On multiple discount rates

C. Chambers F. Echenique
Georgetown Caltech

Columbia Sept. 15 2017

This paper

A theory of intertemporal decision-making that is robust to the discount rate.

Motivation

Problem:

- Economists use present-value calculations to make decisions.
- Project evaluation or cost-benefit analysis.
- Calculations are very sensitive to the assumed discount rate.

Motivation

Weitzman (AER, 2001):
Cost-benefit analysis is now used to analyze environmental projects "the effects of which will be spread over hundreds of years ..."
"The *most critical single problem with discounting future benefits and costs is that no consensus now exists, or for that matter has ever existed, about what actual rate of interest to use.*"

Motivation: Climate change

Tony asks a question.

Nicholas gives an answer.

Stern report (2006) on global warming.

Climate change

Example: Stern report (2006) on global warming.

Hal Varian (NYT, 2006):
"should the social discount rate be 0.1 percent, as Sir Nicholas Stern, . . . would have it, or 3 percent as Mr. Nordhaus prefers?"

N. Stern: 0.1%

W. Nordhaus: 3%

Example: Stern report (2006) on global warming.

Hal Varian (NYT, 2006):

"There is *no definitive answer* to this question because it is inherently an ethical judgment that requires comparing the well-being of different people: those alive today and those alive in 50 or 100 years."

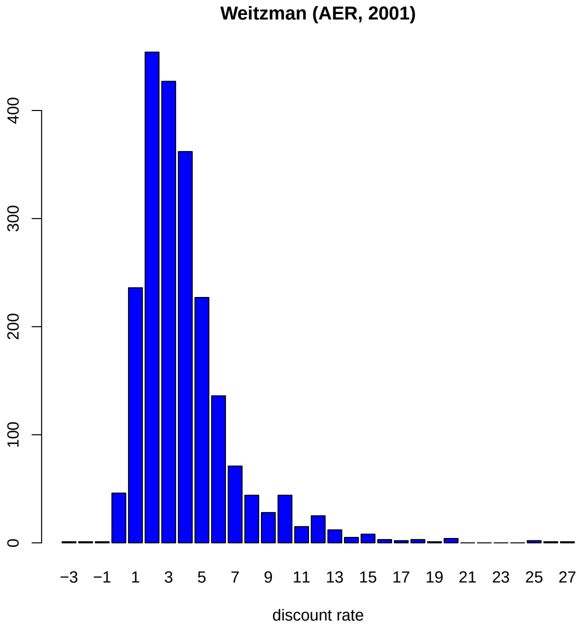
Motivation

- Not only ethical judgement.
- Also economic considerations, theoretical and empirical:
- What is the right model think about intertemporal tradeoffs? What is the right savings rate; growth rate; role of uncertainty, etc.

Weitzman (2001)

Survey of 2,160 Ph.D-level economists.

- "what real interest rate should be used to discount over time the benefits and costs of projects being proposed to mitigate the possible effects of global climate change."
- use "professionally considered gut feeling"



Weitzman (AER, 2001)
0
100
200
300
400
−3
−1
1
3
5
7
9
11
13
15
17
19
21
23
25
27
discount rate

Weitzman (2001)

Survey of 2,160 Ph.D-level economists:

- Mean rate : 3.96 %
- StdDev: 2.94 %

Smaller survey: 50 leading economists *(incl. G. Akerlof, K. Arrow, G. Becker, P. Krugman, D. McFadden, R. Lucas, R. Solow, J. Stiglitz, J. Tobin . . .)*

- Mean rate : 4.09 %
- StdDev: 3.07 %

Project evaluation

US Office of Management and Budget recommends:

Use discount rate between 1% and 7%, when evaluating "intergenerational benefits and costs."

Problem:

A decision maker

has to make a decision

Her advisors

have a set $D \subset (0,1)$ discount rates.

Primitives of our model.

- A set X $(= \ell_\infty)$ of sequences $\{x_n\}_{n=0}^{\infty}$.
- A (closed) set $D \subseteq (0,1)$ of discount factors.

- Sequences should be interpreted as *utility streams*.
- D could come from a survey (like Weitzman) or a government agency like the US Office of Management and Budget

Two criteria:

- Utilitarian

$$U(x) = \sum_{t=0}^{\infty} \left(\int_D (1-\delta)\delta^t d\mu(\delta) \right) x_t$$

where μ is a prob. measure on D.
(favored by Weitzman; analyzed recently by Jackson and Yariv)

Two criteria:

- Utilitarian

$$U(x) = \sum_{t=0}^{\infty} \left(\int_D (1-\delta)\delta^t d\mu(\delta) \right) x_t$$

where μ is a prob. measure on D.
(favored by Weitzman; analyzed recently by Jackson and Yariv)

- Maxmin

$$U(x) = \min\{(1-\delta)\sum_{t=0}^{\infty} \delta^t x_t : \delta \in \hat{D}\}$$

for $\hat{D} \subseteq D$.
(used for robustness in analogous situations with uncertainty).

How to think about utilitarian and maxmin

Example 1

Utilitarian with $D = \{\frac{1}{10}, \frac{9}{10}\}$ and uniform μ. Then

$$(1, -5.55, 0, 0, \ldots) \sim (0, 0, \ldots)$$

while

$$(0, 0, \ldots) \succ (\underbrace{0, \ldots, 0,}_{\text{9 periods}} 1, -5.55, 0, 0, \ldots)$$

(Issue highlighted by Weitzman and Jackson-Yariv)

Ruled out by maxmin.

How to think about utilitarian and maxmin

Example 2

$$(10, 8, 0, \dots) \succ (14, 4, 0 \dots)$$

while

$$(14, 1004, 0 \dots) \succ (10, 1008, 0, \dots).$$

Ruled out by utilitarian; allowed by maxmin.

How to think about utilitarian and maxmin

Example 3

$$(0,1,0,\ldots) \succ (0,0,2,0,\ldots)$$

while

$$(5,0,2,\ldots) \succ (5,1,0,\ldots)$$

(a failure of *separability*)

Ruled out by utilitarian; allowed by maxmin.

How to think about utilitarian and maxmin

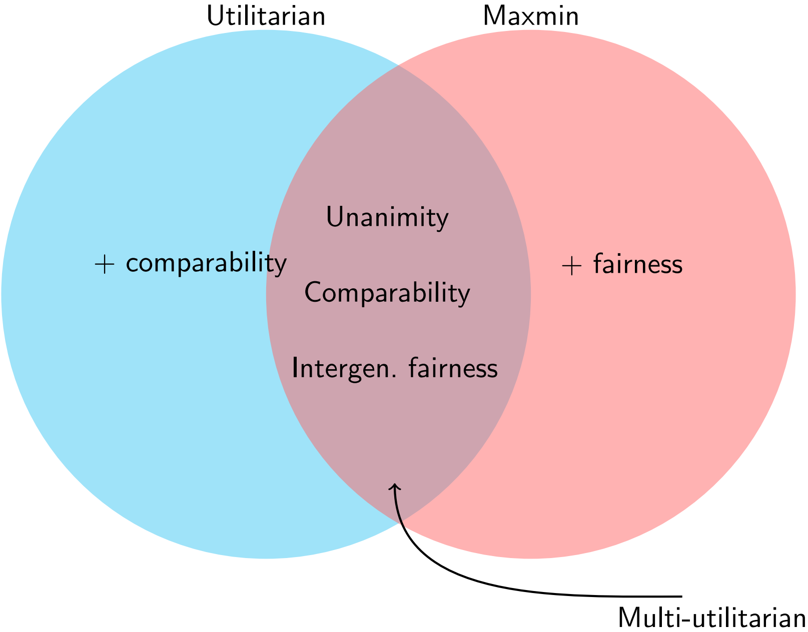
In *common:*

- Unanimity
- Intergenerational comparability.
- Intergenerational fairness.

Give rise to a new *multi-utilitarian* criterion.

Special about utilitarian: + Intergenerational comparability.

Special about maxmin: + Intergenerational fairness.


Utilitarian
Maxmin
+ comparability
Unanimity
Comparability
Intergen. fairness
+ fairness
Multi-utilitarian

Utilitarian and maxmin have in common:

A unanimity axiom.

If all experts recommend x over y, then choose x over y.

Utilitarian and maxmin have in common:

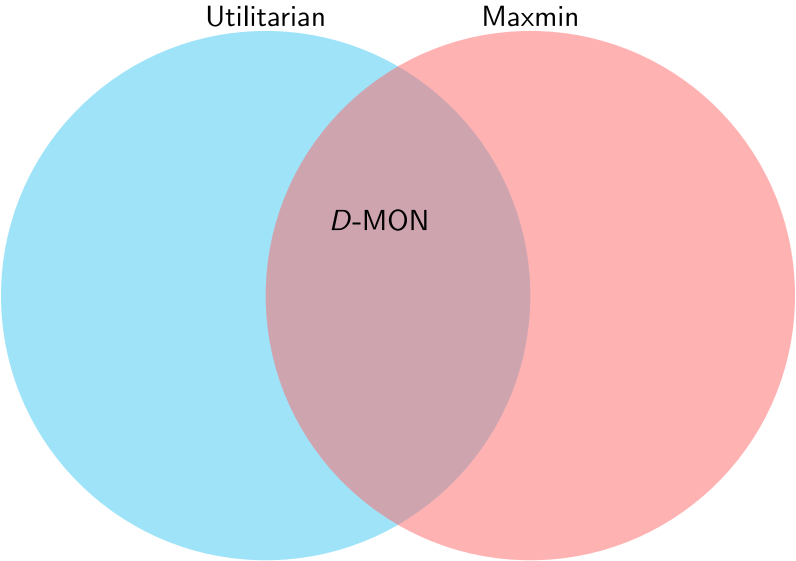
A unanimity axiom.

D-monotonicity:

$$(\forall \delta \in D) \quad (1-\delta)\sum_t \delta^t x_t \geq (1-\delta)\sum_t \delta^t y_t \Longrightarrow x \succeq y;$$

and

$$(\forall \delta \in D) \quad (1-\delta)\sum_t \delta^t x_t > (1-\delta)\sum_t \delta^t y_t \Longrightarrow x \succ y;$$


Utilitarian
Maxmin
D-MON

Utilitarian and maxmin have in common:

Intergenerational comparability of utility.

Co-cardinality (COC): For any $a > 0$ and constant seq. θ,

$$x \succeq y \text{ iff } ax + \theta \succeq ay + \theta.$$

How to think about COC:

- Wish to avoid the conclusion in Arrow's thm.
- Arrow's IIA says that only information on pairwise comparisons matter.
- Arrow: When comparing policies A and B, only generations' ordinal ranking of A and B is allowed to matter.
- To relax IIA, d'Aspremont and Gevers (1977), (formalizing Sen) propose COC.

Intergenerational comparability – COC

How to think about COC:

- Wish to avoid the conclusion in Arrow's thm.
- When comparing policies A and B, also *utility levels* may matter.
- But not when utilities result from *the same* affine transformation.

COC: Constrain choice when all generations' utilities are measured in the same units.

Intergenerational comparability – COC

- In comparing policies A and B, consider generation t's utility $U(A, t)$ and $U(B, t)$.
- Allow social decision to depend on utilities: weaken Arrow's IIA.
- Utility function $V(Z, t) = a + bU(Z, t)$ $(b > 0)$ represents the same preferences as U.

Intergenerational comparability – COC

- In comparing policies A and B, consider generation t's utility $U(A, t)$ and $U(B, t)$.
- Allow social decision to depend on utilities: weaken Arrow's IIA.
- Utility function $V(Z, t) = a + bU(Z, t)$ $(b > 0)$ represents the same preferences as U.
- COC says that social decisions are invariant to common affine transformations.
- Ex: $b = 1$. Then $V(A, t) - U(A, t) = V(B, t) - U(B, t) = a$ for all generations t.
- So decision on A vs. B should be the same.

Intergenerational comparability – COC

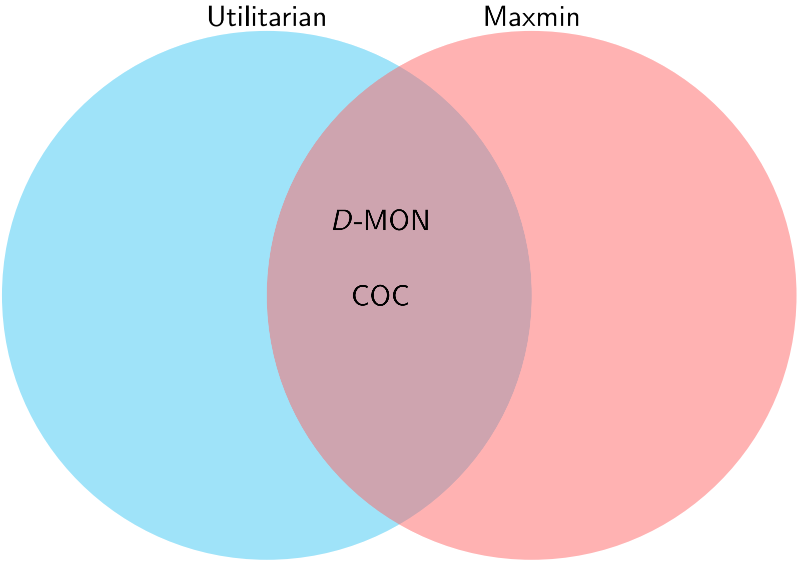
When COC fails.

Suppose

$$(10, 8, 0, \dots) \succ (14, 4, 0 \dots)$$

while

$$(1014, 1004, 1000 \dots) \succ (1010, 1008, 1000, \dots).$$

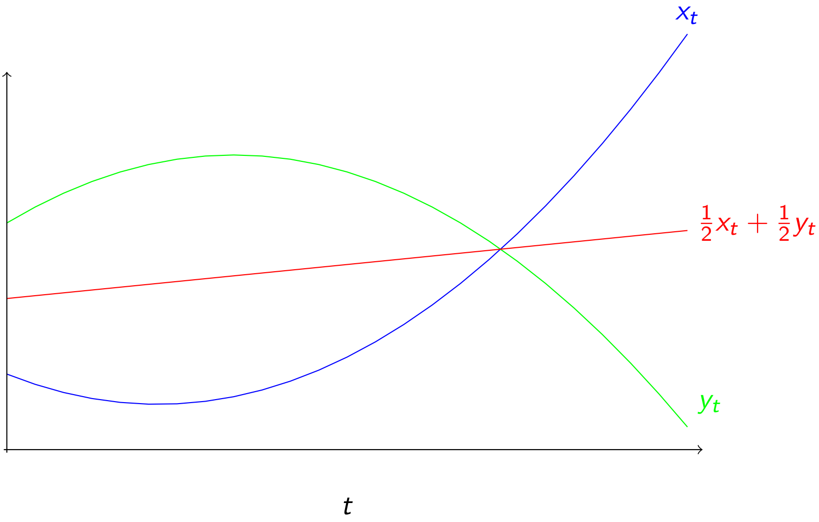

Utilitarian
Maxmin
D-MON
COC

Utilitarian and maxmin have in common:

Intergenerational fairness.

Convexity (CVX):

$$\left.\begin{array}{l} x \succeq \theta \\ y \succeq \theta \end{array}\right\} \Longrightarrow \lambda x + (1-\lambda) y \succeq \theta \quad \forall \lambda \in (0,1)$$


x_t
$\frac{1}{2}x_t + \frac{1}{2}y_t$
y_t
t

CVX

Note: CVX is an intrinsic preference for intertemporal smoothing.

Utilitarian and maxmin have in common:

D-MON, COC and CVX give rise to a *multi-utilitarian* criterion.

Utilitarian and maxmin have in common:

Theorem

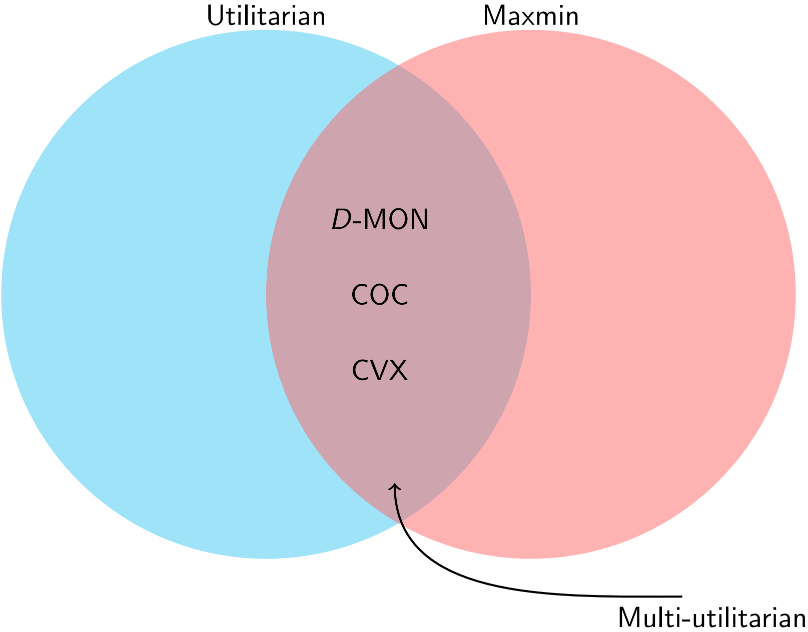
$\succeq$ satisfies

- *D-MON,*
- *COC,*
- *CVX,*
- *CONT*

iff $\exists$ a convex set $\Sigma \subseteq \Delta(D)$ s.t.

$$U(x) = \min_{\mu \in \Sigma} \sum_{t=0}^{\infty} \left(\int_D (1-\delta)\delta^t d\mu(\delta) \right) x_t$$

represents $\succeq$.


Utilitarian
Maxmin
D-MON
COC
CVX
Multi-utilitarian

What is special about Utilitarianism?

What is special about Utilitarianism?

Invariance with respect to individual origins of utilities (IOU):

$$x \succeq y \Longrightarrow x + z \succeq y + z.$$

What is special about Utilitarianism?

Theorem

$\succeq$ *satisfies the axioms in Theorem 1 and IOU iff* $\exists \mu \in \Delta(D)$ *s.t.*

$$U(x) = \sum_{t=0}^{\infty} \left(\int_D (1-\delta)\delta^t d\mu(\delta) \right) x_t$$

represents $\succeq$*.*

Intergenerational comparability – IOU

- In comparing policies A and B, consider generation t's utility $U(A,t)$ and $U(B,t)$.
- Suppose that $V(A,t) - U(A,t) = V(B,t) - U(B,t) = a_t$ for all generations t.
- No longer a common scale as in COC.
- Allow social decision to depend on the change in generations' utilities.

Intergenerational comparability – IOU

When IOU fails:

$$(10, 8, 0, \ldots) \succ (14, 4, 0 \ldots)$$

while

$$(14, 1004, 0 \ldots) \succ (10, 1008, 0, \ldots).$$

Intergenerational comparability – IOU

When IOU fails:

$$(0,1,0,\ldots) \succ (0,0,2,0,\ldots)$$

while

$$(5,0,2,\ldots) \succ (5,1,0,\ldots)$$

violates IOU because

$$(5,0,2,\ldots)-(0,1,0,\ldots) = (5,0,2,\ldots)-(0,0,2,\ldots) = (5,0,0,\ldots).$$

(a failure of *separability*)

What is special about Maxmin?

What is special about Maxmin?

Invariance to stationary relabeling (ISTAT):

For all $t \in \mathbf{N}$ and all $\lambda \in [0, 1]$,

$$x \sim \theta \Longrightarrow \lambda x + (1 - \lambda)(\underbrace{\theta, \ldots, \theta}_{t \text{ times}}, x) \sim \theta.$$

What is special about Maxmin?

Theorem

$\succeq$ satisfies the axioms in Theorem 1, STAT and COMP iff $\exists \hat{D} \subseteq D$ (nonempty and closed) s.t.

$$U(x) = \min\{(1-\delta)\sum_{t=0}^{\infty} \delta^t x_t : \delta \in \hat{D}\}$$

represents $\succeq$.

Meaning of ISTAT

Recall "anonymity," a basic notion of fairness: Social decisions shouldn't depend on agents' names.

Should we impose anonymity?

Meaning of ISTAT

Recall “anonymity,” a basic notion of fairness: Social decisions shouldn’t depend on agents’ names.

Should we impose anonymity?

We may have:

$$(-1, 3, 3, -1, 0, \ldots) \sim 0$$

and

$$0 \succ (-1, -1, 3, 3, 0, \ldots),$$

a violation of anonymity but natural in the intertemporal context.

Meaning of ISTAT :

ISTAT says that $(-1,3,3,-1,0,\ldots) \sim 0$ implies

$$(0,-1,3,3,-1,0,\ldots) \sim 0.$$

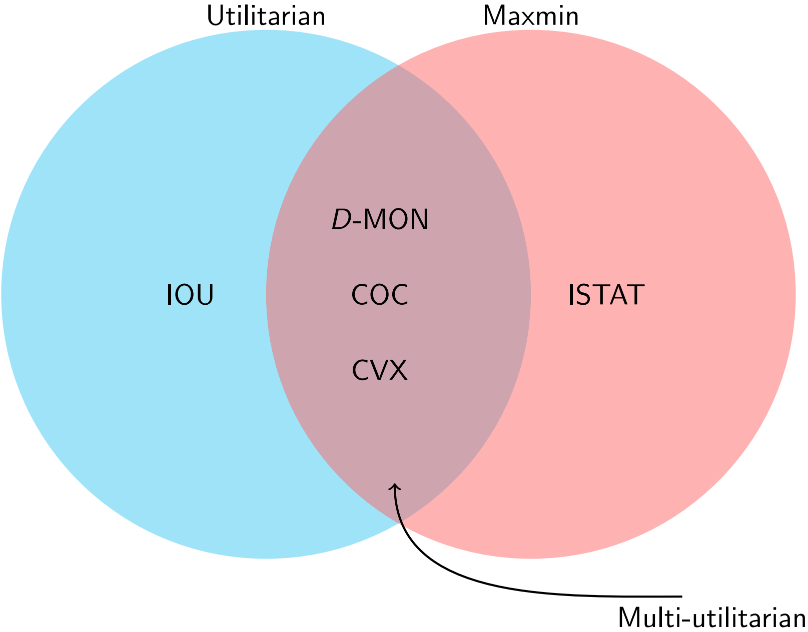
Note $(0,-1,3,3,-1,0,\ldots)$ results from $(-1,3,3,-1,0,\ldots)$ by treating (or "relabeling") generation t as $t-1$, for $t \geq 1$.

Meaning of ISTAT :

ISTAT says that if $x \sim \theta$ then

$$x' = (\underbrace{\theta, \ldots, \theta}_{T \text{ times}}, x) \sim x.$$

- x' results from x by a relabeling of agents' names: $x'_t = x_{t-1}$ $(t \geq T+1)$.
- This is a relabeling of generations $t = T+1, T+2, \ldots$
- Generations $t = 0, \ldots T$ receive θ, the same as they would receive under the alternative stream θ.


Utilitarian
Maxmin
D-MON
IOU
COC
ISTAT
CVX
Multi-utilitarian

Notation

- ℓ_1 set of all *absolutely summable* sequences.
- ℓ_∞ set of all *bounded* sequences.
- $\mathbf{1} = (1, 1, \ldots)$
- For $\theta \in \mathbf{R}$, we denote by θ the seq. $\theta\mathbf{1}$.

Notation

$$\|x\|_1 = \sum_{t=0}^{\infty} |x_t|$$

$$\|x\|_\infty = \sup\{|x_t| : t = 0, 1, \ldots\}$$

Ideas in the proofs.

We look at the set $P = \{x \in \ell^\infty : x \succeq 0\}$.

We want to characterize this as having the form:

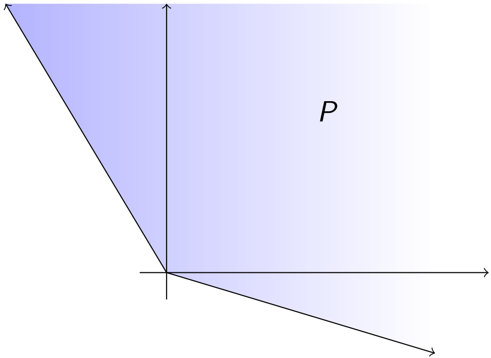
$$\bigcap_{\delta \in D} \{x : (1-\delta) \sum_t \delta^t x_t \geq 0\}$$

The rest are details.

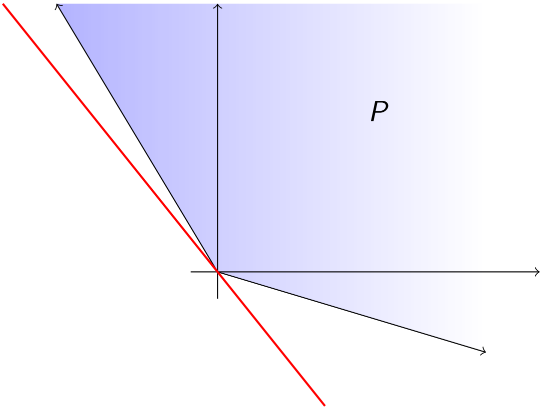
Ideas in the proofs.

The set $P = \{x \in \ell^\infty : x \succeq 0\}$ is a closed, convex cone.

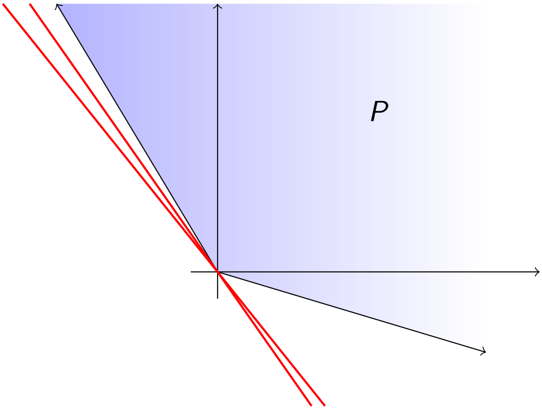
P is a closed cvx. cone



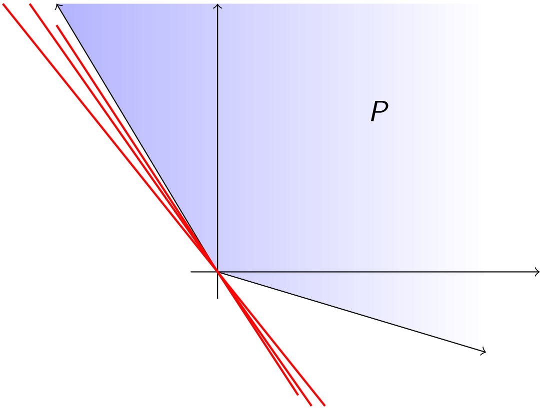
P is a closed cvx. cone



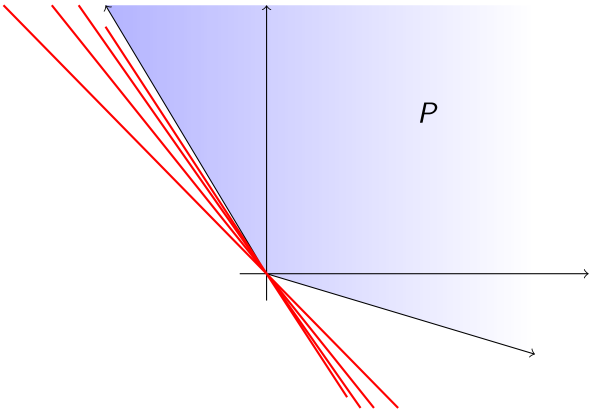
P is a closed cvx. cone



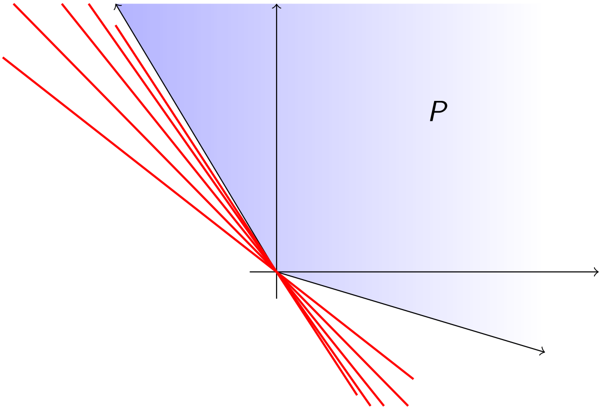
P is a closed cvx. cone



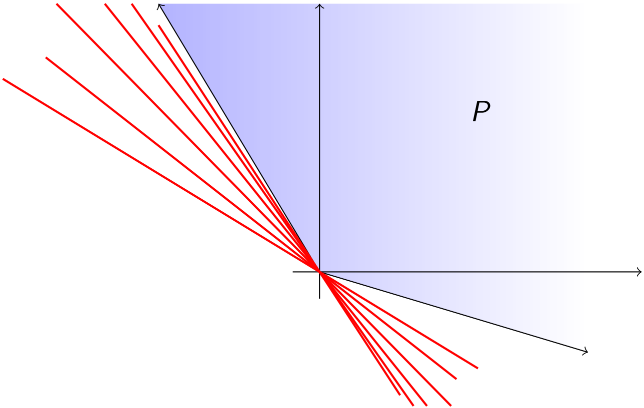
P is a closed cvx. cone



P is a closed cvx. cone



P is a closed cvx. cone



By duality, and using cont. at infinity:

$$P = \bigcap_{m \in M} \{x : x \cdot m \geq 0\}$$

for some set M of prob. measures on $\{0, 1, 2, \dots\}$.

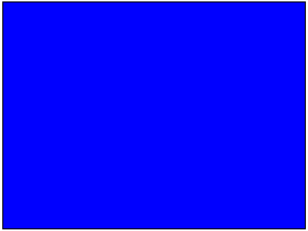
A multiple-prior representation.

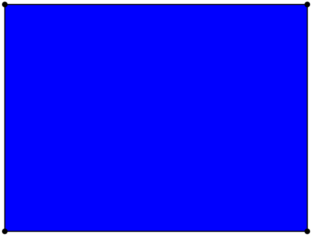
To say something about M, natural approach is:

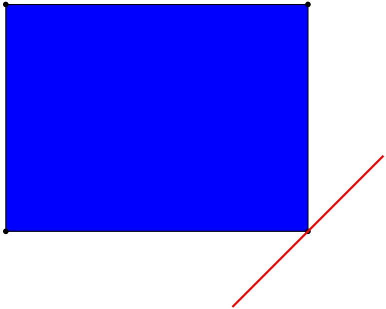
$$\min m \cdot z$$
$$\text{s.t.} m \in M$$

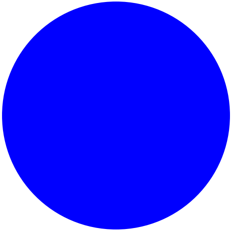
Solutions are extreme points of M.

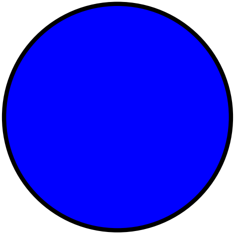
Challenge: work with extreme points of M isn't enough. We need *unique* solutions.

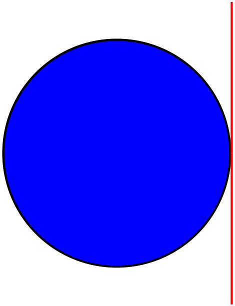


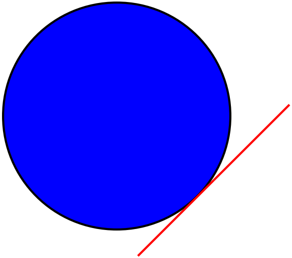


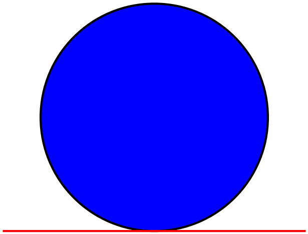


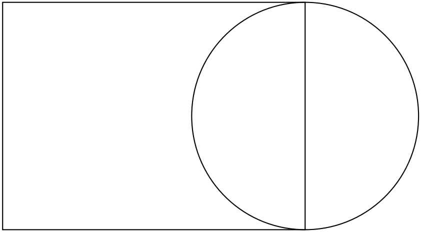


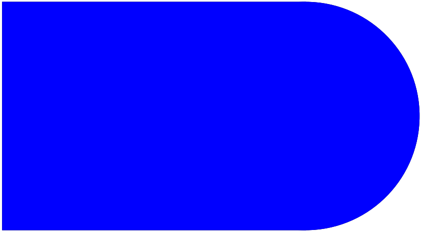


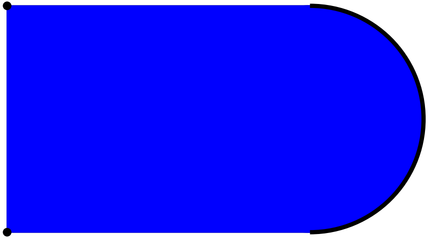












An *exposed point* of M is a point $m' \in M$ such that there is some x for which $x \cdot m' < x \cdot m$ for all $m \in M \backslash \{m'\}$.

A result of Lindenstrauss and Troyanski in our context:

Theorem

In our context, a weakly compact convex set is the (weakly) closed convex hull of its strongly exposed points (and hence exposed points).

Since M consists of prob measures, any exposed point m' can be chosen with corresponding x satisfying $x \cdot m' = 0$.

Hence for such x (in the maxmin case) $x \sim 0$.

But $x \sim 0$ implies $x + (0, 0, \dots, 0, x) \sim 0$ by stationarity

Then since we have indifference, there is also a supporting $m_x \in M$ for which $0 = m_x \cdot x + m_x \cdot (0, 0, \dots, 0, x) \leq m_x \cdot y$ for all $y \in P$.

Also by stationarity, $(0, 0, \dots, 0, x) \sim 0$. So :

$$\begin{cases} (0, 0, \dots, 0, x) \in P \Longrightarrow m_x \cdot (0, 0, \dots, 0, x) & \geq 0 \\ x \in P \Longrightarrow m_x \cdot x & \geq 0 \end{cases}$$

Conclude $m_x = m$, since x exposes m.

But $x \sim 0$ implies $x + (0, 0, \ldots, 0, x) \sim 0$ by stationarity

Then since we have indifference, there is also a supporting $m_x \in M$ for which $0 = m_x \cdot x + m_x \cdot (0, 0, \ldots, 0, x) \leq m_x \cdot y$ for all $y \in P$.

Also by stationarity, $(0, 0, \ldots, 0, x) \sim 0$. So :

$$\begin{cases} (0, 0, \ldots, 0, x) \in P \Longrightarrow m_x \cdot (0, 0, \ldots, 0, x) & \geq 0 \\ x \in P \Longrightarrow m_x \cdot x & \geq 0 \end{cases}$$

Conclude $m_x = m$, since x exposes m.

But $x \sim 0$ implies $x + (0, 0, \ldots, 0, x) \sim 0$ by stationarity

Then since we have indifference, there is also a supporting $m_x \in M$ for which $0 = m_x \cdot x + m_x \cdot (0, 0, \ldots, 0, x) \leq m_x \cdot y$ for all $y \in P$.

Also by stationarity, $(0, 0, \ldots, 0, x) \sim 0$. So :

$$\begin{cases} (0, 0, \ldots, 0, x) \in P \Longrightarrow m_x \cdot (0, 0, \ldots, 0, x) & = 0 \\ x \in P \Longrightarrow m_x \cdot x & = 0 \end{cases}$$

Conclude $m_x = m$, since x exposes m.

Let m^T be the "updated" m

$$m^T = \frac{(m(T-1), m(T), m(T+1), \ldots)}{m(\{T-1, \ldots\})}.$$

Then: $0 = m_x \cdot (0, 0, \ldots, 0, x) = m \cdot (0, 0, \ldots, 0, x)$ means that $m^T \cdot x = 0$

Whenever $p \in P$, $(0, 0, \ldots, 0, p) \in P$ (again by stationarity). So $m \cdot (0, 0, \ldots, 0, p) \geq 0$ and thus $m^T \cdot p \geq 0$

Conclude $m^T \in M$.

So $m^T \in M$ and $m^T \cdot x = 0$ gives $m^T = m$ since x exposes m.

Characterizes the geometric distribution.

Literature

- Karcher and Trannoy (1999); Foster and Mitra (2003); Wakai (2008), Drugeon, Thai and Hanh (2016).
- Debate on δ/multiple δ in project evaluation: Ramsey (1928), Weizman (2001), Stern (2006), Nordhaus (2007), Dasgupta (2007), Feng and Ke (2017).
- Aggregation necessitates unique δ: Marglin (1963), Feldstein (1964), Zuber (2011), Jackson-Yariv (2014), Jackson-Yariv (2014).
- Multiple priors literature: Bewley (1986; 2002), Gilboa-Schmeidler (1989), Chateauneuf, Maccheroni, Marinacci, and Tallon (2005).
- Random δ: Higashi, Hyogo, Takeoka, (2009), Pennesi (2014), Lu-Saito (2015).