Investor Behavior and Portfolio Optimization: Implications from Stochastic Dominance and Prospect theory

Nikolai Sheung-Chi Chow

Australian National University

If you come across any problems, contact me at nikolaichow0809@gmail.com

To my family and friends

Preface

This all starts from a book titled "Economics Explanation" by Steven Ng-Sheong Cheung. This book, written in Chinese, explores different concepts in economics and the author's own thoughts on the subject. Cheung's book opens by arguing that it is not just life, but the world and nature at large, have patterns and are regulated by rules. He states that these patterns and rules can be understood and used to explain or predict phenomena. After being convinced, I began to have this desire to understand more about these principles, particularly those influencing human behavior, and understand their potential applications. This thesis is an attempt to satisfy that desire.

This thesis, "Investor Behavior and Portfolio Optimization: Implications from Stochastic Dominance and Prospect Theory," represents a journey, an intellectual odyssey filled with countless hours of work, moments of doubts and uncertainty. It stands as a testament to the collective wisdom, guidance, and encouragement of mentors, peers, and colleagues who have generously contributed their time and expertise towards this endeavor. This preface is an opportunity to acknowledge these invaluable contributions.

I would like to first express my deep appreciation to Raymond Honfu Chan, Simon Grant, Timothy Kam, Shu-Kam Lee, and Wing Keung Wong. Their continuous guidance and encouragement have been instrumental in shaping this work. Their wisdom, patience, and commitment to the highest standards of scholarship have inspired me throughout this journey. The depth of their knowledge and their commitment to teaching are truly inspiring.

Throughout my undergraduate journey, Dr. Shu-Kam Lee is my most important mentor. As the first individual to introduce me to the vast realm of academic research, Shu-Kam played a important role in shaping my scholarly trajectory. His encouragement and assistance have been instrumental in fostering my growth and academic pursuits.

Professor Alan Wing-Keung Wong has been a beacon of enlightenment, illuminating my path towards an international-level academic pursuit. It was through his guidance that I grasped the exquisite concept of "stochastic dominance," a concept that has since shaped my research focus. In a testament to his faith in my potential, he offered me an invaluable opportunity by extending an employment offer even before the completion of my master's degree. His belief in me has been a significant catalyst in my academic journey.

Professor Raymond Honfu Chan stands as an academic pillar from whom one can expect nothing but unwavering reliability. He showed me that a scholar's work can indeed influence and transform the world. A piece of wisdom he imparted resonates deeply with me to this day - ”Problems may appear complex, but by deconstructing them into smaller components, we may find a way to solve them.” This approach, rooted in the paradigm of ’first principles thinking,’ encourages the reduction of problems to their foundational elements. This methodology has since evolved into a fundamental cornerstone of my intellectual instincts.

I am incredibly fortunate to have Professor Simon Grant as my supervisor at Australian National University. His receptive attitude towards my unconventional ideas, along with his invaluable advice from various perspectives, significantly streamlined the process. For me, he is a figure representing the word “professional”, providing a benchmark that I strive to emulate. His relentless passion, emblematic of the most dedicated scholars, is a constant source of inspiration. His commitment to research is such that he often immerses himself in scholarly work during the early hours of the morning, typically between 2 and 5 am. This profound dedication underscores the indomitable pursuit of knowledge. Beyond this, his prowess in bridging the concepts of decision theory, microeconomics, and behavioral economics is unparalleled. His exceptional talent for fostering these connections not only broadens my academic perspective but also intensifies my dedication to my own scholarly pursuits.

Dr. Timothy Kam is another figure who has been instrumental in my journey. His comprehensive understanding of economics never ceases to amaze me, consistently sparking my curiosity and motivating me to delve deeper into the subject. His propensity for sharing invaluable information and insights has consistently guided me. Tim has always endeavored to be forthright, providing constructive and thoughtful suggestions. His honesty and consistent support have greatly enriched my academic pursuit. Beyond his intellectual prowess, his kindness and emotional intelligence stand as attributes I continuously strive to emulate.

I am indebted to John Stachurski and Jingni Yang. Professor John Stachurski is the one showing me that regardless of one's origins, it is sheer diligence and earnest dedication that can truly push the boundaries of any field. I learnt a lot from him, especially from some of his amazing books, which are the most user-friendly books you could find if you are interested in the subjects they cover. Dr. Jingni Yang, on the other hand, is an amazing friend. Jingni is also a source of constant inspiration. She has a never-give-up spirit that always motivates me to do my best. She has also been a valuable sounding board for my ideas. Her insightful suggestions and unyielding support have been invaluable. I am honoured to have her as a friend.

I also need to thank Han Bleichrodt, Haim Levy, Sam Ng and Tom Smith for their invaluable suggestions and constructive critiques on the part of the thesis, which have greatly enhanced the quality of this work. Their perspectives and insights have enriched my understanding and broadened my intellectual horizons, which has undoubtedly contributed to the quality of this thesis. I am also grateful for the enlightening discussions I have had with Rabah Amir, Damien Eldridge, Ruitian Lang, Frank Liu, Juergen Meinecke, Idione Meneghel, Ronald Stauber and others.

Lastly, I offer special thanks to my dear friends I met on this academic journey, including Yunshan Dai, Lawrence Guo , Shu Hu, Haiyan Lin, Wending Liu, Jianhua Mei, Sam Ng, Darapheak Tin, Kazuki Tomioko, Zeming Wang, Man Lung Yau, Chien Yeh, Thomas Yang, and others. Your friendship and support have been pillars during my journey.

In this journey, every conversation, every moment of shared insights, and every word of encouragement has played a pivotal role in making this thesis a reality. To all those mentioned and the countless others who have contributed to this journey in their own unique ways, I extend my heartfelt gratitude.

The journey towards the completion of this thesis has been challenging and rewarding, a rigorous intellectual exercise that has profoundly shaped me as a researcher. As I stand at this juncture, looking back at the journey undertaken and forward to the many adventures that await, I am filled with deep gratitude for the constellation of brilliant minds that guided me through this journey. This work is as much theirs as it is mine.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Stochastic Dominance (SD) is a compelling concept. It allows us to compare the expected utility of diverse risky options by just comparing the cumulative distributions (CDFs) of the corresponding random variables. This becomes particularly significant given the challenge posed by eliciting individuals' utility in numerous situations. In addition, by comparing CDFs, the conclusions reached are applicable to a class of utility functions, rather than being restricted to a specific function. As such, analysis based on SD demand fewer assumptions than those based on a singular utility function. Importantly, this means the conclusions we obtain from these analyses are likely to exhibit less disparity across different individuals.

An alternative famous approach, which shares the advantageous characteristics discussed above, is to formulate an analysis based on the Mean-Variance (MV) framework. MV analysis, developed by Harry Markowitz, allows one to evaluate an investment by just examining its expected return and variance of returns. Nevertheless, utilizing MV necessitates additional assumptions relative to SD, either on the distribution or the utility function.

Despite its merits, SD is not without its assumptions. In particular, the second-order SD assumes risk aversion. This leads us to question the validity of assuming humans are risk-averse everywhere. In practice, the implications of this assumption do not always align with our lived experiences, which has sparked significant debate within the spheres of behavioral economics and finance. In this thesis, we delve into the influence of the behavioral economics and finance revolution, with a particular emphasis on the prospect theory developed by Daniel Kahneman and Amos Tversky.

This thesis explores the aspects of investor behavior and portfolio optimization. We will start by discussing some implications and applications of SD, followed by discussions on incorporating concepts in prospect theory with SD and its implications.

Chapter 2 begins our exploration into the connections between Central Moments and different SD concepts. The chapter further provides a comprehensive analysis of diversification as a practical application to illustrate these connections. An inquiry of general and polynomial utility functions provides the theoretical backbone for these ideas. The chapter then addresses the moment rule tests and further elaborates on the importance of these ideas.

Our investigation then shifts to an empirical phenomenon known as the Eastern Halloween effect, through the lens of Stochastic Dominance in Chapter 3. An extensive literature review is followed by an analysis of data, with specific focus on the Halloween period and non-Halloween period. The chapter delves into international markets and explains the Eastern Halloween effect's risk implications, tying this phenomenon back to the concepts of market efficiency and stochastic dominance.

In Chapter 4, we revisit the subject of portfolio optimization. We show the connections between Partial Moments, Mean-Variance Analysis, and Prospect Theory. The relationships between Prospect Stochastic Dominance (PSD) and moments are examined, along with the impact on Mean-Variance Analysis. The chapter concludes by presenting a method for Portfolio Optimization based on PSD.

The penultimate chapter, Chapter 5, extends the discussion by focusing on loss aversion, a key characteristic identified in Prospect Theory. It outlines the implications on moments and Portfolio Choice in the context of loss aversion. In Chapter 4 and 5, linear probability transformations are assumed throughout

Finally, the thesis concludes with Chapter 6, which offers a synthesis of the topics discussed and implications for future research.

Overall, this thesis aims to provide a rigorous, comprehensive exploration of investor behavior and portfolio optimization, uncovering valuable insights into the practical and theoretical implications of Stochastic Dominance and Prospect Theory. Through these interconnected inquiries, this thesis aims to bridge the gap between traditional portfolio theories and the complexities of investor behavior. It also strives to generate insights that can be readily applied by portfolio managers and other financial practitioners, enhancing the efficiency of investment decision-making processes and contributing to better financial outcomes for clients.

Ultimately, the work undertaken in this thesis is a testament to the ever-evolving field of finance, where classical theories and contemporary behavioral finance intersect to offer a more comprehensive understanding of the market and investor behavior. By embracing this multidisciplinary approach, the thesis aspires to enrich the dialogue within academia and practice alike, forging a path towards a more nuanced understanding of the dynamic world of finance.

Chapter 2

Central Moments, Stochastic Dominance, Moment Rule, and Diversification with an Application

Joint work with Raymond Chan, Xu Guo, & Wing-Keung Wong

Published in Chaos, Solitons & Fractals

In this chapter, we first develop some properties to state the relationships among central moments, stochastic dominance (SD), risk-seeking stochastic dominance (RSD), and integrals for the general utility functions and the polynomial utility functions of both risk averters and risk seekers. We then introduce the moment rule and establish some necessary and/or sufficient conditions between stochastic dominance and the moment rule for the general utility functions and the polynomial utility functions of both risk averters and risk seekers without imposing the same-location-scale-family condition. Thereafter, we apply the moment rules to develop some properties of portfolio diversification for the general utility functions and the polynomial utility functions for both risk averters and risk seekers. The findings in our paper enable academics and practitioners to draw preferences of both risk averters and risk seekers on their choices of portfolios or assets by using different moments. We illustrate this by using the moment rule tests to compare excess return of 49 industry portfolios from Kenneth French's online data library.

2.1 Overview

Lehmann (1951, 1952, 1955) first compare several sets of distributions that lead to the development of the theory of stochastic dominance (SD). Afterwards, Hadar and Russell (1969, 1971), Hanoch and Levy (1969), and others establish some basic relationships between SD and the preference of different assets to the second order. Extensions of the SD theory to the higher-order include Whitmore (1970), Ekern (1980), Bawa, *et al.* (1985), Muliere and Scarsini (1989), Wong and Li (1999), Li and Wong (1999), Denuit, *et al.* (2013), Post and Koppa (2013), Fang and Post (2017, 2022), Post and Kopa (2017) and many others. Jean (1975) is one of the first few papers that express the moments in terms of successive integrals of a probability density function so that the SD rankings can be compared with moment rankings. Other studies, for example, Brockett and Garven (1998), establish some properties for the relationship between risk, return, skewness, and utility preferences. Studying the high-order moments is important because many studies, for example, Kraus and Litzenberger (1976), Scott and Horvath (1980), Astebro (2003), Cvitanić, Polimenis, and Zapatero (2008), Choi and Nam (2008), Chiu (2010), and Astebro, Mata, and Santos-Pinto (2015) have found that preference of high-order moments plays an important role in asset pricing and many other areas in finance and economics. In this chapter, we first extend their work by developing some properties to state the relationships between the n^{th}-order (central) moments, the n^{th}-order SD, the n^{th}-order risk-seeking SD (RSD), and the n^{th}-order [reversed] integrals for both n^{th}- and $(n+1)^{th}$-order [R]SD for general risk-averse [risk-seeking] utility functions and the polynomial utility functions of both risk averters and risk seekers for any order n, including $n = 2, 3$, and 4 as the special cases.

Markowitz (1952a) first introduces the mean-variance (MV) rule for risk averters, and Wong (2007) and others introduce the MV rule for risk seekers because it is well-known that the MV rule for risk averters cannot handle some situations, see, for example, Copeland *et al.* (2005) and Levy (2015) for more information. Studies like Brockett and Garven (1998), and Meyer, Li, and Rose (2005) prefer other rules than the MV rule. For example, Brockett and Garven (1998) suggest that the SD rule is better than the MV rule. One limitation of the MV rule is that it has not measured any information from high moments, while measuring information from high moments is very important in many empirical data.[1] To circumvent the limitations of the MV rule, In this chapter, we extend the MV rule by introducing the moment rules to acquire information from higher moments in the comparison, including the mean-variance-skewness and mean-variance-skewness-kurtosis rules for both risk averters and risk seekers. As far as we know, our paper is the first paper in the literature to introduce the moment rules for both risk averters and risk seekers by establishing some necessary and/or sufficient conditions between SD and the moment rule under some conditions, inferring that the moment rule could be as good as the SD rule under some conditions. In addition, we extend the theory further by removing the same-location-scale-family condition to establish some necessary conditions between SD and the moment rule for both risk averters and risk seekers under some conditions.

Diversification is one of the most important areas in finance. Many studies[2] have developed some properties of diversification related to the MV and SD rules. In this chapter, we extend their theories by applying the moment rules to develop some properties of portfolio diversification to compare the preferences of two sets of assets for the general utility functions and the polynomial utility functions of both risk averters and risk seekers. In addition, applying the moment rules to develop some properties of portfolio diversification to compare the preferences between an individual asset, a completely diversified portfolio, and a partially diversified portfolio. We also develop some properties of portfolio diversification that could allow us to compare preferences of some pairs of partially diversified portfolios for any set of independent assets or some sets of dependent assets and the general utility functions of both risk averters and risk seekers. The findings in our paper enable academics and practitioners to draw preferences of both risk averters and risk seekers on their choices of portfolios or assets by using different moments. We illustrate the applications of the moment rules introduced In this chapter by using the tests of the moment rules to compare the excess return of 49 industry portfolios from Kenneth French's online data library. We find that the results are reasonably stable from Jan 1992 to Dec 2021. First, around 30% of the portfolios that are dominated by an industry portfolio under a moment rule are also dominated by an industry portfolio under

[1] See, for example, Lim (1989), Perez-Quiros and Timmermann (2001), Jondeau and Rockinger (2003, 2006), Cvitanić, Polimenis, and Zapatero (2008), Choi and Nam (2008), Harvey, Liechty, Liechty, and Muller (2010), Grigoletto and Lisi (2011), Buckle, Chen, and Williams (2016), and Do, Brooks, Treepongkaruna, and Wu (2016) for more information.

[2] See, for example, Hadar and Russell (1971), Tesfatsion (1976), Li and Wong (1999), Wong (2007), Egozcue and Wong (2010), Guo and Wong (2016), and Chan, *et al.* (2020) for more information.

a moment rule from Jan 2002 to Dec 2011. Second, around 50% of the portfolios that are dominated by an industry portfolio under a moment rule from Jan 2002 to Dec 2011 are also dominated by an industry portfolio under a moment rule in the period from Jan 2012 to Dec 2021. Third, around 50% of the portfolios that are dominated by an industry portfolio under a moment rule from Jan 1992 to Dec 2001 are also dominated by an industry portfolio under a moment rule from Jan 2012 to Dec 2021.

The paper is organized as follows. We introduce some definitions and notations in the next section. Section 2.3 develops some properties on the relationships among central moments, stochastic dominance, and expected utility. Section 2.4 introduces the moment rule and develops some properties for the moment rule. Section 2.5 develops some properties of portfolio diversification for the general utility functions and the polynomial utility functions of both risk averters and risk seekers. Section 2.6 develops some properties for the preferences between some partially-diversified portfolios. Section 2.7 provides testing procedures for the moment rule. Section 2.8 illustrates the applicability of the theory developed in our paper by using real-life data. Section 2.9 concludes our findings.

2.2 Definitions and notations

Let $\overline{\mathbb{R}}$ be the set of extended real numbers and $\Omega = [a, b]$ be a subset of $\overline{\mathbb{R}}$ in which $a < b$. For random variable $Z = X$ and Y with "*cumulative distribution function*" (CDF) $H = F$ and G and probability density function $h = f$ and g, the mean of Z is defined as $\mu_Z = \mu_H$. The CDF $H_1(x) \equiv H(x) \equiv \mu[a, x]$ of the measure μ is defined on the support $\Omega = [a, b] \subset \overline{\mathbb{R}}$ with $\mu(\Omega) = 1$. For any integer $n > 1$, we define the following:

$$C_H^{(n)} = \int_a^b (x - \mu_H)^n dH(x) \ , \quad H_n(x) = \int_a^x H_{n-1}(t)\, dt \quad , \quad H_n^R(x) = \int_x^b H_{n-1}^R(t)\, dt \ . \qquad (2.2.1)$$

We note that $H_0(x) = H_0^R(x) = h(x)$, $C_H^{(2)} = \sigma_H^2$ is the variance of H, and $C_H^{(n)}$ is the n^{th}-order central moment for any integer $n \geq 2$. We also note that for any variable Z with CDF H, we will use both $C_Z^{(n)} = C_H^{(n)}$ to be the n^{th}-order central moment of Z for any integer $n \geq 2$. We further let γ_H as the skewness, and κ_H as the kurtosis of H respectively. We call H_n (H_n^R) the n^{th}-order (reversed) integral. We note that H_n is used in the development of the SD theory for risk averters while H_n^R is used in the development of the SD theory for risk seekers, see, for example, Quirk and Saposnik (1962), Hanoch and Levy (1969), Hammond (1974), Levy (2015), and Guo and Wong (2016) and the references therein for more information.

In this chapter, we first extend the MV rule[3] for both risk averters and risk seekers introduced by Markowitz (1952a), Wong (2007), and others. It is well-known that the mean-variance rule for risk averters cannot handle some situations, see, for example, the paradox used in the example in p66-67 of Copeland, *et al.* (2005). We modify the example as follows:

Example 2.1. *For any pair of two prospects, X and Y, with probability functions, P_X and P_Y, respectively, such that $P_X(z) = 0.2$ for $z = 3, 5, 7, 9,$ and 11, and $P_Y(z) = 0.2$ for $z = 3, 4, 5, 6,$ and 7.*

One could easily find that X and Y as stated in Example 2.1 do not dominate each other by using the MV rule for risk averters. However, it is obvious that X is preferred to Y. To solve the problem, one could apply the MV rule for risk seekers introduced by Wong (2007) and others. Now, when one applies the mean-variance rule for risk seekers, one will conclude that X dominates Y, and thus, conclude that risk seekers will prefer X to Y. However, there are still some limitations of using the mean-variance rules for both risk averters and risk seekers because by using the rules, one could only conclude that risk seekers will prefer X to Y but cannot conclude that risk averters will also prefer X to Y in the paradox as stated in Copeland, *et al.* (2005) and Example 2.1 but, in fact, it is well-known that both risk averters and risk seekers will prefer X to Y in this example. To circumvent the limitation, a well-known solution is to apply the theory of stochastic dominance (SD). To do so, Levy (2015), Guo and Wong (2016), Bai, *et al.* (2021), and others define the n^{th}-order SD. We state the rule briefly here. One may refer to Guo and Wong (2016) for the full definition of the rules.

Definition 2.1. *For any integer n and for any pair of random variables, X and Y, with CDFs, F and G, respectively,*

1. *X is said to dominate Y by the n^{th}-order stochastic dominance for risk averters for $n \geq 1$, denoted by $X \succeq_n Y$, or $F \succeq_n G$ if and only if $F_n(x) \leq G_n(x)$ for each x and $F_k(b) \leq G_k(b)$ for $k = 1, \cdots, n-1$ if $n > 1$, and*

[3]One may refer to Definition 8 in Wong (2007) for the rule.

2. *X is said to dominate Y by the n^{th}-order risk-seeking stochastic dominance (RSD)*[4] *for $n \geq 1$, denoted by $X \succeq_n^R Y$ or by $F \succeq_n^R G$, if and only if $F_n^R(x) \geq G_n^R(x)$ for each x in $[a, b]$, and $F_k^R(b) \geq G_k^R(b)$ for $k = 1, \cdots, n-1$ if $n > 1$.*

Applying Definition 2.1 to the paradox as stated in Copeland, *et al.* (2005) and Example 2.1, one will conclude that both risk averters and risk seekers will prefer X to Y and, in fact, conclude that all investors with increasing utility, including both risk averters and risk seekers, will prefer X to Y because we have both $X \succeq_1 Y$ and $X \succeq_1^R Y$.

Stochastic dominance is useful for ranking prospects with uncertainty because ranking prospects is equivalent to maximizing the expected utility preferences.[5] To show the advantages of applying SD, we first define utility functions for risk averters and risk seekers (Levy, 2015; Guo and Wong, 2016; Bai, *et al.,* 2020) as the following:

Definition 2.2. *Sets of utility functions, U_n and U_n^R , for risk averters and risk seekers are:*

$$U_n = \{u : (-1)^i u^{(i)} \leq 0 \, , \, i = 1, \cdots, n\} \quad \textit{and} \quad U_n^R = \{u : u^{(i)} \geq 0 \, , \, i = 1, \cdots, n\}, \tag{2.2.2}$$

respective, where $u^{(i)}$ is the i^{th} derivative of the utility function u.

In addition, in this chapter, we will develop a theory related to the n^{th}-order polynomial utility function. For this purpose, we define the n^{th}-order polynomial utility function as the following:

Definition 2.3. *If $u \in U_n$ or U_n^R and $u^{(n)}$ is a nonzero constant, then $u \in U_{np}$ or U_{np}^R is a n^{th}-order polynomial utility function for risk averters and risk seekers, respectively.*

We note that it is easy to extend the theory to include non-differentiable utilities.[6] For any investor with $u \in U_n$, it is well known that a negative second derivative for the utility function infers that investors are risk-averse and a positive third derivative for the utility function is a necessary, but not sufficient condition for decreasing absolute risk aversion (DARA).

So far, it is well known that the SD rule is more superior to the MV rule. Thus, this paper aims to improve the MV rule. To do so, we extend the MV rule to be the moment rule and we will

[4]We note that Levy (2015) and others call it RSSD while we follow Guo and Wong (2016) and others to call it RSD.

[5]We follow von Neumann-Morgenstern (1944) to compare the preference among prospects. For an empirical test on SD see Post(2003)

[6]See Wong and Ma (2008).

study the relationships between the n^{th}-order central moments with the n^{th}-order (reversed) integrals, the n^{th}-order (risk-seeking) SD, and the moment rule. We discuss the theory in the next section.

2.3 Theory

In this section, we develop some properties on the relationships among central moments, stochastic dominance, and expected utility. These properties have implications for the moment rule in Section 2.4 and portfolio diversification in Section 2.5. We first develop some properties for general utility functions defined in Definition 2.2 and develop properties for polynomial utility functions thereafter in this section.

2.3.1 General utility functions

We first develop some properties on the relationships among central moments, stochastic dominance, and expected utility for general utility functions defined in Definition 2.2. Since

$$\begin{aligned} \mu_H &= \int_a^b t dH(t) = tH(t)\Big|_a^b - \int_a^b H(t)dt \\ &= b - H_2(b) = a + H_2^R(a) \ , \end{aligned} \tag{2.3.1}$$

where $H = F$ or G, we have

$$\mu_F - \mu_G \;=\; G_2(b) - F_2(b) = F_2^R(a) - G_2^R(a). \tag{2.3.2}$$

Chan, *et al.* (2020) extend the work by Jean (1975) and others by establishing the following result:

$$G_3(b) - F_3(b) = G_3^R(a) - F_3^R(a) = \frac{1}{2}\left(\sigma_G^2 - \sigma_F^2\right) \ , \tag{2.3.3}$$

given $\mu_F = \mu_G$.

In this chapter, we first extend their results by establishing the following theorem to examine the relationship between the n^{th}-order (central) moments and the n^{th}-order integrals for both n^{th}- and $(n+1)^{th}$-order SD:

Theorem 2.4. *Let $C_H^{(k)}$ be the k^{th}-order central moment for any integer $k \geq 2$. For any given $n \geq 2$, if $C_F^{(k)} = C_G^{(k)}$ for all $2 \leq k < n$ and $\mu_F = \mu_G$, the following statements are equivalent:*

1.
$$G_{n+1}(b) - F_{n+1}(b) \;=\; \frac{(-1)^n}{n!}\left(C_G^{(n)} - C_F^{(n)}\right) \; ; \tag{2.3.4}$$

2. *$G_{n+1}(b) \geq F_{n+1}(b)$ if and only if $(-1)^n C_G^{(n)} \geq (-1)^n C_F^{(n)}$;*

3. *If $F \succeq_n G$, then $(-1)^n C_F^{(n)} < (-1)^n C_G^{(n)}$; and*

4. *If $F \succeq_{n+1} G$, then $(-1)^n C_F^{(n)} \leq (-1)^n C_G^{(n)}$.*

The proof of Theorem 2.4 is shown in the appendix. Since $F \succeq_n G$ implies $F \succeq_{n+1} G$, one may believe that it is not necessary to have Part 4 of Theorem 2.4. We note that this is not true because Part 4 of Theorem 2.4 includes the case in which $F \succeq_n G$ does not hold but we still have both $F \succeq_{n+1} G$ and $(-1)^n C_G^{(n)} \geq (-1)^n C_F^{(n)}$. Part 3 of Theorem 2.4 does not cover this situation. Thus, we still require to have Part 4 in Theorem 2.4. We note that Fishburn (1980) has derived Part 3 of Theorem 2.4. From Theorem 2.4, we can obtain the following corollary to compare the second-order central moments and SD:

Corollary 2.5. *If $\mu_F = \mu_G$, then*

1. *$G_3(b) \geq F_3(b)$ if and only if $\sigma_F^2 \leq \sigma_G^2$;*

2. *if $F \succeq_2 G$, then $\sigma_F^2 < \sigma_G^2$; and*

3. *if $F \succeq_3 G$, then $\sigma_F^2 \leq \sigma_G^2$,*

where σ_H^2 is the variance of H for $H = F$ or G.

We note that Equation (2.3.3) is a special case of Theorem 2.4. Now, we turn to develop the following corollary to compare the third-order central moments and SD:

Corollary 2.6. *If $\mu_F = \mu_G$ and $\sigma_F^2 = \sigma_G^2$, then*

1. *$G_4(b) \geq F_4(b)$ if and only if $\gamma_G \leq \gamma_F$;*

2. *if $F \succeq_3 G$, then $\gamma_F > \gamma_G$; and*

3. *if $F \succeq_4 G$, then $\gamma_F \geq \gamma_G$,*

where γ_H is the skewness of H for $H = F$ or G.

Part 1 of Corollary 2.6 shows the necessary and sufficient conditions for the magnitude of the skewnesses with that of the fourth-order integrals for any two distributions, while Part 2 of Corollary 2.6 shows the relationship of the skewnesses with the third-order SD of any two distributions under the conditions of both equal mean and equal variance. From Theorem 2.4, we can also obtain the following corollary to show the relationships between the SD and the kurtosises for any two distributions under the conditions of equal mean, equal variance, and equal skewness:

Corollary 2.7. *If $\mu_F = \mu_G$, $\sigma_F^2 = \sigma_G^2$, and $\gamma_G = \gamma_F$, then*

1. *$G_5(b) \geq F_5(b)$ if and only if $\kappa_G \geq \kappa_F$;*

2. *if $F \succeq_4 G$, then $\kappa_G > \kappa_F$; and*

3. *if $F \succeq_5 G$, then $\kappa_G \geq \kappa_F$,*

where κ_H is the kurtosis of H for $H = F$ or G.

Part 1 of Corollary 2.7 shows the necessary and sufficient conditions for the magnitude of the kurtosises with that of the fifth-order integrals for any two distributions, while Part 2 of Corollary 2.6 shows the relationship between the kurtosises and the fourth-order SD for two different distributions under the conditions of equal mean, equal variance, and equal skewness.

We turn to develop the following theorem as a complement of Theorem 2.4 to show the relationship among the n^{th}-order (central) moments, the n^{th}-order reversed integrals, and the n^{th}- and $(n+1)^{th}$-order RSD:

Theorem 2.8. *Let $C_H^{(k)}$ be the k^{th}-order central moment for any integer $k \geq 2$. For any given $n \geq 2$, if $C_F^{(k)} = C_G^{(k)}$ for all $2 \leq k < n$ and $\mu_F = \mu_G$, the following statements are equivalent:*

1.
$$F_{n+1}^R(a) - G_{n+1}^R(a) \;=\; \frac{1}{n!}\left(C_F^{(n)} - C_G^{(n)}\right) \; ; \tag{2.3.5}$$

2. $F_{n+1}^R(a) \geq G_{n+1}^R(a)$ *if and only if* $C_F^{(n)} \geq C_G^{(n)}$;
3. *If* $F \succeq_n^R G$, *then* $C_F^{(n)} > C_G^{(n)}$; *and*
4. *If* $F \succeq_{n+1}^R G$, *then* $C_F^{(n)} \geq C_G^{(n)}$.

The proof of Theorem 2.8 is shown in the appendix. Similar to Theorem 2.4, since $F \succeq_n^R G$ implies $F \succeq_{n+1}^R G$, one may believe that it is not necessary to have Part 4 of Theorem 2.8. We note that this is not true because Part 4 of Theorem 2.8 includes the case in which $F \succeq_n^R G$ does not hold but we still have both $F \succeq_{n+1}^R G$ and $(-1)^n C_G^{(n)} \geq (-1)^n C_F^{(n)}$. From Theorem 2.8, we obtain the following corollary to compare the second-order central moments with RSD:

Corollary 2.9. *If $\mu_F = \mu_G$, then*

1. $F_3^R(a) \geq G_3^R(a)$ *if and only if* $\sigma_F^2 \geq \sigma_G^2$;
2. *if* $F \succeq_2^R G$, *then* $\sigma_F^2 > \sigma_G^2$;
3. *if* $F \succeq_3^R G$, *then* $\sigma_F^2 \geq \sigma_G^2$; *and*

where σ_H^2 is the skewness of H for $H = F$ or G.

We note that Equation (2.3.3) is a special case of both Theorems 2.4 and 2.8.

Now, we turn to develop the following corollary to compare the third-order central moments with RSDs:

Corollary 2.10. *If $\mu_F = \mu_G$ and $\sigma_F^2 = \sigma_G^2$, then we have*

1. $F_4^R(a) \geq G_4^R(a)$ *if and only if* $\gamma_F \geq \gamma_G$;

2. *if* $F \succeq_3^R G$, *then* $\gamma_F > \gamma_G$; *and*

3. *if* $F \succeq_4^R G$, *then* $\gamma_F \geq \gamma_G$,

where γ_H *is the skewness of* H *for* $H = F$ *or* G.

Part 1 of Corollary 2.10 shows the necessary and sufficient conditions on the magnitude of the skewnesses with the fourth-order reverse integrals for any two distributions while Part 2 of Corollary 2.10 shows the relationship of the skewnesses with the third-order RSD for any two distributions under the conditions of equal mean and equal variance.

From Theorem 2.8, we can also obtain the following corollary to show the relationships between the kurtosises and RSDs for any two distributions under the conditions of equal means, equal variances, and equal skewnesses:

Corollary 2.11. *If* $\mu_F = \mu_G$, $\sigma_F^2 = \sigma_G^2$, *and* $\gamma_G = \gamma_F$, *then we have*

1. $F_5^R(a) \geq G_5^R(a)$ *if and only if* $\kappa_F \geq \kappa_G$;

2. *if* $F \succeq_4^R G$, *then* $\kappa_F > \kappa_G$; *and*

3. *if* $F \succeq_5^R G$, *then* $\kappa_F \geq \kappa_G$;

where κ_H *is the kurtosis of* H *for* $H = F$ *or* G.

Part 1 of Corollary 2.10 shows the necessary and sufficient conditions on the magnitude of the kurtosises for any two distributions with the fifth-order reverse integrals while Part 2 of Corollary 2.10 shows the relationship of the kurtosises for two different distributions with the fourth-order RSD under the conditions of equal mean, variance, and skewness.

Guo *et al.* (2014) and others comment that there is no equivalence relationship between moments and SD. In Section 2.3.1, we find the sufficient condition but not the necessary condition from SD to moment. Is it possible to get the necessary condition, and thus, obtain the equivalence relationship between moments and SD? We explore the answer in the next section.

2.3.2 Polynomial utility functions

We turn to develop some properties on relationships among central moments, stochastic dominance, and expected utility for the polynomial utility functions defined in Definition 2.3. We first develop the following theorem as a complement of Theorem 2.4 to establish some relationships between the n^{th}-order (central) moments and n^{th}-order SD under n^{th}-order polynomial utility functions.

Theorem 2.12. *Let $C_H^{(k)}$ be the k^{th}-order central moment for any integer $k \geq 2$. For any given $n \geq 2$, if $C_F^{(k)} = C_G^{(k)}$ for all $2 \leq k < n$, and $\mu_F = \mu_G$, then for all n^{th}-order polynomial utility function $u \in U_{np}$, the following statements are equivalent:*

1. $F \succeq_n G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $(-1)^n C_F^{(n)} \leq (-1)^n C_G^{(n)}$.

By applying Theorem 2.4, one could conclude that $G_{n+1}(b) \geq F_{n+1}(b)$ is equivalent to $(-1)^n C_G^{(n)} \geq (-1)^n C_F^{(n)}$ under the assumption of the theorem. Thus, we just need to prove that $G_{n+1}(b) \geq F_{n+1}(b)$ is equivalent to $Eu(F) \geq Eu(G)$ under the assumption of the theorem. The proof of Theorem 2.12 is shown in appendix. Thereafter, by using both Theorem 2.12 and Equation (2.3.2), we can obtain the following corollaries:

Corollary 2.13. *Suppose $\mu_F = \mu_G$. For any quadratic utility function $u \in U_{2p}$, the following statements are equivalent*

1. $F \succeq_2 G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $\sigma_F^2 \leq \sigma_G^2$.

We now construct a simple counterexample to show that if the utility function u is cubic, even $\mu_F = \mu_G = \mu$ and $\sigma_F^2 \leq \sigma_G^2$, we may still have $Eu(F) < Eu(G)$.

Example 2.2. *For a cubic utility function, we have:*

$$u(x) = u(\mu) + u'(\mu)(x-\mu) + \frac{u''(\mu)}{2}(x-\mu)^2 + \frac{u'''(\mu)}{6}(x-\mu)^3.$$

As a result, we have:

$$Eu(F) = u(\mu) + \frac{u''(\mu)}{2}\sigma_F^2 + \frac{u'''(\mu)}{6}C_F^{(3)}.$$

Then, we get:

$$Eu(F) - Eu(G) = \frac{u''(\mu)}{2}(\sigma_F^2 - \sigma_G^2) + \frac{u'''(\mu)}{6}(C_F^{(3)} - C_G^{(3)}).$$

This implies that the sign of $Eu(F) - Eu(G)$ depends on the sign of $u''(\mu), u'''(\mu)$, $\sigma_F^2 - \sigma_G^2$, and $C_F^{(3)} - C_G^{(3)}$. Although under the assumed condition, the first term $\frac{u''(\mu)}{2}(\sigma_F^2 - \sigma_G^2)$ is positive, the second term may be negative and this would finally lead the whole term $Eu(F) - Eu(G)$ be negative or positive. For example, consider $u''(\mu) = -2, u'''(\mu) = 6, \sigma_F^2 = 1, \sigma_G^2 = 2, C_F^{(3)} = 1, C_G^{(3)} = 2.5$, we get $Eu(F) - Eu(G) = -0.5 < 0$.

From Theorem 2.12, we obtain the following corollary for any cubic utility function:

Corollary 2.14. *Suppose $\mu_F = \mu_G$ and $\sigma_F^2 = \sigma_G^2$. For any cubic utility function $u \in U_{3p}$, the following statements are equivalent*

1. *$F \succeq_3 G$,*
2. *$Eu(F) \geq Eu(G)$, and*
3. *$\gamma_F \geq \gamma_G$.*

From Theorem 2.12, we obtain the following corollary for any quartic utility function:

Corollary 2.15. *Suppose $\mu_F = \mu_G$, $\sigma_F^2 = \sigma_G^2$, and $\gamma_G = \gamma_F$. For any quartic utility function $u \in U_{4p}$, the following statements are equivalent*

1. *$F \succeq_4 G$,*
2. *$Eu(F) \geq Eu(G)$, and*

3. $\kappa_F \leq \kappa_G$.

Corollary 2.13 tells us that if the means are equal, then under the quadratic utility functions, the preference of the second-order SD is equivalent to the preference (smaller) of the variance (for risk averters); Corollary 2.14 tells us that if the means and variance are equal, then under the cubic utility functions, the preference of the third-order SD is equivalent to the preference (bigger) of the skewness (for risk averters); Corollary 2.15 tells us that if the means, variance, and skewness are equal, then under the quartic utility functions, the preference of the fourth-order SD is equivalent to the preference (smaller) of the kurtosis (for risk averters). In general, Theorem 2.12 tells us that under the conditions that all moments less than n are equal for two assets and under the n^{th}-order polynomial utility functions, then the preference of the n^{th}-order SD is equivalent to the preference of the n^{th}-order central moments (for risk averters). We note that this is very strong result. So far, in literature, for example, Guo *et al.* (2014) comment that there is no equivalence relationship between moments and SD and we believe that our paper is the first paper finds that under some conditions (that is, all moments less than n are equal for two assets and under polynomial utility functions), then the preference of the n^{th}-order SD is equivalent to the preference of the n^{th}-order central moments (for risk averters). In this chapter, we hypothesize that this result only holds for polynomial utility functions, not any other nontrivial utility functions. Now, we turn to develop the results for convex utility functions. We first state the following theorem:

Theorem 2.16. *Let $C_H^{(k)}$ be the k^{th}-order central moment for any integer $k \geq 2$. For any given $n \geq 2$, if $C_F^{(k)} = C_G^{(k)}$ for all $2 \leq k < n$, and $\mu_F = \mu_G$. Then, for all n^{th}-order polynomial utility function $u \in U_{np}^R$, the following statements are equivalent:*

1. $F \succeq_n^R G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $C_F^{(n)} \geq C_G^{(n)}$.

The proof of Theorem 2.16 is shown in appendix. Combining Equation (2.3.2) and Theorem 2.16, we can obtain the following corollaries. We first obtain the following corollary for any quadratic utility function:

Corollary 2.17. *Suppose $\mu_F = \mu_G$. For any quadratic utility function $u \in U_{2p}^R$, the following statements are equivalent*

1. $F \succeq_2^R G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $\sigma_F^2 \geq \sigma_G^2$.

We then obtain the following corollary for any cubic utility function:

Corollary 2.18. *Suppose $\mu_F = \mu_G$ and $\sigma_F^2 = \sigma_G^2$. For any cubic utility function $u \in U_{3p}^R$, the following statements are equivalent:*

1. $F \succeq_3^R G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $\gamma_F \geq \gamma_G$.

And obtain the following corollary for any quartic utility function:

Corollary 2.19. *Suppose $\mu_F = \mu_G$, $\sigma_F^2 = \sigma_G^2$, and $\gamma_G = \gamma_F$. For any quartic utility function $u \in U_{4p}^R$, the following statements are equivalent:*

1. $F \succeq_4^R G$,
2. $Eu(F) \geq Eu(G)$, *and*
3. $\kappa_F \geq \kappa_G$.

2.4 Moment rule

There are many applications of the theory developed in Section 2.3. In this section, we discuss the applications on the extension of the mean-variance (MV) rule and we call it the moment rule. We also develop some properties for the moment rule in this section.

We first modify the MV rule (for risk averters) as follows:

Definition 2.20. For any two prospects X and Y with means μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively, X is said to dominate Y by the MV rule for risk averters, denoted by $X\ MV_{RA}\ Y$, if $\mu_X \geq \mu_Y$ and $\sigma_X \leq \sigma_Y$, in which the inequality holds in at least one of the two.

It is well-known that the mean-variance rule cannot handle some situations. For example, the paradox as shown in Example in p66-67 of Copeland *et al.* (2005). There are several solutions to the paradox.[7] Among them, Wong (2007) introduces the MV rule for risk seekers to solve the paradox. We modify it as follows:

Definition 2.21. For any two prospects X and Y with means μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively, X is said to dominate Y by the MV rule for risk seekers, denoted by $X\ MV_{RS}\ Y$, if $\mu_X \geq \mu_Y$ and $\sigma_X \geq \sigma_Y$, in which the inequality holds in at least one of the two.

We note that Meyer, Li, and Rose (2005) use stochastic dominance to examine whether adding internationally based assets to a wholly domestic portfolio generates diversification benefits for an investor. They conclude that stochastic dominance is superior to MV rule. To circumvent the limitation of the MV rule, we first extend the MV rule for risk averters introduced by Markowitz (1952a) and the MV rule for risk seekers introduced by Wong (2007) and others to obtain the following mean-variance-skewness rule for both risk averters and risk seekers to check their preferences on assets based on the first three moments of the distributions:

Definition 2.22. For any two prospects X and Y with means μ_X and μ_Y, standard deviations σ_X and σ_Y, and skewnesses γ_X and γ_Y, respectively,

1. X is said to dominate Y by the mean-variance-skewness rule for risk averters, denoted by $X\ MVS_{RA}\ Y$, if $\mu_X \geq \mu_Y$, $\sigma_X \leq \sigma_Y$, and $\gamma_X \geq \gamma_Y$, and
2. X is said to dominate Y by the mean-variance-skewness rule for risk seekers, denoted by $X\ MVS_{RS}\ Y$, if $\mu_X \geq \mu_Y$, $\sigma_X \geq \sigma_Y$, and $\gamma_X \geq \gamma_Y$,

in which the inequality holds in at least one of the three.

[7] See Levy (2015) for more information.

Brockett and Garven (1998) comment that it is possible to have X and Y with positive and equal means, X having a larger variance and lower positive skewness than Y, and yet X has a larger expected utility than Y, implying skewness preference for risk averters. Definition 2.22 and the corresponding results in this section could be used to address some concerns for Brockett and Garven (1998). In this chapter, we further extend the rule to the following mean-variance-skewness-kurtosis rule for both risk averters and risk seekers:

Definition 2.23. For any two prospects X and Y with means μ_X and μ_Y, standard deviations σ_X and σ_Y, skewnesses γ_X and γ_Y, and kurtosises, κ_X and κ_Y, respectively,

1. X is said to dominate Y by the mean-variance-skewness-kurtosis rule for risk averters, denoted by $X \ MVSK_{RA} \ Y$, if $\mu_X \geq \mu_Y$, $\sigma_X \leq \sigma_Y$, $\gamma_X \geq \gamma_Y$, and $\kappa_X \leq \kappa_Y$,
2. X is said to dominate Y by the mean-variance-skewness-kurtosis rule for risk seekers, denoted by $X \ MVSK_{RS} \ Y$, if $\mu_X \geq \mu_Y$, $\sigma_X \geq \sigma_Y$, $\gamma_X \geq \gamma_Y$, and $\kappa_X \geq \kappa_Y$,

in which the inequality holds in at least one of the four.

We can also extend the rule further to the following mean-variance-skewness-kurtosis-$\cdots$-n^{th}-order central moment rule (we call it first n^{th}-order moments rule or, in short, n-moment rule, or just moment rule) for both risk averters and risk seekers for their preferences on assets based on the first n^{th}-order moments of the distributions:

Definition 2.24. For any two prospects X and Y with means μ_X and μ_Y, and the k^{th}-order central moments $C_X^{(k)}$ and $C_Y^{(k)}$, respectively, for any $2 \leq k \leq n$,

1. X is said to dominate Y by the n-moment rule for risk averters, denoted by $X \ M_{RA}^n \ Y$, if $\mu_X \geq \mu_Y$, and $(-1)^k C_X^{(k)} \leq (-1)^k C_Y^{(k)}$ for any $2 \leq k \leq n$, and
2. X is said to dominate Y by the n-moment rule for risk seekers, denoted by $X \ M_{RS}^n \ Y$, if $\mu_X \geq \mu_Y$, and $C_X^{(k)} \geq C_Y^{(k)}$ for any $2 \leq k \leq n$,

in which the inequality holds in at least one of the above.

Consider the following conjecture for the preferences of both risk averters and risk seekers for the n-moment rule:

Conjecture 1. For any two prospects X and Y with means μ_X and μ_Y and the k^{th}-order central moments $C_X^{(k)}$ and $C_Y^{(k)}$, respectively, for any $2 \leq k \leq n$, *under some conditions* we have

1. if $X \ M_{RA}^n \ Y$, then $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-averse investor with the utility function $u \in U_n$, and
2. if $X \ M_{RS}^n \ Y$, then $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-seeking investor with the utility function $u \in U_n^R$.

One could easily set an example that Conjecture 1 does not hold if we do not impose any condition. Readers may refer to Chapter 3.13 in Levy (2015) for such a counterexample. In fact, Levy (2015) has shown that $X \ M_{RA}^2 \ Y$ is neither sufficient nor necessary for the SSD relationship. Could Conjecture 1 hold true under some conditions? These rules are useful because Wong (2006) and Wong (2007) establish the results[8] to get the necessary conditions between stochastic dominance and the mean-variance rules for risk averters and risk seekers.

Theorem 5 in Wong (2007) tells us that if both X and Y belong to the same location-scale family or the same linear combination of location-scale families, then Conjecture 1 holds for risk averters ($n = 2$) and for risk seekers ($n = 2$). However, the condition that both X and Y belong to the same location-scale family or the same linear combination of location-scale families is very strong. Not many real-life data follow the same-location-scale-family condition. Could we relax in this condition?

Readers may believe that one could use the result from Theorems 2.4 and 2.8 to remove the same-location-scale-family condition. That is true in the sense that they can be used to obtain the necessary but not sufficient conditions for Conjecture 1 as stated in the following theorem:

Theorem 2.25. *For any two prospects X and Y with means μ_X and μ_Y and the k^{th}-order central moments $C_X^{(k)}$ and $C_Y^{(k)}$, respectively, for any $2 \leq k \leq n$, if $\mu_X = \mu_Y$ and if $C_X^{(k)} = C_Y^{(k)}$ for any $2 \leq k \leq n-1$, then we have*

1. *if $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-averse investor with the utility function $u \in U_n$, then $X \ M_{RA}^n \ Y$, and*

[8] Readers may refer to Theorem 5 in Wong (2007).

2. *if* $E\,[u(X)] \geq E\,[u(Y)]$ *for any risk-seeking investor with the utility function* $u \in U_n^R$*, then* $X\ M_{RS}^n\ Y$*.*

We note that applying the result from Theorems 2.4 and 2.8 could help us to obtain the necessary but not sufficient conditions for Conjecture 1. On the other hand, employing the result from Theorems 2.12 and 2.16 could get us both necessary and sufficient conditions for Conjecture 1 as the following:

Theorem 2.26. *For any two prospects* X *and* Y *with means* μ_X *and* μ_Y *and the* k^{th}*-order central moments* $C_X^{(k)}$ *and* $C_Y^{(k)}$*, respectively, for any* $2 \leq k \leq n$*, if* $\mu_X = \mu_Y$ *and if* $C_X^{(k)} = C_Y^{(k)}$ *for any* $2 \leq k \leq n-1$*, then we have*

1. $X\ M_{RA}^n\ Y$ *if and only if* $E\,[u(X)] \geq E\,[u(Y)]$ *for any risk-averse investor with the utility function* $u \in U_{np}$*, and*

2. $X\ M_{RS}^n\ Y$ *if and only if* $E\,[u(X)] \geq E\,[u(Y)]$ *for any risk-seeking investor with the utility function* $u \in U_{np}^R$*.*

Applying the results from Theorems 2.12 and 2.16 does getting us a very nice result that both necessary and sufficient conditions for Conjecture 1 are satisfied. Nonetheless, we need to impose another very strong (equal-$n-1$-moments) assumption of first $n-1$ moments being equal that most real-life data do not satisfied. In addition, it may be the case that practitioners are only interested in knowing the sufficient condition of Conjecture 1, but not the necessary condition of Conjecture 1. In view of this, we develop the following theorem to relax both the same-location-scale-family assumption and the equal-$n-1$-moments assumption:

Theorem 2.27. *For any two prospects* X *and* Y *with means* μ_X *and* μ_Y*, and the* k^{th}*-order central moments* $C_X^{(k)}$ *and* $C_Y^{(k)}$*, respectively, for any* $2 \leq k \leq n$*, if* $\mu_X = \mu_Y$*, then*

1. *if* $X\ M_{RA}^n\ Y$*, then* $E\,[u(X)] \geq E\,[u(Y)]$ *for any risk-averse investor with the utility function* $u \in U_{np}$*, and*

2. *if* $X\ M_{RS}^n\ Y$*, then* $E\,[u(X)] \geq E\,[u(Y)]$ *for any risk-seeking investor with the utility function* $u \in U_{np}^R$*.*

Theorem 2.27 does relax both the strong same-location-scale-family assumption and the equal-$n-1$-moments assumption that most real-life data do not satisfied. However, Theorem 2.27 requires another assumption that $u \in U_{np}$ or U_{np}^R. it does not require assumption on the distribution (except equal mean), but it is restricted to the types of investors that only belong to polynomial utility functions.

Academics and practitioners may ask: is it possible we do not impose the strong same-location-scale-family assumption, the equal-$n-1$-moments assumption, and we do not restrict to the types of investors that only belong to polynomial utility functions? We get the following theorem for this purpose:

Theorem 2.28. *For any two prospects X and Y with means μ_X and μ_Y, and the k^{th}-order central moments $C_X^{(k)}$ and $C_Y^{(k)}$, respectively, for any $2 \leq k \leq n$, if $\mu_X = \mu_Y$ and if the summation of all the terms in Taylor expansion of the utility after the n term is ignorable, then*

1. *if $X \ M_{RA}^n \ Y$, then $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-averse investor with the utility function $u \in U_n$, and*

2. *if $X \ M_{RS}^n \ Y$, then $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-seeking investor with the utility function $u \in U_n^R$.*

We note that Theorem 2.28 does relax both the strong same-location-scale-family assumption and the equal-$n-1$-moments assumption that most real-life data do not satisfied. In addition, it does not restrict $u \in U_{np}$ or U_{np}^R. But then, there is no free lunch. The price to pay is that we need to impose the assumption that *the summation of all the terms in Taylor expansion of the utility after the n term is ignorable* that makes Theorem 2.28 "ugly".

Could we relax the strong "equal-mean" condition? The answer is "YES" as we have got two as following:

Theorem 2.29. *For any two prospects X and Y with means μ_X and μ_Y ($\mu_X \neq \mu_Y$), and the k^{th}-order central moments $C_X^{(k)}$ and $C_Y^{(k)}$, respectively, for any $k = 2, 3$, for any $n = 2, 3$ we have*

1. *if $X \ M_{RA}^n \ Y$, then $E\left[u(X)\right] \geq E\left[u(Y)\right]$ for any risk-averse investor with the utility function $u \in U_{np}$, and*

2. *if* $X\ M^n_{RS}\ Y$*, then* $E\left[u(X)\right] \geq E\left[u(Y)\right]$ *for any risk-seeking investor with the utility function* $u \in U^R_{np}$.

Theorem 2.30. *For any two prospects* X *and* Y *with means* μ_X *and* μ_Y *(* $\mu_X \neq \mu_Y$*), and the* k^{th}*-order central moments* $C^{(k)}_X$ *and* $C^{(k)}_Y$*, respectively, for any* $2 \leq k \leq n$*,*

1. *if* $X\ M^n_{RA}\ Y$*, and* $C^{(k)}_Y \leq 0$ *for any odd integer* $k \geq 3$*, then* $E\left[u(X)\right] \geq E\left[u(Y)\right]$ *for any risk-averse investor with the utility function* $u \in U_{np}$ *and*

2. *if* $X\ M^n_{RS}\ Y$ *and* $C^{(k)}_X \geq 0$ *for any odd integer* $k \geq 3$*, then* $E\left[u(X)\right] \geq E\left[u(Y)\right]$ *for any risk-seeking investor with the utility function* $u \in U^R_{np}$.

We note that in Part 1 of Theorem 2.30, we only require $C^{(k)}_Y \leq 0$ for any odd integer $k \geq 3$, while the sign of $C^{(k)}_X$ is not required. It can be positive as long as $C^{(k)}_X \geq C^{(k)}_Y$ holds. Similarly, in Part 2 of Theorem 2.30, we only require $C^{(k)}_X \geq 0$ for any odd integer $k \geq 3$, the sign of $C^{(k)}_Y$ is not required. It can be negative as long as $C^{(k)}_X \geq C^{(k)}_Y$ holds.

2.5 Portfolio Diversification

We now extend the theory developed in Section 2.4 to develop some properties of portfolio diversification for the general utility functions and the polynomial utility functions of both risk averters and risk seekers. We will apply some results in Hadar and Russell (1971), Tesfatsion (1976), Li and Wong (1999), Wong (2007), Guo and Wong (2016), Chan, *et al.* (2020), and others. First, we apply Theorem 5 in Wong (2007) and Theorem 2.25 to develop the following properties of portfolio diversification to compare the preferences of two sets of assets for the general utility functions of both risk averters and risk seekers:

Theorem 2.31. *For any* $i = 1, \cdots, m$*, let* $\{X_i\}$ *and* $\{Y_i\}$ *be two sets of independent variables with means* μ_{X_i} *and* μ_{Y_i} *and the* k^{th}*-order central moments* $C^{(k)}_{X_i}$ *and* $C^{(k)}_{Y_i}$*, respectively, for any* $2 \leq k \leq n$*, if both* X_i *and* Y_i *belong to the same location-scale family or the same linear combination of location-scale families, then, for any* $2 \leq k \leq n$*, if* $\mu_{X_i} = \mu_{Y_i}$ *and if* $C^{(k)}_{X_i} = C^{(k)}_{Y_i}$ *for any* $2 \leq k \leq n-1$*, then we have*

1. *if* $E\left[u(X_i)\right] \geq E\left[u(Y_i)\right]$ *for any risk-averse investor with the utility function* $u \in U_n$*, then* $\sum_{i=1}^m \alpha_i X_i\ M^n_{RA}\ \sum_{i=1}^m \alpha_i Y_i$*, and*

2. *if* $E\ [u(X_i)] \geq E\ [u(Y_i)]$ *for any risk-seeking investor with the utility function* $u \in U_n^R$*, then* $\sum_{i=1}^m \alpha_i X_i\ M_{RS}^n\ \sum_{i=1}^m \alpha_i Y_i$*,*

for any $\alpha_i \geq 0, i = 1, \cdots, m$.

Next, we apply Theorem 5 in Wong (2007) and Theorem 2.26 to obtain the following theorem to compare the preferences two sets of assets for the polynomial utility functions of both risk averters and risk seekers:

Theorem 2.32. *For any* $i = 1, \cdots, m$*, let* $\{X_i\}$ *and* $\{Y_i\}$ *be two sets of independent variables with means* μ_{X_i} *and* μ_{Y_i} *and the* k^{th}*-order central moments* $C_{X_i}^{(k)}$ *and* $C_{Y_i}^{(k)}$*, respectively, for any* $2 \leq k \leq n$*, if* $\mu_{X_i} = \mu_{Y_i}$ *and if* $C_{X_i}^{(k)} = C_{Y_i}^{(k)}$ *for any* $2 \leq k \leq n-1$*, then we have*

1. $\sum_{i=1}^m \alpha_i X_i\ M_{RA}^n\ \sum_{i=1}^m \alpha_i Y_i$ *if and only if* $E\ [u(\sum_{i=1}^m \alpha_i X_i)] \geq E\ [u(\sum_{i=1}^m \alpha_i Y_i)]$ *for any risk-averse investor with the utility function* $u \in U_{np}$*, and*

2. $\sum_{i=1}^m \alpha_i X_i\ M_{RS}^n\ \sum_{i=1}^m \alpha_i Y_i$ *if and only if* $E\ [u(\sum_{i=1}^m \alpha_i X_i)] \geq E\ [u(\sum_{i=1}^m \alpha_i Y_i)]$ *for any risk-seeking investor with the utility function* $u \in U_{np}^R$*,*

for any $\alpha_i \geq 0, i = 1, \cdots, m$.

We then apply Theorem 5 in Wong (2007) and Theorem 2.27 to obtain the following theorem to compare the preferences two sets of assets for the polynomial utility functions of both risk averters and risk seekers:

Theorem 2.33. *For any* $i = 1, \cdots, m$*, let* $\{X_i\}$ *and* $\{Y_i\}$ *be two sets of independent variables with means* μ_{X_i} *and* μ_{Y_i} *and the* k^{th}*-order central moments* $C_{X_i}^{(k)}$ *and* $C_{Y_i}^{(k)}$*, respectively, for any* $2 \leq k \leq n$*, if* $\mu_{X_i} = \mu_{Y_i}$*, then*

1. *if* $\sum_{i=1}^m \alpha_i X_i\ M_{RA}^n\ \sum_{i=1}^m \alpha_i Y_i$*, then* $E\ [u(\sum_{i=1}^m \alpha_i X_i)] \geq E\ [u(\sum_{i=1}^m \alpha_i Y_i)]$ *for any risk-averse investor with the utility function* $u \in U_{np}$*, and*

2. *if* $\sum_{i=1}^m \alpha_i X_i\ M_{RS}^n\ \sum_{i=1}^m \alpha_i Y_i$*, then* $E\ [u(\sum_{i=1}^m \alpha_i X_i)] \geq E\ [u(\sum_{i=1}^m \alpha_i Y_i)]$ *for any risk-seeking investor with the utility function* $u \in U_{np}^R$*,*

for any $\alpha_i \geq 0, i = 1, \cdots, m$.

In addition, applying Theorem 5 in Wong (2007) and Theorem 2.28, we obtain the following theorem to compare the preferences two sets of assets for the general utility functions of both risk averters and risk seekers:

Theorem 2.34. *For any $i = 1, \cdots, m$, let $\{X_i\}$ and $\{Y_i\}$ be two sets of independent variables with means μ_{X_i} and μ_{Y_i} and the k^{th}-order central moments $C_{X_i}^{(k)}$ and $C_{Y_i}^{(k)}$, respectively, for any $2 \leq k \leq n$, if $\mu_{X_i} = \mu_{Y_i}$ and if the summation of all the terms in Taylor expansion of the utility after the n term is ignorable, then*

1. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RA}^n \ \sum_{i=1}^m \alpha_i Y_i$, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-averse investor with the utility function $u \in U_n$, and*
2. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RS}^n \ \sum_{i=1}^m \alpha_i Y_i$, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-seeking investor with the utility function $u \in U_n^R$,*

for any $\alpha_i \geq 0, i = 1, \cdots, m$.

Moreover, applying Theorem 5 in Wong (2007) and Theorem 2.29, we obtain the following theorem to compare the preferences of two sets of assets for the polynomial utility functions of both risk averters and risk seekers:

Theorem 2.35. *For any $i = 1, \cdots, m$, let $\{X_i\}$ and $\{Y_i\}$ be two sets of independent variables with means μ_{X_i} and μ_{Y_i} and the k^{th}-order central moments $C_{X_i}^{(k)}$ and $C_{Y_i}^{(k)}$, respectively, for any $k = 2, 3$, for any $n = 2, 3$ we have*

1. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RA}^n \ \sum_{i=1}^m \alpha_i Y_i$, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-averse investor with the utility function $u \in U_{np}$, and*
2. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RS}^n \ \sum_{i=1}^m \alpha_i Y_i$, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-seeking investor with the utility function $u \in U_{np}^R$,*

for any $\alpha_i \geq 0, i = 1, \cdots, m$.

Last, we apply Theorem 5 in Wong (2007) and Theorem 2.30 to obtain the following theorem to compare the preferences of two sets of assets for the polynomial utility functions of both risk averters and risk seekers:

Theorem 2.36. *For any $i = 1, \cdots, m$, let $\{X_i\}$ and $\{Y_i\}$ be two sets of independent variables with means μ_{X_i} and μ_{Y_i} and the k^{th}-order central moments $C_{X_i}^{(k)}$ and $C_{Y_i}^{(k)}$, respectively, for any $2 \leq k \leq n$,*

1. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RA}^n \ \sum_{i=1}^m \alpha_i Y_i$, and $C_{Y_i}^{(k)} \leq 0$ for any odd integer $k \geq 3$ for any i, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-averse investor with the utility function $u \in U_{np}$ and*

2. *if $\sum_{i=1}^m \alpha_i X_i \ M_{RS}^n \ \sum_{i=1}^m \alpha_i Y_i$ and $C_{X_i}^{(k)} \geq 0$ for any odd integer $k \geq 3$ for any i, then $E\left[u(\sum_{i=1}^m \alpha_i X_i)\right] \geq E\left[u(\sum_{i=1}^m \alpha_i Y_i)\right]$ for any risk-seeking investor with the utility function $u \in U_{np}^R$.*

Theorems 2.31 to 2.36 establishe some necessary and sufficient conditions between the preferences of portfolio diversification by using the moment rules and by using expected utility for both risk averters and risk seekers. We note that Theorems 2.31 to 2.36 compare two sets of assets with the same non-negative weight, α_i, in the corresponding i^{th} assets. We turn to establish some necessary and sufficient conditions between the preferences of portfolio diversification by using the moment rules to compare two sets of assets with different weights in the corresponding assets. To do so, we first define

$$\Lambda_n^0 = \left\{ (\lambda_1, \lambda_2, \cdots, \lambda_n)' \in \mathbb{R}^n : 0 \leq \lambda_i \leq 1 \text{ for any } i \ , \ \sum_{i=1}^n \lambda_i = 1 \right\} \tag{2.5.1}$$

and call X_i be an *an individual asset*, $\frac{1}{n}\sum_{i=1}^n X_i$ be the *completely diversified portfolio*, and $\sum_{i=1}^n \lambda_i X_i$ be a *partially diversified portfolio* if there exists i such that $0 < \lambda_i < 1$ and $\lambda_i \in \Lambda_n^0$. We then extend Theorem 12 in Li and Wong (1999) to obtain the following theorems to compare preference among an individual asset, a completely diversified portfolio, and a partially diversified portfolio for the general utility functions of both risk averters and risk seekers:

Theorem 2.37. *For any $X_1, \cdots, X_n$ with $n \geq 2$, if $X_1, \cdots, X_n$ are i.i.d., then we have*

1. *$\frac{1}{n}\sum_{i=1}^n X_i \ M_{RA}^2 \ \sum_{i=1}^n \lambda_i X_i \ M_{RA}^2 \ X_i$ and $E\left[u\left(\frac{1}{n}\sum_{i=1}^n X_i\right)\right] \geq E\left[u\left(\sum_{i=1}^n \lambda_i X_i\right)\right] \geq E\left[u\left(X_i\right)\right]$ for any risk-averse investor with the utility function $u \in U_2$, and*

2. *$X_i \ \ M_{RS}^2 \sum_{i=1}^n \lambda_i X_i M_{RS}^2 \frac{1}{n}\sum_{i=1}^n X_i$ and $E\left[u\left(X_i\right)\right] \geq E\left[u\left(\sum_{i=1}^n \lambda_i X_i\right)\right] \geq E\left[u\left(\frac{1}{n}\sum_{i=1}^n X_i\right)\right]$ for any risk-seeking investor with the utility function $u \in U_2^R$, and*

for any $(\lambda_1, \cdots, \lambda_n) \in \Lambda_n$ *defined in Equation (2.5.1).*

Theorem 2.37 establishes the preferences among individual asset, partially-diversified portfolio, and completely-diversified portfolio by using both moment rule and expected-utility rule for both risk averters and risk seekers.

2.6 Majorization

Using the results in Section 2.5, one can compare the preferences among individual asset, partially-diversified portfolio, and completely-diversified portfolio, but cannot compare the preferences between two partially-diversified portfolios. To circumvent the limitation, we apply the results in Egozcue and Wong (2010) to compare the preferences between some partially-diversified portfolios in this section. To do so, we first define

$$\Lambda_n = \left\{ (\lambda_1, \lambda_2, \cdots, \lambda_n)' \in \mathbb{R}^n : 1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \right\} . \tag{2.6.1}$$

We then follow Hardy, *et al.* (1934) and others to make the following definition:

Definition 2.38. Let $\vec{a}_n, \vec{b}_n \in \Lambda_n$ defined in (2.6.1). b_n is defined to majorize $\vec{a}_n$, represented by $\vec{b}_n \succeq_M \vec{a}_n$, if $\sum_{i=1}^{k} b_i \geq \sum_{i=1}^{k} a_i$, for any $k = 1, 2, \cdots, n$.

Applying Theorem 2.26, Theorem 7 in Egozcue and Wong (2010), and Theorem 3.8 in Guo and Wong (2016), we obtain the following theorem to compare preference of any two partially diversified portfolios for any set of independent assets and for the general utility functions of both risk averters and risk seekers:

Theorem 2.39. *For* $n > 1$*, let* $\vec{a}_n, \vec{b}_n \in \Lambda_n$ *and* $\vec{X}_n = (X_1, \cdots, X_n)'$ *in which* $X_1, \cdots, X_n$ *are i.i.d, if* $\vec{b}_n \succeq_M \vec{a}_n$*, then*

1. $\vec{a}'_n \vec{X}_n \; M^2_{RA} \; \vec{b}'_n \vec{X}_n$ *and* $E\left[u\left(\vec{a}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{b}'_n \vec{X}_n\right)\right]$ *for any risk-averse investor with the utility function* $u \in U_2$*, and*
2. $\vec{b}'_n \vec{X}_n \; M^2_{RS} \; \vec{a}'_n \vec{X}_n$ *and* $E\left[u\left(\vec{b}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{a}'_n \vec{X}_n\right)\right]$ *for any risk-seeking investor with the utility function* $u \in U_2^R$.

Can we drop the i.i.d. assumption to compare non-i.i.d. portfolio? Samuelson (1967) argues that, in general, we can't, but he does obtain some results to drop the i.i.d. assumption. To follow Samuelson's idea to drop the i.i.d. assumption, we apply Theorem 2.39, Corollary 9 in Egozcue and Wong (2010), and Corollary 3.2 in Guo and Wong (2016) to obtain the following corollary to compare preferences of some pairs of two partially diversified portfolios for some sets of dependent assets and for the general utility functions of both risk averters and risk seekers:

Corollary 2.40. *For $n > 1$, $\vec{X}_n = (X_1, \cdots, X_n)'$ is a series of dependent or independent random variables, For any $\vec{a}_n, \vec{b}_n \in \Lambda_n$, if there exist i.i.d. $\vec{Y}_n = (Y_1, \cdots, Y_n)'$ and P_{nn} in which $\vec{X}_n = P_{nn}\vec{Y}_n$ such that $\vec{b}'_n P_{nn} \succeq_M \vec{a}'_n P_{nn}$ with $\vec{a}'_n P_{nn}$, then*

1. *$\vec{a}'_n \vec{X}_n \ M^2_{RA} \ \vec{b}'_n \vec{X}_n$ and $E\left[u\left(\vec{a}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{b}'_n \vec{X}_n\right)\right]$ for any risk-averse investor with the utility function $u \in U_2$, and*
2. *$\vec{b}'_n \vec{X}_n \ M^2_{RS} \ \vec{a}'_n \vec{X}_n$ and $E\left[u\left(\vec{b}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{a}'_n \vec{X}_n\right)\right]$ for any risk-seeking investor with the utility function $u \in U^R_2$.*

Last, we turn to apply Theorem 2.39, Corollary 12 in Egozcue and Wong (2010), and Corollary 3.3 in Guo and Wong (2016) to obtain the following corollary to compare preferences of some pairs of two partially diversified portfolios for some sets of dependent assets and for the general utility functions of both risk averters and risk seekers:

Corollary 2.41. *For $n > 1$, $\vec{X}_n = (X_1, \cdots, X_n)'$ and $\vec{Y}_n = (Y_1, \cdots, Y_n)'$ are two series of dependent or independent random variables and $\vec{V}_n = (V_1, \cdots, V_n)'$ and $\vec{W}_n = (W_1, \cdots, W_n)'$ are two series of i.i.d. random variables. For any $\vec{a}_n, \vec{b}_n \in \Lambda_n$, if there exist P_{nn} and Q_{nn} such that $\vec{a}'_n P_{nn}$, $\vec{b}'_n Q_{nn} \in \Lambda_n$ and $\vec{b}'_n Q_{nn} \succeq_M \vec{a}'_n P_{nn}$, $\vec{X}_n = P_{nn}\vec{V}_n$, $\vec{Y}_n = Q_{nn}\vec{W}_n$, $V_i \succeq_2 W_i$ for all $i = 1, 2, ..., n$; then*

1. *$\vec{a}'_n \vec{X}_n \ M^2_{RA} \ \vec{b}'_n \vec{Y}_n$ and $E\left[u\left(\vec{a}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{b}'_n \vec{Y}_n\right)\right]$ for any risk-averse investor with the utility function $u \in U_2$, and*
2. *$\vec{b}'_n \vec{X}_n \ M^2_{RS} \ \vec{a}'_n \vec{Y}_n$ and $E\left[u\left(\vec{b}'_n \vec{X}_n\right)\right] \geq E\left[u\left(\vec{a}'_n \vec{Y}_n\right)\right]$ for any risk-seeking investor with the utility function $u \in U^R_2$.*

2.7 Moment rule tests

In this section, we describe the testing procedure to test the performance of assets by using the moment rule up to the fourth order. To do so, we let null hypothesis H_0^1 be $H_0^1 : \mu_X < \mu_Y$, H_0^2 be $H_0^2 : \sigma_X^2 > \sigma_Y^2$, H_{0S}^2 be $H_{0S}^2 : \sigma_X^2 < \sigma_Y^2$, H_0^3 be $H_0^3 : \gamma_X < \gamma_Y$, H_0^4 be $H_0^4 : \kappa_X > \kappa_Y$, H_{0S}^4 be $H_{0S}^4 : \kappa_X < \kappa_Y$. Further, we let H_A^i be alternative hypothesis of H_0^i for $i = 1, 2, 3, 4$ and H_{AS}^i be alternative hypothesis of H_{0S}^i for $i = 2, 4$.

To test whether $X \ MV_{RA} \ Y(X \ MV_{RS} \ Y)$, we simply test the joint null hypotheses $H_0^{MV} : H_0^1 \cup H_0^2 (H_{0S}^{MV} : H_0^1 \cup H_{0S}^2)$, if we reject the joint null hypotheses, we will conclude that $X \ MV_{RA} \ Y(X \ MV_{RS} \ Y)$ because the joint alternative hypotheses are $H_A^{MV} : H_A^1 \cap H_A^2 (H_{AS}^{MV} : H_A^1 \cap H_{AS}^2)$. On the other hand, to test whether $X \ MVS_{RA} \ Y(X \ MVS_{RS} \ Y)$, we simply test the joint null hypotheses $H_0^{MVS} : H_0^1 \cup H_0^2 \cup H_0^3 (H_{0S}^{MV} : H_0^1 \cup H_{0S}^2 \cup H_0^3)$, if we reject the joint null hypotheses, we will conclude that $X \ MVS_{RA} \ Y(X \ MVS_{RS} \ Y)$ because the joint alternative hypotheses are $H_A^{MVS} : H_A^1 \cap H_A^2 \cap H_A^3 (H_{AS}^{MVS} : H_A^1 \cap H_{AS}^2 \cap H_A^3)$. Last, but not the least, to test whether $X \ MVSK_{RA} \ Y(X \ MVSK_{RS} \ Y)$, we simply test the joint null hypotheses $H_0^{MVSK} : H_0^1 \cup H_0^2 \cup H_0^3 \cup H_0^4 (H_{0S}^{MVSK} : H_0^1 \cup H_{0S}^2 \cup H_0^3 \cup H_{0S}^4)$. If we reject the joint null hypotheses, we will conclude that $X \ MVSK_{RA} \ Y(X \ MVSK_{RS} \ Y)$ because the joint alternative hypotheses are $H_A^{MVSK} : H_A^1 \cap H_A^2 \cap H_A^3 \cap H_A^4 (H_{AS}^{MVSK} : H_A^1 \cap H_{AS}^2 \cap H_A^3 \cap H_{AS}^4)$.

Because H_0^{MV}, H_{0S}^{MV}, H_0^{MVS}, H_{0S}^{MVS}, H_0^{MVSK}, or H_{0S}^{MVSK}, can be viewed as a composite of multiple hypotheses, we can test those individual hypotheses separately and see whether all of them are rejected in order to reject H_0^{MV}, H_{0S}^{MV}, H_0^{MVS}, H_{0S}^{MVS}, H_0^{MVSK}, or H_{0S}^{MVSK}. For convenient, we let the p-values of H_0^{MV}, H_{0S}^{MV}, H_0^{MVS}, H_{0S}^{MVS}, H_0^{MVSK}, and H_{0S}^{MVSK} as the largest p-values for their corresponding multiple hypotheses. Take H_0^{MVS} as an example, we can find out the p-values for testing H_0^1, H_0^2, and H_0^3, say p_1, p_2 and p_3, then we reject H_0^{MVS} at α significant level if $p^{MVS} = max\{p_i\} \leq \alpha$ for $i = 1, 2, 3$. In order to test H_0^1, H_0^2, and H_{0S}^2 we use the commonly-used T- and F tests. To test H_0^3 b, H_0^4 be and H_{0S}^4, we use the bootstrap method (Efron and LePage, 1992; Shao and Tu, 2012).

Since the existing statistical tests for SD, see, for example, Davidson and Duclos (2000), Barrett and Donald (2003), Post and Versijp (2007), Bai, *et al.* (2011, 2015), and Ng, *et al.* (2017), are generally computationally intensive due to the need to use resampling methods, researchers may not want to use the SD tests. In this situation, they could just apply the moment rule tests. Using the moment tests, readers should be able to draw similar conclusions in Vinod (2004), Tsang, *et al.* (2016). Chan, *et al.* (2020), Lv, *et al.* (2021), and many others when one uses their data.

2.8 Applications

In this section, we illustrate the applicability of the theory developed in our paper by using real-life data.[9] We use the excess return of 49 industry average value-weighted portfolios from Kenneth French's online data library to illustrate the statistical test on H_0^{MVS}, H_{0S}^{MVS}, H_0^{MV}, and H_{0S}^{MV}. The portfolios we used in our analysis contain almost all common stocks listed on NYSE, AMEX, and NASDAQ, which are formed by using the four-digit standard industrial classification code.[10]

We denote the portfolios by their names in Kenneth French's online data library. In addition, we use the one-month US Treasury bill as a risk-free rate to help us calculate the excess returns. In this chapter, we mainly focus on the monthly excess return from Jan 1992 to December 2021. All data is pulled directly from Kenneth French's online data library.

We first divide our data into three sub-periods, including 1) Jan 1992 to Dec 2001, 2) Jan 2002 to Dec 2011, and 3) Jan 2012 to Dec 2021. We then compare the portfolios by using the relevant test on H_0^{MVS}, H_{0S}^{MVS}, H_0^{MV}, and H_{0S}^{MV}. For any portfolio that is dominated by any other portfolio in MV_{RA}, MV_{RS}, MVS_{RA}, or MVS_{RS} sense for at least two periods, we will show them in Table 4.1. In Table 4.1, the second column represents the minimum p-value among all the 48 p-values for testing hypothesis H_0^{MVSK} in which Y is the excess return of the industry portfolio denoted in the first column, and X is the excess return of another industry portfolio. To be more specific, under MVS_{RA}, we will perform 48 times of the test for each Y

[9]There are various studies that try to apply the SD approach in the real data. For example, Qiao, *et al.* (2014) and Clark, *et al.* (2016) find that risk averters prefer investing spot to futures while risk seekers prefer investing futures to spot, Wong, *et al.* (2008) conclude that third-order risk averters prefer investing in some Asian hedge funds to other Asian hedge funds, Chan, *et al.* (2020) find that the third-order risk averters prefer investing in the S&P 500 index to the Nasdaq 100 index and the third-order risk seekers prefer investing in the Nasdaq 100 index to the S&P 500 index. Vinod (2004) examines mutual funds and finds fourth-order SD dominance among the funds. Kallio and Hardoroudi (2019) obtain new results for fourth and fifth-order stochastic dominance. On the other hand, Hoang, *et al.* (2015) find that risk-averse investors prefer not to include gold while risk-seeking investors prefer to include it in their portfolios. Chui, *et al.* (2020) use SD test to check whether the market is efficient.

[10]See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html for more details about the portfolios.

(in order to compare Y with all the other 48 portfolios) and display the smallest p-value among the tests in the second column when the p-value is smaller than 10%. The same logic is applied to the third, fourth, and fifth columns, which represent the minimum p-values among all the 48 p-values for testing hypothesis H_{0S}^{MVS}, H_0^{MV}, and H_{0S}^{MV}, respectively, in which Y is the excess return of the industry portfolio denoted in the first column and X is the excess return of another industry portfolio when the minimum p-value is smaller than 10%. We also compares all the relevant moments for MV_{RA}, MV_{RS}, MVS_{RA}, or MVS_{RS} by using the corresponding 90% bootstrap confidence interval (CI). We will denote "All" if the confidence interval results are consistent with the test results in all of our hypotheses, "Partial" if the confidence interval results are consistent with one or some but not all the test results in our hypotheses, and "None" if the confidence interval results are not consistent with any of the test results in all of our hypotheses.

From Table 4.1, we first discuss the results for the period from Jan 1992 to Dec 2001. The results show that "Other" is dominated by at least one industry portfolio in the sense of MVS_{RS}; that is, there exists an industry portfolio that has a significantly higher return, higher variance, and higher skewness compared to "Other". For the MV_{RA} relationship, we find that "Toys", "Gold", "Boxes", and "Other" are dominated by at least one industry portfolio; that is, there exists an industry portfolio that has a significantly higher return and lower variance compared to "Toys", "Gold", "Boxes", or "Other". For the MV_{RS} relationship, we find that "Food", "Toys", "Util", "PerSv", "Boxes", "Whlsl", and "Other" are dominated by at least one industry portfolio; that is, there exists an industry portfolio has significantly higher return and higher variance compared to "Food", "Toys", "Util", "PerSv", "Boxes", "Whlsl", and "Other". The results from the confidence intervals are all consistent with the results from the hypothesis test.

We then discuss the results for the period from Jan 2002 to Dec 2011. From the table, "Books" and "Other" are found to be dominated by at least one industry portfolio in the sense of MV_{RA}. For the MV_{RS} relationship, all portfolios displayed in Table 4.1 in the period are dominated by at least one industry portfolio. On the other hand, in the period from Jan 2012 to Dec 2021, "Gold" is dominated by at least one industry portfolio in the sense of MVS_{RA}; that is, there exists an industry portfolio that has a significantly higher return, lower variance, and higher skewness compared to "Gold"; for the MVS_{RS} relationship, "Books", "Util", "Telem", "PerSv", and "Paper" are dominated by at least one industry portfolio; for

TABLE 2.1: Moment rule tests

Industry	MVS_{RA}	MVS_{RS}	MV_{RA}	MV_{RS}	CI
		Jan 1992 to Dec 2001			
Food	n.a.	n.a.	n.a.	0.0778*	All
Toys	n.a.	n.a.	0.0825*	0.0407**	All
Gold	n.a.	n.a.	0.0912*	n.a.	All
Util	n.a.	n.a.	n.a.	0.0811*	All
PerSv	n.a.	n.a.	n.a.	0.0903*	All
Boxes	n.a.	n.a.	0.0546*	0.0573*	All
Whlsl	n.a.	n.a.	n.a.	0.0625*	All
Other	n.a.	0.0834*	0.0339**	0.0395**	All
		Jan 2002 to Dec 2011			
Food	n.a.	n.a.	n.a.	0.0907*	All
Beer	n.a.	n.a.	n.a.	0.0964*	All
Toys	n.a.	n.a.	n.a.	0.0968*	All
Books	n.a.	n.a.	0.0610	0.0394**	All
Hshld	n.a.	n.a.	n.a.	0.0908*	All
Drugs	n.a.	n.a.	n.a.	0.0395**	All
Telcm	n.a.	n.a.	n.a.	0.0582*	All
PerSv	n.a.	n.a.	n.a.	0.0971*	All
Paper	n.a.	n.a.	n.a.	0.0960*	All
Other	n.a.	n.a.	0.0811*	0.0480**	All
		Jan 2012 to Dec 2021			
Food	n.a.	n.a.	n.a.	0.0194**	All
Beer	n.a.	n.a.	n.a.	0.0568*	All
Books	n.a.	0.0898*	n.a.	0.0898*	Partial
Hshld	n.a.	n.a.	n.a.	0.03808**	All
Drugs	n.a.	n.a.	n.a.	0.0850*	All
Gold	0.0782*	n.a.	0.0689*	n.a.	Partial
Util	n.a.	0.0793*	n.a.	0.0167**	All
Telcm	n.a.	0.0943*	n.a.	0.0339**	All
PerSv	n.a.	0.0567*	0.0650*	0.0567*	Partial
Paper	n.a.	0.0833*	n.a.	0.0297**	All
Boxes	n.a.	n.a.	n.a.	0.0737*	All
Whlsl	n.a.	n.a.	n.a.	0.0740*	None
Other	n.a.	n.a.	n.a.	0.0385**	All

The *, and ** denote the significance at 10%, and 5%, respectively.

the MV_{RA} relationship, both “Gold” and “PerSv” are dominated by at least one industry portfolio; and for the MV_{RS} relationship, except “Gold”, all portfolios shown in Table 4.1 in the period are dominated by at least one industry portfolio. The results from the confidence interval are all consistent with those from the hypothesis test.

In summary from all the results, we notice that around 30% of the portfolios are dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS}, MV_{RA}, or MV_{RS} sense in the period from Jan 1992 to Dec 2001 and are dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS}, MV_{RA}, or MV_{RS} sense during the period from Jan 2002 to Dec 2011. Second, around 50% of the portfolios are dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS},

MV_{RA}, or MV_{RS} sense in the period from Jan 2002 to Dec 2011 and are also dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS}, MV_{RA}, or MV_{RS} sense in the period from Jan 2012 to Dec 2021. Third, around 50% of the portfolios are dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS}, MV_{RA}, or MV_{RS} sense in the period from Jan 1992 to Dec 2001 and are dominated by at least one industry portfolio in MVS_{RA}, MVS_{RS}, MV_{RA}, or MV_{RS} sense in the period from Jan 2012 to Dec 2021. Most of the results from the confidence interval are consistent with those from the hypothesis test, except they only partially support the results for "Books", "Gold", and "PerSv". The result of the confidence interval does not support that of the hypothesis test for "Whlsl".

Kindly noted that the empirical application is not conclusive. The set of industries that is classified as significantly dominated seems not robust over time in this study. One possible interpretation is that the location, dispersion and/or shape of the probability distributions of the industries could change over time. Future research could focus on the evaluation of the statistical properties of the proposed tests using Monte Carlo simulations, and the development of universally valid statistical inference methods based on statistical subsampling and moment estimation methods.

2.9 Concluding remarks

In this chapter, we first extend the work of Markowitz (1952a), Tobin (1958), Chan, *et al.* (2020), and others by developing some theorems to state the relationships among central moments, stochastic dominance (SD), risk-seeking stochastic dominance (RSD), and integrals and establishing the relationship between the n^{th}-order (central) moments and the n^{th}-order [reversed] integrals for both n^{th}- and $(n+1)^{th}$-order [R]SD for general risk-averse [risk-seeking] utility functions and the polynomial utility functions of both risk averters and risk seekers for any order n, including $n = 2, 3$, and 4 as the special cases. We then apply the relationships to extend the mean-variance (MV) rule established by Markowitz (1952a), Wong (2007), and others by introducing the moment rule. As far as we know, our paper is the first introduces it in the literature, including the mean-variance-skewness rule, the mean-variance-skewness-kurtosis rule, for both risk averters and risk seekers.

Wong (2006, 2007) establish the necessary conditions between stochastic dominance and the mean-variance rules for both risk averters and risk seekers when the assets belong to the same location-scale family or the same linear combination of location-scale families. In this chapter, we extend the theory further by removing the same-location-scale-family condition to establish some necessary conditions between SD and the moment rule for both risk averters and risk seekers under some conditions. Thereafter, we apply the moment rules to develop some properties of portfolio diversification for the general utility functions and the polynomial utility functions of both risk averters and risk seekers. Last, we incorporate the idea of majorization with the moment rules to develop some properties of portfolio diversification for the general utility functions to compare the preferences between two partially diversified portfolios.

The findings in our paper enable academics and practitioners to draw preferences of both risk averters and risk seekers on their choices of portfolios or assets by using different moments. We illustrate this point by using the moment rule tests to compare the excess return of 49 industry portfolios from Kenneth French's online data library. We find that the results are reasonably stable from Jan 1992 to Dec 2021. First, around 30% of the portfolios that are dominated by an industry portfolio under a moment rule are also dominated by an industry portfolio under a moment rule from Jan 2002 to Dec 2011. Second, around 50% of the portfolios that are dominated by an industry portfolio under a moment rule from Jan 2002 to Dec 2011 are also dominated by an industry portfolio under a moment rule in the period from Jan 2012 to Dec 2021. Third, around 50% of the portfolios that are dominated by an industry portfolio under a moment rule from Jan 1992 to Dec 2001 are also dominated by an industry portfolio under a moment rule from Jan 2012 to Dec 2021. The set of industries that are classified as significantly dominated seems not robust over time in this study. One possible interpretation is that the location, dispersion, and/or shape of the probability distributions of the industries have changed over time. Future research could focus on the evaluation of the statistical properties of the proposed tests using Monte Carlo simulations, and the development of universally valid statistical inference methods based on statistical subsampling and moment estimation methods.

Higher orders stochastic dominance and risk measure are useful in many empirical studies, see, for example, Tehranian (1980) and Ogryczak and Ruszczyński (1999). Meyer, Li, and Rose (2005) use stochastic dominance to examine whether adding internationally-based assets to a wholly domestic portfolio generates diversification benefits for an investor. They conclude that stochastic dominance is superior to the mean-variance rule. Further extensions could examine whether stochastic dominance is superior to the moment rule.

Extensions of our paper could also include developing properties between central moments with prospect and Markowitz stochastic dominance for investors with S-shaped or reverse S-shaped utility function, developing the moment rule, establishing the necessary and/or sufficient conditions between prospect and Markowitz stochastic dominance, and developing some properties of portfolio diversification for investors with S-shaped or reverse S-shaped utility function.[11] Readers may read Levy and Wiener (1998), Levy and Levy (2002, 2004), Post and Levy (2005), and Wong and Chan (2008) for more information on prospect and Markowitz stochastic dominance, read Egozcue, *et al.* (2011) and Ortobelli Lozza, *et al.* (2018) for the diversification properties for other types of investors, including investors with S-shaped or reverse S-shaped utility function. In addition, our paper develops some properties by using the information of higher-order moments but has not used any information on the joint effects from moments. Thus, extensions of our paper could study the joint effects of moments, see, for example, Martellini and Ziemann (2009), Kostakis, Muhammad, and Siganos (2012), Lambert and Hubner (2013), Vo and Tran (2020), and many others.

[11]S-shaped utility function is proposed by Kahneman and Tversky (1979) and reverse S-shaped utility function is proposed by Markowitz (1952b)

2.10 Appendix

Proof of Theorem 2.4.

We prove Part 1 of the theorem by induction. Under the assumption that $\mu_F = \mu_G$, Chan, *et al.* (2020) have already proved formula (2.3.4) for $n = 2$ (see (2.3.3)). Now we prove that if the theorem holds for all $2 \le k \le n-1$, then it also holds for n. By the induction hypothesis, we have

$$G_{k+1}(b) - F_{k+1}(b) = \frac{(-1)^k}{k!}\big(C_G^{(k)} - C_F^{(k)}\big), \quad \forall 2 \le k \le n-1.$$

The assumption of the theorem for n implies that $C_G^{(k)} = C_F^{(k)}$ for all $2 \le k < n$. Hence we have $G_k(b) = F_k(b)$ for all $2 \le k \le n$. We also have $G_1(b) = 1 = F_1(b)$.

To get (2.3.4) for n, we note that for H where H can be F or G:

$$\begin{aligned}
C_H^{(n)} &= \int_a^b \big(x-\mu_H\big)^n dH(x) \\
&= \big(x-\mu_H\big)^n H(x)\Big|_a^b - n\int_a^b \big(x-\mu_H\big)^{n-1} H(x)dx \\
&= \big(b-\mu_H\big)^n - n\big(x-\mu_H\big)^{n-1}H_2(x)\Big|_a^b + \frac{n!}{(n-2)!}\int_a^b \big(x-\mu_H\big)^{n-2}H_2(x)dx \\
&= \sum_{k=1}^{n} \frac{(-1)^{k+1}n!}{(n+1-k)!}\big(b-\mu_H\big)^{n+1-k}H_k(b) + \frac{n!}{(-1)^n}H_{n+1}(b).
\end{aligned}$$

Hence

$$\begin{aligned}
& C_G^{(n)} - C_F^{(n)} \\
= \; & \sum_{k=2}^{n} \frac{(-1)^{k+1}n!}{(n+1-k)!}\big(b-\mu\big)^{n+1-k}[G_k(b) - F_k(b)] + \frac{n!}{(-1)^n}[G_{n+1}(b) - F_{n+1}(b)] \\
= \; & \frac{n!}{(-1)^n}[G_{n+1}(b) - F_{n+1}(b)]
\end{aligned}$$

which is precisely (2.3.4) for n. Thus, the assertion of Part 1 of Theorem 2.4 holds. Part 2 and part 4 of Theorem 2.4 can be obtained from using the result from Part 1 of Theorem 2.4 holds. We turn to prove Part 3 of Theorem 2.4.

To prove Part 3, we will try to prove $F \succeq_n G$ implies $G_{n+1}(b) > F_{n+1}(b)$ under the assumption of Theorem 2.4. Since $G_n(x) \ge F_n(x)$ for each x and $G_n(x) > F_n(x)$ for at least one x_0 in $[a, b]$

implies $\int_a^b G_n(x)dx > \int_a^b F_n(x)dx$, that is $G_{n+1}(b) > F_{n+1}(b)$. Thus, the Part 3 of Theorem 2.4 holds. □

Proof of Theorem 2.8. We prove Part 1 of the theorem by induction. Under the assumption that $\mu_F = \mu_G$, Chan, *et al.* (2020) have already proved formula (2.3.5) for $n = 2$ (see (2.3.3)). Now we prove that if the theorem holds for all $2 \le k \le n-1$, then it also holds for n. By the induction hypothesis, we have

$$F^R_{k+1}(a) - G^R_{k+1}(a) = \frac{1}{k!}\big(C^{(k)}_F - C^{(k)}_G\big), \quad \forall 2 \le k \le n-1.$$

To get (2.3.5) for n, we note that

$$\begin{aligned}
& C^{(n)}_F - C^{(n)}_G \\
= \; & -n\int_a^b (x-\mu)^{n-1}[F(x) - G(x)]dx \\
= \; & n\int_a^b (x-\mu)^{n-1}[F^R_1(x) - G^R_1(x)]dx \\
= \; & n(x-\mu)^{n-1}[G^R_2(x) - F^R_2(x)]\Big|_a^b + \frac{n!}{(n-2)!}\int_a^b (x-\mu)^{n-2}[F^R_2(x) - G^R_2(x)]dx \\
= \; & \sum_{k=2}^{n} \frac{n!}{(n+1-k)!}\big(a-\mu\big)^{n+1-k}[F^R_k(a) - G^R_k(a)] + n![F^R_{n+1}(a) - G^R_{n+1}(a)]
\end{aligned}$$

Hence

$$\begin{aligned}
C^{(n)}_F - C^{(n)}_G &= n![F^R_{n+1}(a) - G^R_{n+1}(a)] \\
F_{n+1}(a) - G_{n+1}(a) &= \frac{1}{n!}(C^{(n)}_F - C^{(n)}_G)
\end{aligned}$$

which is precisely (2.3.5) for n. Thus, the assertion of Part 1 of Theorem 2.8 holds. Part 2 and part 3 of Theorem 2.8 can be obtained from using the result from Part 1 of Theorem 2.8 holds. We turn to prove Part 4 of Theorem 2.8.

To prove Part 3, we will try to prove $F \succeq^R_n G$ implies $F^R_{n+1}(a) > G^R_{n+1}(a)$ under the assumption of Theorem 2.8. Since $F^R_n(x) \ge G^R_n(x)$ for each x and $F^R_n(x) > G^R_n(x)$ for at least one x_0

in $[a,b]$ implies $\int_a^b F_n^R(x)dx > \int_a^b G_n^R(x)dx$, that is $F_{n+1}^R(a) > G_{n+1}^R(a)$. Thus, the Part 3 of Theorem 2.8 holds. □

Proof of Theorem 2.12.

By theorem 2.4, $G_{n+1}(b) \geq F_{n+1}(b)$ is equivalent to $(-1)^n C_G^{(n)} \geq (-1)^n C_F^{(n)}$ under the assumption of this theorem. Thus, we just need to prove that $G_{n+1}(b) \geq F_{n+1}(b)$ is equivalent to $Eu(F) \geq Eu(G)$ under the assumption of this theorem.

$$\begin{aligned}
\Delta Eu &\equiv Eu(F) - Eu(G) \equiv \int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) \\
&= [G_2(b) - F_2(b)]u^{(1)}(b) - \int_a^b [G_2(x) - F_2(x)]u^{(2)}(x)dx \\
&= [G_2(b) - F_2(b)]u^{(1)}(b) - [G_3(b) - F_3(b)]u^{(2)}(b) + \int_a^b [G_3(x) - F_3(x)]u^{(3)}(x)dx \\
&= \sum_{k=2}^{n} (-1)^i [G_i(b) - F_i(b)]u^{(i-1)}(b) + (-1)^{n+1} \int_a^b [G_n(x) - F_n(x)]u^{(n)}(x)dx.
\end{aligned}$$

By our assumptions , Theorem 2.4 and Equation 2.3.2, we get $G_k(b) = F_k(b)$ for all $2 \leq k \leq n$. We have

$$\Delta Eu = (-1)^{n+1} \int_a^b [G_n(x) - F_n(x)]u^{(n)}(x)dx.$$

By the assumption, we have nonzero constant $u^{(n)}(x) = u^{(n)}$, thus

$$\Delta Eu = (-1)^{n+1} u^{(n)} [G_{n+1}(b) - F_{n+1}(b)].$$

Since $(-1)^{n+1}u^{(n)}$ is always positive, $[G_{n+1}(b) - F_{n+1}(b)] \geq 0$ implies $\Delta Eu \geq 0$ and $\Delta Eu \geq 0$ implies $[G_{n+1}(b) - F_{n+1}(b)] \geq 0$. Thus, the Theorem 2.12 holds. □

Proof of Theorem 2.16.

By theorem 2.12, $F_{n+1}^R(a) \geq G_{n+1}^R(a)$ is equivalent to $C_F^{(n)} \geq C_G^{(n)}$ under the assumption of this theorem. Thus, we just need to prove that $F_{n+1}^R(a) \geq G_{n+1}^R(a)$ is equivalent to $Eu(F) \geq Eu(G)$

under the assumption of this theorem.

$$\begin{aligned}
\Delta Eu &\equiv Eu(F) - Eu(G) \equiv \int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) \\
&= \int_a^b [F_1^R(x) - G_1^R(x)]u^{(1)}(x)dx \\
&= [F_2^R(a) - G_2^R(a)]u^{(1)}(a) + \int_a^b [F_2^R(x) - G_2^R(x)]u^{(2)}(x)dx \\
&= [F_2^R(a) - G_2^R(a)]u^{(1)}(a) + [F_3^R(a) - G_3^R(a)]u^{(2)}(a) + \int_a^b [F_3^R(x) - G_3^R(x)]u^{(3)}(x)dx \\
&= \sum_{k=2}^{n} [F_i^R(a) - G_i^R(a)]u^{(i-1)}(a) + \int_a^b [F_n^R(x) - G_n^R(x)]u^{(n)}(x)dx.
\end{aligned}$$

By our assumptions, Theorem 2.12 and Equation 2.3.2, we get $F_k^R(a) = G_k^R(a)$ for all $2 \leq k \leq n$. We have

$$\Delta Eu = \int_a^b [F_n^R(x) - G_n^R(x)]u^{(n)}(x)dx.$$

By the assumption, we have nonzero constant $u^{(n)}(x) = u^{(n)}$, thus

$$\Delta Eu = u^{(n)}[F_{n+1}^R(a) - G_{n+1}^R(a)].$$

Since $u^{(n)}$ is always positive, $[F_{n+1}^R(a) - G_{n+1}^R(a)] \geq 0$ implies $\Delta Eu \geq 0$ and $\Delta Eu \geq 0$ implies $[F_{n+1}^R(a) - G_{n+1}^R(a)] \geq 0$. Thus, the Theorem 2.16 holds. □

Proof of Theorems 2.27 and 2.28:

Using the Taylor formula, we can express the $u(F)$ as:

$$u(F) = u(\mu) + (x-\mu)u^{(1)}(\mu) + \frac{1}{2}(x-\mu)^2 u^{(2)}(\mu) + ... + \frac{1}{n!}(x-\mu)^n u^{(n)}(\mu) + R_X.$$

Assuming that the remainder R_X is negligible (or $u^{(n+1)} = 0$), then the expected utility could be expressed as:

$$Eu(F) = u(\mu) + \frac{1}{2}C_F^{(2)}u^{(2)}(\mu) + ... + \frac{1}{n!}C_F^{(n)}u^{(n)}(\mu) + R_n^F.$$

Similarly, for $u(G)$ we have:

$$Eu(G) = u(\mu) + \frac{1}{2}C_G^{(2)}u^{(2)}(\mu) + ... + \frac{1}{n!}C_G^{(n)}u^{(n)}(\mu) + R_n^G.$$

Under Theorem 2.27, $u^{(i)} = 0$ for any $i > n$ and thus the difference of expected utility between F and G could be expressed as:

$$Eu(F) - Eu(G) = \frac{1}{2}u^{(2)}(\mu)(C_F^{(2)} - C_G^{(2)}) + \cdots + \frac{1}{n!}u^{(n)}(\mu)(C_F^{(n)} - C_G^{(n)}).$$

Thus, $X\ M_{RA}^n\ Y$ or $X\ M_{RS}^n\ Y$ implies $Eu(F) - Eu(G) \geq 0$ under its corresponding utility assumption.

Under Theorem 2.28, both R_n^F and R_n^G are ignorable. Hence,

$$Eu(F) - Eu(G) \approx \frac{1}{2}u^{(2)}(\mu)(C_F^{(2)} - C_G^{(2)}) + \cdots + \frac{1}{n!}u^{(n)}(\mu)(C_F^{(n)} - C_G^{(n)}),$$

and thus, $X\ M_{RA}^n\ Y$ or $X\ M_{RS}^n\ Y$ implies $Eu(F) - Eu(G) \geq 0$ under its corresponding utility assumption.

Proof of Theorem 2.29:

Consider first the M_{RA}^2 rule. Assume that $u^{(3)} = 0$. Then we have $u^{(2)}(\mu_X) = u^{(2)}(\mu_Y) = c < 0$ and further:

$$\begin{aligned} Eu(F) &= u(\mu_X) + \frac{1}{2}C_F^{(2)}u^{(2)}(\mu_X); \\ Eu(G) &= u(\mu_Y) + \frac{1}{2}C_G^{(2)}u^{(2)}(\mu_Y). \end{aligned}$$

Thus the difference of expected utility between F and G could be expressed as:

$$Eu(F) - Eu(G) = u(\mu_X) - u(\mu_Y) + \frac{1}{2}[C_F^{(2)} - C_G^{(2)}]c.$$

Since XM_{RA}^2Y, we have $\mu_X \geq \mu_Y$ and $C_F^{(2)} \leq C_G^{(2)}$, thus we have $Eu(F) \geq Eu(G)$.

Now we turn to consider M_{RA}^3 rule. Assume now that $u^{(4)} = 0$. Then we have $u^{(3)}(\mu_X) = u^{(3)}(\mu_Y) = c > 0$ and further:

$$\begin{aligned} Eu(F) &= u(\mu_X) + \frac{1}{2}C_F^{(2)}u^{(2)}(\mu_X) + \frac{1}{3!}C_F^{(3)}u^{(3)}(\mu_X); \\ Eu(G) &= u(\mu_Y) + \frac{1}{2}C_G^{(2)}u^{(2)}(\mu_Y) + \frac{1}{3!}C_F^{(3)}u^{(3)}(\mu_Y). \end{aligned}$$

Thus the difference of expected utility between F and G could be expressed as:

$$Eu(F) - Eu(G) = u(\mu_X) - u(\mu_Y) + \frac{1}{2}[C_F^{(2)}u^{(2)}(\mu_X) - C_G^{(2)}u^{(2)}(\mu_Y)] + \frac{1}{3!}c(C_F^{(3)} - C_G^{(3)}).$$

Since $u^{(3)} > 0$ and $\mu_X \geq \mu_Y$, we have $u^{(2)}(\mu_X) \geq u^{(2)}(\mu_Y)$. It follows that

$$\begin{aligned} & C_F^{(2)}u^{(2)}(\mu_X) - C_G^{(2)}u^{(2)}(\mu_Y) \\ = \; & C_F^{(2)}(u^{(2)}(\mu_X) - u^{(2)}(\mu_Y)) + u^{(2)}(\mu_Y)(C_F^{(2)} - C_G^{(2)}) \geq 0. \end{aligned}$$

Since XM_{RA}^3Y, we have $\mu_X \geq \mu_Y$, $0 \leq C_F^{(2)} \leq C_G^{(2)}$ and $C_F^{(3)} \geq C_G^{(3)}$, thus we have $Eu(F) \geq Eu(G)$. The proofs for Part 2 in Theorem 2.29 are similar and thus omitted here. □

Proof of Theorem 2.30:

Similar to the derivations in the proof for Theorem 2.29, we can have:

$$\begin{aligned} Eu(F) - Eu(G) &= u(\mu_X) - u(\mu_Y) + \frac{1}{2}[C_F^{(2)}u^{(2)}(\mu_X) - C_G^{(2)}u^{(2)}(\mu_Y)] + \cdots \\ & \quad + \frac{1}{n!}[C_F^{(n)}u^{(n)}(\mu_X) - C_G^{(n)}u^{(n)}(\mu_Y)]. \end{aligned}$$

We need to consider the sign of the terms $A_i = C_F^{(i)}u^{(i)}(\mu_X) - C_G^{(i)}u^{(i)}(\mu_Y), i = 2, \cdots, n$. First consider the even numbers i. In this situation, we have

$$u^{(i)} \leq 0, u^{(i+1)} \geq 0, C_F^{(i)} \leq C_G^{(i)} \text{ and } C_F^{(i)} \geq 0.$$

Then we get:

$$A_i = C_F^{(i)}(u^{(i)}(\mu_X) - u^{(i)}(\mu_Y)) + u^{(i)}(\mu_Y)(C_F^{(i)} - C_G^{(i)}) \geq 0.$$

Now assume that i is odd number. We have:

$$u^{(i)} \geq 0, u^{(i+1)} \leq 0, C_F^{(i)} \geq C_G^{(i)}.$$

Then we get:

$$A_i = u^{(i)}(\mu_X)(C_F^{(i)} - C_G^{(i)}) + C_G^{(i)}(u^{(i)}(\mu_X) - u^{(i)}(\mu_Y)).$$

A sufficient condition for A_i being positive is that $C_G^{(i)} \leq 0$. As a result, Part 1 of Theorem 2.30 holds. The proofs for Part 2 in Theorem 2.30 are similar and thus omitted here. □

2.11 References

Astebro, T. (2003). The return to independent invention: evidence of unrealistic optimism, risk seeking or skewness loving?. Economic Journal 113(484), 226-239.

Astebro, T., Mata, J., & Santos-Pinto, L. (2015). Skewness seeking: risk loving, optimism or overweighting of small probabilities?. Theory and Decision 78, 189-208.

Bai, Z.D., Guo, X., Li, H., & Wong, W.K. (2022). Stochastic Dominance with Applications in Economics, Finance, and Income Inequality, World Scientific Publishing, forthcoming.

Bai, Z.D., Li, H., Liu, H.X., Wong, W.K., 2011. Test statistics for prospect and Markowitz stochastic dominances with applications. Econometrics Journal, 122, 1-26.

Bai, Z.D., Li, H., McAleer, M., Wong, W.K., 2015. Stochastic dominance statistics for risk averters and risk seekers: An analysis of stock preferences for USA and China. Quantitative Finance, 15(5), 889-900.

Barrett, G. F., & Donald, S. G. (2003). Consistent tests for stochastic dominance. Econometrica, 71(1), 71-104.

Bawa, V.S., Bodurtha Jr.J.N., Rao, M.R., & Suri, H.L. (1985). On determination of stochastic dominance optimal sets. Journal of Finance, 40(2), 417-431.

Borch, K. (1979). Utility and stochastic dominance. In Expected utility hypotheses and the Allais paradox (pp. 193-201). Springer, Dordrecht.

Brockett, P.L., & Garven, J.R. (1998). A Reexamination of the Relationship Between Preferences and Moment Orderings by Rational Risk-Averse Investors, GENEVA Papers on Risk and Insurance Theory, 23(2), 127-137.

Buckle, M., Chen, J., & Williams, J. M. (2016). Realised higher moments: Theory and practice. European Journal of Finance 22(13), 1272-1291.

Chan, R.H., Clark, E., Guo, X., & Wong, W.K. (2020). New Development on the Third Order Stochastic Dominance for Risk-Averse and Risk-Seeking Investors with Application in Risk Management, Risk Management 22, 108-132.

Charness, G., Karni, E., & Levin, D. (2007). Individual and group decision making under risk: An experimental study of Bayesian updating and violations of first-order stochastic dominance. Journal of Risk and uncertainty, 35(2), 129-148.

Chui, D., Cheng, W. W., Chow, S. C., & Ya, L. I. (2020). Eastern Halloween effect: A stochastic dominance approach. Journal of International Financial Markets, Institutions and Money, 68, 101241.

Chiu, W. H. (2010). Skewness Preference, Risk Taking and Expected Utility Maximisation, Geneva Risk and Insurance Review 35(2), 108-129.

Choi, P., & Nam, K. (2008). Asymmetric and leptokurtic distribution for heteroscedastic asset returns: The SU-normal distribution. Journal of Empirical Finance,15(1), 41-63.

Clark, E., Qiao, Z., & Wong, W. K. (2016). Theories of risk: Testing investor behavior on the taiwan stock and stock index futures markets. Economic Inquiry, 54(2), 907-924.

Copeland, T.E., Weston, J.F., & Shastri, K. (2005). Financial Theory and Corporate Policy, 4th Edition, Addison-Wesley.

Cvitanić, J., Polimenis, V., & Zapatero, F. (2008). Optimal portfolio allocation with higher moments. Annals of Finance,4(1), 1-28.

Davidson, R., & Duclos, J. Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. Econometrica, 68(6), 1435-1464.

Denuit, M., Eeckhoudt, L., Tsetlin, I., & Winkler, R.L. (2013). Multivariate concave and convex stochastic dominance. Biagini, F., Richter, A., Schlesinger, H., eds. Risk Measures and Attitudes (Springer, London), 11-32).

Do, H.X., Brooks, R., Treepongkaruna, S., & Wu, E. (2016). Stock and currency market linkages: New evidence from realized spillovers in higher moments. International Review of Economics Finance 42, 167-185.

Egozcue, M., Fuentes García, L., Wong, W.K., & Zitikis, R. (2011). Do Investors Like to Diversify? A Study of Markowitz Preferences, European Journal of Operational Research 215(1), 188-193.

Egozcue, M., & Wong, W.K. (2010). Gains from diversification: A majorization and stochastic dominance approach. European Journal of Operational Research, 200(3), 893-900.

Ekern, S. (1980). Increasing Nth degree risk, Economics Letters 6(4), 329-333.

Fang, Y., & Post, T. (2017). Higher-degree stochastic dominance optimality and efficiency. European Journal of Operational Research, 261(3), 984-993.

Fang, Y., & Post, T. (2022). Optimal portfolio choice for higher-order risk averters. Journal of Banking & Finance, 106429.

Fishburn, P. C. (1980). Stochastic dominance and moments of distributions. Mathematics of Operations Research, 5(1), 94-100.

Grigoletto, M., & Lisi, F. (2011). Practical implications of higher moments in risk management. Statistical Methods and Applications,20(4), 487-506.

Guo, X., Post, T., Wong, W.K., & Zhu, L.X. (2014), Moment Conditions for Almost Stochastic Dominance, Economics Letters, 124(2), 163-167.

Guo, X., & Wong, W.K. (2016). Multivariate Stochastic Dominance for Risk Averters and Risk Seekers, RAIRO - Operations Research 50(3), 575-586.

Hadar, J., & Russell, W.R. (1969). Rules for Ordering Uncertain Prospects, American Economic Review 59, 25-34.

Hadar, J., & Russell, W.R. (1971). Stochastic Dominance and Diversification. Journal of Economic Theory 3(3), 288-305.

Hanoch G., & Levy, H. (1969). The Efficiency Analysis of Choices Involving Risk. Review of Economic studies 36, 335-346.

Hardy, G.H., Littlewood, J.E., & Pólya, G. (1934). Inequalities, Cambridge University Press, Cambridge, MA.

Harvey, C.R., Liechty, J.C., Liechty, M.W., & Müller, P. (2010). Portfolio selection with higher moments. Quantitative Finance 10(5), 469-485.

Hoang, V.T.H., Wong, W.K., & Zhu, Z.Z., 2015. Is gold different for risk-averse and risk-seeking investors? An empirical analysis of the Shanghai Gold Exchange. Economic Modelling, 50, 200-211.

Jean, W.H. (1975). Comparison of moment and stochastic dominance ranking methods, Journal of Financial and Quantitative Analysis 10(1), 151-161.

Jondeau, E., & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: Existence, persistence, and comovements. Journal of Economic dynamics and Control 27(10), 1699-1737.

Jondeau, E., & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. European Financial Management,12(1), 29-55.

Kallio, M., Hardoroudi, N. D. (2019). Advancements in stochastic dominance efficiency tests. European Journal of Operational Research, 276(2), 790-794.

Kostakis, A., & Muhammad, K., Siganos, A. (2012). Higher co-moments and asset pricing on London stock exchange. Journal of Banking Finance 36, 913-922.

Kraus, A., & Litzenberger, R. (1976). Skewness preference and the valuation of risk assets. The Journal of Finance,31, 1085-1100.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-292.

Lambert, M., & Hubner, G. (2013). Comoment risk and stock returns. Journal of Empirical Finance 23, 191-205.

Lehmann, E.L. (1951). Consistency and Unbiasedness of Certain Nonparametric Tests, Annals of Mathematical Statistics 22(2), 165-179.

Lehmann, E.L. (1952). Testing Multiparameter Hypotheses, Annals of Mathematical Statistics 23(4), 541-552.

Lehmann, E.L. (1955). Ordered Families of Distributions, Annals of Mathematical Statistics 26(3), 399-419.

Levy, H. (2015). Stochastic dominance: Investment decision making under uncertainty. Third Edition, Springer, New York.

Levy, H., & Levy, M. (2004). Prospect Theory and Mean-Variance Analysis. The Review of Financial Studies, 17(4), 1015-1041.

Levy, H., & Wiener, Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. Journal of Risk and Uncertainty, 16(2), 147-163.

Levy, M., & Levy, H. (2002). Prospect Theory: Much Ado About Nothing? Management Science, 48(10), 1334-1349.

Li, C.K., & Wong, W.K. (1999). Extension of stochastic dominance theory to random variables. RAIRO - Operations Research 33(4), 509-524.

Lim, K. (1989). A new test of the three-moment capital asset pricing model. Journal of Financial and Quantitative Analysis,24, 205-216.

Lv, Z., Chu, A. M., Wong, W. K., & Chiang, T. C. (2021). The maximum-return-and-minimum-volatility effect: evidence from choosing risky and riskless assets to form a portfolio. Risk Management, 23(1), 97-122.

Ma, C., & Wong, W.K. (2010). Stochastic dominance and risk measure: A decision-theoretic foundation for VaR and C-VaR. European Journal of Operational Research, 207(2), 927-935.

Markowitz, H.M. (1952a). Portfolio selection. Journal of Finance, 7, 77-91.

Markowitz, H. (1952b). The utility of wealth. Journal of political Economy, 60(2), 151-158.

Martellini, L., & Ziemann, V. (2009). Improved estimates of higher-order comoments and implications for portfolio selection. Review of Financial Studies 23(4), 1467-1502.

Meyer, T.O., Li, X.M., & Rose, L.C. (2005). Comparing mean variance tests with stochastic dominance tests when assessing international portfolio diversification benefits. Financial Services Review 14(2), 149-168.

Muliere, P., & Scarsini, M. (1989). A note on stochastic dominance and inequality measures? Journal of Economic Theory 49(2), 314-323.

Ng, P., Wong, W. K., & Xiao, Z. (2017). Stochastic dominance via quantile regression with applications to investigate arbitrage opportunity and market efficiency. European Journal of Operational Research, 261(2), 666-678.

Ogryczak, W., & Ruszczyński, A. (1999). From stochastic dominance to mean-risk models: Semideviations as risk measures. European Journal of Operational Research 116(1), 33-50.

Ortobelli Lozza, S., Wong, W. K., Fabozzi, F. J., & Egozcue, M. (2018). Diversification versus optimality: is there really a diversification puzzle?. Applied Economics, 50(43), 4671-4693.

Perez-Quiros, G., & Timmermann, A. (2001). Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities. Journal of Econometrics 103(1), 259-306.

Post, T. (2003). Empirical tests for stochastic dominance efficiency. Journal of Finance, 58(5), 1905-1931.

Post, T., & Kopa, M. (2017). Portfolio choice based on third-degree stochastic dominance. Management Science, 63(10), 3381-3392.

Post, T., & Kopa, M. (2013). General linear formulations of stochastic dominance criteria. European Journal of Operational Research, 230(2), 321-332.

Post, T., & Levy, H. (2005). Does risk seeking drive stock prices? A stochastic dominance analysis of aggregate investor preferences and beliefs. Review of Financial Studies, 18(3), 925-953.

Post, T., & Versijp, P. (2007). Multivariate tests for stochastic dominance efficiency of a given portfolio. Journal of Financial and Quantitative Analysis, 42(2), 489-515.

Qiao, Z., Clark, E., & Wong, W. K. (2014). Investors¡¦ preference towards risk: Evidence from the Taiwan stock and stock index futures markets. Accounting & Finance, 54(1), 251-274.

Ritchken, P., & Kuo, S. (1989). On stochastic dominance and decreasing absolute risk averse option pricing bounds. Management Science, 35(1), 51-59.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. Journal of Economic theory, 2(3), 225-243.

Samuelson, P.A., 1967. General Proof that Diversification Pays. Journal of Financial and Quantitative Analysis 2(1), 1-13.

Scott, R. C., & Horvath, P. A. (1980). On the direction of preference for moments of higher order than the variance. Journal of Finance 35(4), 915-919.

Slovic, P. (1964). Assessment of risk taking behavior. Psychological Bulletin, 61(3), 220-233.

Tehranian, H. (1980). Empirical Studies in Portfolio Performance Using Higher Degrees of Stochastic Dominance, Journal of Finance 35(1), 159-171.

Tesfatsion, L.(1976). Stochastic Dominance and Maximization of Expected Utility. Review of Economic Studies 43, 301-315.

Tobin, J. (1958). Liquidity Preference as Behavior towards Risk. Review of Economic Studies 25,65-86.

Tsang, C. K., Wong, W. K., & Horowitz, I. (2016). Arbitrage opportunities, efficiency, and the role of risk preferences in the Hong Kong property market. Studies in Economics and Finance.

Vinod, H. D. (2004). Ranking mutual funds using unconventional utility theory and stochastic dominance. Journal of Empirical Finance, 11(3), 353-377.

Vo, X.V., & Tran, T.T.A. (2020). Higher-order comoments and asset returns: evidence from emerging equity markets. Annals of Operations Research, forthcoming.

von Neumann, J., & Morgenstern, O. (1944). Theory of Games and Economic Behavior. Princeton University Press, Princeton N.J.

Whitmore G.A. (1970). Third-degree Stochastic Dominance. American Economic Review 60, 457-459.

Wong, W.K. (2006), Stochastic Dominance Theory for Location-Scale Family. Journal of Applied Mathematics and Decision Sciences, now Advances in Decision Sciences 2006, 1-10.

Wong, W.K. (2007). Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. European Journal of Operational Research, 182(2), 829-843.

Wong, W.K., & Chan, R. (2008). Prospect and Markowitz Stochastic Dominance. Annals of Finance, 4(1), 105-129.

Wong, W.K., & Li., C.K. (1999). A note on convex stochastic dominance. Economics Letters 62, 293-300.

Wong, W.K., & Ma, C. (2008). Preferences over Location-Scale Family, Economic Theory 37(1), 119-146.

Wong, W. K., Phoon, K. F., & Lean, H. H. (2008). Stochastic dominance analysis of Asian hedge funds. Pacific-Basin finance journal, 16(3), 204-223.

Yalonetzky, G. (2013). Stochastic dominance with ordinal variables: Conditions and a test. Econometric Reviews, 32(1), 126-163.

Chapter 3

Eastern Halloween effect: A stochastic dominance approach

Joint work with David Chui, Wui Wing Cheng, and Ya LI

Published in Journal of International Financial Markets, Institutions and Money

The Eastern market wisdom of "May is poor, June is bleak, and July will turn around" unveils an international stock markets dynamic that lower returns in May followed by even worse return in June but rebounding back to an upward trend in July. This wisdom is termed as "Eastern Halloween" effect in this chapter which has some similarities with the traditional Halloween Effect. However, the traditional risk approaches fail to explain this Eastern Halloween effect satisfactorily. We apply stochastic dominance method to examine the Eastern Halloween effect and provide solid evidence of existence of Eastern Halloween effect.

3.1 Overview

Bouman and Jacobsen (2002) first recorded the existence of a seasonal pattern based on an old market wisdom "sell in May and go away" or commonly known as Halloween effect. The study focused in the U.S. stock market as well as 35 other world markets including developed and emerging ones. Strong Halloween Effect was observed for 10 markets with high significance level of one percent. Also, average market return during winter period (namely from November to April cycle), was found to be consistently higher than that during summer period (namely from May to October cycle). The finding of such market inefficiency pointed to a simple timing strategy that investors would be fully invested for December to April (six month) period and out of the stock market for May to October (the other six months) period to yield a superior return with the inclusion of transaction costs. In a subsequent paper, Jacobsen and Visaltanachoti (2009) further document the disper sions between summer period returns and winter period returns in a micro level study across various U.S. industries where consumer sectors tend to outperform the market during summer period while production sectors tend to outperform during winter period. In a recent study, Andrade, Chhaochharia and Fuerst (2013) update Bouman and Jacobsen (2002) and confirm the Halloween effect still exists in the sample period 1998 to 2012.

In a different approach, Kamstra et al. (2003) used a seasonal affective disorder (SAD) factor to draw relevancy of investors' sentiment to the change of seasons. Depressed investors became more risk averse during fall season or earlier winter season due to lack of sunlight hence higher expected return is required for late winter. Cao and Wei (2005), on the other hand, argued that apathy dominates aggression when temperature is high which lead to lower contemporaneous average return during the summer. However, Jacobsen and Marquering (2008, 2009) showed that the simple summer/winter seasonal indicator subsumes the return predictability of the complex SAD and weather variables. This adds flavor to the conventional market wisdom interpretation that most investors have summer vacation suggested by Bouman and Jacobsen (2002). Indeed, Hong and Yu (2009) found that stock market trading volume is lower in summer. If investor participation is lower in summer, market wide risk bearing capacity drops and generate higher expected future return for winter. When most investors resume trading in winter and participation becomes higher, overall risk bearing capacity rises and generate lower expected future return for summer.

While the Halloween effect has become a well-established phenomenon and a focal point of study, the East has a similar adage to describe the market calendar condition, namely "May is poor, June is bleak, and July will turn around". This saying holds that investors can outperform a simple buy-and-hold strategy by selling stocks at the beginning of May and buying them back at the beginning of July or termed as "Eastern Halloween" effect in this chapter.

As Eastern Halloween effect has yet to be a recognized market anomaly, there has been limited study in this front. This wisdom (or saying) is originated from Hong Kong when institutional analysts find the pattern that low returns in May and June are followed by an upward trend in July. One of the mainstream explanations is that companies in Hong Kong normally distribute dividend in May and June. The fiscal year end for most companies in Hong Kong lies on March 31. Firms announce their dividend policy and make payment two months later. According to the price behavior, the stock price will naturally fall until the investors expected the next interim dividend after the summer vacation.

In this chapter, we firstly examine the Eastern Halloween effect to "Eastern" markets that has a proximity to Greater China region and then extend the Eastern Halloween study to Western markets. In the chosen six Eastern markets and six Western markets, we observe 4 out of 6 markets in each region are significant at 10-percent level. Within this portfolio of 8 markets, Hong Kong, Singapore, France, Germany, Italy and UK have a 5-percent significant level. A natural follow up question to ask is "Is the higher risk in July to April compensated for the higher return?" To do this, we used volatility index and GARCH approach to model the risk. Our results reveal some interesting insights. Although the returns differ in Eastern Halloween period and the rest of the calendar year, the standard deviation in these two distinct periods remains fairly constant. The U.S. is the only market displaying significant different risk levels measured by volatility index. In analyzing this further, we used GARCH (1, 1) and optimal GARCH (p, q) approach to model time-varying volatility more explicitly. In both measures, we cannot find discernable risk difference in the two distinct periods. This inconsistency between return and volatility is also noted in Fama and French (2018) where they found the discrepancy between equity premium and volatility in the market return.

Overall, the evidence is discouraging for efficient market advocates. In other words, the failure of existing risk story to explain the Eastern Halloween effect. As the Eastern Halloween effect cannot be explained by risk, we use stochastic dominance approach to seek if it can be a better tool to explain the effect. Fong et al. (2005) first apply the stochastic dominance approach to explain international momentum strategies. The primary advantage of stochastic dominance approach is that it provides a very general framework to assess market dynamics without the need for asset pricing benchmarks. In addition, stochastic dominance theory makes no assumptions about the distribution of stock market returns and minimal assumptions about investors' utility functions. Using stochastic dominance, the testing hypothesis is "If May or June return is stochastically dominated by other months at first order, then it is unlikely that the problem is due to omitted risk factors but more likely that Eastern Halloween reflects market inefficiency". If such hypothesis is true, investors should buy the stochastically dominant portfolio and sell the stochastically dominated one to increase their expected utility. Specifically, we find that May and June are stochastically dominated by February, March, April, July, October, November and December.

The contributions to the literature are twofold. To the best of our knowledge, it is the first paper to examine the Eastern Halloween effect in an international context including Eastern and Western stock markets. The traditional Halloween effect in US and international markets has been well-researched. However, little attention has been paid to the Eastern wisdom of the market dynamics. Our study shows that the Eastern wisdom exists not only in Eastern markets, but also in Western markets. We have demonstrated that a simple random walk model does not hold. Monetary policy makers need to pay attention to the underlying stock return stability and monthly distribution. Murtazashvili and Vozlyublennaia (2012) document the seasonal impact on Capital Asset Pricing Model (CAPM) parameters. Second, we believe this paper is the first to apply stochastic dominance approach to explain the Eastern Halloween effect where risk-based explanations fail to account fully for the Eastern Halloween effect. Using stochastic dominance theory, this paper attempts to shed light on this new market dynamics by seeking if utility theory rather than risk theory is relevant in explaining the effect.

The rest of this paper is organized in the following order. Section 3.2 provides literature review on Halloween effect and efficient market hypothesis. Section 3.3 describes the data and shows the profitability of Eastern Halloween wisdom in the international market and the risk explanation. An overview of stochastic dominance theory is presented in Section 3.5.1. Results of stochastic dominance tests for Eastern Halloween effect are reported in Section 3.6. Section 3.7 summarizes the findings and makes concluding remarks.

3.2 Literature review

Bouman and Jacobsen (2002) first rigorously analyze the Halloween Effect for 37 international stock markets from January 1970 to August 1998. They found that the sell-in-May effect was present in 36 out of 37 markets and was statistically significant in 20 out of 37 markets. Bouman and Jacobsen showed a return in the November-to-April time frame 11% higher than in the May-to-October time frame in the US market; the return difference is 24% in the UK which went as far back to the year 1694! The sell-in-May effect has been around for a very long time, and, as it requires only two trades per year, it persists even after trading costs.

However, Maberly and Pierce (2003) documented that the Halloween effect disappeared in the US stock market once the months of October 1987 and August 1998 of Sven Bouman and Ben Jacobsen's original dataset were excluded. In a decade later, Andrade, Chlaochharia and Fuerst gave the counter argument in 2013. Testing the sell-in-May effect with out-of sample data from November 1998 through April 2012, they found that in the 14 years since the publication of Bouman and Jacobsen's original analysis, the indicator did not disappear. In fact, on average, across the 37 markets studied, the out-performance in the winter months was still about 10% higher than in the summer months. They also found that the effect did not come in lumps. The effect existed in three out of four years and did not depend on specific industries, markets, or months. It was clear that the effect was both real and persistent.

Calendar-based anomaly study has a long history in the Finance literature. Earlier studies focus on the January effect, a monthly seasonality by which the stock return tends to be higher in January than other months (e.g., Rozeff and Kinney, 1976; Keim, 1983; Reinganum, 1983; Tinic and West, 1984). Later, the literature extends to the Halloween effect, a semiannual seasonality by which stock return tends to be higher in the six-month winter period than in the six-month summer period (e.g., Bouman and Jacobsen, 2002; Kamstra et al., 2003). More recently, Heston and Sadka (2008) establish a calendar effect by which stocks tend to have relatively high (low) returns every year in the same calendar month. Heston and Sadka (2010) provide out-of-sample evidence towards this pattern from 14 international markets. Keloharju, Linnainmaa, and Nyberg (2016) further document the same-calendar-month effect in portfolios sorted by anomaly variables as well as commodities.

The January effect is a long-lasting seasonal anomaly in the US stock market whereby the returns in January are higher than in any other month (Wachtel, 1942; Officer, 1975). Rozeff and Kinney (1976) analyze the data from 1904 to 1974 and conclude that the average return for stocks during January is seven times greater than any other month during the year and the trend is particularly phenomenal in small-capitalization stocks. Keim (1983), Roll (1983), and Lakonishok and Smidt (1984) demonstrate that the abnormal performance of small stocks is what largely contributes to large January returns. The January effect is mainly focused in smaller firms. However, Riepe (1998, 2001) investigates the continuance of the January effect and finds diminishing trends of excess returns in January.

Bouman and Jacobsen (2002) first document high stock returns during May to October and low returns during November to April in the US stock market and 35 other markets, including European and emerging markets, out of 37 countries they examine. The findings point to a simple timing strategy that refrains from the equity market at the end of May until October to improve capital gains. Jacobsen and Visaltanachoti (2009) further document the dispersion between summer returns and winter returns in various industries in the United States. Returns during winter are higher than returns during summer for 48 out of 49 market sectors, where consumer sectors tend to outperform the market during summer while production sectors tend to outperform during winter. Unlike other anomalies, the effect has existed since it was first documented. For example, Andrade et al. (2013) identify the Halloween effect in the sample period 1998–2012. However, Lucey and Zhao (2008) show that the Halloween effect is most likely driven by January returns and may not be present in US equities in the long term.

To rationalize the Halloween effect, Kamstra et al. (2003) propose the seasonal affective disorder (SAD), where depressed investors become more risk averse during fall or early winter due to lack of sunlight, and hence a higher expected return is required for late winter. Alternatively, Cao and Wei (2005) argue that investors become less risk averse due to aggression driven by low temperatures in winter, and therefore the contemporaneous average return is higher. However, Jacobsen and Marquering (2008, 2009) show that the simple summer/winter seasonal indicator subsumes the return predictability of the complex SAD and weather variables and favors the conventional market wisdom interpretation that most investors have summer vacation, as suggested by Bouman and Jacobsen (2002). Indeed, Hong and Yu (2009) find that stock market trading volume is lower in summer.

Heston and Sadka (2008) present a new periodic oscillating pattern in stock returns in which the outperformance of winning stocks lasts in the same calendar month up to 20 years. Jegadeesh (1990) shows the positive annual returns for lags of 12, 24, and 36 months. In contrast to previous results of long-term reversal, Heston and Sadka (2008) document a general pattern lasting up to 240 months. Heston and Sadka (2010) extend their study to 14 international stock markets and document significant seasonalities. Li et al. (2018) study the seasonality in stock returns in both advanced markets and emerging markets. Despite much effort to document calendar anomalies, no studies have examined the seasonal pattern in stock market from a culture and geographic region approach such as "Eastern" approach.

In explaining market anomalies, behavioral finance offers an alternative theory to EMH. Shiller (2015) pointed out departures from efficiency in financial markets due to cognitive biases such as overconfidence and over-reaction, information bias and other human errors in reasoning and processing information. On the other hand, defenders of EMH argue that the biases are in individuals and not in the efficient market (Malkiel, 2003, Chap. 11). An attempt to reconcile the economic theories, based on EMH and behavioral finance, respectively, is the adaptive market hypothesis introduced by Lo (2004) who regards it as a modification of EMH derived from evolutionary principles. Earlier Lo and MacKinlay (1989) argued against random walk models assumed by EMH.

Under the evolutionary principles in an adaptive market, prices reflect as much information as dictated by the combination of economic conditions and the number and nature of different groups of market participants: pension fund managers, hedge fund managers, market makers, retail investors, etc. If multiple members of a single group are competing for scarce resources in a single market, then the market is likely to exhibit high efficiency (e.g., the Treasury notes market). On the other hand, if a small number of members are competing for abundant resources, the market is likely to be inefficient. The degree of market efficiency can change with time and is related to "environmental" factors characterizing "market ecology", such as the number of competitors in the market, profit opportunities, and how the market participants adapt to changing environments. Under the adaptive market hypothesis, much of what behavioral finance cites as counterexamples to economic rationality is consistent with the evolutionary principles of financial interactions: competition, adaptation and natural selection; see Lo (2004).

3.3 Data, profitability and volatility analysis

3.3.1 The Halloween period and summer vacation

Following Bouman and Jacobsen (2002) approach, we define the exogenous Halloween period as the period November and April. It follows that summer vacation is the period May to October. Alternatively, we follow Eastern wisdom which defines the Eastern Halloween period as the period July to April.

3.3.2 Market selection

We select G7 countries into our choice of six western markets including developed countries such as US, UK, Germany, France, Italy, and Canada. To capture emerging and mature economy in Eastern markets, we include China, Hong Kong, Korea and Singapore and Taiwan and Japan in as our choice of six eastern markets. The overall stock market performance is based on Morgan Stanley Capital Index (MSCI) World Index and the period of study covers monthly data from 1998 to 2018.[1]

3.4 International study

An interesting question that arises is whether these low returns during the period May-June are more or less evenly spread over these months in all markets. In Table 3.1, we report the difference between average monthly returns in summer and the rest of the year in six eastern markets and six western markets. Let $r_{i,t}$ be the stock monthly return, μ_i be a constant, $\varepsilon_{i,t}$ be the usual error term, and $EH_{i,t}$ be a dummy variable that takes the value 1 for the July-April period and 0 for the May-June period. Using pooled data, α is the coefficient on the Eastern Halloween dummy variable. When α is significantly positive, this rejects the null hypothesis of no Eastern Halloween Effect. Taking variation over time into consideration, we add year

[1] Using monthly data is a common practice in testing Halloween effect. Examples are Bouman and Jacobsen (2002), Andrade, Chhaochharia and Fuerst (2013), Guo, Luo and Zhang (2014), Dichtl and Drobetz (2015), etc. One of the reasons for using monthly data is that the data is less noisy compare to daily data and still could provide reasonable sample size compare to use returns of multiple months.

dummies to control for the year fixed effect. When dummies for time are included (in this case: year dummies), we estimate α only within-year variation. In the overall regression, we add both year-fixed effect and country-fixed effect.

$$r_{i,t} = \mu_i + \alpha EH_{i,t} + \varepsilon_{i,t}$$

We observe that the returns tend to be below average from May to June in Hong Kong, Japan, Singapore, Taiwan, France, Germany, Italy, and UK. Four out of six Eastern markets have significantly lower returns in May and June. Similarly, the mean return from July to April is larger than from May to June in four western markets. The overall returns of the May-June period are negative except for China, whereas the average monthly returns are positive in the July-Apr period. The results are significant at the 5% level for Hong Kong, Singapore, France, Germany, Italy, and UK.

Table 3.1: Market by Market Eastern Halloween Effect

	N	Average monthly return	Average monthly return of May-June period	Average monthly return of July-Apr period	Eastern Halloween Effect	t-value of Eastern Halloween Effect	Year fixed Effect	Country fixed Effect
Eastern Markets								
Overall	1326	0.0072	-0.0061	0.0562	0.0623	2.03***	Y	Y
China	181	0.0101	0.0040	0.0658	0.0618	0.70	Y	
Hong Kong	229	0.0058	-0.0025	0.0746	0.0771	2.32***	Y	
Japan	229	0.0027	-0.0043	0.0327	0.0370	1.86*	Y	
Korea	229	0.0162	-0.0198	0.0807	0.1005	1.42	Y	
Singapore	229	0.0058	-0.0068	0.0795	0.0863	2.13**	Y	
Taiwan	229	0.0033	-0.0067	0.0134	0.0201	1.69*	Y	
Western Markets								
Overall	1335	0.0031	-0.0185	0.0216	0.0401	2.32***	Y	Y
Canada	229	0.0065	-0.0016	0.0378	0.0394	0.91	Y	
France	216	0.0027	-0.0121	0.0479	0.0600	2.43***	Y	
Germany	216	0.0038	-0.0245	0.0296	0.0541	2.36***	Y	
Italy	216	-0.0008	-0.0278	0.0118	0.0396	3.27***	Y	
UK	229	0.0012	-0.0205	0.0125	0.0330	3.01***	Y	
USA	229	0.0047	-0.0040	0.0195	0.0235	1.25	Y	

This table reports annual average monthly returns and average monthly return during Eastern Halloween and Non-Eastern Halloween period. We also report the estimated slope coefficients of the monthly time-series regression where $EH_{i,t}$ is a dummy variable that equals one for July-April period and zero otherwise, taking account of year fixed effect. $r_{i,t} = \mu_i + \alpha EH_{i,t} + \varepsilon_{i,t}$. t-statistics based on Newey and West (1987) standard errors are shown in parentheses. ***, **, and * correspond to two-tail significance levels of 1%, 5%, and 10%, respectively. We add year-fixed effect and country-fixed effect for overall regression. Year-fixed effect is also added for each individual market.

3.4.1 Risk explanation for the Eastern Halloween effect

Hong and Yu (2009) first used time-varying volatility to explain the lower returns in summer period. They defined summer as the third quarter (July, August, and September) for Northern Hemisphere markets. In this chapter, volatility indexes from UK, US, France, Germany, Hong Kong, Korea and Japan are chosen to analyze market time varying volatility from a top-down approach and the criteria is simply a widely recognized and tradable benchmark. The Chicago Board Options Exchange volatility index (VIX) is a measure of the US stock market's expectation of volatility implied by S&P 500 index options which is a tradable volatility instrument since January 1990. It measures one standard deviation of the normal probability curve (68% confidence level) of the expected range of movement in the S&P 500 index. VIX is often called the "fear index" which is a decent indicator of market sentiment. Similarly, the FTSE Implied Volatility Index Series (VFTSE) is a series of end-of-day indices that measure the implied volatility of the FTSE 100 and FTSE MIB indices since January 2000. It is a forward-looking index which provides valid information about investors risk appetite.

CAC 40 VOLA index (VCAC) is based on CAC 40 which consists of the largest 40 companies listed on Euronext Paris since January 2000. The index is calculated based on VIX methodology and enhanced to account for specific characteristics of the underlying option markets. The VDA-New Index (VDAX) measures the implied volatility based on the next 30 days for the DAX option contracts which is a decent proxy of market volatility of German stock market since January 1992. In Eastern markets, HSI Volatility Index (VHSI) measures the volatility in Hong Kong stock market since January 2001. KOSPI Volatility Index (VKOSPI) measures the over market risk sentiment in Korea since January 2003. Nikkei Volatility Index (VNKY) is a decent indicator of Japanese overall stock market volatility since January 2001. Appendix 1 shows the chosen volatility indexes corresponding to each market.

Where $VOL_{i,t}$ is stock monthly volatility, μ_i a constant, $\varepsilon_{i,t}$ is the usual error term, $EH_{i,t}$ is a dummy variable takes the value 1 for July-April period and zero for May-June period. Using pooled data, α is the coefficient on Eastern Halloween dummy variable. When α is significantly positive, this rejects the null hypothesis of no Eastern Halloween Effect. Taking variation over time into consideration, we add year dummies to control for the year-fixed effect. When dummies for time are included (in this case: year dummies), we estimate α only within-year variation. In the overall regression, we add both year-fixed effect and country-fixed effect.

$$\mathrm{VOL}_{i,t} = \mu_i + a\mathrm{EH}_{i,t} + \varepsilon_{i,t}$$

The key risk results of the Eastern Halloween effect are presented in Table 3.2. The highest average monthly volatility index is in Japan at 25.31, while the lowest is in the US at 19.68. Also, we find that the volatility during the summer is not significantly lower than during the rest of the year except for the US market ($t = 2.34$).

TABLE 3.2: Market by Market Risk and Return (Volatility Index)

	N	Average monthly volatility	Average monthly volatility of May-June period	Average monthly volatility of July-Apr period	Eastern Halloween Effect	t-value of Eastern Halloween Effect	Year fixed Effect	Country fixed Effect
Eastern Markets								
Overall	555	23.45	22.43	23.88	1.45	0.89	Y	Y
Hong Kong	193	23.23	21.82	23.71	1.89	1.18	Y	
Japan	193	25.31	24.73	25.75	1.02	0.77	Y	
Korea	169	21.58	20.57	21.69	1.12	0.70	Y	
Western Markets								
Overall	1036	20.93	19.98	21.87	1.89	1.48	Y	Y
France	205	23.26	22.08	23.48	1.40	0.79	Y	
Germany	301	21.35	19.79	21.42	1.63	1.85	Y	
UK	205	19.94	18.92	20.34	1.42	1.16	Y	
USA	325	19.68	17.98	20.15	2.17	2.34***	Y	

This table reports annual average monthly volatility index and average monthly volatility index during Eastern Halloween and Non-Eastern Halloween period. We also report the estimated slope coefficients of the monthly time-series regression where $EH_{i,t}$ is a dummy variable that equals one for July-April period and zero otherwise. $VOL_{i,t} = \mu_i + \alpha EH_{i,t} + \varepsilon_{i,t}$. t-statistics based on Newey and West (1987) standard errors are shown in parentheses. ***, **, and * correspond to two-tail significance levels of 1%, 5%, and 10%, respectively. We add year-fixed effect and country-fixed effect for overall regression. Year-fixed effect is also added for each individual market.

We have also conducted additional analysis to discern whether there is lower volatility during summer. To generate a trading strategy based on the volatility prediction, we use a generalized autoregressive conditional heteroskedasticity (GARCH) model to predict stock market index volatility. We use past 10-day rolling window returns to generate GARCH (1,1) prediction and also use full sample returns to find the fitted GARCH (1,1) value Hansen and Lunde (2005) compared a large number of parametric volatility models in an extensive empirical study and found that no other model provides significantly better forecasts than the GARCH (1,1) model. In their paper, they found no evidence that a GARCH (1,1) is outperformed by more sophisticated model in their analysis.

Table 3.3: Market by Market Risk and Return GARCH (1,1)

	N	Average monthly volatility	Average monthly volatility of May-June period	Average monthly volatility of July-Apr period	Eastern Halloween Effect	t-value of Eastern Halloween Effect	Year fixed Effect	Country fixed Effect
Eastern Markets								
Overall	1326	7.20	7.22	7.16	-0.06	-0.18	Y	Y
China	181	7.72	7.77	7.60	-0.17	-0.45	Y	
Hong Kong	229	7.62	7.61	7.48	-0.13	-0.49	Y	
Japan	229	5.38	5.30	5.32	0.02	0.05	Y	
Korea	229	9.68	9.68	9.62	-0.06	-0.21	Y	
Singapore	229	7.59	7.64	7.77	-0.17	-0.46	Y	
Taiwan	229	7.56	7.43	7.32	-0.11	-0.61	Y	
Western Markets								
Overall	1335	5.78	5.82	5.72	-0.10	-0.29	Y	Y
Canada	229	5.81	5.79	5.64	-0.15	-0.80	Y	
France	216	5.94	5.67	5.52	-0.15	-0.52	Y	
Germany	216	6.95	6.89	6.77	-0.12	-0.38	Y	
Italy	216	6.79	6.56	6.65	0.09	0.11	Y	
UK	229	4.90	4.68	4.70	0.02	0.09	Y	
USA	229	4.39	4.52	4.38	-0.14	-0.83	Y	

This table reports annual average monthly GARCH(1,1) volatility and average monthly volatility during Eastern Halloween and Non-Eastern Halloween period. We also report the estimated slope coefficients of the monthly time-series regression where $EH_{i,t}$ is a dummy variable that equals one for July-April period and zero otherwise. The equation used in the regression is: $VOL_{i,t} = \mu_i + \alpha EH_{i,t} + \varepsilon_{i,t}$. The t-statistics are based on Newey and West (1987) standard errors, and they are shown in parentheses. The significance levels are represented as ***, **, and * for two-tail significance at 1%, 5%, and 10%, respectively. For the overall regression, we add year-fixed effect and country-fixed effect. Year-fixed effect is also added for each individual market.

However, Leeb and Pötscher (2008) argued against Hansen and Lunde (2005) by using Akaike Information Criterion (AIC) to produce the optimal subset model of GARCH (p,q) for forecasting purpose. Bayesian Information Criterion (BIC), on the other hand, can produce the optimal subset model of GARCH (p, q) for explanatory modelling. Based on the above work, our forecasting model uses lower order GARCH (p, q) models than a high order ARCH (p) for better numerical stability of estimation. Therefore, our composed model is based on GARCH (1,1) as the optimal model for stock returns in Germany, France, Italy, Canada, China, Korea, Taiwan and Japan while GARCH (2,1), GARCH (4,2), GARCH (2,3) and GARCH (3,3) as the optimal GARCH model for Hong Kong, Singapore, UK and USA respectively. The results in Tables 3.3 and 3.4 consistently show that there is no discernable difference in volatility between summer and winter seasons. Based on the sign of Eastern Halloween dummy variable, the GARCH model based volatility is lower during July to April when the returns are higher.

In the Appendix 2, we document the corresponding AIC and BIC for GARCH models. We also adopt CUSUM tests to identify whether there is a break in the volatility. Market volatility may change for a host of reasons including (both local and global) social, political or economic

Table 3.4: Market by Market Risk and Return optimal GARCH(p,q)

	N	Average monthly volatility	Average monthly volatility of May-June period	Average monthly volatility of July-Apr period	Eastern Halloween Effect	t-value of Eastern Halloween Effect	Year fixed Effect	Country fixed Effect
Eastern Markets								
Overall	1326	7.19	7.10	6.98	-0.12	-0.20	Y	Y
China	181	7.72	7.82	7.64	-0.18	-0.51	Y	
Hong Kong	229	7.62	7.45	7.26	-0.19	-1.04	Y	
Japan	229	5.38	5.97	5.98	0.01	0.02	Y	
Korea	229	9.68	9.62	9.50	-0.12	-0.28	Y	
Singapore	229	7.55	7.87	7.78	-0.09	-0.18	Y	
Taiwan	229	7.56	7.85	7.64	-0.21	-0.72	Y	
Western Markets								
Overall	1335	5.76	5.79	5.70	-0.09	-0.22	Y	Y
Canada	229	5.81	5.97	5.68	-0.29	-0.91	Y	
France	216	5.94	5.67	5.60	-0.07	-0.40	Y	
Germany	216	6.95	6.87	6.76	-0.11	-0.35	Y	
Italy	216	6.79	6.89	6.91	0.02	0.09	Y	
UK	229	4.90	4.67	4.82	0.15	0.83	Y	
USA	229	4.39	4.56	4.44	-0.12	-0.81	Y	

This table reports annual average monthly optimal GARCH(p,q) volatility and average monthly volatility during Eastern Halloween and Non-Eastern Halloween period. GARCH(1,1) is the optimal model for Germany, France, Italy, Canada, China, Korea, Taiwan, and Japan, while GARCH(2,1), GARCH(4,2), GARCH(2,3), and GARCH(3,3) are the optimal models for Hong Kong, Singapore, UK, and USA, respectively.We also report the estimated slope coefficients of the monthly time-series regression where $EH_{i,t}$ is a dummy variable that equals one for July-April period and zero otherwise. $VOL_{i,t} = \mu_i + \alpha EH_{i,t} + \varepsilon_{i,t}$. The t-statistics are based on Newey and West (1987) standard errors, and they are shown in parentheses. The significance levels are represented as ***, **, and * for two-tail significance at 1%, 5%, and 10%, respectively. For the overall regression, we add year-fixed effect and country-fixed effect. Year-fixed effect is also added for each individual market.

events. As a result, we observe sudden and substantial changes in volatility in each market. In the Appendix 3, we adopt the Bai and Perron (2003) multiple break tests. We identify the major breaks in the volatility. The major breaks happen during the 2003 SARS outbreak, 2008 global financial crisis and 2012 European sovereign debt crisis. Chen et al. (2010) deploy a simple, theoretical, single-beta CAPM model to explain the puzzle of low returns for high idiosyncratic volatility stocks. Chung and Chuwonganant (2018) also document the negative relationship between market volatility and stock returns. They emphasize more on the volatility shock.

3.5 Market efficiency and stochastic dominance

3.5.1 Concept of stochastic dominance

Consider prospect X and Y takes value over the real number space $\Omega = [a, b]$. To facilitate exposition, we let X and Y be random returns. For any outcome x, we define the s-order cumulative distribution functions $F^{(s)}$ and $G^{(s)}$ of X and Y to be:

$$F^{(s)}(x) = \int_a^x F^{(s-1)}(t)dt \qquad \text{and} \qquad G^{(s)}(x) = \int_a^x G^{(s-1)}(t)dt,$$

where $F^{(1)}(x) = F(x)$ and $G^{(1)}(x) = G(x)$. We now define stochastic dominance (SD) as follows:[2]

Definition 3.1. X is said to first (second)-order dominate Y by SD, denoted by $X \succ_1^A Y$ or $F \succ_1^A G$ $\big(X \succ_2^A Y$ or $F \succ_2^A G\big)$ if and only if $F^{(1)}(x) \leq G^{(1)}(x)$ $\big(F^{(2)}(x) \leq G^{(2)}(x)\big)$ for all $x \in [a, b]$.

We denote findings of first-, and second-order stochastic dominance by FSD and SSD. In comparing return distributions, the findings of FSD, for example FSD of X over Y $(X \succ_1^A Y)$ means that, for any return level x, distribution F will always have a lower return to get return less than x compared to distribution G. On the other hand, SSD of X over Y $(X \succ_2^A Y)$ means that, for any return level x, the integral of the cumulative probability of X is less than that of Y. SD findings are highly relevant to expected utility analysis. Hadar and Russell (1969) and Hanoch and Levy (1969) show that for the class of all von Neumann–Morgenstern utility functions, the following statement holds:

$$F \succ_s^A G \quad \text{if and only if} \quad E_F u(x) \geq E_G u(x) \ \text{ for all } \ u \in \mathcal{U}_s^A \ \text{ with } \ s = 1, 2.$$

Here, $E_F u(x)$ and $E_G u(x)$ are the expected utility of choosing X and Y respectively, $\mathcal{U}_1^A$ is defined for all $u' \geq 0$, and $\mathcal{U}_2^A \subset \mathcal{U}_1^A$ is defined by $u' \geq 0$ and $u'' \leq 0$. Because of this above statement, Falk and Levy (1989) argued that market efficiency could be tested by detecting the stochastic dominance relations. These rules help us to look at the whole distribution of returns

[2]The concept of stochastic dominance is proposed by Hadar and Russell (1969); Hanoch and Levy (1969), and Rothchild and Stiglitz (1970).

and not at certain parameters, e.g., mean and variance. With risk-averse assumption, Falk and Levy (1989) claimed that the existence of stochastic dominance relations did not conform to market efficiency. Since investors can buy the stochastically dominant portfolio and sell the stochastically dominated one, thereby increasing their expected utility, the market is said to be "inefficient" if such investment opportunities exist, and the market is said to be "efficient" if it is impossible for investors, independent of their utility function, to increase their respective expected utility. In an efficient market, we expect that the price of a dominant investment will instantaneously increase, and hence the return decreases, such that inequality of expected utility will not hold for any investor.

3.5.2 Hypothesis test with stochastic dominance

Having established the relationship between stochastic dominance and market efficiency in Section 3.5.1, the hypotheses are constructed for market efficiency testing:

$H_0^{i,j}$: Month i stochastic dominates month j

for $i = 1, 2, 3, 4, 11, 12$ and $j = 5, 6, 7, 8, 9, 10$ if Halloween effect is tested. And for $i = 1, 2, 3, 4, 7, 8, 9, 10, 11, 12$ and $j = 5, 6$ if eastern Halloween effect is tested. For $k = i, j$, let $F_k(x)$ be the distribution of month k's returns, and let $D_k^{(1)}(x) = F_k(x)$. Then recursively define $D_k^{(s)}(x) = \int_{-\infty}^{x} D_k^{s-1}(t)dt$ for $s \geq 2$. Let $\Delta_{i,j}^{(s)}(x) = D_i^{(s)}(x) - D_j^{(s)}(x)$, and $d_{i,j}^{(s)} = \sup_{x \in X} \Delta_{i,j}^{(s)}(x)$. The hypothesis $H_0^{i,j}$ and its corresponding alternative could be stated as:

$$H(s)_0^{i,j} : d_{i,j}^{(s)} \leq 0 \text{ vs. } H(s)_1^{i,j} : d_{i,j}^{(s)} > 0 \tag{A.1}$$

Define $\{X_{kt} : t = 1, \ldots, N, k = 1, \ldots, 12\}$ as the dataset. The test statistic for the above hypothesis is based on Linton et al. (2005) and Cho et al. (2007) as the following:[3]

$$D(s)_N^{i,j} = \sup_{x \in X} \sqrt{N} \left[\overline{D}_N^{(s)}\left(x; \overline{F}_i\right) - \overline{D}_N^{(s)}\left(x; \overline{F}_j\right) \right]$$

where

$$\overline{D}_N^{(s)}\left(x; \overline{F}_k\right) = \frac{1}{N(s-1)!} \sum_{i=1}^{N} \mathbf{1}(X_{ki} \leq x)(x - X_{ki})^{s-1} \quad \text{for} \quad k = 1, ..., 12$$

[3]There is other stochastic dominance test which also has interesting properties for example Bai et al. (2015).

According to Linton et al. (2005), the limiting distributions depend on the unknown true distributions of the dataset. Cho et al. (2007) suggest estimating the asymptotic p-values using resampling schemes such as bootstrapping and subsampling. The (re-centered overlapping) bootstrap procedure suggested by Cho et al. (2007) is described as follows:

1. Let $W_N^{i,j} = D(s)_N^{i,j}$ be the test statistic of the hypothesis in A.1 using the original full sample.

2. Generate subsamples (or blocks) of size b. That is, let $\{X_{kt} : t = 1, \ldots, N-b+1, k = 1, \ldots, 12\}$ be the first subsample, $\{X_{kt} : t = 2, \ldots, N-b+2, k = 1, \ldots, 12\}$ be the second subsample, and so on, until $\{X_{kt} : t = b, \ldots, N, k = 1, \ldots, 12\}$ is the last subsample.

3. Generate the bootstrap sample by sampling the $N-b+1$ overlapping blocks and laying them end-to-end in the order sampled. Repeat this for the number of bootstrap samples.

4. Compute re-centered test statistics W_N^{i,j^*} using the bootstrap sample as

$$W_N^{i,j^*} = D(s)_N^{i,j^*} = \sup_{x \in X} \sqrt{N}\left[\overline{D}_N^{(s)^*}\left(x; \overline{F}_i\right) - \overline{D}_N^{(s)^*}\left(x; \overline{F}_j\right)\right],$$

where

$$\overline{D}_N^{(s)^*}\left(x; \overline{F}_k\right) = \frac{1}{N(s-1)!}\sum_{i=1}^{N}\left\{\mathbf{1}(X_{ki} \le x)(x - X_{ki})^{s-1} - \omega(i,b,N)\mathbf{1}(X_{ki} \le x)(x - X_{ki})^{s-1}\right\},$$

for $k = 1, ..., 12$ and

$$\omega(i,b,N) = \begin{cases} i/b & \text{if } i \in [1, b-1] \\ 1 & \text{if } i \in [b, N-b+1] \\ (N-i+1)/b & \text{if } i \in [N-b+2, N] \end{cases} \tag{A.2}$$

Then, repeat calculate the re-centered test statistic using one of the unused bootstrap samples and the asymptotic p-value will be approximated by calculating the proportion of $W_N^{ij^*}$'s that exceeds W_N^{ij} in the repetitions. In this paper, we will focus on the case where $s = 1$, which is the case of FSD.

3.6 Eastern Halloween examined by stochastic dominance

Based on Linton et al. (2005); Cho et al. (2007), we apply stochastic dominance testing on Eastern Halloween and Western Halloween effects for both Eastern and Western markets. The first hypothesis is to test whether the monthly returns for 'participation period' (R_{pp}) stochastically dominates the monthly returns for 'go away period' (R_{gap}). If R_{pp} is not stochastically dominating R_{gap}, the hypothesis is 'Rejected' outright. To ensure the robustness of our testing, for those non-rejected monthly return under the first hypothesis test, we further test the hypothesis whether the monthly returns for 'go away period' (R_{gap}) stochastically dominates the monthly returns for 'participation period' (R_{pp}). In case under the second hypothesis testing, R_{gap} stochastically dominates R_{pp}, which shows that it contradicts to the first hypothesis, we then classified it as 'Inconclusive'.

However, if under the second hypothesis, R_{gap} stochastically dominates R_{pp}, is being rejected, we list out the dominance relationship in the table with the notation $F \succ_K^A G$ indicates the return distribution of F is stochastically dominates the distribution of G in k order. The results of stochastic dominance testing are shown in Tables 3.5, 3.6, 3.7 and 3.8. Tables 3.5 and 3.6 show the stochastic dominance tests of Western Halloween effect on Eastern and Western markets respectively.

Monthly return for i month (M_i) of the MSCI Index is compared to monthly return for j month (M_j) of the MSCI index where $i = 1, 2, 3, 4, 11, 12$ (Western 'participation period') and $j = 5, 6, 7, 8, 9, 10$ (Western 'go away period'). For Eastern markets in Table 3.5, it indicates that 19 out of the 36 relationships, month i is stochastically dominant month j, while 12 out of 36 results are found to be 'Inconclusive'. 5 out of 36 findings are 'Rejected'. The result indicates that the monthly returns within the participation period generally stochastic dominate the monthly returns within the 'go away period'. This implies that investors able to increase their utility by 'buy and hold' during the participation period and sell away at the beginning of the 'go away' period. The stochastic dominance result for Western Halloween Effect on Western markets is shown in Table 3.6. Similar to the findings on Eastern markets, 22 out of 36 relationships, and month i is found to stochastically dominates month j, while 6 out of 36 results are found to be 'Inconclusive'. 8 out 36 findings are 'Rejected'. This also implies that investors able to increase their utility by investing the stochastic dominated one in the participation period and sell away the one being stochastic dominated at the beginning of the 'go away period'.

TABLE 3.5: Stochastic Dominance tests of Eastern markets (Western Halloween effect)

$F \backslash G$	M_5	M_6	M_7	M_8	M_9	M_{10}
M_1	Inconclusive	Rejected	Rejected	$M_1 \succ_1^A M_8$	Inconclusive	Rejected
M_2	$M_2 \succ_1^A M_5$	Inconclusive	Rejected	$M_2 \succ_1^A M_8$	$M_2 \succ_1^A M_9$	Rejected
M_3	$M_3 \succ_1^A M_5$	$M_3 \succ_1^A M_6$	Inconclusive	$M_3 \succ_1^A M_8$	$M_3 \succ_1^A M_9$	Inconclusive
M_4	$M_4 \succ_1^A M_5$	$M_4 \succ_1^A M_6$	Inconclusive	$M_4 \succ_1^A M_8$	$M_4 \succ_1^A M_9$	Inconclusive
M_{11}	$M_{11} \succ_1^A M_5$	Inconclusive	Inconclusive	$M_{11} \succ_1^A M_8$	$M_{11} \succ_1^A M_9$	Inconclusive
M_{12}	$M_{12} \succ_1^A M_5$	$M_{12} \succ_1^A M_6$	Inconclusive	$M_{12} \succ_1^A M_8$	$M_{12} \succ_1^A M_9$	Inconclusive

Western Halloween period is defined as November to April. We test the stochastic dominance relationship for Eastern markets. The first hypothesis is to test whether the monthly returns for 'participation period' (R_{pp}) stochastically dominates the monthly returns for 'go away period' (R_{gap}). If R_{pp} is not stochastically dominating R_{gap}, the hypothesis is 'Rejected' outright. To ensure the robustness of our testing, for those non-rejected monthly returns under the first hypothesis test, we further test the hypothesis whether the monthly returns for 'go away period' (R_{gap}) stochastically dominates the monthly returns for 'participation period' (R_{pp}). In case under the second hypothesis testing, Rgap stochastically dominates Rpp, which shows that it contradicts the first hypothesis, we then classify it as 'Inconclusive'. However, if under the second hypothesis, Rgap stochastically dominates R_{pp}, is being rejected, we list out the dominance relationship in the table with the notation $F \succ_K^A G$ indicating that the return distribution of F stochastically dominates the distribution of G in k order.

The stochastic dominance result for Western Halloween Effect on Western markets is shown in Table 3.6. Similar to the findings on Eastern markets, 22 out of 36 relationships, and month i is found to stochastically dominates month j, while 6 out of 36 results are found to be 'Inconclusive'. 8 out 36 findings are 'Rejected'. This also implies that investors able to increase their utility by investing the stochastic dominated one in the participation period and sell away the one being stochastic dominated at the beginning of the 'go away period'.

As for Eastern Halloween Effect, through applying the same methodology as above, the monthly return for i month (M_i) of the MSCI Index is compared to monthly return for j month (M_j) of the MSCI index where $i = 1, 2, 3, 4, 7, 8, 9, 10, 11, 12$ (Eastern 'participation period') and $j = 5, 6$ (Eastern 'go away period'). It is shown in Table 3.7 and 3.8 for Eastern markets and Western markets respectively. In Table 3.7, there are 14 out of 20 relationships of month i stochastically dominated month j for Eastern markets, while only 3 out of 20 relationships are 'Inconclusive', and the number of 'Rejected' is also 3 out 20. As for Western markets, 15 out of 20 relationships shown that month i stochastically dominated month j under the Eastern Halloween effect. There are only 2 'Inconclusive' and 3 'Rejected' out of the 20 relationships.

Table 3.6: Stochastic Dominance tests of Western markets (Western Halloween effect)

$F \backslash G$	M_5	M_6	M_7	M_8	M_9	M_{10}
M_1	Rejected	Rejected	Rejected	Rejected	Inconclusive	Rejected
M_2	$M_2 \succ_1^A M_5$	$M_2 \succ_1^A M_6$	Inconclusive	$M_2 \succ_1^A M_8$	$M_2 \succ_1^A M_9$	Rejected
M_3	$M_3 \succ_1^A M_5$	$M_3 \succ_1^A M_6$	Inconclusive	$M_3 \succ_1^A M_8$	$M_3 \succ_1^A M_9$	Inconclusive
M_4	$M_4 \succ_1^A M_5$	$M_4 \succ_1^A M_6$	$M_4 \succ_1^A M_7$	$M_4 \succ_1^A M_8$	$M_4 \succ_1^A M_9$	Inconclusive
M_{11}	$M_{11} \succ_1^A M_5$	$M_{11} \succ_1^A M_6$	Inconclusive	$M_{11} \succ_1^A M_7$	Inconclusive	Rejected
M_{12}	$M_{12} \succ_1^A M_5$	$M_{12} \succ_1^A M_6$	$M_{12} \succ_1^A M_7$	$M_{12} \succ_1^A M_8$	$M_{12} \succ_1^A M_9$	$M_{12} \succ_1^A M_{10}$

Western Halloween period is defined as November to April. We test the stochastic dominance relationship for Western markets. The first hypothesis is to test whether the monthly returns for 'participation period' (R_{pp}) stochastically dominates the monthly returns for 'go away period' (R_{gap}). If R_{pp} is not stochastically dominating R_{gap}, the hypothesis is 'Rejected' outright. To ensure the robustness of our testing, for those non-rejected monthly returns under the first hypothesis test, we further test the hypothesis whether the monthly returns for 'go away period' (R_{gap}) stochastically dominates the monthly returns for 'participation period' (R_{pp}). In case under the second hypothesis testing, Rgap stochastically dominates Rpp, which shows that it contradicts the first hypothesis, we then classify it as 'Inconclusive'. However, if under the second hypothesis, Rgap stochastically dominates R_{pp}, is being rejected, we list out the dominance relationship in the table with the notation $F \succ_K^A G$ indicating that the return distribution of F stochastically dominates the distribution of G in k order.

Table 3.7: Stochastic Dominance tests of Eastern markets (Eastern Halloween Effect)

$F \backslash G$	M_5	M_6
M_1	Inconclusive	Rejected
M_2	$M_2 \succ_1^A M_5$	$M_2 \succ_1^A M_6$
M_3	$M_3 \succ_1^A M_5$	$M_3 \succ_1^A M_6$
M_4	$M_4 \succ_1^A M_5$	$M_4 \succ_1^A M_6$
M_7	$M_7 \succ_1^A M_5$	$M_7 \succ_1^A M_6$
M_8	Inconclusive	Rejected
M_9	Inconclusive	Rejected
M_{10}	$M_{10} \succ_1^A M_5$	$M_{10} \succ_1^A M_6$
M_{11}	$M_{11} \succ_1^A M_5$	$M_{11} \succ_1^A M_6$
M_{12}	$M_{12} \succ_1^A M_5$	$M_{12} \succ_1^A M_6$

Eastern Halloween period is defined as July to April. We test the stochastic dominance relationship for Eastern markets. The first hypothesis is to test whether the monthly returns for 'participation period' (R_{pp}) stochastically dominates the monthly returns for 'go away period' (R_{gap}). If R_{pp} is not stochastically dominating R_{gap}, the hypothesis is 'Rejected' outright. To ensure the robustness of our testing, for those non-rejected monthly returns under the first hypothesis test, we further test the hypothesis whether the monthly returns for 'go away period' (R_{gap}) stochastically dominates the monthly returns for 'participation period' (R_{pp}). In case under the second hypothesis testing, Rgap stochastically dominates Rpp, which shows that it contradicts the first hypothesis, we then classify it as 'Inconclusive'. However, if under the second hypothesis, Rgap stochastically dominates R_{pp}, is being rejected, we list out the dominance relationship in the table with the notation $F \succ_K^A G$ indicating that the return distribution of F stochastically dominates the distribution of G in k order.

Table 3.8: Stochastic Dominance tests of Western market (Eastern Halloween Effect)

$F \backslash G$	M_5	M_6
M_1	Rejected	Rejected
M_2	$M_2 \succ_1^A M_5$	$M_2 \succ_1^A M_6$
M_3	$M_3 \succ_1^A M_5$	$M_3 \succ_1^A M_6$
M_4	$M_4 \succ_1^A M_5$	$M_4 \succ_1^A M_6$
M_7	$M_7 \succ_1^A M_5$	$M_7 \succ_1^A M_6$
M_8	Inconclusive	$M_8 \succ_1^A M_6$
M_9	Inconclusive	Rejected
M_{10}	$M_{10} \succ_1^A M_5$	$M_{10} \succ_1^A M_6$
M_{11}	$M_{11} \succ_1^A M_5$	$M_{11} \succ_1^A M_6$
M_{12}	$M_{12} \succ_1^A M_5$	$M_{12} \succ_1^A M_6$

Eastern Halloween period is defined as July to April. We test the stochastic dominance relationship for Western markets. The first hypothesis is to test whether the monthly returns for 'participation period' (R_{pp}) stochastically dominates the monthly returns for 'go away period' (R_{gap}). If R_{pp} is not stochastically dominating R_{gap}, the hypothesis is 'Rejected' outright. To ensure the robustness of our testing, for those non-rejected monthly returns under the first hypothesis test, we further test the hypothesis whether the monthly returns for 'go away period' (R_{gap}) stochastically dominates the monthly returns for 'participation period' (R_{pp}). In case under the second hypothesis testing, Rgap stochastically dominates Rpp, which shows that it contradicts the first hypothesis, we then classify it as 'Inconclusive'. However, if under the second hypothesis, Rgap stochastically dominates R_{pp}, is being rejected, we list out the dominance relationship in the table with the notation $F \succ_K^A G$ indicating that the return distribution of F stochastically dominates the distribution of G in k order.

All the results shown that investors able to buy the stochastically dominant investment at the beginning of Eastern ‘participation period’, sell the one stochastically dominated at the beginning of the Eastern ‘go away period’ to increase their expected utility. This demonstrates that stochastic dominance for Eastern Halloween hypothesis can have better explanatory power than the traditional Halloween hypothesis. Indeed, stochastic dominance as a general approach based on expected utility maximization has been widely used in finance research. Clark and Kassimatis (2014) construct a portfolio-based investment strategy based on second and third degree stochastic dominance.

As a robustness check, test for the equality of the distribution of returns are performed and the results are shown in Table 3.9 and 3.10 by using the Kolmogorov-Smirnov test.

Table 3.9 shows that, for Eastern counties, 39 out of 44 results are consistent with the results of stochastic dominance test. That is, when the results of stochastic dominance test imply distribution of returns are different, most of the Kolmogorov-Smirnov test results agree. Table 3.10 shows a similar pattern, 37 out of 44 results are consistent with the results of stochastic

Table 3.9: Two sample Kolmogorov-Smirnov tests (Eastern markets)

$F\backslash G$	M_5	M_6	M_7	M_8	M_9	M_10
M_1	0.119	0.220***	0.239***	0.202**	0.101	0.211**
M_2	0.165*	0.138	0.156	0.193**	0.193**	0.138
M_3	0.156	0.183**	0.110	0.257***	0.183**	0.073
M_4	0.202**	0.220***	0.165*	0.294***	0.229***	0.092
M_7	0.220***	0.239***	NA	NA	NA	NA
M_8	0.119	0.147	NA	NA	NA	NA
M_9	0.073	0.239***	NA	NA	NA	NA
M_{10}	0.202**	0.193**	NA	NA	NA	NA
M_{11}	0.193**	0.138	0.147	0.202**	0.239***	0.147
M_{12}	0.266***	0.229***	0.110	0.294***	0.275***	0.128

As a robustness check, test for the equality of the distribution of returns are performed for Eastern markets. The results are shown by using the Kolmogorov-Smirnov test. ***, ** and * correspond to two-tail significance levels of 1%, 5% and 10%, respectively.

Table 3.10: Two sample Kolmogorov-Smirnov tests (Western markets)

$F\backslash G$	M_5	M_6	M_7	M_8	M_9	M_{10}
M_1	0.093	0.148	0.167*	0.120	0.093	0.343***
M_2	0.213***	0.306***	0.120	0.306***	0.157	0.194**
M_3	0.204**	0.278***	0.083	0.231***	0.194**	0.204**
M_4	0.296***	0.380***	0.185**	0.352***	0.269***	0.102
M_7	0.185**	0.260***	NA	NA	NA	NA
M_8	0.111	0.130	NA	NA	NA	NA
M_9	0.111	0.185**	NA	NA	NA	NA
M_{10}	0.389***	0.463***	NA	NA	NA	NA
M_{11}	0.185**	0.25***	0.093	0.176*	0.213***	0.259***
M_{12}	0.296***	0.370***	0.213***	0.287***	0.333***	0.185**

As a robustness check, test for the equality of the distribution of returns are performed for Western markets. The results are shown by using the Kolmogorov-Smirnov test. ***, ** and * correspond to two-tail significance levels of 1%, 5% and 10%, respectively.

dominance test, which means that results of Kolmogorov-Smirnov test generally agree with stochastic dominance test results.

3.7 Testing the trading rule

Following Bouman and Jacobsen (2002), we compare annual returns of the Eastern Halloween strategy and a Buy and Hold strategy. As for Eastern Halloween strategy, we assume that an investor who would like to make profit from the effect by buying a market portfolio at the end of June and sells this portfolio at the beginning of next May. This investor will then invest in a risk-free asset (short-term Treasury bonds) for 2 months (May and June). On the other hand, Buy and Hold strategy holds the stock market portfolio throughout the year. Table 3.11 contains the average annual returns and the standard deviation of the two strategies. The bold numbers are higher returns.

TABLE 3.11: Buy and Hold Versus Eastern Halloween

	Buy and Hold Strategy		*Eastern Halloween Strategy*	
	Mean	Standard Deviation	Mean	Standard Deviation
Eastern Markets				
China	**4.61**	26.53	2.26	26.93
Hong Kong	6.42	18.59	**8.95**	19.60
Japan	1.50	22.90	**2.18**	23.07
Korea	**12.49**	32.26	12.22	28.81
Singapore	4.92	22.90	**5.78**	21.70
Taiwan	1.69	17.47	**2.09**	17.41
Western Markets				
Canada	4.58	18.02	**4.67**	17.61
France	2.12	21.04	**4.66**	19.70
Germany	1.61	22.15	**4.37**	21.64
Italy	-1.87	21.99	**3.01**	19.80
UK	1.69	13.12	**3.53**	13.23
USA	4.70	16.08	**5.54**	15.56

Following Bouman and Jacobsen (2002), we compare annual returns of the Eastern Halloween strategy and a Buy and Hold strategy. As for Eastern Halloween strategy, we assume that an investor who would like to make profit from the effect by buying a market portfolio at the end of June and sells this portfolio at the beginning of next May. This investor will then invest in a risk-free asset (short-term Treasury bonds) for 2 months (May and June). On the other hand, Buy and Hold strategy holds the stock market portfolio throughout the year.

Interestingly, the Eastern Halloween strategy outperforms the Buy and Hold strategy in all the mature markets. This indicates an overwhelming evidence of such anomaly across most of the markets studied. However, there are exceptions such as China and Korea from the Eastern market bundle. This is indeed an interesting result as much as why such anomaly does exist in most of the Western markets considering such anomaly originated from a Hong Kong market adage with specific economic relevance to Hong Kong. The authors are of the opinion that the existence of such calendar effects could be influenced by the western (or international) financial practice noticeably on the accounting standard front and dividend distribution period over the March to June cycle. For the western markets, as long as it is consistent, corporations are more liberal to choose their fiscal years and most likely have a financial year ending on March

31st. For the Eastern markets, some Asian markets follow the fiscal year from April 1st to March 31st and announce their annual performance in March and then give out subsequent dividends in the next 2 months or so. May is the time when the stock prices decrease after the dividend payment. However, China and Korea follow a calendar year in which January 1st to December 31st is the fiscal year. While we also consider the migration of Asia culture as a potential influencer to this anomaly, it is beyond the scope of this paper to rationalize a sociocultural framework to investigate if such an anomaly claim can be substantiated.

3.8 Concluding remarks

This paper presents evidence focusing on one form of calendar anomalies that violate the efficient market hypothesis. Based on the well-known Halloween effects, we examine a similar effect (known as Eastern Halloween effect) that originates from Hong Kong but have not recognized as an established anomaly. Empirical research should pay attention to the stock return distributions, especially if monthly data are used. We show that Eastern Halloween effect exists in Hong Kong, Japan, Singapore, Taiwan, France, Germany, Italy and UK. However, risk evaluation approaches such volatility index and GARCH model cannot explain the Eastern Halloween effects. The same inconsistency between return and risk was also reported by Fama and French (2018) in a recent study. Therefore, this paper abjures this risk analysis approach and applies stochastic dominance criteria for testing the Eastern Halloween effect. The main feature of the stochastic dominance approach is that it makes minimal assumptions about the distribution of returns or investor risk preferences. Based on this method, formal tests confirm that May to June calendar period have been stochastically dominated by the rest of the calendar months during the year period. This implies that investors can increase their expected utility by buy the stochastically dominant portfolio and sell the stochastically dominated one. We also test such market inefficiency to see if this timing strategy based on Eastern Halloween effect can yield superior returns. We find the outperformance of this timing strategy over a Buy and Hold strategy in all covered markets except for China and Korea markets. In conclusion, this paper has provided evidence that the new defined Eastern Halloween effect is a market anomaly which can add to the calendar anomaly portfolio. This phenomenon is not just for the Eastern market studied but also for the Western markets as well.

3.9 Appendix 1

Table 3.12: Summary of Volatility Index

Volatility Index	Market	Description	Available Date
VHSI	Hong Kong	The index aims to measure the 30 calendar-day expected volatility of the Hang Seng Index implicit in the prices of near-term and next-term Hang Seng Index Options which are now trading on the Hong Kong Exchanges and Clearing Limited's derivatives market.	January 2001
VNKY	Japan	The index is calculated by using prices of Nikkei 225 futures and Nikkei 225 options on the Osaka Securities Exchange.	January 2001
VKOSPI	Korea	The index represents the volatility of the future (30-day maturity) KOSPI 200 Index as predicted by investors of the KOSPI 200 option market based on the KOSPI 200 option price	January 2003
VCAC	France	The index captures implied volatility embedded in prices of out of the money index call and put options available on Liffe. It is based on CAC 40 which consists of the largest 40 companies listed on Euronext Paris. It is calculated in a transparent manner, using only one external parameter, the risk-free interest rate.	Jan 2000
VDAX	Germany	The index indicates the fluctuations in DAX expected in the derivatives market for the following 45 days.	January 1992
VFTSE	UK	The index represents the imped volatility on the FTSE 100. It allows market participants to anticipate on the variations of the UKX in 30 calendar days. It is calculated using the prices of the UKX options covering the Out Of The Money strike prices for the near and next term maturities.	January 2000
VIX	US	The Chicago Board Options Exchange Volatility Index reflects a market estimate of future volatility, based on the weighted average of the implied volatilities for a wide range of strikes. 1st and 2nd month expirations are used until 8 days from expiration, then the 2nd and 3rd are used.	January 1990

3.10 Appendix 2

Table 3.13: Volatility Test

	GARCH(1,1)		Optimal GARCH(p,q) based on AIC			Volatility CUSUM test	
	AIC	BIC	Optimal order	AIC	BIC	t-stat	p-vale
Eastern Markets							
Overall							
China	2067.6515	2082.3172	(1,1)	2067.6515	2082.3172	4.5870	0.0000
Hong Kong	3635.7046	3652.6657	(2,1)	3634.9860	3656.1874	4.8053	0.0000
Japan	3516.2285	3533.4827	(1,1)	3516.2285	3533.4827	4.5886	0.0000
Korea	2513.3514	2528.7717	(1,1)	2513.3514	2528.7717	3.3325	0.0000
Singapore	2910.9867	2927.2604	(4,2)	2905.0739	2937.6213	2.4197	0.0000
Taiwan	2493.0565	2508.4768	(1,1)	2493.0565	2508.4768	5.6972	0.0000
Western Markets							
Overall							
Canada	3442.2730	3459.5272	(1,1)	3442.2730	3459.5272	1.6384	0.0093
France	1373.144	1386.645	(1,1)	1373.144	1386.645	2.6019	0.0000
Germany	1444.9924	1458.4936	(1,1)	1444.9924	1458.4936	2.3307	0.0000
Italy	1432.0981	1445.5992	(1,1)	1432.0981	1445.5992	4.0449	0.0000
UK	3468.1957	3485.4499	(2,3)	3467.7895	3497.9844	5.7641	0.0000
USA	3232.7883	3250.1285	(3,3)	3226.5264	3261.2069	2.0338	0.0005

For Volatility CUSUM test, see Ploberger, Werner, and Walter Kramer. "The Cusum Test with OLS Residuals." Econometrica 60, no. 2 (March 1992): 271–285.

3.11 Appendix 3

Table 3.14: Volatility Breaks Test

Break Test1					
	1st F-statistics	2nd F-statistics	3rd F-statistics	4th F-statistics	Break Dates
Eastern Markets					
China	25.2*	227.14*	42.87*	10.45	2000M11, 2008M05, 2010M05
Hong Kong	14.53*	179.46*	54.02*	37.79*	2000M11, 2003M10, 2008M01, 2012M12
Japan	10.76*	78.64*	42.67*	0.00	2003M01, 2008M06, 2011M12
Korea	36.06*	249.19*	24.77*	7.79	2002M01, 2008M10, 2012M08
Singapore	38.89*	126.12*	65.83*	7.57	2002M07, 2007M11,2010M09
Taiwan	98.69*	213.2*	95.41*	1.38	2003M10, 2008M01, 2012M09
Western Markets					
Canada	38.28*	41.19*	72.79*	0.74	2002M01, 2008M01, 2012M08
France	31.03*	159.93*	133.37*	0.00	2004M01,2008M06, 2012M12
Germany	36.25*	57.61*	112.22*	0.00	2004M02,2008M10, 2013M03
Italy	95.61*	171.25*	125.85*	9.39	2004M02,2008M10, 2014M01
UK	59.43*	105.63*	70.96*	4.37	2004M02, 2008M01, 2013M03
USA	59.82*	125.29*	62.26*	0.00	2003M08, 2008M01, 2012M08

3.12 References

Andrade, S. C., Chhaochharia, V., & Fuerst, M. E. (2013). "Sell in May and Go Away" Just Won't Go Away. Financial Analysts Journal, 69(4), 94-105.

Bouman, S., & Jacobsen, B. (2002). The Halloween indicator," Sell in May and go away": Another puzzle. The American Economic Review, 92(5), 1618-1635.

Bai, J., & Perron, P. (2003). Critical values for multiple structural change tests. The Econometrics Journal, 6(1), 72-78.

Bai, Z., H. Li, M. McAleer, W.K.Wong (2015), "Stochastic dominance statistics for risk averters and risk seekers: an analysis of stock preferences for USA and China." Quantitative Finance, 15(5): 889-900.

Chen, J., Chollete, L., & Ray, R. (2010). Financial distress and idiosyncratic volatility: An empirical investigation. Journal of Financial Markets, 13(2), 249-267.

Cho, Y. H., Linton, O., & Whang, Y. J. (2007). Are there Monday effects in stock returns: A stochastic dominance approach. Journal of Empirical Finance, 14(5), 736-755.

Chung, K. H., & Chuwonganant, C. (2018). Market volatility and stock returns: The role of liquidity providers. Journal of Financial Markets, 37, 17-34.

Clark, E., & Kassimatis, K. (2014). Exploiting stochastic dominance to generate abnormal stock returns. Journal of Financial Markets, 20, 20-38.

Dichtl, H., & Drobetz, W. (2015). Sell in May and Go Away: Still good advice for investors?. International Review of Financial Analysis, 38, 29-43.

Falk, H., & Levy, H. (1989). Market reaction to quarterly earnings' announcements: A stochastic dominance based test of market efficiency. Management Science, 35(4), 425-446.

Fama, E. F., & French, K. R. (2018). Volatility Lessons. Financial Analysts Journal, 74(3), 42-53.

Fong, W. M., Wong, W. K., & Lean, H. H. (2005). International momentum strategies: a stochastic dominance approach. Journal of Financial Markets, 8(1), 89-109.

Guo, B., Luo, X., & Zhang, Z. (2014). Sell in May and go away: Evidence from China. Finance Research Letters, 11(4), 362-368.

Hanoch, G., & Levy, H. (1969). The Efficiency Analysis of Choices Involving Risk. The Review of Economic Studies, 36(3), 335-346.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. Journal of Applied Econometrics, 20(7), 873-889.

Haug, M., & Hirschey, M. (2006). The January effect. Financial Analysts Journal, 62(5), 78-88.

Heston, S. L., Sadka, R. (2010). Seasonality in the Cross Section of Stock Returns: The International Evidence. Journal of Financial and Quantitative Analysis 45, 1133-1160.

Hong, Harrison, & Jialin Yu (2009), Gone fishin': Seasonality in trading activity and asset prices, Journal of Financial Market 12, 672–702.

Hadar, J., & Russell, W. R. (1969). Rules for ordering uncertain prospects. The American Economic Review, 59(1), 25-34.

Jacobsen, B., Visaltanachoti, N. (2009). The Halloween effect in US sectors. Financial Review, 44, 437-459.

Kamstra, M. J., Kramer, L. A., Levi, M. D. (2003). Winter blues: A SAD stock market cycle. The American Economic Review, 93, 324-343.

Keim, D. B. (1983). Size related firm characteristic factors and stock return seasonality: further empirical evidence. Journal of Financial Economics 12, 13–32.

Keloharju, M., Linnainmaa, J. T., Nyberg, P. (2016). Return seasonalities. Journal of Finance, 71, 1557-1590.

Lakonishok, J., Smidt, S. (1984). Volume and turn-of-the-year behavior. Journal of Financial Economics 13, 435-455.

Li, F., Zhang, H., & Zheng, D. (2018). Seasonality in the cross section of stock returns: Advanced markets versus emerging markets. Journal of Empirical Finance, 49, 263-281.

Linton, O., Maasoumi, E., & Whang, Y. J. (2005). Consistent testing for stochastic dominance under general sampling schemes. The Review of Economic Studies, 72(3), 735-765.

Lucey, B. M., Zhao, S. (2008). Halloween or January? Yet another puzzle. International Review of Financial Analysis 17, 1055-1069.

Maberly, E. D., & Pierce, R. M. (2004). Stock market efficiency withstands another challenge: Solving the "sell in May/buy after Halloween" puzzle. Econ Journal Watch, 1(1), 29-46.

Maberly, E. D., & Pierce, R. M. (2003). The Halloween effect and Japanese equity prices: myth or exploitable anomaly. Asia-Pacific Financial Markets, 10(4), 319-334.

Murtazashvili, I., & Vozlyublennaia, N. (2012). The role of data limitations, seasonality and frequency in asset pricing models. Journal of International Financial Markets, Institutions and Money, 22(3), 555-574.

Officier, P. R (1975). Seasonality in Australian Capital Markets. Journal of Financial Economics 2, 29-51.

Ploberger, W., & Krämer, W. (1992). The CUSUM test with OLS residuals. Econometrica: Journal of the Econometric Society, 271-285.

Reinganum, M. R. (1983). The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects. Journal of Financial Economics 12, 89-104.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. Journal of Economic theory, 2(3), 225-243.

Rozeff, M. S., Kinney, W. R. (1976). Capital market seasonality: The case of stock returns. Journal of Financial Economics 3, 379-402.

Wooldridge, Jeffrey, M. (2001). Applications of Generalized Method of Moments Estimation. Journal of Economic Perspectives, 15 (4): 87-100.

Zhang, C. Y., & Jacobsen, B. (2012). Are monthly seasonals real? A three century perspective. Review of Finance, 17(5), 1743-1785.

Chapter 4

Revisiting Portfolio Optimization: The Interplay of Partial Moments, Mean-Variance Analysis, and Prospect Theory

We consider the use of mean-variance (MV) analysis and portfolio optimization for investors with preferences in line with prospect theory. We explore the connection between MV analysis, portfolio optimization and prospect theory by utilizing the concept of prospect stochastic dominance (PSD) so that our results are applicable to a wide class of investors. This paper first shows that the partial moments as return-risk measures are consistent with PSD and can be used to eliminate investments which are PSD-inefficient. We then show that by incorporating partial moments, we can derive a simple rule similar to the MV rule to rank investments for different value functions under prospect theory. This rule can help us identify the exact segment on the MV frontier which is PSD-efficient so that one can find out the subset of MV portfolio choices which is both MV-efficient and PSD-efficient. Using partial moments, we also provide a portfolio optimization method for constructing a portfolio that dominates a target investment in the PSD sense.

4.1 Overview

Prospect theory, developed by Kahneman and Tversky (1979) and Tversky and Kahneman (1992), is a widely recognized theory in the field of behavioral finance. It has played a significant role in the study of investment decision-making, particularly in: the explanation of market anomalies [e.g. Barberis et al. (2001) and Barberis et al. (2021)]; the examination of whether financial decisions are based on prospect theory [e.g. Ljungqvist and Wilhelm Jr (2005) and Abdellaoui et al. (2013)]; and the examination of how prospect theory impacts existing finance tools [e.g. Levy and Levy (2004) and Zakamouline and Koekebakker (2009)]. These studies offer different insights compared to those derived from "classical" theories, such as expected utility theory, due to their incorporation of key features of human decision-making, including the consideration of changes relative to a status quo, risk aversion in the gain domain and risk-seeking in the loss domain, the preference for avoiding losses over acquiring equivalent gains, and the subjective transformation of objective probabilities.

While there have been numerous studies that have attempted to apply prospect theory to finance research, few have focused on developing tools to assist investors who conform to prospect theory in making more efficient investment decisions. In this chapter, we will focus on modifying and creating tools for such investors, as well as examining the impact of prospect theory on existing tools and its implications. We will particularly focus on the value function. To clarify in terms of prospect theory, we're working under the assumption of linear probability transformations. This will allow us to create a benchmark when we start to assume human do not use objective probability directly when they try to make decisions and it can also help us to isolate the effect from different elements of prospect theory if we want to look at the implications of it.

In this chapter, we demonstrate that partial moments, as return-risk measures, can be used as measurement tools that are consistent with prospect theory. To the best of our knowledge, the only previous work in this area has been conducted by Zakamouline and Koekebakker (2009), who used the value functions proposed by prospect theory to approximate the expected value function with a function of mean and partial moments, and then constructed a portfolio performance measure for piecewise-quadratic value functions. In contrast, our analysis of the connection between partial moments and prospect theory does not require the assumption of a piecewise-quadratic value function. Instead, we utilize the concept of "prospect stochastic dominance" (PSD) as proposed in Levy and Wiener (1998) and which has been further discussed in Baucells and Heukamp (2006), enabling our results to be applied to a wide class of investors who behave as if they are using different value functions under prospect theory. As it turns out, these results can also be used to eliminate investments which are PSD-inefficient.[1]

Furthermore, we demonstrate that by incorporating partial moments, we can connect the mean-variance (MV) rule with prospect theory and extend it to rank investments for different value functions under prospect theory. The MV rule, proposed by Markowitz (1952), itself is simple and elegant. If one risky investment has a higher expected return and lesser return variance than another, it should be preferred. This rule suggests that when assessing two risky investments for different individuals, the calculation of expected utility after eliciting each individual's utility function isn't necessarily required. Instead, a more straightforward method would be to directly compare the expected returns and return variance of the investments.[2]

The MV rule is particularly significant when it is linked to expected utility theory, which states that MV analysis is consistent with the framework in three situations: 1) the outcomes of the investments follow some specific distributions and the utility function is concave; 2) the utility function is quadratic; or 3) the utility function is concave and is well approximated by a function of mean and variance. However, these situations all require the utility function to be increasing and concave over all outcomes, an assumption that is frequently challenged in the real world. One classic example discussed by Friedman and Savage (1948) is the simultaneous purchase of actuarially unfair insurance and actuarially unfair lottery tickets, which seems to involve both risk aversion and risk-seeking behavior, conflicting with the assumption of risk aversion implied by a concave utility function.

The S-shaped value function was first proposed by Kahneman and Tversky (1979) in their development of prospect theory. This value function is convex in the loss domain and concave in the gain domain. The theory underscores a psychological response that manifests as a concave function relative to the magnitude of physical change. When this principle is applied to both gains and losses, this principle suggests that the perceived value difference between a gain (or loss) of 100 and 200 is more pronounced than that between 1,100 and 1,200. In economic

[1]The PSD rule is generalized in Baucells and Heukamp (2006) by including loss aversion and decision weighting.

[2]Similar results are also found in Chambers et al. (2014) when uncertainty is involved.

parlance, this means that the marginal "utility" of gains is decreasing and the marginal "disutility" of losses is also decreasing. This implies that people are risk-seeking in the loss domain and risk-averse in the gain domain.

The disposition effect observed in securities trading is commonly cited as evidence of this risk-seeking behavior in the loss domain, as highlighted by studies such as Weber and Camerer (1998) and Frazzini (2006). A trading strategy that mirrors this sentiment is the martingale approach. In this tactic, traders double their investment following a loss, anticipating that a subsequent win will not only recoup prior losses but also generate a profit. This aligns with the observation that individuals often take greater risks when trying to prevent imminent losses.

Levy and Levy (2004) were among the first to explore how the new paradigm may affect the MV rule, using the PSD rule. Counterintuitively, they showed that a portion of the MV-efficient set is also PSD-efficient, and this set is located in the segment between the minimum variance portfolio and the point of tangency from the origin to the frontier. In this chapter, our results demonstrate that incorporating partial moments in MV analysis can make it consistent with choices made using an S-shaped value function, and provide a way to rank investments, even if some have both a higher mean and higher variance. This means that modified MV analysis may still be applicable even when investors are risk-seeking on losses. Our results also help to identify the exact segment on the MV frontier that is PSD-efficient, allowing for the creation of a modified MV frontier that is both MV-efficient and PSD-efficient.

Notably, our study also unveils an intriguing outcome: an investment with both a higher mean and increased variance can be deemed more favourable than another for the majority of S-shaped value functions. Specifically, we discern that, even with the heightened volatility of the higher variance investment, downside risk stands out as the dominant factor influencing preferences between the two investments, especially in the context of most S-shaped value functions.

Using partial moments, we also present a method for constructing a portfolio that dominates a target investment in the PSD sense through portfolio optimization. There have been several papers that examine portfolio optimization under prospect theory, such as Berkelaar et al. (2004), which analyzed the optimal investment strategy for loss-averse investors, and He and Zhou (2011), who conducted an analytical treatment of a single-period portfolio choice model for agents with features suggested by prospect theory. While these works are valuable, they often assume a specific utility function to obtain their results. It is unclear whether a single functional form can accurately represent a large group of investors. Hence, a method for constructing a suitable portfolio for a class of investors is necessary.

The remainder of the chapter is organized as follows: in Section 4.2, we discuss the concept of PSD; in Section 4.3, we present the relationships between PSD and partial moments; in Section 4.4, we demonstrate how to incorporate partial moments in the MV analysis to make it consistent with choices made using an S-shaped value function; in Section 4.5, we provide a portfolio optimization method based on PSD; and in Section 4.6, we conclude.

4.2 Prospect stochastic dominance

Let X and Y be two random variables (or risky investments) with continuous *cumulative distribution functions* (CDFs), F and G and *Probability density function* (PDFs) f and g. Hereafter, we shall refer to the value function as the utility function. Subsequently, we define several classes of utility functions in the following.

Definition 4.1 (S-shaped utility function)**.** Given CDFs F and G, let U_2^A, U_2^D, U^S, and $U^S(F,G)$ be the sets of the utility functions u such that:

$$
\begin{aligned}
U_2^A &= \{u : (-1)^{i+1}u^{(i)} \geq 0 \ , \ i = 1, 2\}, \\
U_2^D &= \{u : u^{(i)} \geq 0 \ , \ i = 1, 2\}, \\
U^S &= \{u : (u^+ \in U_2^A) \quad \text{and} \quad (u^- \in U_2^D)\}, \\
U^S(F,G) &= \{u : (u \in U^S) \quad \text{and} \quad (\textstyle\int udF \ \& \ \int udG \text{ are finite})\},
\end{aligned}
$$

where $u^{(i)}$ is the i^{th} derivative of the utility function u, $u^+ = u$ restricted for $x > 0$, and $u^- = u$ restricted for $x < 0$.

Noted that investors with utility function in U_2^A are risk-averse while investors with utility function in U_2^D are risk-seeking. Investors with utility function in U^S are risk-averse on gains but risk-seeking on losses. Note that U^S contains all S-shaped utility functions that are continuous and piecewise smooth; thus, it also contains all such S-shaped utility functions that will imply loss aversion behaviour. With the assumption that the expected utility is well defined for the CDFs under consideration, the utility function proposed in Kahneman and Tversky (1979) belongs to set $U^S(F,G)$. Accordingly, for an agent with behaviour consistent with utility function u, F is preferred to G if

$$E_F u(x) - E_G u(x) = \int u(x)dF(x) - \int u(x)dG(x) \geq 0.$$

In order to compare risky investments for a wide class of investors who behave as if they are using expected utility with S-shaped utility function, we employ the concept of Prospect Stochastic Dominance (PSD) proposed by Levy and Wiener (1998). PSD allows us to compare the expected utility of different choices by only comparing the CDFs of the corresponding random variables. This allows us to make conclusions about the relative utility of different options without specifying a particular utility function. PSD is expressed in the following definition.

Definition 4.2 (Prospect Stochastic Dominance)**.** Given CDFs F and G with bounded or unbounded domain, F dominates G by PSD (denoted by $F \succeq_P G$) if $\int_{x_1}^{x_2} [G(z) - F(z)]\, dz \geq 0$ for all $x_1 \leq 0 \leq x_2$.[3]

One can view the PSD condition as applying a second order stochastic dominance condition to the positive domain and applying a risk-seeking stochastic dominance condition to the negative domain. The relationship between PSD and expected utility with S-shaped utility functions is studied in Levy and Wiener (1998), Baucells and Heukamp (2006), and Wong and Chan (2008). We formally extend this relationship for CDFs with an unbounded domain as follows.[4]

[3] An example is that for random variables X and Y that are normally distributed, if the mean of F is 0.08 with a standard deviation of 0.10, and the mean of G is 0.04 with a standard deviation of 0.07, then it follows that $F \succeq_P G$. For a more detailed discussion on this, see Levy and Levy (2004).

[4] In Theorem 4 of Levy and Wiener (1998), U_p is the set of all S-shaped utility functions as suggested by Kahneman and Tversky (1979). $u \in U_p$ is equivalent to $u'(x) \geq 0$ for all x, $u''(x) \leq 0$ for all $x > 0$, and $u''(x) \geq 0$ for all $x < 0$. Consider u_0 as follows:

$$u_0(x) = \begin{cases} x^{0.88} & \text{if } x \geq 0 \\ -2.25(-x)^{0.88} & \text{if } x < 0 \end{cases}$$

which is the one estimated by Tversky and Kahneman (1992) and $u_0 \in U_p$. Now take $F(x)$ and $G(x)$ as follows:

$$F(x) = G(x) = \begin{cases} 1 - \frac{1}{x^{0.38}} & \text{if } x \geq 1 \\ 0 & \text{if } x < 1 \end{cases}$$

Theorem 4.3. *Given CDFs F and G with bounded or unbounded domain, $F \succeq_P G$ if and only if $E_F u(x) \geq E_G u(x)$ for all $u \in U^S(F,G)$.*

Namely, $F \succeq_P G$ implies that F is preferred to G for any S-shaped utility function in $U^S(F,G)$ and the converse also holds.

4.3 Relationships between PSD and moments

4.3.1 Theoretical relationships

Relationships between PSD and moments are developed in this section. Some of the results will be useful for the analysis in other sections. They can also be used to detect the absence of a PSD relationship between distributions, and they can be used to reduce the number of pairwise comparisons if we want to know the PSD relationship between many distributions. It is useful in the sense that comparing distributions with the PSD condition involves comparing integrals of distributions for a large number of outcomes, see Linton et al. (2005) as an example, which will create computational burden. These conditions are also useful if we only have limited information on distributions; for example, we only have information about partial moments. This chapter will label existing results as propositions and new results as lemmas or theorems.

Let the means of random variables X and Y be denoted by μ_X and μ_Y, respectively. Similarly, the variances of X and Y are represented as σ_X^2 and σ_Y^2. From now on, it is assumed that both $E[X^2]$ and $E[Y^2]$ are finite. We begin by discussing an established result concerning PSD and the mean of returns:

Proposition 4.4. *For random variables X and Y, $F \succeq_P G$ implies $\mu_X \geq \mu_Y$.*

Proposition 4.4 is a direct corollary of Theorem 4.3 since the identity function is contained in $U_S(F,G)$. The proposition simply says that if a CDF dominates the other CDF in a PSD sense, then the mean of the first one will always be weakly higher than the mean of the other. Note

Then $\int_{x_1}^{x_2}[G(z) - F(z)] = 0$ for all $x_1 \leq 0 \leq x_2$, that is, $F \succeq_P G$. However, the difference between $E_F u(x)$ and $E_G u(x)$ in this case is not well defined since both of these are equal to $\int_1^\infty \frac{0.38}{\sqrt{x}} dx$ which is infinite.

that it is also the case for first order stochastic dominance (FSD) and second order stochastic dominance (SSD). [5]

To delve deeper into the implications of the PSD relationship, consider that the mean of a random variable can be divided into its two first partial moments. For $Z = X$ or Y, and $H = F$ or G, we define the following:

$$\mu_Z = \mu_Z^+ - \mu_Z^- \quad \text{such that} \quad \mu_Z^+ = \int_0^\infty x dH(x) \quad \text{and} \quad \mu_Z^- = \int_{-\infty}^0 -x dH(x), \tag{A.1}$$

where μ_Z^+ is the upper first partial moment (or measure of expected gain) and μ_Z^- is the lower first partial moment (or measure of expected loss).

Some examples of using partial moments in finance include Holthausen (1981), Anthonisz (2012), and Liu and O'Neill (2018). It turns out that we can derive the relationship between PSD and the first partial moment for the general distributions. Before we derive the general relationship between PSD and the first partial moment, we establish the following technical results:

Lemma 4.5. *For random variables X and Y, the following statements are true:*

1. $\mu_X^+ - \mu_Y^+ = \int_0^\infty [G(x) - F(x)]dx.$
2. $\mu_Y^- - \mu_X^- = \int_{-\infty}^0 [G(x) - F(x)]dx.$

The first part of lemma 4.5 states that the difference between the first upper partial moments associated with the CDFs F and G, $\mu_X^+ - \mu_Y^+$, is equal to the integral of the difference between the CDFs, $G(x) - F(x)$, over the interval $[0, \infty)$. The second part states that the difference between the first lower partial moments associated with the CDFs F and G, $\mu_Y^- - \mu_X^-$, is equal to the integral of the difference between the CDFs, $G(x) - F(x)$, over the interval $(-\infty, 0]$. Using this Lemma, necessary conditions for PSD that involve first partial moments are given in Theorem 4.6.

Theorem 4.6. *For random variables X and Y, the following statements are true:*

1. *If $F \succeq_P G$, then $\mu_X^+ \geq \mu_Y^+$.*

[5] For discussion between moments and n^{th}-order stochastic dominance see Chan et al. (2022).

2. *If $F \succeq_P G$, then $\mu_X^- \leq \mu_Y^-$.*

Theorem 4.6 states that PSD will not only affect the mean, but it will also affect both the first upper partial moment and the first lower partial moment. Theorem 4.6 also implies that if we compare two distributions, say F and G, if a distribution F has either a smaller first upper partial moment or a larger first lower partial moment, we will never get $F \succeq_P G$ (denote as $F \not\succeq_P G$). We illustrate this in the following example:

Example 4.1. *Consider a pair of normal distributions, F and G, with $\mu_X > \mu_Y$, $\sigma_X > \sigma_Y$ and $\mu_X\sigma_Y = \mu_Y\sigma_X$. The normal distributions will cross at most once and, in this case, G will cross F from below at $x_0 = 0$.*[6] *Since $\int_{-\infty}^{0}[F(z) - G(z)]\,dz > 0$, we know that $\mu_X^- > \mu_Y^-$ by Lemma 4.5. In addition, since $\mu_X > \mu_Y$, we have $\mu_X^+ > \mu_Y^+$. Now, we consider the PSD relationship. In order to obtain $F \succeq_P G$ or $G \succeq_P F$, we must have $\int_{x_1}^{x_2}[G(z) - F(z)]\,dz \geq 0$, or $\int_{x_1}^{x_2}[F(z) - G(z)]\,dz \geq 0$ for all $x_1 \leq 0 \leq x_2$, but F and G will cross at most one time and G will cross F from below at $x_0 = 0$. In this case, we have $\int_{-\infty}^{0}[F(z) - G(z)]\,dz > 0$ and $\int_{0}^{\infty}[G(z) - F(z)]\,dz > 0$ which means that neither $F \succeq_P G$ nor $G \succeq_P F$ will happen.*

The results could also help develop risk-return analysis tools if we require $F \succeq_P G$, which means that the distribution is preferred over others for all investors with S-shaped utility functions, see the following example for more information.

Example 4.2 (Omega ratio)**.** *The omega ratio was first introduced by Keating and Shadwick (2002). For $Z = X$ or Y with $H = F$ or G, according to proof of Proposition 1 in Guastaroba et al. (2016), using zero as the return threshold yields the following:*

$$\Omega_Z(0) = 1 + \frac{\mu_Z}{\int_{-\infty}^{0} H(x)dx} \; .$$

We then consider the difference of two omega ratios for two distributions as shown in the following:

$$\Omega_X(0) - \Omega_Y(0) = \frac{\mu_X}{\int_{-\infty}^{0} F(x)dx} - \frac{\mu_Y}{\int_{-\infty}^{0} G(x)dx} \; .$$

Since μ_Z^- is finite, we know that $\int_{-\infty}^{0}[G(z) - F(z)]dz$ is finite from Lemma 4.5, and therefore, we have $\int_{-\infty}^{0}[G(z) - F(z)]dz = \int_{-\infty}^{0} G(x)dx - \int_{-\infty}^{0} F(x)dx$. Since $F \succeq_P G$ implies $\mu_X \geq \mu_Y$

[6]See Properties 2 and 3 of Levy (2011) on p.87-88, F and G will cross exactly one time and G will cross F from below if $\sigma_X^2 > \sigma_Y^2$ and the normal distributions will cross at $x_0 = \frac{\mu_X\sigma_Y - \mu_Y\sigma_X}{\sigma_Y - \sigma_X}$.

and $\int_{x_1}^{x_2}[G(z) - F(z)]dz \geq 0$ *for all pairs of* x_1 *and* x_2 *such that* $x_1 \leq 0 \leq x_2$*. In this sense,* $F \succeq_P G$ *implies* $\Omega_X(0) \geq \Omega_Y(0)$ *for* $\mu_X \geq 0$*.*

For the risk analysis tool, the second moments of distribution with the corresponding random variable could also be partitioned into two second partial moments (or semi-variances from zero). For $Z = X$ or Y, and $H = F$ or G, we define the following:

$$\tau_Z = \tau_Z^+ + \tau_Z^- \quad \text{such that} \quad \tau_Z^+ = \int_0^\infty x^2 dH(x) \quad \text{and} \quad \tau_Z^- = \int_{-\infty}^0 x^2 dH(x), \tag{A.2}$$

where τ_Z is the the second moment, τ_Z^+ is the second upper partial moment and τ_Z^- is the second lower partial moment. The upper second partial moment and second lower partial moment are special cases of semivariance. The semivariance as a measurement tool is first proposed in Markowitz (1959). The relationship between it and concave utility functions (or second order stochastic dominance) is further investigated by Ogryczak and Ruszczyński (1999). Before we go to the relationship between PSD and second partial moments, we establish the following technical results.

Lemma 4.7. *For random variables* X *and* Y*, the following statements are true:*

1. $\tau_Y^+ - \tau_X^+ = 2\int_0^\infty \left(\int_0^x [G(t) - F(t)]dt + \mu_Y^+ - \mu_X^+\right)dx$.
2. $\tau_Y^- - \tau_X^- = 2\int_{-\infty}^0 \left(\mu_Y^- - \mu_X^- - \int_x^0 [G(t) - F(t)]dt\right)dx$.

The first part of the Lemma states that the difference between the second upper partial moments associated with the CDFs F and G, $\tau_Y^+ - \tau_X^+$, is equal to twice the integral of the sum of the inner integral, $\int_0^x [G(t) - F(t)]dt$, and the difference between the first upper partial moments, $\mu_Y^+ - \mu_X^+$, over the interval $[0, \infty)$. The second part states that the difference between the second lower partial moments associated with the CDFs F and G, $\tau_Y^- - \tau_X^-$, is equal to twice the integral of the difference between the first lower partial moments, $\mu_Y^- - \mu_X^-$, and the inner integral, $\int_x^0 [G(t) - F(t)]dt$, over the interval $(-\infty, 0]$. Using this Lemma, we derive the necessary conditions for the PSD that involve the first and second partial moments and display the results in the following theorem.

Theorem 4.8. *For random variables* X *and* Y*, the following statements are true:*

1. *Given that* $\mu_X^+ = \mu_Y^+$, *if* $F \succeq_P G$, *then* $\tau_X^+ \leq \tau_Y^+$.

2. *Given that* $\mu_X^- = \mu_Y^-$, *if* $F \succeq_P G$, *then* $\tau_X^- \geq \tau_Y^-$.

The first part of Theorem 4.8 states that if F dominates G in the sense of the PSD and the first upper partial moment of two distributions are the same, then it must be the case that the upper second partial moment of F lies below the upper second partial moment of G. On the other hand, the second part of Theorem 4.8 states that if F dominates G in the sense of the PSD and the first lower partial moment of two distributions are the same, then it must be the case that the lower second partial moment of F is on top of the lower second partial moment of G.

Instead of using zero as the target return, we can also use any amount x as the target return to define the first lower partial moment $\mathcal{L}_Z(x)$ and the first upper partial moment $\mathcal{U}_Z(x)$, for $H = F$ or G, $Z = X$ or Y, as shown in the following:

$$\mathcal{L}_Z(x) := E[(x - Z)_+] = \int_{-\infty}^{x} (x - z)dH(z) \ ,$$

$$\mathcal{U}_Z(x) := E[(Z - x)_+] = \int_{x}^{\infty} (z - x)dH(z) \ ,$$

where $(x)_+ = \max\{x, 0\}$. Noted that $\mathcal{L}_Z(x)$ could be considered as expected loss below the target return x and $\mathcal{U}_Z(x)$ could be considered as expected gain above the target return x. The relationship between these partial moments and cumulative distribution has been well-studied. We modify the results in Bawa et al. (1985) and Proposition 2 in Ogryczak and Ruszczyński (1999) to obtain the following.

Proposition 4.9. *If* $E\{|Z|\} < \infty$, *then the following statements are true:*

- $\mathcal{L}_Z(x) = \int_{-\infty}^{x} H(z)dz$.

- $\mathcal{U}_Z(x) = \int_{-\infty}^{x} H(z)dz - (x - \mu_H)$.

The first part states that the first lower partial moment, $\mathcal{L}_Z(x)$, is equal to the integral of the CDF $H(z)$ over the interval $(-\infty, x]$. This means that the first lower partial moment captures the probability mass of the random variable Z up to the value x. The second part states that the first upper partial moment, $\mathcal{U}_Z(x)$, is equal to the difference between the integral of the

CDF $H(z)$ over the interval $(-\infty, x]$ and the term $(x - \mu_H)$, where μ_H is the mean value of the random variable Z. This means that the first upper partial moment captures the remaining probability mass of the random variable Z beyond the value x, after accounting for the mean value of the distribution. Using the above results, we can get the following technical result.

Lemma 4.10. *For any pair of random variables X and Y with CDFs F and G, the following statements are true:*

1. $\mathcal{L}_Y(x) - \mathcal{L}_X(x) = \int_{-\infty}^{x} [G(z) - F(z)]dz.$
2. $\mathcal{U}_X(x) - \mathcal{U}_Y(x) = \int_{x}^{\infty} [G(z) - F(z)]dz.$
3. $\big[\mathcal{L}_Y(0) - \mathcal{L}_X(0)\big] - \big[\mathcal{L}_Y(x) - \mathcal{L}_X(x)\big] = \int_{x}^{0} [G(z) - F(z)]dz, \quad \forall x \leq 0.$
4. $\big[\mathcal{U}_X(0) - \mathcal{U}_Y(0)\big] - \big[\mathcal{U}_X(x) - \mathcal{U}_Y(x)\big] = \int_{0}^{x} [G(z) - F(z)]dz, \quad \forall x \geq 0.$

The first part of Lemma 4.10 states that the difference between the first lower partial moments, $\mathcal{L}_Y(x) - \mathcal{L}_X(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $(-\infty, x]$. The second part states that the difference between the first upper partial moments, $\mathcal{U}_X(x) - \mathcal{U}_Y(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[x, \infty)$. The third part states that the difference between the first lower partial moments evaluated at 0, $\mathcal{L}_Y(0) - \mathcal{L}_X(0)$, minus the difference between the first lower partial moments evaluated at x, $\mathcal{L}_Y(x) - \mathcal{L}_X(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[x, 0]$ for all $x \leq 0$. The fourth part states that the difference between the first upper partial moments evaluated at 0, $\mathcal{U}_X(0) - \mathcal{U}_Y(0)$, minus the difference between the first upper partial moments evaluated at x, $\mathcal{U}_X(x) - \mathcal{U}_Y(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[0, x]$ for all $x \geq 0$. Combining Definition 4.2 and Lemma 4.10, we obtain the following corollary.

Corollary 4.11 (Safety and aggressiveness at the same time)**.** *$F \succeq_P G$ if and only if*

$$\min\Big\{ \big[\mathcal{L}_Y(0) - \mathcal{L}_Y(x_1)\big] - \big[\mathcal{L}_X(0) - \mathcal{L}_X(x_1)\big], \big[\mathcal{U}_X(0) - \mathcal{U}_X(x_2)\big] - \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(x_2)\big] \Big\} \geq 0$$
$$\forall \;\; x_1 \leq 0 \leq x_2 \; . \tag{A.3}$$

The corollary states that $F \succeq_P G$ if and only if the minimum of two differences is greater than or equal to 0 for all $x_1 \le 0 \le x_2$. The first difference is the difference between the first lower partial moments evaluated at 0 and x_1, i.e., $\big[\mathcal{L}_Y(0) - \mathcal{L}_Y(x_1)\big] - \big[\mathcal{L}_X(0) - \mathcal{L}_X(x_1)\big]$. The second difference is the difference between the first upper partial moments evaluated at 0 and x_2 for random variable X and Y, i.e., $\big[\mathcal{U}_X(0) - \mathcal{U}_X(x_2)\big] - \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(x_2)\big]$. Corollary 4.11 gives us a representation equivalent to $F \succeq_P G$. This result will be useful when we construct our portfolio optimization method in Section 4.5.

4.3.2 A simple method to reduce the number of comparisons

Comparing distributions with stochastic dominance conditions normally involves comparing integrals of distributions for a huge amount of outcomes which is the same for comparing PSD relationships, see Linton et al. (2005) as an example. This is not a quick task in general. For example, assume that you have 25 investments to consider. Then, fully understanding the relationship between these investments requires 25×24 times pairwise comparisons with each comparison involving the comparisons of integrals of distributions for a huge amount of outcomes. In this section, we make use of the results in Section 4.3.1 and illustrate that the number of comparisons can be reduced significantly by using a simple method. To do so, we first consider the following immediate consequence of both Theorems 4.6 and 4.8.

Corollary 4.12. *For random variables X and Y, the following statements are true:*

1. *If $\mu_X^+ < \mu_Y^+$, then $F \not\succeq_P G$.*
2. *If $\mu_X^- > \mu_Y^-$, then $F \not\succeq_P G$.*
3. *Given that $\mu_X^+ = \mu_Y^+$, if $\tau_X^+ > \tau_Y^+$, then $F \not\succeq_P G$.*
4. *Given that $\mu_X^- = \mu_Y^-$, if $\tau_X^- < \tau_Y^-$, then $F \not\succeq_P G$.*

We can utilize the aforementioned results to formulate the following hypotheses: $H_0^+ : \mu_X^+ \ge \mu_Y^+$, $H_0^- : \mu_X^- \le \mu_Y^-$, $H_0^{2+} : \tau_X^+ \le \tau_G^+$, and $H_0^{2-} : \tau_X^- \ge \tau_G^-$. Rejecting H_0^+ or H_0^- is equivalent to reject $F \succeq_P G$. When we compare the two distributions, if both H_0^+ and H_0^- are not rejected, then we can further check whether H_0^{2+} or H_0^{2-} is rejected, in this situation, again, it is equivalent to check whether $F \succeq_P G$ is rejected. By using this procedure, we can reduce the number of comparisons before we go to compare the integral of distributions for a huge amount of outcomes. To demonstrate this, we present an empirical example utilizing real-life data in Kenneth French's online data library. The data consists of the excess returns of 25 portfolios formed based on size and Book-to-Market ratio (BE/ME). Portfolios are constructed annually at the close of each June by combining two distinct criteria: size, represented by market equity (ME), and the ratio of book equity to market equity (BE/ME). The size-based classification hinges on the New York Stock Exchange (NYSE) market equity quintiles as of the end of June in year t, yielding five distinct portfolios sorted by size. For the BE/ME-based classification, the ratio for June of year t is derived from the book equity of the last fiscal year, which concludes in $t-1$, divided by the market equity (ME) from December of $t-1$. This methodology, grounded on NYSE quintiles, similarly results in five portfolios. The final portfolios are essentially the intersections of these size and BE/ME-based portfolios. [7] The portfolios we used in our analysis contain almost all common stocks listed on NYSE, AMEX, and NASDAQ.[8]

We denote the portfolios by using their names in Kenneth French's online data library. In addition, we use the one-month US Treasury bill as a risk-free rate to calculate the excess returns. This paper mainly focuses on the monthly excess return from July 1963 to December 2021. All data is directly obtained from Kenneth French's online data library. To test H_0^+, H_0^-, H_0^{2+}, and H_0^{2-}, we use the bootstrap method in Efron and Tibshirani (1994) which will be further elaborated in Appendix 4.8. We perform these tests to the empirical distribution for each of the 600 pairs of two portfolios and exhibit the results in Table 4.1.

[7] These portfolios have been analysed by Post and Levy (2005).

[8] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html for more details about the portfolios.

TABLE 4.1: Redundant comparisons suggested by the corresponding hypotheses

Size	BE/ME	H_0^+	H_0^-	H_0^{2+}	H_0^{2-}	Total
1	1	0	24	21	0	24
1	2	1	21	15	2	23
1	3	5	16	11	3	20
1	4	6	8	9	5	16
1	5	1	10	9	2	13
2	1	0	23	18	1	24
2	2	3	16	11	3	19
2	3	10	8	6	7	18
2	4	10	4	5	9	17
2	5	3	15	12	1	18
3	1	3	21	13	2	24
3	2	10	10	5	5	20
3	3	15	2	2	11	17
3	4	15	1	3	11	18
3	5	5	8	9	2	14
4	1	7	15	4	4	21
4	2	15	7	2	7	22
4	3	17	2	2	10	19
4	4	15	1	3	9	18
4	5	6	10	8	2	16
5	1	21	1	0	15	22
5	2	21	0	0	18	21
5	3	22	0	0	20	22
5	4	21	1	0	5	22
5	5	11	9	4	2	20

Note: We use 5% as the significance level in this table.

Table 4.1 provides a comprehensive overview of the results. The first column shows the sizes of the companies we considered in the portfolio; the second column shows the BE/ME of the companies; the numbers in columns 3, 4, 5, and 6 correspond to the number of times we rejected H_0^+, H_0^-, H_0^{2+}, and H_0^{2-}. These rejections occur when comparing the portfolio indicated in Columns 1 and 2 (denoted as F) with one of the other 24 portfolios (denoted as G). The last column displays the total number of rejections for $F \succeq_P G$. This encompasses cases rejecting either H_0^+ or H_0^-. Only when both H_0^+ and H_0^- are not dismissed do we proceed to test for H_2^+ or H_2^- rejections.

The outlined approach significantly reduces the number of comparisons. On average, it reduces 19.5 out of the 24 comparisons, resulting in a requirement to conduct only $(24 - 19.5)/24 = 18.75\%$ of the original comparisons. This reduction in the number of comparisons enhances efficiency and streamlines the analysis process.[9]

4.4 PSD and Mean-Variance Analysis

This section is devoted to the relationship between PSD and MV Analysis under the premise of elliptically distributed asset returns. We will first define the elliptical distribution, followed by a discussion of results related to our analysis. Subsequently, we will delve into the connection between the MV rule and PSD. In the final part of this section, we will extend the findings of Levy and Levy (2004), exploring the link between MV-efficient and PSD-efficient portfolios.

4.4.1 Elliptical Distribution

Elliptical distributions, a class of probability distributions that includes the multivariate normal distribution and the multivariate t-distribution, are critical in finance due to their adaptability and capability to handle heavy-tailed data, which are common in financial returns. Financial returns often exhibit 'fat tails' or heavy tails, which means that extreme events (like significant stock market crashes or unprecedented price jumps) occur more frequently than what would be predicted by the standard Gaussian distribution. By using distributions that can cater to these heavy tails, we can model and predict these extreme events with higher accuracy. The contours of these distributions, which form ellipses (in the bivariate case), allow for a broad range of shapes beyond the bell curve, accommodating different types of data. For application examples, consider works such as Xiao and Valdez (2015), and Landsman et al. (2018). We define an elliptically distributed random vector as follows.

Definition 4.13. Let $\boldsymbol{Z}$ be any N-dimensional random vector. $\boldsymbol{Z}$ is said to be "elliptically distributed" with mean vector $\boldsymbol{\mu_Z}$, dispersion matrix $\boldsymbol{\Sigma_Z}$, characteristic generator ϕ, denoted

[9] There are other potential applications for this approach. One such example could be employing a PSD concept to analyze market efficiency like Cho et al. (2007) and Chui et al. (2020) utilized an SSD concept in their respective market efficiency analyses.

as $\boldsymbol{Z} \sim \mathcal{E}_N(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$, if and only if

$$\boldsymbol{Z} \overset{d}{=} \boldsymbol{\mu_Z} + \boldsymbol{\Lambda}\mathcal{R}\boldsymbol{U^{(k)}}$$

where $\boldsymbol{U^{(k)}}$ is a k-dimensional random vector uniformly distributed on $S^{k-1} := \{\boldsymbol{z} \in \mathbb{R}^k : ||\boldsymbol{z}||_2 = 1\}$ and $||\cdot||_2$ is the Euclidean norm. $\mathcal{R}$ is a nonnegative random variable that is stochastically independent of $\boldsymbol{U^{(k)}}$, $\boldsymbol{\mu_Z} \in \mathbb{R}^N$, $\boldsymbol{\Lambda} \in \mathbb{R}^{N\times k}$ with $\text{rank}(\boldsymbol{\Lambda}) = k$ such that $\boldsymbol{\Sigma_Z} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}'$ and characteristic generator $\phi : \mathbb{R}^+ \to \mathbb{R}$ is a function such that the characteristic function $\boldsymbol{t} \mapsto \varphi_{\boldsymbol{Z}-\boldsymbol{\mu_Z}}(\boldsymbol{t})$ of $\boldsymbol{Z} - \boldsymbol{\mu_Z}$ corresponds to $\boldsymbol{t} \mapsto \phi(\boldsymbol{t}'\boldsymbol{\Sigma_Z}\boldsymbol{t})$, $\boldsymbol{t} \in \mathbb{R}^N$.

Before delving further into our analysis, it's important to acknowledge the relevant properties of elliptical distributions. These properties, which we will utilize in our study, were previously discussed in detail by Frahm (2004) and Gupta et al. (2013).[10]

Proposition 4.14. *Let* $\boldsymbol{Z}$ *be an* N*-dimensional random vector such that* $\boldsymbol{Z} \sim \mathcal{E}_N(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$ *with* $\text{rank}(\boldsymbol{\Sigma_Z}) = k$*, then the following statements are true:*

1. $E[\boldsymbol{Z}] = \boldsymbol{\mu_Z}$.
2. $Cov(\boldsymbol{Z}) = \frac{E(\mathcal{R}^2)}{k}\boldsymbol{\Sigma_Z}$.
3. $\boldsymbol{\Lambda}^{-1}(\boldsymbol{Z} - \boldsymbol{\mu_Z}) \sim \mathcal{E}_k(\boldsymbol{0}, \boldsymbol{I_k}, \phi)$.

where $\boldsymbol{\Lambda}^{-1}$ *is the generalized inverse of* $\boldsymbol{\Lambda}$*, and* $\boldsymbol{I_k}$ *is the* k*-dimensional identity matrix.*

The random vector $\boldsymbol{Z}$ possesses a scale-location-parameter distribution, with its mean vector $\boldsymbol{\mu_Z}$, and its covariance matrix $Cov(\boldsymbol{Z})$ equal to the second moment of random variable $\mathcal{R}$ divided by $\text{rank}(\boldsymbol{\Sigma_Z})$ times the dispersion matrix $\boldsymbol{\Sigma_Z}$. By applying a transformation using the generalized inverse (sometimes referred to as the Moore–Penrose inverse) of matrix $\boldsymbol{\Lambda}$ (denoted as $\boldsymbol{\Lambda}^{-1}$) to the vector equal to $\boldsymbol{Z}$ subtracting the mean vector $\boldsymbol{\mu_Z}$, the random vector follows an elliptically symmetric distribution $\mathcal{E}_k(\boldsymbol{0}, \boldsymbol{I}_k, \phi)$, with a zero mean vector and k-dimensional identity matrix as the dispersion matrix, standardizing $\boldsymbol{Z}$.[11]

[10] See p.11-12 of Frahm (2004) and p.17 of Gupta et al. (2013).
[11] Another related discussion is in Owen and Rabinovitch (1983).

In this section, we consider random vectors such that their probability density function exists and their support is $\mathbb{R}^N$. Examples of distributions that satisfy these conditions include multivariate normal distribution, multivariate t-distribution, multivariate Cauchy, multivariate logistic distributions, etc. Suppose we have a random vector $\boldsymbol{Z}$ such that $\boldsymbol{Z} \sim \mathcal{E}_N(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$, with probability density function of $\boldsymbol{Z}$ exists and support of $\boldsymbol{Z}$ is $\mathbb{R}^N$, we will denote it as $\boldsymbol{Z} \sim \mathcal{E}_N^*(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$. In this section, we will first discuss random variables X and Y such that $X \sim \mathcal{E}_1^*(\mu_X, \Sigma_X, \phi)$ and $Y \sim \mathcal{E}_1^*(\mu_Y, \Sigma_Y, \phi)$ with corresponding CDFs F and G. In the one-dimensional case, matrix $\boldsymbol{\Lambda}$ becomes the square root of dispersion parameter, that is $\sqrt{\Sigma_Z}$, and the standardized random variables will become $(Z - \mu_Z)/\sqrt{\Sigma_Z}$ for $Z = X$ and Y.

Some of our results will be related to the intersection of the cumulative distribution functions of X and Y. We derive a related Lemma as follows.

Lemma 4.15 (Single-crossing Lemma)**.** *Let $X \sim \mathcal{E}_1^*(\mu_X, \Sigma_X, \phi)$ and $Y \sim \mathcal{E}_1^*(\mu_Y, \Sigma_Y, \phi)$, then the following statements are true:*

1. *F and G will cross if and only if $\sigma_Y^2 \neq \sigma_X^2$.*
2. *F and G will cross exactly one time and G will cross F from above if $\sigma_Y^2 > \sigma_X^2$.*

For CDFs that cross only once, this characteristic is often described as the "single-crossing property" in various works. Manski (1988) posits that if function G intersects function F from above, under specific conditions, it can be inferred that G embodies greater risk than F. In addition, de Castro and Galvao (2019) demonstrate an application of the single-crossing property in a dynamic context.

4.4.2 Modified Mean-Variance Analysis

Consider normally distributed returns, one of the common assumptions people will use when applying the MV analysis. One unexpected finding regarding PSD on comparing assets is the lack of any dominance between two normal distributions, say F and G, when both have the same expected return $\mu_X = \mu_Y$. This contrasts with the traditional theory of stochastic dominance for risk-averse and risk-seeking individuals, where dominance between F and G can be established despite equal expected returns.

Recall from Section 2.3 that $\mu_X - \mu_Y = \int_{-\infty}^{\infty}[G(x) - F(x)]dx$. If we assume $\mu_X = \mu_Y$, then it must be the case that the area between F and G when $F(x) > G(x)$ is exactly the same as the area between F and G when $F(x) < G(x)$; that is, $\int_{-\infty}^{\infty}[G(x) - F(x)]dx = 0$. In such scenarios, no matter where F and G cross, there is no way we can get $\int_{x_1}^{x_2}[G(z) - F(z)]\, dz \geq 0$ or $\int_{x_1}^{x_2}[G(z) - F(z)]\, dz \leq 0$ for all $x_1 \leq 0 \leq x_2$. This idea is illustrated in Figure 4.1. The $+$ and $-$ signs correspond to the area $\int_{-\infty}^{x_0}[G(t) - F(t)]dt$ and $\int_{x_0}^{\infty}[G(t) - F(t)]dt$. Since $\int_{-\infty}^{0}[G(t) - F(t)]dt \geq 0$, we cannot get $G \succeq_P F$. However, when we try to integrate the difference between F and G from zero to some positive value, we know that the $+$ area after the zero line is less than the $-$ area; therefore, $\int_0^x[G(t) - F(t)]dt \leq 0$ for some x, and thus, it is impossible to get $F \succeq_P G$.This result, as it turns out, could be extended for elliptical distributions, and we state this formally as the following theorem.

Theorem 4.16 (PSD Impossibility Theorem)**.** *Let* $X \sim \mathcal{E}_1^*(\mu, \Sigma_X, \phi)$ *and* $Y \sim \mathcal{E}_1^*(\mu, \Sigma_Y, \phi)$. *Then, we have* $F \not\succeq_P G$ *and* $G \not\succeq_P F$, *unless* $F = G$ *almost everywhere.*

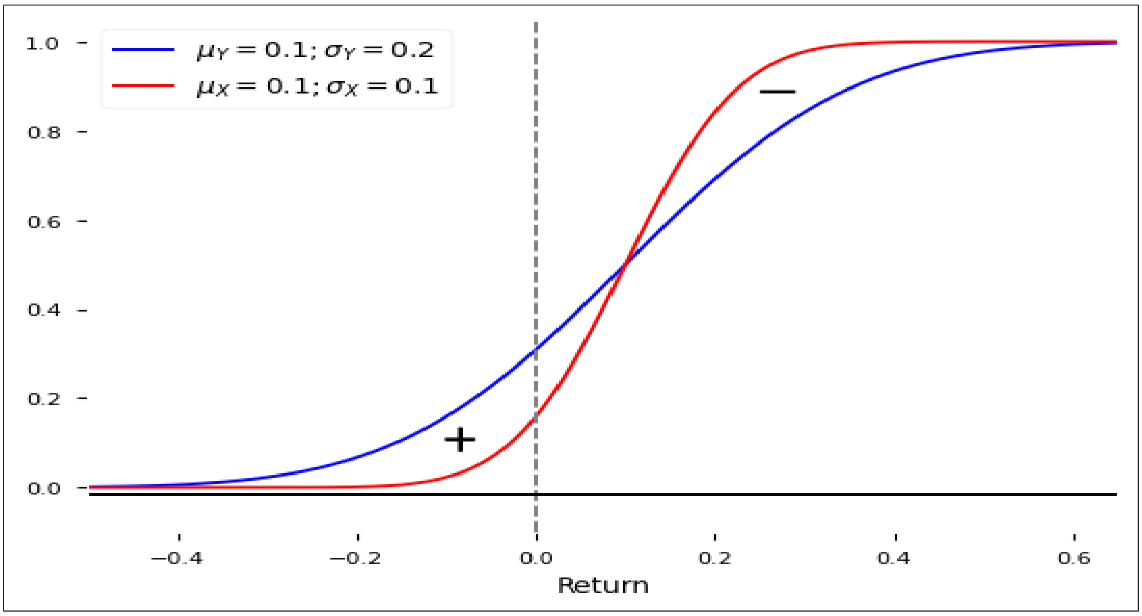


Figure 4.1: Example of PSD Impossibility Theorem

Theorem 4.16 establishes that given two distributions with equal means, there exists an S-shaped utility function that prefers the first distribution over the second, and another S-shaped utility function that favors the second distribution over the first. Consequently, it's indeterminable to universally prefer one distribution over the other across all S-shaped utility functions. This theorem provides additional justification for the potential need to adjust MV analysis when catering to investors whose preferences align with prospect theory. Recall that the MV rule states that if one risky investment has a higher expected return and lesser return variance than another, it should be preferred given that it either has a strictly higher expected return or a strictly lower return variance. However, the above theorem tells us that a strictly higher expected return is necessary for investors with an S-shaped utility function to prefer it over other investments.

In order to develop a modified MV analysis for investors with an S-shaped utility function, it is crucial to first understand the cross-over point of CDFs. It turns out that the cross-over point can be characterized by the first partial moments and variance associated with the respective distributions. We establish the following lemma that describes the cross-over point of two CDFs.

Lemma 4.17. *Let* $X \sim \mathcal{E}_1^*(\mu_X, \Sigma_X, \phi)$ *and* $Y \sim \mathcal{E}_1^*(\mu_Y, \Sigma_Y, \phi)$ *where* $\sigma_X^2 < \sigma_Y^2$*. Then the following statements are true:*

1. *F will cross G exactly for one time from below at x_0 such that $x_0 > 0$ if $\mu_X^+ \geq \mu_Y^+$.*

2. *F will cross G exactly for one time from below at x_0 such that $x_0 < 0$ if $\mu_X^- \geq \mu_Y^-$.*

The above lemma provides an interesting insight into the behavior of two cumulative distribution functions. We have shown that the cross-over point depends on the first partial moments and variance of the two distributions. This result is particularly important when we discuss the concept of PSD. This lemma plays a vital role in understanding and simplifying the conditions for PSD. It also allows us to restate or reframe previous theorems and lemmas in a more straightforward and compact form. It is important to note that the converses are false. For example, suppose X and Y have the same mean, $E[Z] = \mu_Z^+ - \mu_Z^-$ for $Z = X$ or Y. It is not hard to see that it could be the case that $\mu_X^+ < \mu_Y^+$ and $\mu_X^- < \mu_Y^-$ irrespective of the value of x_0. In the following, we illustrate how the lemma allows us to revisit and reinterpret the lemma from Levy and Levy (2004) in a more succinct and illuminating way.

With Lemmas 4.5, 4.5, and 4.17, one could restate some of the lemmas or theorems in related literature with a more compact form. For example, the important lemma on p.1035 of Levy and Levy (2004) can be restated as the following.

Example 4.3. *Assume that both F and G are normal CDFs such that $\mu_X < \mu_Y$ and $\sigma_X < \sigma_Y$. Then, G dominates F by PSD iff $\mu_X^- \geq \mu_Y^-$.*

In this example, two normal CDFs, F and G, are compared. Despite X having a lower expected return and less return standard deviation than Y, the dominance in terms of PSD isn't determined solely by these factors. Instead, G dominates F by PSD if and only if the expected loss for X is at least as large as that for Y. This implies that, all else being equal, a risk-averse decision-maker would prefer Y over X due to its lower expected loss. We generalized the above results as shown in the following theorem.

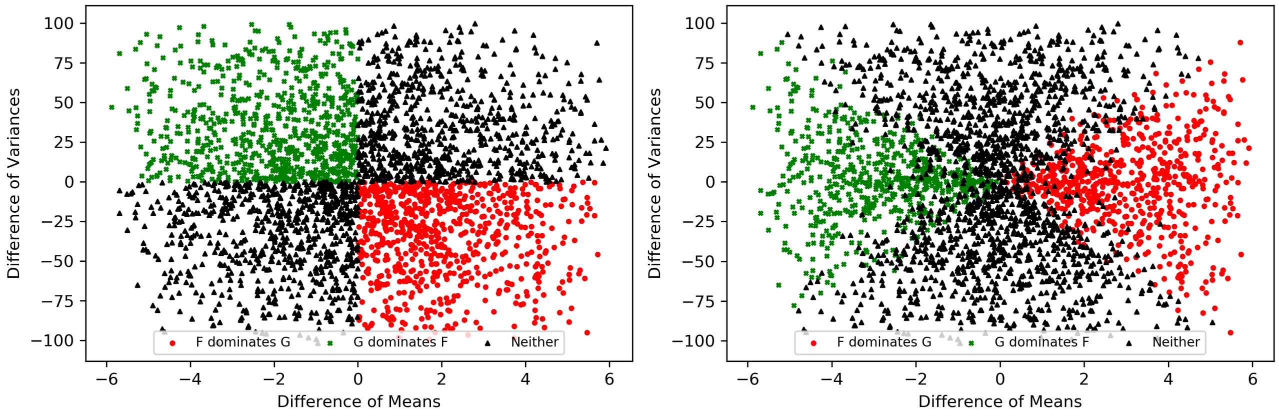
Theorem 4.18. *Let $X \sim \mathcal{E}_1^*(\mu_X, \Sigma_X, \phi)$ and $Y \sim \mathcal{E}_1^*(\mu_Y, \Sigma_Y, \phi)$ such that $\sigma_X^2 \leq \sigma_Y^2$. Then, the following statements are true:*

1. *$F \succeq_P G$ if and only if $\mu_X > \mu_Y$ and $\mu_X^+ \geq \mu_Y^+$.*
2. *$G \succeq_P F$ if and only if $\mu_X < \mu_Y$ and $\mu_X^- \geq \mu_Y^-$.*

unless $F = G$ almost everywhere.

The theorem posits that, given that volatility associated with X is lower, F is preferred over G when the expected value of X exceeds that of Y, and the expected gain associated with X is not less than that of Y. This insight is particularly crucial in financial decision-making, suggesting that a financial asset with a superior expected value combined with a comparable potential for upside is inherently more desirable. On the converse, given that volatility associated with Y is higher,G is preferred over F when X has a lower expected value than Y but carries a comparable (or potentially greater) risk of downside, as captured by the measure of expected loss. This finding is enlightening. To rephrase, it implies that despite the elevated volatility linked to Y, the downside risk plays a more decisive role in determining preferences between X and Y for any u in $U^S(F, G)$.

The second part of the Theorem 4.18 is also relevant to Levy and Levy (2004) because they show that a segment of the MV frontier may be PSD-inefficient. With Theorem 4.18, one could just compare the portfolios in that segment with other portfolios on the MV frontier and find out which portfolios are PSD-efficient, meaning that there is no other portfolio that dominates it in the PSD sense. We will illustrate this in Section 4.4.3.



(A) MV rule (B) modified MV rule

FIGURE 4.2: Comparison between MV rule and modified MV rule

Using Theorem 4.18, we can get a rule to compare normal distributions in the sense of PSD. We call it the modified MV rule which can be stated formally in the following corollary.

Corollary 4.19 (Modified MV rule for prospect theory)**.** *Let* $X \sim \mathcal{E}_1^*(\mu_X, \Sigma_X, \phi)$ *and* $Y \sim \mathcal{E}_1^*(\mu_Y, \Sigma_Y, \phi)$*. Then, the following statements are true:*

1. *If* $\mu_X > \mu_Y$*,* $\mu_X^+ \geq \mu_Y^+$*, and* $\sigma_X^2 \leq \sigma_Y^2$*, then* $E_F u(x) \geq E_G u(x)$*.*
2. *if* $\mu_X > \mu_Y$*,* $\mu_X^- \leq \mu_Y^-$*, and* $\sigma_X^2 \geq \sigma_Y^2$*, then* $E_F u(x) \geq E_G u(x)$*.*

for all $u \in U^S(F, G)$*.*

With Theorem 4.18 and Corollary 4.19, one may still want to see the relationship between the MV rule and PSD. Recalled that if F dominates G by the MV rule (denoted by $F \succeq_{MV} G$) it means that $\mu_X \geq \mu_Y$ and $\sigma_X^2 \leq \sigma_Y^2$, with at least one of the inequalities being strict. In order to have a better understanding on the relationship between the MV rule and PSD, we generate 3,000 pairs of normally distributed series with random mean and random variance and use them to compare the MV rule with the modified MV rule, as stated in Corollary 4.19. The comparison is shown in Figure 4.2.

In Figure 4.2, we refer to the first distribution in a pair as F and the second distribution as G. We calculate the means and variances for all 3,000 pairs of distributions and plot the difference $\mu_X - \mu_Y$ and $\sigma_X^2 - \sigma_Y^2$ of each pair in both Figures 4.2a and 4.2b. In Figure 4.2a, we use red to represent the situation that F has a larger mean and smaller variance and we use green to represent the situation that G has a larger mean and smaller variance. In short, we denote F dominates G under the MV rule by red and G dominates F under the MV rule by green in Figure 4.2a. In Figure 4.2b, we use red to represent the situation in which F has a strictly larger mean, larger upper first partial moment, and smaller variance, or when F has a strictly larger mean, smaller lower first partial moment, and higher variance. We use green to represent the situation in which G has a strictly larger mean, larger upper first partial moment, and smaller variance, or when G has a strictly larger mean, smaller lower first partial moment, and higher variance. In short, we denote F dominates G under the modified MV rule by red and G dominates F under the modified MV rule by green in Figure 4.2b.[12] In Figure 4.2, we see that when a distribution dominates another distribution by the MV rule (e.g. $F \succeq_{MV} G$), it is not necessary that the distribution also dominates another distribution by the PSD rule (e.g. $F \succeq_P G$). The reason is that, given that the distribution has a higher mean and smaller variance, say $\mu_X \leq \mu_Y$ and $\sigma_X^2 \leq \sigma_Y^2$, it may not be the case that the distribution has a strictly higher mean or higher upper first partial moment; that is, $\mu_X = \mu_Y$ or $\mu_X^+ < \mu_Y^+$, which implies $F \not\succeq_P G$ by Theorem 4.18.

On the other hand, Figure 4.2 shows that we will never get a "wrong" result when we use the MV rule in a sense that a distribution dominates another distribution by the MV rule (e.g. $F \succeq_{MV} G$). It will never be the case in which the distribution is dominated by another distribution by the PSD rule (e.g. $G \succeq_P F$) and vice versa. This implies that the discussion on whether concave utility functions or S-shaped utility functions better describe human behaviour might not be that important when we try to compare normal distributions. One can just use the first part of the modified MV rule and if we find out, for example, F dominates G by the first part of the modified MV rule, then F also dominates G by the MV rule which means F is preferred by all investors with a concave utility function and all investors with S-shaped utility functions.

4.4.3 Mean-Variance Portfolio Choice and Prospect Theory

Levy and Levy (2004) is one of the very first studies to examine how prospect theory may

[12] We draw the mean μ_i for distribution i from a uniform distribution such that $\mu_i \in [-3, 3]$. The graph doesn't change much when we change it to $\mu_i \in [0, 3]$.

affect the MV rule by using the PSD rule. [13] Under the assumptions that returns follow a multivariate normal distribution, portfolios can be formed without restrictions, and no two assets are perfectly correlated. Levy and Levy (2004) find out that, with diversification, the elements of an MV-efficient set which is excluded from the PSD-efficient set is at most one segment on the MV frontier.[14] A portfolio is said to be MV-efficient if there is no other portfolio that will dominate it by the MV-rule. Such MV-efficient portfolios can be identified by employing the standard approaches to discover the Markowitz Efficient Frontier. A portfolio is said to be PSD-efficient if there is no other portfolio that will dominate it in the PSD sense. In this section, we make use of the results in Section 4.4, and illustrate that we could identify which portfolios are actually excluded from the PSD-efficient set and which are not.

A nice characteristic of multivariate normally distributed random variables is that their linear combination also yields a normally distributed random variable. This suggests that conclusions drawn from analyzing assets with normally distributed returns can easily extend to portfolios composed of assets with multivariate normally distributed returns. Interestingly, this holds true even when we examine an elliptical distribution. Adapting Theorem 2.16 from Fang et al. (2018), we arrive at the following modified version.

Proposition 4.20. *Assume that $\boldsymbol{Z} \sim \mathcal{E}_N(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$ with rank($\boldsymbol{\Sigma_Z}$) $= k$, and $\theta \in \mathbb{R}^N$, then*

$$\theta' \boldsymbol{Z} \sim \mathcal{E}_1(\theta' \boldsymbol{\mu_X}, \theta' \boldsymbol{\Sigma} \theta, \phi)$$

Let $\theta_1, ..., \theta_N$ as portfolio weights on assets $1, ..., N$ with corresponding random returns $Z_1, ..., Z_N$ where $\boldsymbol{Z} = (Z_1, ..., Z_N)$. In this section, we assume that only a finite quantity of each asset can be shorted, thus the set of possible asset allocations is defined as:

$$\Theta = \Big\{ \theta \in \mathbb{R}^N : \sum_{i=1}^{N} \theta_i = 1, \theta_n \geq M \ \text{ for } \ n = 1, ..., N \Big\}.$$

[13] Levy et al. (2012) demonstrates that the Sharpe-Lintner-Mossin homogeneous expectation capital asset pricing model's security market line theorem remains consistent within the cumulative prospect theory framework. Meanwhile, Levy and Levy (2021) highlights that prospect theory suggests a substantial and abrupt shift in optimal allocation as the investment horizon expands.

[14] These assumptions need to be checked when dealing with real word data. For example, the assumptions that returns follow a multivariate normal distribution could be checked by using the method proposed in Richardson and Smith (1993).

where M is any negative number that takes a finite value, that is $M \in \mathbb{R}_{\leq 0}$. Our analysis will primarily concentrate on portfolios' return $\theta'\boldsymbol{Z}$ such that $\theta \in \Theta$. Furthermore, we assume that $\boldsymbol{Z} \sim \mathcal{E}_N^*(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$ and that no pair of asset returns are perfectly correlated.

In the context of portfolio distributions, Chamberlain (1983) posits that if asset returns are jointly elliptically distributed, the mean and variance alone can describe these distributions. This infers the validity of MV analysis for an elliptical distribution.[15]

Equipped with this understanding, we're now poised to examine the connection between MV-efficient portfolios and PSD-efficient portfolios, thereby expanding on the insights presented by Levy and Levy (2004).

Theorem 4.21 (Levy & Levy Revised)**.** *Consider N assets with corresponding return vector $\boldsymbol{Z} \sim \mathcal{E}_N^*(\boldsymbol{\mu_Z}, \boldsymbol{\Sigma_Z}, \phi)$ and no pair of asset returns being perfectly correlated. For portfolios generated by these N assets with different weight vector $\theta \in \Theta$, then the following statements are true:*

1. *The PSD-efficient set is a subset of the MV-efficient set.*
2. *The portfolio in the MV-efficient set is PSD-efficient if and only if no other portfolio in the MV-efficient set offers a strictly higher expected return and weakly lower first lower partial moment of return.*

The portfolios on the MV frontier will be those hard to rank by the MV rule, meaning that there is no other portfolio that can offer a higher expected return and lower return variance compared to that portfolio. On the MV frontier, we have portfolios with higher expected return and higher return variance compared to some other portfolios on the frontier, and the Theorem 4.18 will be useful in this case. When considering two portfolios, each represented by the return distributions F and G such that $\sigma_X^2 \leq \sigma_Y^2$, if $\mu_X < \mu_Y$ and $\mu_X^- \geq \mu_Y^-$ are observed, the portfolio corresponding to F is deemed PSD-inefficient. This inefficiency implies that a preferable portfolio exists for investors with preferences in line with prospect theory.

Given this consideration, we suggest the implementation of a modified MV frontier. This revision maintains the original MV frontier while excluding the segment of portfolios that are

[15] Schuhmacher et al. (2021) gives some justification that MV analysis is valid even for a specific skew-elliptical distribution.

not part of the PSD-efficient set. Theorem 4.21 refined the results of Levy and Levy (2004) in significant ways. First, rather than strictly using the multivariate normal distribution, our work hinges on the more expansive elliptically distributed random vector, which captures a broader range of return distributions. Second, while Levy and Levy (2004) suggests the PSD-efficient set might exclude one segment from the MV-efficient set, our research clearly delineates criteria for determining when MV-efficient portfolios align with PSD-efficiency, providing sharper insights for investment strategies.

We illustrate the modified MV frontier using 100 independent normally distributed return series, along with the scalarization method for the multi-criterion problem in Boyd and Vandenberghe (2004).[16] The resulting MV frontier and modified MV frontier are shown in Figure 4.3.

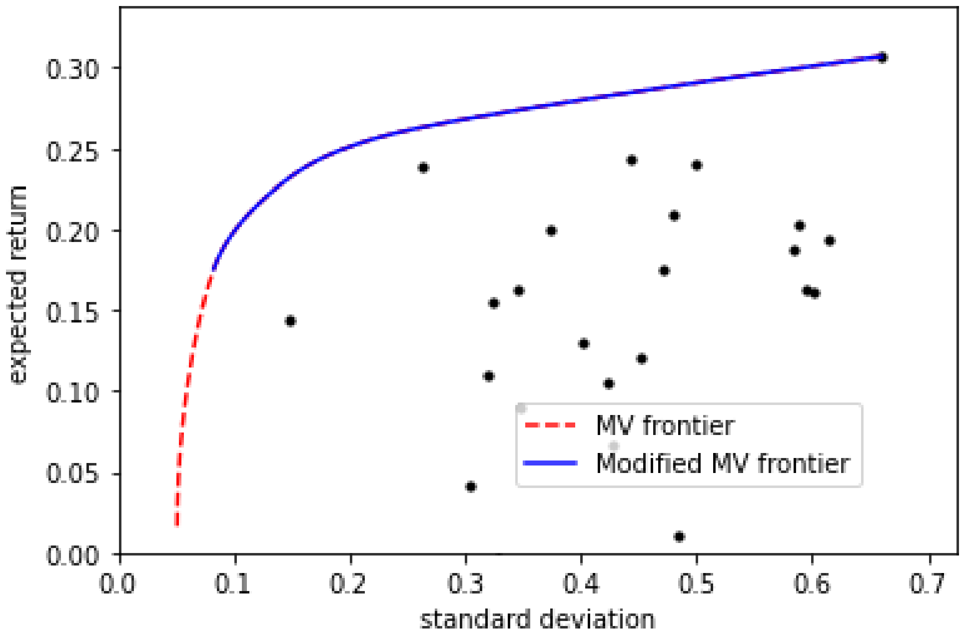


FIGURE 4.3: Standard Deviation-Return Trade-off

In Figure 4.3, the red line represents the segment of the MV frontier which is PSD-inefficient and the blue line represents the modified MV frontier which is both MV and PSD-efficient. Using the Theorem 4.18 improve the result in Levy and Levy (2004) since it could actually help us to identity which portfolios are PSD-inefficient while Levy and Levy (2004) give out which

[16]Some books call it the weighted sum method. Another famous method is the ε-constrained method. In the current setting, there is not much difference between them. See Theorem 4.6 in Ehrgott (2005) for their connection.

portfolios are the potential one.[17] Portfolios on the modified MV frontier are both MV-efficient and PSD-efficient.

4.5 Portfolio Optimization based on PSD

In the preceding section, we delved into the subject of PSD-efficient portfolios, assuming these portfolios were formulated from base assets with elliptically distributed returns. Without specific assumptions regarding the distribution, identifying PSD-efficient portfolios can be quite a complex task. This complexity arises due to $\succeq_P$ being a partial order, which typically hinders the use of a real value function for ranking the available choices. Consequently, we cannot leverage traditional optimization techniques to pinpoint portfolios bearing the highest "PSD rank". In this section, our focus will shift towards a distinct portfolio optimization problem that can integrate the concept of PSD.

4.5.1 General Case

Let $R_1, ..., R_N$ be random returns for assets $1, ..., N$, with joint cumulative distribution function $\mathcal{F}(\mathbf{r})$ where $\mathbf{r} \in R_1 \times ... \times R_N$. Assume that $E[{R_n}^2] < \infty$ for all $n = 1, ..., N$, and further let $w_1, ..., w_N$ as portfolio weights on assets $1, ..., N$. Without short-selling, the set of possible asset allocations is defined as:

$$W = \left\{ w \in \mathbb{R}^N : \sum_{i=1}^{N} w_i = 1, w_n \geq 0 \;\text{ for }\; n = 1, ..., N \right\}.$$

The marginal CDF $\mathcal{F}_w(\eta)$ for asset allocation $w \in W$ is defined as:

$$\mathcal{F}_w(\eta) = \int_{\{w'\mathbf{r} \leq \eta\}} d\mathcal{F}(\mathbf{r})$$

where the corresponding portfolio return is denoted as $R_w = R_1 w_1 + ... + R_N w_N$. Letting Y be a target random return with finite expected value and corresponding CDF G. Consider the distribution G as a benchmark. A compelling inquiry arises: Given a group agents with

[17] The graph seems to violate the condition $G \succeq_P F$ implies $\frac{\mu_Y}{\sigma_Y} > \frac{\mu_X}{\sigma_X}$ mentioned in Levy and Levy (2004). The reason is that the return on the portfolio that is generated may not be sufficiently "normal" to show that conditions will also appear.

preferences shaped by an S-shaped utility function, how might we identify a portfolio that outperforms this benchmark? Under these settings, we can form the following problem:

$$\begin{aligned} &\max E[R_w] \\ \text{subject to } &\mathcal{F}_w \succeq_P G, \\ &w \in W. \end{aligned} \tag{A.4}$$

where $E[R_w]$ is the expected portfolio return. A similar problem for risk averters (based on SSD) is first considered in Dentcheva and Ruszczyński (2003) and further analyzed in Post and Kopa (2017), Post et al. (2018), and Fang and Post (2022). The problem identified in this literature can be explained as follows: given the expected utility of choosing portfolio random return R_w is higher than the target random return Y for all utility functions $u \in U^S(F, G)$, the objective function $E[R_w]$ is maximized. Since $\succeq_P$ is a partial order, this denotes the absence of a single objective function to represent it in general. Typically, optimization under these conditions presents substantial challenges. Nevertheless, the proposed formulation circumvents these inherent challenges, introducing a viable avenue for optimization. For the above problem, it is not clear how the assets allocations change will affect the cumulative distribution function of the corresponding portfolio return R_w. In order to solve the above problem, we use Corollary 4.11 and transform problem (A.4) to an equivalent problem. This is shown below.

$$\max \; E[R_w]$$

$$\text{subject to } \; \big[\mathcal{L}_Y(0) - \mathcal{L}_Y(\eta_1)\big] - \big[\mathcal{L}_{R_w}(0) - \mathcal{L}_{R_w}(\eta_1)\big] \;\geq\; 0, \quad \forall \eta_1 \leq 0, \tag{A.5}$$

$$\big[\mathcal{U}_{R_w}(0) - \mathcal{U}_{R_w}(\eta_2)\big] - \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(\eta_2)\big] \;\geq\; 0, \quad \forall \eta_2 \geq 0, \tag{A.6}$$

$$w \in W.$$

Consider that we have many finite states $t = 1, ..., T$, let r_{nt} denote the return of asset n in period t, p_t as probabilities of these realizations where $p_t = 1/T$. Assume random variables under concerns are uniform-bounded, then $\mathcal{L}_Z(\eta)$ could be express as follows:

$$\mathcal{L}_Z(\eta) = \sum_{t=1}^{T} p_t S_t(\eta)$$

with variables $S_t(\eta)$ and corresponding binary variables b_t that satisfy the following inequalities:

$$\begin{aligned}
&S_t(\eta) \geq \eta - \sum_{n}^{N} w_n r_{nt}, && t = 1, ..., T,\\
&S_t(\eta) \leq \eta - \sum_{n}^{N} w_n r_{nt} - K_L b_t(\eta), && t = 1, ..., T,\\
&S_t(\eta) \geq 0, && t = 1, ..., T,\\
&S_t(\eta) \leq K_U(1 - b_t(\eta)), && t = 1, ..., T,\\
&b_t(\eta) \in \{0, 1\}, && t = 1, ..., T.
\end{aligned} \tag{A.7}$$

where $K_L = \min\{y_t, r_{nt}\} - \max\{y_t, r_{nt}\}$ and $K_U = \max\{y_t, r_{nt}\} - \min\{y_t, r_{nt}\}$. The upper bound of $\mathcal{L}_Z(\eta)$ could be expressed as

$$\mathcal{L}_Z(\eta) \leq \sum_{t=1}^{T} p_t S_t^*(\eta)$$

with variables $S_t^*(\eta)$ that satisfies the following inequalities:

$$\begin{aligned}
&S_t^*(\eta) \geq \eta - \sum_{n}^{N} w_n r_{nt}, && t = 1, ..., T,\\
&S_t^*(\eta) \geq 0, && t = 1, ..., T.
\end{aligned} \tag{A.8}$$

For $\mathcal{U}_H(\eta)$, its upper bound could be defined in a similar way by replacing $\eta - \sum_n^N w_n r_{nt}$ with $\sum_n^N w_n r_{nt} - \eta$. We also replace $S_t(\eta)$ with $D_t(\eta)$, $S_t^*(\eta)$ with $D_t^*(\eta)$ and $b_t(\eta)$ with $c_t(\eta)$ to denote new decision variables. The above formation makes use of the standard technique on handling max operator in mixed integer linear programming and linear programming. It is common in the literature to use ideas related to (A.7) to solve portfolio optimization problems with second or higher order stochastic dominance constraints, including Dentcheva and Ruszczyński (2003), Dentcheva and Ruszczyński (2006), Post et al. (2018), and Fang and Post (2022). Using these concepts, one can replace (A.5) and (A.6) with the following inequalities.

$$
\begin{aligned}
&\big[\mathcal{L}_Y(0) - \mathcal{L}_Y(\eta_1)\big] \geq \sum_{t=1}^{T} p_t \big[S_t^*(0) - S_t(\eta_1)\big], && \forall \eta_1 < 0,\\
&\big[\mathcal{U}_Y(0) - \mathcal{U}_Y(\eta_2)\big] \leq \sum_{t=1}^{T} p_t \big[D_t(0) - D_t^*(\eta_2)\big], && \forall \eta_2 > 0,\\
&S_t(\eta_1) \geq \eta_1 - \sum_{n=1}^{N} w_n r_{nt}, && \forall \eta_1 < 0, t = 1, ..., T,\\
&S_t(\eta_1) \leq \eta_1 - \sum_{n=1}^{N} w_n r_{nt} - K_L b_t(\eta_1), && \forall \eta_1 < 0, t = 1, ..., T,\\
&S_t^*(0) \geq -\sum_{n=1}^{N} w_n r_{nt}, && t = 1, ..., T,\\
&D_t(0) \geq \sum_{n=1}^{N} w_n r_{nt}, && t = 1, ..., T,\\
&D_t(0) \leq \sum_{n=1}^{N} w_n r_{nt} - K_L c_t(0), && t = 1, ..., T,\\
&D_t^*(\eta_2) \geq \sum_{n=1}^{N} w_n r_{nt} - \eta_2, && \forall \eta_2 > 0, t = 1, ..., T,\\
&S_t(\eta_1) \leq K_U(1 - b_t(\eta_1)), && \forall \eta_1 < 0, t = 1, ..., T,\\
&D_t(0) \leq K_U(1 - c_t(0)), && t = 1, ..., T,\\
&S_t(\eta_1), S_t^*(0), D_t(0), D_t^*(\eta_2) \geq 0, && \forall \eta_1 < 0 < \eta_2, t = 1, ..., T,\\
&b_t(\eta_1), c_t(0) \in \{0, 1\}, && \forall \eta_1 < 0, t = 1, ..., T. \qquad \text{(A.9)}
\end{aligned}
$$

4.5.2 Illustration

For practical purposes, we can set up $2M + 1$ return levels such that $\eta_1^M < \eta_1^{M-1} < ... < \eta_1^1 < \eta_1^0 = 0 < \eta_2^1 < ... < \eta_2^{M-1} < \eta_2^M$ and use them to replace the condition $\forall \eta_1 \leq 0 \leq \eta_2$ in order to get an approximate solution. Target random return Y has a discrete distribution with realizations $y_t, t = 1, ..., T$. The realizations could be separated into two parts y_i^d for $i = 1, ..., T_1$ and y_j^a for $j = 1, ..., T_2$, so that $y_{T_1}^d \leq ... \leq y_2^d \leq y_1^d < 0 \leq y_1^a \leq y_2^a \leq ... \leq y_{T_2}^a$, and $T_1 + T_2 = T$. Under this setting, one can get the following.

Theorem 4.22. *If Y has a discrete distribution with realizations $y_t, t = 1, ..., T$, then following inequality imply (A.6) holds:*

$$\big[\mathcal{U}_{R_w}(0) - \mathcal{U}_{R_w}(y_j^a)\big] - \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(y_j^a)\big] \;\; \geq \;\; 0, \quad j = 1, ..., T_2. \tag{A.10}$$

With the above theorem, let $L_i = \mathcal{L}_Y(0) - \mathcal{L}_Y(\eta_1^i)$, $U_j = \mathcal{U}_Y(0) - \mathcal{U}_Y(y_j^a)$, $s_{it} = S_t(\eta_1^i)$, $d_{0t} = D_t(0)$, $d_{jt} = D_t(y_j^a)$, $K_L = \min\{y_t, r_{nt}\} - \max\{y_t, r_{nt}\}$, $K_U = \max\{y_t, r_{nt}\} - \min\{y_t, r_{nt}\}$, $b_{it} = b_t(\eta_1^i)$, and $c_{jt} = c_t(y_j^a)$ for $i = 0, ..., M$ and $j = 1, ..., T_2$, then we can get a mixed integer programming problem as follow:

$$\begin{aligned}
\max_{w,s,s^*,d,d^*,b,c} \quad & \sum_{t=1}^{T}\sum_{n=1}^{N} p_t w_n r_{nt} & \\
\text{subject to} \quad & \sum_{t=1}^{T} p_t\big[s_{0t}^* - s_{it}\big] \leq L_i, & i = 1, ..., M,\\
& \sum_{t=1}^{T} p_t\big[d_{0t} - d_{jt}^*\big] \geq U_j, & j = 1, ..., T_2,\\
& \sum_{n=1}^{N} w_n r_{nt} + s_{it} \geq \eta_1^i, & i = 1, ..., M, t = 1, ..., T,\\
& \sum_{n=1}^{N} w_n r_{nt} + s_{it} + K_L b_{it} \leq \eta_1^i, & i = 1, ..., M, t = 1, ..., T,\\
& s_{it} + K_U b_{it} \leq K_U, & i = 1, ..., M, t = 1, ..., T,\\
& \sum_{n=1}^{N} w_n r_{nt} + s_{0t}^* \geq \eta_1^0, & t = 1, ..., T,\\
& -\sum_{n=1}^{N} w_n r_{nt} + d_{0t} \geq 0, & t = 1, ..., T,\\
& -\sum_{n=1}^{N} w_n r_{nt} + d_{0t} + K_L c_{0t} \leq 0, & t = 1, ..., T,\\
& d_{0t} + K_U c_{0t} \leq K_U, & t = 1, ..., T,\\
& -\sum_{n=1}^{N} w_n r_{nt} + d_{jt}^* \geq -y_j^a, & j = 1, ..., T_2, t = 1, ..., T,\\
& s_{it}, s_{0t}^*, d_{0t}, d_{jt}^* \geq 0, & i = 1, ..., M, j = 1, ..., T_2, t = 1, ..., T,\\
& b_{it}, c_{0t} \in \{0, 1\}, & i = 1, ..., M, t = 1, ..., T,\\
& \sum_{n=1}^{N} w_n = 1, & \\
& w_n \geq 0, & n = 1, ..., N.
\end{aligned} \tag{A.11}$$

One can solve the above problem with a standard mixed integer linear programming solver and the solution of the above problem will be an approximate solution for the original problem. [18] We illustrate this by using the data from Dentcheva and Ruszczyński (2003) and letting the target portfolio as an equal-weighted one. The optimal portfolio w_p which is based on PSD is $w_p = (0, 0, 0, 0.311, 0, 0.060, 0.369, 0.261)$. We compare it with the optimal portfolio w_s based on SSD $w_s = (0, 0, 0.068, 0.188, 0, 0.391, 0.231, 0.122)$. The probability distribution of these portfolios' returns are plotted in Figure 4.4. The optimal portfolio based on PSD has a mean return of 11.78% and the optimal portfolio based on SSD has a mean return of 11.01%. We denote returns of w_p as PSD and w_s as SSD in the figure.

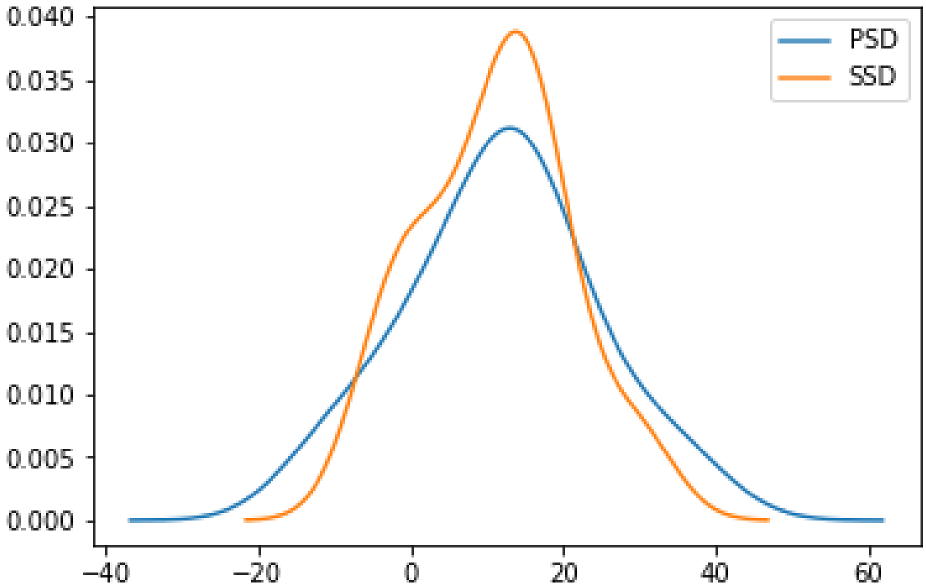


FIGURE 4.4: Comparing portfolio optimization based on PSD with the one based on SSD

There are two interesting observations to be made here. It seems that an optimal portfolio based on PSD will normally get us a higher return and volatility compared to an optimal portfolio based on SSD. The reason is quite intuitive since the first one doesn't require risk averse everywhere. Thus, we can get a higher return by not considering inappropriate "risk". The ability to disregard certain forms of risk, particularly those deemed 'inappropriate' or 'irrelevant', enables us to reap greater rewards. This flexibility, embedded in the PSD model, provides opportunities for increased returns while accepting higher volatility as part of the "risk-return trade-off". Another observation is that the optimal portfolio, which is based on PSD, seems to get us a portfolio which is less diversified compared to an optimal portfolio

[18] For details related to mixed integer linear programming one could read Bertsimas and Tsitsiklis (1997).

which is based on SSD. This discrepancy indicates that the selection and weighting of assets in a portfolio based on PSD leans towards a more concentrated selection, despite the commonly held belief that diversification is key to risk mitigation.

4.6 Concluding remarks

In this study, we explored the use of MV analysis and portfolio optimization for investors with diverse preferences in line with prospect theory. To do so, we employed the concept of PSD and incorporated partial moments in the analysis. Our results demonstrated that partial moments, as return-risk measures, are consistent with PSD and can be utilized to eliminate PSD-inefficient investments.

The relationship between prospect theory and MV analysis is not straightforward, as the two theories differ significantly. It may seem that prospect theory undermines the validity of MV analysis, but our results, along with those of Levy and Levy (2004), showed that a portion of the MV-efficient set is also PSD-efficient. This portion of the results relates to a simple rule we call the modified MV rule. By including partial moments, the modified MV rule allows for the ranking of investments for different value functions under prospect theory. This also resolves an issue with the original MV rule, which cannot rank investments with both higher mean and higher variance, a common occurrence in the real world. Empirical studies, including those conducted by Kahneman and Tversky (1979), Fiegenbaum (1990), Sinha (1994), and Duxbury and Summers (2004), have shown that people often prefer higher variance when faced with losses. Using the modified MV rule, we identified the specific segment on the MV frontier that is PSD-efficient, enabling the creation of a modified MV frontier that is both MV-efficient and PSD-efficient.

In addition to this, we presented a portfolio optimization method, based on partial moments, for constructing a portfolio that dominates a target investment in the PSD sense. This method does not rely heavily on distributional assumptions and can be solved using standard mixed integer linear programming solvers. An optimal portfolio constructed using the PSD approach tends to yield higher returns and volatility compared to an optimal portfolio based on the SSD approach. This outcome can be attributed to the intuitive reason that the PSD model does not impose risk aversion uniformly across all scenarios. As a result, by disregarding certain types of risk that are deemed inappropriate, we can achieve higher returns. The ability to selectively ignore specific forms of risk allows us to capitalize on opportunities for increased returns while acknowledging the presence of higher volatility as part of the risk-return trade-off. Additionally, it is worth noting that the optimal portfolio derived from the PSD seems to exhibit lower diversification compared to an optimal portfolio based on the SSD model. This observation highlights the trade-off between diversification and potential returns within the PSD framework. Overall, this simple framework can assist investors who adhere to prospect theory in making more efficient investment decisions.

Lastly, we want to end with a remark that the tools developed in this chapter might also be applicable to institutional investors. This assertion stems from the observation that these investors occasionally exhibit risk-seeking behaviors in the face of losses, contrary to the conventional belief. A case in point is the employment of the Martingale strategy, where investment values are systematically increased post-losses, which is an implication of risk-seeking on losses. This behavior aligns with the experimental findings such as those presented in Abdellaoui et al. (2013), wherein financial professionals exhibited tendencies consistent with the prospect theory. The tools developed in this study can also be viewed as valuable instruments for situations where a degree of risk-seeking is permissible or even desirable.

4.7 Appendix A: Proofs of Lemmas and Theorems

Proof of Theorem 4.3

Sufficiency: Step 1. Let $E = \{x : G(x) - F(x) \geq 0\}$, and $E^c = \{x : G(x) - F(x) < 0\}$, with their corresponding indicator functions $I_E(x)$ and $I_{E^c}(x)$. Define a function $Q : \mathbb{R}^- \to \mathbb{R}^-$ as

$$Q(x) = \sup\left\{x' \leq 0 \;\middle|\; \int_{x'}^{0}[G(z) - F(z)]I_E(z)\,dz = \int_{x}^{0}[F(z) - G(z)]I_{E^c}(z)dz\right\}$$

The exists of x' for each x is guarantee from the fact that $\int_{x'}^{0}[G(z) - F(z)]I_E(z)$ is a continuous function of x' and $\int_{x}^{0}[G(z) - F(z)]\,dz \geq 0$ for all $x \leq 0$. Since the supremum is unique, Q is well-defined. Note that under this definition, $F \succeq_P G$ implies $Q(x) \geq x$ for all $x \leq 0$. Also note that $[F(z) - G(z)]I_{E^c}(z) \geq 0$ for all z, thus $Q(x)$ is increase with x. Since $Q(x)$ is monotone, it is also continuous and differentiable almost everywhere.

Differentiating $\int_{Q(x)}^{0}[G(z) - F(z)]I_E(z)\,dz = \int_{x}^{0}[F(z) - G(z)]I_{E^c}(z)dz$ at any point when $Q'(x)$ exists, then one can get $I_E(Q(x))[G(Q(x)) - F(Q(x))]Q'(x) = [F(x) - G(x)]I_{E^c}(x)$ to hold almost everywhere over R^-. Suppose that $u \in U^S(F, G)$, then for any $x \leq 0$ we have:

$$\int_{x}^{0} udF = uF\Big|_{x}^{0} - \int_{x}^{0} Fdu \quad \text{and} \quad \int_{x}^{0} udG = uG\Big|_{x}^{0} - \int_{x}^{0} Gdu.$$

For $\int_{x}^{0} |G - F|du$, we have:

$$\int_{x}^{0} |G - F|du \leq \int_{x}^{0} Gdu + \int_{x}^{0} Fdu \leq uG\Big|_{x}^{0} + uF\Big|_{x}^{0} - \Big[\int_{x}^{0} udG + \int_{x}^{0} udF\Big] < \infty.$$

Then $\int_{x}^{0}[G - F]du$ for any $x \leq 0$ could be expressed as:

$$\begin{aligned}\int_{x}^{0}[G - F]du &= \int_{x}^{0}[G(z) - F(z)]I_E(z)du(z) + \int_{x}^{0}[G(z) - F(z)]I_{E^c}(z)du(z) \\ &= \int_{x}^{0}[G(z) - F(z)]I_E(z)du(z) - \int_{x}^{0} I_E(Q(z))[G(Q(z)) - F(Q(z))]Q'(z)du(z)\end{aligned} \tag{A.12}$$

The last term of the above equation could be expressed as:

$$\begin{aligned}
\int_x^0 I_E(Q(z))[G(Q(z)) - F(Q(z))]Q'(z)du(z) &= \int_x^0 I_E(Q(z))[G(Q(z)) - F(Q(z))]Q'(z)u'(z)dz \\
&\leq \int_x^0 I_E(Q(z))[G(Q(z)) - F(Q(z))]Q'(z)u'(Q(z))dz \\
&= \int_x^0 I_E(Q(z))[G(Q(z)) - F(Q(z))]du(Q(z)) \\
&= \int_{Q(x)}^0 [G(z) - F(z)]I_E(z)du(z)
\end{aligned} \tag{A.13}$$

where the inequality holds because diminishing sensitivity $u^{''}(x) \geq 0$ and $Q(x) \geq x$ for all $x \leq 0$. Then substitute this into equation (A.12) will get us the following:

$$\begin{aligned}
\int_x^0 [G - F]du &\geq \int_x^0 [G(z) - F(z)]I_E(z)du(z) - \int_{Q(x)}^0 [G(z) - F(z)]I_E(z)du(z) \\
&= \int_x^{Q(x)} [G(z) - F(z)]I_E(z)du(z) \geq 0
\end{aligned} \tag{A.14}$$

Step 2. Let $M : R^+ \to R^+$ as:

$$M(x) = \infty \left\{ x' \geq 0 \,\middle|\, \int_0^{x'} [G(z) - F(z)]I_E(z)dz = \int_0^x [F(z) - G(z)]I_{E^c}(z)dz \right\} \tag{A.15}$$

The existence of x' for each $x \geq 0$ is guaranteed from the fact that the first term of the equation is a continuous function of x' and $\int_0^x [G(z) - F(z)]dz \geq 0$ for all $x \geq 0$. Since the infimum is unique, M is well-defined. Note that under this definition, $M(x) \leq x$ for all $x \geq 0$ and $M(x)$ is increase with $x \geq 0$. Since $M(x)$ is monotone, it is also continuous and differentiable almost everywhere.

Differentiating $\int_0^{M(x)} [G(z) - F(z)]I_E(z)dz = \int_0^x [F(z) - G(z)]I_{E^c}(z)dz$ at x when $M'(x)$ exists, one can get $[F(x) - G(x)]I_{E^c}(x) = I_E(M(x))[G(M(x)) - F(M(x))]M'(x)$ which holds almost everywhere over R^+. Suppose that $u \in U^S(F,G)$, then for any $x \geq 0$ using similar argument in Step 1, we have:

$$\int_0^x [G-F]du = \int_0^x [G(z)-F(z)]I_E(z)du(z) + \int_0^x [G(z)-F(z)]I_{E^c}(z)du(z) \tag{A.16}$$

Then $\int_0^x [G-F]du$ could be expressed as:

$$\begin{aligned}\int_0^x [G-F]du &= \int_0^x [G(z)-F(z)]I_E(z)u'(z)dz \\ &\quad - \int_0^x I_E(M(z))[G(M(z))-F(M(z))]M'(z)u'(z)dz\end{aligned} \tag{A.17}$$

Denote the last term as T then it could be expressed as:

$$\begin{aligned}T &\leq \int_0^x I_E(M(z))[G(M(z))-F(M(z))]M'(z)u'(M(z))dz \\ &= \int_0^{M(x)} [G(z)-F(z)]I_E(z)du(z)\end{aligned} \tag{A.18}$$

where the inequality holds because diminishing sensitivity $u''(x) \leq 0$, with the fact that $x \geq M(x) \geq 0$ for all $x \geq 0$. Then substitute this into equation (A.17) for $x \geq 0$ will get us the following:

$$\begin{aligned}\int_0^x [G-F]du &\geq \int_0^x [G(z)-F(z)]I_E(z)du(z) - \int_0^{M(x)} [G(z)-F(z)]I_E(z)du(z) \\ &= \int_{M(x)}^x [G(z)-F(z)]I_E(z)du(z) \geq 0\end{aligned} \tag{A.19}$$

Step 3. Adding (A.14) and (A.19) together will get us the following:

$$\int_{x_1}^{x_2} [G(z)-F(z)]du \geq \int_{x_1}^{Q(x_1)} [G(z)-F(z)]I_E(z)du(z) + \int_{M(x_2)}^{x_2} [G(z)-F(z)]I_E(z)du(z) \tag{A.20}$$

which implies $\int_{x_1}^{x_2}[G(z)-F(z)]du(z) \geq 0$ for all $x_1 \leq 0 \leq x_2$, and $\int_{-\infty}^{\infty}[G(z)-F(z)]du(z) \geq 0$.

Step 4. By Lemma 1 in Hanoch and Levy (1969), and Lemma 1* in Tesfatsion (1976), we have $\int udF - \int udG = \int [G-F]du$, and therefore $E_F u(x) \geq E_G u(x)$ for all $u \in U^S(F,G)$.

Necessity: Suppose that there exists $x' \geq 0$, such that $\int_{-x'}^{0}[G(z)-F(z)]dz < 0$ or $\int_0^{x'}[G(z)-F(z)]dz < 0$. If such x' does not exist, it will means that $\int_{x_1}^{x_2}[G(z)-F(z)]dz \geq 0$ for all $x_1 \leq 0 \leq x_2$ which means $F \succeq_P G$. For $\int_{-x'}^{0}[G(z)-F(z)]dz < 0$, choose $u_1(x) = 0$ for $x \geq 0$, $u_1(x) = x$ for $-x' < x < 0$, and $u_1(x) = -x'$ for $x \leq -x'$. For $\int_0^{x'}[G(z)-F(z)]dz < 0$, choose $u_2(x) = x'$ for $x \geq x'$, $u_2(x) = x$ for $0 < x < x'$, and $u_2(x) = 0$ for $x \leq 0$. By Lemma 1 in Hanoch and Levy (1969), and Lemma 1* in Tesfatsion (1976), we have $\int_{-\infty}^{\infty} u_1 dF - \int_{-\infty}^{\infty} u_1 dG = \int_{-x'}^{0}[G-F]du_1 < 0$ for the first case and $\int_{-\infty}^{\infty} u_2 dF - \int_{-\infty}^{\infty} u_2 dG = \int_0^{x'}[G-F]du_2 < 0$ for the second case. □

Proof of Lemma 4.5. This is a special case of Lemma 4.10, thus the proof will be omitted here.

Proof of Theorem 4.6. $F \succeq_P G$ implies $\int_0^{\infty}[G(x)-F(x)]dx \geq 0$ which, in turn, implies $\mu_X^+ \geq \mu_Y^+$ by Lemma 4.5. Thus, the assertion in first part of Theorem 4.6 holds. Second, $F \succeq_P G$ implies $\int_{-\infty}^{0}[G(x)-F(x)]dx \geq 0$ which, in turn, implies $\mu_X^- \leq \mu_Y^-$ by Lemma 4.5. Thus, the assertion in second part of Theorem 4.6 holds. □

Proof of Lemma 4.7. According to Proposition 3 in Ogryczak and Ruszczyński (1999), we know the following:

$$\tau_Z^+ \;=\; 2\int_0^{\infty}\Big(\int_{-\infty}^{x} H(t)dt - x + \mu_H\Big)dx \;, \tag{A.21}$$

for $H = F$ and G, then we can get:

$$\tau_Z^+ \;=\; 2\int_0^{\infty}\Big(\int_0^{x} H(t)dt + \int_{-\infty}^{0} H(t)dt - x + \mu_H\Big)dx \;. \tag{A.22}$$

Since both τ_X^+ and τ_Y^+ are finite, the difference between them could be expressed as:

$$\tau_Y^+ - \tau_X^+ = 2\int_0^{\infty}\Big(\int_0^{x}[G(t)-F(t)]dt + \int_{-\infty}^{0}[G(t)-F(t)]dt + \mu_Y - \mu_X\Big)dx \;.$$

Thereafter, applying Lemma 4.5, we could get $\tau_Y^+ - \tau_X^+ = 2\int_0^\infty \Big(\int_0^x [G(t)-F(t)]dt + \mu_Y^+ - \mu_X^+ \Big) dx$. Thus, the first part of Lemma 4.7 holds. Second, according to Proposition 3 in Ogryczak and Ruszczyński (1999), we get the following:

$$\tau_Z^- \;=\; 2\int_{-\infty}^{0} \Big(\int_{-\infty}^{x} H(t)dt \Big) dx \tag{A.23}$$

for $H = F$ and G. Since both τ_X^- and τ_Y^- are finite, the difference between them could be expressed as:

$$\begin{aligned}
\tau_Y^- - \tau_X^- &= 2\int_{-\infty}^{0} \Big(\int_{-\infty}^{x} [G(t) - F(t)]dt \Big) dx \\
&= 2\int_{-\infty}^{0} \Big(\int_{-\infty}^{0} [G(t) - F(t)]dt - \int_{x}^{0} [G(t) - F(t)]dt \Big) dx
\end{aligned} \tag{A.24}$$

then using Lemma 4.5 we could get $\tau_Y^- - \tau_X^- = 2\int_{-\infty}^0 \Big(\mu_Y^- - \mu_X^- - \int_x^0 [G(t) - F(t)]dt \Big) dx$. Thus, the second part of Lemma 4.7 holds. $\square$

Proof of Theorem 4.8. By Lemma 4.7, $\tau_Y^+ - \tau_X^+ = 2\int_0^\infty \int_0^x [G(t) - F(t)]dtdx$ when $\mu_Y^+ = \mu_X^+$. $F \succeq_P G$ implies $\int_0^x [G(t) - F(t)]dt \geq 0$ for all $x \geq 0$ which, in turn, implies $\tau_X^+ \leq \tau_Y^+$. Thus, the assertion in first part of Theorem 4.8 holds. Second, $\tau_Y^- - \tau_X^- = 2\int_{-\infty}^0 \Big(-\int_x^0 [G(t) - F(t)]dt \Big) dx$ when $\mu_Y^- = \mu_X^-$. $F \succeq_P G$ implies $-\int_x^0 [G(t) - F(t)]dt \leq 0$ for all $x \leq 0$ which, in turn, implies $\tau_X^- \geq \tau_Y^-$. Thus, the assertion in second part of Theorem 4.8 holds. $\square$

Proof of Lemma 4.10 . We will only show the proof of part 2. For part 1, one could see Proposition 2 in Ogryczak and Ruszczyński(1999). For part 3 and 4, they are just immediate consequences of part 1 and 2. *Part 2.* By Property 4.9, we can get

$$\mathcal{U}_X(x) - \mathcal{U}_Y(x) = \int_{-\infty}^{x} F(z)dz - (x - \mu_X) - \int_{-\infty}^{x} G(z)dz - (x - \mu_Y)$$

Using the fact that $\mu_X - \mu_Y = \int_{-\infty}^{\infty}[G(x) - F(x)]dx$, we can get the following:

$$\begin{aligned}\mathcal{U}_X(x) - \mathcal{U}_Y(x) &= \int_{-\infty}^{x} [F(z) - G(z)]dz + \int_{-\infty}^{\infty} [G(z) - F(z)]dz \\ &= \int_{x}^{\infty} [G(z) - F(z)]dz\end{aligned}$$

□

Proof of Lemma 4.15.

Sufficiency of Part 1. Suppose $\sigma_Y^2 = \sigma_X^2$, we want to show that F and G will not cross. First, note that $\sigma_Y^2 = \sigma_X^2$ if and only if $\sqrt{\Sigma_Y} = \sqrt{\Sigma_X}$ by Property 4.14, thus it is sufficient to show that $\sqrt{\Sigma_Y} = \sqrt{\Sigma_X}$ implies G will not cross F. By Property 4.14 and Definition 4.13, the transformed variables $(X - \mu_X)/\sqrt{\Sigma_X}$ and $(Y - \mu_Y)/\sqrt{\Sigma_Y}$ are both elliptically distributed with mean 0, dispersion parameter 1, and characteristic generator ϕ, that is $\mathcal{E}_1^*(0, 1, \phi)$. Let H as the CDF and h as the PDF for all random variables that are elliptically distributed with mean 0, dispersion parameter 1, and characteristic generator ϕ, for $\mu_X \neq \mu_Y$, we can get:

$$F(x) = H\Big(\frac{x - \mu_X}{\sqrt{\Sigma_X}}\Big) \neq H\Big(\frac{x - \mu_Y}{\sqrt{\Sigma_Y}}\Big) = G(x)\ , \forall x \in \mathbb{R}$$

since H is strictly monotone, and it is either:

1. $\frac{x-\mu_X}{\sqrt{\Sigma_X}} > \frac{x-\mu_Y}{\sqrt{\Sigma_Y}}$, $\forall x \in \mathbb{R}$, or
2. $\frac{x-\mu_X}{\sqrt{\Sigma_X}} < \frac{x-\mu_Y}{\sqrt{\Sigma_Y}}$, $\forall x \in \mathbb{R}$,

when $\mu_X \neq \mu_Y$. For $\mu_X = \mu_Y$, it implies that$(x-\mu_X)/\sqrt{\Sigma_X} = (x-\mu_Y)/\sqrt{\Sigma_Y}$,$\forall x \in \mathbb{R}$,thus we can get:

$$F(x) = H\Big(\frac{x-\mu_X}{\sqrt{\Sigma_X}}\Big) = H\Big(\frac{x-\mu_Y}{\sqrt{\Sigma_Y}}\Big) = G(x) \ ,\forall x \in \mathbb{R}.$$

Thus F and G will not cross when $\sigma_Y^2 = \sigma_X^2$.

Necessity of Part 1. For G will cross F when $\sigma_Y^2 \neq \sigma_X^2$, we will prove by cases. Without loss of generality, assume $\sigma_Y^2 > \sigma_X^2$. First, note that $\sigma_Y^2 > \sigma_X^2$ if and only if $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$ by Property 4.14, thus it is sufficient to show that $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$ implies G will cross F. Suppose that $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$ and G doesn't cross F, then there are only three possible cases:

1. $H\Big(\frac{x-\mu_X}{\sqrt{\Sigma_X}}\Big) > H\Big(\frac{x-\mu_Y}{\sqrt{\Sigma_Y}}\Big)$,$\forall x \in \mathbb{R}$;
2. $H\Big(\frac{x-\mu_X}{\sqrt{\Sigma_X}}\Big) < H\Big(\frac{x-\mu_Y}{\sqrt{\Sigma_Y}}\Big)$,$\forall x \in \mathbb{R}$;
3. $H\Big(\frac{x-\mu_X}{\sqrt{\Sigma_X}}\Big) = H\Big(\frac{x-\mu_Y}{\sqrt{\Sigma_Y}}\Big)$,$\forall x \in \mathbb{R}$.

For case 1, suppose that $H\Big(\frac{x-\mu_X}{\sqrt{\Sigma_X}}\Big) > H\Big(\frac{x-\mu_Y}{\sqrt{\Sigma_Y}}\Big)$,$\forall x \in \mathbb{R}$ then we can get:

$$(x-\mu_X)\frac{\sqrt{\Sigma_Y}}{\sqrt{\Sigma_X}} > (x-\mu_Y).$$

If $\mu_X = \mu_Y$, the left side of the equation will be smaller than the right side when $x < \mu_X$, thus a contradiction. If $\mu_X > \mu_Y$ or $\mu_X < \mu_Y$, consider x such that:

$$\frac{x-\mu_X}{\sqrt{\Sigma_X}} < \frac{x-\mu_Y}{\sqrt{\Sigma_Y}},$$

that is

$$x < \frac{E[X]\sqrt{\Sigma_Y} - E[Y]\sqrt{\Sigma_X}}{\sqrt{\Sigma_Y} - \sqrt{\Sigma_X}},$$

will implies:

$$(x-\mu_X)\frac{\sqrt{\Sigma_Y}}{\sqrt{\Sigma_X}} < (x-\mu_Y),$$

which is a contradiction. The existence of such x is guaranteed by the fact that $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$ and the domain of H is unbounded. The proof for case 2 is similar to case 1.

Consider case 3 that F and G will cross at x_0, that is $F(x_0) = G(x_0)$, then we can get the following:

$$F(x_0) = H\Big(\frac{x_0 - \mu_X}{\sqrt{\Sigma_X}}\Big) = H\Big(\frac{x_0 - \mu_Y}{\sqrt{\Sigma_Y}}\Big) = G(x_0),$$

since H is strictly monotone, we can get:

$$\frac{x_0 - \mu_X}{\sqrt{\Sigma_X}} = \frac{x_0 - \mu_Y}{\sqrt{\Sigma_Y}}.$$

Solving x_0 gets us the following:

$$x_0 = \frac{\mu_X\sqrt{\Sigma_Y} - \mu_Y\sqrt{\Sigma_X}}{\sqrt{\Sigma_Y} - \sqrt{\Sigma_X}}.$$

Since there is only one solution for x_0 when $\sigma_Y^2 > \sigma_X^2$, and we know $\sigma_Y^2 > \sigma_X^2$ if and only if $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$ by Property 4.14, we conclude that F and G will cross at most one time. Thus a contradiction. The analysis of this case also shows that if F and G cross, then F and G will cross at most one time when $\sigma_Y^2 > \sigma_X^2$.

Part 2. Based on the above, we conclude that G will cross F when $\sigma_Y^2 > \sigma_X^2$ and they cross exactly for one time. Since G will cross F exactly for one time when $\sigma_Y^2 > \sigma_X^2$, it is sufficient to prove G will also cross F from above by checking the first derivative of the following equation:

$$H\Big(\frac{x_0 - \mu_X}{\sqrt{\Sigma_X}}\Big) - H\Big(\frac{x_0 - \mu_Y}{\sqrt{\Sigma_Y}}\Big),$$

where $x_0 \in \mathbb{R}$ such that $F(x_0) = G(x_0)$. The first derivative of the above equation will be strictly larger than zero as follows:

$$\frac{1}{\sqrt{\Sigma_X}}h\Big(\frac{x_0 - \mu_X}{\sqrt{\Sigma_X}}\Big) - \frac{1}{\sqrt{\Sigma_Y}}h\Big(\frac{x_0 - \mu_Y}{\sqrt{\Sigma_Y}}\Big) > 0,$$

since $\sqrt{\Sigma_Y} > \sqrt{\Sigma_X}$. This implies that $F(x) > G(x)$ for $x > x_0$ and $F(x) < G(x)$ for $x < x_0$ which mean G will also cross F from above when $\sigma_Y^2 > \sigma_X^2$. $\square$

Proof of Theorem 4.16. Without loss of generality, we assume that $\sigma_X^2 < \sigma_Y^2$. By Lemma 4.15, G will cross F exactly for one time and G will cross F from above in a point x_0.[19] This implies that $G(x) > F(x)$ for all x such that $-\infty > x > x_0$ and $G(x) < F(x)$ for all x such that $x_0 < x < \infty$. With these two conditions, it is impossible to get either $\int_{x_1}^{x_2}[G(z) - F(z)]\, dz \geq 0$ or $\int_{x_1}^{x_2}[G(z) - F(z)]\, dz \leq 0$ for all $x_1 \leq 0 \leq x_2$, no matter x_0 is located in the negative domain, positive domain, or exactly at zero. Consider that x_0 is located in positive domain, then we will get $\int_0^{x_0}[G(z) - F(z)]dz \geq 0$ and $\int_{x_0}^{\infty}[G(z) - F(z)]dz \leq 0$.[20] This implies that G will not dominate F by PSD. Recall that $\mu_X - \mu_Y = \int_{-\infty}^{\infty}[G(x) - F(x)]dx$, by $\mu_X - \mu_Y = 0$, we have $\int_{-\infty}^{\infty}[G(x) - F(x)]dx = 0$;[21] that is, $\int_{-\infty}^{x_0}[G(z) - F(z)]\, dz$ must be equal to $-\int_{x_0}^{\infty}[G(z) - F(z)]\, dz$. Thus, it is impossible to have F dominates G by PSD since $\int_0^{x_0}[G(z) - F(z)]\, dz < -\int_{x_0}^{\infty}[G(z) - F(z)]\, dz$, and thus, $\int_0^{\infty}[G(z) - F(z)]dz < 0$. Similar results could be derived when x_0 is located in the negative domain or exactly at zero. Therefore, the assertion of Theorem 4.16 holds. □

Proof of Lemma 4.17.

Part 1. Consider that $x_0 \leq 0$. Since $\sigma_X^2 < \sigma_Y^2$, G will cross F from above and cross exactly for one time only according to Lemma 4.15. This implies that $F(x) > G(x)$ for all x such that $0 < x < \infty$. In this case, by Lemma 4.5 we must have $\mu_X^+ - \mu_Y^+ = \int_0^{\infty}[G(x) - F(x)]dx < 0$.Thus, the assertion in Part 1 holds.

Part 2. The proof is similar to the one in Part 1. Consider $x_0 \geq 0$. Since $\sigma_X^2 < \sigma_Y^2$, G will cross F from above and cross exactly for one time only. This implies that $F(x) < G(x)$ for all x such that $-\infty < x < 0$. In this case, by Lemma 4.5, we must have $\mu_Y^- - \mu_X^- = \int_{-\infty}^{0}[G(x) - F(x)]dx > 0$. Thus, the assertion in Part 2 holds.

□

[19] For normal distributions as an example, see Properties 2 and 3 in p.87-88 of Levy (2011) in which F and G will cross exactly one time and G will cross F from above since $\sigma_Y^2 > \sigma_X^2$.

[20] See the graph for an example of this case

[21] See p.59 in Levy(2016) for example.

Proof of Theorem 4.18.

Sufficiency of Part 1. Consider the case in which $\sigma_X < \sigma_Y$. Since the variance of G is higher and $\mu_X^+ \geq \mu_Y^+$, by Lemma 4.17, F will cross G from below at x_0 such that $x_0 > 0$. This implies that $F(x) < G(x)$ for all x such that $-\infty < x \leq 0$ which means $\int_{x_1}^{0}[G(z) - F(z)]dz \geq 0$ for all x_1 such that $x_1 < 0$. Since $\mu_X^+ \geq \mu_Y^+$ and $x_0 > 0$, by Lemma 4.5 we can get $\int_0^{x_0}[G(z) - F(z)]dz + \int_{x_0}^{\infty}[G(z) - F(z)]dz \geq 0$, where the first term is positive and the second term is negative. This implies that $\int_0^{x_2}[G(z) - F(z)]dz \geq 0$ for all x_2 such that $x_2 > 0$, and thus, F dominates G by PSD. For the case in which $\sigma_X = \sigma_Y$, since F and G will not intercept, $\mu_X > \mu_Y$ is equivalence to $\mu_X^+ > \mu_Y^+$.[22] Noted that $\mu_X > \mu_Y$ implies that $F(x) < G(x)$ for all x such that $a < x < b$ which also implies that F dominates G by PSD.

Necessity of Part 1. We want to prove that if $\mu_X \leq \mu_Y$ or $\mu_X^+ < \mu_Y^+$, then F will not dominate G by PSD. Consider $\sigma_X^2 < \sigma_Y^2$. The case of $\mu_X = \mu_Y$ has already been considered in Theorem 4.16. For $\mu_X < \mu_Y$, this implies that $\int_{-\infty}^{x_0}[G(z) - F(z)]dz + \int_{x_0}^{\infty}[G(z) - F(z)]dz < 0$ where the first term is positive and the second term is negative. Then, no matter where x_0 is located, it is impossible to have F dominates G by PSD. For $\mu_X^+ < \mu_Y^+$, by Lemma 4.5, we have $\int_0^{\infty}[G(z) - F(z)]dz < 0$. Then, again, no matter where x_0 is located, it is impossible to have F dominates G by PSD. Now, consider $\sigma_X^2 = \sigma_Y^2$. Since $\mu_X < \mu_Y$ is equivalence to $\mu_X^+ < \mu_Y^+$ and implies G dominates F by PSD, F will not dominate G by PSD, and thus, the assertions hold.

Sufficiency of Part 2. The proof is similar to Part 1. Consider the case in which $\sigma_X < \sigma_Y$. Since the variance of G is higher and $\mu_X^- \geq \mu_Y^-$, by Lemma 4.17, F will cross G from below at x_0 such that $x_0 < 0$. This implies that $F(x) > G(x)$ for all x such that $0 \geq x > \infty$ which means that $\int_0^{x_2}[G(z) - F(z)]dz \leq 0$ for all x_2 such that $x_2 > 0$. Since $\mu_X^- \geq \mu_Y^-$ and $x_0 < 0$, by Lemma 4.7, we can get $\int_{-\infty}^{x_0}[G(z) - F(z)]dz + \int_{x_0}^{0}[G(z) - F(z)]dz \leq 0$, where the first term is positive and the second term is negative. This implies that $\int_{x_1}^{0}[G(z) - F(z)]dz \leq 0$ for all x_1 such that $x_1 < 0$, and thus, G dominates F by PSD. For the case in which $\sigma_X = \sigma_Y$, since F and G will not intercept, $\mu_X < \mu_Y$ is equivalence to $\mu_X^- > \mu_Y^-$.[23] Noted that $\mu_X < \mu_Y$ implies that $F(x) > G(x)$ for all x such that $-\infty < x < \infty$ which also implies that G dominates F by PSD.

[22]The case $\mu_X^+ = \mu_Y^+$ will be ignored, since it implies $F = G$.
[23]We skip the discussion for the case $\mu_X^- = \mu_Y^-$ because it implies $F = G$.

Necessity of Part 2. We want to prove that if $\mu_X \geq \mu_Y$ or $\mu_X^- < \mu_Y^-$, then G will not dominate F by PSD. To prove this result, we first consider $\sigma_X^2 < \sigma_Y^2$. The case of $\mu_X = \mu_Y$ has already been considered in Theorem 4.16. For $\mu_X > \mu_Y$, this implies that $\int_{-\infty}^{x_0}[G(z) - F(z)]dz + \int_{x_0}^{\infty}[G(z) - F(z)]dz > 0$ in which the first term is positive and the second term is negative. Then, no matter where x_0 is located, it is impossible to have G dominates F by PSD. For $\mu_X^- < \mu_Y^-$, by Lemma 4.7, we have $\int_{-\infty}^{0}[G(z) - F(z)]dz > 0$. Again, no matter where x_0 is located, it is impossible to have G dominates F by PSD. Now consider $\sigma_X^2 = \sigma_Y^2$. Since $\mu_X > \mu_Y$ is equivalence to $\mu_X^- < \mu_Y^-$ and implies F dominates G by PSD, G will not dominate F by PSD, and thus, the assertion holds. $\square$

Proof of Theorem 4.21.

Part 1. Consider MV-efficient portfolios identified by maximising the expected return of the portfolio given a target portfolio return variance and related constraints. For any portfolio that does not lie on the MV-efficient frontier, there exists a portfolio with an equivalent return variance and higher expected return. Lemma 4.15 implies that the CDF of these two portfolios will not cross. Using the property from Lemma 1 in Hanoch and Levy (1969) and Lemma 1* in Tesfatsion (1976) (e.g $\mu_X - \mu_Y = \int_{-\infty}^{\infty}[G(x) - F(x)]dx$), we can conclude that the portfolio that is not on the MV-efficient frontier will be dominated by the portfolio which has same variance of return and higher mean return in FSD sense which implies the first portfolio is dominated by the latter one in PSD sense. Thus, the PSD-efficient set is a subset of the MV-efficient set.

Part 2. There doesn't exist a portfolio that will offer strictly higher expected return and weakly lower variance of return compared to portfolios in the MV-efficient set. Thus, Theorem 4.18 implies Part 2 holds. $\square$

Proof of Theorem 4.22.

Let $H_2^a(\eta_2) = \int_0^{\eta_2} H(z)dz$ for $0 \leq \eta_2$, and $H = F$ and G. It is sufficient to show (A.10) implies $F_2^a(\eta_2) \leq G_2^a(\eta_2)$ for all $\eta_2 \geq 0$.

For (A.6), given (A.10) holds, we consider three cases. First, consider the case that $0 \leq \eta_2 \leq y_1^a$, we have

$$
\begin{aligned}
F_2^a(\eta_2) &\leq \theta_1 F_2^a(0) + (1-\theta_1)F_2^a(y_1^a) \\
&\leq \theta_1 G_2^a(0) + (1-\theta_1)G_2^a(y_1^a) \\
&= G_2^a(\eta_2)
\end{aligned}
$$

for $\theta_1 = (y_1^a - \eta_2)/(y_1^a)$. The first inequality come from the fact that $H_2^a(\cdot)$ is convex and the last equality holds because $G_2^a(\cdot)$ is piecewise linear. For the case that $y_{j+1}^a \leq \eta_2 \leq y_{j+1}^a$, we have

$$
\begin{aligned}
F_2^a(\eta_2) &\leq \theta_2 F_2^a(y_j^a) + (1-\theta_2)F_2^a(y_{j+1}^a) \\
&\leq \theta_2 G_2^a(y_j^a) + (1-\theta_2)G_2^a(y_{j+1}^a) \\
&= G_2^a(\eta_2)
\end{aligned}
$$

for $\theta_2 = (y_{j+1}^a - \eta_2)/(y_{j+1}^a - y_j^a)$.The first inequality come from the fact that $H_2^a(\cdot)$ is convex and the last equality holds because $G_2^a(\cdot)$ is piecewise linear. Last, consider the case that $y_{T_2}^a \leq \eta_2$, we have

$$
\begin{aligned}
F_2^a(\eta_2) &= F_2^a(y_{T_2}^a) + \int_{y_{T_2}^a}^{\eta_2} F(z)dz \\
&\leq G_2^a(y_{T_2}^a) + \int_{y_{T_2}^a}^{\eta_2} 1dz \\
&= G_2^a(\eta_2)
\end{aligned}
$$

Combining these three cases, we have (A.10) implies $F_2^a(\eta_2) \leq G_2^a(\eta_2)$ for all $\eta_2 \geq 0$ which equivalent to (A.6). $\square$

4.8 Appendix B: Details of Hypothesis tests

Bootstrap is a statistical technique that involves resampling a dataset multiple times to estimate a statistic. In the code, the bootstrap method is used to generate multiple resamples of the data and to compare the partial first and second moments of the resampled data with the partial first and second moments of the original data to test different hypotheses.

The bootstrap method is used to test four hypotheses. The first two hypotheses test whether the partial mean of one variable is greater or less than the partial mean of another variable. The third and the last hypotheses test whether the partial second moment of one variable is less than the partial second moment of another variable.

To test each hypothesis, we generate a resample of the data and calculate the partial first and second moments of the resampled data. Then, it compares the resampled partial moments with the partial moments of the original data to determine whether the null hypothesis is rejected. This process is repeated a large number of times (3,000 in our case) to estimate the p-value of the null hypothesis. If the p-value is less than a pre-specified significance level (0.05 in our case), then the null hypothesis is rejected. We state the steps for testing H_0^+ and H_0^{2+} in the following:

1. Pick a pair of portfolios, denoted as X and Y. Transform the data to X' and Y' so that the first upper partial moments of X' and Y' become the same.
2. Resample X' and Y' to get new data X^* and Y^*.
3. Compare the difference of first upper partial moment from the new sample, $\mu_{X^*}^+ - \mu_{Y^*}^+$, with the one from the old sample, $\mu_X^+ - \mu_Y^+$. Repeat this process 3000 times.
4. Reject the null hypothesis H_0^+ if the difference of first upper partial moment from the new sample $\mu_{X^*}^+ - \mu_{Y^*}^+$ is greater than the difference of first upper partial moment from the old sample $\mu_X^+ - \mu_Y^+$ more than 95% of the time, indicating that our result is statistically significant at the 5% level.
5. Transform the data X and Y to X'' and Y'' so that the second upper partial moments of X'' and Y'' become the same.

6. Resample X'' and Y'' to get new data X^{**} and Y^{**}.

7. Compare the ratio of second upper partial moment from the new sample, $\tau^+_{X^{**}}/\tau^+_{Y^{**}}$, with the one from the old sample, τ^+_X/τ^+_Y. Repeat this process 3000 times.

8. If H_0^+ is not rejected, then reject the null hypothesis H_0^{2+} if the ratio of partial second upper moment from the new sample $\frac{\tau^+_{X^{**}}}{\tau^+_{Y^{**}}}$ is less than that from the old sample $\frac{\tau^+_X}{\tau^+_Y}$ more than 95% of the time, again indicating statistical significance at the 5% level.

The procedure for testing H_0^- and H_0^{2-} mirrors the one described above.

4.9 References

Abdellaoui, M., Bleichrodt, H., and Kammoun, H. (2013). Do financial professionals behave according to prospect theory? an experimental study. *Theory and Decision*, 74:411–429.

Anthonisz, S. A. (2012). Asset pricing with partial-moments. *Journal of Banking & Finance*, 36(7):2122–2135.

Barberis, N., Huang, M., and Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics*, 116(1):1–53.

Barberis, N., Jin, L. J., and Wang, B. (2021). Prospect theory and stock market anomalies. *The Journal of Finance*, 76(5):2639–2687.

Baucells, M. and Heukamp, F. H. (2006). Stochastic dominance and cumulative prospect theory. *Management Science*, 52(9):1409–1423.

Bawa, V. S., Bodurtha Jr, J. N., Rao, M., and Suri, H. L. (1985). On determination of stochastic dominance optimal sets. *The Journal of Finance*, 40(2):417–431.

Berkelaar, A. B., Kouwenberg, R., and Post, T. (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics*, 86(4):973–987.

Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to linear optimization*, volume 6. Athena scientific Belmont, MA.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Chamberlain, G. (1983). A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory*, 29(1):185–201.

Chambers, R. G., Grant, S., Polak, B., and Quiggin, J. (2014). A two-parameter model of dispersion aversion. *Journal of Economic Theory*, 150:611–641.

Chan, R. H., Chow, S. C., Guo, X., and Wong, W. K. (2022). Central moments, stochastic dominance, moment rule, and diversification with an application. *Chaos, Solitons & Fractals*, 161:112251.

Cho, Y.-H., Linton, O., and Whang, Y.-J. (2007). Are there monday effects in stock returns: A stochastic dominance approach. *Journal of Empirical Finance*, 14(5):736–755.

Chui, D., Cheng, W. W., Chow, S. C., and Ya, L. (2020). Eastern halloween effect: A stochastic dominance approach. *Journal of International Financial Markets, Institutions and Money*, 68:101241.

de Castro, L. and Galvao, A. F. (2019). Dynamic quantile models of rational behavior. *Econometrica*, 87(6):1893–1939.

Dentcheva, D. and Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566.

Dentcheva, D. and Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30(2):433–451.

Duxbury, D. and Summers, B. (2004). Financial risk perception: Are individuals variance averse or loss averse? *Economics Letters*, 84(1):21–28.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Ehrgott, M. (2005). *Multicriteria optimization*, volume 491. Springer Science & Business Media.

Fang, K. W. T., Kotz, S., and Ng, K. W. (2018). *Symmetric multivariate and related distributions*. CRC Press.

Fang, Y. and Post, T. (2022). Optimal portfolio choice for higher-order risk averters. *Journal of Banking & Finance*, 137:106429.

Fiegenbaum, A. (1990). Prospect theory and the risk-return association: An empirical examination in 85 industries. *Journal of Economic Behavior & Organization*, 14(2):187–203.

Frahm, G. (2004). *Generalized elliptical distributions: theory and applications*. PhD thesis, Universität zu Köln.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, 61(4):2017–2046.

Friedman, M. and Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, 56(4):279–304.

Guastaroba, G., Mansini, R., Ogryczak, W., and Speranza, M. G. (2016). Linear programming models based on omega ratio for the enhanced index tracking problem. *European Journal of Operational Research*, 251(3):938–956.

Gupta, A. K., Varga, T., and Bodnar, T. (2013). *Elliptically contoured models in statistics and portfolio theory*. Springer.

Hanoch, G. and Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, 36(3):335–346.

He, X. D. and Zhou, X. Y. (2011). Portfolio choice under cumulative prospect theory: An analytical treatment. *Management Science*, 57(2):315–331.

Holthausen, D. M. (1981). A risk-return model with risk and return measured as deviations from a target return. *The American Economic Review*, 71(1):182–188.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–291.

Keating, C. and Shadwick, W. F. (2002). A universal performance measure. *Journal of Performance Measurement*, 6(3):59–84.

Landsman, Z., Makov, U., and Shushi, T. (2018). A multivariate tail covariance measure for elliptical distributions. *Insurance: Mathematics and Economics*, 81:27–35.

Levy, H. (2011). *The capital asset pricing model in the 21st century: Analytical, empirical, and behavioral perspectives*. Cambridge University Press.

Levy, H., De Giorgi, E. G., and Hens, T. (2012). Two paradigms and nobel prizes in economics: a contradiction or coexistence? *European Financial Management*, 18(2):163–182.

Levy, H. and Levy, M. (2004). Prospect theory and mean-variance analysis. *Review of Financial Studies*, 17(4):1015–1041.

Levy, H. and Levy, M. (2021). Prospect theory, constant relative risk aversion, and the investment horizon. *PloS one*, 16(4):e0248904.

Levy, H. and Wiener, Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. *Journal of Risk and Uncertainty*, 16(2):147–163.

Linton, O., Maasoumi, E., and Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *The Review of Economic Studies*, 72(3):735–765.

Liu, Z. and O'Neill, M. J. (2018). Partial moment volatility indices. *Accounting & Finance*, 58(1):195–215.

Ljungqvist, A. and Wilhelm Jr, W. J. (2005). Does prospect theory explain ipo market behavior? *The Journal of Finance*, 60(4):1759–1790.

Manski, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision*, 25:79–104.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. Wiley.

Ogryczak, W. and Ruszczyński, A. (1999). From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research*, 116(1):33–50.

Owen, J. and Rabinovitch, R. (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *The Journal of Finance*, 38(3):745–752.

Post, T., Karabatı, S., and Arvanitis, S. (2018). Portfolio optimization based on stochastic dominance and empirical likelihood. *Journal of Econometrics*, 206(1):167–186.

Post, T. and Kopa, M. (2017). Portfolio choice based on third-degree stochastic dominance. *Management Science*, 63(10):3381–3392.

Post, T. and Levy, H. (2005). Does risk seeking drive stock prices? a stochastic dominance analysis of aggregate investor preferences and beliefs. *The Review of Financial Studies*, 18(3):925–953.

Richardson, M. and Smith, T. (1993). A test for multivariate normality in stock returns. *Journal of Business*, pages 295–321.

Schuhmacher, F., Kohrs, H., and Auer, B. R. (2021). Justifying mean-variance portfolio selection when asset returns are skewed. *Management Science*, 67(12):7812–7824.

Sinha, T. (1994). Prospect theory and the risk return association: another look. *Journal of Economic Behavior & Organization*, 24(2):225–231.

Tesfatsion, L. (1976). Stochastic dominance and the maximization of expected utility. *The Review of Economic Studies*, 43(2):301–315.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Weber, M. and Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, 33(2):167–184.

Wong, W.-K. and Chan, R. H. (2008). Prospect and markowitz stochastic dominance. *Annals of Finance*, 4(1):105–129.

Xiao, Y. and Valdez, E. A. (2015). A black–litterman asset allocation model under elliptical distributions. *Quantitative Finance*, 15(3):509–519.

Zakamouline, V. and Koekebakker, S. (2009). A generalisation of the mean-variance analysis. *European Financial Management*, 15(5):934–970.

Chapter 5

Portfolio Optimization for Most Investors with Loss Aversion

This paper explores the significant role of loss aversion, a psychological concept embedded in prospect theory. Traditional portfolio theories often fail to consider the impact of loss aversion on investment decisions, leading to an inadequate reflection of an investor's behavioral pattern. To address this, we introduce a novel approach for integrating loss aversion into portfolio optimization. Our method embraces an extended version of Prospect Stochastic Dominance (PSD), enhanced by the inclusion of Loss Aversion, which we term Prospect Stochastic Dominance with Loss Aversion (PSD-L). Analysis based on PSD-L could allow us to accommodate a broader spectrum of investors who employ various value functions with a loss aversion feature under prospect theory. We then employ the application of partial moments and show that partial moments as performance tools are consistent with PSD-L. Using partial moments, we also provide a portfolio optimization method for constructing a portfolio that dominates a target investment in the PSD-L sense. By doing so, we establish a method for constructing portfolios that are more consistent with investor behavioral tendencies, which can give us a more nuanced understanding of investor behavior and enhance the efficiency of investment decisions.

5.1 Overview

Loss aversion, originating from prospect theory in Kahneman and Tversky (1979) and Tversky and Kahneman (1992), refers to the psychological phenomenon where individuals feel the pain of a loss more than the pleasure of a comparable gain. In finance, loss aversion is instrumental in understanding several market behaviors and anomalies. For instance, it explains the endowment effect proposed in Thaler (1980), signifying higher valuation of owned objects compared to their acquisition cost. Furthermore, loss aversion affects individuals' risk management approaches, influencing investment strategies based on their loss sensitivity.

Incorporating loss aversion into portfolio optimization aligns investment strategies with investor behavioral tendencies. Traditional portfolio theories often address loss aversion inadequately, leading to an incomplete understanding of how investors feel the sting of losses more acutely than they enjoy equivalent gains. Acknowledging loss aversion provides a new perspective to understand investment decisions. This perspective can assist in separating the influences of various components of prospect theory if we aim to examine its implications. Furthermore, it will also provide tools to enhance the efficiency of investment decisions.

In this study, we establish the application of partial moments as consistent performance measurement tools within the context of prospect theory with an emphasis on loss aversion. We generalized the framework of "Prospect Stochastic Dominance with Loss Aversion" (PSD-L) proposed in Baucells and Heukamp (2006), which extended from the Prospect Stochastic Dominance (PSD) proposed by Levy and Wiener (1998). This method allows our findings to be relevant for a broad spectrum of investors who seemingly employ varied value functions under the prospect theory.

Through the use of partial moments, we introduce a methodology to construct a portfolio that holds dominance over a target investment in terms of PSD-L, accomplished via the portfolio optimization. Numerous studies, such as Berkelaar et al. (2004) and Jarrow and Zhao (2006), have explored portfolio optimization under prospect theory and specifically investigated optimal investment strategies for loss-averse investors. He and Zhou (2011), on the other hand, offered an analytical approach to a single-period portfolio choice model for agents exhibiting features recommended by prospect theory. While these contributions are valuable, they often rely on a specified utility function to derive their findings, leaving a gap in understanding whether a single functional form can accurately portray a diverse group of investors. This points to the necessity of a method for formulating an appropriate portfolio for a class of investors.

The remainder of the paper is organized as follows: in Section 5.2, we discuss the concept of PSD-L; in Section 5.3, we present the relationships between PSD-L and partial moments; in Section 5.4, we demonstrate how to incorporate partial moments into the portfolio optimization problem; and in Section 5.5, we conclude.

5.2 Prospect Stochastic Dominance and Loss Aversion

Let X and Y be two random variables (or risky investments) with continuous *cumulative distribution functions* (CDFs), F and G, and *probability density functions* (PDFs) f and g. The means of X and Y are denoted by μ_X and μ_Y, respectively. Similarly, the variances of X and Y are represented as σ_X^2 and σ_Y^2, respectively. In this chapter, we assume that both $E[X^2]$ and $E[Y^2]$ are finite. From this point on, we will refer to the value function as the utility function. We will define several classes of utility functions in the subsequent sections.

Definition 5.1 (S-shaped utility function)**.** Given CDFs F and G. Let U_2^A, U_2^D, U^S, and $U^S(F,G)$ be the sets of the utility functions u such that:

$$
\begin{aligned}
U_2^A &= \Big\{u: \ (-1)^{i+1}u^{(i)} \geq \ 0 \ , \ i=1,2\Big\}, \\
U_2^D &= \Big\{u: \ u^{(i)} \geq \ 0 \ , \ i=1,2\Big\}, \\
U^S &= \Big\{u: \ \big(u^{+} \in U_2^A\big) \quad \text{and} \quad \big(u^{-} \in U_2^D\big)\Big\}, \\
U^S(F,G) &= \Big\{u: \ \big(u \in U^S\big) \quad \text{and} \quad \big(\textstyle\int udF \ \& \ \int udG \text{ are finite}\big)\Big\},
\end{aligned}
$$

where $u^{(i)}$ is the i^{th} derivative of the utility function u, $u^{+} = u$ restricted for $x > 0$, and $u^{-} = u$ restricted for $x < 0$.

It is noted that investors with a utility function in U_2^A are risk-averse, while those with a utility function in U_2^D are risk-seeking. Investors with a utility function in U^S are risk-averse on gains but risk-seeking on losses.

For loss aversion, a utility function displaying this feature will typically have a kink at the origin (representing the reference point or status quo) and be steeper for losses than for gains.

This implies that the marginal disutility of a loss is greater than the marginal utility of a gain of the same size. Loss aversion is characterized by the condition $u(y) - u(x) \leq u(-x) - u(-y)$ for all $y \leq x \leq 0$ in Wakker and Tversky (1993), which is equivalent to $u'(-x) \geq u'(x)$ for all $x \geq 0$ in our setting. The related classes of utility functions will be defined in the subsequent sections.

Definition 5.2. *Let U^L and $U^{SL}(F,G)$ be the sets of the utility functions u such that:*

$$
\begin{aligned}
U^L &= \Big\{u: \ u'(-x) \geq u'(x) \quad \textit{for all} \quad x \geq 0\Big\} \\
U^{SL} &= \Big\{u: \ \big(u \in U^L\big) \quad \textit{and} \quad \big(u \in U^S\big)\Big\} \\
U^{SL}(F,G) &= \Big\{u: \ \big(u \in U^{SL}\big) \quad \textit{and} \quad \big(\textstyle\int udF \ \& \ \int udG \ \textit{are finite}\big)\Big\}.
\end{aligned}
$$

Note that U^L contains all utility functions with loss aversion feature. With the assumption that the expected utility is well defined for the CDFs under consideration, the utility functions we concern belong to set $U^{SL}(F,G)$. Accordingly, for agent with behaviour consistent with utility function u, F is preferred to G if

$$
E_F u(x) - E_G u(x) = \int u(x)dF(x) - \int u(x)dG(x) \geq 0.
$$

In order to compare risky investments for a wide class of investors who behave as if they are using expected utility with S-shaped utility function with loss aversion feature, we employ the concept of PSD-L proposed by Baucells and Heukamp (2006), which extended from the PSD proposed by Levy and Wiener (1998). PSD-L allows us to compare the expected utility of different choices by only comparing the CDFs of the corresponding random variables. This allows us to make conclusions about the relative utility of different options without specifying a particular utility function. PSD is first expressed in the following definition:

Definition 5.3 (Prospect Stochastic Dominance)**.** Given CDFs F and G, F dominates G by PSD (denoted by $F \succeq_P G$) if $\int_{x_1}^{x_2} [G(z) - F(z)]\, dz \geq 0$ for all $x_1 \leq 0 \leq x_2$.

The relevance of PSD in the analysis of risk-associated choices becomes apparent when examining diverse investment comparisons among a broad category of investors. These investors exhibit behaviors as if they are using different value functions under prospect theory. In specific, it can be prove that $F \succeq_P G$ if and only if $E_F u(x) \geq E_G u(x)$ for all $u \in U^S(F,G)$.

Baucells and Heukamp (2006) extend the idea of PSD to incorporate the effect of loss aversion and probability weighting in the stochastic dominance condition. For PSD concept that incorporates the effect of loss aversion, the PSD-L, we defined it as follows.

Definition 5.4 (PSD-L). F dominates G by PSD-L (denoted by $F \succeq_{PL} G$) if and only if

$$\int_{-x}^{0} [G(z) - F(z)]dz \geq \max\left\{0, \int_{0}^{x} [F(z) - G(z)]dz\right\} \quad \text{for all} \quad x \geq 0.$$[1]

For PSD-L to establish dominance of F over G, the integrated difference from zero to any amount of loss $-x$ should be at least as large as the greater value between 0 and the integrated difference from zero to an equivalent gain x. The distinction between PSD and PSD-L arises from the understanding that decision-makers might place greater weight on losses than on equivalent gains, a phenomenon deeply rooted in the foundations of prospect theory.

What truly sets PSD-L apart from PSD, and standard second-order stochastic dominance is its emphasis on the loss aversion principle. By accounting for the likelihood that decision-makers assign more weight to losses than to equivalent gains, PSD-L offers a tool that facilitates a more detailed analysis. This can be particularly relevant in situations where negative outcomes (or losses) have an outsized impact on decision-making processes.

Note that Definition 5.4, is just a more compact way to express conditions in Definition 5 in Baucells and Heukamp (2006) without probability weighting functions. Baucells and Heukamp (2006) then show that for any CDFs F^* and G^* with bounded domain, $F^* \succeq_{PL} G^*$ implies $E_{F^*}u(x) \geq E_{G^*}u(x)$, for all $u \in U^{SL}(F, G)$ on Proposition 7 in Baucells and Heukamp (2006). We formally extend this relationship for CDFs with unbounded domain as follows:

Theorem 5.5. *Given CDFs F and G, then $F \succeq_{PL} G$ if and only if $E_F u(x) \geq E_G u(x)$ for all $u \in U^{SL}(F, G)$.*

Namely, $F \succeq_{PL} G$ implies that F is preferred to G for any utility function in $U^{SL}(F, G)$ and the converse also holds.

5.3 Implications on moments

In this section, we delve into the exploration of the intricate relationships between PSD-L and moments. The derived results bear relevance not only for the analysis carried out in subsequent sections but also hold significant potential to enhance efficiency in evaluating PSD-L relationships between various distributions. The presented relationships between PSD-L and moments provide valuable tools for simplifying the process of pairwise comparisons when there is a need to discern PSD-L relationships across multiple distributions. This is of substantial relevance given the computational burdens associated with comparing the integrals of distributions for numerous outcomes, akin to the process involved in PSD, as illustrated in Linton et al. (2005). Moreover, these conditions play an essential role when dealing with scenarios where information on distributions is limited. These conditions are also useful if we only have limited information on distributions; for example, we only have information about partial moments. This paper will label existing results as propositions and new results as lemmas or theorems.

Consider the mean of a random variable $Z = X$ and Y with CDF $H = F$ and G. The expected value of these random variables can be dissected into two distinct components, often referred to as the first partial moments as follows.

$$\mu_Z = \mu_Z^+ - \mu_Z^- \quad \text{such that} \quad \mu_Z^+ := \int_0^\infty x dH(x) \quad \text{and} \quad \mu_Z^- := \int_{-\infty}^0 -x dH(x), \tag{A.1}$$

where μ_Z^+ is the upper first partial moment (or measure of expected gain) and μ_Z^- is the lower first partial moment (or measure of expected loss).

It turns out that we can derive the relationship between PSD-L and the first partial moment for the general distributions. Before we derive the general relationship between PSD-L and the first partial moment, we state the following useful Property by Chow (2023):

Property 5.3.1. For random variables X and Y, the following statements are true:

1. $\mu_X^+ - \mu_Y^+ = \int_0^\infty [G(x) - F(x)]dx$.

2. $\mu_Y^- - \mu_X^- = \int_{-\infty}^0 [G(x) - F(x)]dx$.

The first part states that the difference between the first upper partial moments associated with the CDFs F and G, $\mu_X^+ - \mu_Y^+$, is equal to the integral of the difference between the CDFs,

$G(x) - F(x)$, over the interval $[0, \infty)$. The second part states that the difference between the first lower partial moments associated with the CDFs F and G, $\mu_Y^- - \mu_X^-$, is equal to the integral of the difference between the CDFs, $G(x) - F(x)$, over the interval $(-\infty, 0]$. Using this Property, necessary conditions for PSD that involve first partial moments are given in Theorem 5.6.

Theorem 5.6. *For random variables X and Y such that $F \succeq_{PL} G$, the following are true:*

1. $\mu_Y^- - \mu_X^- \geq \max\left\{0, \mu_Y^+ - \mu_X^+\right\}$.
2. $\mu_X \geq \mu_Y$.

The theorem states that if F dominates G in PSD-L sense, then the difference between the expected losses of G and F is at least as large as the greater of zero and the difference between the expected gains of G and F. In practical terms, this theorem suggests that if the anticipated gain for investment Y supersedes that of investment X, it is implausible for all investors (especially those possessing an S-shaped utility function with loss aversion characteristics) to prefer X over Y, unless the expected loss for investment X is less than that of investment Y, and the difference, $\mu_Y^- - \mu_X^-$, exceeds $\mu_Y^+ - \mu_X^+$. The theorem also states that if F dominates G in PSD-L sense, it is necessary for investment X to offer a higher expected return compared to investment Y.

For the risk analysis tool, the second moment of distribution $H = F$ and G with the corresponding random variable $Z = X$ and Y (denoted by τ_Z) could also be partitioned into two second-partial moments (or semi-variances from zero); that is,

$$\tau_Z = \tau_Z^+ + \tau_Z^- \quad \text{such that} \quad \tau_Z^+ := \int_0^\infty x^2 dH(x) \quad \text{and} \quad \tau_Z^- := \int_{-\infty}^0 x^2 dH(x), \tag{A.2}$$

where τ_Z^+ is the second upper partial moment and τ_Z^- is the second lower partial moment. The upper second partial moment and second lower partial moment are special cases of semivariance. The semivariance as a measurement tool is first proposed in Markowitz (1959). The relationship between it and concave utility functions (or second-order stochastic dominance) is further investigated by Ogryczak and Ruszczyński (1999). Before we go to the relationship between PSD-L and second partial moments, we state the following useful Property by Chow (2023):

Property 5.3.2. For random variables X and Y, the following statements are true:

1. $\tau_Y^+ - \tau_X^+ = 2\int_0^\infty \left(\int_0^x [G(t) - F(t)]dt + \mu_Y^+ - \mu_X^+ \right) dx.$
2. $\tau_Y^- - \tau_X^- = 2\int_{-\infty}^0 \left(\mu_Y^- - \mu_X^- - \int_x^0 [G(t) - F(t)]dt \right) dx.$

The first part of the Property states that the difference between the second upper partial moments associated with the CDFs F and G, $\tau_Y^+ - \tau_X^+$, is equal to twice the integral of the sum of the inner integral, $\int_0^x [G(t) - F(t)]dt$, and the difference between the first upper partial moments, $\mu_Y^+ - \mu_X^+$, over the interval $[0, \infty)$. The second part states that the difference between the second lower partial moments associated with the CDFs F and G, $\tau_Y^- - \tau_X^-$, is equal to twice the integral of the difference between the first lower partial moments, $\mu_Y^- - \mu_X^-$, and the inner integral, $\int_x^0 [G(t) - F(t)]dt$, over the interval $(-\infty, 0]$. Using this Property, we derive the necessary conditions for the PSD that involve the first and second partial moments and display the results in the following theorem:

Theorem 5.7. *For random variables X and Y, if $F \succeq_{PL} G$, and $\mu_X^- = \mu_Y^-$, then $\tau_X^- \geq \tau_Y^-$.*

Theorem 5.7 states that if F dominates G in the sense of the PSD-L and the first lower partial moment of two distributions are the same, then it must be the case that the lower second partial moment of F is on top of the lower second partial moment of G.

Instead of using zero as the target return, we can also use any amount x as the target return to define first lower partial moment $\mathcal{L}_Z(x)$ and first upper partial moment $\mathcal{U}_Z(x)$, for $H = F$ and G, $Z = X$ and Y, as shown in the following:

$$\mathcal{L}_Z(x) := E[(x - Z)_+] = \int_{-\infty}^x (x - z)dH(z) \ ,$$

$$\mathcal{U}_Z(x) := E[(Z - x)_+] = \int_x^\infty (z - x)dH(z) \ ,$$

where $(x)_+ = \max\{x, 0\}$. Noted that $\mathcal{L}_Z(x)$ could be considered as expected return below target x and $\mathcal{U}_Z(x)$ could be considered as expected return above target x. The relationship between these partial moments and cumulative distribution has been well studied in Bawa et al. (1985) and Ogryczak and Ruszczyński (1999). Some related useful properties by Chow (2023) are stated as follows:

Property 5.3.3. For any pair of random variables X and Y with CDFs F and G, the following statements are true:

1. $\mathcal{L}_Y(x) - \mathcal{L}_X(x) = \int_{-\infty}^{x}[G(z) - F(z)]dz.$
2. $\mathcal{U}_X(x) - \mathcal{U}_Y(x) = \int_{x}^{\infty}[G(z) - F(z)]dz.$
3. $\big[\mathcal{L}_Y(0) - \mathcal{L}_X(0)\big] - \big[\mathcal{L}_Y(x) - \mathcal{L}_X(x)\big] = \int_{x}^{0}[G(z) - F(z)]dz, \quad \forall x \leq 0.$
4. $\big[\mathcal{U}_X(0) - \mathcal{U}_Y(0)\big] - \big[\mathcal{U}_X(x) - \mathcal{U}_Y(x)\big] = \int_{0}^{x}[G(z) - F(z)]dz, \quad \forall x \geq 0.$

The first part of Property 5.3.3 states that the difference between the first lower partial moments, $\mathcal{L}_Y(x) - \mathcal{L}_X(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $(-\infty, x]$. The second part states that the difference between the first upper partial moments, $\mathcal{U}_X(x) - \mathcal{U}_Y(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[x, \infty)$. The third part states that the difference between the first lower partial moments evaluated at 0, $\mathcal{L}_Y(0) - \mathcal{L}_X(0)$, minus the difference between the first lower partial moments evaluated at x, $\mathcal{L}_Y(x) - \mathcal{L}_X(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[x, 0]$ for all $x \leq 0$. The fourth part states that the difference between the first upper partial moments evaluated at 0, $\mathcal{U}_X(0) - \mathcal{U}_Y(0)$, minus the difference between the first upper partial moments evaluated at x, $\mathcal{U}_X(x) - \mathcal{U}_Y(x)$, is equal to the integral of the difference between the CDFs, $G(z) - F(z)$, over the interval $[0, x]$ for all $x \geq 0$. Combining Definition 4.2 and Lemma 5.3.3, we obtain the following corollary:

Corollary 5.8 (Safety overrides aggressiveness). *$F \succeq_{PL} G$ if and only if*

$$\begin{aligned} & \big[\mathcal{L}_Y(0) - \mathcal{L}_Y(-x)\big] - \big[\mathcal{L}_X(0) - \mathcal{L}_X(-x)\big] \\ \geq \quad & \max\Big\{0, \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(x)\big] - \big[\mathcal{U}_X(0) - \mathcal{U}_X(x)\big]\Big\} \end{aligned} \tag{A.3}$$

for all $x \geq 0$.

In essence, this corollary can be interpreted as "safety overrides aggressiveness", as it implies that when comparing two portfolios, "safety" is important since $\big[\mathcal{L}_Y(0) - \mathcal{L}_Y(-x)\big] - \big[\mathcal{L}_X(0) - \mathcal{L}_X(-x)\big] \geq 0$, and downside risk is more important compared to upside potential, that is $\big[\mathcal{L}_Y(0) - \mathcal{L}_Y(-x)\big] - \big[\mathcal{L}_X(0) - \mathcal{L}_X(-x)\big] \geq \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(x)\big] - \big[\mathcal{U}_X(0) - \mathcal{U}_X(x)\big]$. This reflects the behavior of loss-averse investors, who prioritize avoiding losses over obtaining equivalent gains, a key aspect of prospect theory.

5.4 Portfolio Choice based on PSD-L

Without specific assumptions regarding the distribution, identifying PSD-L-efficient portfolios can be quite a complex task. This complexity arises due to $\succeq_{PL}$ is a partial order, which typically hinders the use of a real value function for ranking the available choices. Consequently, we cannot leverage traditional optimization techniques to pinpoint portfolios bearing the highest "PSD-L rank". In this section, our focus will shift towards a distinct portfolio optimization problem that can integrate the concept of PSD-L.

Let $R_1, ..., R_N$ be random returns for assets $1, ..., N$, with joint cumulative distribution function $\mathcal{F}(\mathbf{r})$ where $\mathbf{r} \in R_1 \times ... \times R_N$. Assume that $E[{R_n}^2] < \infty$ for all $n = 1, ..., N$, and further let $w_1, ..., w_N$ as portfolio weights on assets $1, ..., N$. Without short selling, the set of possible asset allocations is defined as:

$$W = \Big\{ w \in \mathbb{R}^N : \sum_{i=1}^{N} w_i = 1, w_n \geq 0 \;\text{ for }\; n = 1, ..., N \Big\}.$$

The marginal CDF $\mathcal{F}_w(\eta)$ for asset allocation $w \in W$ is defined as:

$$\mathcal{F}_w(\eta) = \int_{\{w'\mathbf{r} \leq \eta\}} d\mathcal{F}(\mathbf{r})$$

where the corresponding portfolio return is denoted as $R_w = R_1 w_1 + ... + R_N w_N$. Letting Y as target random return with finite expected value and corresponding CDF G. Consider the distribution G as a benchmark. A compelling inquiry arises: Given a group agents with preferences shaped by an S-shaped utility function with loss aversion feature, how might we identify a portfolio that outperforms this benchmark? Under these settings, we can form the following problem:

$$\begin{aligned} &\max E[R_w] \\ \text{subject to } \; &\mathcal{F}_w \succeq_{PL} G, \\ &w \in W. \end{aligned} \tag{A.4}$$

where $E[R_w]$ is the expected portfolio return. A similar problem for risk averters (based on SSD) is first consider in Dentcheva and Ruszczyński (2003) and further analyzed in Post and Kopa (2017), Post et al. (2018), and Fang and Post (2022). The problem here is saying that given the expected utility of choosing portfolio random return R_w is higher than the target random return Y for all utility functions $u \in U^{SL}(F, G)$, we maximize the objective function $E[R_w]$. Since $\succeq_P$ is a partial order, this denotes the absence of a single objective function to represent it in general. Typically, optimization under these conditions presents substantial challenges. Nevertheless, the proposed formulation circumvents these inherent challenges, introducing a viable avenue for optimization. For the above problem, it is not clear how the assets allocations change will affect the cumulative distribution function of the corresponding portfolio return R_w. In order to solve the above problem, we use Corollary 5.8 and transform problem (A.4) to an equivalent problem in below.

$$
\begin{aligned}
&\max\ E[R_w] \\
&\text{subject to } \big[\mathcal{L}_Y(0) - \mathcal{L}_Y(-\eta)\big] - \big[\mathcal{L}_{R_w}(0) - \mathcal{L}_{R_w}(-\eta)\big] \\
&\qquad\qquad \geq \max\Big\{0, \big[\mathcal{U}_Y(0) - \mathcal{U}_Y(\eta)\big] - \big[\mathcal{U}_{R_w}(0) - \mathcal{U}_{R_w}(\eta)\big]\Big\}, \quad \forall \eta \geq 0, \\
&\qquad\qquad w \in W.
\end{aligned} \tag{A.5}
$$

Consider that we have finite many states $t = 1, ..., T$, let r_{nt} denotes the return of asset n in period t, p_t as probabilities of these realizations where $p_t = 1/T$. Assume random variables under concerns are uniformed bounded then $\mathcal{L}_Z(\eta)$ could be express as follows:

$$\mathcal{L}_Z(\eta) = \sum_{t=1}^{T} p_t S_t(\eta)$$

with variables $S_t(\eta)$ and corresponding binary variables b_t that satisfies the following inequalities:

$$
\begin{aligned}
&S_t(\eta) \geq \eta - \sum_{n}^{N} w_n r_{nt}, && t = 1, ..., T,\\
&S_t(\eta) \leq \eta - \sum_{n}^{N} w_n r_{nt} - K_L b_t(\eta), && t = 1, ..., T,\\
&S_t(\eta) \geq 0, && t = 1, ..., T,\\
&S_t(\eta) \leq K_U(1 - b_t(\eta)), && t = 1, ..., T,\\
&b_t(\eta) \in \{0, 1\}, && t = 1, ..., T.
\end{aligned}
\tag{A.6}
$$

where $K_L = \min\{y_t, r_{nt}\} - \max\{y_t, r_{nt}\}$ and $K_U = \max\{y_t, r_{nt}\} - \min\{y_t, r_{nt}\}$.

For $\mathcal{U}_H(\eta)$, it could be defined in a similar way by replacing $\eta - \sum_n^N w_n r_{nt}$ with $\sum_n^N w_n r_{nt} - \eta$. We also replace $S_t(\eta)$ with $D_t(\eta)$ and $b_t(\eta)$ with $c_t(\eta)$ to denote new decision variables. The above formation make use of the standard technique on handling max operator in mixed integer linear programming and linear programing. It is common in the literature use idea related to (A.6) to solve portfolio optimization problem with second or higher order stochastic dominance constraint, including Dentcheva and Ruszczyński (2003), Dentcheva and Ruszczyński (2006), Post et al. (2018), and Fang and Post (2022). Using these concepts, one can transform the orgrinal problem into the following.

$$
\begin{aligned}
&\max\ E[R_w] \\
&\text{subject to } \big[\mathcal{L}_Y(0)-\mathcal{L}_Y(-\eta)\big]-\big[E[S_t(0)]-E[S_t(-\eta)]\big] \\
&\qquad \geq \big[\mathcal{U}_Y(0)-\mathcal{U}_Y(\eta)\big]-\big[E[D_t(0)]-E[D_t(\eta)]\big], && \forall\eta\geq 0, \\
&\qquad \big[\mathcal{L}_Y(0)-\mathcal{L}_Y(-\eta)\big]-\big[E[S_t(0)]-E[S_t(-\eta)]\big]\geq 0, && \forall\eta\geq 0, \\
&\qquad S_t(\eta)\geq \eta-\sum_n^N w_n r_{nt}, && \forall\eta\leq 0, t=1,...,T, \\
&\qquad S_t(\eta)\leq \eta-\sum_n^N w_n r_{nt}-K_L b_t(\eta), && \forall\eta\leq 0, t=1,...,T, \\
&\qquad S_t(\eta)\leq K_U(1-b_t(\eta)), && \forall\eta\leq 0, t=1,...,T, \\
&\qquad D_t(\eta)\geq \sum_n^N w_n r_{nt}-\eta, && \forall\eta\geq 0, t=1,...,T, \\
&\qquad D_t(\eta)\leq \sum_{n=1}^N w_n r_{nt}-\eta-K_L c_t(\eta), && \forall\eta\geq 0, t=1,...,T, \\
&\qquad D_t(\eta)\leq K_U(1-c_t(\eta)), && \forall\eta\geq 0, t=1,...,T, \\
&\qquad S_t(-\eta), D_t(\eta)\geq 0, && \forall\eta\geq 0, t=1,...,T, \\
&\qquad b_t(-\eta), c_t(\eta)\in\{0,1\}, && \forall\eta\geq 0, t=1,...,T. \\
&\qquad w\in W.
\end{aligned}
\tag{A.7}
$$

One can solve the above problem with a standard mixed integer linear programming solver and the solution of the above problem will be an approximate solution for the original problem.

5.5 Concluding remarks

This paper has highlighted the profound influence of loss aversion, a psychological principle deeply embedded within the prospect theory, on investors' decision-making processes. Traditional portfolio theories, which are commonly employed, often neglect the ramifications of loss aversion, leading to a deficiency in accurately mirroring an investor's behavioral patterns. To address this shortfall, we have developed an innovative approach to incorporate loss aversion into portfolio optimization.

Our methodology employs an enhanced version of PSD, supplemented with the aspect of Loss Aversion. This is hereby referred to as Prospect Stochastic Dominance with Loss Aversion (PSD-L). By utilizing the PSD-L framework, our analysis caters to a wide range of investors who utilize distinct value functions under the prospect theory, with a particular emphasis on the feature of loss aversion.

Key to our approach is the application of partial moments as consistent performance measurement tools aligned with the PSD-L framework. The use of partial moments offers a robust methodology for constructing a portfolio that exhibits dominance over a target investment in the PSD-L sense, achieved through meticulous portfolio optimization.

This work provides a refined understanding of investor behavior by considering the pervasive influence of loss aversion. It further establishes a methodology for devising portfolios that more accurately represent investor behavioral tendencies. Our findings have significant implications for enhancing the efficiency of investment decision-making processes, bridging the existing gap in understanding whether a single functional form can accurately portray a diverse group of investors.

In conclusion, our exploration into the intricate interplay of loss aversion, prospect theory, and portfolio optimization highlights the pivotal role of psychology in finance. It pushes the boundaries of traditional portfolio theories, offering a more comprehensive framework that accommodates the behavioral tendencies of investors, ultimately contributing to more effective and tailored investment strategies.

5.6 Appendix: Proofs of Lemmas and Theorems

Proof of Theorem 5.5. *Sufficiency*: Step 1. Let $E = \{x : G(x) - F(x) \geq 0\}$, and $E^c = \{x : G(x) - F(x) < 0\}$, with their corresponding indicator functions $I_E(x)$ and $I_{E^c}(x)$. Define a function $Q : \mathbb{R}^- \to \mathbb{R}^-$ as

$$Q(x) = \sup\left\{x' \leq 0 \;\middle|\; \int_{x'}^{0} [G(z) - F(z)] I_E(z)\, dz = \int_{x}^{0} [F(z) - G(z)] I_{E^c}(z) dz\right\}$$

The exists of x' for each x is guarantee from the fact that $\int_{x'}^{0}[G(z) - F(z)]I_E(z)$ is a continuous function of x' and $\int_{x}^{0}[G(z) - F(z)]\, dz \geq 0$ for all $x \leq 0$. Since the supremum is unique, Q is well-defined. Note that under this definition, $F \succeq_{PL} G$ implies $Q(x) \geq x$ for all $x \leq 0$. Also note that $[F(z) - G(z)]I_{E^c}(z) \geq 0$ for all z, thus $Q(x)$ is increase with x. Since $Q(x)$ is monotone, it is also continuous and differentiable almost everywhere.

Differentiating $\int_{Q(x)}^{0}[G(z) - F(z)]I_E(z)\, dz = \int_{x}^{0}[F(z) - G(z)]I_{E^c}(z)dz$ at any point when $Q'(x)$ exists, then one can get $I_E(Q(x))[G(Q(x)) - F(Q(x))]Q'(x) = [F(x) - G(x)]I_{E^c}(x)$ holds almost everywhere over R^-. Suppose that $u \in U^{SL}(F, G)$, then for any $x \leq 0$ we have:

$$\int_{x}^{0} |G - F| du \leq \int_{x}^{0} G du + \int_{x}^{0} F du < \infty$$

by Lemma 1 in Hanoch and Levy (1969), and Lemma 1* in Tesfatsion (1976) that $\int u dF - \int u dG = \int [G - F] du$. Then $\int_{x}^{0}[G - F]du$ for any $x \leq 0$ could be expressed as:

$$\begin{aligned}
\int_{x}^{0} [G - F] du &= \int_{x}^{0} [G(z) - F(z)] I_E(z) du(z) + \int_{x}^{0} [G(z) - F(z)] I_{E^c}(z) du(z) \\
&= \int_{x}^{0} [G(z) - F(z)] I_E(z) du(z) - \int_{x}^{0} I_E(Q(z))[G(Q(z)) - F(Q(z))]Q'(z) du(z)
\end{aligned} \tag{A.8}$$

The last term of the above equation could be expressed as:

$$\begin{aligned}
\int_x^0 I_E(Q(z))[G(Q(z))-F(Q(z))]Q'(z)du(z) &= \int_x^0 I_E(Q(z))[G(Q(z))-F(Q(z))]Q'(z)u'(z)dz \\
&\leq \int_x^0 I_E(Q(z))[G(Q(z))-F(Q(z))]Q'(z)u'(Q(z))dz \\
&= \int_x^0 I_E(Q(z))[G(Q(z))-F(Q(z))]du(Q(z)) \\
&= \int_{Q(x)}^0 [G(z)-F(z)]I_E(z)du(z)
\end{aligned} \tag{A.9}$$

where the inequality holds because diminishing sensitivity $u^{''}(x) \geq 0$ and $Q(x) \geq x$ for all $x \leq 0$. Then substitute this into equation (A.8) will get us the following:

$$\begin{aligned}
\int_x^0 [G-F]du &\geq \int_x^0 [G(z)-F(z)]I_E(z)du(z) - \int_{Q(x)}^0 [G(z)-F(z)]I_E(z)du(z) \\
&= \int_x^{Q(x)} [G(z)-F(z)]I_E(z)du(z) \geq 0
\end{aligned} \tag{A.10}$$

Step 2. Let $\bar{Q}(x) = Q(-x)$ for all $x \geq 0$ then define a function $M : R^+ \to R^+$ as:

$$\begin{aligned}
M(x) = \infty\Big\{x' \geq 0 \,\Big|\, \int_{-x'}^{\bar{Q}(x)} [G(z)-F(z)]I_E(z)dz + \int_0^{x'} [G(z)-F(z)]I_E(z)dz \\
= \int_0^x [F(z)-G(z)]I_{E^c}(z)dz\Big\}
\end{aligned} \tag{A.11}$$

The exists of x' for each $x \geq 0$ is guarantee from the fact that the first two terms of the equation is a continuous function of x' and $\int_{-x}^{x}[G(z)-F(z)]dz \geq 0$ for all $x \geq 0$. Since the infimum is unique, M is well-defined. Note that under this definition, $M(x) \leq x$ for all $x \geq 0$ and $M(x)$ is increase with $x \geq 0$. Since $M(x)$ is monotone, it is also continuous and differentiable almost everywhere. To simplify notation, let $\bar{M}(x) = -M(x)$ then differentiating $\int_{\bar{M}(x)}^{\bar{Q}(x)}[G(z)-F(z)]I_E(z)dz + \int_0^{M(x)}[G(z)-F(z)]I_E(z)dz = \int_0^x[F(z)-G(z)]I_{E^c}(z)dz$ at x when

$M'(x)$ exists, one can get:

$$\begin{aligned}[F(x) - G(x)]I_{E^c}(x) &= I_E(\bar{Q}(x))[G(\bar{Q}(x)) - F(\bar{Q}(x))]\bar{Q}'(x) \\ &- I_E(\bar{M}(x))[G(\bar{M}(x)) - F(\bar{M}(x))]\bar{M}'(x) \\ &+ I_E(M(x))[G(M(x)) - F(M(x))]M'(x)\end{aligned} \tag{A.12}$$

which holds almost everywhere over R^+. Suppose that $u \in U^{SL}(F, G)$, then for any $x \geq 0$ using similar argument in Step 1, we have:

$$\int_0^x [G - F]du = \int_0^x [G(z) - F(z)]I_E(z)du(z) + \int_0^x [G(z) - F(z)]I_{E^c}(z)du(z) \tag{A.13}$$

Then $\int_0^x [G - F]du$ could be expressed as:

$$\begin{aligned}\int_0^x [G - F]du &= \int_0^x [G(z) - F(z)]I_E(z)u'(z)dz \\ &- \Big[\int_0^x I_E(\bar{Q}(z))[G(\bar{Q}(z)) - F(\bar{Q}(z))]\bar{Q}'(z)u'(z)dz \\ &- \int_0^x I_E(\bar{M}(z))[G(\bar{M}(z)) - F(\bar{M}(z))]\bar{M}'(z)u'(z)dz \\ &+ \int_0^x I_E(M(z))[G(M(z)) - F(M(z))]M'(z)u'(z)dz\Big]\end{aligned} \tag{A.14}$$

Let's denote the three terms within the bracket as T_1, T_2 and T_3 then they could be expressed as:

$$\begin{aligned}
T_1 - T_2 + T_3 &\leq \int_0^x I_E(\bar{Q}(z))[G(\bar{Q}(z)) - F(\bar{Q}(z))]\bar{Q}'(z)u'(\bar{Q}(z))dz \\
&\quad - \int_0^x I_E(\bar{M}(z))[G(\bar{M}(z)) - F(\bar{M}(z))]\bar{M}'(z)u'(\bar{M}(z))dz \\
&\quad + \int_0^x I_E(M(z))[G(M(z)) - F(M(z))]M'(z)u'(M(z))dz \\
&= \int_0^{\bar{Q}(x)} [G(z) - F(z)]I_E(z)du(z) - \int_0^{\bar{M}(x)} [G(z) - F(z)]I_E(z)du(z) \\
&\quad + \int_0^{M(x)} [G(z) - F(z)]I_E(z)du(z) \\
&= \int_{\bar{M}(x)}^{\bar{Q}(x)} [G(z) - F(z)]I_E(z)du(z) + \int_0^{M(x)} [G(z) - F(z)]I_E(z)du(z)
\end{aligned} \tag{A.15}$$

where the inequality holds because loss aversion $u'(-x) \geq u'(x)$, diminishing sensitivity $u''(x) \leq 0$ and $u''(-x) \geq 0$, with the fact that $0 \leq -\bar{Q}(x) \leq x$, $0 \leq -\bar{M}(x) \leq x$, and $x \geq M(x) \geq 0$ for all $x \geq 0$. Then substitute this into equation (A.14) for $x \geq 0$ will get us the following:

$$\begin{aligned}
\int_0^x [G - F]du &\geq \int_0^x [G(z) - F(z)]I_E(z)du(z) \\
&\quad - \left[\int_{-M(x)}^{Q(-x)} [G(z) - F(z)]I_E(z)du(z) + \int_0^{M(x)} [G(z) - F(z)]I_E(z)du(z) \right] \\
&= \int_{M(x)}^x [G(z) - F(z)]I_E(z)du(z) - \int_{-M(x)}^{Q(-x)} [G(z) - F(z)]I_E(z)du(z)
\end{aligned} \tag{A.16}$$

Step 3. Adding (A.10) and (A.16) together will get us the following:

$$\int_{-x}^x [G(z) - F(z)]du \;\; \geq \;\; \int_{M(x)}^x [G(z) - F(z)]I_E(z)du(z) + \int_{-x}^{-M(x)} [G(z) - F(z)]I_E(z)du(z) \tag{A.17}$$

which implies that $\int_{-x}^x [G(z) - F(z)]du \geq 0$ for $x \geq 0$ and $\int_{-\infty}^{\infty} [G(z) - F(z)]du \geq 0$.

Step 4. By Lemma 1 in Hanoch and Levy (1969), and Lemma 1* in Tesfatsion (1976), we have $\int udF - \int udG = \int [G - F]du$, and therefore $E_F u(x) \geq E_G u(x)$ for all $u \in U^{SL}(F, G)$.

Necessity: Suppose that there exists $x' \geq 0$, such that $\int_{-x'}^{0}[G(z) - F(z)]dz < 0$ or $\int_{-x'}^{x'}[G(z) - F(z)]dz < 0$. For $\int_{-x'}^{0}[G(z) - F(z)]dz < 0$, choose $u_1(x) = 0$ for $x \geq 0$, $u_1(x) = x$ for $-x' < x < 0$, and $u_1(x) = -x'$ for $x \leq -x'$. For $\int_{-x'}^{x'}[G(z)-F(z)]dz < 0$, choose $u_2(x) = x'$ for $x \geq x'$, $u_2(x) = x$ for $-x' < x < x'$, and $u_2(x) = -x'$ for $x \leq -x'$. By Lemma 1 in Hanoch and Levy (1969), and Lemma 1* in Tesfatsion (1976), we have $\int_{-\infty}^{\infty} u_1 dF - \int_{-\infty}^{\infty} u_1 dG = \int_{-x'}^{0}[G-F]du_1 < 0$ for the fist case and $\int_{-\infty}^{\infty} u_2 dF - \int_{-\infty}^{\infty} u_2 dG = \int_{-x'}^{x'}[G - F]du_2 < 0$ for the second case. □

Proof of Theorem 5.6. Part 1. By Propetry 5.3.1, we know that $\mu_X^+ - \mu_Y^+ = \int_0^{\infty}[G(x) - F(x)]dx$ and $\mu_Y^- - \mu_X^- = \int_{-\infty}^{0}[G(x) - F(x)]dx$. Since, $F \succeq_{PL} G$ means that $\int_{-x}^{0}[G(z) - F(z)]\, dz \geq \max\left\{0, -\int_0^x [G(z) - F(z)]\, dz\right\}$ for all $x > 0$, therefore $\mu_Y^- - \mu_X^- \geq \max\left\{0, \mu_Y^+ - \mu_X^-\right\}$.

Part 2. Since $\mu_Z = \mu_Z^+ - \mu_Z^-$ fo $Z = X$ and Y. Thus this Part is an immediate consequence of Part 1. □

Proof of Theorem 5.7. By Property 5.3.2, we know that if $\mu_F^- = \mu_G^-$, then $\int_x^0 [G(t) - F(t)]dt \geq 0$ for all $x \leq 0$ implies $\tau_X^- \geq \tau_Y^-$. Since, $F \succeq_{PL} G$ implies that $\int_x^0 [G(t) - F(t)]dt \geq 0$ for all $x \leq 0$. Thus, the Theorem 5.7 holds. □

5.7 References

Baucells, M. and Heukamp, F. H. (2006). Stochastic dominance and cumulative prospect theory. *Management Science*, 52(9):1409–1423.

Bawa, V. S., Bodurtha Jr, J. N., Rao, M., and Suri, H. L. (1985). On determination of stochastic dominance optimal sets. *The Journal of Finance*, 40(2):417–431.

Berkelaar, A. B., Kouwenberg, R., and Post, T. (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics*, 86(4):973–987.

Chow, N. S.-C. (2023). Revisiting portfolio optimization: The interplay of partial moments, mean-variance analysis, and prospect theory. *Mean-Variance Analysis, and Prospect Theory (May 23, 2023)*.

Dentcheva, D. and Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566.

Dentcheva, D. and Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30(2):433–451.

Fang, Y. and Post, T. (2022). Optimal portfolio choice for higher-order risk averters. *Journal of Banking & Finance*, 137:106429.

Hanoch, G. and Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, 36(3):335–346.

He, X. D. and Zhou, X. Y. (2011). Portfolio choice under cumulative prospect theory: An analytical treatment. *Management Science*, 57(2):315–331.

Jarrow, R. and Zhao, F. (2006). Downside loss aversion and portfolio management. *Management Science*, 52(4):558–566.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–291.

Levy, H. and Wiener, Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. *Journal of Risk and Uncertainty*, 16(2):147–163.

Linton, O., Maasoumi, E., and Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *The Review of Economic Studies*, 72(3):735–765.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment.* Wiley.

Ogryczak, W. and Ruszczyński, A. (1999). From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research*, 116(1):33–50.

Post, T., Karabatı, S., and Arvanitis, S. (2018). Portfolio optimization based on stochastic dominance and empirical likelihood. *Journal of Econometrics*, 206(1):167–186.

Post, T. and Kopa, M. (2017). Portfolio choice based on third-degree stochastic dominance. *Management Science*, 63(10):3381–3392.

Tesfatsion, L. (1976). Stochastic dominance and the maximization of expected utility. *The Review of Economic Studies*, 43(2):301–315.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior & organization*, 1(1):39–60.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Wakker, P. and Tversky, A. (1993). An axiomatization of cumulative prospect theory. *Journal of risk and uncertainty*, 7:147–175.

Chapter 6

Conclusion

This thesis represents an exploration into the intricacies of financial theory and investor behavior, based on the implications of Stochastic Dominance and Prospect Theory.

The initial chapters focused on extending stochastic dominance principles and their implications and applications. An innovative extension to the mean-variance rule, the moment rule, was proposed to better predict the portfolio or asset choices of both risk-averse and risk-seeking investors. The work on calendar anomalies added an essential element to our understanding of market dynamics, challenging the efficient market hypothesis and providing important insights into their potential use for trading strategies.

The core of the thesis involved a deep dive into portfolio optimization, notably accommodating concepts in prospect theory. By acknowledging the loss aversion and risk seeking aspect in investor decision-making, the research takes a significant step toward a more nuanced understanding of financial markets and investment strategies.

The development of a modified mean-variance rule that incorporates partial moments is a substantial contribution, allowing for more accurate ranking and selection of investments under various value functions within prospect theory. Further, the generalization of the Prospect Stochastic Dominance PSD) and Prospect Stochastic Dominance with Loss Aversion (PSD-L) framework is a novel contribution to the field. By using these concepts, we offer more models for optimizing portfolios that more closely align with real-world investor behavior.

Finally, the findings of this thesis have important implications for the practice of portfolio management. The tools and models developed can assist portfolio managers in devising investment strategies that not only consider expected returns and variance of return but more objectives that in line with their client's preference. These tools can enhance the efficiency of the decision-making process and lead to more satisfactory investment outcomes for clients.

In conclusion, the journey taken by this thesis through the realms of stochastic dominance and prospect theory has led to a richer, more nuanced understanding of investor behavior. By modifying traditional financial theories and pushing the boundaries of current thinking, this work contributes to a more comprehensive and accurate view of financial markets and portfolio management. The thesis calls for a continued exploration into the complex interplay of behavioral tendencies and financial theory, for it is only through such research that we can better comprehend and navigate the complex world of finance.