STOCHASTIC DOMINANCE: AN OVERALL REVIEW AND AN EMPIRICAL EVALUATION ON THE EQUITY PREMIUM PUZZLE

EVANTHIA K. ZERVOUDI

ABSTRACT. This paper presents all the Stochastic Dominance (SD) rules (including Markowitz and Prospect SD Rule), situations relating to these rules such as arbitrage, and several classes of preferences consistent to the SD approach. It also contains various algorithms for testing SD relations and necessary and sufficient conditions that may improve the efficiency of each SD Rule. Finally, it is examined the existence of the well-known equity premium puzzle in several economic environments, under different economic conditions and varying time horizons. The analysis is developed in a Stochastic Dominance framework using original evidence from three markets: US, UK and Germany market. The results show that stocks stochastically dominate bonds at second order (and at any higher order) for any time horizon, under different economic situations for all international markets (US, UK, Germany). This implies that stocks outperform bonds in a great percentage and the equity premium puzzle is real and robust.

1. Introduction

In 1970, Rothschild and Stiglitz introduced the notion of *Stochastic Dominance* (SD) as an efficient way for pairwise comparison of random variables; for these comparisons Rothschild and Stiglitz used stochastic orderings i.e. binary relations defined on classes of probability distributions. A first characteristic of Stochastic Dominance (SD) that makes it very popular in decision making analysis, is that it is a nonparametric decision making approach. SD rules do not require any specification on investors' preferences or on assets' probability distribution functional forms but rely only on common preferences and rational beliefs. Each of the SD rules represents a specific type of investors (i.e. rational, risk averse etc.) who has her preferences. Thus, even if the Stochastic Dominance (SD) is a nonparametric decision making approach, there are classes of preferences that are consistent with the SD rules; under these types of preferences the Stochastic Dominance relations remain unaffected. Another desirable feature of the SD rules is that they can be defined either in terms of net wealth or in terms of change of wealth. This makes the Stochastic Dominance approach consistent with the classical Expected Utility framework (where the preferences are function of final net wealth) and with the Prospect Theory framework (where the preferences are function of change of

Date: Date of Submission January 15, 2019, Communicated to Prof. Yuri E. Gliklikh ; Date of Acceptance March 1, 2019.

2000 *Mathematics Subject Classification.* Primary 60G05, 60G07, 60G20; Secondary 60H30, 60H35, 65C60.

Key words and phrases. Stochastic Dominance, Equity premium puzzle, Prospect Theory, Decision Theory, Effectiveness,Utility Function.

wealth). Finally, the Stochastic Dominance (SD) rules can be applied without problem either on mixed bets or on bets with only positive or only negative returns (as several classical decision making approaches). This is a quite important property of the Stochastic Dominance approach since prospects with either positive or negative returns concern a more realistic framework in finance while bets with only positive or only negative returns are extremely rare.

A first goal of this paper is to present all the basic Stochastic Dominance (SD) rules and their relation with well-known situations such as the arbitrage. The classical Stochastic Dominance (SD) rules are completed by adding the new Prospect SD rules introduced by Levy and Wiener (1998) and Levy and Levy (2002) which concern the relation between the Stochastic Dominance and the Prospect Theory. They are also presented the classes of preferences that are consistent with the SD framework and it is analyzed how each of these preferences are related with each of the SD rules. Another axis of this study is the effectiveness of the Stochastic Dominance (SD) rules and how this could be improved. For this purpose, necessary and sufficient conditions are added in the analysis in order to improve the efficiency of each rule. In the final part of this paper, it is examined the well-known equity premium puzzle. There are several studies which support that this puzzle is not real; in the present paper it is proved that this puzzle does really exist and it is not a random phenomenon; the examination if this puzzle is real or not, is done in a Stochastic Dominance framework using empirical distribution functions (real observed data).

2. Stochastic Dominance Rules

2.1. **First Order Stochastic Dominance (SD1) Rule.** The first order stochastic dominance (SD1) criterion is appropriate for rational investors, who prefer more to less, independently of their risk attitude. The most appropriate utility function for rational agents is a nondecreasing function (without any additional requirement for concavity or convexity). Graphically, this rule means that the cumulative distribution function F is located totally below the cumulative distribution function G. These two cumulative distribution functions must not cross but they may tangent.

Definition 1. *A cumulative distribution function* F dominates a G by first order stochastic dominance *i.e.* F$(SD1)$G *if and only if*

$$F(X) \leqslant G(X) \quad \forall X \textit{ possible outcome} \tag{2.1}$$
$$\textit{with at least one strict inequality}$$

Remark 1. *The* first order stochastic dominance *(SD1) criterion is appropriate for all rational investors, who prefer more money to less, regardless if they are risk averters or risk lovers i.e. independently of their risk attitude. The utility function of rational agents must be nondecreasing without any additional requirement for concavity or convexity.*

Remark 2. *Graphically, this rule means that the cumulative distribution function F is totally located below the cumulative distribution function G. These two cumulative distribution functions must not cross but they may tangent.*

Remark 3. *If returns are normally distributed and $\sigma_F = \sigma_G$ then the* Mean Variance rule coincides with the first order stochastic dominance *rule i.e.* $SD1 \Leftrightarrow MV$ *.*

2.1.1. ***First order Stochastic Dominance (SD1) and Arbitrage.*** Arbitrage is a situation where an investor can achieve sure profit without any risk taking. Jarrow (1986) and Levy (1992) studying the relation between the arbitrage and the First order Stochastic Dominance, conclude that *arbitrage implies First order Stochastic Dominance (SD1) but not necessarily the inverse.* The only case where one could say that arbitrage implies First order Stochastic Dominance (SD1) and vice versa is when the cumulative distribution functions of the risky assets are *perfectly correlated* or the *one risky asset is a monotone function of the other* even with imperfect correlation. A characteristic example where arbitrage implies First order Stochastic Dominance and vice versa has to do with the derivatives and their underlying assets.

Proposition 1. *In the case of two uncertain risky assets with cumulative distribution functions F and G and returns R_F and R_G respectively an arbitrage situation implies the relation:*

$$R_F - R_G \geqslant 0 \forall R_F, R_G \tag{2.2}$$

with at least one strict inequality.

Proposition 2. *If we don't know the correlation between F and G this relation is also ensured by another important condition:*

$$\min(R_F) - \max(R_G) > 0 \tag{2.3}$$

which also implies First order Stochastic Dominance.

Conclusion 1. *If an agent obtains arbitrary profit by holding a portfolio R_F then its cumulative distribution function is first order stochastic dominant (SD1). Therefore, arbitrage implies First order Stochastic. However, the inverse is not necessarily true.*

2.2. **Second Order Stochastic Dominance (SD2) Rule.**

Definition 2. *A cumulative distribution function* F dominates a G by second order stochastic dominance *i.e.* F($SD2$)G *if and only if*

$$\int_{-\infty}^{x} [G(t) - F(t)\ ]\, dt \geqslant 0\ \forall X \tag{2.4}$$

with at least one strict inequality

Remark 4. *This rule concerns rational investors who are also risk averters (they are willing to sacrifice a larger return in order to avoid a great risk exposure). Thus, in the second order stochastic dominance (SD2) criterion there is the additional assumption that of the* global risk aversion. *In this case, investors' preferences must be nondecreasing and concave in order to capture the (additional) characteristic of risk aversion. Using the second order stochastic dominance criterion one can achieve a more sensible selection of investments for people who prefer a less profitable investment in order to avoid a great risk.*

Remark 5. *Graphically, this rule means that the area between the cumulative distribution function G and the cumulative distribution function F from -∞ to any positive x is positive. Even if there is an area where $[G(t) - F(t)]$ is negative then there must be a preceding larger area where $[G(t) - F(t)\ ]$ is positive.*

Remark 6. *If returns are normally distributed then the Mean Variance rule coincides with the second order stochastic dominance rule i.e.* $SD2 \Leftrightarrow MV$ *.*

2.3. **Second Order Stochastic Dominance (SD2*) Rule for risk-seeking investors.**

Levy and Wiener (1998) proposed an additional Stochastic Dominance rule, also called "second order stochastic dominance rule", which however concerns investors who are risk seekers and not risk averters as in the classical second order stochastic dominance rule.

Definition 3. *Specifically, a cumulative distribution function* F dominates a G by second order stochastic dominance *i.e.* F$(SD2^*)$G *if and only if*

$$\int_x^{\infty} [G(t) - F(t)\ ]\, dt \geqslant 0\ \forall x \tag{2.5}$$

with at least one strict in equality

Remark 7. *This rule concerns all rational investors who are risk seekers. Their only concern is about the return in the sense that they are willing to undertake a greater risk in order to have a higher return. Investors' preferences must be nondecreasing and convex in order to capture the additional characteristic of risk seeking. If* X$(SD2^*)$Y *for two uncertain risky assets X,Y then* $(-$Y$\,(SD2) -$ X.)

Remark 8. *Graphically, this rule means that the area between the cumulative distribution function G and the cumulative distribution function F from the* ***greatest x*** *to $+\infty$ is positive. It is possible to exist an area where $[G(t) - F(t)\ ]$ is negative before the last positive area (for the greatest x) but this negative area must be smaller than the last positive one. In other words, for a negative area between the cdf F and G there must be a larger positive area to the right of it.*

Remark 9. *If returns are normally distributed then the Mean Variance rule coincides with the second order stochastic dominance rule i.e.* $SD2^* \Leftrightarrow MV$ *.*

2.3.1. ***First Order (SD1) and Second Order (SD2) Stochastic Dominance Rules in terms of total wealth and in terms of change of wealth.***

The first (SD1) and the second order (SD2, SD2*) Stochastic Dominance rules can be expressed either in terms of total wealth (t+w) or in terms of change of wealth (t). The first and the second order stochastic dominance rules remain unaffected in both expressions. The only mathematical difference between the two expressions is a "shift" of the cumulative distribution functions F and G by a constant term w which is the initial wealth level. More specifically:

Definition 4.

$$SD1 : F(X) \leqslant G(X)\ \Leftrightarrow F(X + w) \leqslant G(X + w)\ \forall X, w \tag{2.6}$$

Definition 5 (SD2)**.**

$$\int_{-\infty}^{x} [G(t) - F(t)\ ]\, dt \geqslant 0 \Leftrightarrow \int_{-\infty}^{x} [G(t + w) - F(t + w)\ ]\, dt \geqslant 0\ \forall X, w \tag{2.7}$$

Definition 6 (SD2*).

$$\int_x^{\infty} [G(t) - F(t)\ ]\, dt \geqslant 0 \Leftrightarrow \int_x^{\infty} [G(t+w) - F(t+w)\ ]\, dt \geqslant 0 \ \forall X, w \tag{2.8}$$

Remark 10. *Since the $SD1$ and the $SD2$ rules are independent of the initial wealth level w, they can be defined either in terms of "change of wealth" x or in terms of the total wealth $x + w$.*

2.4. **Third Order Stochastic Dominance (SD3) Rule.**

Definition 7. *A cumulative distribution function* F dominates a G by third order stochastic dominance *i.e.* F($SD3$)G *if and only if*

$$\int_{-\infty}^{x} \int_{-\infty}^{t} [G(u) - F(u)\ ]\, dudt \geqslant 0 \ \forall X \ \textit{and} \ E_F(X) \geqslant E_G(X) \tag{2.9}$$

with at least one strict inequality

Remark 11. *This rule concerns all rational investors who are risk averters and additionally take into account the skewness of the cumulative distribution function of each prospect. In most empirical researches, investors seem to have a clear preference to positively skewed assets which offer a protection against great losses.*

2.5. **N-th Order Stochastic Dominance (SDn) Rule.** The first three degrees of Stochastic Dominance (SD1, SD2 and SD3) are the most meaningful rules in finance since they capture widely observed investors' behaviors. However, there are more than three degrees of Stochastic Dominance.

Definition 8. *The general definition of the n-th order stochastic dominance arises from the third order stochastic dominance rule by replacing the double integral by n-1 integrals. Thus, a cdf* F dominates *a cdf* G by nth order stochastic dominance *i.e.* F(SDn)G *if and only if*

$$\int_{-\infty}^{x} \int_{-\infty}^{t_1} ... \int_{-\infty}^{t_{n-2}} [G(u) - F(u)\ ]\, dudt_{n-2}...dt_1 \geqslant 0 \ \forall X \text{ and } E_F(X) \geqslant E_G(X) \tag{2.10}$$

with at least one strict inequality

Remark 12. *As we see, at each higher degree of Stochastic Dominance a new condition is added. Hence, one can understand that if the first order Stochastic Dominance holds then the Stochastic Dominance of any higher degree holds as well. Thus, the $SD1$ implies the $SD2$ which implies the $SD3$. Consequently, the efficient sets arising by SDi rules i=2,3,.. are subsets of the $SD1$ rule. This happens because any additional condition in Stochastic Dominance rules removes element from their efficient sets. Hence, the $SD3$ efficient set is subset of the $SD2$ efficient set which is subset of the $SD1$ efficient set.*

2.6. **Prospect Stochastic Dominance (PSD) Rule.**

Definition 9. *A cumulative distribution function* F dominates *a cumulative distribution function* G by Prospect Stochastic Dominance *i.e.* F(PSD)G *if and only if*

$$\int_y^x [G(t) - F(t)\ ]\,dt \geqslant 0 \text{ for all pairs } x > 0 \text{ and } y < 0 \tag{2.11}$$
$$\text{with at least one strict inequality}$$

Remark 13. *This rule concerns all rational investors who are risk averters for gains i.e. for the integral* $[0, x]$ *and risk seekers for losses i.e. for the integral* $[y, 0]$ *. Such type of investors is the Prospect Theory investors whose preferences are described by an S-shaped value function i.e. nondecreasing function, concave for gains (returns greater than a reference point) and convex for losses (returns smaller than a reference point) (see Kahneman and Tversky (1979)).*

Remark 14. *The Prospect Stochastic Dominance (PSD) rule can be also expressed as a combination of the SD2 and the* $SD2^*$ *rules. This becomes clearer in the formulation of the Prospect Stochastic Dominance rule that Levy and Levy (2002) propose. In this formulation the integral* $\int_y^x [G(t) - F(t)\ ]\,dt$ *breaks into two integrals by separating the returns into gains and losses. Specifically:*

$$\int_y^0 [G(t) - F(t)\ ]\,dt \geqslant 0 \text{ for all } y \leqslant 0 \tag{2.12}$$
$$\text{and} \int_0^x [G(t) - F(t)\ ]\,dt \geqslant 0 \text{ for all } x \geqslant 0$$
$$\text{with at least one strict inequality}$$

Remark 15. *The first integral concerns the risk seeking for losses and it corresponds to the SD2* rule; the second integral concerns the risk aversion for gains and it corresponds to the SD2 rule .*

Remark 16. *The graphical explanation of this PSD rule is expressed as a combination of the SD2 (second integral) and the SD2* (first integral) rules' graphical explanations (see Tversky, Kahneman (1992), Baucells, Heukamp (2006)). If we isolate each integral and analyze it independently then we have for the second integral the SD2 rule's graphical interpretation and for the first integral the SD2* rule's graphical interpretation. The first integral regards the risk seeking for losses and the second one regards the risk aversion for gains.*

Remark 17. *Since the Prospect Stochastic Dominance (PSD) is a combination of SD2 and SD2* rules one can understand that the PSD rule is independent of the initial wealth level w as well. Of course, this is normal sinse the Prospect Stochastic Dominance is developed in the Prospect Theory framework where the value function is defined on the change of wealth and it is independent of the initial wealth level.*

2.7. **Markowitz Stochastic Dominance (MSD) Rule.** Finally, in Levy and Levy's [2002] we find another *Stochastic Dominance rule* which concerns *Markowitz type investors* whose preferences are described by an *inverse S-shaped value function*. This type of preferences is nondecreasing, concave for losses and convex for gains and represents agents who are risk averse for losses and risk seeking for gains.

By analyzing several hypothetical gambles, Markowitz concluded that individuals tend to be risk averse for losses and risk seeking for gains, as long as the possible outcomes are not very extreme.

Definition 10. *For extreme outcomes, Markowitz argues that individuals become risk averse for gains and risk seeking for losses. Thus, a* cdf F dominates a cdf G in Markowitz sense *i.e.* F(MSD)G *if and only if*

$$\int_{y}^{x} [G(t) - F(t)\ ]\,dt \geqslant 0 \text{ for all pairs } x \geq 0 \text{ and } y \leq 0 \tag{2.13}$$

with at least one strict inequality

Remark 18. *This rule concerns Markowitz type rational investors who are risk averters for losses (interval $(-\infty, y)$) and risk seekers for gains (interval $(x, +\infty)$).*

Remark 19. *The Markowitz Stochastic Dominance (MSD) rule, as the Prospect Stochastic Dominance (PSD), can be also expressed as a combination of the SD2 and the SD2^* rules. Levy and Levy (2002) propose another formulation of the Markowitz Stochastic Dominance (MSD) rule where the integral $\int_y^x [G(t) - F(t)\ ]\,dt$ breaks into two integrals by separating the returns into gains and losses. Specifically:*

$$\begin{aligned} \int_{-\infty}^{y} [G(t) - F(t)\ ]\,dt &\geqslant 0 \text{ for all } y \leqslant 0 \\ \text{and} \int_{x}^{\infty} [G(t) - F(t)\ ]\,dt &\geqslant 0 \text{ for all } x \geqslant 0 \end{aligned} \tag{2.14}$$

with at least one strict inequality

Remark 20. *The first integral concerns the risk aversion for losses and it corresponds to the SD2 rule and the second integral concerns the risk seeking for gains and it corresponds to the SD2* rule.*

Remark 21. *The graphical explanation of this rule is a combination of SD2* and SD2 rules' graphical explanations. If each integral is isolated and analyzed independently then the second integral is related to the SD2* rule's graphical interpretation and the risk seeking for gains while the first integral is related to the SD2 rule's graphical interpretation and the risk aversion for losses.*

2.7.1. ***Relation between the Markowitz and the Prospect Stochastic Dominance Rule.*** The Markowitz Stochastic Dominance and the Prospect Stochastic Dominance seem to be inverse i.e. if *the cumulative distribution function F dominates the cumulative distribution function G* in the Prospect Stochastic Dominance sense i.e. $F(PSD)G$ then *the cumulative distribution function G dominat*es *the cumulative distribution function F* in the Markowitz Stochastic Dominance sense i.e. $G(MSD)F$. But, this is not necessarily true. The explanation is simple. The **necessary condition** for all the Stochastic Dominance rules is the dominant cumulative distribution function to have higher mean than the dominated cumulative distribution function. Thus, if the cumulative distribution function F dominates the cumulative distribution function G in the Prospect Stochastic Dominance sense($F(PSD)G$)this means that the cumulative distribution function F has higher mean than the cumulative distribution function G; but this implies that the cumulative distribution function G cannot dominate the cumulative distribution function F in the Markowitz Stochastic Dominance sense $(G(MSD)F)$ since G has not higher mean than F. Hence, the only case where Markowitz Stochastic Dominance and Prospect Stochastic Dominance rule are inverse is when F and G have the *same mean i.e.* $E_F(x) = E_G(x)$. If this condition holds then $F(PSD)\,G \Leftrightarrow \; G(MSD)\,F$.

3. Stochastic Dominance Criteria and Utility function classes

Stochastic Dominance (SD) is in general a nonparametric decision making approach. SD rules do not require any specification on investors' preferences or assets' probability distribution functional forms but rely only on common preferences and rational beliefs. However, there are classes of utility functions that are valid to the Stochastic Dominance approach and strictly related to the Stochastic Dominance rules. These classes of preferences represent agents with specific characteristics. Levy (1992) in his paper is referred to such classes of preferences U_i where i=1,2,3... and their relation with the SD rules:

(1) $\mathbf{U}_1$ the class of *increasing* utility functions e.g. functions u with positive first derivative $u' \geqslant 0$. This type of preferences represents *rational* agents who prefer more than less.
(2) $\mathbf{U}_2$ the class of *increasing concave* utility functions e.g. functions u with $u' \geqslant 0$ and $u'' \leqslant 0$. This type of preferences represents agents who are *rational* and *risk averters* in the sense that they prefer more money to less but also they dislike risk and they are willing to sacrifice return in order to avoid a greater risk.
(3) $\mathbf{U}_2^*$ the class of *increasing convex* utility functions e.g. functions u with $u' \geqslant 0$ and $u'' \geqslant 0$. This type of preferences represents agents who are *rational* and *risk seekers* in the sense that they prefer more money to less but they are also willing to undertake a greater risk in order to have a greater return. This type of investors prefer prospects whose return is above their certain mean.
(4) $\mathbf{U}_3$ the class of *increasing concave* utility functions with *positive third derivative* e.g. functions u with $u' \geqslant 0$, $u'' \leqslant 0$ and $u''' \geqslant 0$. This type of preferences represents agents who are *rational* and *risk averters* in the sense that they prefer more money to less but also they dislike risk and they are willing to sacrifice return in order to avoid a greater risk but they also have a clear preference to positively skewed assets (positive third derivative) in order to avoid too large losses.
(5) $\mathbf{U}_{DARA}$ the class of *decreasing absolute risk aversion* utility functions with $p = -\frac{u''}{u'} \Rightarrow p' \leqslant 0$ e.g. *increasing concave* utility functions $u \in \mathbf{U}_2$ with $u' \neq 0$ and $p' \leqslant 0$. For this type of preferences holds $\mathbf{U}_{DARA} \subset \mathbf{U}_3$.
(6) $\mathbf{U}_{SRA}$ the class of *standard risk aversion* utility functions with $p = -\frac{u''}{u'} \Rightarrow p' \leqslant 0$ and $\phi = -\frac{u'''}{u''} \Rightarrow \phi' \leqslant 0$ e.g. *increasing concave* utility functions with *positive third derivative* utility functions $u \in \mathbf{U}_3$ with $u' \neq 0$, $u'' \neq 0$ and $p' \leqslant 0$, $\phi' \leqslant 0$. This ϕ is the absolute prudence index which measures

the *prudence*, a property of decision makers' preferences that has to do with the precautionary savings of agents. The decreasing absolute prudence is a necessary and sufficient condition which guarantees that the savings of agents who are rich enough are not so sensitive to the risk related to the future income. It also means that the fourth derivative of the utility function is negative, a characteristic which called *temperance*.

(7) In general, $\mathbf{U}_n$ is the class of utility functions u with $u^{(2n)} \leqslant 0$ and $u^{(2n+1)} \geqslant 0$ assuming that these derivatives exist.

(8) $\mathbf{V}_p$ the class of all *S-shaped* utility functions V with $V' \geqslant 0\ \forall x \neq 0$ and $V'' \leqslant 0$ for $x > 0$ and $V'' \geqslant 0$ for $x < 0$. This class of preferences concerns all *rational investors* who are *risk averters* for gains *and risk seekers* for losses.

(9) $\mathbf{V}_M$ the class of all *reverse S-shaped* value functions V with $V' \geqslant 0\ \forall x \neq 0$ and $V'' \leqslant 0$ for $x < 0$ and $V'' \geqslant 0$ for $x > 0$. This type of preferences is concave for losses and convex for gains and represents *rational* agents who are *risk averters* for losses *and risk seekers* for gains.

4. Relations between Classes of Preferences and SD rules

4.1. First Order Stochastic Dominance (SD1) Rule.

$$\begin{aligned} &\mathbf{F}\,(SD1)\,\mathbf{G}\text{:}\ F(X) \leqslant G(X)\ \forall X \\ &\Leftrightarrow E_F U(X) \geqslant E_G U(X)\ \forall u \in \mathbf{U}_1 \end{aligned} \tag{4.1}$$

This means that if the cdf F stochastically dominates the cdf G at first order then F will be preferred to G for any increasing utility function $u \in \mathbf{U}_1$ i.e. any rational agent will prefer F instead of G.

4.2. Second Order Stochastic Dominance (SD2) Rule.

$$\begin{aligned} &\mathbf{F}\,(SD2)\,\mathbf{G}\text{:} \int_{-\infty}^{x} [G(t) - F(t)\ ]\, dt \geqslant 0\ \forall X \\ &\Leftrightarrow E_F U(X) \geqslant E_G U(X)\ \forall u \in \mathbf{U}_2 \end{aligned} \tag{4.2}$$

This means that if the cdf F stochastically dominates the cdf G at second order then F will be preferred to G for any increasing and concave utility function $u \in \mathbf{U}_2$ i.e. any rational and risk averter agent will prefer F instead of G.

4.3. Second Order Stochastic Dominance (SD2*) Rule for risk-seeking investors.

$$\begin{aligned} &\mathbf{F}\,(SD2^*)\,\mathbf{G}\text{:}\ \int_{x}^{\infty} [G(t) - F(t)\ ]\, dt \geqslant 0\ \forall X \\ &\Leftrightarrow E_F U(X) \geqslant E_G U(X)\ \forall u \in \mathbf{U}_2^* \end{aligned} \tag{4.3}$$

This means that if the cdf F stochastically dominates the cdf G at second order then F will be preferred to G for any increasing and convex utility function $u \in \mathbf{U}_2^*$ i.e. any rational and risk seeker agent will prefer F instead of G.

4.4. **Third Order Stochastic Dominance (SD3) Rule.**

$$\begin{gathered}\mathbf{F}\,(SD3)\,\mathbf{G}\text{: } \int_{-\infty}^{x}\int_{-\infty}^{t}[G(u)-F(u)\,]\,du dt \geqslant 0\ \forall X \text{ and } E_F(X) \geqslant E_G(X)\\ \Leftrightarrow E_F U(X) \geqslant E_G U(X)\ \forall u \in \mathbf{U}_3\end{gathered} \tag{4.4}$$

This means that if the cumulative distribution function F stochastically dominates the cumulative distribution function G at third order then F will be preferred to G for any increasing and concave utility function with positive third derivative $u \in \mathbf{U}_3$ i.e. any rational and risk averter agent who prefer positively skewed assets will choose F instead of G.

4.5. **Prospect Stochastic Dominance (PSD) Rule.**

$$\begin{gathered}\mathbf{F}\,(PSD)\,\mathbf{G}\text{: } \int_{x}^{y}[G(t)-F(t)\,]\,dt \geqslant 0 \text{ for all pairs } x>0 \text{ and } y<0\\ \Leftrightarrow E_F V(X) \geqslant E_G V(X)\quad \forall V \in \mathbf{V}_p\end{gathered} \tag{4.5}$$

This means that if the cumulative distribution function F stochastically dominates the cumulative distribution function G in the Prospect Theory sense then F will be preferred by G for any S-shaped value function $V \in \mathbf{V}_p$ i.e. any Prospect Theory investor will prefer F instead of G (see Baucells, Heukamp (2006)).

Remark 22. *It is important here to mention the relation (in terms of utility functions) between the Prospect Stochastic Dominance and the classical expected utility framework:A cumulative distribution function F stochastically dominates another cumulative distribution function G in the Prospect Theory sense if and only if F dominates G in the expected utility framework for all S-shaped utility functions* $V_P(w{+}x)$.

Remark 23. *The use of the total wealth (w+x) instead of the change of wealth x is because expected utility maximizers take their decisions based on the final net wealth and not the change of wealth. Thus, the Prospect Stochastic Dominance can be considered as a special case of the Expected Utility Theory where the classical concave utility function is replaced by an S-shaped value function concave for gains (x>0) and convex for losses (x<0) using as reference point the current wealth w.*

4.6. **Markowitz Order Stochastic Dominance (MSD) Rule.**

$$\begin{gathered}\mathbf{F}\,(MSD)\,\mathbf{G}\text{: } \int_{-\infty}^{y}[G(t)-F(t)\,]\,dt \geqslant 0 \text{ for all } y \leqslant 0\\ \text{and} \int_{x}^{\infty}[G(t)-F(t)\,]\,dt \geqslant 0 \text{ for all } x \geqslant 0\\ \Leftrightarrow E_F V(X) \geqslant E_G V(X)\ \forall V \in \mathbf{V}_M\end{gathered} \tag{4.6}$$

This means that if the cumulative distribution function F stochastically dominates the cumulative distribution function G in the Markowitz sense then F will be preferred by G for any inverse S-shaped value function $V \in U_M$ i.e. any Markowitz type investor will prefer F instead of G.

5. SD Rules' Effectiveness

An important issue in decision making analysis is the *effectiveness* of a decision rule such as the Stochastic Dominance rules; the effectiveness of a rule is given by the fraction of the size of its efficient set to the size of its feasible set. Thus, the effectiveness of a decision rule can be improved significantly by reducing the size of its feasible set (see Kuosmanen (2004) among others). One way to do this is by adding necessary conditions in the analysis; some portfolios won't meet the additional requirements and they will be excluded from the resulting feasible sets. But, there are some important issues that must be taken into account when an additional restriction is imposed on a decision making rule; the computational cost of this additional restriction, the number and the characteristics of the portfolios that are excluded from the initial original feasible sets are some of these issues under consideration.

Concerning the SD rules, it holds that at each Stochastic Dominance rule of a higher order *a new condition is added*. More specifically, the first order stochastic dominance (SD1) rule is the simplest criterion in the SD framework,the only requirement is rationality (this rule represents rational agents). The efficient set resulting from SD1 rule is too large (the greatest of all the higher order's SD rules' efficient sets) and quite similar to the initial feasible set. This fact makes the SD1 efficient set uninformative and the first order stochastic dominance (SD1) rule to seem ineffective. The SD2 rule is more effective than SD1 rule. The SD2 rule requires one more necessary condition, that of risk aversion (this rule represents rational and risk averse agents). The SD2* rule is also more effective than the SD1 rule. The SD2* rule's additional necessary condition is that of risk seeking (this rule represents rational and risk seeker agents). These additional assumptions reduce the size of the resulting feasible sets and make the SD2 and the SD2* rules to seem more effective. In the SD3 rule there is the additional requirement of the positive third derivative (this rule represents rational and risk averse agents who take into account the skewness of the assets) etc. The PSD rule and the MSD rule require a great number of assumptions, a fact that reduces very much the size of their feasible sets and makes these rules to seem quite effective. In general, the more the assumptions required by a rule the smaller the resulting feasible set and the greater the effectiveness of this rule.

If a cumulative distribution function F stochastically dominates a cumulative distribution function G at first order ($F(SD1)G$) then F stochastically dominates G at any higher order, as well i.e. SD1 implies SD2 implies SD3 etc ($SD1 \Longrightarrow SD2 \Longrightarrow SD3 \Longrightarrow SDi \quad \forall i > 3$). This means that the efficient sets resulting by higher order SD rules are subsets of the efficient sets resulting by lower order SD rules i.e. $SDi \subseteq ... \subseteq SD3 \subseteq SD2 \subseteq SD1 \quad \forall i > 3$. The importance of necessary conditions which may improve significantly the effectiveness of a Stochastic Dominance rule will be clearer in the empirical study below.

5.1. **Necessary and sufficient conditions for each SD rule.** It is remarkable that the pairwise comparisons of the two empirical distribution functions of bonds and stocks are time consuming and increase the computational cost of the algorithms. So, it is important to reduce these comparisons in order to improve the efficiency of the corresponding algorithms and save resources (by decreasing the time and the number of the pairwise comparisons). The necessity to reduce the number of comparisons becomes crucial if one thinks that every stochastic dominance rule requires $\frac{T!}{(T-2)!}$ pairwise comparisons for T alternative risky assets each of which has n monthly observations equally weighted with probability $\frac{1}{n}$. These comparisons demand a great amount of resources; a reduction in comparisons implies a significant decrease in demanding computing time. In order to do this, it is essential to use necessary and sufficient conditions that each Stochastic Dominance rule must satisfy; using these conditions the number of comparisons for each SD rule is reduced, the decision making analysis is simplified while the speed and the computational cost of the algorithms (corresponding to these SD rules) is improved significantly. Such general necessary conditions are given by Levy (1992) and De Nadai and Pianca (2007). In Levy [1992] we find some important necessary conditions that each SD rule must satisfy. Using these *necessary* conditions we can simplify enough the decision making analysis and improve the speed and the computational cost of an algorithm. More specifically, according to Levy (1992):

(1) An uncertain prospect is dominant if it has *greater or equal mean* to its alternatives. In other words, an option (portfolio, asset etc) dominates another if it has a *higher or equal mean*.
(2) Between two prospects with *equal means* that with the *greater variance* cannot be the dominant.
(3) The prospect which contains the lowest observaion cannot be the dominant.
(4) If a prospect is stochasticly dominated at any order i then it is excluded from the efficient set resulting from this SD_i rule without the need for further comparisons with other prospects. The efficient frontier created by a SD_i rule contains all efficient portfolios that are dominant at this order i.
(5) The first order stochastic dominance (SD1) implies the second order stochastic dominance (SD2), the third order stochastic dominance (SD3) and the stochastic dominance of any higher order:$SD1 \Longrightarrow SD2 \Longrightarrow SD3 \Longrightarrow SDi \quad \forall i > 3$. This relation means that a portfolio which is SD1 dominated, it is also SD2 dominated and SD3 dominated. On the other hand, if a prospect is not dominant by third order ($SD3$) then the same prospect cannot be dominant by second ($SD2$) or by first order ($SD1$) as well. Generally speaking, if an option is stochastic dominant at a lower order then it will be stochastic dominant at any higher order as well without need for further examination and if an cumulative distribution function F stochastically dominates another cumulative distribution function G by first order then it stochastically dominates G by any higher order and in the prospect sense as well i.e.$SD1 \Rightarrow SDi \forall i > 1$ and $SD1 \Rightarrow PSD$. If F and G have the same mean then F stochastically dominates G in the prospect sense e.g. $F\,(PSD)\,G$ if and only if G stochastically dominates F in the Markowitz sense e.g. $G\,(MSD)\,F$. This is a quite important necessary condition for this empirical study since if a cumulative distribution function is first order (SD1) stochastic dominant it is also second order (SD2or SD2*) stochastic dominant and if it is second order (SD2or SD2*) stochastic dominant it is also third order (SD3) and any higher order ($SDi\ \forall i > 3$) stochastic dominant without need additional conditions and algorithms.

These **necessary conditions** are employed in this paper in order to reduce the pairwise comparisons of the two empirical distribution functions of bonds and stocks. The more the necessary and sufficient conditions that are employed the greater the effectiveness of the SD rules and the smaller the computational cost of the corresponding algorithms. Some other important necessary and sufficient conditions for SD1, SD2 and SD2* rules, that can also be used in this study, are given by De Nadai and Pianca (2007). More specifically:

(1) The **sufficient condition** for the SD1 of a cumulative distribution function F over another cumulative distribution function G i.e. $F\,(SD1)\,G$ is given by: $\min_x F(x) \leq \max_x G(x)$. This condition implies that the cumulative distribution function F stochastically dominates the cumulative distribution function G by first order if the minimum of the cumulative distribution function F is not greater than the maximum of the cumulative distribution function G. Graphically, this means that the cumulative distribution function F stochastically dominates the cumulative distribution function G by first order if the cumulative distribution function F is located totally below or at most it tangents the cumulative distribution function G.

(2) The **necessary conditions** for the first order stochastic dominance of a cumulative distribution function F over a cumulative distribution function G are given by $E_F(X) > E_G(Y)$ and by the "left tail condition" $\min_x F(x) \geq \min_x G(x)$. The "left tail problem" is a sufficient condition for non dominance of the thicker left tail cumulative distribution function over its alternatives. Thus, if the cumulative distribution function F stochastically dominates the cumulative distribution function G at first order then the cumulative distribution function F has no thicker left tail than the cumulative distribution function G ($F\,(SD1)\,G$).

(3) A cumulative distribution function (cdf) F stochastically dominates the cumulative distribution function G at second order if the minimum of the cdf F is smaller than the maximum of the cdf G i.e. $F(SD2)G$ if $\min_x F(x) < \max_x G(x)$. This condition implies that the cumulative distribution function F stochastically dominates the cumulative distribution function G by second order if the minimum of the cumulative distribution function F is smaller than the maximum of the cumulative distribution function G. Graphically, this means that the cumulative distribution function F stochastically dominates the cumulative distribution function G by second order if the cumulative distribution function F is located below or intersects the cumulative distribution function G. Another sufficient condition for the second order stochastic dominance (for both risk averse and risk seeking investors) is the first order stochastic dominance rule SD1 i.e. $SD1 \Rightarrow SD2 \; and \; SD2^*$.

(4) The **necessary conditions** for the second order stochastic dominance of a cumulative distribution function F over a cumulative distribution function G are given by $E_F(X) \geqslant E_G(Y)$ (for both risk averse and risk seeking investors) and by the "left tail condition" $\min_x G(x) \geq \min_x F(x)$ (for risk averse investors). The "left tail problem" has to do with the fact that the thicker left tail of a cumulative distribution function is a sufficient condition for no dominance of this cumulative distribution function over the

alternatives. This condition implies that the cumulative distribution function G must have thicker left tail than the cumulative distribution function F so as to the cumulative distribution function F stochastically dominates the cumulative distribution function G by second order i.e. $F\,(SD2)\,G$. On other hand, the necessary conditions for the second order stochastic dominance (for risk seeking investors) of a cumulative distribution function F over a cumulative distribution function G is the "right tail condition" $\max_x F(x) > \max_x G(x)$. This condition implies that the cumulative distribution function F must have thicker right tail than the cumulative distribution function G in order to stochastically dominates the cumulative distribution function G at second order (SD2*).

Remark 24. *The use of these necessary conditions reduces the comparisons between the ex-post returns of bonds and stocks and improves significantly the speed and the effectiveness of the employing algorithms. The analysis starts by examining which of these necessary conditions are satisfied (so as to avoid redundant comparisons) and it continues by employing the previous algorithms. Firstly, it is examined if the stocks' empirical distribution stochastically dominate bonds' empirical distribution at first order (SD1) and if this relation does not hold the examination continues with the second order (SD2), the third order SD relations etc. If stocks stochastically dominate bonds at a lower order then they stochastically dominate bonds at any higher order without need any further comparison* ($SD1 \Rightarrow SD2 \Rightarrow SD3 \Rightarrow SDi\ \forall i > 3$)*. Since the Prospect Stochastic Dominance (PSD) is a combination of SD2 and SD2* rules then if stocks' empirical distribution is stochastically dominant at first order (SD1) or at second order (SD2 and SD2*) over bonds' empirical distribution then it will be Prospect stochastically dominant as well* ($SD1 \Rightarrow SD2 \Rightarrow PSD$)*. After that, the first two moments (the mean and the variance) of stocks and bonds' empirical distributions are compared; the option (stocks or bonds) which has a greater or equal mean and not greater variance than its alternative is the dominant one. This dominant option cannot contain the lowest observation; so another quick checking concerns which of these empirical distributions contains the lowest observed return at each investment period. Finally, if there is not a SD1 option, the left and the right tail conditions are examined.*

6. **Empirical Study**

6.1. **Equity premium puzzle and issues under consideration.** One of the most interesting puzzles in financial economics is the equity premium puzzle which is based on the observation that equity returns tend to be higher than bond returns. Mehra and Prescott (1985) first report that in the US between 1889 and 1978 the average real rate of return on T-Bills is 0.80% per year while the average real rate of return on equities is 6.98% per year; in other words the equity premium is 6.18% per year. Later studies use data sets that start as early as 1802 and span until 2005, and still find that the average US equity return (inflation adjusted) has been approximately 7.67% while the return on the risk free asset has been 1.31%; thus the average premium has been approximately 6.36% (Mehra, 2006; p.5). This means that stocks outperform bonds in a very high percentage. The equity premium can be considered as a compensation of investors for holding the risky stocks instead of the "riskless" government bonds. However , this premium is significantly greater than the premium that can be expected from standard neoclassical models; for instance a relative risk aversion coefficient of over 30 could explain the puzzle, however, theoretical and empirical estimates indicate that the coefficient is around 2 (see, Benartzi and Thaler, 1995). Thus, there is not a reasonable risk aversion parameter that can explain the puzzle (see also, Siegel, 1999; Kandel and Stambaugh, 1991).

Many studies attempt to provide an explanation for the puzzle, using a range of theoretical approaches, testing methodologies, and data. Kandel and Stambaugh (1991) support that the large equity premium may arise from the fact that consumers are too averse to small negative consumption's shocks that may be caused by stock markets fluctuations. Epstein and Zin (1990) emphasize the role of 'first-order risk aversion' as an explanation to the puzzle. A different approach is adopted by Constantinides (1990) who argues that the puzzle is resolved if one relaxes the time separability of the classical expected utility theory preferences in order to allow for consumption complementarity , i.e. use a utility function where the utility is derived by a comparison between the current consumption and prior levels of consumption; this behavior is described as habit persistence. In similar spirit, Otrok, Ravikumar, and Whiteman (2002) provide an explanation of the puzzle employing an intertemporal consumption-CAPM with habit formation, while Meyer and Meyer (2005) show that a habit formation utility function may eliminate the puzzle. In order to face the weaknesses of the models that used the habit formation for explaining the equity premium puzzle, other models proposed not to compare the current consumption with prior levels of consumption but the current consumption with others' levels of consumption. Abel (1990) used the term "catch-up with the Joneses" to describe this behavior. However, even in this case the risk aversion level that may explain the equity premium puzzle remains very high. Another effort to solve the equity premium puzzle is done by Mankiw and Zeldes (1991) who support that this puzzle arises from the aggregation of stockholders' and no stockholders' consumption; most of the people prefer not to hold stocks but for those who decide to hold stocks their consumption is extremely sensitive to stock market fluctuations. Thus, the equity premium puzzle is explained as a strong incentive for agents to invest in stocks. However, this study demands a very high risk aversion level as well. Other studies indicate that market frictions (i.e. inability of investors to diversify their portfolios) and informational asymmetry explain a significant proportion of the premium (Zhou, 1999).

Longstaff and Piazzesi (2004) present a model where the equity premium reflects three types of risk (consumption-risk, event-risk, corporate-risk) and show that their model implies an equity premium much larger that the premium implied by standard models. Ang, Bekaert, Liu (2005) use a model that assumes disappointment aversion preferences and asymmetric aversion to gains versus losses; they show that the large equity premium is reconciled with a typical asset allocation to equities of about 60% (see also, Gul, 1991). The robustness of this puzzle was also examined by Siegel (1999) who found that the real equity returns were stable for several decades while the short-term government bonds followed a downward trend. They tested the equity premium for several investment periods and they concluded that agents who chose stocks for long investment periods achieve much greater returns than those who chose bonds. This result shows the important role of the investment period in this puzzle. Benartzi and Thaler (1995) developed one of the most important studies in equity premium puzzle analysis. They tried to solve this puzzle using two significant notions. The first notion was the loss aversion according to which agents are more sensitive to losses than to gains. The second notion was that of mental accounting. This behavioral concept is related with the aggregation rules that agents follow when they take their decisions. Benartzi and Thaler concluded that loss averse investors are more willing to undertake a risk when the evaluation of their investment is not frequent. This means that a risky asset is more attractive to a loss averse investor if she wants to hold it for a long period and as the evaluation of her investment is not frequent. This behavior is characterized by Benartzi and Thaler (1995) as Myopic Loss Aversion. The Myopic Loss Aversion is the key that Benartzi and Thaler (1995) used to explain the equity premium puzzle. According to Benartzi and Thaler (1995) as the investment horizon or the evaluation period increases, the stocks become more attractive to a rational loss averse investor. The evaluation period at which a Prospect Theory investor becomes indifferent between stocks and bonds is not a unique for all investors but the most representative evaluation period is one year. In 2016, Zervoudi and Spyrou presented new evidence on the optimal evaluation period of Benartzi and Thaler (1995), over-time and across different markets. Their results indicate that the optimal holding period is now formulated to 7 months (shorter than 12 months) and to 4 or 5 months during periods of economic crisis (much shorter to that reported by Benartzi and Thaler). This optimal holding period is in accordance with the Myopic Loss Aversion hypothesis, remains insensitive to the value function used in the analysis but it is quite sensitive to economic conditions. Zervoudi and Spyrou (2016) conclude that the globalization, the intensity, and the frequency of financial market crises during the recent decades could explain this shorter optimal holding period and the tendency of investors to evaluate their portfolios more frequently.

All these studies begin from the common base that this puzzle is real. However, there are studies which support that the equity premium puzzle does not exist but it is a random phenomenon or the consequence of a statistical bias. Specifically, Siegel and Thaler (1997) support that stocks markets experienced a "good luck" for many years (they had great returns) and even in periods of "bad luck" the equity returns were always higher than that of short-term fixed income securities. Thus, the equity premium was very high for too long periods something that justifies its magnitude. An alternative explanation of the equity premium puzzle is based on the survivorship bias. If the data used for the estimation of the equity risk premium are returns "survived" by an economic catastrophe then the results must be necessarily distorted by the survivorship bias. Such an economic catastrophe was the stock market crash and the resultant Great Depression of 1929 in US. Hence, the equity risk premium calculated using only US data of this period is necessarily distorted since the calculation is done for a survivor.

In this paper it is examined if the equity premium puzzle is real and if it is robust under different conditions; the important in the present paper is that this examination is done using the SD rules applied on empirical distribution functions (EDFs) of stocks and bonds. Specifically, it is examined if the stocks' empirical distributions stochastically dominate bonds' empirical distributions in different markets, for different investment periods and under different economic conditions i.e. when economic crises are included in the analysis.

The using real empirical data are derived from several markets such as the international financial markets of UK, Germany and US. The entire time period under consideration is the 1980-2014 period. The recent period contains events such as the stock market crash in 1987, the Gulf war (1990), the peso crisis (1994), the Barings collapse (1995), the Asian financial crisis (1997), the collapse of LTCM hedge fund (1998), the Russian debt moratorium (1998), the stock market bubble (1995-2000), the subprime crisis and the European financial crisis, among others. The resulting question is how these crises may influence the stochastic dominance relation between stocks and bonds i.e. how the economic conditions may affect their stochastic dominance relation? Moreover, how this SD relation is formulated when it is employed the entire period and how when sub periods (which may contain an economic crisis) are put under consideration i.e. how the time horizon may influence the SD relation? Finally, does the same SD relation hold for every market or does it change from market to market i.e. does the different economic environment play a role in the SD relation's formulation? This paper tries to give answers to these questions using the SD rules. Briefly, it is examined the existence and the robustness of the equity premium puzzle in several economic environments, time horizons and economic conditions in a stochastic dominance framework.

6.2. **Data and Testing Methodology.** In empirical researches, the comparison between alternative investments in a Stochastic Dominance framework is traditionally based on observed historical data i.e. on discrete empirical distributions. The importance of using empirical distributions in decision making analysis was also pointed out by Bawa et al. (1979). From a Bayesian perspective, when true distributions are unknown the use of empirical distribution functions is necessary and completely justified by von Neumann - Morgenstern - Savage axioms. But it is also the dynamic of SD rules that imposes the use of discrete empirical distributions; using continuous distributions the SD criteria are reduced into a simple mean-variance rule based on only two moments, the mean and the variance; this is not a desirable situation since there are higher order's moments, like skewness and kurtosis that investors take into account during the portfolio selection process. In this paper, empirical data are used in order to compute stocks and bonds' historical monthly returns and generate each asset's empirical distribution function (without pre-assuming a specific distribution functional form i.e. normal or lognormal); it is also employed a resampling method in order to eliminate any correlation between bonds and stocks. The SD rules are applied on the empirical distributions resulting from the ex-post returns of the two assets under consideration (stocks and bonds).

In this empirical analysis, Matlab algorithms were developed and the practical algorithms of Levy (1992) and De Nadai and Pianca (2007) were adapted in order to examine the stochastic dominance (or not) of stocks over bonds. The data used in this study are a representative stock and a representative bond portfolio for three important international markets: the US (the S&P 500 Composite Index and the US Benchmark 10-Year DS Government Bond Index), the UK (the FTSE All Share Index and the UK Benchmark 10-Year DS Government Bond Index) and the Germany (the DAX 30 Performance Index and the BD Benchmark 10-Year DS Government Bond Index) for the period between January 1980 and April 2014 (411 monthly price observations). Returns are defined as the first difference of the log price levels.

Table 1 presents the **descriptive statistics for the sample data**

	S&P 500	US BOND	FTSE	UK BOND	DAX 30	GERMANY BOND
MEAN	0.0068	0.0010	0.0065	0.0019	0.0072	0.0011
VAR	0.0022	0.0006	0.0025	0.0006	0.0039	0.0003
STDEV	0.0473	0.0249	0.0497	0.0246	0.0627	0.0174
KURTOSIS	4.7577	3.2726	3.2142	2.3552	4.9883	1.4340
SKEWNESS	-1.2424	0.0535	-1.1114	-0.0057	-1.4219	-0.3721
SHARPE RATIO	0.1449	0.0419	0.1299	0.0782	0.1152	0.0660
MIN	-0.2732	-0.1282	-0.2407	-0.1054	-0.3201	-0.0699
MAX	0.1153	0.1127	0.1245	0.1068	0.1754	0.0616

Notes to Table 1: The table presents **descriptive statistics for representative stock and bond portfolia** for three important international markets, the US (the S&P 500 Composite Index and the US Benchmark 10-Year DS Government Bond Index), the UK (the FTSE All Share Index and the UK Benchmark 10-Year DS Government Bond Index), and Germany (the DAX 30 Performance Index and the BD Benchmark 10-Year DS Government Bond Index) for the period between January 1980 and April 2014 (411 monthly price observations).

Remark 25. *Note that the sample period is selected in order to contain important crises that generated significant volatility in financial markets, such as the DOT.com bubble (1995-2000), the subprime crisis (2006-2009), and the European financial crisis (2009-2013). In order to gain deeper insight into investor behavior during crisis the full sample is also divided into smaller 5-year sub-samples and repeat the empirical analysis both for the entire sample period and for each of these sub-periods separately, i.e. start with the full 411 months and then separate the period into 72- month, 60- month and 64-month sub-periods.*

6.2.1. ***Algorithms for the first order and the second order SD rules***. The SD relation between the bonds and the stocks is examined through the practical algorithms of Levy (1992) and De Nadai and Pianca (2007) (among others). The algorithms given by De Nadai and Pianca (2007) concerns the first order (SD1), the second order (SD2) and the second order for risk seeking agents (SD2*) SD rules and they are applied on the empirical distribution functions of bonds and stocks employed in this paper. Suppose $x_1 \leq x_2 \leq \cdots \leq x_n$ the ex-post rates of returns of stocks' index and $y_1 \leq y_2 \leq \cdots \leq y_n$ the ex-post rates of returns of bonds. SD1 and SD2 algorithms require an increasing ranking of all observation for both empirical distributions.

Algorithm 1. ***SD1 algorithm:*** *the empirical distribution of stocks dominates the empirical distribution of bonds at first order (SD1) if and only if* $y_i \leq x_i$ $\forall i = 1, 2, \ldots$ *with at least one strict inequality*

Algorithm 2. ***SD2 algorithm****: the empirical distribution of stocks dominates the empirical distribution of bonds at second order (SD2) if and only if* $Y_i \leq X_i$ $\forall i = 1, 2, \ldots$ *where* Y_i*and* X_i *are the cumulative observations of the empirical distributions i.e.*$Y_i = \sum_{j=1}^{i} y_j$ *with* $Y_1 = y_1$ *and* $Y_n = \sum_{j=1}^{n} y_j$ *and* $X_i = \sum_{j=1}^{i} x_j$ *with* $X_1 = x_1$ *and* $X_n = \sum_{j=1}^{n} x_j$ *with at least one strict inequality*

Algorithm 3. ***SD2* algorithm****: the empirical distribution of stocks dominates the empirical distribution of bonds at second order for risk seeking agents (SD2*)if and only if* $\bar{Y}_i \leq \bar{X}_i$ $\forall i = 1, 2, \ldots$ *where* $\bar{Y}_i$*and* $\bar{X}_i$ *are the cumulative observations of the empirical distributions i.e.* $\bar{Y}_i = y_n + y_{n-1} + \cdots + y_{n-i+1}$ *with* $\bar{Y}_1 = y_n$ *and* $\bar{Y}_n = y_n + y_{n-1} + \cdots + y_1$ *and* $\bar{X}_i = x_n + x_{n-1} + \cdots + x_{n-i+1}$ *with* $\bar{X}_1 = x_n$ *and* $X_n = x_n + x_{n-1} + \cdots + x_1$ *with at least one strict inequality.*

Remark 26. *The SD2 and the SD2* require the calculation of the cumulative observations of the empirical distributions which is graphically interpreted as the area enclosed between the two empirical distributions.*

7. Results of the Empirical Study

A first general result is that there is not a first order stochastic dominance (SD1) relation between the empirical distributions of bonds and stocks in any case (case of the whole period or of the sub-periods) since graphically there is not a cumulative distribution which is totally below or it is tangent to the other. So, the left tail condition of the SD1 is not satisfied by any empirical distribution in all cases. The two empirical distributions have a cross-point near to the reference point (x=0) that is graphically consistent with the SD2 and the $SD2^*$ rules. The analysis continues by testing the necessary and sufficient conditions for the SD2 and the $SD2^*$ relations (since SD1 relation does not exist).

The matrices in all cases show that the mean and the variance (the first two moments of the empirical distribution functions) of the stocks are greater than that of bonds. If the two assets had the same mean then the asset with the higher variance cannot be the stochastically dominant; this is not the case, so the stocks can be the dominant asset. Moreover, the minimum observed return of the stocks' empirical distribution is always smaller than that of the bonds' empirical distribution while the maximum observation of the stocks' empirical distribution is always greater than that of the bonds. Thus, the necessary conditions for the SD2 and the $SD2^*$ relations hold for all periods under consideration. The next step is to employ the algorithms given by De Nadai and Pianca (2007) in order to examine the SD2 and the $SD2^*$ relations between the two empirical distributions.

The results show that the area enclosed between the empirical distributions of stocks and bonds (cumulative stocks' returns $X_i \leqslant$ cumulative bonds' returns Y_i) is smaller than the area enclosed between the empirical distributions of bonds and stocks (cumulative stocks' returns $X_i \geqslant$ cumulative bonds' returns Y_i) in all periods; this practically means that the greater part of the stocks' empirical distribution is under the empirical distribution of bonds. This implies that the stocks stochastically dominate bonds at second order (SD2 and $SD2^*$). But since the SD2 relation implies any higher order stochastic dominance and the Prospect SD (PSD) as well, then stocks (blue line) stochastically dominate bonds (green line) at second order and at any higher order and in PSD sense ($SD2 \Rightarrow SD3 \Rightarrow SDi \Rightarrow PSD\ \forall i > 2$). This means that there is a quite strong stochastic dominance of stocks over the bonds which implies that the equity premium puzzle is real and it appears for any investment period and under different economic situations such as various economic crises (DOT.com Bubble, subprime crisis, EU crisis).

However, it is important to examine if this puzzle is robust from market to market i.e. if stocks stochastically dominate bonds for any period and under any economic condition not only for the US market but for other markets as well; two major international markets, the UK and the Germany market, are employed for this examination and the same process is repeated firstly for the whole period (1980-2014) and after for the (i) 1995-2000 sub-period (DOT.com Bubble) (ii) the 2005-2009 sub-period (subprime crisis) and (iii) the 2009-2014 (EU crisis).

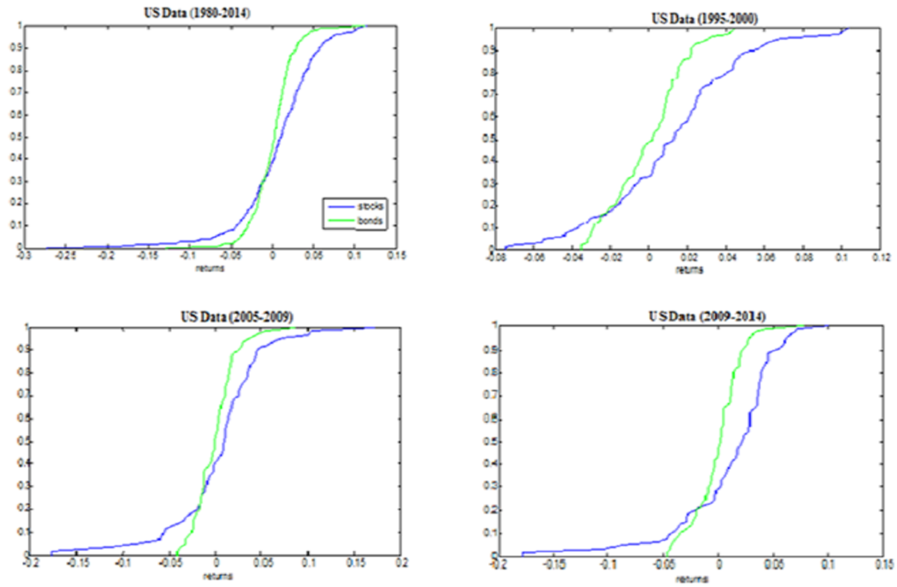
7.1. **US Data**

	S&P 500 COMPOSITE STOCKS INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0068	0,0142	0,0061	0,0117
VAR	0,0022	0,0014	0,0026	0,0020
MIN	-0,2732	-0,0747	-0,1771	-0,1790
MAX	0,1153	0,1042	0,1732	0,1008

	US BENCHMARK 10 YEAR DS GOVT. INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0010	0,0019	0,0005	-0,0001
VAR	0,0006	0,0004	0,0005	0,0005
MIN	-0,1282	-0,0350	-0,0409	-0,0468
MAX	0,1127	0,0448	0,0863	0,0776

Notes to Table 2: The table presents the mean and the variance (the first two moments of the empirical distribution functions), the minimum and the maximum observed return of a representative stock and bond portfolio for the US market (the S&P 500 Composite Index and the US Benchmark 10-Year DS Government Bond Index), for the whole period between January 1980 to April 2014 (411 monthly price observations), the1995-2000 sub-period (DOT.com Bubble), the 2005-2009 sub-period (subprime crisis) and the 2009-2014 sub-period (EU crisis).

FIGURE 1 (US market): This figure presents the empirical distribution functions of a representative stock and bond portfolio for the US market (the S&P 500 Composite Index and the US Benchmark 10-Year DS Government Bond Index), for the whole period between January 1980 to April 2014, the1995-2000 sub-period (DOT.com Bubble), the 2005-2009 sub-period (subprime crisis) and the 2009-2014 sub-period (EU crisis).



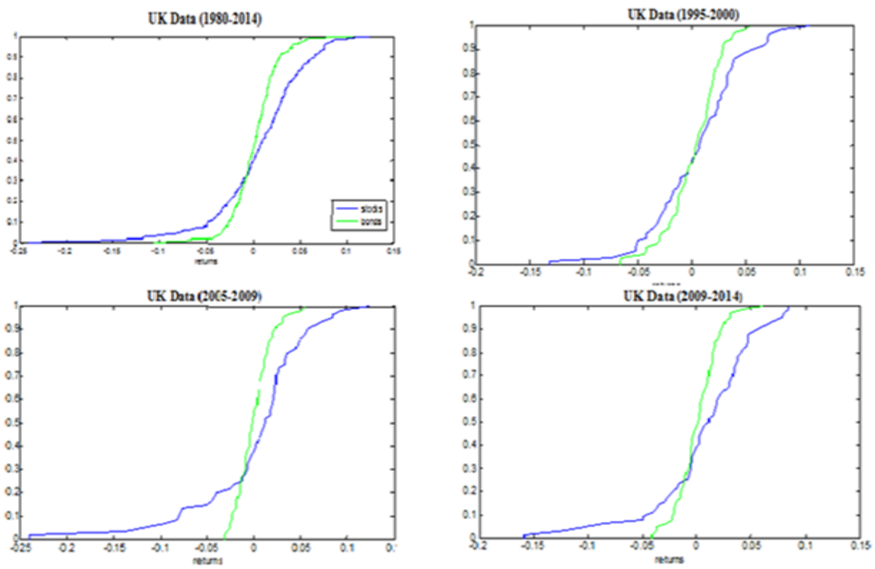
7.2. **Major International Markets: UK and Germany**

	FTSE ALL SHARE INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0065	0,0091	0,0022	0,0079
VAR	0,0025	0,0014	0,0036	0,0022
MIN	-0,2407	-0,1141	-0,2407	-0,1594
MAX	0,1245	0,0890	0,1245	0,0846

	UK BENCHMARK 10 YEAR DS GOVT. INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0019	0,0030	0,0010	0,0018
VAR	0,0006	0,0003	0,0003	0,0004
MIN	-0,1054	-0,0354	-0,0303	-0,0411
MAX	0,1068	0,0428	0,0561	0,0612

Notes to Table 3: The table presents mean, variance, min and max of a representative stock and a bond portfolio for the UK international market (FTSE All Share Index and UK Benchmark 10-Year DS Government Bond Index), for the whole period 1980-2014 (411 monthly observations), 1995-2000 sub-period (DOT.com Bubble), 2005-2009 sub-period (subprime crisis), 2009-2014 sub-period (EU crisis).

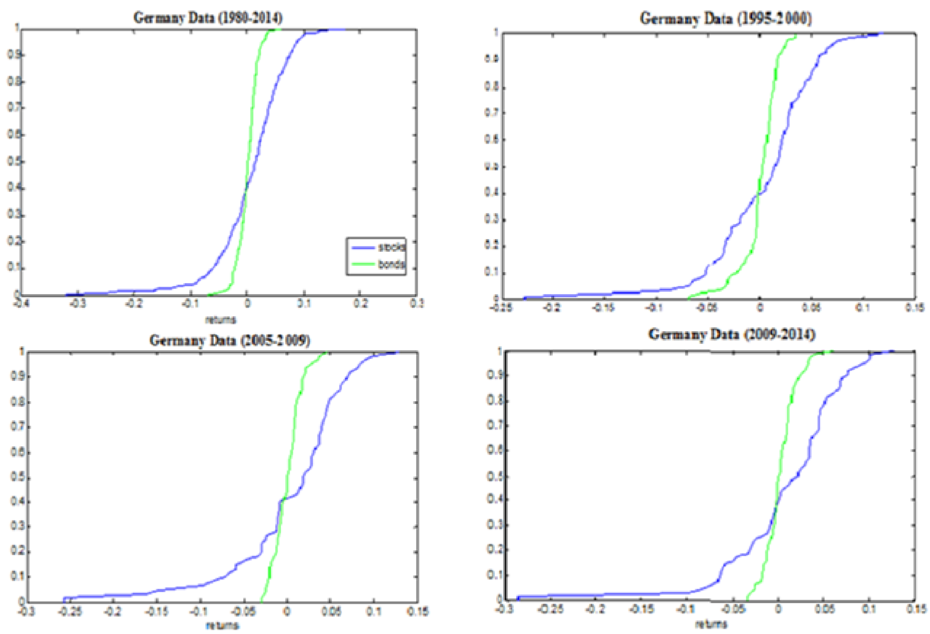
FIGURE 2 (UK market) : This figure presents the empirical distribution functions of a representative stock and bond portfolio for the UK market (the FTSE All Share Index and the UK Benchmark 10-Year DS Government Bond Index), for the whole period between January 1980 to April 2014, the1995-2000 sub-period (DOT.com Bubble), the 2005-2009 sub-period (subprime crisis) and the 2009-2014 sub-period (EU crisis).



	DAX 30 PERFORMANCE STOCKS INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0072	0,0151	0,0057	0,0109
VAR	0,0039	0,0028	0,0044	0,0039
MIN	-0,3201	-0,1774	-0,2575	-0,2854
MAX	0,1754	0,1268	0,1292	0,1272
	BD BENCHMARK 10 YEAR DS GOVT. INDEX			
	Period 1980-2014	Sub-Period 1995-2000	Sub-Period 2005-2009	Sub-Period 2009-2014
MEAN	0,0011	0,0022	0,0009	0,0029
VAR	0,0003	0,0002	0,0003	0,0003
MIN	-0,0699	-0,0432	-0,0289	-0,0342
MAX	0,0616	0,0323	0,0475	0,0616

Notes to Table 4: The table presents mean, variance, min and max of a representative stock and a bond portfolio for the Germany international market (the DAX 30 Performance Index and the BD Benchmark 10-Year DS Government Bond Index), for the whole period 1980-2014, 1995-2000 sub-period (the DOT.com Bubble), 2005-2009 sub-period (subprime crisis) and 2009-2014 sub-period (EU crisis).

FIGURE 3 (Germany market): This figure presents the empirical distribution functions of a representative stock and bond portfolio for the Germany international market (the DAX 30 Performance Index and the BD Benchmark 10-Year DS Government Bond Index), for the whole period between January 1980 to April 2014, the1995-2000 sub-period (DOT.com Bubble), the 2005-2009 sub-period (subprime crisis) and the 2009-2014 sub-period (EU crisis).



Conclusion 2. *The graphs and the matrices for the UK and the Germany markets have the same characteristics with those of the US market in all periods. Specifically, the empirical distributions of bonds and stocks have a cross-point which implies that there is not a first order stochastic dominance (SD1) relation between the two empirical distributions but they are graphically consistent with SD2 and SD2*.*

Conclusion 3. *The matrices in all cases show that the mean and the variance of the stocks are greater than the mean and the variance of bonds, the minimum return of the stocks' empirical distribution is always smaller than that of the bonds' empirical distribution while the maximum observation of the stocks' empirical distribution is greater than that of the bonds' empirical distribution.*

Conclusion 4. *The necessary conditions which concern the SD2 and the SD2* relations hold for the whole period and for all the sub-periods under consideration, for both markets. Once more, the above algorithms are used in order to examine the SD2 and the SD2* relations between the two empirical distributions. The results are the same with those of the US market in all periods i.e. the stocks (blue line) stochastically dominate bonds (green line) at second order (SD2 and SD2*) but also at any higher order and in PSD sense, for both markets the UK and the Germany market. This means that the equity premium puzzle is a real and a very robust puzzle which appears for any investment period, under different economic conditions and in any economic environment (US, UK or Germany market).*

8. Conclusion

This paper presents all the Stochastic Dominance (SD) rules (including Markowitz and Prospect SD Rule), situations relating to these rules such as arbitrage, and several classes of preferences consistent with these SD rules. They are also proposed necessary and sufficient conditions that could improve the effectiveness of each Stochastic Dominance rule reducing significantly the computational cost of the algorithm used for testing these SD relations. In the final part of the paper it is examined the existence of the well-known equity premium puzzle in several economic environments, under different economic conditions and varying time horizons. The analysis is developed in a Stochastic Dominance framework using original evidence from three markets: US, UK and Germany market. The results show that stocks stochastically dominate bonds at second order (SD2 and SD2*) and at any higher order and in PSD sense. This holds for any time period, under different economic conditions and for all economic environments (US, UK and Germany market). Hence, it is proved that the equity premium puzzle is real and quite robust since it holds in any market for any time horizon and economic situation. Another element that indicates the power of the equity premium puzzle is that stocks dominate bonds at second order (and at any higher order) which implies that stocks outperform bonds in a great percentage. In general, the Stochastic Dominance is a wide and useful framework within which one can develop any type of analysis and make comparisons without requirements for further specifications of functional types or parameter estimation.

References

[1] Abel, A.: Asset Prices under Habit Formation and Catching up with the Joneses, *The American Economic Review* **80(2)** (1990) 38-42.

[2] Ang, A., Bekaert, G. and Liu, J.: Why stocks may disappoint, *Journal of Financial Economics* **76(3)** (2005) 471–508.

[3] Baucells, M. and Heukamp, F.: Stochastic Dominance and Cumulative Prospect Theory, *Management Science* **52(9)** (2006) 1409-1423.

[4] Bawa, V.S., Lindenberg, E. and Rafsky L.: An algorithm to determine stochastic dominance admissible sets, *Management Science* **25(7)** (1979) 609–622.

[5] Benartzi, S. and Thaler, R.H.: Myopic Loss Aversion and the Equity Premium Puzzle, *The Quarterly Journal of Economics* **110(1)** (1995) 73-92.

[6] Constantinides, G.M.: Habit Formation: A Resolution of the Equity Premium Puzzle, *Journal of Political Economy* **98(3)** (1990) 519-543,.

[7] De Nadai, G. and Pianca, P.: Cumulative prospect theory and second order stochastic dominance criteria: an application to mutual funds performance, in: *Convegno: Statistics for funds management*, Working Paper **157** (2007) 7-18, Department of Applied Mathematics, University Ca'Foscari Venezia .

[8] Epstein, L.G. and Zin, S.E.: First-order risk aversion and the equity premium puzzle, *Journal of Monetary Economics* **26** (1990) 387-407.

[9] Gul, F., A.: Theory of Disappointment Aversion, *Econometrica* **59(3)** (1991) 667-686.

[10] Haigh, M.S. and List, J.A.: Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis, *The Journal of Finance* **60(1)** (2005) 523–534.

[11] Jarrow, R.: The Relationship between Arbitrage and First order Stochastic Dominance, *The Journal of Finance* **41** (1986) 915-921.

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

[13] Kandel, S. and Stambaugh, R.: Asset returns and intertemporal preferences, *Journal of Monetary Economics* **27(1)** (1991) 39–71.

[14] Kuosmanen, T: Efficient Diversification According to Stochastic Dominance Criteria, *Management Science* **50(10)** (2004) 1390–1406.
[15] Levy, H.: Stochastic Dominance and Expected Utility: Survey and Analysis, *Management Science* **38(4)** (1992) 555–593.
[16] Levy, H. and Wiener, Z.: Stochastic Dominance and Prospect Dominance with Subjective Weighting Functions, *Journal of Risk and Uncertainty* **16(2)** (1998) 147–163.
[17] Levy, M. and Levy, H.: Prospect Theory: Much Ado about Nothing?, *Management Science* **48(10)** (2002) 1334–1349.
[18] Longstaff, F.A. and Piazzesi, M.: Corporate earnings and the equity premium, *Journal of Financial Economics* **74(3)** (2004) 401–421.
[19] Mankiw, G. and Zeldes, S.: The consumption of stockholders and nonstockholders, *Journal of Financial Economics* **29(1)** (1991) 97-112.
[20] Mehra, R.: The Equity Premium Puzzle: A Review, *Foundations and Trends in Finance* **2(1)** (2006) 1–81.
[21] Mehra, R. and Prescott, E.C.: The Equity Premium: A Puzzle, *Journal of Monetary Economics* **15(2)** (1985) 145–161.
[22] Meyer, D.J. and Meyer, J.: Risk preferences in multi-period consumption models, the equity premium puzzle, and habit formation utility, *Journal of Monetary Economics* **52(8)** (2005) 1497–1515.
[23] Otrok, C. Ravikumar, B. and Whiteman, C.H.: Habit formation: a resolution of the equity premium puzzle?, *Journal of Monetary Economics* **49(6)** (2002) 1261–1288.
[24] Pirvu, T.A. and Schulze, K.: Multi-stock portfolio optimization under prospect theory, *Mathematics and Financial Economics* **6(4)** (2012) 337-362 .
[25] Rothschild, M. and Stiglitz, J.: Increasing risk: I. A definition, *Journal of Economic Theory* **2(3)** (1970) 225-243, .
[26] Siegel, J.J. and Thaler, R.H.: Anomalies: The Equity Premium Puzzle, *The Journal of Economic Perspectives* **11(1)** (1997) 191-200, .
[27] Siegel, J.J.: The Shrinking Equity Premium, *The Journal of Portfolio Management* **26(1)** (1999) 10-17, .
[28] Tversky, A. and Kahneman, D.: Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty* **5(4)** (1992) 297–323, .
[29] Von Neuman, J. and Morgenstern, O.: *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, (2nd edition) 1947, (3rd edition) 1953.
[30] Zervoudi, E. and Spyrou, S.: The Equity Premium Puzzle: New Evidence on the Optimal Holding Period and Optimal Asset Allocation, *Review of Behavioral Finance* **8(1)** (2016) 39-57.
[31] Zhou, C.: Informational Asymmetry and Market Imperfections: Another Solution to the Equity Premium Puzzle, *Journal of Financial and Quantitative Analysis* **34(4)** (1999) 445-464.

E. K. Zervoudi: Athens University of Economics and Business, Athens, 10434, Greece
E-mail address: `zervoudiev@aueb.gr`