Rank-SizeAnalysis of Optimal Portfolio Weights Across Portfolio Optimization Models

R. Cerqueti[1] **F. Cesarone**[2] **A. Di Paolo**[2] **V. Ficcadenti**[3]

[1]**Sapienza University of Rome** - *roy.cerqueti@uniroma1.it*
[2]**Roma Tre University** - *francesco.cesarone@uniroma3.it*, *alessio.dipaolo@uniroma3.it*
[3]**London South Bank University** - *ficcadv2@lsbu.ac.uk*

EURO 2025

Outline

Portfolio Selection: From Assets to Weights

Input
Set of Assets
history and/or distribution parameters

→

Portfolio Optimisation Models

MV – Mean-Variance model H. Markowitz, 1952; H. M. Markowitz, 1959

CVaR – Conditional Value-at-Risk Rockafellar, Uryasev, et al., 2000

RP – Risk Parity portfolio Maillard et al., 2010; Roncalli, 2013

MDP – Most Diversified Portfolio Choueifaty, 2008; Choueifaty et al., 2013

→

Output
Vector of Weights

Each model implements a different trade-off between expected gain and risk, based on historical data and specific risk measures. The resulting weight vectors are then subject to structural analysis via rank-size laws.

Minimum Risk Strategies -– Variance and CVaR Optimisation

General Minimum-Risk Portfolio Problem:

$$\min_{x \in \Delta} \text{Risk}(x), \quad \text{where } \Delta = \left\{ x \in \mathbb{R}^n_+ \,\middle|\, \sum_{i=1}^{n} x_i = 1; x_i \geq 0 \right\}$$

where n is the number of assets, x_i is the quota of investment in asset i and no short selling is allowed.

Minimum Variance (MV):

$$\min_{x \in \Delta} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j$$

Conditional Value-at-Risk (CVaR):

$$\min_{x,\zeta,d} \zeta + \frac{1}{\varepsilon T} \sum_{t=1}^{T} d_t \quad \text{s.t.} \begin{cases} d_t \geq -\sum_{i=1}^{n} r_{it} x_i - \zeta, \text{ t = 1,} \ldots \text{,T} \\ d_t \geq 0, \text{ t = 1,} \ldots \text{,T} \\ x \in \Delta, \quad \zeta \in \mathbb{R} \end{cases}$$

Risk Diversification Strategies – RP and MD

Goal: Avoid concentration by distributing risk evenly or diversifying it effectively.

Two main approaches:

- **Risk Parity (RP):** Equalise the Total Risk Contribution (TRC) across assets:

$$TRC_i(x) = x_i \cdot \frac{\partial \sigma(x)}{\partial x_i} \quad \Rightarrow \quad TRC_i(x) = TRC_j(x) \quad \forall i, j$$

  where

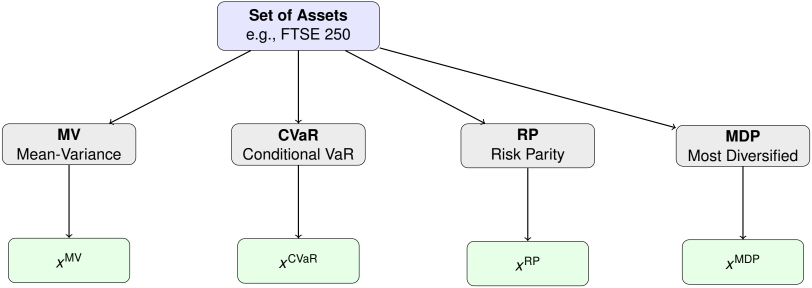
$$\sigma(x) = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}, \quad \frac{\partial \sigma(x)}{\partial x_i} = \frac{1}{\sigma(x)} (\Sigma x)_i$$

- **Most Diversified Portfolio (MDP):** Maximise the diversification ratio:

$$DR(x) = \frac{\sum_{i=1}^{n} \sigma_i x_i}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}}$$

From Assets to Portfolio Weight Vectors

We apply portfolio selection models to the same set of assets (e.g., the **FTSE 250**) to obtain four vectors of optimal weights x, each representing a different allocation strategy.



Each model yields an allocation vector $x = (x_1, \dots, x_n) \in \Delta$, where $\Delta = \{x \in \mathbb{R}^n : \sum x_i = 1, x_i \geq 0\}$ is the unit simplex.

Our aim

We want to

1. explore a Family of **Rank-Size laws** to model optimal portfolio weights resulting from different models;
2. use a **stochastic dominance method** for the optimal selection of law(s) that fits the data across models and/or dataset
3. understand the **informative power** of the parameters of the selected laws.

Rank-size analysis

Statistical methodology that allows one to derive a **unique system** from **disaggregated** and **properly ordered data** Mandelbrot, 1961; Zipf, 1949.

- **Starting point**: collection of *n* observations of a quantitative phenomenon - the size
- Observations ranked in **decreasing order** - highest size has rank 1 and lowest value has rank *n*
- Approximation of the **descending** scatter plot with a properly selected **decreasing curve** through a best-fit procedure

Rank-size and valid analysis

- Crucial point for a valid analysis: **statistically satisfactory fit**
- Selection of the **right law** - guess after ranked data visualisation.
- Analysis of the **calibrated parameters** - get insights on the system's structure described by the ranked data.

Famous laws and their applications are:

- Pareto, 1896 - Wealth distribution
- Auerbach, 1913 - Urban development via cities' size
- Zipf, 1949 - Words' frequencies
- Ausloos and Cerqueti, 2016 - Number of cities in Italian Provinces

Rank-size laws selected in this work

$$y = A\exp(-\alpha r) \qquad \text{Exponential law: } A, \alpha \geq 0 \qquad \text{(EXP)}$$

$$y = \frac{B}{r^{\beta}} + C \qquad \text{Power law: } B, \beta \geq 0,\ C \in \mathbb{R} \qquad \text{(POW - POW2)}$$

$$y = \frac{D(R+1-r)^{\gamma}}{r^{\xi}} \qquad \text{DGBD: } D, \gamma, \xi \geq 0 \qquad \text{(DGBD)}$$

$$y = \frac{E}{(r+\varepsilon)^{\lambda}} \qquad \text{Zipf-Mandelbrot law: } E, \varepsilon, \lambda \geq 0 \qquad \text{(ZM)}$$

$$y = p_1 r^3 + p_2 r^2 + p_3 r + p_4 \qquad \text{Polynomial: } p_i \in \mathbb{R},\ i = 1, \ldots, 4 \qquad \text{(POLY)}$$

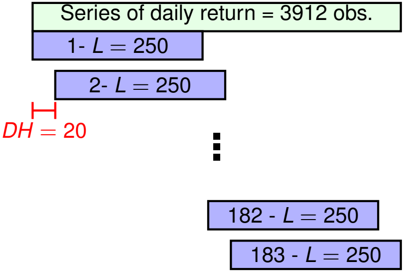
$$y = F\frac{(R+1-r+\Psi)^{\zeta_1}}{(R(r+\Phi))^{\zeta_2}} \qquad \text{Universal law: } F, \Psi, \zeta_1, \Phi, \zeta_2 \geq 0 \qquad \text{(UL)}$$

Where, R is the maximum rank in the considered empirical case, y and r are the variables and the rest are parameters to be calibrated.

Dataset & Rolling Time Window

We apply the strategies to the following real-world datasets

Index	# Assets	# Days	Time interval	Abbreviation
Dow Jones	29	3912	Jan 2009-Dec 2023	DJI
Euro Stoxx 50	47	3912	Jan 2009-Dec 2023	STOXX50E
NASDAQ 100	72	3912	Jan 2009-Dec 2023	NDX
FTSE 250	166	3912	Jan 2009-Dec 2023	FTSE



Some examples from FTSE 250 (ZIPF fit)

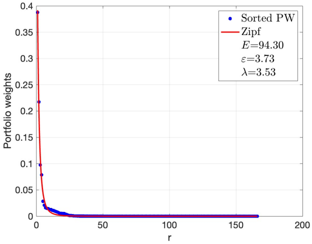


Figure: A Zipf fit on W. from MV model

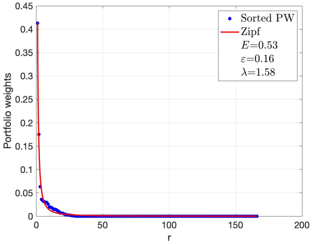


Figure: A Zipf fit on W. from CVaR model

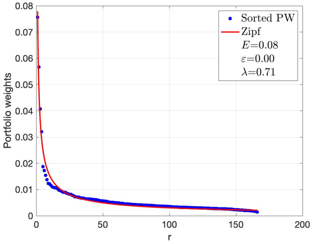


Figure: A Zipf fit on W. from RP model

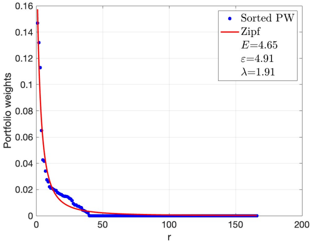


Figure: A Zipf fit on W. from MD model

Rank-Size best fits

Some examples from FTSE 250 (POLY fit)

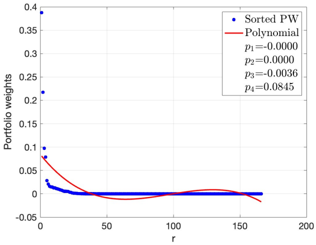


Figure: A POLY fit on W. from MV model

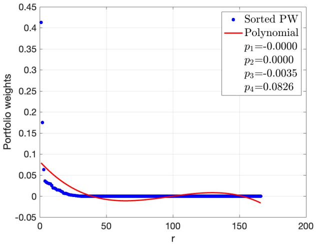


Figure: A POLY fit on W. from CVaR model

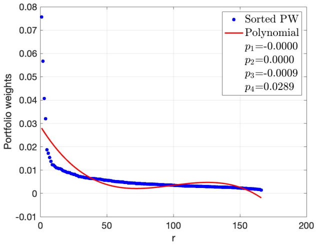


Figure: A POLY fit on W. from RP model

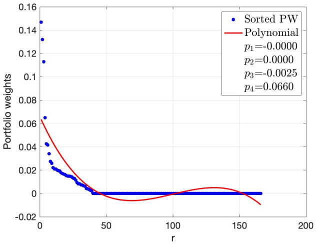


Figure: A POLY fit on W. from MD model

FTSE 250 empirical results: MV model example

Table: R^2 average results MV

R^2	μ	σ	*Me*	min	max
Exponential	0.9700	0.0219	0.9761	0.8849	0.9986
Power law2	0.9222	0.0678	0.9389	0.7427	0.9991
Power law	0.9403	0.0563	0.9561	0.7735	0.9992
DGBD	0.9430	0.0543	0.9613	0.7810	0.9991
Zipf	0.9776	0.0245	0.9862	0.8739	0.9996
Polynomial	0.4222	0.1550	0.4003	0.1898	0.7727
Universal	0.9644	0.0390	0.9771	0.7776	0.9992

Table: RMSE average results MV

RMSE	μ	σ	*Me*	min	max
Exponential	0.0051	0.0021	0.0047	0.0015	0.0152
Power law2	0.0074	0.0039	0.0066	0.0012	0.0194
Power law	0.0066	0.0038	0.0056	0.0012	0.0193
DGBD	0.0065	0.0038	0.0055	0.0012	0.0193
Zipf	0.0041	0.0025	0.0032	0.0008	0.0159
Polynomial	0.0251	0.0093	0.0248	0.0093	0.0454
Universal	0.0052	0.0031	0.0042	0.0011	0.0172

Table: MAE average results MV

MAE	μ	σ	*Me*	min	max
Exponential	0.0014	0.0005	0.0014	0.0004	0.0027
Power law2	0.0039	0.0018	0.0038	0.0006	0.0075
Power law	0.0033	0.0017	0.0032	0.0005	0.0071
DGBD	0.0025	0.0012	0.0023	0.0006	0.0057
Zipf	0.0013	0.0006	0.0012	0.0002	0.0029
Polynomial	0.0115	0.0028	0.0120	0.0053	0.0159
Universal	0.0020	0.0013	0.0017	0.0003	0.0099

Table: AIC results MV

AIC	μ	σ	*Me*	min	max
Exponential	-1779.58	143.54	-1779.13	-2167.07	-1388.80
Power law2	-1676.99	190.30	-1663.52	-2223.04	-1306.23
Power law	-1718.58	195.24	-1721.10	-2227.04	-1307.46
DGBD	-1725.14	194.62	-1726.98	-2221.13	-1308.20
Zipf	-1871.53	180.21	-1902.68	-2353.07	-1371.60
Polynomial	-1244.92	136.39	-1222.89	-1550.37	-1022.43
Universal	-1792.06	196.95	-1829.76	-2282.86	-1238.16

FTSE 250 empirical results: RMSE across models

Table: RMSE average results MV

RMSE	μ	σ	*Me*	min	max
Exponential	0.0051	0.0021	0.0047	0.0015	0.0152
Power law2	0.0074	0.0039	0.0066	0.0012	0.0194
Power law	0.0066	0.0038	0.0056	0.0012	0.0193
DGBD	0.0065	0.0038	0.0055	0.0012	0.0193
Zipf	0.0041	0.0025	0.0032	0.0008	0.0159
Polynomial	0.0251	0.0093	0.0248	0.0093	0.0454
Universal	0.0052	0.0031	0.0042	0.0011	0.0172

Table: RMSE average results CVaR

RMSE	μ	σ	*Me*	min	max
Exponential	0.0052	0.0020	0.0051	0.0012	0.0109
Power law2	0.0077	0.0034	0.0074	0.0018	0.0170
Power law	0.0069	0.0033	0.0062	0.0018	0.0164
DGBD	0.0067	0.0033	0.0061	0.0018	0.0162
Zipf	0.0044	0.0021	0.0039	0.0013	0.0115
Polynomial	0.0250	0.0092	0.0238	0.0077	0.0516
Universal	0.0061	0.0047	0.0048	0.0017	0.0341

Table: RMSE average results RP

RMSE	μ	σ	*Me*	min	max
exp	0.0032	0.0010	0.0030	0.0015	0.0055
Power law2	0.0011	0.0009	0.0008	0.0004	0.0045
Power law	0.0008	0.0003	0.0007	0.0004	0.0017
DGBD	0.0010	0.0009	0.0007	0.0004	0.0045
Zipf	0.0009	0.0009	0.0006	0.0004	0.0045
Polynomial	0.0030	0.0019	0.0024	0.0009	0.0099
Universal	0.0012	0.0009	0.0009	0.0003	0.0048

Table: RMSE average results MD

RMSE	μ	σ	*Me*	min	max
Exponential	0.0041	0.0021	0.0038	0.0014	0.0102
Power law2	0.0061	0.0015	0.0059	0.0035	0.0114
Power law	0.0043	0.0017	0.0038	0.0022	0.0112
DGBD	0.0039	0.0017	0.0034	0.0017	0.0111
Zipf	0.0033	0.0016	0.0030	0.0014	0.0101
Polynomial	0.0091	0.0050	0.0080	0.0026	0.0272
Universal	0.0034	0.0016	0.0031	0.0015	0.0112

RMSE distributions

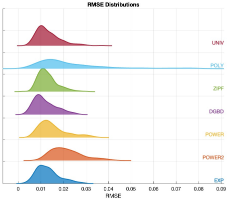


Figure: RMSE distributions MV

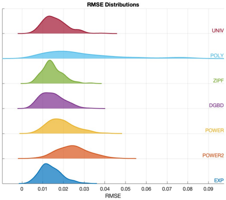


Figure: RMSE distributions CVaR

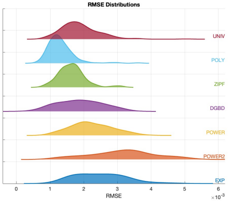


Figure: RMSE distributions RP

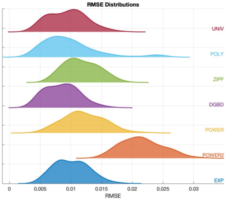


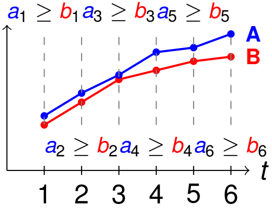
Figure: RMSE distributions MD

Stochastic Dominance Criteria: ZSD, FSD, SSD

ZSD (Zero-order)

$\Pr(A - B \leq 0) = 0$

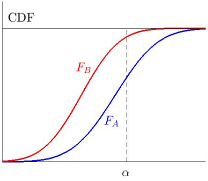
$\Rightarrow a_t \geq b_t \; \forall t$



Dominance in all states.

FSD (First-order)

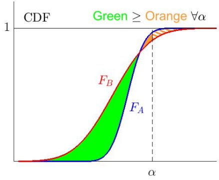
$F_A(\alpha) \leq F_B(\alpha) \; \forall \alpha$



CDF of A lies below B

SSD (Second-order)

$$\int_{-\infty}^{\alpha} F_A(\tau) d\tau \leq \int_{-\infty}^{\alpha} F_B(\tau) d\tau$$



Integrates area under CDFs.

Equivalently, SSD $\iff \mathrm{CVaR}_\beta(A) \leq \mathrm{CVaR}_\beta(B) \; \forall \beta \in (0, 1]$

RMSE–Based Model Evaluation via Stochastic Dominance

$RMSE = (RMSE_t)_{t=1}^{T}$ and $RMSE' = (RMSE'_t)_{t=1}^{T}$ is a sequence of RMSEs obtained from two different rank-size functions in rolling windows T.
Assume uniform probability: $\pi_t = 1/T$.

First-Order Stochastic Dominance (FSD)

$RMSE$ FSD-dominates $RMSE'$ iff:

$$\sum_{RMSE_t \leq \alpha} \pi_t \geq \sum_{RMSE'_t \leq \alpha} \pi_t \quad \forall \alpha \in \mathbb{R}.$$

Equivalent: $RMSE_{(j)} \leq RMSE'_{(j)} \quad \forall j = 1, \ldots, T.$

Second-Order Stochastic Dominance (SSD)

$RMSE$ SSD-dominates $RMSE'$ iff:

$$\sum_{t=1}^{j} RMSE_{(t)} \leq \sum_{t=1}^{j} RMSE'_{(t)} \quad \forall j = 1, \ldots, T.$$

Reflects dominance in cumulative error.

CVaR and Tail-Based Risk Measures

Define tail and CVaR operators:

$$Tail_\beta(RMSE) = \sum_{RMSE_t \geq \eta} \pi_t RMSE_t,$$

$$CVaR_\beta(RMSE) = \frac{1}{\beta} Tail_\beta(RMSE).$$

If $\pi_t = 1/T$ and $\beta = j/T$:

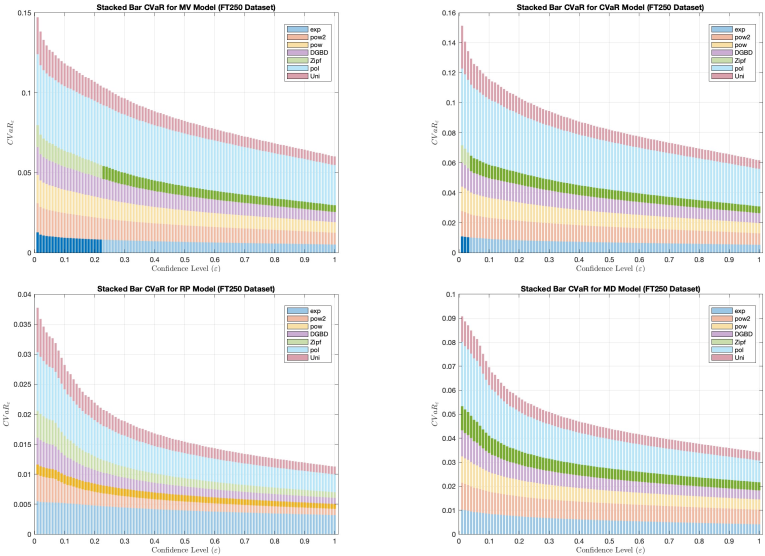
$$CVaR_{j/T}(RMSE) = \frac{1}{j} \sum_{t=1}^{j} RMSE_{(t)},$$

$$Tail_{j/T}(RMSE) = \frac{j}{T} CVaR_{j/T}(RMSE).$$

$CVaR$ dominance is equivalent to comparing tail mass:

$$CVaR_\beta(RMSE) \leq CVaR_\beta(RMSE') \quad \forall \beta \in (0,1].$$

CVaR across Confidence Levels — FTSE250 Case



Cumulative ε-Stochastic Dominance ($CS\varepsilon SD$)

From Bruni et al., 2017: Approximate SSD with tolerance $\varepsilon > 0$.

Let $\Delta_j = CVaR_{j/T}^{(RMSE)} - CVaR_{j/T}^{(RMSE')}$ and sort: $\Delta^{(1)} \geq \cdots \geq \Delta^{(T)}$.

Cumulative difference:

$$C_t := \sum_{j=1}^{t} \Delta^{(j)} \quad \forall j = 1, \ldots, T.$$

Max cumulative CVaR disadvantage:

$$C^{\max}(RMSE, RMSE') := \max_t C_t.$$

Decision rule:

$$\min\left\{C^{\max}(RMSE, RMSE'),\ C^{\max}(RMSE', RMSE)\right\}$$

CVaR differences

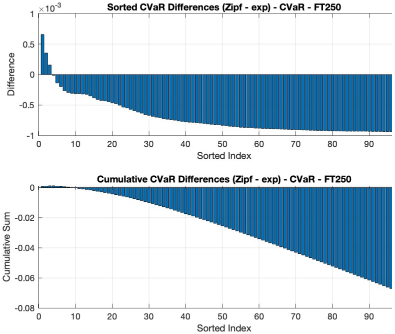


Figure: CVaR differences Zipf-Exp

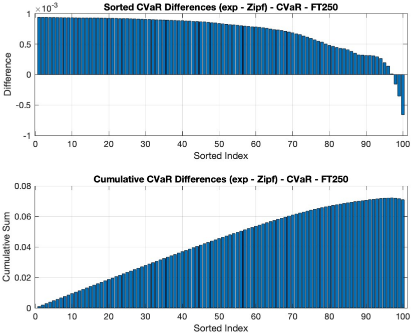


Figure: CVaR differences Exp-Zipf

Ranking Functions via Pairwise $CS\varepsilon SD$ Scoring

Objective: Identify the best-fitting *rank-size* function in each model/dataset by pairwise comparisons under $CS\varepsilon SD$.

Step 1 – Pairwise comparison: For each unordered pair $(RMSE, RMSE')$, evaluate:

$$C^{\max}(RMSE, RMSE') := \max_t \sum_{j=1}^{t} \left(CVaR_{j/T}^{(RMSE)} - CVaR_{j/T}^{(RMSE')} \right)^{\downarrow}$$

Step 2 – Assign score:

$$\begin{aligned} +1 \quad &\text{if} \quad C^{\max}(RMSE, RMSE') < C^{\max}(RMSE', RMSE) \\ +1 \quad &\text{if} \quad C^{\max}(RMSE', RMSE) < C^{\max}(RMSE, RMSE') \end{aligned}$$

No points assigned in case of equality.

Step 3 – Aggregated score: Each function receives a total score equal to the number of functions it dominates:

$$\text{Score}(f) := \sum_{f' \neq f} \mathbb{I}\left[C^{\max}(f, f') < C^{\max}(f', f)\right]$$

Interpretation: Score ranges in $\{0, \dots, m-1\}$. A function scoring $m-1$ dominates all others in that model/dataset.

Score Results

Table: Scores for FTSE250

FT250	MV	CVaR	MD	RP	MEAN
Exponential	5	5	2	1	3.25
Power law2	1	1	1	3	1.5
Power law	2	2	3	6	3.25
DGBD	3	3	4	4	3.5
Zipf	6	6	6	5	5.75
Polynomial	0	0	0	0	0
Universal	4	4	5	2	3.75

Table: Scores DJ

DJ	MV	CVaR	MD	RP	MEAN
Exponential	4	6	4	1	3.75
Power law2	1	1	0	0	0.5
Power law	2	2	1	2	1.75
DGBD	6	4	6	3	4.75
Zipf	5	5	3	5	4.5
Polynomial	0	0	2	6	2
Universal	3	3	5	4	3.75

Table: Scores for NS

NS	MV	CVaR	MD	RP	MEAN
Exponential	5	6	3	0	3.5
Power law2	1	1	1	2	1.25
Power law	2	2	2	4	2.5
DGBD	3	3	4	6	4
Zipf	6	5	6	5	5.5
Polynomial	0	0	0	1	0.25
Universal	4	4	5	3	4

Table: Scores for EX

EX	MV	CVaR	MD	RP	MEAN
Exponential	6	6	3	0	3.75
Power law2	1	1	0	2	1
Power law	2	2	2	4	2.5
DGBD	3	3	6	6	4.5
Zipf	5	5	5	5	5
Polynomial	0	0	1	1	0.5
Universal	4	4	4	3	3.75

Averaged Scores

Table: Score averaged across datasets

Across Dataset	MV	CVaR	MD	RP	MEAN
Exponential	4.6	5.6	3.4	0.4	3.5
Power law2	1	1	0.6	2	1.15
Power law	2	2	2	4.2	2.55
DGBD	3.8	3.2	4.6	5	4.15
Zipf	5.6	5.4	5.2	4.8	5.25
Polynomial	0	0	0.6	1.8	0.6
Universal	4	3.8	4.6	2.8	3.8

Conclusions

- → Several **rank-size functions** can be identified that can capture the shape of the distribution of optimal portfolio weights.
- → The ϵ**– stochastic dominance criterion** from Bruni et al., 2017, makes it possible to order the most suitable function.
- → The estimated parameters show **regularity** and **persistence** of the information across functions. The information power of rank-size functions allows us to characterise and represent a complex system of n assets with intrinsic movement and co-movement.

Future research

- How can rank-size parameter information be used to build **predictive models** on portfolio selection strategies?
- Is it possible to use the information content provided by rank-size functions to **guess the shape** of the distribution of future optimal weights?
- Can **portfolio concentration** indices be constructed based on the parameters of the rank-size function?

Thank you for your attention

References I

Auerbach, F. (1913). Das gesetz der bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, *59*, 74–76.

Ausloos, M., & Cerqueti, R. (2016). A universal rank-size law. *PloS one*, *11*(11), e0166011.

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research*, *259*(1), 322–329.

Choueifaty, Y. (2008). Towards maximum diversification. *Available at SSRN 4063676*.

Choueifaty, Y., Froidure, T., & Reynier, J. (2013). Properties of the most diversified portfolio. *Journal of investment strategies*, *2*(2), 49–70.

References II

- Maillard, S., Roncalli, T., & Teıletche, J. (2010).The properties of equally weighted risk contribution portfolios. *Journal of Portfolio Management*, *36*(4), 60.
- Mandelbrot, B. (1961).On the theory of word frequencies and on related markovian models of discourse. *Structure of language and its mathematical aspects*, *12*, 190–219.
- Markowitz, H. (1952).Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. Retrieved October 15, 2024, from http://www.jstor.org/stable/2975974
- Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. Yale University Press. http://www.jstor.org/stable/j.ctt1bh4c8h
- Pareto, V. (1896). *Cours d'économie politique* (Vol. 1). F. Rouge.

References III

Rockafellar, R. T., Uryasev, S., et al. (2000). Optimization of conditional value-at-risk. *Journal of risk*, *2*, 21–42.

Roncalli, T. (2013). *Introduction to risk parity and budgeting*. CRC Press.

Zipf, G. K. (1949). Human behaviour and the principle of least-effort. cambridge ma edn. *Reading: Addison-Wesley*, *24*.

Some optimization approaches to asset allocation

- ***Minimize risk*** (VaR, CVaR, Variance, MAD, ...)
- ***Optimize some performance measure***
  (Sharpe Ratio, Omega Ratio, Rachev Ratio, etc.)

ADVANTAGES:

- "Optimal" solution
- Coherence with decision theory
- Long history [De Finetti 1940, Markowitz 1952, Markowitz 1959]

DISADVANTAGES:

- Highly sensitive to data
- Optimization process amplifies estimation errors
- Not balanced in terms of risk allocation to each asset

Portfolio selection: minimum-risk strategies

- **Minimize Risk** (**Variance**, **CVaR**, MAD, Entropy, Drawdown, ...)

$$\begin{cases} \min\limits_{x} \quad \text{Risk(x)} = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j \\ \text{s.t.} \\ \qquad x \in \Delta \end{cases} \tag{1}$$

  - σ_{ij} is the covariance between asset returns
  - x_i is the percentage of capital invested in asset i, $i = 1, \ldots, n$
  - we consider the standard simplex $\Delta = \{x_i \in \mathbb{R} : \sum_{i=1}^{n} x_i = 1, x_i \geq 0\}$, with budget and no short-selling constraints.

Conditional Value-at-Risk

Definition (Conditional Value at Risk)

$$CVaR_{\varepsilon}(x) = \mathbb{E}[L_P(x) | L_P(x) \geq VaR_{\varepsilon}(x)]$$

where $VaR_{\varepsilon}(x) = -Q_{\varepsilon}[R_P(x)]$.
From the computational viewpoint, the Conditional Value-at-Risk (CVaR) of a portfolio x is equal to the minimum value on ζ of the auxiliary function $F_{\varepsilon}(x, \zeta)$ [Rockafellar-Uryasev 2000]

$$CVaR_{\varepsilon}(x) = \min_{\zeta \in \mathbb{R}} F_{\varepsilon}(x, \zeta), \tag{2}$$

where

$$F_{\varepsilon}(x, \zeta) = \zeta + \frac{1}{\varepsilon} \sum_{t=1}^{T} p_t [l(x, t) - \zeta]^+. \tag{3}$$

Furthermore, minimizing $CVaR_{\varepsilon}(x)$ w.r.t. x is equivalent to minimizing $F_{\varepsilon}(x, \zeta)$ w.r.t. x and ζ as follows:

$$\min_{x \in C} CVaR_{\varepsilon} = \min_{x \in C} \min_{\zeta \in \mathbb{R}} F_{\varepsilon}(x, \zeta) = \min_{(x, \zeta) \in C x \mathbb{R}} F_{\varepsilon}(x, \zeta). \tag{4}$$

Conditional Value-at-Risk: linear formulation

Furthermore, considering the losses as negative outcomes of portfolio return and assuming that all scenarios $l(x,t)$ are equally likely, we have $l(x,t) = -r_P(x,t) = -\sum_{i=1}^{n} r_{it}x_i$ and $p_t = \frac{1}{T}$. Hence, Expression (3) becomes

$$F_\varepsilon(x,\zeta) = \zeta + \frac{1}{\varepsilon T}\sum_{t=1}^{T}\left[\sum_{i=1}^{n} -r_{it}x_i - \zeta\right]^+. \tag{5}$$

Following [Rockafellar-Uryasev, 2000], we can introduce T additional variables $d_t = \max(-\sum_{i=1}^{n} r_{it}x_i - \zeta, 0)$, with $d_t \geq 0$, $d_t \geq -\sum_{i=1}^{n} r_{it}x_i - \zeta$, thus obtaining the following Linear Programming problem

$$\begin{cases} \min\limits_{x} & CVaR_\varepsilon(x) \\ \text{s.t.} & \\ & x \in \Delta \end{cases} \Rightarrow \begin{cases} \min\limits_{x,\zeta,d} & \zeta + \frac{1}{\varepsilon}\frac{1}{T}\sum_{t=1}^{T} d_t \\ \text{s.t.} & \\ & d_t \geq -\sum_{i=1}^{n} r_{it}x_i - \zeta \qquad t = 1,\ldots,T \\ & d_t \geq 0 \qquad t = 1,\ldots,T \\ & x \in \Delta \\ & \zeta \in \mathbb{R} \end{cases}$$

Some diversification approaches to asset allocation

- Equally Weighted (EW) (also called naïve or uniform) portfolio

ADVANTAGES:

- No sensitivity to estimation errors
- Seems to yield good performance in practice [De Miguel, 2009])
- Very long history (**Babylonian Talmud** ca. 1500 years ago: *A man should always place his money, one third in land, a third into merchandise and keep a third in hand*)

DISADVANTAGES:

- No attempt to search for a “best” solution
- No use of past experience

Portfolio selection: risk-diversification strategies

Asset allocation based on risk-diversification:

- **Risk diversification**: Risk Parity (Equal Risk Bounding) [Maillard 2010, Roncalli 2014, Cesarone-Tardella 2016]
  - Effective number of bets [Meucci 2009, 2015]
  - CVaR Equal Risk Contribution [Cesarone-Colucci 2017]
  - Expectiles Risk Parity [Bellini-Cesarone-Tardella 2019]
- **Maximum diversification** [Choueifaty et al. 2013].

Risk-parity objective

Assuming that the risk of a portfolio of assets is measured by the volatility of its returns

$$\sigma(x) = \sqrt{x^T \Sigma x} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}$$

The **marginal risk contribution** of asset i is

$$\sigma_{x_i}(x) = \frac{\partial \sigma(x)}{\partial x_i} = \frac{x_i \sigma_i^2 + \sum_{j \neq i} x_j \sigma_{ij}}{\sigma(x)} = \frac{1}{\sigma(x)} (\Sigma x)_i \quad (6)$$

For a positively homogeneous risk measure as volatility, a reasonable measure of total risk contribution of each asset to the total risk of the portfolio is provided by

$$TRC_i(x) = x_i \sigma_{x_i}(x) \quad (7)$$

The **Risk Parity** (RP) portfolio is characterized by the requirement of having equal total risk contribution from each asset

$$TRC_i(x) = TRC_j(x) \quad \forall i, j$$

Risk-parity portfolio

A possible approach to solve the RP problem is the **least squares model**

$$\begin{cases} \min\limits_{x} & \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}(x_i(\Sigma x)_i - x_j(\Sigma x)_j)^2 \\ \text{s.t.} & \\ & x \in \Delta \end{cases} \tag{8}$$

Clearly, we obtain the RP portfolio only if $f(x^*) = 0$, i.e., $x_i(\Sigma x)_i = x_j(\Sigma x)_j$.

Most Diversified portfolio

The **Most Diversified portfolio** (MDP) is obtained maximizing the **Diversification ratio**, $DR(x)$, or minimizing its reciprocal. Using volatility as a risk measure, $DR(x)$ is the ratio of the weighted average of the stocks volatilities σ_i (the worst case) divided by the portfolio volatility [Choueifaty et al., 2008], $DR(x) = \frac{\sum_{i=1}^{n} \sigma_i x_i}{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}}$.

$$
\begin{cases}
\min\limits_{x} \dfrac{1}{DR(x)} = \dfrac{\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}}{\sum_{i=1}^{n} x_i \sigma_i} \\
\text{s.t.} \\
\qquad x \in \Delta
\end{cases}
\Longrightarrow
\begin{cases}
\min\limits_{y} \quad \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} y_i y_j \sigma_{ij} \\
\text{s.t.} \\
\qquad \sum\limits_{i=1}^{n} \sigma_i y_i = 1 \\
\qquad y_i \geq 0 \quad i = 1, \ldots, n
\end{cases}
$$

Thus, the normalized optimal portfolio weights are $x_i^{MDP} = \dfrac{y_i^*}{\sum\limits_{i=1}^{n} y_i^*}$.

Zero-order Stochastic Dominance (ZSD)

Let A and B be two random variables with distribution functions

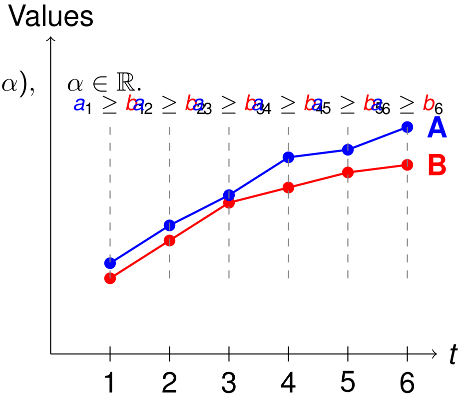
$$F_A(\alpha) = \Pr(A \leq \alpha), \quad F_B(\alpha) = \Pr(B \leq \alpha), \quad \alpha \in \mathbb{R}.$$

ZSD: A dominates B iff:

$$F_{A-B}(0) = \Pr(A - B \leq 0) = 0$$

In empirical terms, if we denote the realizations by a_t and b_t, then:

$$a_t \geq b_t \quad \text{almost everywhere.}$$

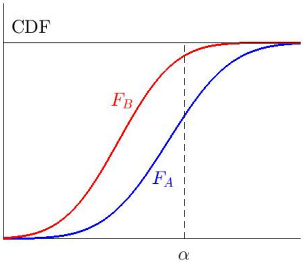


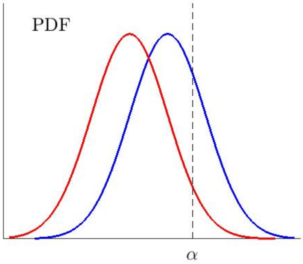
First-order Stochastic Dominance (FSD)

A is preferred to *B* w.r.t. FSD iff

$$F_A(\alpha) \leq F_B(\alpha) \quad \forall \alpha \in \mathbb{R}. \tag{9}$$

and the inequality is strict for at least one α





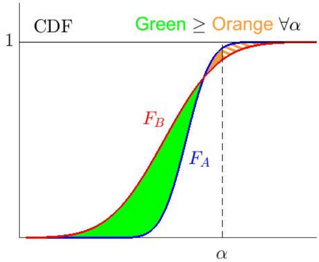
In other words, the probability that *A* takes a value less than α is always lower than (or equal to) the probability that *B* takes a vale less than α.

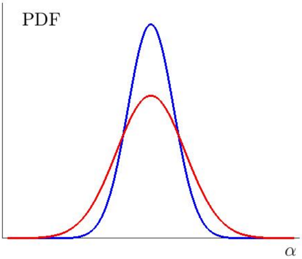
Second-order Stochastic Dominance (SSD)

A is preferred to B w.r.t. SSD iff

$$\int_{-\infty}^{\alpha} F_A(\tau)d\tau \leq \int_{-\infty}^{\alpha} F_B(\tau)d\tau \quad \forall\alpha \in \mathbb{R}. \tag{10}$$

and the inequality is strict for at least one α





Second-order Stochastic Dominance (SSD)

The following alternative criterion can express the SSD relationships A is preferred to B w.r.t. SSD iff

$$CVaR_\beta(A) \leq CVaR_\beta(B) \quad \forall \beta \in (0, 1]. \tag{11}$$

Linear regression: HI vs Exponential parameters

$$HI = \sum_{i=1}^{n} x_i^2, \qquad y = \hat{A}\exp(-\hat{\alpha} r)$$

Linear regression

$$HI_t = \hat{\beta}_0 + \hat{\beta}_1 A_t + \hat{\beta}_2 \alpha_t + \varepsilon_t,$$

where ε_t is the residual associated with the observation in period t the test statistic is:

$$d = \frac{\sum_{t=2}^{T} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{T} \varepsilon_t^2}.$$

Durbin-Watson tests the null hypothesis that the residuals are uncorrelated:

- $d = 2$, no correlation in the residuals
- $d \approx 0$, residuals positively correlated
- $d > 2$, residuals negatively correlated

Autocorrelation in residuals

Since, from the observed values, the residuals of the regression appear autocorrelated, we apply the first difference and then the linear regression on the differentiated values. We denote by

$$\Delta HI = HI_t - HI_{t-1},$$
$$\Delta A = A_t - A_{t-1},$$
$$\Delta \alpha = \alpha_t - \alpha_{t-1},$$

we have

$$\Delta HI = \hat{\beta}'_0 + \hat{\beta}'_1 \Delta A_t + \hat{\beta}'_2 \Delta \alpha_t + \varepsilon'_t.$$

Regression results

Table: Regression Results for MD (FT250)

Variable	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	R^2	d	$\hat{\beta}'_0$	$\hat{\beta}'_1$	$\hat{\beta}'_2$	R^2	d
Estimate	0.02289	0.39561	0.13126	0.97744	0.39701	0.00002	0.39574	0.15188	0.92190	1.98754
p-value	0.00000	0.00000	0.00000		0.00000	0.93125	0.00000	0.00000		0.92695

Table: Regression Results for CVaR (FT250)

Variable	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	R^2	d	$\hat{\beta}'_0$	$\hat{\beta}'_1$	$\hat{\beta}'_2$	R^2	d
Estimate	0.06896	0.02117	-0.21962	0.84731	0.54484	-0.00001	0.02191	-0.16208	0.71792	2.32450
p-value	0.00000	0.18505	0.00000		0.00000	0.99623	0.06061	0.00000		0.02597