Testing for Bivariate Stochastic Dominance Using Inequality Restrictions

Thanasis Stengos*and Brennan S. Thompson†

1st April 2009

Abstract

In this paper, we propose of a test of bivariate stochastic dominance using a generalized framework for testing inequality constraints. This test has the advantage of taking the covariance structure of the estimates of the joint distribution functions into consideration. Our proposed methods are illustrated with a Monte Carlo experiment and an empirical application which utilizes Canadian household survey data on income and educational attainment.

Keywords: Stochastic dominance, inequality restrictions, multidimensional poverty

*Department of Economics, University of Guelph, Guelph, ON N1G 2W1. E-mail: tstengos@uoguelph.ca

†Department of Economics, Ryerson University, Toronto, ON M5B 2K3. E-mail: brennan@ryerson.ca

1 Introduction

Stochastic dominance is a criterion used extensively in welfare economics and other fields, such as financial economics, to compare a set of distributions. In the simplest case, one univariate distribution, A, is said to (weakly) first-degree stochastically dominate another, B, if, for any value z, the cumulative distribution function (CDF) for A at z is less than or equal to the CDF for B at z (formal definitions are given in Section 2).

In the past two decades, a number of statistical tests of stochastic dominance have been put forth in the literature. These tests can broadly be divided into two general categories. Tests in category one, which include those proposed by Anderson (1996), Fisher et al. (1998), Davidson and Duclos (2000), and Davidson and Duclos (2007), all of which are applicable only to univariate distributions, involve evaluating each CDF at a finite number of points.[1] Tests in category two are based on evaluations over the entire support of each CDF. This category includes the univariate tests of McFadden (1989), Kaur et al. (1994), Maasoumi and Heshmati (2000), Barrett and Donald (2003), and Linton et al. (2005), as well as the multivariate tests of McCaig and Yatchew (2007) and Anderson (2008).

Tests in category one have the disadvantage of requiring the researcher to specify a set of arbitrary evaluation points. As suggested by Davidson

[1] Related tests in this category are the tests of Lorenz dominance by Beach and Davidson (1983), Beach and Richmond (1985), Bishop et al. (1993), and Dardanoni and Forcina (1999). Davidson and Duclos (2000) discuss the relation between Lorenz dominance and stochastic dominance.

and Duclos (2000) and Barrett and Donald (2003), these tests might, as a result, be inconsistent. However, these tests have the advantage of making use of the covariances between the estimates made at each of the evaluation points (see Davidson and Duclos, 2000). Tests in category two ignore this covariance structure.

Simulation evidence provided in Tse and Zhang (2004) suggests that, in finite samples, ignoring the dependence between test points will tend result in rejection frequencies under the null which are greater than the nominal size of the test. Specifically, these authors find that, when taking this dependence into account, a test suggested in Davidson and Duclos (2000) outperforms those of Kaur et al. (1994) and Anderson (1996), both of which ignore this dependence, in terms of both size and power.

In this paper, we propose a test for bivariate stochastic dominance which involves evaluating each CDF at a finite number of points (i.e., over a set of grid points). This test, belonging to category one, can be seen as a simple extension of the methods of Fisher et al. (1998) and Davidson and Duclos (2000) to the bivariate case. While a partial extension of these methods was considered by Duclos et al. (2006), these authors do not account for the covariance structure between estimates at each grid point. We are able to do by using the general methods of Kodde and Palm (1986) and Wolak (1989) for testing vectors of inequality constraints.

The remainder of this paper is organized as follows. In Section 2, we provide a formal definition of stochastic dominance (of any order) and briefly discuss how these relations can be related to poverty orderings. In Section 3 we discuss how stochastic dominance relations can be estimated, and provide the asymptotic distribution of these estimates. Section 4 discusses a hypothesis test based on these results, and contrasts this test with that of McCaig and Yatchew (2007), which belongs to category two. These two tests are then compared in a Monte Carlo simulation in Section 5. In Section 6, we present an empirical application of our proposed methods using Canadian household survey data on income and educational attainment. Section 7 concludes.

2 Stochastic dominance and poverty orderings

Let F_A and F_B denote two right-continuous d-dimensional CDFs. We say that distribution F_A (weakly) dominates distribution F_B stochastically at order s (an integer) if $D_A^s(\mathbf{z}) \leq D_B^s(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}_+^d$, where, for $K = A, B$, $D_K^1(\mathbf{z}) = F_K(\mathbf{z})$ and $D_K^s(\mathbf{z})$ is defined recursively as

$$D_K^s(\mathbf{z}) = \int_0^{\mathbf{z}} D_K^{s-1}(\mathbf{u}) d\mathbf{u}, \qquad s \geq 2.$$

In what follows, we will denote this relation by $F_A \succeq_s F_B$.

Alternatively, rather than considering *any* given level $\mathbf{z}$, we may focus on some specific range of values, say the hyperrectangle $[z_1^-, z_1^+] \times \cdots \times [z_d^-, z_d^+]$. In this case, if $D_A^s(\mathbf{z}) \leq D_B^s(\mathbf{z})$ for all $\mathbf{z} \in [z_1^-, z_1^+] \times \cdots \times [z_d^-, z_d^+]$, we may say that distribution F_A dominates distribution F_B stochastically at order

s on a *restricted* basis. Such restricted stochastic dominance relations are particularly useful when considering poverty orderings.

For example, suppose that F_A and F_B denote (univariate) income distributions for countries A and B, respectively.[2] In this case, if the proportion of the population in country A with income at or below any point in the interval $[z^-, z^+]$, which represents a range of admissible poverty lines, is less than or the same as in country B, then we can conclude that the level of poverty in country A is (weakly) lower than in country B for *any* poverty line in this interval as measured by the headcount ratio. That is, denoting by $H_K(z) = F_K(z)$, the headcount ratio for country $K = A, B$, we have $H_A(z) \leq H_B(z)$ for all $z \in [z^-, z^+]$. This, of course, is equivalent to saying that F_A dominates distribution F_B stochastically at first order on a restricted basis.[3]

3 Estimation and inference

While Duclos et al. (2006) extend some of the results from Davidson and Duclos (2000) to the bivariate case, they do not consider the covariance

[2]The relation of bivariate restricted stochastic dominance to poverty orderings is discussed in some detail in Section **??**.

[3]Higher orders of restricted stochastic dominance can also be interpreted in terms of poverty orderings. In this example, if distribution F_A dominates distribution F_B stochastically at order s on a restricted basis, than poverty in country A is (weakly) lower than in country B for any poverty line in the interval $[z^-, z^+]$ as measured by the Foster et al. (1984) measure with poverty aversion parameter $s-1$ (see Foster and Shorrocks, 1988 and Davidson and Duclos, 2000). The interpretation of higher order degrees of stochastic dominance as poverty orderings in the multivariate case is not so straightforward; see the discussion in Atkinson (2003) on multidimensional poverty measures.

between estimates at different points of the distributions. In this section, we discuss how these covariances can be estimated in the bivariate case. In the next section, we show how these covariance estimates can be incorporated into hypothesis tests for bivariate stochastic dominance.

To simplify notation, we will focus on just a single sample in this section (the hypothesis tests introduced in the next section will be based on two independent samples). Specifically, suppose we have a sample of n independent and identically distributed (IID) observations drawn from the joint distribution of X and Y. In what follows, we will obtain estimates on a $J \times J$ set of grid points. Along the support of X, we have the points $\lambda_X = (\lambda_{X,1}, \ldots, \lambda_{X,J})'$, and along the support of Y, we have the points and $\lambda_Y = (\lambda_{Y,1}, \ldots, \lambda_{Y,J})'$. We denote by λ a J^2-vector of all of the unique grid points.

Following Davidson and Duclos (2000), it will be convenient, in the bivariate case, to rewrite $D^s(\mathbf{z}) = D^s(z_X, z_Y)$ as

$$D^s(z_X, z_Y) = \frac{1}{(s-1)!} E[(z_X - X)_+^{(s-1)} (z_Y - Y)_+^{(s-1)}], \qquad s \geq 1.$$

where $\phi_+ = \max(0, \phi)$.

A natural estimate of $D^s(\lambda_{X,j}, \lambda_{Y,k})$ is

$$\hat{D}^s(\lambda_{X,j}, \lambda_{Y,k}) = \frac{1}{n(s-1)!} \sum_{i=1}^{n} (\lambda_{X,j} - x_i)_+^{s-1} (\lambda_{Y,k} - y_i)_+^{s-1}, \tag{1}$$

for $j, k = 1, \ldots, J$.

Since each of these estimates is just a sum of IID random variables, we can apply a multivariate central limit theorem to find its asymptotic distribution. Specifically, $\sqrt{n}[\hat{D}^s(\lambda) - D^s(\lambda)]$ will converge in distribution to a normal random vector with mean vector zero and $(J^2 \times J^2)$ covariance matrix $\text{Cov}[\hat{D}^s(\lambda)]$, which has typical element

$$\begin{aligned}
&\text{Cov}[\hat{D}^s(\lambda_{X,j}, \lambda_{Y,k}), \hat{D}^s(\lambda_{X,l}, \lambda_{Y,m})] \\
&= \frac{1}{[(s-1)!]^2} E[(\lambda_{X,j} - X)_+^{s-1}(\lambda_{Y,k} - Y)_+^{s-1}(\lambda_{X,l} - X)_+^{s-1}(\lambda_{Y,m} - Y)_+^{s-1}] \\
&\quad - D^s(\lambda_{X,j}, \lambda_{Y,k}) D^s(\lambda_{X,l}, \lambda_{Y,m}),
\end{aligned}$$

for $j, k, l, m = 1, \ldots, J$.

A consistent estimate of $\text{Cov}[\hat{D}^s(\lambda_{X,j}, \lambda_{Y,k}), \hat{D}^s(\lambda_{X,l}, \lambda_{Y,m})]$ can be obtained using

$$\begin{aligned}
&\hat{\text{Cov}}[\hat{D}^s(\lambda_{X,j}, \lambda_{Y,k}), \hat{D}^s(\lambda_{X,l}, \lambda_{Y,m})] \\
&= \frac{1}{n[(s-1)!]^2} \sum_{i=1}^{n} [(\lambda_{X,j} - x_i)_+^{s-1}(\lambda_{Y,k} - y_i)_+^{s-1}(\lambda_{X,l} - x_i)_+^{s-1}(\lambda_{Y,m} - y_i)_+^{s-1}] \\
&\quad - \hat{D}^s(\lambda_{X,j}, \lambda_{Y,k}) \hat{D}^s(\lambda_{X,l}, \lambda_{Y,m}).
\end{aligned}$$

In the following section, we show how these results can be used to test for bivariate stochastic dominance between two populations.

4 Hypothesis testing

To test for bivariate stochastic dominance between populations, we use the general approach to testing multivariate inequality restrictions of Kodde and Palm (1986) and Wolak (1989). This approach has also been used for tests for of univariate stochastic dominance by Fisher et al. (1998) and Davidson and Duclos (2000).

Specifically, we are interested in testing hypotheses of the form

$$H_0 : F_A \succeq_s F_B$$

against an unrestricted alternative. Letting $\Delta = D_B^s(z_X, z_Y) - D_A^s(z_X, z_Y)$, we can rewrite the null hypothesis above as

$$H_0 : \Delta \geq 0.$$

The unrestricted estimate of Δ is $\hat{\Delta} = \hat{D}_B^s(\lambda) - \hat{D}_A^s(\lambda)$, where $\hat{D}_K(\lambda)$ is the estimator given in the previous section for population $K = A, B$. The restricted estimate of Δ can be found as the solution to the following minimization problem:

$$\min_{\Delta \geq 0} (\hat{\Delta} - \Delta)' \hat{\Sigma}^{-1} (\hat{\Delta} - \Delta), \tag{2}$$

where $\hat{\Sigma}$ is an estimate the asymptotic covariance matrix of $\hat{\Delta}$. Under the assumption that A and B represent two independent samples, we have

$$\hat{\Sigma} = \hat{\text{Cov}}[\hat{D}^s_B(\lambda)]/n_B - \hat{\text{Cov}}[\hat{D}^s_A(\lambda)]/n_A,$$

where $\hat{\text{Cov}}[\hat{D}^s_K(\lambda)]$ is an estimate of the asymptotic covariance matrix of $\sqrt{n_K}(\hat{D}^s_K - D^s_K)$, for $K = A, B$ (see Section 3 for details).

Solving for Δ in (2) is a straightforward quadratic programming (QP) problem.[4] Denoting the solution by $\tilde{\Delta}$, we have the Wald-type (or Hausman-type) test statistic

$$W = (\hat{\Delta} - \tilde{\Delta})'\hat{\Sigma}^{-1}(\hat{\Delta} - \tilde{\Delta}).$$

As shown by Kodde and Palm (1986), under the null, W will converge in distribution to a mixture of J χ^2 distributions.

To avoid the complexities associated with computing the critical values for W, we advocate the use the bootstrap. Specifically, we combine samples of observations on (X_A, Y_A) and (X_B, Y_B) into pooled sample (which is of length $n_A + n_B$). Resampling (in pairs) n_K observations from this pooled sample produces the bootstrap sample (X^*_K, Y^*_K), for $K = A, B$.

Using the two bootstrap samples, we calculate the bootstrap test statistic, W^*, in a matter analogous to that for the original test statistic, W (see above). Repeating this process some large number of times, the bootstrap p-value for W is the proportion of times that W^* exceeds W.

---

[4]We briefly discussed this problem in the appendix to the previous chapter (Section **??**).

We now briefly contrast this approach with that of McCaig and Yatchew (2007), who consider tests of multivariate stochastic dominance of the category two type. Their test statistic is

$$T = \left\{ \int [\psi^s(\mathbf{u})]^2 d\mathbf{u} \right\}^{1/2}$$

where $\psi^s(\mathbf{u}) = \max\{D_A^s(\mathbf{u}) - D_B^s(\mathbf{u}), 0\}$. Of course, when the null is true, T is equal to zero.

In practice, this test involves estimating T and testing whether it is statistically different from zero. Specifically, in the bivariate case, McCaig and Yatchew (2007) estimate T by

$$\hat{T} = \left\{ \sum_{j=1}^{J} \sum_{l=1}^{J} [\hat{\psi}^s(\lambda_{X,j}, \lambda_{Y,l})]^2 \right\}^{1/2},$$

where

$$\hat{\psi}^s(\lambda_{X,j}, \lambda_{Y,l}) = \max\{\hat{D}_A^s(\lambda_{X,j}, \lambda_{Y,l}) - \hat{D}_B^s(\lambda_{X,j}, \lambda_{Y,l}), 0\},$$

and $\hat{D}_A^s(\lambda_{X,j}, \lambda_{Y,l})$ and $\hat{D}_B^s(\lambda_{X,j}, \lambda_{Y,l})$ are obtained using the estimator in (1). As in our approach, $\lambda_X = (\lambda_{X,1}, \ldots, \lambda_{X,J})'$ is a set of J points along the support of X and $\lambda_Y = (\lambda_{Y,1}, \ldots, \lambda_{Y,J})'$ is a set of J points along the support of Y.

Thus, in practice, this test would seem to fall in the same category as our proposed one. However, there is nothing inhibiting the use of an extremely large number of grid points (perhaps every unique point supported by the sample). That being said, McCaig and Yatchew use $J = 25$ in their simulations and empirical applications. While this would be quite computationally demanding for our approach (requiring, e.g., the inversion of a $25^2 \times 25^2$ matrix), it would not be out of the question.

To estimate the null distribution of $\hat{T}$, McCaig and Yatchew (2007) use a bootstrap procedure which is analogous to the one we have described above for our proposed Wald-type test statistic, W.

5 Simulation evidence

To compare our proposed test to that of McCaig and Yatchew (2007), we use a simple Monte Carlo experiment. For each simulation we conduct, 1,000 sets of two independent samples are generated from the bivariate lognormal distribution, with $E(X) = \mu_X$, $E(Y) = \mu_Y$, $\text{Var}(X) = \sigma_X^2$, $\text{Var}(Y) = \sigma_Y^2$, and $\text{Cov}(X, Y) = \sigma_{X,Y}$.[5] The values of these parameters for two alternative data generating processes we consider (denoted A and B) are given in Table 1. The size of the samples we use are $n_A = n_B = n = 50$ and 500.

We consider two different testing scenarios. In the first, we generate both samples from the same distribution (distribution A), and test the null hypothesis that one stochastically dominates the other at first order. Since

[5] Note that these are the parameters for lognormally distributed random variables, *not* the underlying normally distributed random variables (from which the lognormally distributed random variables are generating through exponentiation).

this hypothesis is (weakly) true, we would expect to reject it at the nominal level of the test, which we set at 5%.

In the second scenario, we generate one sample from each of the distributions, and test the null hypothesis that distribution A stochastically dominates distribution B at first order, which is, in fact, false. Thus, the proportion of rejections observed for either test is an indication of the power of that test.

The rejection frequencies for these two scenarios, based on 99 bootstrap replications, are reported in Table 2. For our Wald-type test, we try values of both 5 and 10 for J (the number of evaluation points in each dimension, which we select to be evenly-spaced), so that the total number of grid points are 25 and 100, respectively.

Unfortunately, the size of both the McCaig and Yatchew (2007) test and our Wald-type test (with either value of J) appears to be distorted. This result warrants some further investigation. Some preliminary evidence (not presented in Table 2) suggests that, even with higher values of J, say $J = 20$, there is an improvement, even though our test still tends to over-reject the null.

One notable result of this simulation is that the power of our Wald-type test appears to depend crucially on the number of evaluation points. With $J = 5$, the power of this test appears to be substantially worse than that of the McCaig and Yatchew test. However, with $J = 10$, the power of this test seems to be about the same as, if not better than, the McCaig and Yatchew test.

6 Empirical example

In this section, we consider an empirical example which uses income and education data for two subgroups of the Canadian population. The data for this example is obtained from Statistics Canada's *Survey of Labour and Income Dynamics* (SLID) for 2004. In order to reduce the level of heterogeneity within the sample, only unattached individuals residing in urban areas with populations between 100,000 and 499,999 are included. There are 662 males and 1,151 females. Income is measured using annual after-tax income, and education is measured using years of schooling.

The empirical marginal distribution functions for the income and education variables are plotted in Figures 1 and 2, respectively. Looking at income alone, Figure 1 shows that the empirical marginal distribution function for males intersects that for females from slightly below at a point of approximately $18,000. Examining the education data, the empirical marginal distribution function for females seems to be almost everywhere above that for males, except for a slight dip below around 18 or 19 years of schooling. However, such casual inspection of these estimated marginal distributions is not at all conclusive. First, a formal statistical test would be required to determine if the differences between these estimated functions are significant (this could be done using any of the tests for univariate stochastic dominance discussed in the introduction). Second, and more importantly, these estimates do not capture the joint distribution of income and education.

In order to formally tests for bivariate stochastic dominance, we use the Wald-type test proposed above. In doing so, we first select a 10×10 set of grid points evenly spaced across the empirical support of each variable.

Testing for first-order bivariate stochastic dominance of males over females, our Wald-type test statistic is 0.7792 and the bootstrap p-value (using 99 bootstrap replications) is 0.9596. On the other hand, testing for first-order bivariate stochastic dominance of females over males, our Wald-type test statistic is 72.1589 and the bootstrap p-value is exactly zero. Thus, while we fail to reject the null of first-order bivariate stochastic dominance by males, we can reject the null of first-order bivariate stochastic dominance by females. In other words, we can conclude that the joint distribution of income and education for males first-order stochastically dominates that for females.

7 Conclusion

In this paper, we have proposed a test for bivariate stochastic dominance which involves evaluating each CDF at a finite number of points (i.e., over a set of grid points). Simulation evidence provided here suggests that this test performs about as well as the the test of McCaig and Yatchew (2007).

References

Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. *Econometrica*, 64(5):1183–1193.

Anderson, G. (2008). The empirical assessment of multidimensional welfare, inequality and poverty: Sample weighted multivariate generalizations of the Kolmogorov-Smirnov two sample tests for stochastic dominance. *Journal of Economic Inequality*, 6(1):73–87.

Atkinson, A. B. (2003). Multidimensional deprivation: Contrasting social welfare and counting approaches. *Journal of Economic Inequality*, 1(1):51–65.

Barrett, G. F. and Donald, S. G. (2003). Consistent tests for stochastic dominance. *Econometrica*, 71(1):71–104.

Beach, C. M. and Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares. *Review of Economic Studies*, 50(4):723–735.

Beach, C. M. and Richmond, J. (1985). Joint confidence intervals for income shares and Lorenz curves. *International Economic Review*, 26(2):439–450.

Bishop, J. A., Formby, J. P., and Smith, W. J. (1993). International comparisons of welfare and poverty: Dominance orderings for ten countries. *Canadian Journal of Economics*, 26(3):707–726.

Dardanoni, V. and Forcina, A. (1999). Inference for Lorenz curve orderings. *Econometrics Journal*, 2(1):49–75.

Davidson, R. and Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and the measurement of povery and inequality. *Econometrica*, 68(6):1435–1464.

Davidson, R. and Duclos, J.-Y. (2007). Testing for restricted stochastic dominance. Unpublished manuscript.

Duclos, J.-Y., Sahn, D., and Younger, S. D. (2006). Robust multidimensional poverty comparisons. *Economic Journal*, 116(514):943–968.

Fisher, G., Wilson, D., and Xu, K. (1998). An empirical anlysis of term premiums using significance tests for stochastic dominance. *Economics Letters*, 60(2):195–203.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3):761–766.

Foster, J. E. and Shorrocks, A. F. (1988). Poverty orderings. *Econometrica*, 56(1):173–177.

Kaur, A., Prakasa Rao, B. L. S., and Singh, H. (1994). Testing for second-order stochastic dominance of two distributions. *Econometric Theory*, 10(5):849–866.

Kodde, D. A. and Palm, F. C. (1986). Wald criteria for jointly testing equality and inequality restrictions. *Econometrica*, 54(5):1243–1248.

Linton, O., Maasoumi, E., and Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *Review of Economic Studies*, 72(3):735–765.

Maasoumi, E. and Heshmati, A. (2000). Stochastic dominance amongst Swedish income distributions. *Econometric Reviews*, 19(3):287–320.

McCaig, B. and Yatchew, A. (2007). International welfare comparisons and nonparametric testing of multivariate stochastic dominance. *Journal of Applied Econometrics*, 22(5):951–969.

McFadden, D. (1989). Testing for stochastic dominance. In Fomby, T. B. and Seo, T. K., editors, *Studies in the Economics of Uncertainty*. Springer-Verlag, New York.

Tse, Y. K. and Zhang, X. (2004). A Monte Carlo investigation of some tests for stochastic dominance. *Journal of Statistical Computation and Simulation*, 74(5):361–378.

Wolak, F. A. (1989). Testing inequality constraints in linear econometric models. *Journal of Econometrics*, 41(2):205–235.

Table 1: Data generating processes for simulated data

Distribution	μ_X	μ_Y	σ_X^2	σ_Y^2	$\sigma_{X,Y}$
A	2.801	2.801	3.400	3.400	1.737
B	2.509	2.509	5.645	5.645	1.394

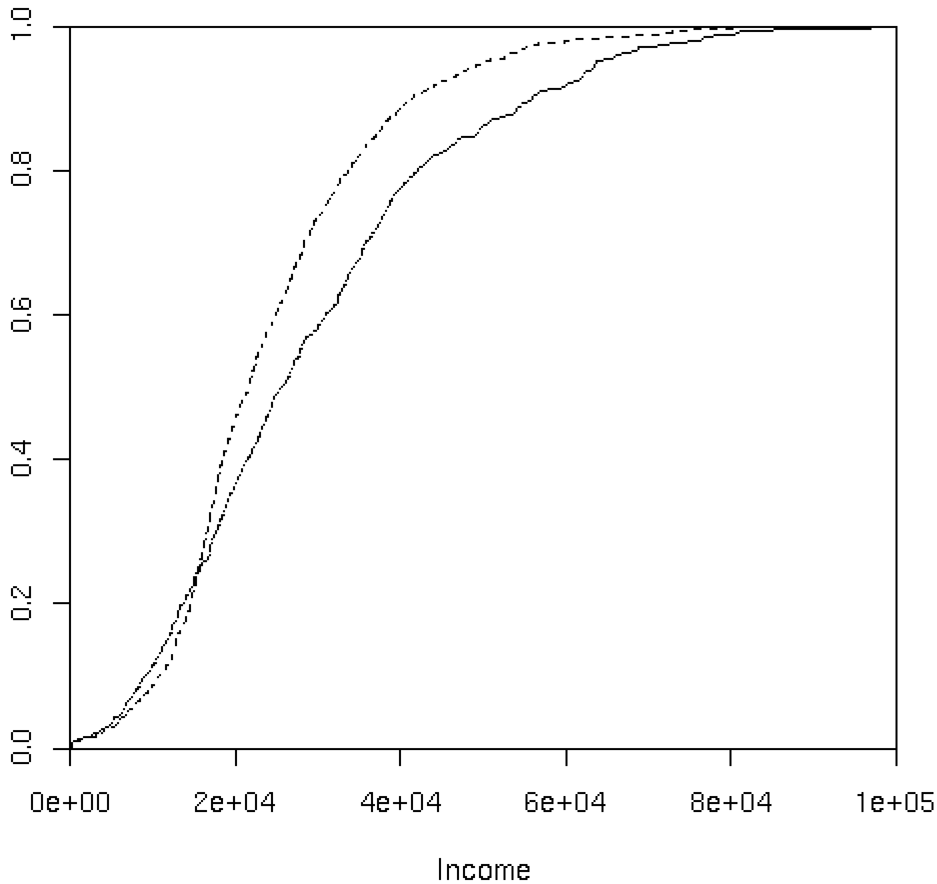
Note: Both distributions belong to the bivariate lognormal family.

Table 2: Rejection frequencies for simulated data

Test statistic	H_0	$n = 50$	$n = 500$
T	$F_A \succeq_1 F_A$	0.045	0.067
	$F_A \succeq_1 F_B$	0.458	1.000
W (with $J = 5$)	$F_A \succeq_1 F_A$	0.051	0.074
	$F_A \succeq_1 F_B$	0.293	0.751
W (with $J = 10$)	$F_A \succeq_1 F_A$	0.067	0.063
	$F_A \succeq_1 F_B$	0.532	1.000

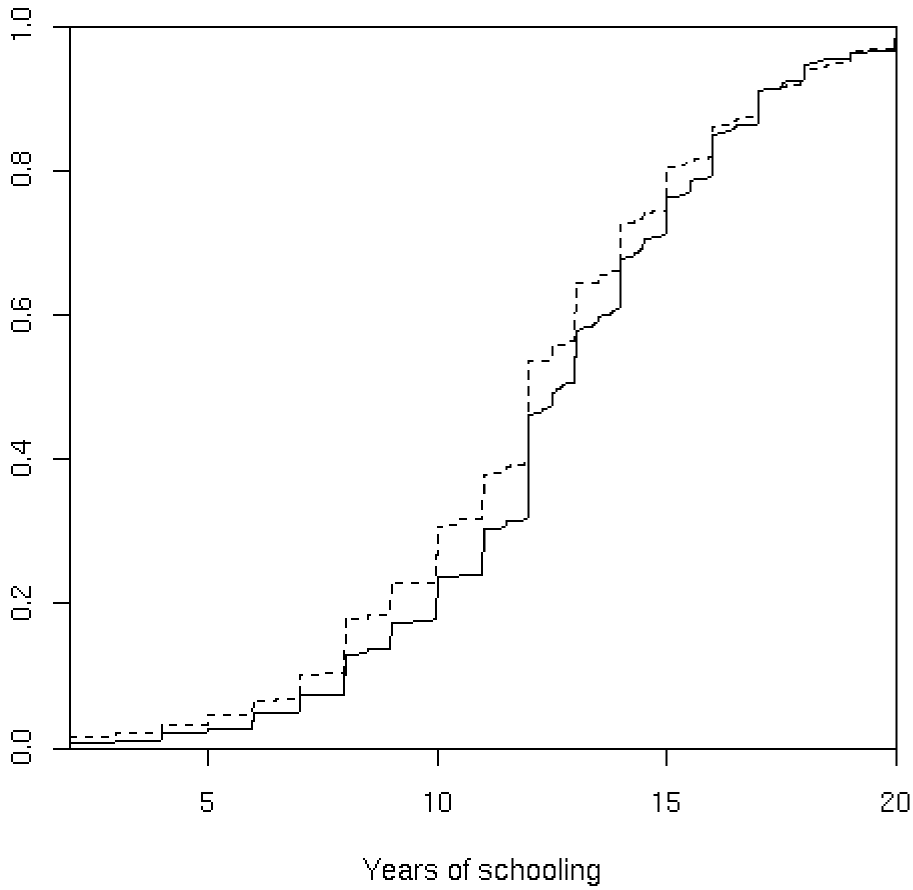
Notes: The nominal level of each test is 0.05. The T test statistic is that of McCaig and Yatchew (2007). W is the Wald-type test statistic proposed here. J is the number of grid points in each dimension (so that J^2 is the total number of grid points).

Figure 1: Empirical marginal distribution functions for income



Note: The solid line is the empirical marginal distribution function for males, and the dashed line is the empirical marginal distribution function for females.

Figure 2: Empirical marginal distribution functions for education



Note: The solid line is the empirical marginal distribution function for males, and the dashed line is the empirical marginal distribution function for females.