THE JOURNAL OF FINANCE • VOL. XXXV, NO. 1 • MARCH 1980

# The Geometric Mean and Stochastic Dominance

WILLIAM H. JEAN*

TWO APPROACHES TO THE choice among risky alternatives that have been developed independently over the past twenty years are the geometric mean criterion and the stochastic dominance decision models. Both can be justified by the expected utility hypothesis, with the geometric mean criterion following as a result of the assumption that the decision-maker has a logarithmic utility function, and the stochastic dominance models requiring the less restrictive assumption of the signs of the first few derivatives of the decision-maker's utility function. In this paper the mathematical relationship between the geometric mean and the integrals of the probability density functions used in stochastic dominance test is derived for a very general class of distributions, and the principal result of the mathematical comparison is that a geometric mean ranking is a necessary condition for a stochastic dominance ranking.

Stochastic dominance showed great promise in the early years of its development as a method for constructing a theory of security prices because of the limited set of assumptions necessary regarding investor preferences. The evolution of a stochastic dominance-based theory of security prices has been stalled because of the frequency with which a no-ranking result occurs, and the lack of efficient algorithms to construct optimum portfolios. Algorithms to construct optimum geometric mean portfolios have been developed and, as a result of the derivations of this paper, those algorithms can be used in the exploration of the optimum stochastic dominance portfolio.

Latané [9] justified the geometric mean criterion on the basis of maximization of terminal wealth in a many-period reinvestment process, the optimal growth approach, and also pointed out that the criterion maximizes expected utility if the Bernoulli [1] logarithmic utility function is assumed. The growth optimal portfolio has been studied extensively since then, with an example being Markowitz's [16] derivation of the asymptotic characteristics of the model when smaller numbers of periods are used. The concept of stochastic dominance has been used for many years, but modern development of the approach is usually credited to Quirk and Saposnik [17], Hadar and Russell [4], Hanoch and Levy [7], and Whitmore [19].

In the case of continuous distributions of returns stochastic dominance ranking models use repeated integrals of the probability density function $f(x)$ to order alternatives. In this paper those repeated integrals are

$$F(x) = \int^{x} f(t)\, dt$$

* The University of Alabama.

$$F_1(x) = \int^x F(t)\,dt$$

$$\text{and} \quad F_k(x) = \int^x F_{k-1}(t)\,dt.$$

When two alternatives exist the probability density function of the first is $f(x)$ and the second is $h(x)$. Dominance conditions will always be stated so the first alternative is preferred to the second. Preference in this paper is always to be understood as weak preference.

Using the condition on the decision-maker's utility function that the first derivative is everywhere non-negative, the necessary and sufficient condition for first degree stochastic dominance (FSD) of the first alternative over the second is

$$F(x) \leq H(x) \qquad \text{for all } x.$$

For second degree stochastic dominance (SSD) a second condition on the utility function is added, that the second derivative is everywhere nonpositive. The necessary and sufficient condition for SSD is

$$F_1(x) \leq H_1(x) \qquad \text{for all } x.$$

With third degree stochastic dominance (TSD) the condition is added that the third derivative of the utility function is everywhere non-negative. Two conditions are necessary and sufficient for TSD, the first that the arithmetic means be ordered

$$E_f(x) \geq E_h(x)$$

and the second

$$F_2(x) \leq H_2(x) \qquad \text{for all } x.$$

When the distributions are defined over a finite range with $b$ as the upper bound value for $x$, the first condition for TSD can be restated as

$$F_1(b) \leq H_1(b)$$

If the signs of the first $n$ derivatives of the utility function can be identified, and if those signs continue to alternate, then the $n - 1$ conditions for $n - th$ degree stochastic dominance (NSD) with finite distributions can be stated as

$$F_k(b) \leq H_k(b) \qquad \text{for } k = 1, 2, \cdots, n-2$$

$$\text{and} \quad F_{n-1}(x) \leq H_{n-1}(x) \qquad \text{for all } x.$$

If any degree of stochastic dominance exists between the two alternatives in the finite distribution case then every pair of the sequence of definite integrals is ordered:

$$F_k(b) \leq H_k(b) \qquad \text{for } k = 1, 2, \cdots$$

Since the integrals used in stochastic dominance are everywhere non-negative and non-decreasing by definition of a distribution function if

$$F_{k-1}(x) \leq H_{k-1}(x)$$

$$\text{then} \quad F_k(x) = \int^x F_{k-1}(t)\,dt \leq \int^x H_{k-1}(t)\,dt = H_k(x)$$

$$\text{for } k = 1, 2, \cdots$$

When FSD exists

$$F_k(x) \leq H_k(x) \qquad \text{for all } x \text{ and } k = 0, 1, 2, \cdots$$

and $b$ is one of the values of $x$. Given SSD

$$F_k(x) \leq H_k(x) \qquad \text{for all } x \text{ and } k = 1, 2, \cdots$$

and $b$ is one of the $x$ values. Under TSD the second condition holding results in

$$F_k(x) \leq H_k(x) \qquad \text{for all } x \text{ and } k = 2, 3, \cdots$$

The first condition for TSD takes care of the case of $k = 1$. With NSD

$$F_k(x) \leq H_k(x) \qquad \text{for all } x \text{ and } k = n-1, n, n+1, \cdots$$

The first $n - 2$ $F_k(b)$ and $H_k(b)$ orderings form the first $n - 2$ conditions for NSD.

The next step is to express the geometric means of the $f$ distribution and $h$ distribution in terms of the $F_k(b)$ and $H_k(b)$.

The geometric mean for a discrete probability distribution is usually expressed as the product

$$G = \prod_i x_i p^{(xi)}$$

The definition of the geometric mean in terms of expectations is that $G$ is the antilogarithm of the expected value of the logarithm of $x$, where the logarithms are the Naperian logarithms. The discrete distribution geometric mean can therefore be stated as

$$G = e^{\sum \ln x_i p(x_i)}$$

For the continuous distributions used in this paper the geometric mean is

$$G = e^{\int \ln x f(x)\,dx}$$

To derive $G$ in terms of the repeated integrals used in the stochastic dominance models the argument of the integral, $\ln xf(x)$, will first be expanded in a power series by application of Taylor's formula. The terms in the power series will be rearranged and simplified, and when the integral of that power series is taken the result will be a series composed of negative multiples of the repeated integrals of the probability density function, the $F_k(b)$.

If it is assumed that the product $\ln xf(x)$ is an analytic function, then it can be expanded in a Taylor series around the fixed point $c$:

$$\ln xf(x) = \sum_{k=0}^{\infty} \frac{\left.\dfrac{d^k[\ln xf(x)]}{dx^k}\right|_{x=c}}{k!} [x-c]^k$$

The first few terms in the series are

$$\ln xf(x) = \ln cf(c) + \left[\ln cf'(c) + \frac{1}{c} f(c)\right][x-c] + \frac{\left[\ln cf''(c) + \frac{2}{c} f'(c) - \frac{1}{c^2} f(c)\right][x-c]^2}{2!} + \cdots$$

The terms in the series can be rearranged to

$$\begin{aligned}\ln xf(x) &= \left[\ln c - \sum_{k=1}^{\infty} \frac{(-1)^k}{k}\left[\frac{1}{c^k}\right](x-c)^k\right] \sum_{t=0}^{\infty} \frac{\left.\dfrac{d^t f(x)}{dx^t}\right|_{x=c}}{t!}[x-c]^t \\ &= \ln c\left[f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots\right] \\ &\quad + \frac{1}{c}(x-c)\left[f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots\right] \\ &\quad - \frac{1}{2}\frac{1}{c^2}(x-c)^2\left[f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots\right] \\ &\quad + \cdots\end{aligned}$$

If $f(x)$ itself is an analytic function the summation over $t$ (the content of the square brackets in the illustration) is just the Taylor series expansion of $f(x)$ and the series expansion of $\ln xf(x)$ can be simplified to

$$\begin{aligned}\ln xf(x) &= \left[\ln c - \sum_{k=1}^{\infty} \frac{(-1)^k}{k}\left[\frac{1}{c^k}\right](x-c)^k\right] f(x) \\ &= \ln cf(x) + \frac{1}{c}(x-c)f(x) - \frac{1}{2}\frac{1}{c^2}(x-c)^2 f(x) + \cdots\end{aligned}$$

Assuming that the series converges uniformly the integration to find $\ln G$ can be performed on each term in the series

$$\begin{aligned}\ln G = \int \ln xf(x)\,dx &= \ln c - \sum_{k=1} \frac{(-1)^k}{k}\left[\frac{1}{c^k}\right] \int (x-c)^k f(x)\,dx \\ &= \ln c + \frac{1}{c}\int (x-c)f(x)\,dx - \frac{1}{2}\frac{1}{c^2}\int (x-c)^2 f(x)\,dx + \cdots\end{aligned}$$

The integrals on the right side will be recognized as the $k-th$ moments around $c$, $M_k(c)$, so the logarithm of the geometric mean can be expressed in terms of those moments:

$$\ln G = \ln c - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \frac{1}{c^k} M_k(c)$$

$$= \ln c + \frac{1}{c} M_1(c) - \frac{1}{2} \frac{1}{c^2} M_2(c) + \cdots$$

If the arbitrary fixed point $c$ is the arithmetic mean of the distribution, $E(x)$, then the logarithm of the geometric mean can be expressed in terms of central moments as

$$\ln G = \ln E(x) - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \frac{1}{(E(x))^k} M_k(E(x))$$

$$= \ln E(x) - \frac{1}{2} \frac{1}{(E(x))^2} M_2(E(x)) + \frac{1}{3} \frac{1}{(E(x))^3} M_3(E(x)) + \cdots$$

This expression is the same as used by Latané, Tuttle, and Jones [12] which they reconciled with the series used by Young and Trent [20].

The fixed point $c$ can be redefined as the upper bound value $b$ for finite distributions. The logarithm of the geometric mean in terms of the $k-th$ moments around $b$, $M_k(b)$, is

$$\ln G = \ln b - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \frac{1}{b^k} M_k(b)$$

The $k-th$ moments around $b$ can be stated in terms of the repeated integrals $F_k(b)$ for the finite distribution as shown by Jean [8]. If $M_k(b)$ is integrated by parts $k$ times with $(x-b)^k$ *as u* and $f(x)$ as $dv$ in the standard formula, the result is

$$M_k(b) \int_a^b (x-b)^k f(x)\, dx = (-1)^k (k!) F_k(b)$$

Substituting the result for $M_k(b)$ in the equation for $\ln G$:

$$\ln G = \ln b - \sum_{k=1}^{\infty} \frac{(k-1)!}{b^k} F_k(b)$$

The logarithm of the geometric mean has been derived in terms of the repeated integrals used in stochastic dominance.

As usual with power series expansions the infinite series expression for the logarithm of the geometric mean can be derived in an alternate way. If

$$\ln G = \int_a^b \ln x f(x)\, dx$$

is integrated by parts an infinite number of times with $\ln x$ as $u$ and $f(x)$ as $dv$ in the standard formula, the result is obtained directly:

$$\ln G = \ln b - \sum_{k=1}^{\infty} \frac{(k-1)!}{b^k} F_k(b)$$

With expansions in infinite series it is wise to check the convergence of the resulting series. $(k-1)!/b_k$ does not converge, but $F_k(b)$ has as a limiting value $(b-a)^k/k!$ for a finite distribution:

$$F(x) \leq 1 \quad \text{for all } x$$

$$F_1(b) = \int_a^b F(x)\,dx \leq \int_a^b 1\,dx = b - a$$

$$F_k(b) = \int_a^b F_{k-1}(x)\,dx \leq \int_a^b \frac{(x-a)^{k-1}}{(k-1)!}\,dx = \frac{(b-a)^k}{k!}$$

Therefore

$$\frac{(k-1)!}{b^k} F_k(b) \leq \frac{(k-1)!}{b^k} \frac{(b-a)^k}{k!} = \left[\frac{b-a}{b}\right]^k \frac{1}{k}$$

For the geometric mean to be well defined $x$ is assumed to take only positive values, $a > 0$. As a result

$$\left[\frac{b-a}{b}\right]^k \frac{1}{k} < \frac{1}{k}$$

which is the term in the harmonic series. The infinite series developed here for the logarithm of the geometric mean therefore converges.

Now the geometric mean criterion and the stochastic dominance models can be compared directly. If any degree of stochastic dominance between two distributions exists, then as shown above

$$F_k(b) \leq H_k(b) \quad \text{for} \quad k = 1, 2, \cdots$$

Each term in the two infinite series expressions for the logarithms of the geometric means is ordered, so the logarithm of the geometric mean of the first distribution is at least as large as the logarithm of the geometric mean of the second distribution:

$$\ln G_f = \ln b - \textstyle\sum_{k=1}^{\infty} \dfrac{(k-1)!}{b^k} F_k(b) \geq \ln b - \textstyle\sum_{k=1}^{\infty} \dfrac{(k-1)!}{b^k} H_k(b) = \ln G_h$$

Since the logarithm function is monotone the geometric means are ordered:

$$G_f \geq G_h$$

The stochastic dominance ranking of any degree results in a geometric mean ranking.

If the geometric means are (definitely) ranked in the opposite direction,

$$G_f < G_h$$

then it would be impossible to have a stochastic dominance ranking of the first investment over the second investment since the set of inequalities

$$F_k(b) \leq H_k(b) \quad \text{for} \quad k = 1, 2, \cdots$$

could not hold for all $k$.

The mathematical results derived therefore show that the geometric mean ranking is a necessary condition for any degree of stochastic dominance ranking. It has not been shown here, and it is not true for general distributions, that the geometric mean ranking is a sufficient condition for a stochastic dominance ranking.

A simple example will demonstrate that the geometric mean ranking is not a sufficient condition for a stochastic dominance ranking. Suppose the first alternative has two possible outcomes, 1 and 10, with respective probabilities of 0.2 and 0.8, and the second alternative has two possible outcomes, 2 and 8, with respective probabilities 0.5 and 0.5.The first alternative's geometric mean is 6.3 and the second's is 4. The first alternative could never dominate the second alternative by any degree of stochastic dominance because it has a non-zero probability outcome lower than the lowest non-zero probability outcome of the second alternative.

There may be some specific distributions where a geometric mean ranking is a sufficient and necessary condition for a stochastic dominance ranking. Levy [13] studied the log-normal distribution and concluded that two conditions, a ranking of the geometric mean and equality of the variances of the logarithms of the variables, were sufficient for stochastic dominance. It is possible that with some distribution the geometric mean ranking alone may be sufficient for a stochastic dominance ranking.

The mathematical results of the present paper provide a partial explanation for some similarities discovered in past studies of the growth optimal portfolio and the multiperiod stochastic dominance portfolio. In Hakansson's 1971 studies, [5] and [6], he found that growth optimal portfolios were efficient when a stochastic dominance test was applied to them. Using the results above, the growth optimal portfolio could not have been inefficient by stochastic dominance. When Levy in 1973 [14] investigated the multiperiod efficient portfolios under stochastic dominance he found the reduction in size of the efficient set as the number of periods increased paralleled Hakansson's earlier results. Had Levy achieved a single member efficient set with stochastic dominance it would necessarily have been the portfolio that maximized the geometric mean.

A more important result of the relationship between the geometric mean criterion and the stochastic dominance decision models is the possibility that geometric mean techniques may provide an approach to the construction of optimum or efficient portfolios by the stochastic dominance criterion. If a stochastic dominance optimum exists, if the efficient set has a single member, it must be the portfolio with the highest geometric mean.

In a "brute force" or iterative method of searching among portfolios, such as with empirical data, each feasible portfolio must be compared at least once with another feasible portfolio. If the number of feasible portfolios is finite, say $n$, then at least $n - 1$ comparisons must be made using the portfolio selection criterion. If the stochastic dominance criterion is used in this comparison all the points in the two distributions must be used for each pair of portfolios. If the geometric mean criterion is used a single number from each portfolio must be compared. Obviously the geometric mean approach will be more efficient than utilizing the entire distributions in stochastic dominance tests.

Several papers have dealt with the optimum geometric mean portfolio, and may be more or less useful in determining an optimum or efficient stochastic dominance portfolio. Latané and Tuttle [10] and Latané and Young [11] used the arithmetic mean-variance approximation of the geometric mean in their portfolio constructions. Gressis, Hayya, and Philippatos [3] and Elton and Gruber [2] assumed log-normal return distributions. Maier, Peterson, and Vander Weide, [15] and [18], took a more general approach, treating both continuous distributions and discrete distributions in a programming approach. It is possible that the latter set of authors have come closest to a general solution for the stochastic dominance portfolio.

## REFERENCES


1. Daniel Bernoulli. "Exposition of a New Theory on the Measurement of Risk." *Econometrica*, Vol. 22 No. 1 (January 1954).
2. Edwin J. Elton, and Martin J. Gruber. "On the Maximization of the Geometric Mean with Lognormal Return Distribution." *Management Science*, Vol. 21 No. 4 (December 1974).
3. Nicolas Gressis, Jack Hayya, and George Philippatos. "Multiperiod Portfolio Efficiency Analysis Via the Geometric Mean." *The Financial Review* (1974).
4. Josef Hadar, and William R. Russell. "Rules for Ordering Uncertain Prospects." *American Economic Review*, Vol. 59 No. 1 (March 1969).
5. Nils H. Hakansson. "Capital Growth and the Mean-Variance Approach to Portfolio Selection." *Journal of Financial and Quantitative Analysis*, Vol. 6 No. 1 (January 1971).
6. ———, "Multi-Period Mean-Variance Analysis: Toward a General Theory of Portfolio Choice." *Journal of Finance*, Vol. 26 No. 4 (September 1971).
7. Giora Hanoch, and Haim Levy. "The Efficiency Analysis of Choices Involving Risk." *Review of Economic Studies*, Vol. 36 No. 2 (July 1969).
8. William H. Jean. "Comparison of Moment and Stochastic Dominance Ranking Methods." *Journal of Financial and Quantitative Analysis*, Vol. 10, No. 1 (March 1975).
9. Henry A. Latané. "Criteria for Choice among Risky Ventures." *Journal of Political Economy*, Vol. 67 No. 2 (April 1959).
10. ———, and Donald L. Tuttle. "Criteria for Portfolio Building." *Journal of Finance*, Vol. 22 No. 3 (September 1967).
11. ———, and William E. Young. "Test of Portfolio Building Rules." *Journal of Finance*, Vol. 24 No. 4 (September 1969).
12. ———, Donald L. Tuttle, and Charles P. Jones. *Security Analysis and Portfolio Management*, 2nd Ed. (Ronald Press, 1975).
13. Haim Levy. "Stochastic Dominance Among Log-Normal Prospects." *International Economic Review*, Vol. 14 No. 3 (October 1973).
14. ———, "Stochastic Dominance, Efficiency Criteria, and Efficient Portfolios: The Multi-Period Case." *American Economic Review*, Vol. 63 No. 5 (December 1973).
15. Steven F. Maier, David W. Peterson, and James W. Vander Weide. "Monte Carlo Investigations of Characteristics of Optimal Geometric Mean Portfolios." *Journal of Financial and Quantitative Analysis*, Vol. 12 No. 2 (June 1977).
16. Harry M. Markowitz. "Investment for the Long Run: New Evidence for an Old Rule." *Journal of Finance*, Vol. 31 No. 5 (December 1976).
17. James Quirk, and Rubin Saposnik, "Admissibility and Measurable Utility Functions." *Review of Economic Studies*, Vol. 29 No. 1 (February 1962).
18. James H. Vander Weide, David W. Peterson, and Steven F. Maier. "A Strategy Which Maximizes the Geometric Mean Return on Portfolio Investments." *Management Science*, Vol. 23 No. 10 (June 1977).
19. G. A. Whitmore. "Third-Degree Stochastic Dominance." *American Economic Review*, Vol. 60 No. 3 (June 1970).
20. William E. Young, and Robert H. Trent. "Geometric Mean Approximations of Individual Security and Portfolio Performance." *Journal of Financial and Quantitative Analysis*, Vol. 4 No. 2 (June 1969).