TESTING FOR INFINITE ORDER STOCHASTIC DOMINANCE WITH APPLICATIONS TO FINANCE, RISK AND INCOME INEQUALITY

By John Knight and Stephen Satchell

Headnote

In this paper we develop a test of infinite degree stochastic dominance based on the use of the empirical moment generating function. Two applications are considered. One uses the income data of Anderson (1996) and derives results consistent with his. In the other application we examine the dominance between the U.S. and U.K. stock markets. Using data on the S&P 500 and the FTALL-Share we show that the U.S. displays infinite degree stochastic dominance of the U.K.

1. Introduction[1]

Economists have always been interested in the measurement and comparison of income inequality across groups and it is well known that difficulties in using inappropriate inequality measures can be avoided by the use of notions of stochastic dominance. Such measurement of inequality across income distributions has close parallels in the measurement of risk in payoffs or returns when we consider problems of decision making under uncertainty. However, the parallel breaks down in the following sense. While it may be natural to assume that the income distribution of, say, Latvia should be uncorrelated with the income distribution of New Zealand, it is not appropriate to assume that their stock indices are independent since they will both be correlated with the global market. Thus, stochastic dominance procedures that are based on the comparison of two marginal distributions and do not utilize the extra information in the joint distribution may well be inferior to procedures based on all the information. We present a dominance testing procedure that allows for cross-correlations. The procedure we discuss may well be useful in inequality studies. In the context of income inequality, a correlated example may be where we wish to compare the wage versus the dividend distribution for the same group of individuals. Indeed, the difficulties with independence have been recognized in recent work by Davidson and Duclos (1997) who present a Lorenz analysis for the correlated case. A number of papers have discussed the econometric techniques appropriate for testing stochastic dominance between two independent income distributions, see Anderson (1996) and Davidson and Duclos (op. cit.). In what follows we shall not only consider testing stochastic dominance over two *correlated* distributions but we shall also refine the definition of stochastic dominance used in our tests.

The notion of stochastic dominance we use is that of infinite degree stochastic dominance. This is well known, see Thistle (1993), to be equivalent to completely monotone marginal utility for wealth, an assumption deemed to be desirable by Scott and Horvath (1980) and Pratt and Zeckhauser (1987) among many others. These

authors present many arguments in favour of the class of utility functions associated with the dominance concept, the simplest being that these utility functions include as examples, those members of the hyperbolic absolute risk aversion (HARA) class that have sensible comparative statics such as decreasing absolute risk aversion. Another argument being what is called proper risk aversion, which is defined as a situation where an undesirable lottery can never be made desirable by the presence of an independent undesirable lottery. It is proved by Pratt and Zeckhauser (1987, Theorem 2, p. 148) that marginal utility being completely monotone, which they call completely proper, will be proper in the above sense. Such an assumption has the behavioural implication that if we can hedge out some unfavourable risks, then agents will be more tolerant to other, independent, unfavourable risks.

In income inequality the precise nature of the social welfare function is not discussed too much, it is required to be convex but its detailed mathematical properties are rarely enunciated. In finance, the agents' utility function analogue is of central importance and a great deal of research, as discussed earlier, has been devoted to checking the implications of assumptions on the optimal portfolio chosen. In Section 2 we present a brief overview of the relevant ideas from utility theory before presenting our testing procedure in Section 3. Conclusions and examples follow in section 4.

2. Utility and Stochastic Dominance

We first define our class of completely proper utility functions. The most general representation for any $u(w) \in U_\infty$ is

$$(1) \qquad u(w) = \int_0^\infty (g(s) - e^{-sw}) dF(s)$$

where g is an arbitrary function, F is non-decreasing and $e^{-sw} = w$ if $s = 0$. An equivalent form to (1) is

$$(2) \qquad u(w) = u(w_1) + \int_0^\infty (e^{-sw}) dF(s)$$

3

for any w_1 where $u(w_1)$ is finite. Differentiating (1), we see that

$$(3) \qquad u'(w) = \int_0^\infty e^{-sw} dF(s)$$

By Bernstein's Theorem (see Feller, 1966, pp. 415-416), $u'(w)$ is completely monotone on (w_0, ∞) if and only if it is the Laplace transform of a measure F on $[0, \infty)$ and the Laplace transform is finite for $w > w_0$. Since von Neumann-Morgenstern utility is invariant to positive affine transformations, we can regard $F(s)$ as a probability distribution of a positive random variable.

It is an immediate consequence of (3) that $u'(w) > 0$ is positive as are all odd derivatives whilst all even derivatives are negative. This would imply that more positive skewness is preferred in returns and income. Whilst the preference for positive skewness in returns is well-understood, it is not so clear to the authors that the same property is deemed desirable for social welfare functions, a topic which we discuss next.

The following considerations generally support the view that skewness is preferred in the social welfare function and that infinite stochastic dominance may be an interesting issue.

Firstly, the empirical prevalence of positive skewness in earnings, see Atkinson (1983, pg. 101, xx 5.2.), for example, suggest that a welfare function which involved higher order terms would have a non-zero contribution die to moments higher then two.

Secondly stochastic dominance where expected utility is now reinterpreted as group welfare is used in inequality economics to order income distributions. So for example, third order stochastic dominance, which implies infinite order stochastic dominance follows from second order stochastic dominance and an additional condition called transfer sensitivity. This is discussed in Sen (1997, pg 138); where he describes transfer sensitivity as the requirement that a fixed-size income transfer should have a greater effect on social welfare when it occurs at a lower level of income, see also Shorrocks and Foster (1987).

Having established that infinite stochastic dominance is an interesting concept

in both decision making under uncertainty and inequality economics we proceed with definitions.

Following Thistle (1993), infinite stochastic dominance is now defined for pairwise comparisons between two distribution functions $F(w)$ and $G(w)$ as follows. We assume that the distributions have support in $w = [w_*, w^*]$ where $w_* > 0$ and w^* may be infinite. We define

$$F_j(w) = \int_{w_*}^{w} F_{j-1}(s)ds, \qquad s = 1, 2, ...$$

where $F_0(w) = F(w)$, with a similar definition for $G_j(w)$. Let $FD_n G$ denote that F, degree n, stochastically dominates G. This can be defined by

a) $F_{n-1}(w) \leq G_{n-1}(w) \quad \forall w \in W$ with $F_{n-1}(w) < G_{n-1}(w)$ for some w and $n \geq 3$.

b) $F_k(w^*) \leq G_k(w^*)$, $k = 1, 2, ..., n - 2$.

We now define infinite stochastic dominance by letting $n \to \infty$, we denote this as $FD_\infty G$. This is equivalent to $E_F(u(w)) \geq E_G(u(w))$ for all completely proper utility functions as defined by 1. Thistle (op. cit., Proposition 4, p. 307) proves that $FD_\infty G$ if and only if: $M_F(-t) < M_G(-t), \quad \forall t \in R_\infty$ where $M_F(-t) = \int_0^\infty e^{-tx} dF(x)$ and $M_G(-t)$ is defined similarly. He also shows that the above applies to the logarithms of $M_F(-t)$ and $M_G(-t)$.

Concluding, our above discussion has shown that a test for infinite degree stochastic dominance can be reduced to testing whether the Laplace transform of one wealth, income or asset price distribution is less than another at all values of t. This is now a problem of basing a test on multiple comparisons which we shall discuss next in section 3.

3. Testing Infinite Degree Stochastic Dominance

With a definition of stochastic dominance there is associated an appropriate class of utility functions and an efficient set. The efficient set is the set of portfolios or distributions which are not dominated by any other distribution. This means that there is no particular distribution preferred to the distribution in the efficient set for all utility functions in the class. It is interesting to note that as we consider higher degrees of stochastic dominance the class of utility functions becomes smaller and the efficient set becomes larger.

Tests of stochastic dominance typically have as their null hypothesis indifference between the two distributions for all utility functions in the class. Thus rejection will not necessarily imply stochastic dominance and we need to refine the tests to build in a decision rule to allow for this. For second order stochastic dominance, Bishop, Formby, and Thistle (1992) suggest that one accept the hypothesis that one distribution dominates another based on the ordinates of the Lorenz curves of the two distributions. They compute t-statistics for the null of indifference. They suggest a procedure of accepting (second degree) dominance if there is at least one significant t-statistic with the right sign and no such significant t-statistic with the "wrong" (non-dominant) sign.

In this section we develop a testing procedure for infinite degree stochastic dominance. Our procedure is based on the empirical moment generating function in keeping with the fact that $FD_\infty G$ if and only if $M_F(-t) < M_G(-t)$, for all $t \in R_+$. Letting

(4) $$z(s_j) = M_F(-s) - M_G(-s)$$

our hypothesis of interest is

(5) $$H_0 : z(s_j) = 0, \quad j = 1, 2, ..., N.$$

Testing such a hypothesis is akin to the multiple comparison testing procedures where the test statistic often used is the maximum "t test".

In our case, since we are really interested in knowing whether $M_F(-s) < M_G(-s)$ or vice versa, formulating our null hypothesis as above may not lead to a definite

conclusion. Consequently, we propose a procedure where the decision rule has four parts and is based on the outcome of both the max t test and min t test. Thus

a) If both max t and min t lead to rejection then the test is inconclusive

b) If max t rejects but min t does not reject then this implies $GD_\infty F$.

c) If max t does not reject but min t does reject then this implies $FD_\infty G$.

d) If both max t and min t do not reject then $\sim (FD_\infty G)$ and $\sim (GD_\infty F)$, i.e., F does not dominate G and G does not dominate F.

We now consider the form of the max t and min t test statistics. Define the empirical mgf functions for F and G as

$$(6) \qquad \hat{M}_F(-s) = \frac{1}{n} \sum_{j=1}^{n} e^{-sy_j}$$

$$\hat{M}_G(-s) = \frac{1}{n} \sum_{j=1}^{n} e^{-sx_j}$$

The empirical counterpart of z_t is given by

$$(7) \qquad \hat{z}(s) = \hat{M}_F(-s) - \hat{M}_G(-s)$$

Letting $\hat{\underset{\sim}{z}}(\underset{\sim}{s})$ be the vector of $\hat{z}(s)$ for N different s values, denoted by the vector $\underset{\sim}{s}$, the following lemma gives the asymptotic distribution of $\hat{\underset{\sim}{z}}(\underset{\sim}{s})$.

Lemma 1 *From the definition of $\hat{M}_F(-s)$ and $\hat{M}_G(-s)$ given in 6 along with $\hat{z}(s)$ in 7 we have via standard Central Limit Theorems that*

$$\sqrt{n}(\hat{\underset{\sim}{z}}(\underset{\sim}{s}) - \underset{\sim}{z}(\underset{\sim}{s})) \xrightarrow{d} N(0, \Sigma)$$

where

$$\Sigma_{ij} = 2M_F(-s_i - s_j) - M_{FG}(-s_i, -s_j) - M_{FG}(-s_j, -s_i),$$

with $M_{FG}(\cdot, \cdot)$ being the join mgf of F and G.

Proof. See Appendix.

From the above distribution, under H_0 $\underset{\sim}{z}(\underset{\sim}{s}) = 0$, i.e., $M_F(-s_j) = M_G(-s_j)$ and we need to adjust the elements of the covariance matrix Σ. That is, we use a pooled estimator for $M_F(-s_i - s_j)$ using both the F and G distributions. Thus under H_0 we estimate Σ_{ij} as

$$\hat{\Sigma}_{ij} = (\hat{M}_F(-s_i - s_j) + \hat{M}_G(-s_i - s_j)) - \hat{M}_{FG}(-s_i, -s_j) - \hat{M}_{FG}(-s_j, -s_i)$$

where

$$\hat{M}_{FG}(-s, -r) = \frac{1}{n} \sum_{j=1}^{n} e^{-sy_j - rx_j}$$

Therefore defining $\underset{\sim}{w}(\underset{\sim}{s})$ as our standardized statistic with unit variance we have

$$\underset{\sim}{w}(\underset{\sim}{s}) = \begin{pmatrix} \frac{\hat{z}(s_1)}{\sqrt{var(\hat{z}(s_1))}} \\ \frac{\hat{z}(s_2)}{\sqrt{var(\hat{z}(s_2))}} \\ \vdots \end{pmatrix} \xrightarrow{d} N(0, \Omega)$$

where

$$\begin{aligned} \Omega_{ij} &= 1 & \text{for } i = j \\ &= \frac{cov(\hat{z}(s_i), \hat{z}(s_j))}{\sqrt{var(\hat{z}(s_i))var(\hat{z}(s_j))}} & \text{for } i \neq j \end{aligned}$$

Since $\underset{\sim}{w}(\underset{\sim}{s})$ has a multivariate Normal distribution, we can easily calculate critical values associated with $\underset{j=1,...,N}{\max} w(s_j)$ as follows:

$$\begin{aligned} P(\max w(s_j) \leq q) &= P(w(s_1) \leq q, w(s_2) \leq q, ...) \\ &= \int_{-\infty}^{q} \cdots \int_{-\infty}^{q} (2\pi)^{-N/2} \mid \Omega \mid^{-1/2} \exp(-\frac{1}{2} \underset{\sim}{w}' \Omega^{-1} \underset{\sim}{w}) d\underset{\sim}{w} \end{aligned}$$

To calculate the critical value for $\min w(s_j)$ we can merely interchange the F and G distributions and use the $\max w(s_j)$ results. Also note that in the special case that F and G are independent then under H_0,

$$\Sigma_{ij} = 2M_F(-s_i - s_j) - 2M_F(-s_j)M_F(-s_i),$$

and there is very little simplification to our procedure.

8

To implement the above testing procedure, the major issue is how many points should we choose for the vector $\underset{\sim}{s}$ and what should be their value. The issue of how many points, i.e., value of N is probably limited by available software to calculate multivariate Normal distribution functions. Which points should be chosen is indeed a difficult question and we can only be guided by the literature on using the empirical moment generating function and characteristic function for estimation. Here the consensus is that they should be sufficiently fine and extended. (See Feuerverger and McDunnough (1981).)

4. Empirical Application

In this section we consider two applications of the testing procedure outlined above. The first application is one from an income distribution comparison where we use the same data as Anderson (1996).[1] The second application is a financial one in which we examine the dominance, if any, between the UK and US stock markets.

From the discussion in the previous section it is clear that to implement the mgf test we need to choose both the number of points for the comparison as well as the value of the points. There is no clearcut way to make these choices so we decided to set the number of points at five and to choose their value by maximizing the value of Σ, the covariance matrix. For the number of points, we are constrained by available software for calculating the orthant probabilities of the multivariate Normal distribution. While the software, for example the NAG Library subroutine G01HBF, allows a dimension up to ten we found that there was very little gain from the test using values close to ten as opposed to the chosen value of five.

4.1 Income Distribution

For the income distribution application we used the same data as in Anderson (1996) and compared the distributions of both pre and post tax incomes for the years a) 1973-1977, b) 1981-1985, c) 1981-1989.

For our pre-tax distribution comparison, the above outlined procedure for the choice of points always gave the same values viz 0.10, 1.90, 3.40, 7.00, 40.00. Table

9

1 gives the "min t" and "max t" values along with their prob. values for the three pairs of yearly comparisons. We notice immediately, for the 77 to 73 comparison that "max t" does not reject but "min t" does which implies 1977 D_∞ 1973. For the 85 to 81 comparison we have the exact opposite result, so that 1981 D_∞ 1985. Both these conclusions are consistent with the findings of Anderson (1996). For the 1989 to 1981 comparison we notice that both "min t" and "max t" reject so that in this instance the test is inconclusive. For this comparison in the Anderson (1996) study he found only marginally significant dominance of 1989 over 1981.

The results for the post-tax comparisons are given in Table 2. (Here we first make the point that our point selection procedure did not produce the same points in each case.) The results for the 77 to 73 and 85 to 81 comparisons are similar to the pre-tax case although the level of significance is greatly reduced. For the 89 to 81 case the results show that there is no stochastic dominance. Again, these results are consistent with those of Anderson (1996).

4.2 Stock Market Comparisons

For the stock market comparisons we wish to compare the S&P 500 index from the U.S. with the FTALL share index from the U.K. We use 373 monthly observations on these indexes excluding dividends from December 1964 up to and including December 1995. Since we need a positive random variable for comparison purposes we consider the arithmetic return $+1$ viz $\frac{P_t}{P_{t-1}}$ for each series. Table 3 gives some summary statistics associated with $\left(\frac{P_t}{P_{t-1}}\right)$ for each series.

Table 4 gives some results associated with the comparison using different numbers of points. We note that there is very little difference in the results from using 3 or 5 points. "min t" rejects consistently while "max t" does not reject, albeit marginally at the 5% significance level. Thus we conclude, although not surprising that SP 500 D_∞ FTALL share, at least for the period examined.

Department of Economics, University of Western Ontario, London, Ontario, Canada

and

Trinity College, University of Cambridge

10

Appendix

Proof of Lemma

Since the elements in the vector $\hat{z}\,(\underline{s})$ are merely sums of differences of exponeni-ated data it is clear that $\sqrt{n}((\hat{\underline{z}}\,(\underline{s}) - (\underline{z}\,(\underline{s}))$ will have a limiting Normal distribution. Therefore our only concern is in calculating the covariance matrix.

Consider first

$$
\begin{aligned}
cov(\hat{z}(r), \hat{z}(s)) &= E[(\hat{z}(r) - E(\hat{z}(r)))(\hat{z}(s) - E(\hat{z}(s)))] \\
&= E[(\hat{M}_F(-r) - \hat{M}_G(-r) - E(\hat{M}_F(-r) - \hat{M}_G(-r))) \\
&\quad \cdot (\hat{M}_F(-s) - \hat{M}_G(-s) - E(\hat{M}_F(-s) - \hat{M}_G(-s)))] \\
&= E[(\hat{M}_F(-r) - E(\hat{M}_F(-r)) - (\hat{M}_G(-r) - E(\hat{M}_G(-r)))) \\
&\quad \cdot (\hat{M}_F(-s) - E(\hat{M}_F(-s)) - (\hat{M}_G(-s) - E(\hat{M}_G(-s))))] \\
&= cov(\hat{M}_F(-r), \hat{M}_F(-s)) - cov(\hat{M}_F(-s), \hat{M}_G(-r)) \\
&\quad - cov(\hat{M}_F(-r), \hat{M}_G(-s)) + cov(\hat{M}_G(-r), \hat{M}_G(-s))
\end{aligned}
$$

Now

$$
\begin{aligned}
& cov(\hat{M}_F(-r), \hat{M}_F(-s)) \\
&= E[\hat{M}_F(-r)\hat{M}_F(-s)] - E[\hat{M}_F(-r)]E[\hat{M}_F(-s)] \\
&= E\left[\frac{1}{n^2}\sum_j \sum_k e^{-ry_j - sy_k}\right] - E\left[\frac{1}{n}\sum e^{-ry_j}\right]E\left[\frac{1}{n}\sum e^{-sy_j}\right]
\end{aligned}
$$

i.e.

$$
cov(\hat{M}_F(-r), \hat{M}_F(-s)) = \frac{1}{n}[M_F(-r-s) - M_F(-r)M_F(-s)]
$$

since

$$
\begin{aligned}
& E\left[\frac{1}{n^2}\left(\sum e^{-(r+s)y_j} + \sum_{j \neq k} \sum e^{-ry_j}e^{-sy_k}\right)\right] \\
&= \frac{1}{n}M_F(-(r+s)) + \frac{(n-1)}{n}M_F(-r)M_F(-s).
\end{aligned}
$$

Similarly,

$$
cov(\hat{M}_F(-s), \hat{M}_G(-r)) = \frac{1}{n}(M_{FG}(-s, -r) - M_F(-s)M_G(-r))
$$

11

$$cov(\hat{M}_F(-r), \hat{M}_G(-s)) = \frac{1}{n}(M_{FG}(-r, -s) - M_F(-r)M_G(-s))$$

$$cov(\hat{M}_G(-r), \hat{M}_G(-s)) = \frac{1}{n}(M_G(-r, -s) - M_G(-r)M_G(-s))$$

i.e.

$$
\begin{aligned}
cov(\hat{z}(r), \hat{z}(s)) &= \frac{1}{n}\{M_F(-r-s) + M_G(-r-s)) \\
&\quad - M_F(-r)M_F(-s) - M_G(-r)M_G(-s) \\
&\quad - M_{FG}(-s, -r) - M_{FG}(-r, -s) \\
&\quad + M_F(-s)M_G(-r) + M_F(-r)M_G(-s)\}
\end{aligned}
$$

Under $H_0 : z(r) = 0$ and $M_F(-r) = M_G(-r)$, thus.

$$cov(\hat{z}(r), \hat{z}(s)) = \frac{1}{n}\{2M_F(-r-s) - M_{FG}(-s, -r) - M_{FG}(-r-s)\}$$

and

$$var(\hat{z}(r)) = \frac{2}{n}(M_F(-2r) - M_{FG}(-r-r))$$

Letting $\hat{\underset{\sim}{z}}(r)$ be the vector of $z(r)$ for N different r values, denoted by the vector $\underset{\sim}{r}$, we have via standard Central Limit Theorems that

$$\sqrt{n}(\hat{\underset{\sim}{z}}(\underset{\sim}{r}) - \underset{\sim}{z}(\underset{\sim}{r})) \xrightarrow{d} N(0, \textstyle\sum)$$

where

$$\textstyle\sum_{ij} = 2M_F(-r_i - r_j) - M_{FG}(-r_j - r_i) - M_{FG}(-r_i - r_j)$$

Note

1. We thank Gordon Anderson for supplying the data.

References

Anderson, G. J. (1996). "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica* 64: 1183-1193.

Atkinson, A. B. (1983). "The Economics of Inequality" and Edition, Clarendon Press, Oxford.

Bishop, J., J. Formby and P. Thistle (1992). "Convergence of the South and Non-South Income Distributions, 1969-1970," *American Economic Review* 82, 262-272.

Davidson, R. and J. Duclos (1997). "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers," *Econometrica* 65, 1453-1465.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed. (New York: John Wiley and Sons).

Feuerverger, A. and P. McDunnough (1981). "On the Efficiency of the Empirical Characteristic Function Procedures," *Journal of the Royal Statistical Society*, Series B 43: 20-27.

Pratt, J. W. and R. J. Zeckhauser (1987). "Proper Risk Aversion," *Econometrica* 55: 143-154.

Scott, R. C. and P. A. Horvath (1980). "On the Direction of Preference for Moments of Higher Order Than the Variance," *Journal of Finance* 35: 915-919.

Sen, A. (1997) "On Economic Inequality", Clarenden Press, Oxford.

Shorrocks, A. F. and J. E. Foster (1987). "Transfer Sensitive Inequality Measures" *Review of Economics Studies* 54: 248-261.

Thistle, P.D. (1993). "Negative Moments, Risk Aversion and Stochastic Dominance," *Journal of Financial and Quantitative Analysis* 28: 301-311.

Table 1

Income Comparisons (Pre-Tax)

	77-73	85-81	89-81
"min t"	-10.268	-0.458	-4.787
Prob	2.05E-4	0.645	1.04E-4
"max t"	-0.823	3.430	-2.073
Prob	0.1398	1.33E-3	1.33E-4

Table 2

Income Comparisons (Post-Tax)

	77-73	85-81	89-81
"min t"	-11.942	-0.999	-2.84E-2
Prob	9.72E-5	0.388	0.763
"max t"	-0.781	4.340	1.210
Prob	2.63E-2	1.26E-4	0.249

Table 3

Summary Statistics

	S&P 500	FTALL Share
mean	4.166E-3	4.473E-3
st. dev.	4.547E-3	5.174E-3
skewness	0.484	0.926
kurtosis	-1.048	0.939
min	0.0	0.0
max	2.038E-2	3.058E-2
Autocorrelations		
lag 1	0.812	0.743
2	0.841	0.741
3	0.816	0.737
4	0.804	0.718
5	0.809	0.718
6	0.809	0.722
7	0.779	0.698
8	0.786	0.714
9	0.768	0.679
10	0.761	0.701

Table 4

Comparison of S&P 500 to FTALL Share

# of pts.	"min t"	Prob	"max t"	Prob
3	-2.88	4.492E-3	-1.00	5.409E-2
4	-2.86	5.047E-3	-1.00	5.390E-2
5	-2.90	8.313E-3	-1.00	5.320E-2