

Does ambiguity /robustness make optimal actions more prudent?

- It is a robustness requirement on optima
- But this does not necessarily mean “more prudent”
- Folk wisdom: sometimes “the best defense is a good offense”

Optima: more prudent?

- Consider the optimum problem

$$\max_{\mathbf{a}\in\mathbf{A}} \sum_{m\in M} \phi\left(U\left(\mathbf{a}, m\right)\right)\mu\left(m\right)$$

  where ϕ and u are twice differentiable, with $\phi', u' > 0$ and $\phi'', u'' < 0$
- Recall that $U\left(\mathbf{a}, m\right) = \sum_{s\in S} u\left(\mathbf{a}\left(s\right)\right) m\left(s\right)$

Optima: more prudent?

- There is a "tilted" prior $\hat{\mu}$, equivalent to μ, such that problems

$$\max_{\mathbf{a}\in\mathbf{A}} \sum_{m\in M} \phi\left(U\left(\mathbf{a}, m\right)\right)\mu\left(m\right) \quad \text{and} \quad \max_{\mathbf{a}\in\mathbf{A}} \sum_{m\in M} U\left(\mathbf{a}, m\right)\hat{\mu}\left(m\right)$$

  have the same solution $\hat{\mathbf{a}}$
- Here

$$\hat{\mu}\left(m\right) = \frac{\phi'\left(U\left(\hat{\mathbf{a}}, m\right)\right)}{\sum_{m\in M}\phi'\left(U\left(\hat{\mathbf{a}}, m\right)\right)\mu\left(m\right)}\mu\left(m\right)$$

Optima: more prudent?

- ϕ' is decreasing
- $\hat{\mu}$ thus alters μ by shifting weight to models m with a lower $U(\hat{\mathbf{a}}, m)$
- $\hat{\mathbf{a}}$ solves EU problem $\max_{\mathbf{a}\in\mathbf{A}} \sum_{m\in M} U(\hat{\mathbf{a}}, m)\,\hat{\mu}(m)$ despite $\hat{\mu}$ handicaps $\hat{\mathbf{a}}$ by overweighting its cons over its pros
- $\hat{\mathbf{a}}$ is a robust solution when compared to the solution of the ordinary EU problem $\max_{\mathbf{a}\in\mathbf{A}} \sum_{m\in M} U(\hat{\mathbf{a}}, m)\,\mu(m)$
- In sum, ambiguity aversion can be interpreted as a desire for robustness on optima

Optima: more prudent?

Back to the monetary policy example. Define $(\mathbf{u}, \boldsymbol{\pi})$ by

$$\begin{aligned} \mathbf{u}(a, w, \varepsilon, \theta) &= \theta_0 + (\theta_{1\pi} + \theta_{1a})\, a + \theta_{1\pi}\theta_3\varepsilon + \theta_2 w \\ \boldsymbol{\pi}(a, w, \varepsilon, \theta) &= a + \theta_3\varepsilon \end{aligned}$$

$$\rho(a, w, \varepsilon, \theta) = (\mathbf{u}(a, w, \varepsilon, \theta), \boldsymbol{\pi}(a, w, \varepsilon, \theta))$$

- Assumptions:
  - shocks are uncorrelated with zero mean and unit variance wrt the known distribution q
  - the policy multiplier is negative, i.e., $\theta_{1\pi} + \theta_{1a} \leq 0$
  - coefficients $\theta_{1\pi}$, θ_2 and θ_3 are known

Optima: more prudent?

- Linear quadratic policy framework
- Objective function $V(a)$ is

$$\sum_{\theta} \phi \left(- \sum_{(w,\varepsilon)} \left(\mathbf{u}^2 (a, w, \varepsilon, \theta) + \boldsymbol{\pi}^2 (a, w, \varepsilon, \theta) \right) q(w, \varepsilon) \right) \mu(\theta)$$

where $\theta = (\theta_0, \theta_{1a}) \in \Theta$

Optima: more prudent?

- If true model economy θ^* is known, the (objectively) optimal policy is

$$a^o = B(\theta^*) = -\frac{\theta_0^* \left(\theta_{1\pi}^* + \theta_{1a}^*\right)}{1 + \left(\theta_{1\pi}^* + \theta_{1a}^*\right)^2}$$

  where $B(\cdot)$ is the best reply function
- If not, the optimal policy is

$$\hat{a} = B(\hat{\mu}) = -\frac{\mathrm{E}_{\hat{\mu}}(\theta_0)\left(\theta_{1\pi}^* + \mathrm{E}_{\hat{\mu}}(\theta_{1a})\right) + Cov_{\hat{\mu}}(\theta_0, \theta_{1a})}{1 + \left(\theta_{1\pi}^* + \mathrm{E}_{\hat{\mu}}(\theta_{1a})\right)^2 + V_{\hat{\mu}}(\theta_{1a})}$$

  where $B(\cdot)$ is the EU best reply function wrt the tilted prior $\hat{\mu}$
- Policy $B(\hat{\mu})$ is the robust version of policy $B(\mu)$ that takes into account ambiguity aversion

Optima: more prudent?

Suppose the monetary authority is dogmatic on θ_{1a}, i.e., there is a value $\bar{\theta}_{1a}$ such that $\mu\left(\bar{\theta}_{1a}\right) = 1$. For example:

- $\bar{\theta}_{1a} = 0$ when dogmatic on a Samuelson-Solow economy
- $\bar{\theta}_{1a} = -\theta^*_{1\pi}$ when dogmatic on a Lucas-Sargent economy

Since $\hat{\mu}$ and μ are equivalent, also $\hat{\mu}\left(\bar{\theta}_{1a}\right) = 1$. Hence,

$$B\left(\hat{\mu}\right) \leq B\left(\mu\right) \Longleftrightarrow \mathrm{E}_{\hat{\mu}}\left(\theta_0\right) \leq \mathrm{E}_{\mu}\left(\theta_0\right)$$

Optima: more prudent?

- The robust policy is more prudent as long as the tilted expected value of θ_0 is lower
- When Lucas-Sargent dogmatic, $B(\mu) = B(\hat{\mu}) = 0$ and so the zero-target-inflation policy is optimal, regardless of any uncertainty
- On "tilted" prudence and ambiguity / robustness
  - Taboga (*FinRL* 2005), Hansen (*AER* 2007), Hansen and Sargent (book 2008), Gollier (*RES* 2011), Collard, Mukerji, Sheppard, Tallon (2012)

Diversification

- Public officials have to decide which treatment $t \in T$ should be administered
- Homogeneous population (same covariate)
- Policy is a distribution $a \in \Delta(T)$, where $a(t)$ is the fraction of the population under treatment t
- $c(t, s)$ is the outcome of treatment t when state s obtains
- $\rho(a, s) = \sum_{t \in T} c(t, s)\, a(t)$ is the average outcome

Diversification

- Policy problem is

$$\max_{a\in\Delta(T)} V(a) = \max_{a\in\Delta(T)} \sum_{m\in M} \phi\left(\sum_{s\in S} \rho(a,s)\, m(s)\right) \mu(m)$$
$$= \max_{a\in\Delta(T)} \sum_{m\in M} \phi\left(\sum_{t\in T} \bar{c}_m(t)\, a(t)\right) \mu(m)$$

where $\bar{c}_m(t) = \sum_{s\in S} c(t,s)\, m(s)$ is the expected outcome of treatment t under model m

Diversification

- Binary case $T = \{t_0, t_1\}$
- Policy $a \in [0,1]$ is the fraction of the population under treatment t_1
- Fractional treatment (and so diversification) if $a \in (0,1)$
- Policy problem is

$$\max_{a\in[0,1]} V(a) = \max_{a\in[0,1]} \mathrm{E}_{\mu}\phi\left((1-a)\,\bar{c}_m(t_0) + a\bar{c}_m(t_1)\right)$$

- If ϕ is linear, $a = 0$ or $a = 1$ unless $\bar{c}_{\bar{\mu}}(t_0) = \bar{c}_{\bar{\mu}}(t_1)$, in which case all $a \in [0,1]$ are optimal
- Under subjective EU fractional treatment is not optimal
- To justify fractional treatment, in a series of papers Charles Manski considered maxmin regret

Diversification

- Suppose ϕ is quadratic
- Set $d_m = \bar{c}_m(t_0) - \bar{c}_m(t_1)$. The optimal policy is

$$\hat{a} = \frac{\mathrm{E}_\mu \mathrm{E}_m \bar{c}_m(t_0)\, d_m}{\mathrm{E}_\mu d_m^2}$$

- $\hat{a} \in (0,1)$ if and only if $|V(0) - V(1)| < \mathrm{E}_\mu d_m^2$
- Fractional treatment may thus emerge when ϕ nonlinear
- In fact, if ϕ concave we have the following convexity property:

$$a \sim b \Longrightarrow \alpha a + (1-\alpha) b \succsim b \qquad \forall \alpha \in [0,1]$$

- First noted by David Schmeidler, who called this property uncertainty aversion

Valuation: static asset pricing

- Two-period economy, with a single consumption good
- Agents decide today c_0 and tomorrow c_1, which is contingent on the state $s \in S = \{s_1, ..., s_k\}$ that tomorrow obtains
- The true probability model is $m^* \in M$
- Consumption pairs $c = (c_0, c_1)$ are ranked by

$$V(c) = \mathrm{E}_{\mu}\phi\left(\mathrm{E}_m u(c)\right)$$

Valuation: static asset pricing

- Agents have an endowment in the two periods, but can also fund their consumption decisions by trading in a frictionless financial market
- A primary asset
$$y = (y_1, ..., y_k)$$
pays out y_i if state s_i obtains
- The Law of one price holds

Valuation: static asset pricing

- If the true model m^* is known, we have the classic pricing formula

$$p_y = \mathrm{E}_{m^*}\left(\frac{\frac{\partial u}{\partial c_1}(\hat{c})}{\frac{\partial u}{\partial c_0}(\hat{c})}y\right)$$

- Risk attitudes affect asset pricing
- In general, we have

$$p_y = \mathrm{E}_{\mu}\left(\frac{\phi'\left(\mathrm{E}_m u(\hat{c})\right)}{\mathrm{E}_{\mu}\phi'\left(\mathrm{E}_m u(\hat{c})\right)}\mathrm{E}_m\left(\frac{\frac{\partial u}{\partial c_1}(\hat{c})}{\frac{\partial u}{\partial c_0}(\hat{c})}y\right)\right)$$

- Both risk and ambiguity attitudes affect pricing
- In a series of papers, Hansen and Sargent study similar formulas and their relevance for some asset pricing empirical puzzles

Valuation: static asset pricing

- Suppose the risk free asset is traded, with (gross) return r_f
- The classic pricing formula can be written as

$$p_y = \frac{1}{r_f} \mathrm{E}_{\hat{m}^*}(y)$$

where $\hat{m}^*$ is the, equivalent, risk neutral version of m^*, given by

$$\hat{m}_i^* = \frac{\frac{\partial u}{\partial c_{i1}}(\hat{c})}{\mathrm{E}_{m^*} \frac{\partial u}{\partial c_1}(\hat{c})} m_i^*$$

Valuation: static asset pricing

- Adjustments for risk, ambiguity and model uncertainty
- Risk: $\hat{m}$ is the risk neutral version of model m given by

$$\hat{m}_i = \frac{\frac{\partial u}{\partial c_{i1}} (\hat{c})}{\mathrm{E}_{m^*} \frac{\partial u}{\partial c_1} (\hat{c})} m_i$$

- Ambiguity: $\hat{\mu}$ is the ambiguity neutral version of prior μ given by

$$\hat{\mu}(m) = \frac{\phi' (\mathrm{E}_m u (\hat{c}))}{\mathrm{E}_\mu (\phi' (\mathrm{E}_m u (\hat{c})))} \mu (m)$$

- Model uncertainty: $\tilde{\mu}$ is given by

$$\tilde{\mu}(m) = \frac{\mathrm{E}_m \frac{\partial u}{\partial c_1} (\hat{c})}{\mathrm{E}_{\hat{\mu}} \mathrm{E}_m \frac{\partial u}{\partial c_1} (\hat{c})} \hat{\mu} (m)$$

Valuation: static asset pricing

- $\hat{m} = m$ under risk neutrality (u linear)
- $\hat{\mu} = \mu$ under ambiguity neutrality (ϕ linear), though possibly $\tilde{\mu} \neq \mu$
- $\tilde{\mu} = \hat{\mu} = \mu$ when the expected marginal utility $\mathrm{E}_m \frac{\partial u}{\partial c_1}(\hat{c})$ is constant, and so model uncertainty is immaterial

Valuation: static asset pricing

- Uncertainty neutral pricing is given by

$$p_w = \frac{1}{r_f} \mathrm{E}_{\tilde{\mu}} \left(\mathrm{E}_{\hat{m}} w \right) = \frac{1}{r_f} \mathrm{E}_{\overline{\tilde{\mu}}} \left(w \right)$$

  where $\overline{\tilde{\mu}}_i = \sum_{m \in M} \hat{m}_i \tilde{\mu}(m)$
- $\overline{\tilde{\mu}}$ is the uncertainty neutral measure on S
- It involves expected marginal utilities, and so in principle it can be estimated from consumption data

Long run: is model uncertainty still relevant?

Does model uncertainty resolve in the long run through learning?

- Consider a recurrent decision problem, in a stationary environment
- What DMs observe depend on the actions they choose
- If the ex post feedback that they receive is partial, a partial identification problem (and so model uncertainty) arises
- It persists at steady state, after DMs learned everything they could (based on the long run frequencies of observations caused by their actions)

Long run: is model uncertainty still relevant?

- Organizing principle: *self-confirming equilibrium*
  - introduced in the early 1990s in the works of Battigalli, Fudenberg and Levine, and Kalai and Lehrer
- DMs best reply to the evidence they collected through their actions
- Steady state actions have to be best replies given the evidence they generated
- The true model being unknown (model uncertainty), prior beliefs might well be not correct
- No longer in a Nash setup where actions are best replies to correct beliefs

Long run: is model uncertainty still relevant?

Consider an urn with 90 Red, or Green, or Yellow balls

- DMs keep betting on Red
- Partial feedback: DMs observe whether or not they won (but not the drawn color)
- Suppose the long run frequency of "wins" is 1/3
- The proportion of Red balls is learned (it is 1/3, i.e., 30 Red balls)
- The proportions of Green and Yellow balls remain unknown
- Partial identification at steady state
- If DMs had observed the colors drawn (perfect feedback), they would have learned the true model (i.e., all colors' proportions)

Long run: is model uncertainty still relevant?

- Steady state betting on Red is only risky (DMs learned the proportion of red balls)
- Steady state betting on other colors remains ambiguous (DMs did not learn anything on their proportions)
- A status quo bias (betting on Red) emerges, captured through ambiguity
- Formally, betting on Red is self-confirming

Long run: is model uncertainty still relevant?

- In general, the bias favors tested alternatives over untested ones
- The higher ambiguity aversion, the higher the bias
- The bias might well trap DMs in self-confirming, but suboptimal (wrt to the true model), actions
- For example, if in the previous urn there are 50 Green balls, the (objectively) optimal action would be to bet on Green, not on Red

Long run: is model uncertainty still relevant?

- In a Game Theoretic setting, this causes a penalization of deviations. As a result, the set of self-confirming equilibria expands (Battigalli, Cerreia, Maccheroni, Marinacci, *AER* 2015)
- Folk wisdom I: "better the devil you know than the devil you do not know"
- Folk wisdom II: "chi lascia la via vecchia per la via nuova, sa quel che perde ma non sa quel che trova" ("*those who leave the old road for a new one, know what they leave but do not know what they will find*")

Long run: a glimpse to learning

- Consider a decision problem over time
- Experimentation is possible
- The degree of ambiguity aversion and of patience affect its option value
- The higher the degree of patience, the higher the value
- The higher the degree of ambiguity aversion, the lower the value
- Ongoing research on this trade-off (Battigalli, Cerreia, Francetich, Maccheroni, Marinacci 2015)

Sources of uncertainty

- We made a distinction between attitudes toward physical and epistemic uncertainty
- A more general issue: do attitudes toward different uncertainties differ?
- Source contingent outcomes: Do DMs regard outcomes (even monetary) that depend on different sources as different economic objects?
- Ongoing research on this subtle topic

Epilogue

- In decision problems with data, it is important to distinguish physical and epistemic uncertainty
- Traditional EU reduces epistemic uncertainty to physical uncertainty, and so it ignores the distinction
- Experimental and empirical evidence suggest that the distinction is relevant and may affect valuation
- We presented two approaches, one Bayesian and one not
- For different applications, different approaches may be most appropriate
- Traditional EU is the benchmark
- Yet, adding ambiguity broadens the scope (empirical and theoretical) and the robustness of results