Received: 4 March 2020 | Revised: 18 December 2020 | Accepted: 20 December 2020
DOI: 10.1002/ijfe.2444

RESEARCH ARTICLE

WILEY

Stochastic dominance algorithms with application to mutual fund performance evaluation

Sree Vinutha Venkataraman | **S. V. D. Nageswara Rao**

School of Management (SOM), IIT Bombay, Mumbai, Powai, India

Correspondence
Sree Vinutha Venkataraman, IIT Campus, House no. 333, Kanpur 208016, India.
Email: sreevibavana@yahoo.com

Abstract

While the possibility for investment A to dominate investment B under the first- and second-order stochastic dominance framework can be tested only at the points of jumps in the probabilities of the distributions, the comparison at interior points is also essential under third-order, due to the non-linearity in the difference in the third-order stochastic dominance integral. Furthermore, it was established that the quantile approach used to test for the first- and second-order dominance does not work under the third-order. If these points are overlooked, it is possible to conclude that an investment is third-order inefficient when it is actually efficient. Also, an inefficient investment may not be relegated to the inefficient set. In this work, to test for third-order efficiency, we derive the expressions for the functions essential for testing the possibility of third-order stochastic dominance at the interior points and arrive at their restrictions on the common grid of the pairwise investments under consideration. We also develop a program to determine the efficient funds and the funds superior and inferior to the benchmark indices at a fast pace. We find that the size of the efficient set reduces drastically under third-order. Also, several funds are found to be superior to the indices under second- and third-order.

KEYWORDS

first-order stochastic dominance algorithm, investment analysis, second-order stochastic dominance algorithm, third-order stochastic dominance algorithm

1 | INTRODUCTION

Information on investor preferences is modelled through assumptions underlying utility functions under different orders of stochastic dominance. The probability distribution of returns is examined to classify investments into two mutually exclusive categories: efficient and inefficient. While first-order stochastic dominance (FSD) assumes that investors prefer more to less, second-order stochastic dominance (SSD) additionally imposes the assumption of risk aversion and third-order stochastic dominance (TSD) further adds the preference for positive skewness.

Studies in the literature are dispersed between analysing optimal selection rules (Hanoch & Levy, 1969), implementing stochastic dominance tests (Jean & Helms Billy, 1986; Porter, Wart, & Ferguson, 1973), evaluating portfolio performance (Joy & Porter, 1974; Meyer, 1977), considering inclusion of riskless asset (Levy & Kroll, 1979), to the more technical aspects of deriving the mathematical relation between the geometric mean and the integrals of the PDF for a general class of distribution

(Jean, 1980), deriving the necessary conditions for dominance (Jean, 1984), testing the validity of the quantile approach (Ng, 2000), modifying the assumptions underlying utility functions (Levy, 2006) and instituting connections among risk measures and stochastic dominance (Ma & Wong, 2010).

The literature on second-order dominance has developed beyond the finite set of comparisons. Kuosmanen (2004), for instance, developed operational tests of portfolio efficiency based on general stochastic dominance criteria that account for infinite set of diversification strategies. Lizyayev and Ruszczyński (2012) discussed testing for portfolio efficiency relative to an infinite portfolio possibilities set.

Kopa and Peter (2008) introduced a linear programming SSD portfolio efficiency test. The authors also derived a necessary condition for SSD efficiency. A dominating portfolio was identified by solving linear programming problem. Dupačová and Kopa (2012) suggested a portfolio efficiency test with respect to SSD criteria. Hodder, Jackwerth, and Olga (2015) compared the performance of portfolio choices based on standard approaches and SSD related strategies. The authors found portfolio choices based on SSD to outperform choices based on standard approaches. Linton, Post, and Whang (2014) proposed a test of portfolio efficiency with respect to stochastic dominance criteria. Bruni, Cesarone, Scozzari, and Tardella (2017) proposed an approximate stochastic dominance rule to find portfolios that approximately dominate a given benchmark. Chen, Mei, and Liu (2019) discussed a stochastic optimization problem with multivariate SSD constraints and developed the necessary and sufficient optimality conditions. Kallio, Hardoroudi, and Dehghan (2018) introduced directional SSD where a portfolio was considered admissible if a step from the benchmark in the direction of the portfolio provided a dominating portfolio.

Several applications of stochastic dominance tests in the context of Finance and Economics are documented in Abhyankar, Keng-Yu, and Zhao (2009), Alkhazali, Lean, and Zoubid (2020), Bawa, Bodurtha, Rao, and Suri (1985), Chan, Clark, Guo, and Wong (2020), Fong, Wong, and Lean (2005), Gasbarro, Wong, and Zumwalt (2007), Hoang, Lean, and Wong (2015), Hoang, Wong, and Zhu (2015), Joy and Porter (1974), Lean, Smyth, and Wong (2007), Post (2003) and Seyhun (1993).

Though the quantile approach can be used to test for the first- and second-order stochastic dominance, Ng (2000), through counter-examples, established the inappropriateness in extending this approach to test for third-order dominance. While the quantile of F can be greater than that of G, F need not dominate G. Also, F can dominate G but the quantile of F can be lower. Furthermore, while it is sufficient to test for the first- and second-order stochastic dominance at points of jumps in probabilities, the comparison at interior points is also essential under third-order, due to the non-linearity in the difference in the third-order dominance integral. If this comparison is omitted, we may declare a fund as third-order inefficient when it is actually not so. These points were overlooked in some earlier studies.

In this work, we derive the functional representations of the second and third degree integrals, to test for the possibility of dominance at the interior points, and we arrive at the restrictions of the functions on the common grid of the pairwise investments under consideration. Furthermore, we develop a program in Matlab to determine the efficient set and the set of funds inferior or superior to benchmark share price indices under all three orders. The processing time is drastically reduced by utilizing the necessary conditions for dominance, along with the information that higher-order efficient sets are subsets of the lower-order efficient sets. The superior and inferior funds are obtained only from those pairs that contain the index as one of the pairs, and with possible dominance after the necessary checks, such that the dominance is not tested in the earlier testing for efficiency. These funds are subsequently integrated with the superior and inferior funds obtained during processing of efficient set. The relative effectiveness of the different orders of the algorithms in reducing the size of the efficient set is discussed.

While the different order algorithms can be applied to any portfolio in general, the third-order dominance is more useful in evaluating the performance of growth funds in particular. Given the objective of capital appreciation, investors in the growth option of the growth fund expect a small probability of making huge returns – the positive skewness criteria of the third-order.

2 | DATA

To illustrate the daily and monthly returns on 50 large cap and diversified equity schemes in India are calculated from the information on NAV, available either in the websites of the funds or the website of Bluechip India, (Historical NAV, Bluechip India) from September

2004 to April 2012. The returns on BSE Sensex and NSE S&P CNX Nifty are obtained from the closing values of the indices displayed in the corresponding websites (For S&P BSE SENSEX, bseindia, markets-historical data and for Nifty 50, nseindia, products-historical index data). Appendix A provides the list of funds and indices. Pairwise comparisons of investments are made to check for dominance under each order. The in-sample set comprises of 60% of the observations and the out of sample period ranges from May 2009 to April 2012.

3 | METHODOLOGY

That investors tend to enhance the utility of wealth rather than actual wealth itself is already documented. However, the functional form of the utility function of the investors' wealth is unknown. Therefore, to model investor preferences under different orders of dominance, the assumptions underlying utility functions are mapped to an equivalent framework pertaining to the more accessible cumulative distribution of returns.

The funds are classified into efficient and inefficient sets based on the definition of dominance.

3.1 | Dominance

Let U denote a continuous non-decreasing utility function, implying that it is differentiable. If U_1 is the set of all non-decreasing utility functions, investment A dominates investment B in U_1, if, for all utility functions $U \in U_1$, $E_AU \geq E_BU$, and for at least one utility function $U_0 \in U_1$, $E_AU_0 > E_BU_0$.

3.2 | Efficient set and inefficient set

Efficient set is the set of all investments that are not dominated by any other investment. Consider two investments A and B. If A does not dominate B and B does not dominate A, investments A and B are declared efficient. In other words, $E_AU_{11} > E_BU_{11}$for $U_{11} \in U_1$and $E_BU_{21} > E_AU_{21}$for $U_{21} \in U_1$, that is, while some investors may prefer A and some prefer B, neither A nor B dominate the other, since, for all utility functions $U \in U_1$, neither $E_AU \geq E_BU$ nor $E_BU \geq E_AU$ holds.

In general, a smaller efficient set in comparison to the feasible set of all investments is preferred. Based on their preferences, investors choose to invest from the pool of efficient investments.

An investment is termed inefficient if it is dominated by at least one investment in the efficient set. The set of all inefficient investments is contained in the inefficient set.

3.3 | First-order stochastic dominance

Investors' preference for more to less implies that as wealth increases, the utility of wealth also increases. Therefore, for $U \in U_1$, $U' \geq 0$ and there exists a range with $U' > 0$. Mapping this to an equivalent definition concerning cumulative distribution functions of returns for funds F and G, we say that:

Theorem 1. *F dominates G by FSD iff $F(x) \leq G(x)$ for all x and $F(x_0) < G(x_0)$ for at least one x_0.*

This is equivalent to saying that $E_FU \geq E_GU$ for all $U \in U_1$ and $E_FU_0 > E_GU_0$ for at least one $U_0 \in U_1$. F(x) must be less than G(x) because, in this case, the probability of obtaining x or a value greater than x is higher under F than G. The theorem can be proved by considering the difference in expected utilities and integrating by parts. For details and proof, refer Levy (2006).

The following necessary rules for dominance are adopted.

3.3.1 | Necessary conditions

Arithmetic, geometric and harmonic means

If fund F with cumulative distribution F(x) dominates fund G with cumulative distribution G(x), the expected value of the return series of F denoted by $E_F(x)$, the geometric mean of the return series of F denoted by $\bar{X}_{GM}(F)$ and harmonic mean of the return series of F denoted by $\bar{X}_{HM}(F)$ must be greater than that of fund G, that is, $E_F(x) > E_G(x)$, $\bar{X}_{GM}(F) > \bar{X}_{GM}(G)$ and $\bar{X}_{HM}(F) > \bar{X}_{HM}(G)$. Details and proof in Jean (1980, 1984) and Levy (2006). Since the geometric mean is defined only for positive numbers, we take x = 1+ r, where r is the return, that is, x is considered as the terminal wealth.

Left tail

If F dominates G, it is necessary for $\text{Min}_F(x) \geq \text{Min}_G(x)$. This implies that G has a thicker left tail (Jean & Helms Billy, 1986).

3.4 | Second-order stochastic dominance

SSD additionally assumes risk aversion. The utility function U is such that $U' \geq 0$, $U'' \leq 0$, and there is at least one point such that $U' > 0$ and one point such that $U'' < 0$. Checking for second-order dominance under cumulative distribution framework is the same as testing for positive difference in the area represented by the second-order integral.

Theorem 2. *Formally, F dominates G by SSD for all risk-averse investors iff* $\int_a^x [G(t) - F(t)]\, dt \geq 0$ *for all* $x \in [a,b]$ *and there is at least one* x_0 *for which the strict inequality holds true.*

This is equivalent to saying that $E_F U \geq E_G U$ for all $U \in U_2$ and $E_F U_0 > E_G U_0$ for at least one $U_0 \in U_2$, where U_2 is the set of all concave utility functions (Levy, 2006). Furthermore, $U_2 \subseteq U_1$.

The necessary rules for SSD are similar to that of FSD, save for the fact that the strong inequality condition is replaced by its more flexible counterpart. For instance, $E_F(x) \geq E_G(x)$ is the necessary condition in terms of arithmetic mean (Levy, 2006).

With SSD efficient set being a subset of FSD efficient set, it is sufficient to check for SSD efficiency on the FSD efficient sets alone.

3.5 | Third-order stochastic dominance

Higher skewness leads to higher expected utility, provided $U''' \geq 0$. To see this, expand $U(w + x)$ about $w + E(x)$, where w represents the initial wealth and x represents the random fluctuation in wealth, into a Taylor series expansion, and take the expected value on both sides. $U''' \geq 0$ will render positivity to the skewness term on the right hand side of the expansion. Adding the assumption of $U''' \geq 0$,

Theorem 3. *F dominates G by TSD if and only if* $\int_a^x \int_a^z [G(t) - F(t)]\, dt\, dz \geq 0$ *for all x and* $\int_a^b [G(x) - F(x)]\, dx \geq 0$ *with at least one strict inequality.*

This is equivalent to saying that $E_F U \geq E_G U$ for all $U \in U_3$ where $U' \geq 0$, $U'' \leq 0$ and $U''' \geq 0$. Details and proof in Levy (2006). The necessary rules under TSD follow that under SSD. Again, TSD efficient set being a subset of SSD efficient set, TSD efficiency check can be done on the SSD efficient set alone.

4 | ALGORITHMS FOR FSD, SSD AND TSD

The initial algorithm for FSD and SSD was developed by Hanoch and Levy (1969). Algorithms for necessary conditions were developed by Porter et al. (1973).

In this paper, the performance of the 50 funds and a benchmark share price index is examined. Each observation is assigned an equal probability. 2,550 permutations are possible because, for each pair of distributions F and G, one needs to check if F dominates G. If no dominance is established, the possibility of G dominating F is to be studied. However, on employing the necessary rule of arithmetic mean, the number of comparisons reduces to one-half or nC_2. Using this rule, if F is found to dominate G, it would be impossible for G to dominate F, since the superior fund must have a higher mean return. On these reduced observations, the geometric mean and subsequently the harmonic mean criteria are introduced. The number of comparisons reduces with the employment of each necessary condition. Finally, on the reduced observations, the left tail condition is enforced. This gives the final set on which FSD is performed. Similarly, necessary conditions of SSD and TSD are evoked before performing the SSD and TSD tests.

4.1 | First-order stochastic dominance

The cumulative distribution framework is easier to implement in terms of quantiles.

Let f and g stand for the rates of return on funds F and G. Assign an equal probability to each observation. Order the rates of returns on funds F and G in ascending order so that,

$$f_1 \leq f_2 \leq f_3 .... \leq f_n$$

and

$$g_1 \leq g_2 \leq g_3 .... \leq g_n$$

$f_i \geq g_i$ for all $i = 1, 2, 3, ...n \ \forall\ i$ implies that the pth quantile of F is more than the pth quantile of G. In this case, the cumulative distribution of F must fall below that of G because, if $f_j < g_j$ for some j, F and G must intercept in which case FSD will not hold.

Thus, F dominates G by FSD iff $f_i \geq g_i$ for all $i = 1, 2, 3, ...n$ and $f_{i_0} > g_{i_0}$ for $i = i_0$.

4.2 | Second-order stochastic dominance

To test if $\int_a^x [G(t) - F(t)]\,dt \geq 0$, it is sufficient to evaluate $\int_a^x [G(t) - F(t)]\,dt$ at points of intersection of F and G, for, if the integral is positive at x_1, an intersection point, it will also be positive for $x < x_1$, since by moving to the left of x_1, the negative area that forms the integral reduces.

Again, the quantile approach is a convenient form of testing for SSD.

Order the rates of returns on funds F and G in ascending order as in FSD. Now, for $\int_a^x [G(t) - F(t)]\,dt \geq 0$ to hold true, it is sufficient to check if the cumulative sum of returns on fund F is more than the corresponding sum on fund G.

Define f_i' as follows:

$$\begin{aligned} f_1' &= f_1 \\ f_2' &= f_1 + f_2 \\ &\vdots \\ f_n' &= \sum_{j=1}^{n} f_j. \end{aligned}$$

Similarly, g_i' is defined as follows:

$$\begin{aligned} g_1' &= g_1 \\ g_2' &= g_1 + g_2 \\ &\vdots \\ g_n' &= \sum_{j=1}^{n} g_j. \end{aligned}$$

F dominates G by SSD iff $f_i' \geq g_i'$ for all $i = 1, 2, 3, \ldots n$ and $f_{i_0}' > g_{i_0}'$ for $i = i_0$.

4.3 | Third-order stochastic dominance

The weakness in the quantile approach for third-order dominance (Theorem 4, Levy, 1992) is documented in Ng (2000). Through examples, Ng (2000) demonstrated that for F to dominate G under third-order, it is not necessary for quantile of F to be greater than quantile of G. Also, quantile of F greater than quantile of G is not identical to F dominating G under third-order. Neither implication is valid. Therefore, if F dominates G rendering G inefficient, the quantile approach may not relegate G to the inefficient set. Similarly, when quantile of F is higher than that of G, we may conclude F dominates G under third-order when it is actually not true. In this case, we may relegate G to the inefficient set when it could have been efficient.

While the difference in the FSD and SSD integrals is linear, the difference in the third-order integral as in Theorem 3 is non-linear. It may turn out to be positive at the probability jump points, but negative at interior points. Therefore, if we omit checking for dominance at the interior points, we would conclude that F dominates G when it is not true. If G never gets dominated in other comparisons, G should be placed in the efficient set. By ignoring dominance check at the interior points, we will be declaring a fund as inefficient when it may not be so. For TSD to hold true, the integral condition should be satisfied at both interior points and probability jump points.

Levy (2006) advocates cumulative distribution framework instead of the quantile approach to test for TSD.

Define,

$F_2(x) = \int_{-\infty}^{x} [F(t)]dt$ and $F_3(x) = \int_{-\infty}^{x} [F_2(t)]dt$ efficient, where F stands for the cumulative distribution of fund F.

$$F(x) = \begin{cases} 0 & ;\ x < x_1 \\ \frac{1}{n} & ;\ x_1 \leq x < x_2 \\ \frac{2}{n} & ;\ x_2 \leq x < x_3 \\ \vdots & \\ \vdots & \\ \frac{n-1}{n} & ;\ x_{n-1} \leq x < x_n \\ 1 & ;\ x_n \leq x \end{cases}.$$

G_2 and G_3 are defined similarly. Given F(x), we derive the expressions for $F_2(x)$ and $F_3(x)$ as follows:

$$F_2(x) = \begin{cases} 0 & ;\ x \leq x_1 \\ \frac{1}{n}(x\text{-}x_1) & ;\ x_1 \leq x \leq x_2 \\ \frac{2}{n}x\text{-}\frac{1}{n}(x_1 + x_2) & ;\ x_2 \leq x \leq x_3 \\ \vdots & \\ \vdots & \\ \frac{k}{n}x\text{-}\frac{1}{n}\left(\sum_{i=1}^{k} x_i\right) & ;\ x_k \leq x \leq x_{k+1} \\ x\text{-}\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) & ;\ x_n \leq x \end{cases}.$$

$$F_3(x)=\begin{cases}0 & ;\ x\le x_1\\[4pt] \frac{1}{2n}(x\text{-}x_1)^2 & ;\ x_1\le x\le x_2\\[4pt] \underbrace{\frac{k}{2n}\left(x^2\text{-}x_k{}^2\right)}_{k=2}\text{-}\underbrace{\frac{1}{n}\left(\sum_{i=1}^{k}x_i\right)(x\text{-}x_k)}_{k=2}+\underbrace{\frac{1}{2n}(x_2\text{-}x_1)^2}_{F_3(x_2)} & ;\ x_2\le x\le x_3\\[4pt] \underbrace{\frac{k}{2n}\left(x^2\text{-}x_k{}^2\right)}_{k=3}\text{-}\underbrace{\frac{1}{n}\left(\sum_{i=1}^{k}x_i\right)(x\text{-}x_k)}_{k=3}+\underbrace{\underbrace{\frac{k}{2n}\left(x_3{}^2\text{-}x_k{}^2\right)}_{k=2}\text{-}\underbrace{\frac{1}{n}\left(\sum_{i=1}^{k}x_i\right)(x_3\text{-}x_k)}_{k=2}+\frac{1}{2n}(x_2\text{-}x_1)^2}_{F_3(x_3)} & ;\ x_3\le x\le x_4\\ \vdots & \\ \underbrace{\frac{k}{2n}\left(x^2\text{-}x_k{}^2\right)}_{k=n}\text{-}\underbrace{\frac{1}{n}\left(\sum_{i=1}^{k}x_i\right)(x\text{-}x_k)}_{k=n}+\underbrace{\underbrace{\frac{k}{2n}\left(x_n{}^2\text{-}x_k{}^2\right)}_{k=n-1}\text{-}\underbrace{\frac{1}{n}\left(\sum_{i=1}^{k}x_i\right)(x_n\text{-}x_k)}_{k=n-1}+\ldots.+\frac{1}{2n}(x_2\text{-}x_1)^2}_{F_3(x_n)} & ;\ x_n\le x\end{cases}$$

The subscripts 2 and 3 denote the second- and third-order dominance integrals.

A fund with cumulative distribution F dominates a fund with cumulative distribution G by TSD if for all $-\infty < x < \infty$,

$F_3(x) \le G_3(x)$ and

$F_3(x_0) < G_3(x_0)$ for $x = x_0$ and

$E_F(x) \ge E_G(x)$.

Denote the returns on fund F and G by x and y. The returns are sorted so that,

$x_1 \le x_2 \le .... \le x_n$ and $y_1 \le y_2 \le .... \le y_n$.

Choose $z_k = x_i$ for some i and $z_k = y_j$ for some j so that $z_1,...z_{2n}$ is the common framework for x and y. For every k, $[z_k, z_{k+1}] \subseteq [x_i, x_{i+1}]$ and $[z_k, z_{k+1}] \subseteq [y_j, y_{j+1}]$.

In this work, for every $z_k \le z \le z_{k+1}$, we arrive at F(z) as the restriction of F(x) from $[x_i, x_{i+1}]$ on $[z_k, z_{k+1}]$ and G(z) as the restriction of G(y) from $[y_j, y_{j+1}]$ on $[z_k, z_{k+1}]$. We have derived the expressions for $F_2(x)$ and $F_3(x)$ in Appendix B.

We illustrate this concept through the following:

Illustration:

For example, say $[z_4, z_5] \subseteq [x_1, x_2]$ and $[z_4, z_5] \subseteq [y_3, y_4]$. We have,

$F(x) = \frac{1}{n}; x_1 \le x < x_2$.

Therefore, $F(z) = \frac{1}{n}; z_4 \le z < z_5$.

$G(y) = \frac{3}{n}; y_3 \le y < y_4$.

Therefore, $G(z) = \frac{3}{n}; z_4 \le z < z_5$.

Similarly, for every $z_k \le z \le z_{k+1}$, $F_2(z)$ is the restriction of $F_2(x)$ from $[x_i, x_{i+1}]$ on $[z_k, z_{k+1}]$, and $G_2(z)$ is the restriction of $G_2(y)$ from $[y_j, y_{j+1}]$ on $[z_k, z_{k+1}]$. Continuing our previous example of $[z_4, z_5] \subseteq [x_1, x_2]$ and $[z_4, z_5] \subseteq [y_3, y_4]$, we have the following:

$F_2(x) = \frac{1}{n}(x - x_1); x_1 \le x \le x_2$.

Therefore, $F_2(z) = \frac{1}{n}(z - x_1); z_4 \le z \le z_5$.

$G_2(y) = \frac{3}{n}y - \frac{1}{n}\sum_{i=1}^{3} y_i; y_3 \le y \le y_4$.

Therefore, $G_2(z) = \frac{3}{n}z - \frac{1}{n}\sum_{i=1}^{3} y_i; z_4 \le z \le z_5$.

Finally, for every $z_k \le z \le z_{k+1}$, $F_3(z)$ is the restriction of $F_3(x)$ from $[x_i, x_{i+1}]$ on $[z_k, z_{k+1}]$ and $G_3(z)$ is the restriction of $G_3(y)$ from $[y_j, y_{j+1}]$ on $[z_k, z_{k+1}]$. Again, continuing our previous example of $[z_4, z_5] \subseteq [x_1, x_2]$ and $[z_4, z_5] \subseteq [y_3, y_4]$, we have.

$F_3(x) = \frac{(x-x_1)^2}{2n}; x_1 \le x \le x_2$.

Therefore, $F_3(z) = \frac{(z-x_1)^2}{2n}; z_4 \le z \le z_5$.

$G_3(y) = G_3(y_3) + \underbrace{\frac{k}{2n}\left(y^2 - y_k{}^2\right)}_{k=3} - \underbrace{\frac{1}{n}\left(\sum_{i=1}^{k} y_i\right)(y - y_k)}_{k=3}$; $y_3 \le y \le y_4$.

Therefore,

$G_3(z) = G_3(y_3) + \underbrace{\frac{k}{2n}\left(z^2 - y_k{}^2\right)}_{k=3} - \underbrace{\frac{1}{n}\left(\sum_{i=1}^{k} y_i\right)(z - y_k)}_{k=3}$; $z_4 \le z \le z_5$.

To determine if F dominates G (assume that G cannot dominate F after the necessary condition checks), define H(x) as $G_3(x)$ - $F_3(x)$. H(x), being the difference of parabolas $F_3(x)$ and $G_3(x)$, is itself a parabola or the segment of a parabola.

For $z_k \le z \le z_{k+1}$, H(z) is a parabola. H(z) is positive under the following scenarios (Levy, 2006):

Case 1. Since parabola has either one minimum or maximum point, if H(z) is an increasing function at the end point z_k, $H'(z_k) \ge 0$. Now, H(z) is ≥0 for all z such that $z_k \le z \le z_{k+1}$.

Case 2. If H(z) is decreasing at the end points z_k, z_{k+1}, both $H'(z_k)$ and $H'(z_{k+1}) \le 0$, in which case, $H(z_k) \ge 0$ and $H(z_{k+1}) \ge 0$ imply $H(z) \ge 0$ for all z such that $z_k \le z \le z_{k+1}$.

Case 3. When H(z) is increasing at z_k and decreasing at z_{k+1}, H(z) attains maximum. Again, $H(z_k) \ge 0$ and $H(z_{k+1}) \ge 0$ imply $H(z) \ge 0$ for all z such that $z_k \le z \le z_{k+1}$.

Case 4. If H(z) is decreasing at z_k but increasing at z_{k+1}, the conditions $H(z_k) \ge 0$ and $H(z_{k+1}) \ge 0$ are not sufficient to conclude that $H(z) \ge 0$ in $z_k \le z \le z_{k+1}$ since the minimum value in the interval (z_k, z_{k+1}) may be a suspect point. This possibility was ignored in the earlier reported algorithm. The recent algorithm for TSD in Levy (2006) checks the value of H(z) also in its minimum point in the entire interval in addition to the end points, as against the earlier algorithm in which the value of H(z) was checked only at z_k, z_{k+1}. Rather than checking the interior points in the interval, it is enough to check if $H(z) \ge 0$ at the minimum point, for positivity at minimum implies positivity in $[z_k, z_{k+1}]$. The minimum point is obtained by equating the first derivative of H(z) to zero. H(z) in the standard form is expressed as $az^2 + bz + c$. $H'(z)$ is zero when $z = \frac{-b}{2a}$. $H\left(\frac{-b}{2a}\right) < 0$ implies there is no TSD.

Hence, for $z \in [z_k, z_{k+1}]$,
$H(z) \ge 0$ iff
$H(z_k) \ge 0$
$H(z_{k+1}) \ge 0$
and
$H'(z_k) \ge 0$ (when H(z) is increasing at end point z_k) or
$H'(z_k) \le 0$ and $H'(z_{k+1}) \le 0$ (when H(z) is decreasing at the end points z_k, z_{k+1})or
$H'(z_k) \ge 0$ and $H'(z_{k+1}) \le 0$ (when H(z) is increasing at end point z_k but decreasing at z_{k+1}) or
$H'(z_k) < 0$ and $H'(z_{k+1}) \ge 0$ (when H(z) decreases at z_k but increases at z_{k+1}) and $H\left(\frac{-b}{2a}\right) \ge 0$.

On completion of necessary checks, the following algorithm is implemented to test if F dominates G by third-order:

Compute $F_2(x_k)$, $G_2(y_k)$, $F_3(x_k)$, $G_3(y_k)$ where $k = 1,2....n$.

Sort the values of z_i, where z_i is some x_i or y_j, that is, $z_1 \le z_2 \le .... \le z_{2n}$.

Compute F(z), G(z), $F_2(z)$, $G_2(z)$, $F_3(z)$, $G_3(z)$ as the restriction explained earlier.

If $F_3(z) \le G_3(z)$, $z \in \{x_i, y_i: i = 1,2,....n\}$ (to check the integral and probability jump points) and.

if for some $k \le 2n-1$, $H'(z_k) = G_2(z_k) - F_2(z_k) < 0$ and $H'(z_{k+1}) = G_2(z_{k+1}) - F_2(z_{k+1}) \ge 0$, where i and j be such that $z_k \le y_j \le z_{k+1}$ and $z_k \le x_i \le z_{k+1}$.

Compute $\frac{-b}{2a}$ as follows:

$$\frac{-b}{2a} = z_k - \frac{H'(z_k)(z_{k+1} - z_k)}{H'(z_{k+1}) - H'(z_k)}.$$

$$H\left(\frac{-b}{2a}\right) = G_3\left(\frac{-b}{2a}\right) - F_3\left(\frac{-b}{2a}\right).$$

Check if $H\left(\frac{-b}{2a}\right) \ge 0$.

If $H(z_k) \ge 0$ for $k = 1,2,..2n$ with at least one strict inequality and $H\left(\frac{-b}{2a}\right) \ge 0$, F dominates G by TSD.

5 | ANALYSIS

Upon application of necessary conditions for FSD dominance, the number of pairwise comparisons reduced from 2,550 to the values in Tables 1 and 2. The results are reported for in-sample (IS) and out-sample (OS) period using daily and monthly frequencies.

Introducing Sensex as the benchmark index, FSD test had to be performed only on 741 combinations as against 2,550 for IS period using monthly frequency. Likewise, the number of combinations to be tested reduced drastically on the employment of necessary conditions for other sample periods and frequencies. All funds are

TABLE 1 Application of necessary conditions (NC) using monthly data

NC	Sensex-IS	S&P CNX Nifty-IS	Sensex-OS	S&P CNX Nifty-OS
AM	1,275	1,275	1,275	1,275
GM	1,204	1,204	1,224	1,223
HM	1,133	1,129	1,172	1,171
LT	741	739	748	743

Abbreviations: AM, Arithmetic mean; GM, Geometric mean; HM, Harmonic mean; LT, Left tail.

TABLE 2 Application of necessary conditions (NC) using daily data

NC	Sensex-IS	S&P CNX Nifty-IS	Sensex-OS	S&P CNX Nifty-OS
AM	1,275	1,275	1,275	1,275
GM	1,176	1,176	1,248	1,249
HM	1,110	1,107	1,220	1,223
LT	714	716	824	827

Abbreviations: AM, Arithmetic mean; GM, Geometric mean; HM, Harmonic mean; LT, Left tail.

TABLE 3 Number of dominance using SSD and TSD

SSD	Sensex-IS	S&P CNX Nifty-IS	Sensex-OS	S&P CNX Nifty-OS
Monthly	533	530	518	515
Daily	487	496	617	619
TSD	**Sensex-IS**	**S&P CNX Nifty-IS**	**Sensex-OS**	**S&P CNX Nifty-OS**
Monthly	5	5	0	0
Daily	8	8	0	0

Abbreviations: SSD, second-order stochastic dominance; TSD, third-order stochastic dominance.

TABLE 4 Percent of superior, inferior and no dominance funds with respect to the indices

	Superior funds %	Inferior funds %	No dominance %
Monthly			
SSD			
IS-Sensex	20	22	58
IS-S&P CNX Nifty	28	8	64
OS-Sensex	52	2	46
OS-S&P CNX Nifty	46	2	52
TSD			
IS-Sensex	8	0	92
IS-S&P CNX Nifty	12	0	88
OS-Sensex	2	0	98
OS-S&P CNX Nifty	4	0	96
Daily			
SSD			
IS-Sensex	12	2	86
IS-S&P CNX Nifty	30	2	68
OS-Sensex	62	2	36
OS-S&P CNX Nifty	66	2	32
TSD			
IS-Sensex	0	14	86
IS-S&P CNX Nifty	0	8	92
OS-Sensex	4	2	94
OS-S&P CNX Nifty	4	2	94

Abbreviations: SSD, second-order stochastic dominance; TSD, third-order stochastic dominance.

found to be efficient under FSD, whereas a much reduced SSD efficient set is visible. Under monthly frequency, only 14 combinations had to be tested for TSD while 41 and 39 combinations had to be tested for TSD using different benchmark indices with daily frequency of observations. Table 3 provides the number of dominance under SSD and TSD.

The entire program including the necessary checks for dominance, the first, second and third degree dominance and the list of superior and inferior funds under all three

TABLE 5 Superior funds with respect to index

FSD superior funds—Nil		
SSD superior funds		
SSD	*Sensex-IS*	*S&P CNX Nifty-IS*
Monthly	BARODA PIONEER GROWTH—G	BIRLA SL FRONTLINE EQ PLAN A–G
	DSP BR TOP 100 EQUITY	BARODA PIONEER GROWTH—G
	FRANK INDIA PRIMA PLUS	DWS ALPHA EQUITY—G
	FRANK INDIA BLUECHIP G	DSP BR TOP 100 EQUITY
	HDFC GROWTH	FRANK INDIA PRIMA PLUS
	HDFC TOP 200—G	FRANK INDIA BLUECHIP G
	HSBC EQUITY	HDFC EQUITY G
	ICICI PRU. TOP 100	HDFC GROWTH
	KOTAK 50—G	HDFC TOP 200—G
	TATA PURE EQUITY	HSBC EQUITY
		ICICI PRU. TOP 100
		KOTAK 50—G
		SBI MAGNUM MULTIPLIER PLUS
		TATA PURE EQUITY
	Sensex-OS	*S&P CNX Nifty-OS*
	BIRLA SL FRONTLINE EQ. PLAN A–G	BIRLA SL FRONTLINE EQ. PLAN A–G
	CANARA ROBECO EQUITY DIV.	CANARA ROBECO EQUITY DIV.
	BNP PARIBAS EQUITY—G	DSP BR TOP 100 EQUITY
	DSP BR TOP 100 EQUITY	DSP BR OPPO. FUND
	DSP BR OPPO. FUND	FRANK INDIA PRIMA PLUS
	FRANK INDIA PRIMA PLUS	FRANK INDIA BLUECHIP G
	FRANK INDIA BLUECHIP G	HDFC CAPITAL BUILDER
	HDFC CAPITAL BUILDER	HDFC GROWTH
	HDFC EQUITY G	HDFC TOP 200—G
	HDFC GROWTH	HSBC INDIA OPPO.
	HDFC TOP 200—G	ICICI PRU. TOP 100
	HSBC INDIA OPPO.	ING CORE EQUITY—G
	ICICI PRU. TOP 100	ING LC EQUITY—G
	ING CORE EQUITY—G	KOTAK 50—G
	ING LC EQUITY—G	L&T GROWTH
	KOTAK 50—G	SBI MAGNUM EQUITY—G
	L&T GROWTH	TATA EQUITY PE FUND G
	SBI MAGNUM EQUITY—G	TATA PURE EQUITY
	TATA EQUITY PE FUND G	TAURUS STARSHARE
	TATA GROWTH	UTI EQUITY
	TATA PURE EQUITY	UTIMNC
	TAURUS STARSHARE	UTI MASTER PLUS
	UTI EQUITY	UTIUNIT SCHEME 1986 MS
	UTIMNC	
	UTI MASTER PLUS	
	UTIUNIT SCHEME 1986 MS	
	Sensex-IS	*S&P CNX Nifty-IS*
Daily	CANARA ROBECO EQUITY DIV.	CANARA ROBECO EQUITY DIV.
	FRANK INDIA PRIMA PLUS	BARODA PIONEER GROWTH—G
	FRANK INDIA BLUECHIP G	DWS ALPHA EQUITY—G
	HDFC EQUITY G	FRANK INDIA PRIMA PLUS
	HDFC GROWTH	FRANK INDIA BLUECHIP G
	HDFC TOP 200—G	TEMPLETON INDIA GROWTH
		HDFC EQUITY G
		HDFC GROWTH
		HDFC TOP 200—G
		HSBC EQUITY
		ICICI PRU. TOP 100
		ICICI PRU. TOP 200
		KOTAK 50—G
		RELIANCE VISION
		SBI MAGNUM MULTIPLIER PLUS

(Continues)

TABLE 5 (Continued)

FSD superior funds—Nil		
	Sensex-OS	*S&P CNX Nifty-OS*
	BIRLA SL EQUITY FUND—G	BIRLA SL EQUITY FUND—G
	BIRLA SL FRONTLINE EQ. PLAN A–G	BIRLA SL FRONTLINE EQ. PLAN A–G
	CANARA ROBECO EQUITY DIV.	CANARA ROBECO EQUITY DIV.
	BNP PARIBAS EQUITY—G	BNP PARIBAS EQUITY—G
	DSP BR TOP 100 EQUITY	DSP BR TOP 100 EQUITY
	DSP BR OPPO. FUND	DSP BR OPPO. FUND
	FRANK INDIA PRIMA PLUS	FRANK INDIA PRIMA PLUS
	FRANK INDIA BLUECHIP G	FRANK INDIA BLUECHIP G
	TEMPLETON INDIA GROWTH	TEMPLETON INDIA GROWTH
	HDFC CAPITAL BUILDER	HDFC CAPITAL BUILDER
	HDFC EQUITY G	HDFC EQUITY G
	HDFC GROWTH	HDFC GROWTH
	HDFC TOP 200—G	HDFC TOP 200—G
	HSBC INDIA OPPO.	HSBC INDIA OPPO.
	ICICI PRU. TOP 100	ICICI PRU. TOP 100
	ICICI PRU. TOP 200	ICICI PRU. TOP 200
	ING CORE EQUITY—G	ING CORE EQUITY—G
	KOTAK 50—G	ING LC EQUITY—G
	MORGAN STANLEY GROWTH—G	KOTAK OPPO.—G
	RELIANCE VISION	KOTAK 50—G
	RELIANCE GROWTH	MORGAN STANLEY GROWTH—G
	SBI MAGNUM MULTIPLIER PLUS	RELIANCE VISION
	SBI MAGNUM EQUITY—G	RELIANCE GROWTH
	TATA EQUITY OPPO. G	SBI MAGNUM MULTIPLIER PLUS
	TATA GROWTH	SBI MAGNUM EQUITY—G
	TATA PURE EQUITY	TATA EQUITY OPPO. G
	TAURUS STARSHARE	TATA GROWTH
	UTI EQUITY	TATA PURE EQUITY
	UTIMNC	TAURUS STARSHARE
	UTI MASTER PLUS	UTI EQUITY
	UTIUNIT SCHEME 1986 MS	UTIMNC
		UTI MASTER PLUS
		UTIUNIT SCHEME 1986 MS
TSD superior funds		
TSD	***Sensex-IS***	***S&P CNX Nifty-IS***
Monthly	DWS ALPHA EQUITY—G	BNP PARIBAS EQUITY—G
	DSP BR OPPO. FUND	DSP BR OPPO. FUND
	SBI MAGNUM MULTIPLIER PLUS	TEMPLETON INDIA GROWTH
	SUNDARAM SELECT FOCUS	ING CORE EQUITY—G
		RELIANCE VISION
		SUNDARAM SELECT FOCUS
	Sensex-OS	*S&P CNX Nifty-OS*
	SUNDARAM INDIA LEADERSHIP	TEMPLETON INDIA GROWTH
		LIC N MF GROWTH FUND—G
Daily	*Sensex-IS*	*S&P CNX Nifty-IS*
	Nil	Nil
	Sensex-OS	*S&P CNX Nifty-OS*
	KOTAK OPPO.—G	TATA EQUITY PE FUND G
	TATA EQUITY PE FUND G	TAURUS BONANZA

Abbreviations: FSD, first-order stochastic dominance; SSD, second-order stochastic dominance; TSD, third-order stochastic dominance; SL, sun life; G, growth option; N, nomura; EQ, equity; BR, blackrock; OPPO, opportunities; PRU, prudential; LC, large cap; MS, mastershare.

orders takes about 4.60–4.67 s for IS monthly data and about 2.5–2.7 s for OS monthly data. For the daily observations, the program takes around 37.9–38.8 min for IS period and runs in around 8.14–8.74 min for the OS period.

6 | RESULTS AND DISCUSSIONS

We find that, under FSD, the efficient set is the same as the feasible set. Levy and Kroll (1979) also documented a similar result in their analysis of performance of mutual funds in comparison to Fisher Index during 1953–1974. Out of the total population of 73 mutual funds, none of them were found to be dominated by the other under FSD. Seyhun (1993) documented that, during 1926–1991, January return corresponding to the smallest decile of NYSE firms dominated the January return of other deciles by FSD. Also, the January returns in other deciles, save for the ninth and tenth decile, dominated the returns for other months under FSD and SSD. Using Asian stock market data from 1988 to 2002, Lean et al. (2007) found day-of-the-week effect to be prevalent in some markets. In Malaysia, investors preferred investment from Friday to Monday, as Friday returns dominated Monday returns by the three orders. Except for Indonesia and Taiwan, Monday returns are dominated by at least one of the other weekday returns. Tuesday returns dominated Monday returns by FSD for Hong Long, Japan and Singapore. Several such instances of FSD dominance are documented.

Under FSD, none of the funds are dominated by (inferior) or dominate (superior) the index. Table 4 gives the percentage of superior, inferior and no dominant funds for both SSD and TSD. Under SSD, the percentage of no dominance funds is more in the IS than OS period under both frequencies. Furthermore, this percentage is consistently high under TSD. Fong et al. (2005) found that winner portfolios dominated loser portfolios under both SSD and TSD during 1989–2001. Gasbarro et al. (2007) analysed the performance of iShares during 1996–2003 and documented that Spain (EWP) dominated Japan (EWJ) under FSD during March 1996–June 1998 and Spain dominated Hong Kong (EWH) under SSD. Lean et al. (2007) found that Monday returns are dominated by Thursday returns in Japan under SSD and TSD and by Wednesday returns in Malaysia and Singapore. Also, Monday returns are dominated by all weekday returns for Taiwan under SSD and TSD. Abhyankar et al. (2009) reported that value stocks dominate growth stock for US, Canada and Japan. Hoang, Lean, and Wong (2015) studied French portfolios from 1949 to 2012 and concluded that stock portfolios with gold dominated those without gold under SSD and TSD and bond portfolios without gold dominated the ones with gold. Furthermore, stock portfolios in London also showed similar results to that of Paris. In another study, Alkhazali, Lean, and Zoubid (2020). Using spot prices of gold and oil from January 1986–June 2018 found that gold–oil portfolio stochastically dominated the one without gold by SSD and TSD.

Tables 5 and 6 list the funds superior and inferior to the indices. Table 7 provides the percentage of efficient funds while Table 8 lists the efficient funds. Both Sensex and S&P CNX Nifty belong to the SSD and TSD inefficient sets. Furthermore, both the indices dominate only the same fund during the OS period while the dominance did not exist in the IS period under SSD. While no inferior funds are common to the IS and OS periods under SSD, several superior funds in the IS period continue to demonstrate superiority in the OS period. The set of inferior funds under TSD is the null set using monthly frequency. Also, only one fund superior to S&P CNX Nifty in the IS under TSD continues to retain superiority in the OS period. While there is variation in the IS and OS results, the efficient set is the same across indices. Post (2003) using value-weighted average of NYSE, AMEX and NASDAQ stocks found this market portfolio to be inefficient under SSD during July 1963–October 2001. Chan et al. (2020) found that while third-order risk-averse investor preferred S&P 500 index to NASDAQ-100 index, they were indifferent between the two indices in the bear market during the financial crisis and in the bull market during dotcom bubble and post-2008 financial crisis.

Under SSD, only few funds are not common to daily and monthly frequencies, suggesting that the frequency of data does not play a major role in deciding on the efficient set. On the contrary, under TSD only few funds are common, suggesting that efficient set varies based on the frequency of observations. Using 2004–2014 data, and Hoang, Wong, and Zhu (2015) found that in most cases, for risk-averse investors, non-gold Chinese portfolios dominate the ones with gold under second and/or third-order using both daily and monthly data and concluded that frequency mattered only for few pairs.

With few exceptions, several funds have performed better than the indices under SSD and TSD, advocating a case for active fund management. Joy and Porter (1974) report that the Dow Jones Industrial Average stochastically dominated six mutual funds under SSD during 1954–1963. We find that funds do not remain inferior to indices for long under SSD. Inferior funds in the OS period were not inferior in the IS period under both SSD and TSD. While few inferior funds in the IS period gained superiority to indices in the OS period under SSD, none of the superior funds became inferior in the latter period across indices under both SSD and TSD. Some degree of consistency in the performance of superior funds is envisaged. Furthermore, the efficient fund in the OS was also efficient in the IS period. While the persistence in the performance of inferior funds fails to exist under SSD, TSD consistently reports empty set of inferior funds using monthly frequency.

Bawa et al. (1985) studied 896 distributions of monthly equity returns from CRSP from January 1960 to December 1965 under their first, second-order and super

TABLE 6 Inferior funds with respect to index

FSD inferior funds—Nil		
SSD inferior funds	***Sensex-IS***	***S&P CNX Nifty-IS***
SSD		
Monthly	BIRLA SL ADV. FUND—G HDFC CAPITAL BUILDER L & T OPPO. FUND CUM L&T GROWTH LIC N MF GROWTH FUND—G LIC N MF EQUITY FUND—G MORGAN STANLEY GROWTH FUND—G PRINCIPAL GROWTH TATA EQUITY OPPO. G TATA GROWTH TAURUS BONANZA	BIRLA SL ADV. FUND—G L&T GROWTH LIC N MF EQUITY FUND—G TAURUS BONANZA
	Sensex-OS	*S&P CNX Nifty-OS*
	UTI TOP 100	UTI TOP 100
Daily		
	Sensex-IS	*S&P CNX Nifty-IS*
	LIC N MF EQUITY FUND—G	LIC N MF EQUITY FUND—G
	Sensex-OS	*S&P CNX Nifty-OS*
	UTI TOP 100	UTI TOP 100
TSD inferior funds		
TSD	***Sensex-IS***	***S&P CNX Nifty-IS***
Monthly	Nil	Nil
	Sensex-OS	*S&P CNX Nifty-OS*
	Nil	Nil
Daily		
	Sensex-IS	*S&P CNX Nifty-IS*
	ESCORTS GROWTH PLAN—G L & T OPPO. FUND CUM L&T GROWTH LIC N MF GROWTH FUND—G MORGAN STANLEY GROWTH FUND G SBI MAGNUM EQUITY—G TAURUS BONANZA	ESCORTS GROWTH PLAN—G LIC N MF GROWTH FUND—G MORGAN STANLEY GROWTH FUND—G SBI MAGNUM EQUITY—G
	Sensex-OS	*S&P CNX Nifty-OS*
	SUNDARAM SELECT FOCUS	SUNDARAM SELECT FOCUS

Abbreviations: FSD, first-order stochastic dominance; SSD, second-order stochastic dominance; TSD, third-order stochastic dominance; SL, sun life; G, growth option; ADV, advantage; N, nomura; OPPO, opportunities; PRU, prudential; LC, large cap; MS, mastershare.

TABLE 7 Number of funds in the efficient set

Efficient set	Sensex-IS	S&P CNX Nifty-IS	Sensex-OS	S&P CNX Nifty- OS
SSD				
Monthly	10 (20%)	10 (20%)	1 (2%)	1 (2%)
Daily	11 (22%)	11 (22%)	1 (2%)	1 (2%)
TSD				
Monthly	7 (14%)	7 (14%)	1 (2%)	1 (2%)
Daily	7 (14%)	7 (14%)	1 (2%)	1 (2%)

Abbreviations: SSD, second-order stochastic dominance; TSD, third-order stochastic dominance.

TABLE 8 List of funds in the efficient set

All funds are FSD efficient.		
SSD efficient funds	***Sensex-IS***	***S&P CNX Nifty-IS***
Monthly	BIRLA SL FRONTLINE EQ. PLAN A–G DSP BR TOP 100 EQUITY HDFC GROWTH HDFC TOP 200—G HSBC EQUITY RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC UTIUNIT SCHEME 1986 MS	BIRLA SL FRONTLINE EQ. PLAN A–G DSP BR TOP 100 EQUITY HDFC GROWTH HDFC TOP 200—G HSBC EQUITY RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC UTIUNIT SCHEME 1986 MS
	Sensex-OS	*S&P CNX Nifty-OS*
	UTIMNC	UTIMNC
Daily		
	Sensex-IS	*S&P CNX Nifty-IS*
	BIRLA SL FRONTLINE EQ. PLAN A–G FRANK INDIA PRIMA PLUS FRANK INDIA BLUECHIP G HDFC CAPITAL BUILDER HDFC EQUITY G HDFC GROWTH HDFC TOP 200—G RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC	BIRLA SL FRONTLINE EQ. PLAN A–G FRANK INDIA PRIMA PLUS FRANK INDIA BLUECHIP G HDFC CAPITAL BUILDER HDFC EQUITY G HDFC GROWTH HDFC TOP 200—G RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC
	Sensex-OS	*S&P CNX Nifty-OS*
	UTIMNC	UTIMNC
TSD efficient set		
TSD	*Sensex-IS*	*S&P CNX Nifty-IS*
Monthly	DSP BR TOP 100 EQUITY RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC UTIUNIT SCHEME 1986 MS HSBC EQUITY	DSP BR TOP 100 EQUITY RELIANCE GROWTH SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC UTIUNIT SCHEME 1986 MS HSBC EQUITY
	Sensex-OS	*S&P CNX Nifty-OS*
	UTIMNC	UTIMNC
Daily		
	Sensex-IS	*S&P CNX Nifty-IS*
	BIRLA SL FRONTLINE EQ. PLAN A–G FRANK INDIA PRIMA PLUS HDFC GROWTH HDFC TOP 200—G SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC	BIRLA SL FRONTLINE EQ. PLAN A–G FRANK INDIA PRIMA PLUS HDFC GROWTH HDFC TOP 200—G SBI MAGNUM MULTIPLIER PLUS UTI EQUITY UTIMNC
	Sensex-OS	*S&P CNX Nifty-OS*
	UTIMNC	UTIMNC

Abbreviations: FSD, first-order stochastic dominance; SSD, second-order stochastic dominance; TSD, third-order stochastic dominance; SL, sun life; G, growth option; EQ, equity; BR, blackrock; MS, mastershare.

third-order convex stochastic dominance procedures and found that the optimal set reduced to 454, 25 and 13, respectively.

In this work, with a smaller TSD efficient set relative to SSD efficient set in the IS period, the TSD algorithm has managed to reduce the size of the efficient set. Finally, the percentage of funds in the efficient set reduced from 100% under FSD to 20% under SSD and 14% under TSD in the IS period.

7 | CONCLUSIONS AND FUTURE SCOPE

In this work, to test for TSD at the interior points in addition to the probability jump points for the distributions of returns of the pairwise investments, we derive the functional forms and arrive at their restrictions on the common grid of the pairwise investments under consideration. Furthermore, we develop a program in Matlab to determine the efficient funds and funds superior and inferior to the benchmark indices under first-, second- and third-order stochastic dominance framework. All funds are found to be FSD efficient. Several funds are found to be superior to the indices under SSD and TSD, advocating a case for active fund management. While few inferior funds in the IS period gained superiority to the indices in the OS period under SSD, none of the superior funds became inferior in the OS period across indices under both SSD and TSD. In comparison to the FSD algorithm, SSD algorithm drastically reduces the size of the efficient set while TSD algorithm provides further reduction. The TSD algorithm gives a reasonable size of efficient set and allows investors to choose comparatively efficient funds.

It would be interesting to take this research forward and integrate portfolio optimization technique in this framework. Given the set of mutual funds, we have classified them as FSD, SSD and TSD efficient funds. We may also want to construct mutual funds in such a way that they are TSD efficient. Furthermore, inclusion of risk-free asset can change the efficient set under the three orders. Development of stochastic dominance criteria to address the case of the borrowing rate being higher than the lending rate and the borrowing rate being an increasing function of the borrowed amount would be useful.

Construction of the efficient sets under the three orders for risk-seekers is another interesting exercise. We may want to consider stochastic dominance analysis for utility functions that are concave in some regions and convex in others. Also, construction of the efficient sets using a conditional stochastic dominance analysis, given past information till the previous time period, is another future scope. Studying the effect of the holding period on the efficient set is also important. Moreover, an uncertain holding period may impose additional challenge to the construction of the efficient sets under FSD, SSD and TSD.

ACKNOWLEDGEMENT

We thank the anonymous reviewer for the useful comments and valuable suggestions.

DATA AVAILABILITY STATEMENT

Data openly available in a public repository that does not issue DOIs. The data that support the findings of this study are openly available in the following sources: (a) Bluechipindia website-Mutual Fund, Historical NAV at https://bluechipindia.co.in/MutualFund/MFInner.aspx?id=2, reference number (Historical NAV). (b) Bseindia website, Markets-historical data at https://www.bseindia.com/Indices/IndexArchiveData.html, reference number (S&P BSE SENSEX). (c) Nseindia website, Products-historical index data at https://www1.nseindia.com/products/content/equities/indices/historical_index_data.htm, reference number (Nifty 50). The citation to the data has been provided in the reference list.

ORCID

Sree Vinutha Venkataraman https://orcid.org/0000-0002-1958-4611

REFERENCES

Abhyankar, A., Keng-Yu, H., & Zhao, H. (2009). International value versus growth: Evidence from stochastic dominance analysis. *International Journal of Finance & Economics*, *14*, 222–232.

Alkhazali, O. M., Lean, H. H., Mirzaei, A., & Zoubid, T. (2020). A comparison of the gold-oil portfolio and oil portfolio: A stochastic dominance approach. *Finance Research Letters*, *2020*, 101670. https://doi.org/10.1016/j.frl.2020.101670

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

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

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. https://doi.org/10.1057/s41283-019-00057-9

Chen, Z., Mei, Y., & Liu, J. (2019). Multivariate robust second-order stochastic dominance and resulting risk-averse optimization. *Optimization*, *68*, 1719–1747.

Dupačová, J., & Kopa, M. (2012). Robustness in stochastic programs with risk constraints. *Annals of Operations Research*, *200*, 55–74.

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

Gasbarro, D., Wong, W.-K., & Zumwalt, K. J. (2007). Stochastic dominance analysis of iShares. *The European Journal of Finance*, *13*, 89–101. https://doi.org/10.1080/13518470601025243

Hanoch, G., & Levy, H. (1969). The efficient analysis of choice involving risk. *Review of Economic Studies*, *36*, 335–346.

Historical NAV, Bluechip India. Historical NAV. Retrieved from https://bluechipindia.co.in/MutualFund/MFInner.aspx?id=2, n.d.

Hoang, T.-H.-V., Lean, H. H., & Wong, W.-K. (2015). Is gold good for portfolio diversification? A stochastic dominance analysis of the Paris stock exchange. *International Review of Financial Analysis*, *42*, 98–108.

Hoang, T.-H.-V., 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.

Hodder, J.,. E., Jackwerth, C. J., & Olga, K. (2015). Improved portfolio choice using second-order stochastic dominance. *Review of Finance*, *19*, 1623–1647.

Jean, W.,. H. (1980). The geometric mean and stochastic dominance. *The Journal of Finance*, *35*, 151–158.

Jean, W.,. H. (1984). The harmonic mean and other necessary conditions for stochastic dominance. *The Journal of Finance*, *39*, 527–534.

Jean, W.,. H., & Helms Billy, P. (1986). Stochastic dominance as a decision model. *Quarterly Journal of Business and Economics*, *25*, 65–101.

Joy, M. O., & Porter, B. R. (1974). Stochastic dominance and mutual fund performance. *The Journal of Financial and Quantitative Analysis*, *9*, 25–31.

Kallio, M., Hardoroudi, H., & Dehghan, N. (2018). Second-order stochastic dominance constrained portfolio optimization: Theory and computational tests. *European Journal of Operational Research*, *264*, 675–685.

Kopa, M., & Peter, C. (2008). A second-order stochastic dominance portfolio efficiency measure. *Kybernetika*, *44*, 243–258.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, *50*, 1390–1406.

Lean, H. H., Smyth, R., & Wong, W.-K. (2007). Revisiting calendar anomalies in Asian stock markets using a stochastic dominance approach. *Journal of Multinational Financial Management*, *17*, 125–141.

Levy, H. (1992). Stochastic dominance and expected utility: Survey and analysis. *Management Science*, *38*, 555–593.

Levy, H. (2006). *Stochastic dominance. Investment decision making under uncertainty* (2nd ed.). New York, NY: Springer.

Levy, H., & Kroll, Y. (1979). Efficiency analysis with borrowing and lending: Criteria and their effectiveness. *The Review of Economics and Statistics*, *61*, 125–130.

Linton, O., Post, T., & Whang, Y.-J. (2014). Testing for the stochastic dominance efficiency of a given portfolio. *Econometrics Journal*, *17*, 59–74.

Lizyayev, A., & Ruszczyński, A. (2012). Tractable almost stochastic dominance. *European Journal of Operational Research*, *218*, 448–455.

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*, 927–935.

Meyer, J. (1977). Further applications of stochastic dominance to mutual fund performance. *The Journal of Financial and Quantitative Analysis*, *12*, 235–242.

Ng, M.-C. (2000). A remark on third degree stochastic dominance. *Management Science*, *46*, 870–873.

Nifty 50, Nseindia; Products-historical index data. Retrieved from https://www1.nseindia.com/products/content/equities/indices/historical_index_data.htm.n.d.

Porter, B. R., Wart, J. R., & Ferguson, D. L. (1973). Efficient algorithms for conducting stochastic dominance tests on large numbers of portfolios. *The Journal of Financial and Quantitative Analysis*, *8*, 71–81.

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

S&P BSE SENSEX, Bseindia; Markets-historical data. Retrieved from; https://www.bseindia.com/Indices/IndexArchiveData.html.n.d.

Seyhun, H. (1993). Can omitted risk factors explain the January effect? A stochastic dominance approach. *The Journal of Financial and Quantitative Analysis*, *28*, 195–212. https://doi.org/10.2307/2331286

How to cite this article: Venkataraman SV, Rao SVDN. Stochastic dominance algorithms with application to mutual fund performance evaluation. *Int J Fin Econ*. 2021;1–18. https://doi.org/10.1002/ijfe.2444

APPENDIX A. LIST OF FUNDS AND INDICES

List of funds	
BARODA PIONEER GROWTH—G	L&T OPPO. FUND CUMULATIVE
BIRLA SL ADV. FUND—G	L&T GROWTH
BIRLA SL EQUITY FUND—G	LIC N MF EQUITY FUND—G
BIRLA SL FRONTLINE EQ. PLAN A–G	LIC N MF GROWTH FUND—G
BNP PARIBAS EQUITY—G	MORGAN STANLEY GROWTH FUND—G
CANARA ROBECO EQUITY DIVERSIFIED	PRINCIPAL GROWTH
DSP BR OPPO. FUND	RELIANCE GROWTH
DSP BR TOP 100 EQUITY	RELIANCE VISION
DWS ALPHA EQUITY—G	SBI MAGNUM EQUITY—G
ESCORTS GROWTH PLAN—G	SBI MAGNUM MULTIPLIER PLUS
FRANK INDIA BLUECHIP G	SUNDARAM GROWTH
FRANK INDIA OPPO. FUND	SUNDARAM INDIA LEADERSHIP
FRANK INDIA PRIMA PLUS	SUNDARAM SELECT FOCUS
HDFC CAPITAL BUILDER	TATA EQUITY OPPO. G
HDFC EQUITY G	TATA EQUITY PE FUND G
HDFC GROWTH	TATA GROWTH
HDFC TOP 200—G	TATA PURE EQUITY
HSBC EQUITY	TAURUS BONANZA
HSBC INDIA OPPO.	TAURUS STARSHARE
ICICI PRU. TOP 100	TEMPLETON INDIA GROWTH
ICICI PRU. TOP 200	UTI EQUITY
ING CORE EQUITY—G	UTI MASTER PLUS
ING LC EQUITY—G	UTIMNC
KOTAK 50—G	UTI TOP 100
KOTAK OPPO.—G	UTIUNIT SCHEME 1986 MS
List of indices	
SENSEX	
S&P CNX NIFTY	

APPENDIX B. DERIVATION FOR THE EXPRESSIONS OF $F_2(X)$ AND $F_3(X)$

Derivation for $F_2(x)$

$F_2(x) = \int_{-\infty}^{x} [F(t)]dt$

$$F(x) = \begin{cases} 0 & ; x < x_1 \\ \frac{1}{n} & ; x_1 \le x < x_2 \\ \frac{2}{n} & ; x_2 \le x < x_3 \\ \vdots & \\ \frac{n-1}{n} & ; x_{n\text{-}1} \le x < x_n \\ 1 & ; x_n \le x \end{cases}.$$

When $x < x_1$, $F(x) = 0$. Therefore, $F_2(x) = 0$.
When $x_1 \le x < x_2$, $F(x) = \frac{1}{n}$.

$F_2(x) = \int_{-\infty}^{x} [F(t)]dt = \int_{-\infty}^{x_1} F(t)dt + \int_{x_1}^{x} F(t)dt = 0 + \int_{x_1}^{x} \frac{1}{n}dt = \frac{1}{n}(x - x_1)$.

Therefore, $F_2(x) = \frac{1}{n}(x - x_1)$.
Consider $x_2 \le x < x_3$.
$F(x) = \frac{2}{n}$.

$$F_2(x) = \int_{-\infty}^{x_2} F(t)dt + \int_{x_2}^{x} F(t)dt$$

$$\int_{-\infty}^{x} F(t)dt = \frac{1}{n}(x - x_1).$$

Therefore, $\int_{-\infty}^{x_2} F(t)dt = \frac{1}{n}(x_2 - x_1)$.

$$\int_{x_2}^{x} F(t)dt = \int_{x_2}^{x} \frac{2}{n}dt = \frac{2}{n}(x - x_2).$$

Hence, $\int_{-\infty}^{x} F(t)dt$ for $x_2 \le x < x_3 = \frac{1}{n}(x_2 - x_1) + \frac{2}{n}(x - x_2) = \frac{2}{n}x - \frac{1}{n}(x_1 + x_2)$ and so $F_2(x) = \frac{2}{n}x - \frac{1}{n}(x_1 + x_2)$.

Consider $x_3 \le x < x_4$.
$F(x) = \frac{3}{n}$.

$$F_2(x) = \int_{-\infty}^{x_3} F(t)dt + \int_{x_3}^{x} F(t)dt$$

But $\int_{-\infty}^{x_3} F(t)dt = \frac{2}{n}x_3 - \frac{1}{n}(x_1 + x_2)$.

$F_2(x) = \frac{2}{n}x_3 - \frac{1}{n}(x_1 + x_2) + \int_{x_3}^{x} \frac{3}{n}dt = \frac{2}{n}x_3 - \frac{1}{n}(x_1 + x_2) + \frac{3}{n}(x - x_3)$.

Therefore, $F_2(x) = \frac{3}{n}x - \frac{1}{n}\sum_{i=1}^{3} x_i$.

Continuing this argument, we get for $x_k \le x \le x_{k+1}$.

$F_2(x) = \frac{k}{n}x-\frac{1}{n}\sum_{i=1}^{k} x_i$ and for $x_{n-1} \le x \le x_n F_2(x) = \left(\frac{n-1}{n}\right) x-\frac{1}{n}\sum_{i=1}^{n-1} x_i$.

Finally, for $x_n \le x$,

$$F_2(x) = \int_{-\infty}^{x_n} F(t)dt + \int_{x_n}^{x} F(t)dt$$

$$\int_{-\infty}^{x_n} F(t)dt = \left(\frac{n-1}{n}\right)x_n - \frac{1}{n}(x_1+x_2+....x_{n-1}).$$

$$\int_{x_n}^{x} F(t)dt = \int_{x_n}^{x} dt = x-x_n.$$

$F_2(x) = \left(\frac{n-1}{n}\right)x_n - \frac{1}{n}(x_1+x_2+....x_{n-1}) + (x-x_n)$. Thus, $F_2(x) = x-\frac{1}{n}\sum_{i=1}^{n} x_i$.

Derivation for $F_3(x)$

$F_3(x) = \int_{-\infty}^{x} [F_2(t)]dt.$

$$F_2(x) = \begin{cases} 0 & ; x \le x_1 \\ \frac{1}{n}(x-x_1) & ; x_1 \le x \le x_2 \\ \frac{2}{n}x-\frac{1}{n}(x_1+x_2) & ; x_2 \le x \le x_3 \\ \vdots & \\ \vdots & \\ \frac{k}{n}x-\frac{1}{n}\left(\sum_{i=1}^{k} x_i\right) & ; x_k \le x \le x_{k+1} \\ x-\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) & ; x_n \le x \end{cases}.$$

Since $F_2(x) = 0$, when $x \le x_1$, $F_3(x)$ is also 0 when $x \le x_1$.
Consider, $x_1 \le x \le x_2$.
$F_2(x) = \frac{1}{n}(x-x_1)$.

$F_3(x) = \int_{-\infty}^{x_1} F_2(t)dt + \int_{x_1}^{x} F_2(t)dt = 0 + \int_{x_1}^{x} \frac{1}{n}(t-x_1)dt = \frac{1}{n}\left[\frac{x^2}{2}+\frac{{x_1}^2}{2}-x_1x\right] = \frac{1}{2n}(x-x_1)^2.$

Therefore, $F_3(x) = \frac{(x-x_1)^2}{2n}$.
Consider $x_2 \le x \le x_3$.

$F_3(x) = \int_{-\infty}^{x_2} F_2(t)dt + \int_{x_2}^{x} F_2(t)dt$

But $\int_{-\infty}^{x_2} F_2(t)dt = \frac{1}{2n}(x_2-x_1)^2$ and $\int_{x_2}^{x} F_2(t)dt = \int_{x_2}^{x} \left[\frac{2}{n}t-\frac{1}{n}(x_1+x_2)\right]dt.$

Therefore, $F_3(x) = \frac{1}{2n}(x_2-x_1)^2 + \frac{2}{n}\left[\frac{1}{2}(x^2-{x_2}^2)\right] - \frac{1}{n}(x_1+x_2)(x-x_2).$

This implies $F_3(x) = \underbrace{\frac{k}{2n}(x^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=2}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x-x_k)}_{\downarrow\, k=2} + \underbrace{\frac{1}{2n}(x_2-x_1)^2}_{\Downarrow\, F_3(x_2)}.$

Consider $x_3 \le x \le x_4$.

$F_3(x) = \int_{-\infty}^{x_3} F_2(t)dt + \int_{x_3}^{x} F_2(t)dt = F_3(x_2) + \underbrace{\frac{k}{2n}({x_3}^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=2}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x_3-x_k)}_{\downarrow\, k=2} + \int_{x_3}^{x} F_2(t)dt.$

$\int_{x_3}^{x} F_2(t)dt = \int_{x_3}^{x}\left[\frac{3t}{n}-\frac{1}{n}(x_1+x_2+x_3)\right]dt = \frac{3}{n}\left[\frac{1}{2}(x^2-{x_3}^2)\right] - \frac{1}{n}(x_1+x_2+x_3)(x-x_3) = \underbrace{\frac{k}{2n}(x^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=3}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x-x_k)}_{\downarrow\, k=3}.$

Therefore, $F_3(x) = \underbrace{\frac{k}{2n}(x^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=3}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x-x_k)}_{\downarrow\, k=3} + \underbrace{\underbrace{\frac{k}{2n}({x_3}^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=2}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x_3-x_k)}_{\downarrow\, k=2} + \frac{1}{2n}(x_2-x_1)^2}_{\Downarrow\, F_3(x_3)}.$

Proceeding on similar lines, we get, for $x_{n-1} \le x \le x_n$,

$$F_3(x) = \underbrace{\frac{k}{2n}(x^2-{x_k}^2)-\frac{1}{n}\left(\sum_{i=1}^{k} x_i\right)(x-x_k)}_{\downarrow\, k=n-1} + \underbrace{\frac{k}{2n}({x_{n-1}}^2-{x_k}^2)-\frac{1}{n}\left(\sum_{i=1}^{k} x_i\right)(x_{n-1}-x_k)}_{\downarrow\, k=n-2} + \ldots + \underbrace{\frac{k}{2n}({x_3}^2-{x_k}^2)-\frac{1}{n}}_{\downarrow\, k=2}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x_3-x_k)}_{\downarrow\, k=2} + \frac{(x_2-x_1)^2}{2n}.$$

Finally, for $x_n \le x$.

$F_3(x) = \int_{-\infty}^{x_n} F_2(t)dt + \int_{x_n}^{x} F_2(t)dt = F_3(x_{n-1}) + \frac{k}{2n}({x_n}^2-{x_k}^2) - \frac{1}{n}(x_1+x_2+....x_{n-1})(x_n-x_k)$ with $k = n-1$.
$\int_{x_n}^{x} F_2(t)dt = \int_{x_n}^{x}\left[t-\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\right]dt = \frac{1}{2}[x^2-{x_n}^2] - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)(x-x_n).$

Therefore, $F_3(x) = F_3(x_n) + \frac{1}{2}[x^2 - x_n{}^2] - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)(x - x_n) = \underbrace{\underbrace{\frac{k}{2n}(x^2\text{-}x_k{}^2)}_{k=n}\text{-}\frac{1}{n}\left(\sum_{i=1}^{k} x_i\right)(x\text{-}x_k)}_{k}$

$$= n + \underbrace{\underbrace{\frac{k}{2n}(x_n{}^2\text{-}x_k{}^2)}_{k=n-1}\text{-}\frac{1}{n}\underbrace{\left(\sum_{i=1}^{k} x_i\right)(x_n\text{-}x_k)}_{k=n-1} + \ldots + \frac{1}{2n}(x_2\text{-}x_1)^2}_{F_3(x_n)}.$$