Portfolio Selection with Objective Functions from Cumulative Prospect Theory

Thorsten Hens and János Mayer

13th International Conference on Stochastic Programming
Bergamo, Italy

July 10, 2013
Portfolio optimization in the Prospect Theory (PT) framework.

Probability distortion and Cumulative Prospect theory (CPT).

How to solve this type of problems numerically?

Algorithm based on adaptive simplicial grid refinement and its implementation.

Basic setup of the numerical experiments for comparing (C)PT and MV.

Data-sets used: basic data set, data set with an added European call option, data-set from a normal distribution.

Numerical results.

Conclusions.
Portfolio optimization in the Prospect Theory (PT) framework.
Probability distortion and Cumulative Prospect theory (CPT).
Portfolio optimization in the Prospect Theory (PT) framework.
Probability distortion and Cumulative Prospect theory (CPT).
How to solve this type of problems numerically?
Contents

- Portfolio optimization in the Prospect Theory (PT) framework.
- Probability distortion and Cumulative Prospect theory (CPT).
- How to solve this type of problems numerically?
- Algorithm based on adaptive simplicial grid refinement and its implementation.

Basic setup of the numerical experiments for comparing (C)PT and MV.

Data-sets used: basic data set, data set with an added European call option, data set from a normal distribution.

Numerical results.

Conclusions.
 Portfolio optimization in the Prospect Theory (PT) framework.
Probability distortion and Cumulative Prospect theory (CPT).
How to solve this type of problems numerically?
Algorithm based on adaptive simplicial grid refinement and its implementation.
Basic setup of the numerical experiments for comparing (C)PT and MV.
Portfolio optimization in the Prospect Theory (PT) framework.
Probability distortion and Cumulative Prospect Theory (CPT).
How to solve this type of problems numerically?
Algorithm based on adaptive simplicial grid refinement and its implementation.
Basic setup of the numerical experiments for comparing (C)PT and MV.
Data–sets used: basic data set, data set with an added European call option, data–set from a normal distribution.
Portfolio optimization in the Prospect Theory (PT) framework.

Probability distortion and Cumulative Prospect theory (CPT).

How to solve this type of problems numerically?

Algorithm based on adaptive simplicial grid refinement and its implementation.

Basic setup of the numerical experiments for comparing (C)PT and MV.

Data–sets used: basic data set, data set with an added European call option, data–set from a normal distribution.

Numerical results.
• Portfolio optimization in the Prospect Theory (PT) framework.
• Probability distortion and Cumulative Prospect theory (CPT).
• How to solve this type of problems numerically?
• Algorithm based on adaptive simplicial grid refinement and its implementation.
• Basic setup of the numerical experiments for comparing (C)PT and MV.
• Data–sets used: basic data set, data set with an added European call option, data–set from a normal distribution.
• Numerical results.
• Conclusions.
According to our knowledge the directly relevant literature is rather scarce. Papers related to the **numerical solution** of CPT optimization problems. Assumption: normally distributed resp. elliptically symmetric returns:


Explicit formulas for the two–assets case with one of them being the risk–free asset:

**The Kahneman–Tversky value function**

\[ v(x) = \begin{cases} 
(x - RP)^{\alpha^+} & \text{if } x \geq RP \\
-\beta(RP - x)^{\alpha^-} & \text{if } x < RP 
\end{cases} \]

- Kahneman and Tversky (1979):
  - risk aversion \( \alpha := \alpha^+ = \alpha^- = 0.88 \);
  - loss-aversion \( \beta = 2.25 \).

Probability distortion function:

\[ w(p; \gamma) = \frac{p^\gamma}{[p^\gamma + (1 - p)^\gamma]^{\frac{1}{\gamma}}}; \]

- \(0 < \gamma \leq 1\) (KT: \(\gamma = 0.65\));
- \(\gamma = 1 \Rightarrow \) no distortion.


Alternative weighting functions were proposed by Lattimore et al. (1992), Wu and Gonzalez (1996), Prelec (1998), Rieger and Wang (2004), ...
PT: Static (one–period) portfolio optimization

$$\max V_{PT}(\lambda) := \sum_{s=1}^{S} w_s \cdot v(x_s(\lambda)) \quad \lambda \in T$$

- $w_s := w(p_s; \gamma)$: distorted probabilities;
- $x_s$: prospects, realizations of portfolio return; $x_s(\lambda) := \sum_{k=1}^{K} R_s^k \lambda_k$;
- $T := \{\lambda \mid \sum_{k=1}^{K} \lambda_k = 1, \lambda_k \geq 0, \forall k\}$, a simplex.
In this case the distorted probabilities are computed as:

\[ w_s(\lambda) := \begin{cases} 
  w_s^+ := w(p_s, \gamma^+) & \text{if } x_s(\lambda) \geq RP, \\
  w_s^- := w(p_s, \gamma^-) & \text{if } x_s(\lambda) < RP.
\end{cases} \]

where now the distorted probabilities themselves also depend on \( \lambda \). This leads to

\[ V_{PT}(\lambda) = \sum_{s=1}^{S} w_s(\lambda) \cdot v(x_s(\lambda)). \]
In this case the distorted probabilities are computed as:

\[ w_s(\lambda) := \begin{cases} 
  w_s^+ := w(p_s, \gamma^+) & \text{if } x_s(\lambda) \geq RP, \\
  w_s^- := w(p_s, \gamma^-) & \text{if } x_s(\lambda) < RP.
\end{cases} \]

where now the distorted probabilities themselves also depend on \( \lambda \). This leads to

\[ V_{PT}(\lambda) = \sum_{s=1}^{S} w_s(\lambda) \cdot v(x_s(\lambda)). \]

**If** \( \gamma^+ \neq \gamma^- \) **then**

\[ p_s = \frac{1}{S}, \quad \forall s \implies \text{the probability weighting matters.} \]
Prospect Theory (PT):

- Small probabilities are overweighted;
- Large probabilities are underweighted,

regardless of the size of the loss or gain.
Prospect Theory (PT):

- Small probabilities are overweighted;
- large probabilities are underweighted, regardless of the size of the loss or gain.

Cumulative Prospect Theory (CPT):
only small probabilities corresponding to extreme losses or gains are overweighted.
**Rank–ordered lottery:** \(( (x_{1(\lambda)}(\lambda), p_{1(\lambda)}), \ldots, (x_{S(\lambda)}(\lambda), p_{S(\lambda)}) ) \), where \((1(\lambda), \ldots, S(\lambda))\) is a permutation of \((1, \ldots, S)\) such that

\[
x_{1(\lambda)}(\lambda) \leq \ldots \leq x_{r(\lambda)}(\lambda) \leq \text{RP} \leq x_{r+1(\lambda)}(\lambda) \leq \ldots \leq x_{S(\lambda)}(\lambda)
\]

holds. 

\((r(\lambda) = 0 \Rightarrow \text{gains in all states}; r(\lambda) = S(\lambda) \Rightarrow \text{losses in all states})\).

**Numerical aspect:** sorting.
We introduce the notation $\bar{x}_s(\lambda) := x_s(\lambda)(\lambda)$, $\bar{p}_s := p_s(\lambda)$ and $\bar{r} := r(\lambda)$. Computing the weights:

$$
\begin{align*}
w^-_1(\lambda) &:= w^-(\bar{p}_1), \\ w^-_s(\lambda) &:= w^-(\bar{p}_1 + \ldots + \bar{p}_s) - w^-(\bar{p}_1 + \ldots + \bar{p}_{s-1}), 1 < s < \bar{r}, \\ w^+_s(\lambda) &:= w^+(\bar{p}_s + \ldots + \bar{p}_S) - w^+(\bar{p}_{s+1} + \ldots + \bar{p}_S), \bar{r} \leq s < S, \\ w^+_S(\lambda) &:= w^+(\bar{p}_S).
\end{align*}
$$
We introduce the notation $\bar{x}_s(\lambda) := x_s(\lambda)(\lambda)$, $\bar{p}_s := p_s(\lambda)$ and $\bar{r} := r(\lambda)$. Computing the weights:

$$
\begin{align*}
  w_1^- (\lambda) &:= w^- (\bar{p}_1), \\
  w_s^- (\lambda) &:= w^- (\bar{p}_1 + \ldots + \bar{p}_s) - w^- (\bar{p}_1 + \ldots + \bar{p}_{s-1}), \quad 1 < s < \bar{r}, \\
  w_s^+ (\lambda) &:= w^+ (\bar{p}_s + \ldots + \bar{p}_S) - w^+ (\bar{p}_{s+1} + \ldots + \bar{p}_S), \quad \bar{r} \leq s < S, \\
  w_S^+ (\lambda) &:= w^+ (\bar{p}_S).
\end{align*}
$$

\begin{itemize}
  \item Both the ordering of the scenarios and $\bar{r}$ depend on $\lambda$;
\end{itemize}
We introduce the notation $\bar{x}_s(\lambda) := x_s(\lambda)(\lambda)$, $\bar{p}_s := p_s(\lambda)$ and $\bar{r} := r(\lambda)$. Computing the weights:

\[
\begin{align*}
  w_1^-(\lambda) & := w^-(\bar{p}_1), \\
  w_s^-(\lambda) & := w^-(\bar{p}_1 + \ldots + \bar{p}_s) - w^-(\bar{p}_1 + \ldots + \bar{p}_{s-1}), 1 < s < \bar{r}, \\
  w_s^+(\lambda) & := w^+(\bar{p}_s + \ldots + \bar{p}_S) - w^+(\bar{p}_{s+1} + \ldots + \bar{p}_S), \bar{r} \leq s < S, \\
  w_S^+(\lambda) & := w^+(\bar{p}_S).
\end{align*}
\]

★ Both the ordering of the scenarios and $\bar{r}$ depend on $\lambda$;  
★ $\bar{p}_1 + \ldots + \bar{p}_s = F_\lambda(\bar{x}_s(\lambda))$;  
★ $\bar{p}_{s+1} + \ldots + \bar{p}_S = 1 - F_\lambda(\bar{x}_s(\lambda))$.  

The objective function in CPT is $V_{\text{CPT}}(\lambda) := \sum_{s=1}^{\bar{r}} w_s^-(\lambda) \cdot v(\bar{x}_s(\lambda)) + \sum_{s=\bar{r}+1}^S w_s^+(\lambda) \cdot v(\bar{x}_s(\lambda))$.  

Computing the weights: $w_1^-(\lambda) := w^-(\bar{p}_1)$,
We introduce the notation $\bar{x}_s(\lambda) := x_s(\lambda)(\lambda)$, $\bar{p}_s := p_s(\lambda)$ and $\bar{r} := r(\lambda)$. Computing the weights:

\[
\begin{align*}
\bar{w}_1^-(\lambda) & := w^-(\bar{p}_1), \\
\bar{w}_s^-(&\lambda) := w^-(\bar{p}_1 + \ldots + \bar{p}_s) - w^-(\bar{p}_1 + \ldots + \bar{p}_{s-1}), \ 1 < s < \bar{r}, \\
\bar{w}_s^+(&\lambda) := w^+(\bar{p}_s + \ldots + \bar{p}_S) - w^+(\bar{p}_{s+1} + \ldots + \bar{p}_S), \ \bar{r} \leq s < S, \\
\bar{w}_S^+(&\lambda) := w^+(\bar{p}_S).
\end{align*}
\]

확률의 순서와 $\bar{r}$는 $\lambda$에 의존합니다;
확률 $\bar{p}_1 + \ldots + \bar{p}_s = F_\lambda(\bar{x}_s(\lambda))$;
확률 $\bar{p}_{s+1} + \ldots + \bar{p}_S = 1 - F_\lambda(\bar{x}_s(\lambda))$.

The objective function in CPT is

\[
V_{CPT}(\lambda) := \sum_{s=1}^{r} w_s^-(\lambda) \cdot v(\bar{x}_s(\lambda)) + \sum_{s=r+1}^{S} w_s^+(\lambda) \cdot v(\bar{x}_s(\lambda)).
\]
PT versus CPT: First impressions (I)

An artificial example with 2 assets and 3 scenarios

<table>
<thead>
<tr>
<th>probability</th>
<th>Asset1</th>
<th>Asset2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>-0.035</td>
<td>-0.01</td>
</tr>
<tr>
<td>0.2</td>
<td>0.03</td>
<td>-0.07</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.025</td>
<td>0.01</td>
</tr>
</tbody>
</table>
PT versus CPT: First impressions (II)

A further artificial example with 2 assets and 3 scenarios

<table>
<thead>
<tr>
<th>probability</th>
<th>Asset1</th>
<th>Asset2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.035</td>
<td>0.005</td>
</tr>
<tr>
<td>0.2</td>
<td>-0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>0.3</td>
<td>0.025</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Parameter settings: $RP = 0$, $\alpha^+ = \alpha^- = 0.88$, $\beta = 2.25$, $\gamma = 0.65$. 
Numerical aspects

The main sources of numerical difficulties:

- Both the PT and the CPT portfolio optimization models are non-convex optimization problems: the objective function is not concave, in general.
- Since the KT value function is not differentiable at the RP, both $V_{PT}$ and $V_{CPT}$ are non-smooth.
- Is the non-smoothness of the KT value function the sole source of the non-smoothness of $V_{CPT}$? Most probably not, because sorting introduces abrupt changes in $w_s(\lambda)$ when $\lambda$ crosses a point where the ordering changes. Is $V_{CPT}$ continuous at all? The answer is probably yes. These issues will be further explored experimentally and theoretically.

To get first ideas on the behavior of $V_{CPT}$:

Why not begin with the beginning? Let us try grid-search approaches for CPT-optimization first! We employ simplicial grids.
Numerical aspects

The main sources of numerical difficulties:

- Both the PT and the CPT portfolio optimization models are non-convex optimization problems: the objective function is not concave, in general.
- Since the KT value function is not differentiable at the RP, both $V_{PT}$ and $V_{CPT}$ are non-smooth.
- Is the non-smoothness of the KT value function the sole source of the non-smoothness of $V_{CPT}$? Most probably not, because sorting introduces abrupt changes in $w_s(\lambda)$ when $\lambda$ crosses a point where the ordering changes. Is $V_{CPT}$ continuous at all? The answer is probably yes. These issues will be further explored experimentally and theoretically.

To get first ideas on the behavior of $V_{CPT}$:

Why not begin with the beginning? Let us try grid-search approaches for CPT-optimization first! We employ simplicial grids.
Simplices

n–dimensional simplex in $\mathbb{R}^N$: the convex hull of $n + 1$ affinely independent points in $\mathbb{R}^N$.

<table>
<thead>
<tr>
<th>dimension of simplex:</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>#vertices:</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>#edges:</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>15</td>
<td>21</td>
<td>28</td>
<td>36</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td>#2-dim facets:</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>20</td>
<td>35</td>
<td>56</td>
<td>84</td>
<td>120</td>
<td>165</td>
</tr>
</tbody>
</table>
Simplices: barycentric coordinates

An \((n - 1)\)-dimensional simplex in \(\mathbb{R}^n\):

\[
T = \{ x \mid x = \sum_{k=1}^{n} \mu_k x^{(k)}, \quad \sum_{k=1}^{n} \mu_k = 1, \quad \mu_k \geq 0, \quad \forall k \}\.
\]

The \((n - 1)\)-dimensional unit simplex in \(\mathbb{R}^n\):

\[
U = \{ x \mid x = \sum_{k=1}^{n} \mu_k e^{(k)}, \quad \sum_{k=1}^{n} \mu_k = 1, \quad \mu_k \geq 0, \quad \forall k \}\ = \{ x \mid \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0, \quad \forall j \}.
\]

- \(x^{(1)}, \ldots, x^{(n)}\) are the vertices of the simplex \(T\);
- \(\mu_1, \ldots, \mu_n\) are the (uniquely determined) barycentric coordinates of \(x \in T\).
- A simplicial partition of the unit simplex corresponds to a simplicial subdivision of \(T\) and vice versa.
A regular simplex is a simplex with all edges having the same length.

A regular simplicial grid or mesh with grid constant $gc \in \mathbb{N}_{++}$ results by partitioning the unit simplex $U$ into congruent regular subsimplices, such that the edges of $U$ become partitioned into $gc$ equal pieces.
Adaptive simplicial grid method

- Primary objects: grid–points in the current mesh.
- Compute the objective function value for each of the grid–points in the current mesh as well as at the centroids of subsimplices. Choose one point $p$ with the highest objective value.
- Look for a subsimplex in a coarser grid (e.g., $g_{c1} < \frac{g_{c2}}{2}$) which contains $p$. Refine the grid in the subsimplex found.
- Repeat the procedure with this subsimplex.
- This procedure is embedded into an outer loop for finding starting points via generating uniformly distributed points over the starting simplex.

![Diagram showing adaptive simplicial grid method](image)

barycentric coordinates $\mapsto$ spatial coordinates  

Authors:

Implementation

Generating a grid:

*Regular grid (mesh)* over the unit–simplex $U$, corresponding to the grid–constant $gc \in \mathbb{N}_+$:

$$T_{gc} := \left\{ x = \frac{1}{gc}(k_0, k_1, \ldots, k_n)^T \mid \sum_{i=0}^{n} k_i = gc, \; k_i \in \mathbb{N}_+, \; \forall i \right\}$$

The grid is computed via *compositions* $(n + 1)$–tuples $(k_0, k_1, \ldots, k_n)$ with $\sum_{i=0}^{n} k_i = gc, \; gc, \; k_i \in \mathbb{N}_+, \; \forall i$. For setting up a list of all compositions, a recursive algorithm of D. Knuth has been employed.

- Primary data structure:
  - A list of all *compositions* for $gc$;
  - the spatial coordinates of the current simplex.

  Subsimplices are constructed only according to the needs of the algorithm.

Crucial ingredient of the algorithm:

For a given point $p$ in the current simplex, the problem is to find the vertices of a subsimplex of the partition, which contains $p$. For this an algorithm of Kuhn and Eaves has been implemented.
Comparing (C)PT and MV

With $V = V_{PT}$ and $V = V_{CPT}$ we compare the solutions of:

$$\max_\lambda V(\xi^T\lambda) \quad \begin{cases} \mathbb{1}^T\lambda = 1 \\ \lambda \geq 0, \end{cases}$$

versus

$$\max_\lambda V(\xi^T\lambda) \quad \lambda \in \Lambda_{MV}^*.$$

Optimal solution: $\lambda^*$.

$\Lambda_{MV}^* \subset \{\lambda \mid \mathbb{1}^T\lambda = 1, \lambda \geq 0\} \Rightarrow V(\xi^T\lambda_{MV}^*) \leq V(\xi^T\lambda^*)$.

Levy and Levy (2004): if $\xi$ has a multivariate normal distribution and short-sales are allowed, then the solutions of the above optimization problems coincide.

Central question in our experiments: Is this so in general? Is it sufficient to optimize $V$ along the MV frontier if we do not have a normal distribution and/or $\lambda \geq 0$ is imposed?
Comparing (C)PT and MV

With $V = V_{PT}$ and $V = V_{CPT}$ we compare the solutions of:

$$\max_{\lambda} \quad V(\xi^T \lambda)$$

$$\begin{cases} 
\mathbb{1}^T \lambda &= 1 \\
\lambda &\geq 0,
\end{cases}$$

versus

$$\begin{cases} 
\max_{\lambda} \quad V(\xi^T \lambda) \\
\lambda &\in \Lambda^*_{MV}
\end{cases}$$

Optimal solution: $\lambda^*$. 

$\Lambda^*_{MV}$: the set of portfolios along the MV efficient frontier.

$$\Lambda^*_{MV} \subset \{ \lambda \mid \mathbb{1}^T \lambda = 1, \lambda \geq 0 \} \Rightarrow V(\xi^T \lambda^*_{MV}) \leq V(\xi^T \lambda^*)$$

Levy and Levy (2004): if $\xi$ has a multivariate normal distribution and short–sales are allowed, then the solutions of the above optimization problems coincide.
Comparing (C)PT and MV

With $V = V_{PT}$ and $V = V_{CPT}$ we compare the solutions of:

$$\max_{\lambda} \quad V(\xi^T \lambda)$$

$$\begin{align*}
\mathbb{1}^T \lambda &= 1 \\
\lambda &\geq 0,
\end{align*}$$

versus

$$\begin{align*}
\max_{\lambda} \quad V(\xi^T \lambda) \\
\lambda &\in \Lambda^*_{MV}
\end{align*}$$

Optimal solution: $\lambda^*$.

$\Lambda^*_{MV}$: the set of portfolios along the MV efficient frontier.

Levy and Levy (2004): if $\xi$ has a multivariate normal distribution and short-sales are allowed, then the solutions of the above optimization problems coincide.

$$\Lambda^*_{MV} \subset \{\lambda \mid \mathbb{1}^T \lambda = 1, \lambda \geq 0\} \quad \Rightarrow \quad V(\xi^T \lambda^*_{MV}) \leq V(\xi^T \lambda^*).$$

Central question in our experiments:

Is this so in general? Is it sufficient to optimize $V$ along the MV frontier if we do not have a normal distribution and/or $\lambda \geq 0$ is imposed?
How to compare the two types of optimal portfolios?

Proximity measures:

- Kroll, Levy and Markowitz (1984)
  \[ I_{OBJR} = \frac{V(\xi^T \lambda^*_MV) - V(\xi^T \lambda_{naive})}{V(\xi^T \lambda^*) - V(\xi^T \lambda_{naive})} \]
  where \( \lambda_{naive} \) is the “naïve” portfolio with equal weights.

- Kallberg and Ziemba (1983)
  \[ I_{CER} = \frac{CE_v[\xi^T \lambda^*] - CE_v[\xi^T \lambda^*_MV]}{CE_v[\xi^T \lambda^*]} \]
  with the certainty equivalent defined as \( CE_v[\xi^T \lambda] = v^{-1} \left( V(\xi^T \lambda) \right) \).

- DeMiguel, Garlappi and Uppal (2009), De Giorgi and Hens (2009)
  \[ I_{CED} := CE_v[\xi^T \lambda^*] - CE_v[\xi^T \lambda^*_MV] \]
  Interpretation: the difference as added value in monetary terms.

In our computational results all three quantities are expressed in % terms with annualized values for \( I_{CED} \).
Basic monthly–returns data set. Monthly data for 8 indices, 1994 February – 2011 May (208 months); we would like to express our thanks to Dieter Niggeler of BhFS for providing us with this data–set.

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSCITR</td>
<td>Goldman Sachs Commodity Index; total return;</td>
</tr>
<tr>
<td>I3M</td>
<td>3 months US Dollar LIBOR interest rate;</td>
</tr>
<tr>
<td>HFRIFFM</td>
<td>Hedge Fund Research International, Fund of Funds, market defensive index;</td>
</tr>
<tr>
<td>MSEM</td>
<td>Morgan Stanley Emerging Markets Index, total return, stocks;</td>
</tr>
<tr>
<td>MXWO</td>
<td>MSCI World Index, total return, stocks</td>
</tr>
<tr>
<td></td>
<td>(developed countries);</td>
</tr>
<tr>
<td>NAREIT</td>
<td>FTSE, US Real Estate Index, total return;</td>
</tr>
<tr>
<td>PE</td>
<td>LPX50, LPX Group Zurich, Listed Private Equities, total return;</td>
</tr>
<tr>
<td>JPMBSD</td>
<td>JP Morgan Bond Index, Developed Markets, total return.</td>
</tr>
</tbody>
</table>

Sample from the corresponding normal distribution. With the empirical $\mu$ and $\Sigma$ from the basic data–set we have generated a multivariate normal sample with sample–size 10’000.

Testing for multivariate normality

\( H_0 \): Multivariate normality holds for the basic data–set. Results from Mardia’s tests involving multivariate skewness and kurtosis:

<table>
<thead>
<tr>
<th></th>
<th>Original Data–Set N=208</th>
<th>Normal sample N=208</th>
<th>Normal sample N=10’000</th>
</tr>
</thead>
<tbody>
<tr>
<td>mv. skewness</td>
<td>15.88</td>
<td>3.08</td>
<td>0.06</td>
</tr>
<tr>
<td>test stat. ( \chi^2(120) )</td>
<td>550.62</td>
<td>106.65</td>
<td>104.14</td>
</tr>
<tr>
<td>( p )-value</td>
<td>0</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td>mv. kurtosis</td>
<td>113.12</td>
<td>77.57</td>
<td>79.74</td>
</tr>
<tr>
<td>test stat. ( \mathcal{N}(0,1) )</td>
<td>18.88</td>
<td>-1.38</td>
<td>-1.02</td>
</tr>
<tr>
<td>( p )-value</td>
<td>0</td>
<td>0.17</td>
<td>0.31</td>
</tr>
</tbody>
</table>

\( p = 0 \) for both cases \( \Rightarrow H_0 \) **can be rejected**, e.g., on a 99.9% significance level.
We wish to investigate also the influence of probability weighting \( \implies \) empirical distributions needed. These we have generated via \( k \)-means clustering of the basic data-sets. \( k = \text{???} \)

**Additional goal:** For achieving a clear deviation from multivariate normality we wish to add a European call option to the empirical distribution, having positive payoff in only one state. We solve the LP:

\[
\begin{align*}
\max_{\pi, \varepsilon} & \quad \varepsilon \\
\text{s.t.} & \quad \sum_{s=1}^{S} \pi_s = 1 \\
& \quad \sum_{s=1}^{S} r_s^k \pi_s = r_f, \quad k = 1, \ldots, K \\
& \quad \pi_s \geq \varepsilon, \quad s = 1, \ldots, S.
\end{align*}
\]

computing state prices with \( r_f = 0.2\% \)

Optimal solution: \( (\pi^*, \varepsilon^*) \). If \( \varepsilon^* > 0 \) holds \( \implies \) we add a call option on the index MXWO such that \( \sum_{s=1}^{S} \hat{r}_s \pi_s^* = r_f \) holds for the added column \( \hat{r} \implies \) the new data-set is also arbitrage-free.

**Trial-and-error** with increasing \( k \) for getting scenarios with \( \varepsilon^* > 0 \) in the LP above \( \implies \) \( k = 15 \) works; the call option can be appended.
Numerical experiments

We have three data sets, each of them consisting of 15 scenarios and corresponding probabilities:

- **Basic scenario data-set.** This is obtained from the basic monthly returns data-set via $k$-means clustering.
- **Scenario data-set with an added option.** Computed on the basis of the previous scenarios by adding a European call option on the index MXWO; having a positive payoff in a single state.
- **Scenario data-set from a normal distribution.** Generated via $k$-means clustering from the normal sample.

For each of these data-sets and for each one of the 48 investors

- we have computed the PT and CPT optimal portfolios as well as their counterparts by optimizing along the MV-frontier.
- For the comparisons the indices $I_{OBJR}$, $I_{CED}$ and $I_{CER}$ have been computed.
- In addition, for the PT and CPT optimal portfolios their geometrical distance to the MV-frontier has been computed.
CPT versus MV (I)

**basic data set**

**call option added**

**normal distribution**

**Observations.** \((\mu, \sigma)\) space: in the first two cases most of the points corresponding to CPT optimal portfolios are quite away from the MV-frontier, whereas in the third case the points are on the frontier. Frequency histograms for the \(I_{CED}\) index: there is a substantial deviation from 0 in the first two cases, contrary to the third case.
CPT versus MV (II)

<table>
<thead>
<tr>
<th></th>
<th>(I_{OBJR})</th>
<th>(I_{CED})</th>
<th>(I_{CER})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-18.46</td>
<td>0.0028</td>
<td>0.0002</td>
</tr>
<tr>
<td>Mean</td>
<td>76.89</td>
<td>2.1838</td>
<td>0.171</td>
</tr>
<tr>
<td>Median</td>
<td>97.07</td>
<td>0.1734</td>
<td>0.014</td>
</tr>
<tr>
<td>Max.</td>
<td>99.43</td>
<td>67.87</td>
<td>5.151</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>38.15</td>
<td>9.877</td>
<td>0.751</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.544</td>
<td>6.265</td>
<td>6.207</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.604</td>
<td>41.61</td>
<td>41.06</td>
</tr>
</tbody>
</table>

Basic data–set

<table>
<thead>
<tr>
<th></th>
<th>(I_{OBJR})</th>
<th>(I_{CED})</th>
<th>(I_{CER})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-21.99</td>
<td>0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>Mean</td>
<td>58.48</td>
<td>13.04</td>
<td>1.047</td>
</tr>
<tr>
<td>Median</td>
<td>73.69</td>
<td>5.238</td>
<td>0.435</td>
</tr>
<tr>
<td>Max.</td>
<td>97.70</td>
<td>83.72</td>
<td>6.525</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>36.40</td>
<td>19.26</td>
<td>1.508</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.906</td>
<td>2.184</td>
<td>2.104</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.396</td>
<td>7.325</td>
<td>6.987</td>
</tr>
</tbody>
</table>

Call option added

<table>
<thead>
<tr>
<th></th>
<th>(I_{OBJR})</th>
<th>(I_{CED})</th>
<th>(I_{CER})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>99.66</td>
<td>-0.001</td>
<td>-0.00009</td>
</tr>
<tr>
<td>Mean</td>
<td>99.99</td>
<td>0.0002</td>
<td>0.00001</td>
</tr>
<tr>
<td>Median</td>
<td>100.00</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Max.</td>
<td>100.00</td>
<td>0.006</td>
<td>0.0005</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.049</td>
<td>0.001</td>
<td>0.00008</td>
</tr>
<tr>
<td>Skewness</td>
<td>-6.616</td>
<td>4.590</td>
<td>4.594</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>44.85</td>
<td>28.16</td>
<td>28.19</td>
</tr>
</tbody>
</table>

Normal approximation

\(I_{CED}\) across the three data-sets
CPT versus MV (III): Optimal portfolios

Call option: Maximizing CPT; heavy investments into the call option

Call option: Maximizing CPT along the MV–frontier; practically no investment into the call
CPT versus MV (IV): Skewness of optimal portfolios

Original data-set: skewness of the optimal portfolios

Call option added: skewness of the optimal portfolios (cf. Barberis and Huang (2008))
Observations (similar as in the CPT case). $(\mu, \sigma)$ space: in the first two cases most of the points corresponding to CPT optimal portfolios are quite away from the MV–frontier, whereas in the third case the points are on the frontier. Frequency histograms for the $I_{CED}$ index: there is a substantial deviation from 0 in the first two cases, contrary to the third case.
### PT versus MV (II)

#### Basic data–set

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-13.80</td>
<td>0.002</td>
<td>0.0002</td>
</tr>
<tr>
<td>Mean</td>
<td>81.24</td>
<td>3.536</td>
<td>0.273</td>
</tr>
<tr>
<td>Median</td>
<td>98.02</td>
<td>0.429</td>
<td>0.036</td>
</tr>
<tr>
<td>Max.</td>
<td>99.88</td>
<td>83.85</td>
<td>6.206</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>34.49</td>
<td>13.30</td>
<td>0.994</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.806</td>
<td>5.158</td>
<td>5.057</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.539</td>
<td>29.67</td>
<td>28.61</td>
</tr>
</tbody>
</table>

#### Call option added

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-0.917</td>
<td>0.002</td>
<td>0.0002</td>
</tr>
<tr>
<td>Mean</td>
<td>75.74</td>
<td>13.75</td>
<td>1.096</td>
</tr>
<tr>
<td>Median</td>
<td>88.58</td>
<td>5.215</td>
<td>0.433</td>
</tr>
<tr>
<td>Max.</td>
<td>99.87</td>
<td>83.85</td>
<td>6.532</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>29.50</td>
<td>20.79</td>
<td>1.610</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.492</td>
<td>2.135</td>
<td>2.033</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.681</td>
<td>7.024</td>
<td>6.559</td>
</tr>
</tbody>
</table>

#### Normal approximation

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>99.98</td>
<td>-0.009</td>
<td>-0.0007</td>
</tr>
<tr>
<td>Mean</td>
<td>99.99</td>
<td>0.00002</td>
<td>0.000002</td>
</tr>
<tr>
<td>Median</td>
<td>100.0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Max.</td>
<td>100.0</td>
<td>0.005</td>
<td>0.0004</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.006</td>
<td>0.002</td>
<td>0.0001</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.379</td>
<td>-1.895</td>
<td>-1.894</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>17.13</td>
<td>11.73</td>
<td>11.72</td>
</tr>
</tbody>
</table>

$I_{CED}$ across the three data-sets
**PT versus CPT (I)**

**basic data set**

**call option added**

**normal distribution**

**Observations.** Frequency histograms for the $I_{CED}$ index: the variation across the three cases is much smaller now as in the previous CPT–MV and PT–MV comparisons. Interestingly, the difference between PT and CPT is noticeably smaller than the difference with respect to MV for both PT and CPT.
## PT versus CPT (II)

### Basic data-set

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-92.65</td>
<td>-0.356</td>
<td>-0.030</td>
</tr>
<tr>
<td>Mean</td>
<td>89.50</td>
<td>0.562</td>
<td>0.046</td>
</tr>
<tr>
<td>Median</td>
<td>99.25</td>
<td>0.089</td>
<td>0.007</td>
</tr>
<tr>
<td>Max.</td>
<td>100.9</td>
<td>11.13</td>
<td>0.921</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>32.23</td>
<td>1.695</td>
<td>0.140</td>
</tr>
<tr>
<td>Skewness</td>
<td>-4.349</td>
<td>5.249</td>
<td>5.254</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>22.99</td>
<td>32.24</td>
<td>32.30</td>
</tr>
</tbody>
</table>

### Call option added

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-56.98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>88.32</td>
<td>0.501</td>
<td>0.041</td>
</tr>
<tr>
<td>Median</td>
<td>99.42</td>
<td>0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>Max.</td>
<td>100</td>
<td>2.853</td>
<td>0.235</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>28.53</td>
<td>0.780</td>
<td>0.064</td>
</tr>
<tr>
<td>Skewness</td>
<td>-3.507</td>
<td>1.621</td>
<td>1.610</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>16.12</td>
<td>4.598</td>
<td>4.551</td>
</tr>
</tbody>
</table>

### Normal approximation

<table>
<thead>
<tr>
<th></th>
<th>$I_{OBJR}$</th>
<th>$I_{CED}$</th>
<th>$I_{CER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>87.07</td>
<td>-0.015</td>
<td>-0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>98.75</td>
<td>0.160</td>
<td>0.013</td>
</tr>
<tr>
<td>Median</td>
<td>99.99</td>
<td>0.0004</td>
<td>0.00004</td>
</tr>
<tr>
<td>Max.</td>
<td>100.0</td>
<td>1.122</td>
<td>0.091</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.934</td>
<td>0.288</td>
<td>0.024</td>
</tr>
<tr>
<td>Skewness</td>
<td>-2.595</td>
<td>2.134</td>
<td>2.11</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>8.813</td>
<td>6.625</td>
<td>6.510</td>
</tr>
</tbody>
</table>

### $I_{CED}$ across the three data-sets
Conclusions (I)

- To the best of our knowledge we are the first to propose and test a general algorithm for computing asset allocations for CPT. This is a numerically hard problem that is of high relevance for finance.
- We compare the results of maximizing CPT along the mean–variance efficient frontier and maximizing it without that restriction. We find:
  - With normally distributed returns the difference is negligible, in accordance with Levy and Levy (2004).
  - Using standard asset allocation data of pension funds the difference is considerable.
  - With derivatives like call options the restriction to the mean-variance efficient frontier results in a sizable loss; the results from the original data-set are magnified.
- Our computational results indicate that assumption of normally distributed returns is an essential presupposition for the application of the method proposed by Levy and Levy (2004).
Our numerical results fully support the theoretical findings of Barberis and Huang (2008) concerning positive skewness preference of (C)PT investors. This is despite the fact that in our case none of the assumptions in that paper hold.

Can the observed phenomena in fact be mainly attributed to deviations from normality? Recall that we exclude short-sales in our computations! Most probably yes, since in the normally distributed case we still have that constraint, nevertheless the optimal (C)PT portfolios are located along the MV frontier!

Considering the relation between optimal PT and CPT portfolios we observed that the difference between PT and CPT is noticeably smaller than the difference with respect to MV for both PT and CPT. We presume that this is due to the unimodal nature of the distribution of our basic data-set, but verifying this requires further investigations.
Conclusions (III)

Limitations

- For normally distributed returns, maximizing a (C)PT objective function along the MV frontier produces more accurate optimal solutions (cf. negative $I_{CED}$ values for some investors).

- The suggested algorithm clearly has its practical limitations regarding the number of assets in the portfolio; it can be used up to 15–20 assets (asset classes).

Running times for 48 investors, in seconds, under Windows XP, 3.16 GHz processor and 3.26 GB RAM:

<table>
<thead>
<tr>
<th></th>
<th>PT</th>
<th>CPT</th>
<th>PT(MV)</th>
<th>CPT(MV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>original data–set</td>
<td>91</td>
<td>432</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>associated normal distr.</td>
<td>90</td>
<td>430</td>
<td>3.4</td>
<td>4.9</td>
</tr>
<tr>
<td>call option added</td>
<td>180</td>
<td>845</td>
<td>3.6</td>
<td>5.4</td>
</tr>
</tbody>
</table>
Limitations: Computational costs

Average elapsed time in seconds for solving a CPT optimization problem, with $n$ denoting the number of assets. Computing environment: Windows XP, 3.16 GHz processor and 3.26 GB RAM.

For the dimensions from 10 to 20, we generated discrete distributions with 15 realizations for monthly assets–returns of assets included in the DJIA index, just for recording the computational times. From 21 assets upwards, the computational time exceeds 1 hour.

<table>
<thead>
<tr>
<th>$n$</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>9</td>
<td>18</td>
<td>31</td>
<td>55</td>
<td>96</td>
<td>164</td>
<td>268</td>
<td>430</td>
<td>673</td>
<td>1029</td>
<td>1575</td>
<td>2287</td>
<td>3319</td>
</tr>
</tbody>
</table>

In contrast, optimizing CPT along the MV frontier took less than 0.1 seconds for all cases.
For the details see: