Preference Robust Utility-based Shortfall Risk Minimization

Erick Delage
HEC Montréal

(joint work with Shaoyan Guo [Dalian U. of Tech.]
and Huifu Xu [U. of Southampton])

Workshop on Robust Optimization, Avignon, June 29th, 2018

HEC MONTRÉAL

Canada Research Chairs
Chaires de recherche du Canada

Canada

Outline

- Why preference robust optimization ?
- PRO for utility-based shortfall risk measures
- Numerical experiments
- Conclusion

The portfolio selection problem

- An individual meets with his financial advisor to tell him he wishes to invest in a given industrial sector, country, etc.
- Since uncertain factors affect performance, a « good » portfolio is one where the risks of losses are best justified by the potential gains

The portfolio selection problem

- An individual meets with his financial advisor to tell him he wishes to invest in a given industrial sector, country, etc.
- Since uncertain factors affect performance, a « good » portfolio is one where the risks of losses are best ju... potential g...

How can we identify optimal investments?

The strength of utility theory

- In 1954, G. Debreu established that if the preference relation is complete, transitive, and continuous, then there exists a ~~utility~~ risk mapping such that

$$X \succeq Y \quad \Leftrightarrow \quad u(X) \geq u(Y)$$

  where X, Y describe two financial positions

- This implies that any such preference relation can be numerically optimized

$$\underset{x \in \mathcal{X}}{\text{maximize}} \quad u(Z(x))$$

The strength of utility theory

- In 1954, G. Debreu established that if the preference relation is complete, transitive, and continuous, then there exists a ~~utility~~ risk mapping such that

$$X \succeq Y \quad \Leftrightarrow \quad \rho(X) \leq \rho(Y)$$

  where X, Y describe two financial positions
- This implies that any such preference relation can be numerically optimized

$$\text{minimize}_{x \in \mathcal{X}} \quad \rho(Z(x))$$

HEC MONTRĒAL

The strength of utility theory

- In 1954, G. Debreu established that if the preference relation is complete, transitive, and continuous, then there exists a ~~utility~~ risk mapping such that

$$X \succeq Y \quad \Leftrightarrow \quad \rho(X) \leq \rho(Y)$$

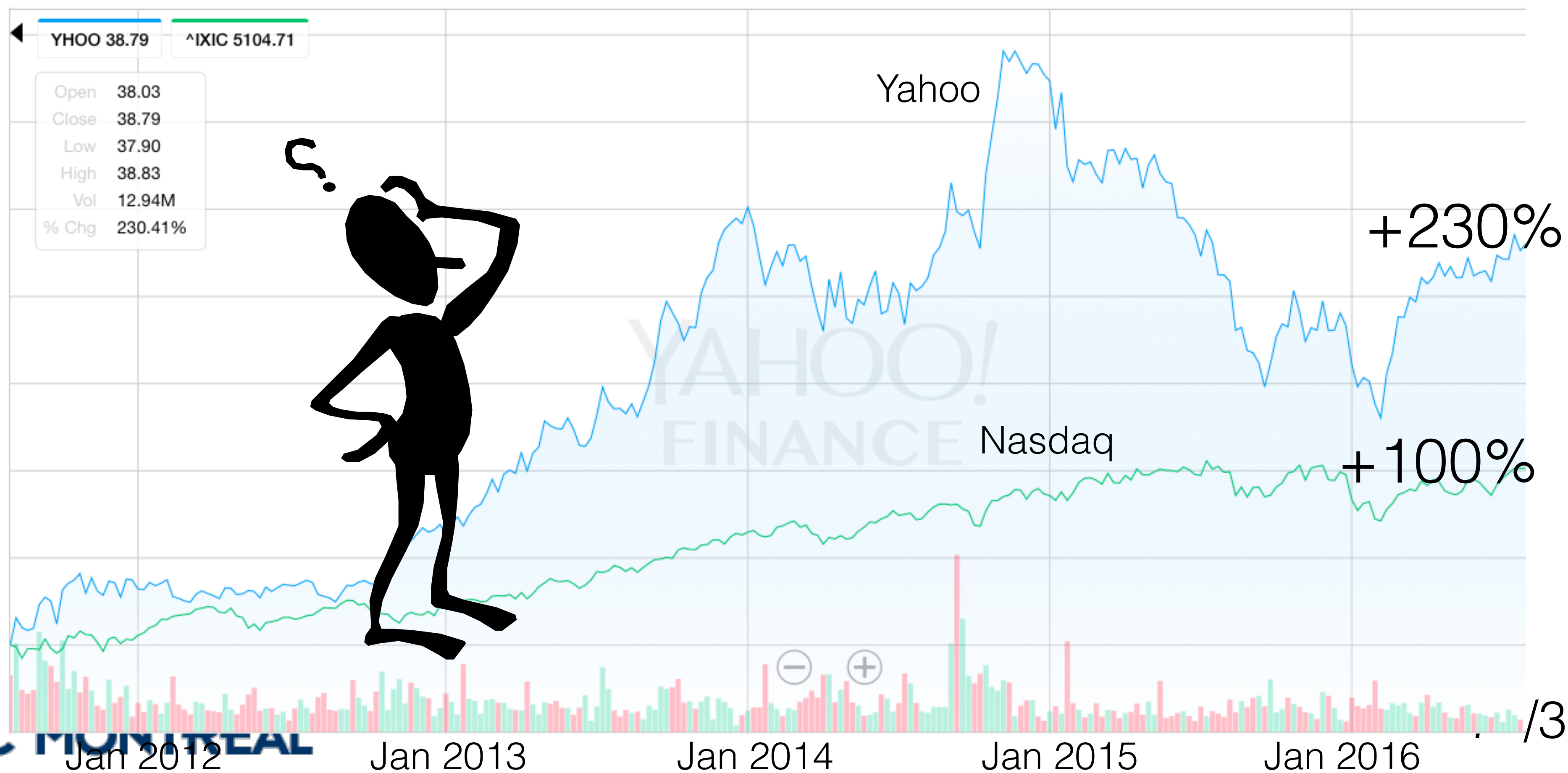
whe … s

- This … on
can

WARNING!
Numerical optimization can only be done once subjective preferences have been fully characterized.

How can we characterize risk preferences?

- An investor can indicate what type of wealth evolution he is comfortable with



How can one assess risk tolerance?

[Grable & Lytton, Financial Services Review (1999)]

1. You have just finished saving for a « once-in-a-lifetime » vacation. Three weeks before you plan to leave, you lose your job. You would:
   A. Cancel the vacation
   B. Take a much more modest vacation
   C. Go as scheduled, reasoning that you need the time to prepare for a job search
   D. Extend your vacation, because this might be your last chance to go first-class

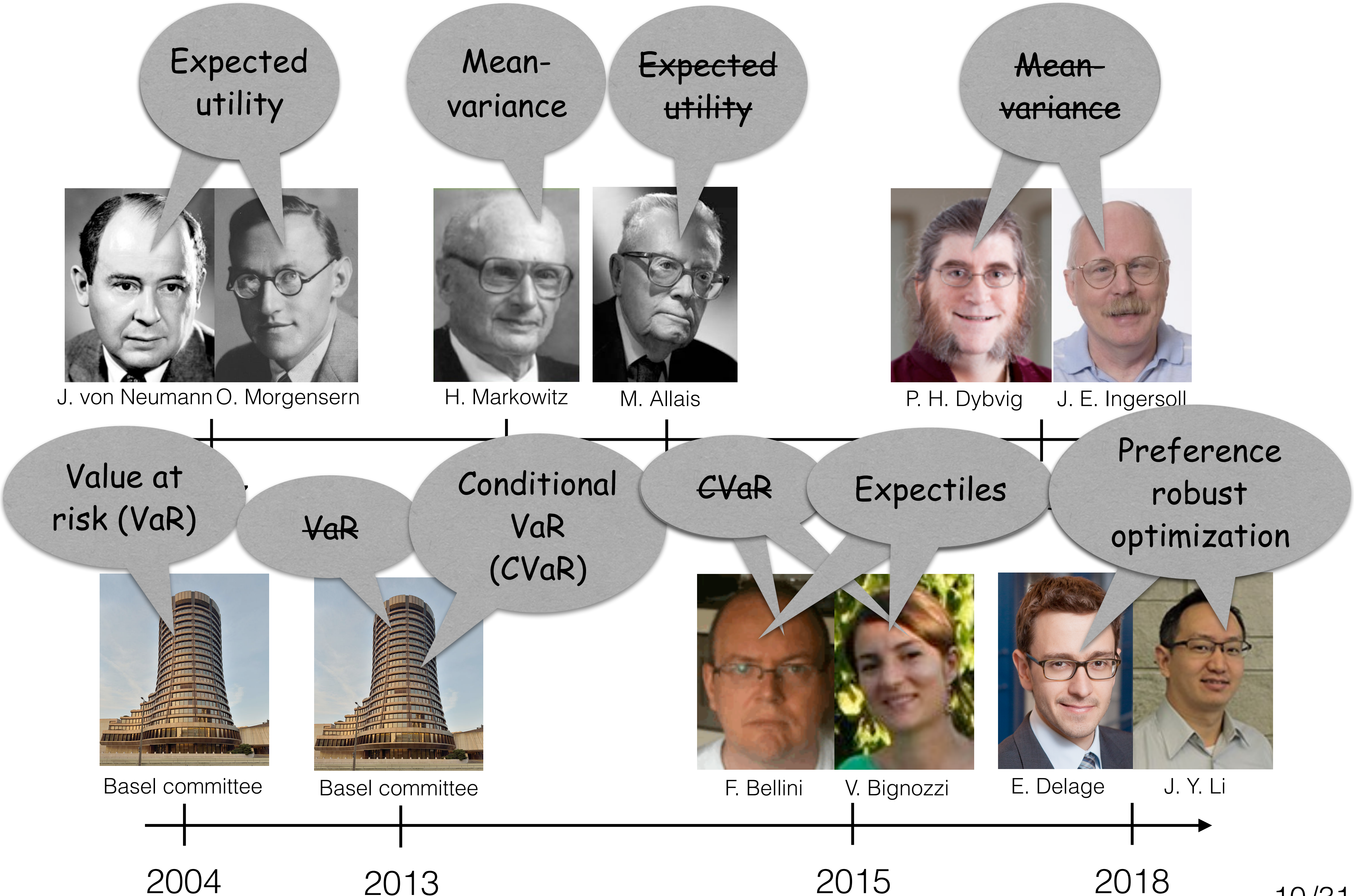
2. You are on a TV game show and can choose one of the following. Which would you take?
   A. $1,000 in cash
   B. A 50% chance at winning $ 5000
   C. A 25% chance at winning $ 10,000
   D. A 5% chance at winning $100,000

The limitations of utility theory

- Issue #1: One cannot make **all possible comparisons**
- Issue #2: One can easily provide **false information** about his preferences (Kahneman & Tversky, 1979)

- Solutions :
  - Make **simplifying assumptions** about the structure of $\rho(\cdot)$ in order to allow **interpolation** and **filter errors**
  - Employ a scheme that **handles uncertainty** about $\rho(\cdot)$

What is the right structure for a risk measure ?



Preference robust optimization
for
utility-based shortfall risk measures

Axiomatic assumptions

Let (Ω, Σ, P) be a probability space with $|\Omega| = M$
and let X, Y, and $Z(x) : \Omega \to \mathbb{R}$ be random variables

- Monotonicity: $X \geq Y \;\Rightarrow\; \rho(X) \leq \rho(Y)$
- Risk Aversion: $\rho(\theta X + (1-\theta)Y) \leq \theta\rho(X) + (1-\theta)\rho(Y)$, $\forall \theta \in [0,\ 1]$
- Law Invariance: $X =_P Y \Rightarrow \rho(Y) = \rho(X)$
- Translation Invariance: $\rho(X + t) = \rho(X) - t$, $\forall t$
- Elicitability (Bellini & Bignozzi, 2015): $\exists$ incentive mechanism for proper reporting
- M+RA+LI+TI = Law invariant **convex** risk measure (Kusuoka, 2001)
  - +Scale Invariance : Law invariant **coherent** risk measure
- M+RA+LI+TI+Elicit. = **Utility-based shortfall** (UBSF) risk measure (Föllmer & Schied, 2002)

What do we know about ρ ?

- The risk measure is a member of the set:

$$\mathcal{R} := \{\rho : \mathcal{L}_p \to \mathbb{R} \mid \rho(0) = 0$$

Monotonicity: $\rho(\cdot)$ non-increasing

Risk aversion: $\rho(\cdot)$ convex

Translation invariance: $\rho(X + t) = \rho(X) - t\,,\ \forall\, X, t$

Scale invariance: $\rho(\alpha X) = \alpha\rho(X)\,,\ \forall\, X, \alpha \geq 0$

Confidence intervals: $\rho(w_k^+) \leq \rho(W_k) \leq \rho(w_k^-)\,,\ \forall k$

Law invariance: (see details in D. & Li, 2018)

Elicitability: $\exists l \in \mathcal{L},\ \rho(X) = \inf\{t : \mathbb{E}_P[l(-X - t)] \leq l(0)\}\,,\ \forall\, X\}$

where $\mathcal{L}$ is the set of convex non-decreasing functions that are strictly increasing for all $y \geq -\epsilon$.

What should we optimize?

- We minimize the preference robust risk measure (D. & Li, 2018):

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{\rho\in\mathcal{R}} \; \rho(Z(x))$$

- Letting $\mathrm{SR}_l^P(X) := \inf\{t : \mathbb{E}_P[l(-X-t)] \leq l(0)\}$, we get:

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{l:\mathrm{SR}_l^P\in\mathcal{R}} \; \mathrm{SR}_l^P(Z(x))$$

- This reduces to:

$$\begin{array}{ll} \underset{x\in\mathcal{X},t}{\text{minimize}} & t \\ \text{subject to} & \mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\, l : \mathrm{SR}_l^P \in \mathcal{R} \end{array}$$

What should we optimize?

- We minimize the preference robust risk measure (D. & Li, 2018):

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{\rho\in\mathcal{R}} \; \rho(Z(x))$$

- Letting $\mathrm{SR}_l^P(X) := \inf\{t : \mathbb{E}_P[l(-X-t)] \leq l(0)\}$, we get:

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{l\in\mathcal{L}(\mathcal{R})} \; \mathrm{SR}_l^P(Z(x))$$

- This reduces to:

$$\begin{aligned} \underset{x\in\mathcal{X},t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\; l\in\mathcal{L}(\mathcal{R}) \end{aligned}$$

What do we know about l ?

- In the case of **convex** UBSF risk measures:

$$\mathcal{L}(\mathcal{R}) := \{l : \mathbb{R} \to \mathbb{R} \mid$$

Non-decreasing:

Convex:

Strictly increasing:

Confidence intervals:

$$\}$$

- In the case of **coherent** UBSF risk measures:

$$\mathcal{L}(\mathcal{R}) := \{l : \mathbb{R} \to \mathbb{R} \mid \exists \tau \in [\tau^-, \tau^+], \ l(x) = \max(\tau x, (1-\tau)x)\}$$

What do we know about l ?

- In the case of **convex** UBSF risk measures:

$\mathcal{L}(\mathcal{R}) := \{l : \mathbb{R} \to \mathbb{R} \mid$

Non-decreasing: $\exists l' : \mathbb{R} \to \mathbb{R} : l'(x) \geq 0 \,, \forall x \in \mathbb{R}$

Convex: $l(y) \geq l(x) + (y - x)l'(x) \,, \forall x, y \in \mathbb{R}$

Strictly increasing: $l(0) = 0, \quad l(-1) = -1$

Confidence intervals: $\mathbb{E}_P[l(-W_k + w_k^-)] \leq l(0)$

$\mathbb{E}_P[l(-W_k + w_k^+)] \geq l(0) \quad \}$

- In the case of **coherent** UBSF risk measures:

$$\mathcal{L}(\mathcal{R}) := \{l : \mathbb{R} \to \mathbb{R} \mid \exists \tau \in [\tau^-, \ \tau^+], \ l(x) = \max(\tau x, \ (1 - \tau)x)\}$$

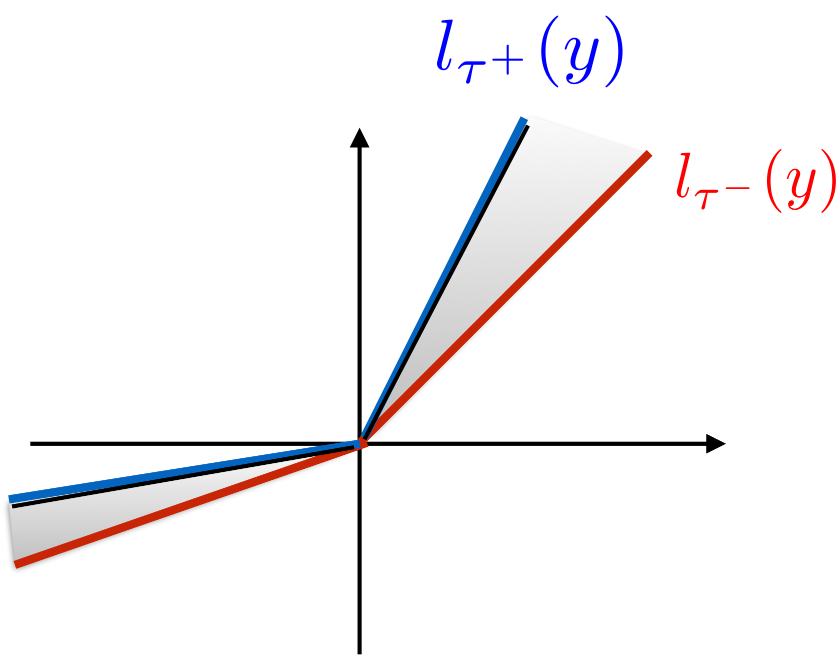
Linear programming representation (I)

- In the case of **coherent** UBSF risk measure:

$$\mathbb{E}_P[l(-Z(x) - t)] \leq l(0)\,, \forall\, l \in \mathcal{L}(\mathcal{R})$$

is shown equivalent to:

$$\mathbb{E}_P[\max(\tau(-Z(x) - t),\ (1-\tau)(-Z(x) - t))] \leq 0\,, \forall\, \tau \in [\tau^-,\ \tau^+]$$



HEC MONTRĒAL

Linear programming representation (I)

- In the case of **coherent** UBSF risk measure:

$$\mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\, l \in \mathcal{L}(\mathcal{R})$$

is shown equivalent to:

$$\mathbb{E}_P[\max(\tau^+(-Z(x)-t),\ (1-\tau^+)(-Z(x)-t))] \leq 0$$

- Hence, equivalent to:

$$\exists v \in \mathbb{R}^M,\ \sum_{\omega\in\Omega} p_\omega v_\omega \leq 0$$

$$v_\omega \geq \tau^+(-Z_\omega(x)-t)\,,\ \forall\, \omega \in \Omega$$

$$v_\omega \geq (1-\tau^+)(-Z_\omega(x)-t)\,,\ \forall\, \omega \in \Omega$$

Linear programming representation (II)

- In the case of **convex** UBSF risk measure:

$$\mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\, l \in \mathcal{L}(\mathcal{R})$$

is shown equivalent to

$$\sup_{l:\mathbb{R}\to\mathbb{R},\, l':\mathbb{R}\to\mathbb{R}} \quad \sum_{\omega\in\Omega} p_\omega l(-Z_\omega(x)-t) \qquad \leq 0$$

$$\text{subject to}$$

$$l(y') \geq l(y) + (y'-y)l'(y)\,, \forall\, y, y' \in \mathbb{R}$$

$$l(0) = 0\,, \quad l(-1) = -1$$

$$l'(y) \geq 0\,, \forall\, y \in \mathbb{R}$$

$$\sum_{y\in\mathcal{S}'} P(-W_k + w_k^- = y)l(y) \leq 0, \forall\, k$$

$$\sum_{y\in\mathcal{S}'} P(-W_k + w_k^+ = y)l(y) \geq 0, \forall\, k.$$

$\mathcal{S}'$ contains the support sets of all $-W_k + w_k^-$, $-W_k + w_k^+$, 0, and -1.

Linear programming representation (II)

- In the case of **convex** UBSF risk measure:

$$\mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\, l \in \mathcal{L}(\mathcal{R})$$

is shown equivalent to

$$\begin{aligned}
\sup_{v\geq 0, w, l:\mathcal{S}'\to\mathbb{R}, l':\mathcal{S}'\to\mathbb{R}} \quad & \sum_{\omega\in\Omega} p_\omega[v_\omega(-Z_\omega(x)-t)+w_\omega] \leq 0 \\
\text{subject to} \quad & v_\omega y + w_\omega \leq l(y)\,, \forall\, y \in \mathcal{S}', \omega \in \Omega \\
& l(y') \geq l(y) + (y'-y)l'(y)\,, \forall\, y, y' \in \mathcal{S}' \\
& l(0) = 0\,, \quad l(-1) = -1 \\
& l'(y) \geq 0\,, \forall\, y \in \mathcal{S}' \\
& \sum_{y\in\mathcal{S}'} P(-W_k + w_k^- = y) l(y) \leq 0, \forall\, k \\
& \sum_{y\in\mathcal{S}'} P(-W_k + w_k^+ = y) l(y) \geq 0, \forall\, k.
\end{aligned}$$

$\mathcal{S}'$ contains the support sets of all $-W_k + w_k^-$, $-W_k + w_k^+$, 0, and -1.

HEC MONTRÉAL

Linear programming representation (II)

- In the case of **convex** UBSF risk measure:

$$\mathbb{E}_P[l(-Z(x)-t)] \leq l(0)\,, \forall\, l \in \mathcal{L}(\mathcal{R})$$

is shown equivalent to

$$\begin{aligned}
\sup_{v\geq 0, w, \alpha, \beta} \quad & \sum_{\omega\in\Omega} p_\omega[v_\omega(-Z_\omega(x)-t)+w_\omega] \leq 0 \\
\text{subject to} \quad & v_\omega y_j + w_\omega \leq \alpha_j\,, \forall\, j=1,\ldots,N, \omega\in\Omega \\
& \alpha_i \geq \alpha_j + (y_i - y_j)\beta_j\,, \forall\, i,j=1,\ldots,N \\
& \alpha_{j_0} = 0\,, \quad \alpha_{j_-} = -1 \\
& \beta_j \geq 0\,, \forall\, j=1,\ldots,N \\
& \sum_{j=1,\ldots,N} P(-W_k + w_k^- = y_j)\alpha_j \leq 0,\ \forall\, k \\
& \sum_{j=1,\ldots,N} P(-W_k + w_k^+ = y_j)\alpha_j \geq 0,\ \forall\, k.
\end{aligned}$$

Legend:

$\alpha_i := l(y_i)$

$\alpha_{j_0} := l(0)$

$\alpha_{j_-} := l(-1)$

$\beta_i := l'(y_i)$

The case of continuous Ω

- Consider the preference robust risk minimization problem:

$$(\vartheta, x^*) := \min_{x\in\mathcal{X},t} \quad t \quad \text{s.t.} \quad \sup_{l\in L} \mathbb{E}_P[l(c(x,\xi)-t)] \leq l(0)$$

- One can approximate this problem with:

$$(\vartheta_N, x_N^*) := \min_{x\in\mathcal{X},t} \quad t \quad \text{s.t.} \quad \sup_{l\in L} \mathbb{E}_{P_N}[l(c(x,\xi)-t)] \leq l(0)$$

discrete approximation of P

- In fact,

Theorem 3:

Under assumptions yet to be described, for any small enough δ and large enough N,

$$\mathbb{P}(|\vartheta_N - \vartheta| \geq \delta) \leq Ce^{-\beta N}.$$

Furthermore, $x_N^* \to x^*$ with probability one.

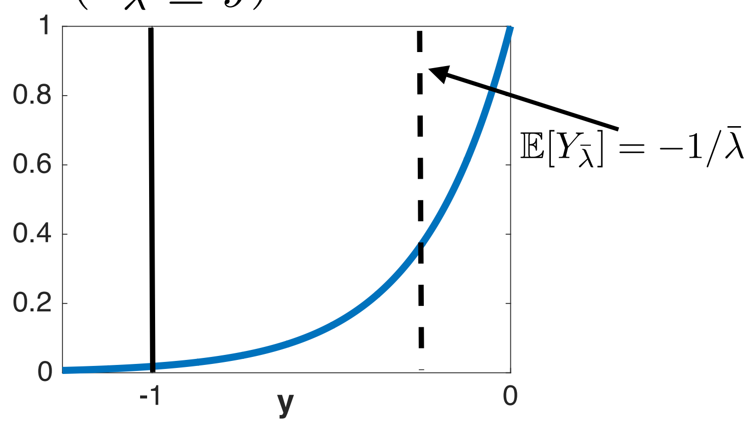
Assumptions needed for Theorem 3

- The set $\mathcal{X}$ is compact
- The function $c(x, \xi)$ is continuous in ξ and Hölder continuous in x

$$|c(x,\xi) - c(x',\xi)| \leq r(\xi)\|x - x'\|^{\nu}, \ \forall x, x' \in \mathcal{X}, \xi \in \Xi$$

- The preference robust risk minimization problem satisfies Slater's condition
- There exists a $\bar{\lambda}$ such that the risk of $Y_{\bar{\lambda}}$ is lower than the risk of a certain loss of 1

$\mathbb{P}(Y_{\bar{\lambda}} \leq y) = 1 - e^{-\bar{\lambda} y}$



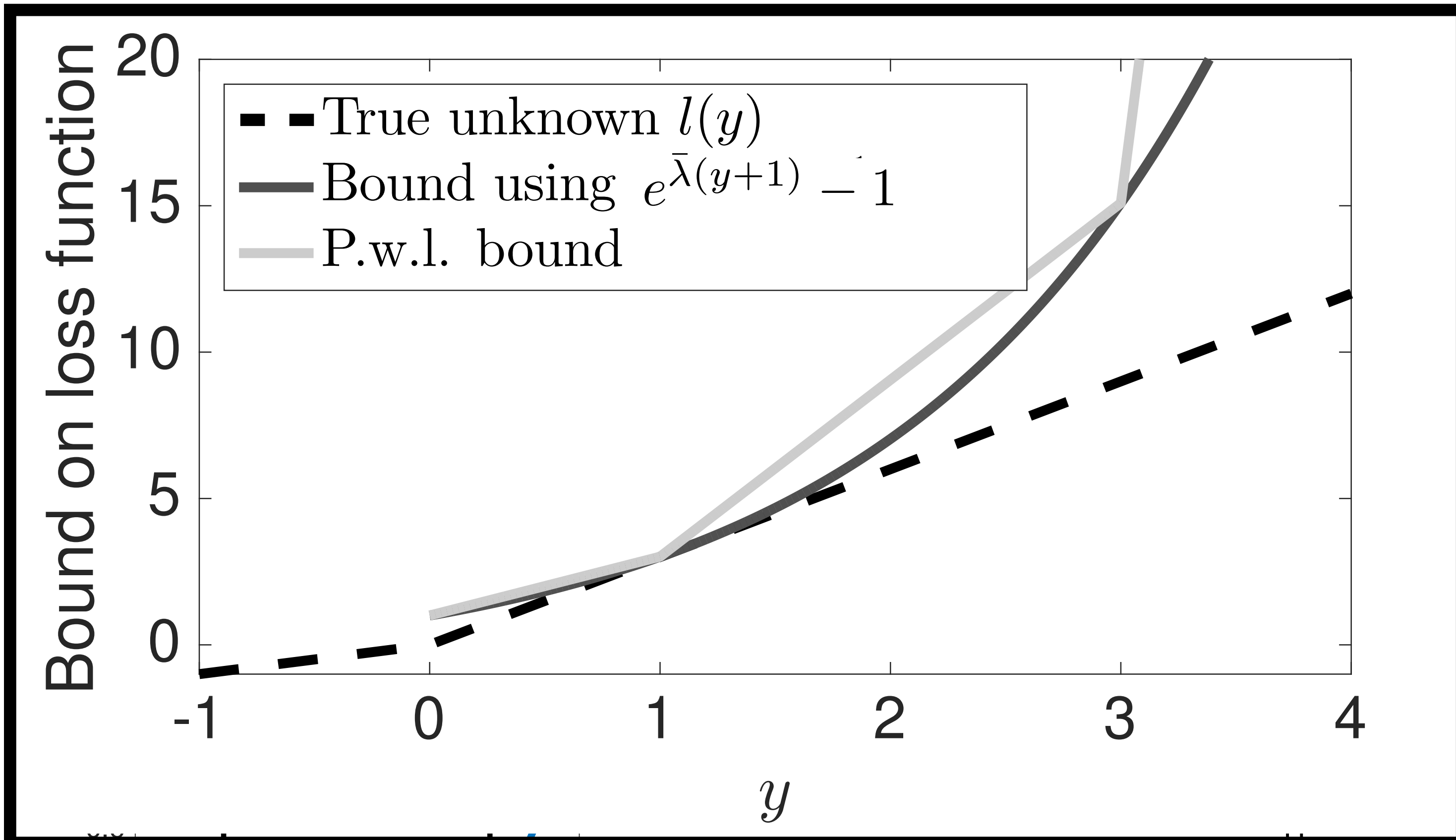
$$\rho(Y_{\bar{\lambda}}) \leq \rho(-1)$$

$$\Downarrow$$

$$l(y) \leq e^{\bar{\lambda}(y+1)} - 1, \ \ \forall y \geq 0$$

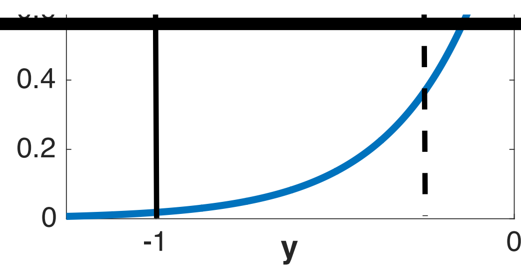
HEC MONTRÉAL

Assumptions needed for Theorem 3



- n
-
-
- er than

)



$\Updownarrow$

$$l(y) \leq e^{\bar{\lambda}(y+1)} - 1\,, \;\; \forall\, y \geq 0$$

Accounting for elicitation errors

- Issue #2: One can easily provide **false information** about his preferences. (Kahneman & Tversky, 1979)

- One can replace the comparison constraint with:

$$\exists \delta \in \mathbb{R}^K,\ \|\delta\|_1 \leq \Gamma,\ \rho(W_k) \leq \rho(Y_k) + \delta_k,\ \forall\, k = 1, \ldots, K$$

- Bertsimas and O'hair (2015) even propose accounting for some preference reversals with :

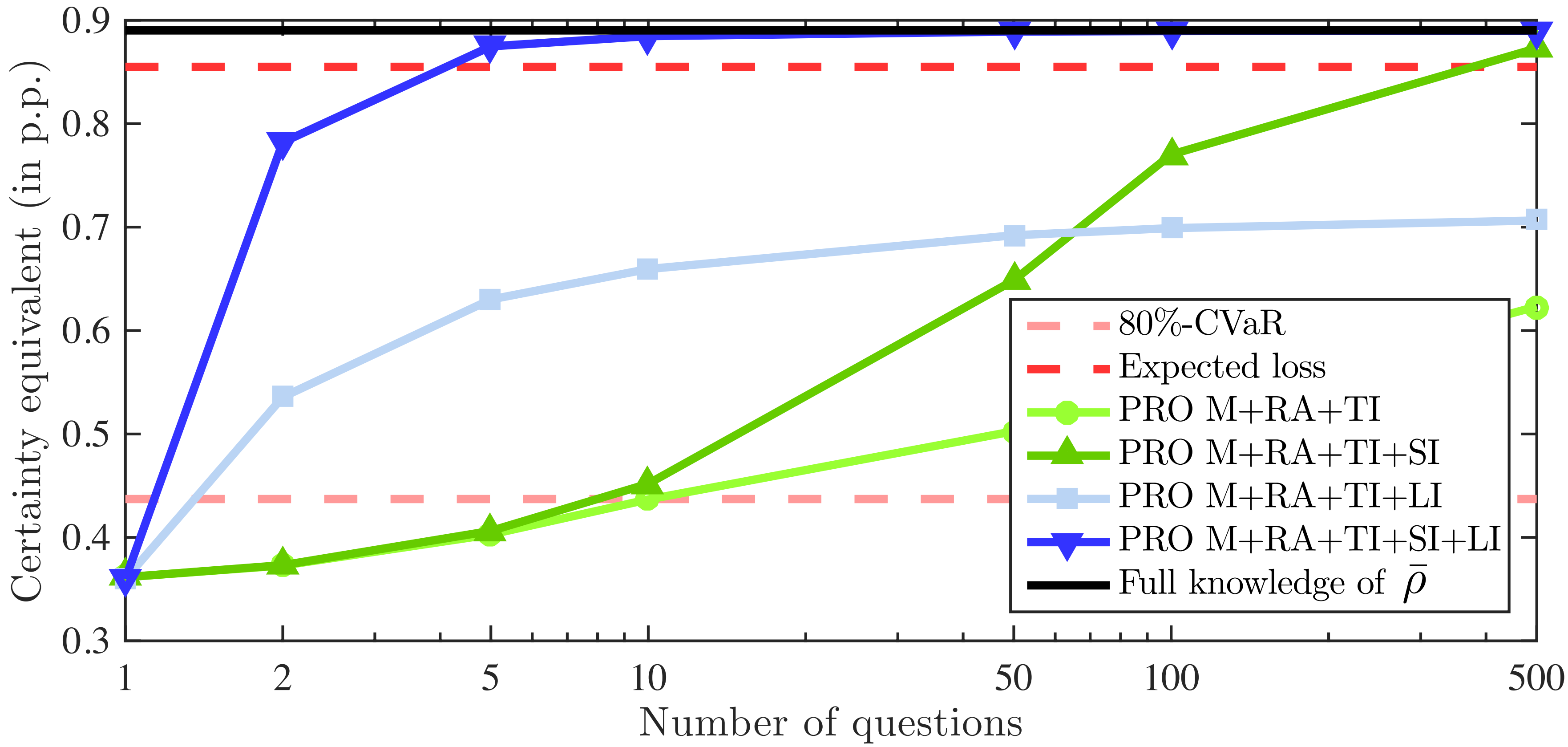
$$\exists z \in \{0,1\}^K,\ \sum z_k \leq \Gamma,\ \left\{ \begin{array}{c} \rho(W_k) \leq \rho(Y_k) + M z_k \\ \rho(Y_k) \leq \rho(W_k) + M(1 - z_k) \end{array} \right\},\ \forall\, k$$

Numerical experiments

Numerical experiments

- Experiments are made using empirical stochastic models based on historical weekly returns from Yahoo Finance
- We create a synthetic decision maker with some choice of $\bar{\rho}$ which is kept hidden
- Information comes from a number of certainty equivalents $\rho(W_k) = \rho(w_k)$ for randomly picked W_k
- Results are averaged over a large number of stochastic models and sets of W_k

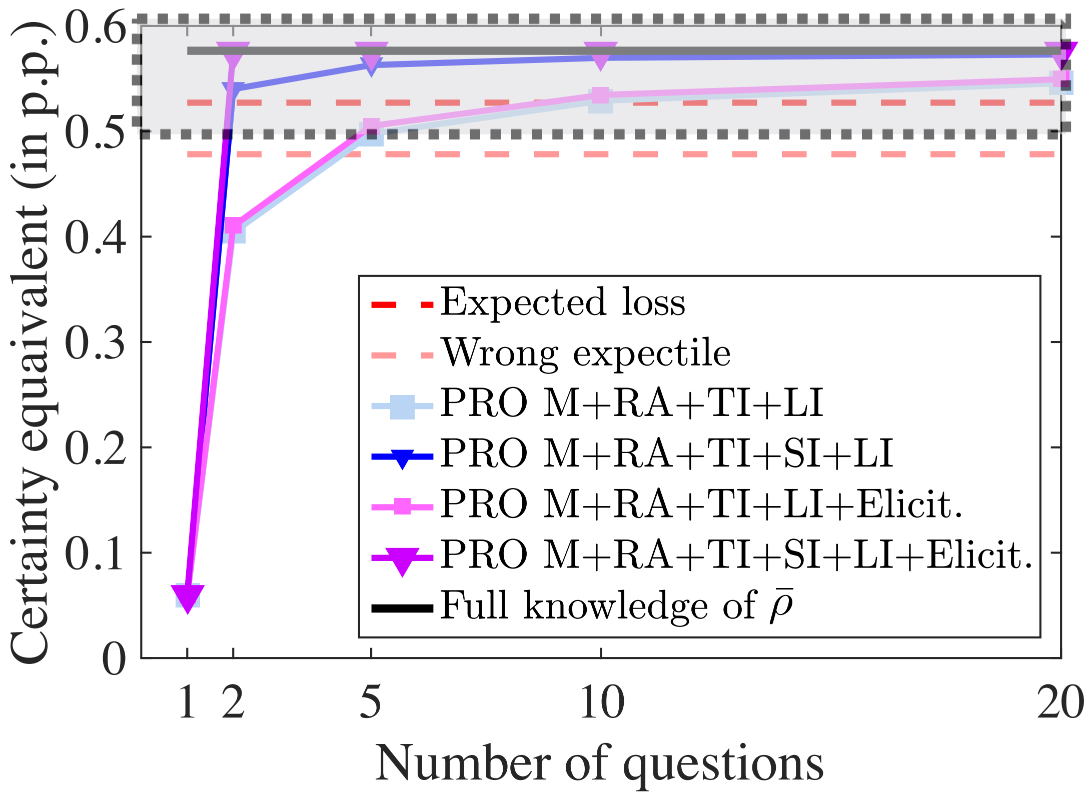
Performance in terms of certainty equivalent*



* Certainty equivalent $= -\bar{\rho}(Z(x))$

[D. & Li, 2017]

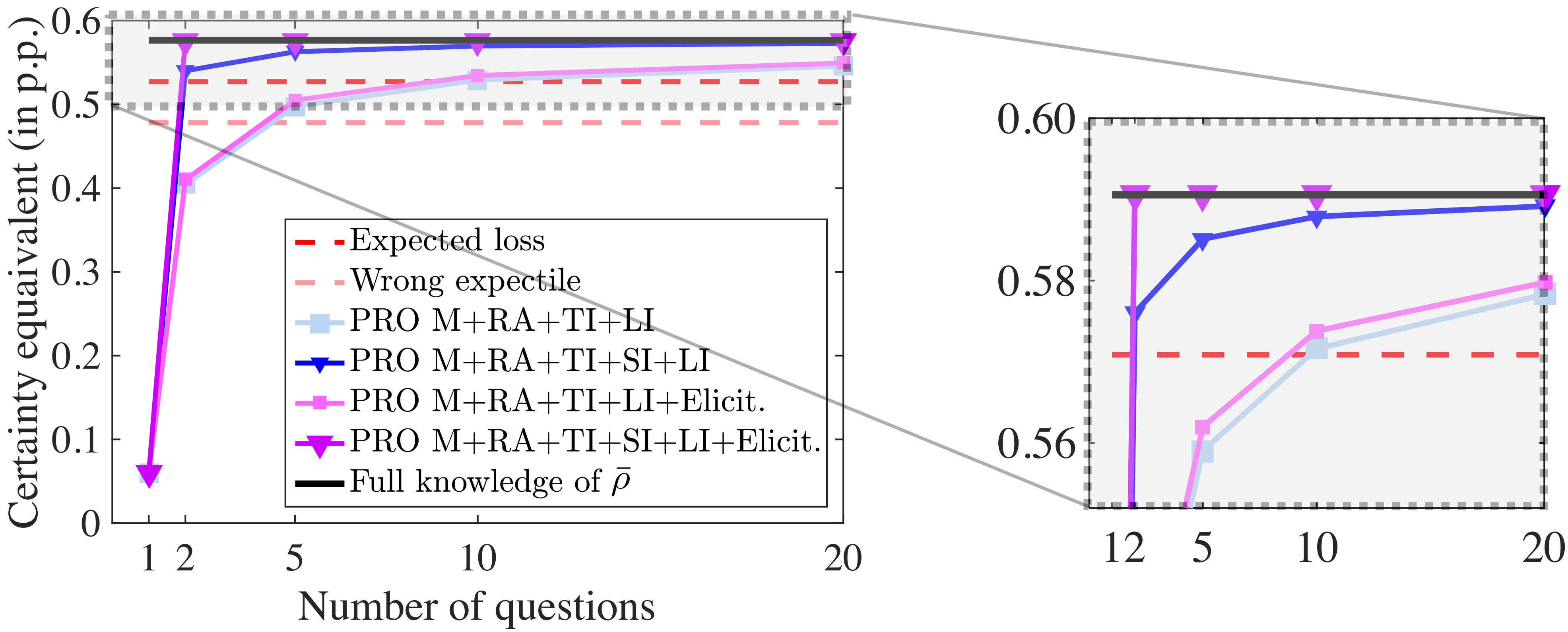
The case of UBSF risk measure



[D., Guo & Xu, 2018]

HEC MONTRÉAL

The case of UBSF risk measure

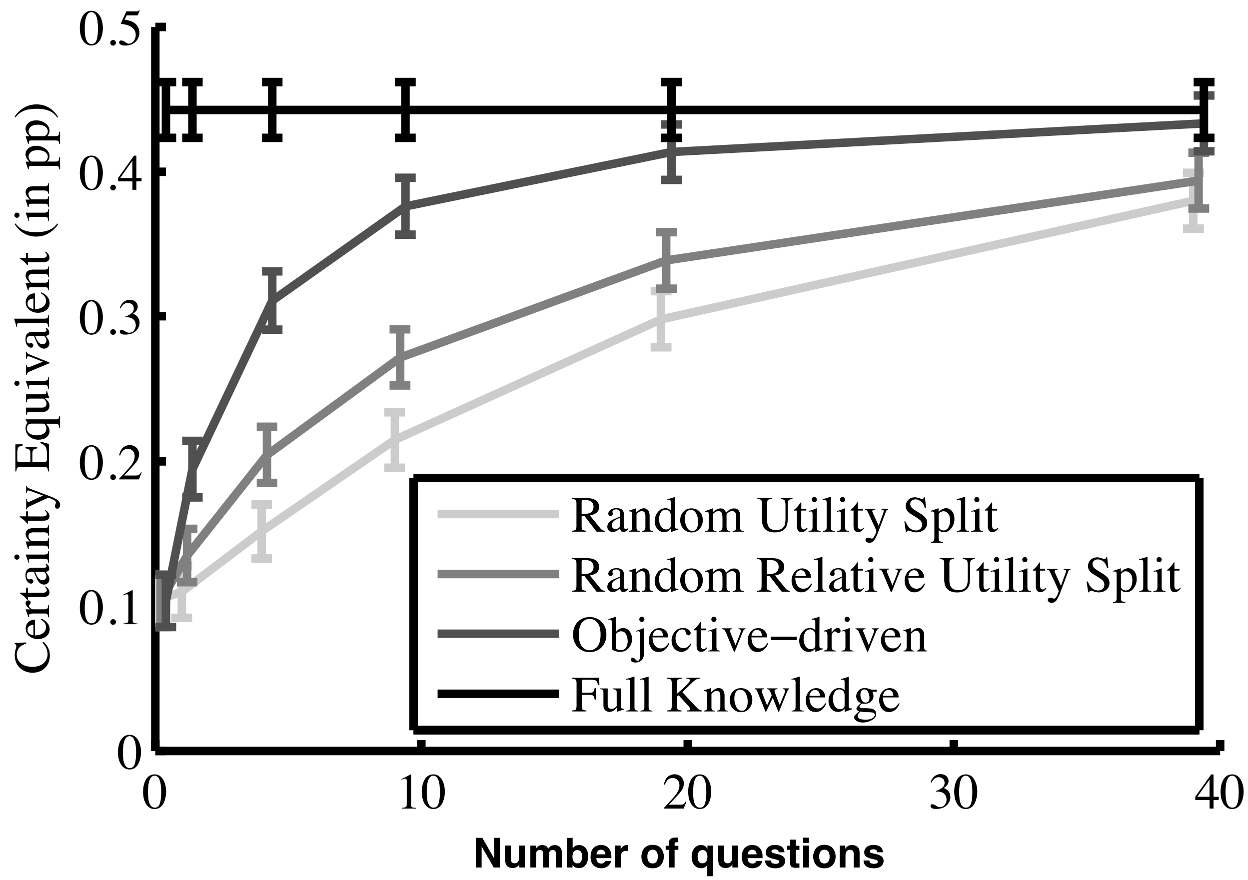


[D., Guo & Xu, 2018]

HEC MONTRÉAL

Effect of elicitation strategy

- One can improve convergence rate by designing effective elicitation strategies



Take-away messages

- Many optimization problems need to reflect the decision maker's risk preferences
- PRO accounts for the limited knowledge about these preferences :

  axioms + confidence intervals
- PRO preserves difficulty of resolution: LP —> LP
- For risk averse optimization, no LP representation for comonotone additivity, i.e. subjective risk measures
- While PRO is currently mostly developed for risk averse optimization, there is great potential for extensions to multi-criteria problems

Bibliography

- **Armbruster, Delage, Decision Making under Uncertainty when Preference Information is Incomplete, Management Science, 2015.**
- Bertsimas, O'hair. Learning Preferences Under Noise and Loss Aversion: An Optimization Approach, Operations Research, 2013.
- Bokrantz. Distributed approximation of Pareto surfaces in multi criteria radiation therapy treatment planning, Physics in Medicine and Biology, 2013
- Boutilier, Patrascu, Poupart, Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. Artificial Intelligence 2006.
- Chajewska, Koller, Parr. Making rational decisions using adaptive utility elicitation. AAAI 2000.
- Chan, Mahmoudzadeh, Purdie. A robust-cvar optimization approach with application to breast cancer therapy. EJOR, 2014.
- **Delage, Guo, Xu, Shortfall Risk Models When Information of Loss Function Is Incomplete, working draft.**
- **Delage, Li, Minimizing Risk Exposure when the Choice of a Risk Measure is Ambiguous. Management Science, 2018.**
- Evren, Ok. On the multi-utility representation of preference relations, Journal of Mathematical Economics, 2011.
- Haskell, Fu, Dessouky. Ambiguity in risk measures in robust stochastic optimization, European Journal of Operational Research, 2016.
- Haskell, Huang, Xu. Preference Elicitation and Robust Optimization with Multi-Attribute Quasi-Concave Choice Functions, 2018.
- Hu, Mehrotra. Robust decision making over a set of random targets or risk-averse utilities with an application to portfolio optimization, IIE Transactions, 2015.
- Hu, Mehrotra. Robust and Stochastically Weighted Multiobjective Optimization Models and Reformulations, Operations Research, 2012.
- Hu, Stepanyan. Reference-Based Almost Stochastic Dominance Rules with Application in Risk-Averse Optimization, 2018.

Thank you for your attention