The role of mixed discounting in risk-averse sequential decision-making

Wenjie Huang

Department of Data and Systems Engineering
Musketeers Foundation Institute of Data Science
The University of Hong Kong

Joint work with Erick Delage (HEC Montréal) and Shanshan Wang (Beihang U.)

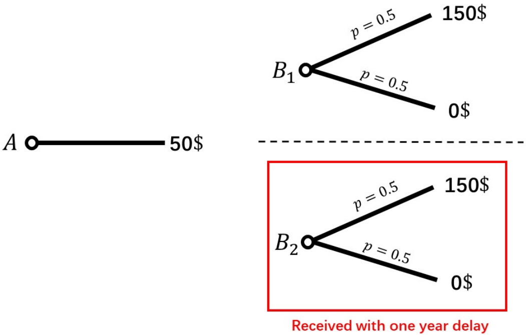
HKU IDS Interdisciplinary Workshop
– Exploring the Foundations: Fundamental AI and Theoretical Machine Learning

May 29, 2025

The Role of Discount Factor

- Optimization: Finite value;
- Finance: Prevailing interest rate – Compute the present value of future cash flows values;
- Behavioral Economics: Time preference/Impulsivity: Favoring present vs. future;

A Lottery Choice Example



- Use power utility model $u(x) = x^\alpha$ where $\alpha = 0.9$.
- Plan A and B_1 have expected utility 33.81 and 45.44, respectively.
- An annual discount factor $\gamma \in (0, 1)$ is introduced. The "discounted" utility is $u(\gamma x) = (\gamma x)^\alpha$. If $\gamma \in [0.72, 1)$, $A \preceq B_2$; otherwise, $A \succeq B_2$.

Basic Models

- Given two potential expenses z_0 and Z_1, with z_0 an immediate deterministic amount while Z_1 is a random one received at time T, the risk-averse exponentially discounted preference is:

$$\text{(D-model)} \qquad \mathcal{R}(z_0, Z_1, T) := \rho(z_0 + \gamma^T Z_1),$$

  where ρ is a risk measure (RM).

- The representation reduces to the well known expected total discounted cost when $\rho(X) := \mathbb{E}[X]$:

$$\text{(E-model)} \qquad \mathcal{R}(z_0, Z_1, T) := \mathbb{E}[z_0 + \gamma^T Z_1] = z_0 + \gamma^T \mathbb{E}[Z_1]$$
$$= \mathbb{E}[z_0 + \mathbf{1}\{\tau \geq T\} Z_1],$$

  where τ is a random interruption time which follows an exponential distribution with mean $-1/\ln(\gamma)$.

[Shwartz 2001, Ermoliev 2010]

Basic Models (Con)

- The random interruption equivalence does not carry through for more general forms of RMs. For instance,

$$\begin{aligned}\mathbb{E}[u(z_0 + \{\tau \geq T\}Z_1)] &= (1-\gamma^T)u(z_0) + \gamma^T[u(z_0 + Z_1)] \\ &\neq [u(z_0 + \gamma^T Z_1)].\end{aligned}$$

  unless $u(\cdot)$ is linear.

- The proposed random interruption model ("RI-model"):

$$\text{(RI-model)} \qquad \mathcal{R}(z_0, Z_1, T) := \rho(z_0 + \mathbf{1}\{\tau \geq T\}Z_1),$$

  and more generally a mixture model ("M-model"),

$$\text{(M-model)} \qquad \mathcal{R}(z_0, Z_1, T) := \rho(z_0 + \mathbf{1}\{\tau \geq T\}\gamma_d^T Z_1),$$

  where τ follows an exponential distribution with mean $-1/\ln(\gamma_r)$.

Empirical Study Supporting the M-model

- Existing field experiment results [Lopez-Guzman 2018] (Acknowledgment: Grateful for the authors providing field experiment data for this research !);
- Subjects were asked to undertake two sessions: Risk Attitude (RA) task and Inter-Temporal Choice (ITC) task;
- Logistic regression;

1. The monthly discounting ranged from 2.9% to 99.5% of the reward, with a median of 60%. From a purely financial point of view, note that a monthly discounting of 2.9% is already equivalent to assuming a yearly interest rate of nearly 42%. In comparison, the average credit card rate in the US over the period of 2014-2018 was below 17%.

$$\text{Discounting} = 1 - \frac{1}{(1 + \text{Interest rate})^{\text{Period}}}$$

2. RI-model achieves a higher goodness-of-fit performance than the D-model, but whether the D-model can indeed safely be rejected, is left for investigation.

Empirical Study Supporting the M-model (Con)

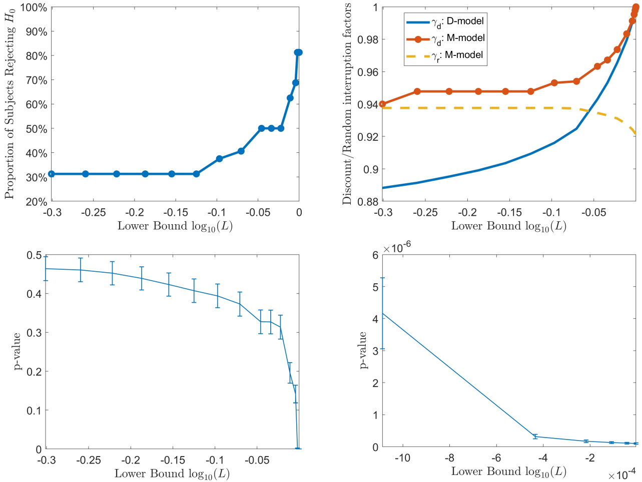
- New features in our analyses:
  1. Use Entropic risk measure: $\rho(X) := (1/\beta)\ln(\mathbb{E}[\exp(\beta X)])$, with $\beta > 0$ as the risk aversion parameter; 32 participants are shown to be risk-averse by RA task.
  2. Credit card assumption: Subjects have access to credit with a daily interest rate less than $(1/L - 1)$ for some $L > 0$ and cannot create value directly from this credit instrument.

     $\Rightarrow$ We necessarily have that $\gamma_d \geq L$ (constrained case).
  3. Conduct likelihood ratio tests for D-model and M-model, under both constrained and unconstrained cases;

- Unconstrained case: Estimated discount factor γ_d is consistent with the findings in [Lopez-Guzman 2018]. For 10 out of the 32 participants, the D-model can safely be rejected in favor of the M-model.

Empirical Study Supporting the M-model (Con)

- Constrained case: 26 subjects (81% of the population) eventually reject D-model as L approaches 1.



Risk Preference Mappings

- Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathfrak{F} := (\mathcal{F}_t)_{t\in\mathbb{N}}$ be a filtration, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ for all $t \in \mathbb{N}$. Consider for all $t \in \mathbb{N}$, a set of random liability $\mathcal{Z}_t \subseteq L_p(\Omega, \mathcal{F}_t, P)$ for some $p \in [1, \infty]$. Consider cash flows $Z_{\mathbb{N}} := (Z_t)_{t\in\mathbb{N}}$ with $Z_t \in \mathcal{Z}_t$ for all $t \geq 0$, adapted to filtration $\mathfrak{F}$. Define a preference mapping $\mathcal{R} : \mathcal{Z}_{\mathbb{N}} \to \mathbb{R}$ that is both monotone, convex, and recursive.

$$\text{(D-model)} \qquad \mathcal{R}_D(Z_{\mathbb{N}}) := \limsup_{T\to\infty} \rho\Big(\sum_{t=0}^{T} \gamma^t Z_t\Big).$$

$$\text{(M-model)} \qquad \mathcal{R}_M(Z_{\mathbb{N}}) := \limsup_{T\to\infty} \rho\Big(\sum_{t=0}^{\min\{\boldsymbol{\tau}, T\}} \gamma_d^t Z_t\Big),$$

where $\mathbb{P}[\boldsymbol{\tau} = t] = (1-\gamma_r)\gamma_r^t$ and $\mathbb{P}[\boldsymbol{\tau} \geq t] = \gamma_r^t$, and $1\{\boldsymbol{\tau} \geq t\}$ is adapted to the filtration $\mathfrak{F}$.

Risk Preference Mappings (Con)

Definition 1 (Recursive preference mapping)

The preference mapping $\mathcal{R}(\cdot)$ is recursive if there exists a preference system $\{\bar{\mathcal{R}}_t\}_{t\in\mathbb{N}}$, with each $\bar{\mathcal{R}}_t : \mathcal{Z}_t \times \mathcal{Z}_{t+1} \to \mathcal{Z}_t$ such that, for all $Z_{\mathbb{N}} \in \mathcal{Z}_{\mathbb{N}}$:

$$\mathcal{R}(Z_{\mathbb{N}}) = \limsup_{T\to\infty} \bar{R}_0(Z_0,\ \bar{R}_1(Z_1, \ldots, \bar{R}_{T-1}(Z_{T-1}, Z_T) \ldots)).$$

Definition 2 (Recursive risk measure)

The risk measure $\rho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \to \mathbb{R}$ is a "recursive risk measure" if there exists a set of risk measures $\{\rho_t\}_{t=0}^{\infty}$ such that

$$\rho(X) = \limsup_{T\to\infty} \rho_0(\rho_1 \cdots \rho_{T-2}(\rho_{T-1}(\mathbb{E}[X|\mathcal{F}_T])) \cdots),$$

where each $\rho_t : \mathcal{L}_p(\Omega, \mathcal{F}_{t+1}, P) \to \mathcal{L}_p(\Omega, \mathcal{F}_t, P)$ is a conditional risk mapping [Ruszczynski 2006].

Risk Preference Mappings (Con)

Theorem 3 (Recursive Formulation of D-model)

D-model satisfies $\bar{\mathcal{R}}_t(Z_t, Z_{t+1}) := Z_t + \gamma^{-t}\rho_t(\gamma^{t+1}Z_{t+1})$*, which reduces to* $\bar{\mathcal{R}}_t(Z_t, Z_{t+1}) := Z_t + \gamma\rho_t(Z_{t+1})$*, when the conditional risk mappings are coherent. In the latter case, we have that:*

$$\mathcal{R}_D(Z_{\mathbb{N}}) = \limsup_{T\to\infty} Z_0 + \gamma\rho_0(Z_1 + \gamma\rho_1(Z_2 + \cdots + \gamma\rho_{T-1}(Z_T)\cdots)). \quad (1)$$

Risk Preference Mappings (Con)

Theorem 4 (Recursive Formulation of M-model)

M-model satisfies $\bar{\mathcal{R}}_t(Z_t, Z_{t+1}) := Z_t + \gamma_d^{-t}\rho_t(0 \oplus_{\gamma_r} \gamma_d^{t+1} Z_{t+1})$*, which reduces to* $\bar{\mathcal{R}}_t(Z_t, Z_{t+1}) := Z_t + \gamma_d \rho_t(0 \oplus_{\gamma_r} Z_{t+1})$*, when the conditional risk mappings are coherent. In the latter case, we have:*

$$\begin{aligned}&\mathcal{R}_M(Z_{\mathbb{N}})\\ =&\limsup_{T\to\infty} Z_0 + \gamma_d \rho_0(0 \oplus_{\gamma_r} (Z_1 + \gamma_d \rho_1(0 \oplus_{\gamma_r} (Z_2 + \cdots + \gamma_d \rho_{T-1}(0 \oplus_{\gamma_r} Z_T)\cdots)))).\end{aligned} \tag{2}$$

An Optimal Stopping Problem Example

- Consider a stochastic process $A_{\mathbb{N}} := (A_t)_{t\in\mathbb{N}}$ of periodical cost (negative of payment) opportunities with each $A_t \in \mathcal{A} \subset \Re_-$. DM must identify when to accept the payment using a controlled stop time (CST) process $\mathbf{s}$ such that $\mathbf{1}\{\mathbf{s} \geq t\}$ is adapted to the natural filtration of the $A_{\mathbb{N}}$ process.
- Given a CST strategy $\mathbf{s}$, the cost flow produced takes the form: $Z_t(\mathbf{s}) := A_{\mathbf{s}}\mathbf{1}\{t = \mathbf{s}\}$ for all $t \in \mathbb{N}$. Letting $\mathcal{S}$ denote the set of all eligible CST strategies, we have the problem:

$$\min_{\mathbf{s}\in\mathcal{S}} \mathcal{R}(Z_{\mathbb{N}}(\mathbf{s})).$$

Assume that $\mathcal{A}$ is bounded (i.e. $\mathcal{A} \subseteq [-\bar{A}, \bar{A}]$) and each A_t are i.i.d.

An Optimal Stopping Problem Example (Con)

Proposition 1

When $\gamma_d < 1$ in general and when $\gamma_r < 1$ if the conditional risk mappings are either comonotone additive or the entropic risk measure and 2) $\mathcal{A} \subset (-\infty, 0]$, the infinite-horizon problem can be approximated to any level of precision using a finite-horizon one.

- Dynamic programming formulation for T horizon problem:

$$V_t^*(A_t) = \min\left\{A_t,\ \gamma_d^{-t}\rho_t(0 \oplus_{\gamma_r} \gamma_d^{t+1} V_{t+1}^*(A_{t+1}))\right\},$$

with $V_T^*(A_T) = \min\{0, A_T\}$. An optimal policy is

$$s_T^* = \inf\{t : V_t^*(A_t) = A_t\} = \inf\{t : A_t \leq R_t\}.$$

- Equivalently, we could solve $R_t = \gamma_d^{-t}\rho(0 \oplus_{\gamma_r} \gamma_d^{t+1} \min\{R_{t+1}, \mathbf{a}\})$, with $R_T = 0$, where $\mathbf{a}$ follows certain distribution on $[-\bar{A}, 0]$.

An Optimal Stopping Problem Example (Con)

- We observe that,

$$\begin{aligned} R_t =& \gamma_d^{-t} \rho(0 \oplus_{\gamma_r} \gamma_d^{t+1} \min\{R_{t+1}, \mathbf{a}\}) \\ =& \frac{1}{\beta\gamma_d^t} \log(\mathbb{E} \exp(0 \oplus_{\gamma_r} (\beta\gamma_d^t)\gamma_d \min\{R_{t+1}, \mathbf{a}\})), \end{aligned}$$

  when $t \to \infty$ and $\gamma_d < 1$, the risk parameter in the conditional entropic risk measure $\beta\gamma_d^t \to 0$ is dissolved, such that we are approaching for solving the risk-neutral equation.

- Under $\mathbf{a} \sim U[-\bar{A}, 0]$, Setting the truncation

$$R_T = \bar{A}(\gamma_d\gamma_r)^{-1}(\sqrt{1 - (\gamma_d\gamma_r)^2} - 1),$$

  instead of $R_T = 0$, yields a better approximation [Hau 2022].

The Properties of D- and M-model

- Finiteness

Proposition 2

Given that $Z_t \in [-B, \ B]$ for all $t \geq 0$, $B \geq 0$, and ρ is a normalized nested risk measure with law-invariant conditional risk mappings, then the M-model is finite if either of the following hold:

(i) $\gamma_d < 1$,

(ii) $\gamma_r < 1$ *and* $Z_t \leq 0$ *for all* $t \geq 0$,

(iii) when $\gamma_r < 1$ *and the nested risk measure is composed of utility-based shortfall conditional risk mappings [Follmer 2002] with loss functions with subdifferentials bounded in a strictly positive interval,*

(iv) when $\gamma_r < 1$ *and the nested risk measure is composed of optimized certainty equivalent [Ben-Tal 1986] conditional risk mappings that employ a surjective and stationary loss function.*

The Properties of D- and M-model (Con)

- Interchangeability

Proposition 3

Given that $Z_t \in [-B, 0]$ *for all* $t \geq 0$, *and* ρ *is a normalized recursive risk measure with subjective conditional measures, then RI- and M- model can be reformulated as D-model with adjusted lower risk measure and lower discount factor.*

- * Counter-example: Entropic risk measure

The Properties of D- and M-model (Con)

- Ordering

Proposition 4

Under a recursive risk measure composed of law-invariant conditional measures, and if $\gamma_d < 1$ and bounded $X_t \geq 0$ for all $t \geq 0$, M-model will lower bound D-model with $\gamma = \gamma_d$.

Definition 1

The conditional mapping ρ_t is "mixture concave" if
$\forall\, \gamma \in [0,\, 1], \forall\, X, Y \in \mathcal{L}_p(\Omega, \mathcal{F}_{t+1}, P)$, $\rho_t(X \oplus_\gamma Y) \geq (1-\gamma)\rho_t(X) + \gamma\rho_t(Y)$ *a.s.*

Proposition 5

Under a recursive risk measure composed of law-invariant and "mixture concave" conditional measures (not necessarily coherent), M-model will upper bound D-model with $\gamma = \gamma_d\gamma_r$ and $\bar{\rho}_t(X) := \gamma_r^{t+1}\rho_t(\gamma_r^{-(t+1)}X)$ as the conditional risk mapping. The latter reduce to $\bar{\rho}_t(X) := \rho_t(X)$ in the case that ρ is coherent.

The Properties of D- and M-model (Con)

- Examples of "mixture concave" convex risk measure:
  1. [Delage 2019] Spectral risk measure, Mean variance, Mean standard deviation, Mean (semi-) deviation from target and Mean weighted mean deviation from quantile.
  2. Optimized certainty equivalent:

$$
\begin{aligned}
&\rho(X \oplus_\gamma Y) \\
=&\inf_{t\in\mathbb{R}}\{t + \mathbb{E}[\ell(X \oplus_\gamma Y - t)]\} \\
=&\inf_{t\in\mathbb{R}}\{t + (1-\gamma)\mathbb{E}[\ell(X - t)] + \gamma\mathbb{E}[\ell(Y - t)]\} \\
\geq&(1-\gamma)\inf_{t\in\mathbb{R}}\{t + \mathbb{E}[\ell(X - t)]\} + \gamma\inf_{t\in\mathbb{R}}\{t + \mathbb{E}[\ell(Y - t)]\} \\
=&(1-\gamma)\rho(X) + \gamma\rho(Y).
\end{aligned}
$$

  3. [Postek 2016] Other minimum of affine function with expected loss

- *Counter-example: Utility-based shortfall with piecewise loss function

Numerical Experiments

- $\bar{A} = 10$; $T = \lceil 1/(1-\gamma_d) \rceil$ for D-model and $T = \lceil 1/(1-\gamma_d \gamma_r) \rceil$ for M-model;

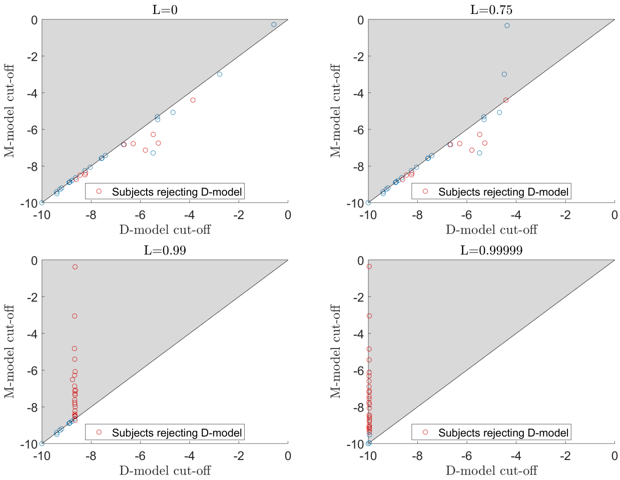


Figure: Distribution of D-model and M-model cut-off values pair

Numerical Experiments (Con)

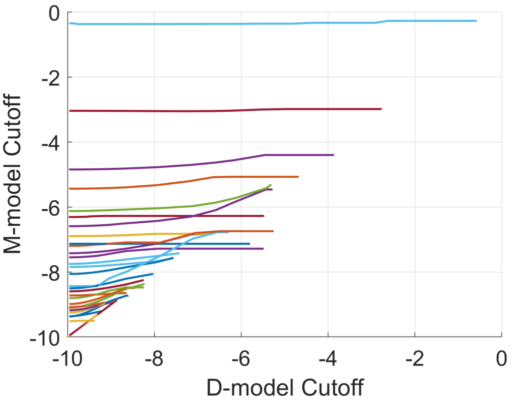


Figure: Cutoffs for D- and M-model as a function of L (Each curve represents a participant of the study)

Numerical Experiments (Con)

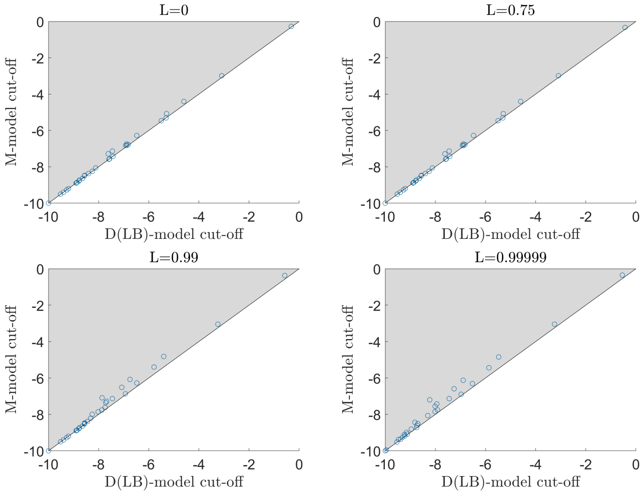
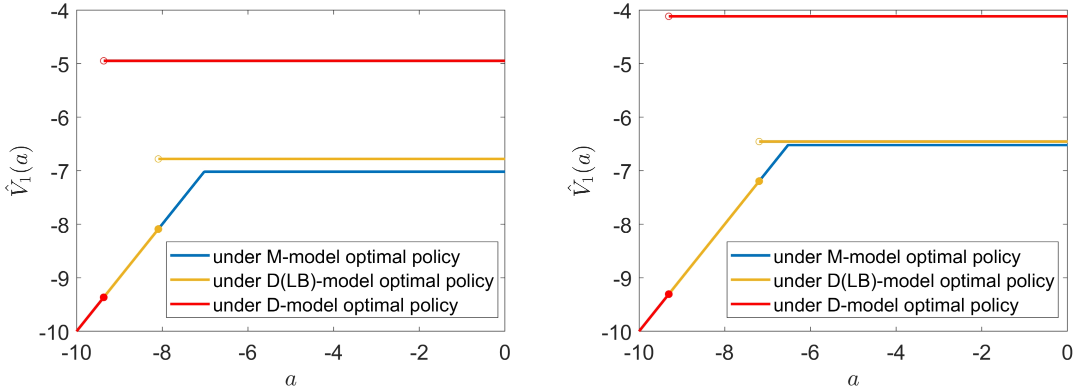


Figure: Distribution of D(LB)-model and M-model cut-off values pair

Numerical Experiments (Con)



(a) Unconstrained participant (b) Constrained participant

Figure: Conditional first-stage risk of running the D-model, D(LB)-model, and M-model policy when true model is an M-model

Numerical Experiments (Con)

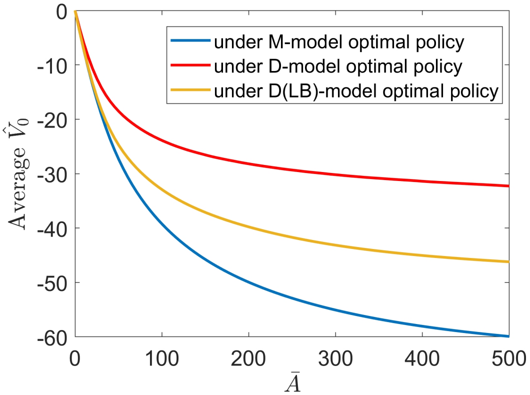


Figure: Average $\hat{V}_0$ for rejected subjects of running D-model, D(LB)-model and M-model policy when the true model is an M-model

Conclusion

1. Present evidences through field experiment data that decision makers are better represented by a propsoed M-model under certain assumptions;
2. Extend the “risk preference mapping” framework [Pichler 2022] to an infinite horizon setting and show how it can be reduced to M-model;
3. Numerical experiments on a simple optimal stopping problem, validate the effects of the proposed model;
4. Address the relationships between D-, and M-model (RI-model);

Reference

[Lopez-Guzman 2018] Lopez-Guzman, S., Konova, A. B., Louie, K., & Glimcher, P. W. (2018). Risk preferences impose a hidden distortion on measures of choice impulsivity. *PLoS One*, 13(1), e0191357.

[Shwartz 2001] Shwartz, A. (2001). Death and discounting. *IEEE Transactions on Automatic Control*, 46(4), 644-647.

[Ermoliev 2010] Ermoliev, Y., Ermolieva, T., Fischer, G., & Makowski, M. (2010). Extreme events, discounting and stochastic optimization. *Annals of Operations Research*, 177(1), 9-19.

[Ruszczyński 2006] Ruszczyński, A., & Shapiro, A. (2006). Conditional risk mappings. *Mathematics of Operations Research*, 31(3), 544-561.

[Hau 2022] Hau, J. L., Petrik, M., Ghavamzadeh, M., Russel, R. (2022). RASR: Risk-Averse Soft-Robust MDPs with EVaR and Entropic Risk. *arXiv preprint arXiv:2209.04067*.

[Follmer 2002] Föllmer, H., & Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and stochastics*, 6, 429-447.

[Ben-Tal 1986] Ben-Tal, A., & Teboulle, M. (1986). Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science*, 32(11), 1445-1466.

[Delage 2019] Delage, E., Kuhn, D., & Wiesemann, W. (2019). "Dice"-sion–making under uncertainty: When can a random decision reduce risk?. *Management Science*, 65(7), 3282-3301.

[Postek 2016] Postek, K., den Hertog, D., Melenberg, B. (2016). Computationally tractable counterparts of distributionally robust constraints on risk measures. *SIAM Review*, 58(4), 603-650.

[Pichler 2022] Pichler, A., Liu, R. P., & Shapiro, A. (2022). Risk-averse stochastic programming: Time consistency and optimal stopping. *Operations Research*, 70(4), 2439-2455.

Thank you !

Q & A

Huang, W., Delage, E., & Wang, S. (2024). The Role of Mixed Discounting in Risk-Averse Sequential Decision-Making. Available at SSRN 5013140.