Journal of Econometrics 162 (2011) 71–78

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Semi-nonparametric test of second degree stochastic dominance with respect to a function

Keith D. Schumann *

Texas A&M University, Department of Agricultural Economics, Agricultural and Food Policy Center, 77843-2124 College Station, TX, United States

ARTICLE INFO

Article history:
Available online 12 October 2009

ABSTRACT

In an expected utility framework, assuming a decision maker operates under utility $k(\cdot|\theta)$, for two risky alternatives X and Y with respective distribution functions F and G, alternative X is said to dominate alternative Y with respect to $k(\cdot|\theta)$ if $\int_{-\infty}^{y}[F(t)-G(t)]\mathrm{d}k(t|\theta) \leq 0$ for all y. Utilizing the empirical distribution functions of F and G, a statistical test is presented to test the null hypothesis of indifference between X and Y given $k(\cdot|\theta)$ against the hypothesis that X dominates Y with respect to $k(\cdot|\theta)$. This is a large sample testing application of stochastic dominance with respect to a function. The asymptotic distribution of the test statistic associated with the null hypothesis given a sub-set of the utility function parameter space is developed. Based on large sample rejection regions, the hypothesis of preference of one alternative over another is demonstrated with an empirical example.

1. Introduction

Risky investment alternatives can be preference-ranked by the ordering of representations that attempt to take risk attitudes into account. Stochastic dominance measures attempt to evaluate preference rankings based on the distributions of the alternatives and an assumed decision mechanism. The goal of this paper is to attempt to extend the literature on statistical methods of testing for second degree stochastic dominance between alternatives. A testing procedure for preference ranking given a utility function as a mechanism for choice will be presented. This procedure generalizes previous methods developed for testing preferences.

The formal expression of the nature of rational decisions in the context of stochastic outcomes was initially presented by Von Neumann and Morgenstern (1953), and later expounded upon, most notably by Friedman and Savage (1948, 1952), Pratt (1964) and Arrow (1965) by way of describing risk aversion. Built on the fundamentals of expected utility, stochastic dominance methods were developed as a technical means to rank risky alternatives. Hadar and Russell (1969) as well as Hanoch and Levy (1969) are commonly attributed as outlining the concepts of first as well as second degree stochastic dominance. Meyer (1975) extended these measures of stochastic dominance by explicitly incorporating the utility function in the comparison rather than general assumptions about the choice mechanism of the decision-making agent. Assuming a strictly increasing utility function, he found that, without loss of generality, for a restricted support random variable, standardized to the support range [0, 1],

$$\int_0^y [F(t)-G(t)]\mathrm{d}r(t) \leq 0$$

for all $y \in [0, 1]$ given an increasing, twice differentiable function $r(\cdot)$ if and only if G is at least as risky as F. The riskiness of a random variable is understood with respect to the utility preference or decision-making mechanism an agent employs. Given equivalent expectations, the more risky prospect is the one that has a higher relative variability. For an agent who is averse to risk, i.e., a person who chooses to insure against risk, minimizing risk or variability in the lower tail of a distribution is a primary concern.

Meyer (1975) formalized the concept of second degree stochastic dominance with respect to a function ($SSD(k_\theta)$ or $SSD(k)$) with an application (Meyer, 1977a) and theoretical development (Meyer, 1977b). For two restricted support random variables standardized to the [0,1] interval with corresponding distribution function, F $SSD(k)$ G if and only if

$$\int_0^y [F(t)-G(t)]\mathrm{d}k(t) \leq 0$$

for $y \in [0, 1]$ and a function $k(\cdot|\theta)$ with a given value of θ. This is a necessary and sufficient condition for F to be preferred to or indifferent to G by all agents with a utility function $u(\cdot)$ that exhibits equal or more risk aversion, or concavity, than the function $k(\cdot)$. The $SSD(k)$ ordering result is unique at y. It can be seen that second degree stochastic dominance (SSD) is a special case of $SSD(k_\theta)$ where $k(x) = x$.

Another generalization of stochastic dominance with respect to a function is called stochastic efficiency with respect to a

* Tel.: +1 979 845 8014.
E-mail address: k-schumann@tamu.edu.

0304-4076/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.jeconom.2009.10.009

function (SERF) (Hardaker et al., 2004). This method seeks to examine the dominance of one alternative over another given a continuous range of a risk aversion measure as originally specified by Meyer (1977b). A potential problem with a particular application of stochastic dominance with respect to a function is that order preferences are evaluated only at specified boundary points of the risk aversion measure. For two alternatives with distribution functions that cross multiple times, this could lead to misidentifications of the efficient set of alternatives over that range. The SERF method evaluates the certainty equivalent ($CE(\theta)$), the fixed amount where a risk averse agent would be indifferent between the random variable and that amount, of each alternative over the relevant parameter space of the utility function.

Several methods have been proposed to develop statistical hypothesis tests for ranking one random alternative over another, or placing alternatives in an efficient set of preferred alternatives. One often used assumption is that the nonparametric estimates of the underlying distributions can be effectively utilized in assessments. In particular, McFadden (1989) and Klecan et al. (1991) developed tests, based on equally numbered observations, which are variations of the Kolmogorov–Smirnov procedure and deal with the issue of sampling independence. Anderson (1996) proposed nonparametric tests for first, second, and third degree stochastic dominance criteria based on analogs of Pearson goodness of fit tests. These tests were shown to be comparable in size and power to generalized Lorenz curve methods (as in Bishop et al., 1989) for comparing distributional differences for wealth. Multivariate extensions to Anderson's test have also been introduced (see Crawford (2005) and Post and Versijp (2005)).

In addition to mean–variance methods and stochastic dominance criteria, efforts to link these and related numerical procedures have been undertaken. Yitzhaki (1982) introduced methods to incorporate Gini's mean difference (GMD) in the analysis of preference ranking, and he elaborated on this work by utilizing resampling methods to calculate the variance of estimates of this type (Yitzhaki, 1991). A comparison of mean-GMD analyses and stochastic dominance results as well as an application in agricultural commodities was illustrated in McDonald et al. (1997). Here it was suggested that mean-GMD methods of analyzing the efficient set of preferred alternatives had superior properties over stochastic dominance methods. Shalit and Yitzhaki (1994) introduced the concept of marginal conditional stochastic dominance (MCSD) and Seiler (2001) proposed a nonparametric test for this preference ranking method.

Another work in the area of hypothesis testing for preferences was Eubank et al. (1993), hereafter referred to as ESY. ESY described a nonparametric method to test for second degree stochastic dominance. This work was a generalization of a methodology developed by Deshpande and Singh (1985), where the distribution of one of the risky prospects was assumed to be known for testing purposes. The asymptotic properties of the SSD test statistic, including asymptotic power, were elicited in ESY. The large sample variance of the statistic had a form that was cumbersome from an estimation perspective, so a resampling method to approximate the variance was suggested. Another variation of this testing procedure was developed by Kaur et al. (1994) which can be implemented across unequal sample sizes but requires a search across a finite number of possibilities to determine the proper test statistic. Davidson and Duclos (2006) evaluate bootstrapping methods that evidence an improvement in the asymptotic efficiency of existing statistics for testing stochastic dominance in their simulated experiments.

2. Hypothesis test

Drawing on many of the former concepts and methodologies, the conditions for test of second degree stochastic dominance with respect to a function are specified in this section. Consider an agent with an initial level of wealth who wishes to invest in a subset of a finite number of risky investment alternatives. The agent's initial wealth will be labeled w_0, and two risky opportunities, X and Y, will be considered for investment. It is assumed that the agent has a single utility function for decision-making involving wealth, measured with a proxy utility function $k(\cdot|\theta)$, where the unknown parameter set θ describes the shape specifications of that utility function.

The risk measure is utility function-specific and is a function of θ as it relates to the variables X and Y. The value for θ is considered to be known a priori. The class of utility functions to be considered will be those that are increasing and twice differentiable with respect to the risky variable. This ensures that the local risk aversion coefficient exists for all X and Y.

Independent samples of size n and m from each of the prospective investments X and Y respectively will be considered as a basis for estimating the distribution functions of each of the alternatives. The observations on the investments will be denoted as (x_i) and (y_j) for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

Given the previous conditions, the expected utility of the first investment, for example, is defined as

$$Ek(X|\theta) = \int_{\chi} k(t|\theta)\mathrm{d}F(t), \tag{1}$$

where F is the cumulative distribution function of the investment X over χ, the support of X. Often, there is not enough information to make parametric distributional assumptions on F. Even so, imposing assumed distributions when making comparisons can be detrimental to a testing procedure employing utility-weighting given the relative emphasis on the tails of the distributions. Since the distribution functions of the alternatives are assumed to be unknown, the empirical distribution function (EDF), F_n, will be considered as an estimate of F. The EDF will be defined as

$$F_n(x) = \frac{1}{n}\sum_{1}^{n} I_{(-\infty,x_i]}(x), \tag{2}$$

where $I_A(z)$ is the indicator function which takes on a value of one if $z \in A$ and zero otherwise.

Referring again to the definition of second degree stochastic dominance with respect to a function, where $F\,SSD(k_\theta)\,G$ if and only if

$$\int_{-\infty}^{x} [F(t) - G(t)]\mathrm{d}k(t|q) \leq 0$$

for all x in the support of X and Y, a function $k(\cdot|\theta)$, and a given value of θ, a formal testing procedure can be elicited. Given that the utility functions in the general class in question are one-to-one and assuming these functions are locally monotonic, it is straightforward to conjecture that the desired result might be attained by transforming the variables to utility measures and then using these as the basis for the quantities of interest in the test procedure for second degree stochastic dominance described in ESY.

Following the previous assumptions, we wish to test hypotheses of the type:

$$H_0 : F = G \quad \forall\theta \in \Theta_1 \tag{3}$$

$$H_1 : F\,SSD(k_\theta)\,G \quad \forall\theta \in \Theta_1, \tag{4}$$

where Θ_1 is the parameterization of interest of the given utility function $k(\cdot|\theta)$. It will be shown that the test statistic described in