

Comparative Case Study on Portfolio Optimization: Modified Continuous Rank Length vs. Stochastic Dominance

Tereza Čapková¹

Abstract: This paper conducts a case study comparing the effectiveness of a modified Continuous Rank Length (CRL) metric with the Stochastic Dominance approach in portfolio optimization. The modified CRL method, designed to assess portfolio performance across the entire time interval, provides a holistic evaluation of rank stability, while Stochastic Dominance offers a robust framework for comparing return distributions. The analysis of simulated portfolios demonstrates that SD consistently identifies portfolios with the highest returns and lowest risks, particularly in early periods of the dataset. Conversely, CRL excels at selecting portfolios with enhanced long-term stability but is less efficient due to computational demands and broader stock inclusion. These findings provide actionable insights into the trade-offs between maximizing returns, ensuring stability, and managing computational resources in portfolio selection.

Keywords: Portfolio Performance, GET package, ERL, CRL, Extreme Rank Length, Continuous Rank Length, Financial Management, Financial Analysis, Stochastic Dominance

JEL Classification: G17

1 Introduction

In portfolio evaluation, researchers and practitioners must select the most suitable method for analyzing performance. With the increasing complexity of financial markets, traditional methods of comparison and evaluation often need to be revised to capture the nuances of portfolio behavior, mainly when long-term stability and risk are of interest. This study provides a comprehensive comparison of two portfolio optimization methods, offering practical insights for those interested in maximizing portfolio stability over time.

The two methods analyzed in this study are stochastic dominance (SD) and Continuous Rank Length (CRL). Stochastic dominance is a well-established approach in portfolio optimization, widely used for identifying portfolios with the best risk-return trade-offs. In contrast, the CRL metric, originally designed to detect outliers in functional data, is adapted here to evaluate portfolio stability over time. By comparing these methods, we aim to explore whether CRL's emphasis on consistent performance can serve as a valuable alternative to SD, particularly for investors who prioritize long-term stability over short-term gains.

2 Methods

We utilized a modified version from the GET package developed by Myllymäki & Mrkvička (2020) for the CRL metric. Due to our limited expertise in linear programming, we based our analysis on a previous case study by Kozmík (2019), which provided all the necessary SD data for comparison.

Before delving into the comparison, we will first define the metrics and methods used for it. Then, very briefly, let us also introduce stochastic dominance and our assumptions about this comparison.

2.1 Continuous Rank Length

CRL metric (Hahn, 2015; Mrkvička et al., 2022) is a method used to assess function behavior while accounting for the spacing between functions. The original formula can be found in Section 2.3.2 of Mrkvička et al. (2022). The standard version of this metric focuses on specific points within the interval rather than evaluating the entire interval. With a simple modification, we extended the CRL by summing these individual ranks, transforming it from an infimal metric to an integral one. The formula for this modification looks like this:

$$c_i = \frac{1}{s+1} \sum C_i(r) \quad \text{with} \quad C_i(r) = s+1 - c_i(r)$$

¹ University of South Bohemia in České Budějovice, Faculty of Economics, Department of Applied Mathematics and Informatics, Studentská 13, 370 05 České Budějovice 5, Czech Republic, capkot00@jcu.cz.

where c_j is the continuous pointwise rank that is a refinement of R_i . For a better understanding, see Section 2.3.2 in Mrkvička et al., 2022.

Summing the ranks allows the metric to reflect the overall performance across the interval rather than concentrating only on the lowest rank. This provides a more well-rounded view of the function's behavior and makes the metric less sensitive to extreme outliers. A single poor rank, which would heavily influence a minimum-based metric, does not overshadow the overall performance in the sum-based approach. This is especially advantageous when evaluating portfolios, as it rewards those that consistently perform well, giving a more precise representation of long-term stability. This improvement addresses the limitations found in other metrics, such as the Extreme Rank Length, which was overly sensitive to extreme variations.

Return

Suppose we have stock i , where $i = 1, 2, \dots, n$ and we have adjusted closing stock prices in period t , where $t = 1, 2, \dots, m$, as initial values IV_{it} of stock i at the start of period t and final value FV_{it} of stock i at the end of period t . The simple return of one stock in one-period R_{it} is calculated using the following formula. To determine the return of a specific portfolio CR_t for period t , the individual simple returns of stocks within the portfolio must be weighed and summed. Assume we have n stocks and let w_i represent the weight of stock i in the portfolio. The weighted return of the portfolio in a period t can be calculated using the following formula. (Stewart et al., 2019)

$$R_{it} = \frac{FV_{it}}{IV_{it}} - 1 \quad CR_t = \sum_{i=1}^n w_i R_{it}$$

Risk

To assess risk, we calculated the standard deviation (SD) of adjusted returns for each quarter using complete daily data from the respective quarters. This allowed for a more precise and granular volatility measurement by capturing the daily variations in returns. This approach ensures a thorough evaluation of risk for each period. Since our software tools require functions to be either maximized or minimized, we used the inverse of the SD for each quarter. This adjustment was necessary to meet the software requirements and facilitate our analysis.

2.2 Stochastic Dominance

Stochastic dominance is a statistical approach used in finance to compare investments by evaluating their entire return distributions. Unlike methods that focus on single points or average returns, stochastic dominance considers cumulative probabilities, allowing investors to understand potential returns and associated risks. This comprehensive view enables tailored decision-making across varying risk tolerance levels, making it a powerful tool for selecting investments in complex financial landscapes.

First-order stochastic dominance (FSD) applies when one investment is preferred across all levels of wealth, regardless of risk tolerance. Investment A dominates investment B in the first order if, for every possible outcome, A offers an equal or greater return than B . This means that $F_A(x) \leq F_B(x)$ for all x , with strict inequality for at least one x , where F represents the cumulative distribution function (CDF) of returns. FSD is particularly relevant for investors who prioritize higher returns without consideration for risk, and it was formalized in early works by Hadar & Russell (1969) and Hanoch & Levy (1969), who developed foundational rules for ranking uncertain prospects.

Second-order stochastic dominance (SSD) refines the comparison by incorporating risk aversion, making it relevant for investors who value both returns and risk mitigation. Investment A is said to dominate investment B at the second order if: $\int_{-\infty}^x F_A(t)dt \leq \int_{-\infty}^x F_B(t)dt$ for all x , with strict inequality for at least one x . This criterion suggests that A provides the same or higher expected returns while reducing potential downside risk compared to B . SSD is foundational in finance and economic theory, with Rothschild & Stiglitz (1970) formalizing its relationship with risk preferences and Levy (1992) connecting it to expected utility theory, thus broadening its application in portfolio selection and risk management.

Third-order stochastic dominance (TSD) further refines the analysis by incorporating preferences for risk aversion and skewness, making it valuable for investors who prefer higher returns and lower risk and favor positive skewness (upside potential). Investment A dominates B at the third order if $\int_{-\infty}^x \int_{-\infty}^t F_A(u) du dt \leq \int_{-\infty}^x \int_{-\infty}^t F_B(u) du dt$ for all x , with strict inequality for at least one x . This criterion accounts for the asymmetry in returns, aligning with investors' preference for distributions with more significant potential for high returns. TSD was extended in Fishburn (1974) to

address skewness preferences. Whitmore & Findlay (1978) provided a detailed analysis, making TSD an essential tool for capturing the full spectrum of investment preferences under uncertainty.

Reflecting on stochastic dominance, our assumption is that strict inequalities must hold across all points, aligning with the equivalence that one portfolio dominates another if, and only if, for all utility functions (within the relevant set for each dominance type), the utility of one portfolio is greater than or equal to the utility of the other. However, our approach makes it possible to identify portfolios that excel across the broader performance spectrum while not meeting this strict requirement at a single point or under a specific function. This could be particularly impactful in cases with non-standard or unusual distributions. For distributions resembling normality, results may appear similar under both methods; however, portfolios with atypical distributions might yield different insights. This approach could provide results that are as strong, or potentially stronger, than FSD. Still, SSD and TSD may present more nuanced outcomes, possibly offering an empirical or graphical perspective on similar findings.

3 Case study

As previously mentioned, we used the same data set for consistency since we based our comparison on Kozmík's (2019) research. The data comprises stock prices from assets in the Dow Jones Industrial Average (DJIA) index, representing some of the largest and most actively traded companies. The data was collected using the quantmod package (Ryan & Ulrich, 2024), which retrieves financial information from Yahoo Finance. Our data set spans from April 2008 to July 2018, explicitly chosen to include Visa's market entry and avoid issues with missing data. This period also excludes the impact of the COVID-19 pandemic, ensuring that that event does not skew the data. While applying this approach in more volatile conditions, such as during the pandemic, is a future research goal and it is beyond the scope of this study.

Table 1 lists all the stock tickers, basic characteristics, and information. The characteristics were calculated quarterly from the quarterly returns. As noted by Kozmík, Apple stands out with the highest expected returns, but it also exhibits one of the highest standard deviations, indicating a higher level of risk.

Table 1 Basic characteristics calculated from quarterly data

Ticker	Name	Mean	Min	Max	SD	Median
APPL	Apple	6,36	-34,93	45,79	0,15	8,27
V	Visa	5,77	-24,48	31,98	0,11	4,64
HD	Home Depot	5,76	-19,92	34,12	0,10	6,63
UNH	United Health	5,58	-27,59	26,53	0,12	5,33
BA	Boeing	5,15	-24,07	30,13	0,14	4,79
NKE	Nike	4,85	-22,54	24,61	0,12	7,21
DD	DuPont	4,57	-51,39	84,95	0,24	4,25
MSFT	Microsoft	4,07	-26,09	25,21	0,12	5,76
CAT	Caterpillar	3,54	-39,75	56,60	0,18	5,98
DIS	The Walt Disney Company	3,53	-24,93	31,48	0,13	5,11
INTC	Intel	3,51	-20,15	19,91	0,11	1,95
TRV	The Travelers Companies	3,46	-16,86	27,38	0,09	4,63
AXP	American Express	3,35	-47,51	62,16	0,18	3,21
MMM	3M	3,31	-25,24	22,39	0,10	2,67
MCD	McDonald's	3,29	-13,60	20,54	0,07	3,23
JPM	JPMorgan Chase & Co	3,22	-35,98	37,27	0,16	4,64
JNJ	Johnson & Johnson	2,46	-12,55	16,01	0,07	2,26
PFE	Pfizer	2,35	-23,85	26,13	0,10	2,82
RTX	United Technologies Corp.	2,33	-21,42	19,28	0,10	4,31
MRK	Merck & Co	2,22	-16,82	20,85	0,09	1,98
CSCO	Cisco Systems	2,13	-25,74	25,21	0,11	2,75
VZ	Verizon Communications	2,11	-11,54	25,77	0,09	2,35
CVX	Chevron	2,08	-16,88	21,23	0,11	-0,26
WBA	Walgreens Boots Alliance	2,06	-29,00	29,72	0,14	0,98
WMT	Walmart	1,93	-12,87	26,54	0,08	2,05
KO	The Coca-Cola Company	1,74	-15,70	17,86	0,07	1,61
GS	Goldman Sachs Groups Inc	1,51	-37,07	49,12	0,17	1,26
IBM	IBM	1,18	-23,15	14,62	0,08	2,17
PG	Protector & Gamble	1,04	-24,49	14,34	0,08	1,51
XOM	Exxon Mobile	0,73	-16,16	19,84	0,09	0,28

Source: Own processing

Table 2 outlines the weight allocations for stocks under FSD, SSD, and TSD criteria, with corresponding returns and risks detailed in Table 5. The FSD portfolio assigns significant weight to high-return stocks, such as AAPL and HD, to maximize overall returns. In contrast, the more restrictive SSD portfolio concentrates on fewer assets, placing greater emphasis on the most profitable stocks, which leads to a slightly higher expected return. The TSD portfolio, while similar to SSD, does not increase the weights on top-performing stocks further. As a result, the expected return for TSD is slightly lower, indicating that relaxing the constraint with TSD did not yield better returns (see Table 5).

Table 2 Weights of portfolios selected by stochastic dominance

	AAPL	CAT	DIS	HD	JPM	MCD	TRV	UNH	V
■ FSD	0,441	0,001	0,015	0,321	0,029	0,037	0,030	0,002	0,124
■ SSD	0,528	0	0	0,472	0	0	0	0	0
■ TSD	0,510	0	0	0,49	0	0	0	0	0

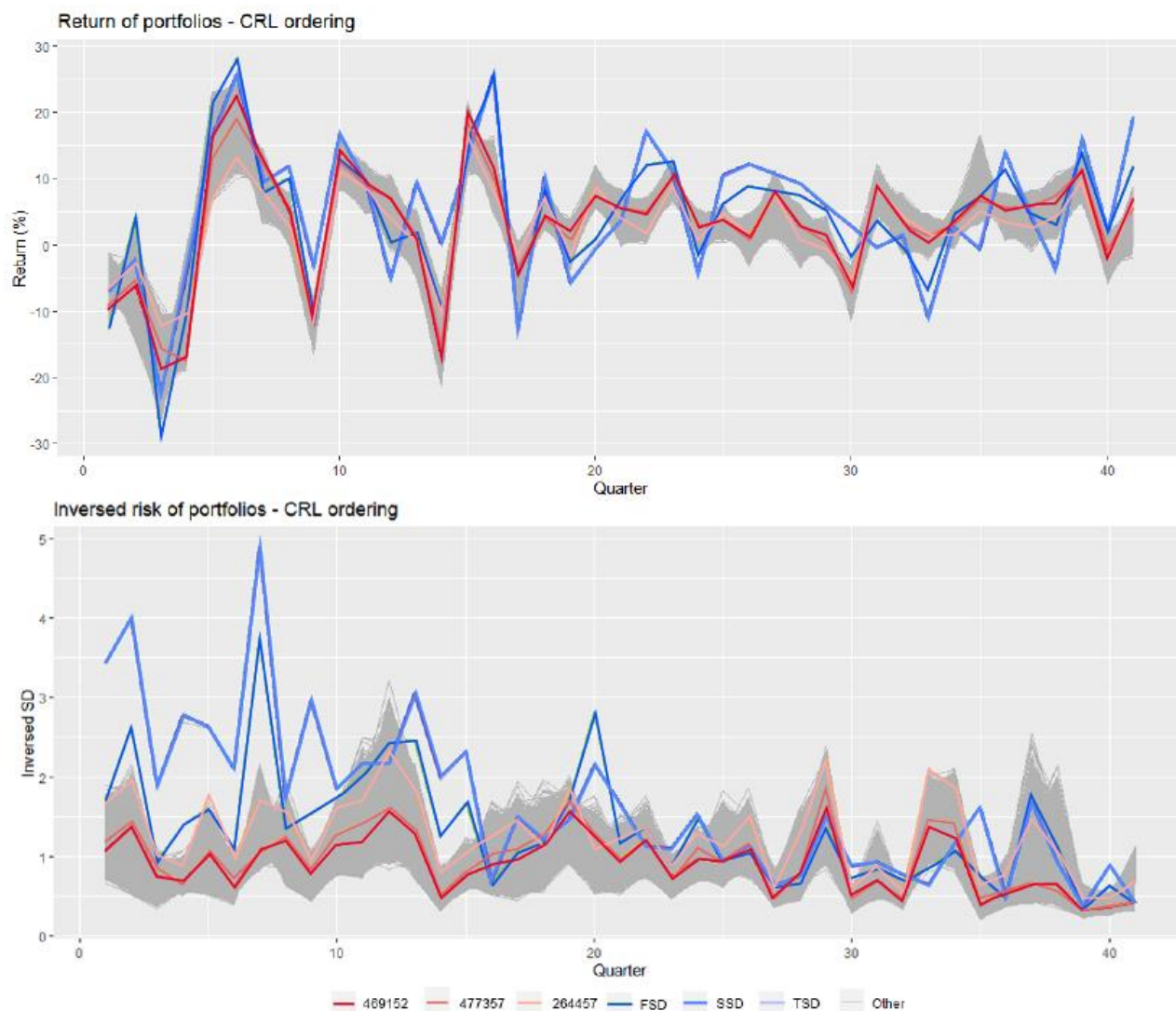
Source: Kozmik, 2019

One of the main advantages of the CRL metric is its ability to provide clear visualization. This allows us to plot all the functions selected by stochastic dominance and those identified as the best by our method within a single graph. This visual representation offers an intuitive comparison between the two approaches and highlights their respective selections.

Figure 1 presents our initial simulation of 500,000 portfolios, analyzed using the CRL metric. Although the number of simulated portfolios might seem relatively low given the 30 stocks involved, the results still offer valuable insights. In specific periods, such as quarters 32 to 35, our portfolios demonstrate notably higher returns and exhibit risk levels

comparable to those selected by stochastic dominance. However, in the earlier quarters, the portfolios chosen by SD outperform ours, particularly in terms of risk management.

Figure 1 First comparison of CRL and SD



Source: own processing

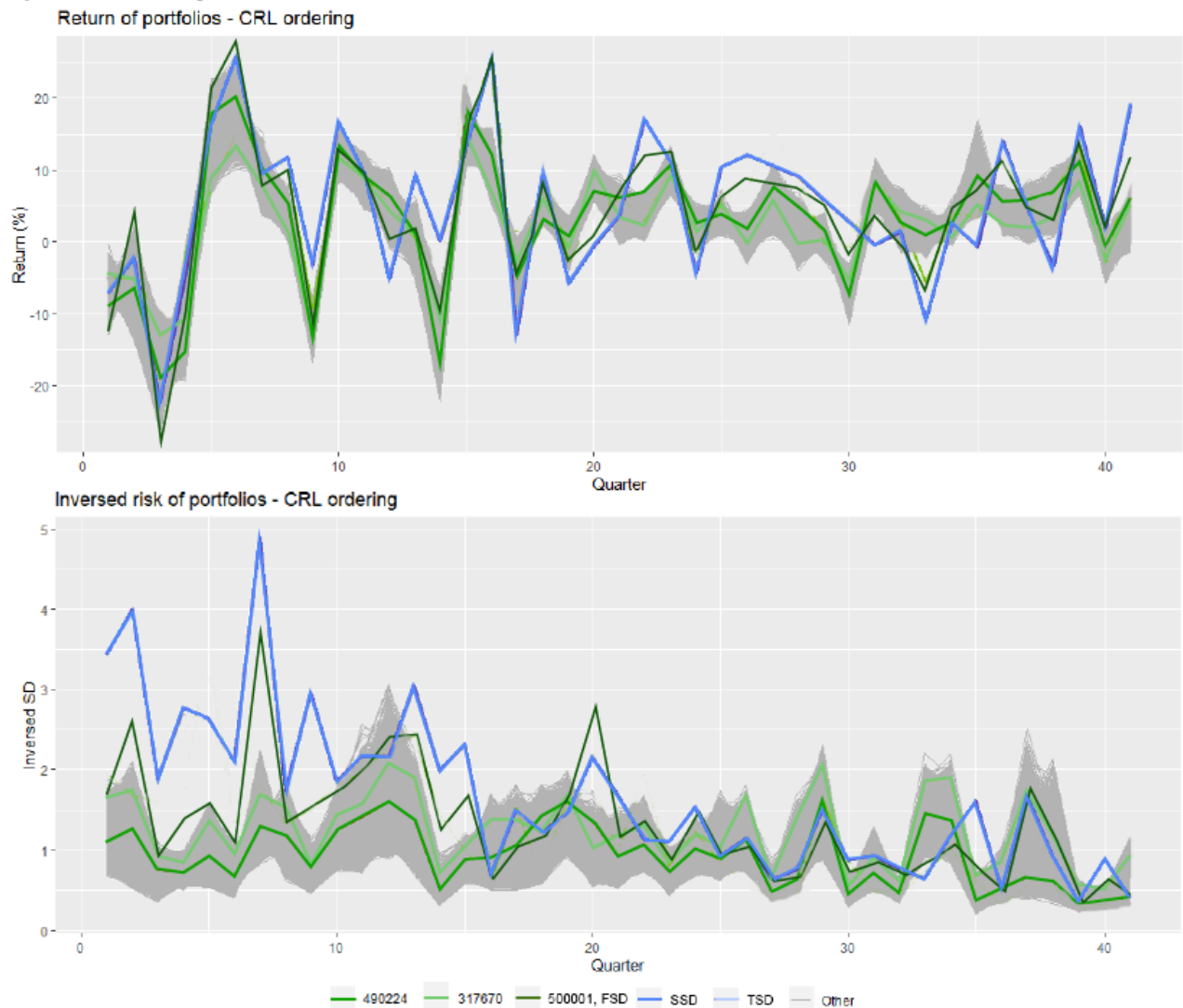
Table 3 displays the weight allocations for portfolios selected by the CRL method. Instead of including portfolio IDs, we used color labeling for clearer visualization. This table highlights a vital disadvantage of the CRL approach: the lack of constraints during simulation. As a result, every stock from the dataset is included, which may yield less efficient and optimal portfolios compared to the more selective nature of SD.

Table 3 Weights of three best portfolios by utilizing CRL metric (first simulation)

	AAPL	V	UNH	HD	NKE	BA	DD	MSFT	DIS	JPM	AXP	CAT	MCD	MMM	INTC
■	0,029	0,025	0,037	0,002	0,025	0,007	0,019	0,021	0,003	0,022	0,029	0,022	0,034	0,008	0,085
■	0,057	0,049	0,058	0,03	0,03	0,049	0,041	0,057	0,041	0,012	0,01	0,061	0,008	0,011	0,054
■	0,011	0,064	0,065	0,071	0,061	0,021	0,073	0,04	0,024	0,02	0,043	0,073	0,02	0,006	0,021
	TRV	PFE	WBA	JNJ	MRK	UTX	VZ	CVX	CSCO	GS	KO	WMT	IBM	PG	XOM
■	0,009	0,013	0,079	0,016	0,066	0,016	0,043	0,025	0,061	0,009	0,075	0,081	0,025	0,051	0,061
■	0,047	0,045	0,02	0,015	0,056	0,059	0,027	0,009	0,031	0,001	0,033	0,02	0,002	0,046	0,021
■	0,053	0,069	0,011	0,001	0,057	0,043	0,037	0,005	0,036	0,01	0,029	0,012	0,003	0,019	0,001

Source: own processing

Since a single simulation is insufficient for robust analysis, we conducted additional simulations to provide a more comprehensive comparison. Figure 2 presents another CRL-based simulation. Interestingly, CRL identified the FSD selected portfolio as the third best in this instance. Despite this overlap, the overall behavior of the portfolios selected by CRL compared to those chosen by stochastic dominance remains consistent: SD performs better in the early quarters of the dataset, while CRL-selected portfolios tend to dominate in the later quarters. This difference is largely due to AAPL's exceptional performance in the earlier quarters, which leads SD to prioritize it. In contrast, CRL is currently unable to focus on high-performing stocks like AAPL because our simulations are entirely random and lack a mechanism for targeted selection.

Figure 2 Second comparison of CRL and SD

Source: own processing

Table 4 presents the weight allocations for portfolios selected by the CRL method. The insights remain consistent with our previous observations: CRL's approach includes a wide range of stocks without constraints, which may lead to greater diversification but raises questions about the method's efficiency compared to stochastic dominance.

Table 4 Weights of three best portfolios by utilizing CRL metric (second simulation)

	AAPL	V	UNH	HD	NKE	BA	DD	MSFT	DIS	JPM	AXP	CAT	MCD	MMM	INTC
■	0,019	0,049	0,003	0,008	0,072	0,009	0,002	0,016	0,018	0,032	0,011	0,011	0,005	0,001	0,059
■	0,086	0,034	0,082	0,021	0,063	0,049	0,086	0,045	0,034	0,026	0	0,033	0,018	0,023	0,079
■	0,441	0,124	0,002	0,321	0	0	0	0	0,015	0,029	0	0,001	0,037	0	0
	TRV	PFE	WBA	JNJ	MRK	UTX	VZ	CVX	CSCO	GS	KO	WMT	IBM	PG	XOM
■	0,027	0,087	0,051	0,071	0,034	0,019	0,086	0,050	0,053	0,013	0,076	0,001	0,010	0,019	0,078
■	0,024	0,044	0,025	0,013	0,028	0,057	0,020	0,004	0,003	0,015	0,022	0,035	0,005	0,007	0,019
■	0,030	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Source: own processing

What is particularly compelling is the comparison of the basic characteristics of the selected portfolios. While stochastic dominance effectively identifies the portfolio with the highest potential value, the CRL method stands out for pinpointing portfolios that demonstrate greater stability over time. This attribute can be particularly valuable for investors prioritizing consistent performance and lower volatility, even if it means forgoing the highest possible returns associated with greater risk. However, it's worth noting that while FSD portfolios offer a similar level of stability, they come with less diversification. This raises the question of whether CRL's significantly higher computation time, which takes nearly a full day compared to less than a minute for SD, is justified for investors seeking a balance between strength and a well-diversified portfolio.

Table 5 Basic characteristics of all selected portfolios

ID	Mean (return)	Min (return)	Max (return)	Mean (risk inversed)
■ ■ FSD, 500001	5,17	-10	22,9	1,29
■ SSD	5,01	-22,40	25,8	0,86
■ TSD	5	-22,40	25,7	0,86
■ 490224	3,4	-19,10	20,27	1,31
■ 469152	3,38	-18,75	22,65	1,32
■ 477357	3,29	-17,48	19,24	1,25
■ 317670	2,34	-13	14,9	0,94
■ 264457	2,34	-14,43	14,43	0,95

Source: own processing

4 Conclusions

This study provides a comparison of the Continuous Rank Length (CRL) method and stochastic dominance (SD) in portfolio selection, highlighting their strengths and limitations. This case study did not validate our initial assumption that CRL might outperform SD due to SD's strict inequality constraints. While CRL can generate a wide variety of portfolios with similar distribution characteristics, we found it challenging to simulate better-performing portfolios.

Despite its computational intensity, the CRL method excels at identifying portfolios that exhibit consistent performance and stability over time. This feature can be particularly valuable to investors prioritizing lower volatility and consistent returns over pursuing the highest possible gains, which often come with higher risk. However, CRL's lack of constraints leads to the inclusion of all stocks in the dataset, raising concerns about efficiency compared to the more selective and computationally efficient SD approach. Stochastic dominance, meanwhile, remains effective at pinpointing optimal portfolios in terms of expected returns, especially early in the dataset, and maintains a clear focus on overall return distribution and risk.

Despite CRL's potential, the substantial computational demands present a significant drawback. While SD calculations take less than a minute, CRL simulations for 500,000 portfolios require nearly a full day. This inefficiency suggests several promising avenues for future research. For instance, optimizing CRL's simulation methods should be a priority. One improvement could involve prioritizing stocks with higher mean returns to streamline the process. Additionally,

incorporating time-weighted adjustments to the metrics could provide a more dynamic reflection of current market conditions, making CRL more adaptable and efficient.

Another intriguing area for exploration is the selection and combination of preferred stocks within CRL-generated portfolios. By strategically combining these preferred stocks with others, we might construct high-performing portfolios that leverage CRL's strengths while addressing its limitations. Further research should also investigate how CRL and SD perform under varying market scenarios, such as economic downturns or periods of high volatility, to evaluate each method's robustness and adaptability.

Ultimately, this research underscores the trade-offs between stability, risk, diversification, and computational efficiency in portfolio selection methods. While SD remains unbeaten in identifying high-value, low-risk portfolios, CRL's focus on stability presents a compelling alternative for risk-averse investors—provided that future developments can mitigate their computational demands.

Acknowledgment

The Grant Agency of South Bohemia University supported this work as an individual grant project under the number 031/2024/S and grant project 129/2022/S. I sincerely appreciate my colleague Jan Fesl for providing the essential computational resources and also my dearest colleague Jana Klicnarová for providing the essential insights.

References

- Kozmík, V. (2019). Robust approaches in portfolio optimization with stochastic dominance [Master's thesis, Charles University]. Charles University Digital Repository.
- Hadar, J., & Russell, W. R. (1969). Rules for ordering uncertain prospects. *The American Economic Review*, 59(1), 25–34.
- Hahn, U. (2015). "A Note on Simultaneous Monte Carlo Tests." Technical report, Centre for Stochastic Geometry and Advanced Bioimaging, Aarhus University.
- Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, 36(3), 335–346.
- Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3), 225–243.
- Levy, H. (1992). Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4), 555–593.
- Fishburn, P. C. (1974). Third-order stochastic dominance and risk. *Management Science*, 21(11), 1341–1346.
- Mrkvíčka, T., Myllymäki, M., Kuronen, M., & Narisetty, N. N. (2022). "New Methods for Multiple Testing in Permutation Inference for the General Linear Model." *Statistics in Medicine*, 41(2), 276–297. doi:10.1002/sim.9236.
- Myllymäki, M., & Mrkvíčka, T. (2020). "GET: Global Envelopes in R." arXiv:1911.06583 [stat.ME].
- Ryan JA, Ulrich JM (2024). *_quantmod: Quantitative Financial Modelling Framework_*. R package version 0.4.26, <<https://CRAN.R-project.org/package=quantmod>>.
- Stewart, S., Piros, C. D., & Heisler, J. (2019). *Portfolio management: Theory and practice*. John Wiley & Sons Inc.
- Whitmore, G. A., & Findlay, M. C. (1978). *Stochastic dominance: An approach to decision making under risk*. D.C. Heath and Company.