Research Paper No. 2005/64

Consistent Testing for Poverty Dominance

Bram Thuysbaert[1] and Ričardas Zitikis[2]

December 2005

Abstract

If uncertainty exists over the exact location of the poverty line or over which measure to use to compare poverty between distributions, one may want to check whether poverty dominance holds. We develop a consistent statistical test to test the null of poverty dominance against the alternative of nondominance. Dominance criteria corresponding to absolute and relative poverty measures are dealt with. The poverty line is allowed to depend on the income distribution. A bootstrap procedure is proposed to estimate critical values for the test. Our results cover both independent and paired samples.

Keywords: poverty, stochastic dominance, random poverty line, bootstrap

JEL classification: C40, I32

[1] Corresponding author: Department of Economics, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium; e-mail: bram.thuysbaert@econ.kuleuven.be

2 Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario Canada.

This study has been prepared within the 2005 UNU-WIDER Internship Programme and is published in the project on New Directions in Development Economics.

UNU-WIDER acknowledges the financial contributions to the research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Norway (Royal Ministry of Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1810-2611 ISBN 92-9190-752-9 (internet version)

Acknowledgements

The research was started during a visit of the second author at the Department of Economics, University of Leuven. The author acknowledges the hospitality of the Department. The first author gratefully acknowledges the hospitality of UNU-WIDER. Both authors are grateful to Geert Dhaene for comments and suggestions

The research of the first author was supported by research grant G.0112.03 from the Flemish Fund for Scientific Research, and the research of the second author was partially supported by a Discovery Research Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by the authors.

1 Introduction

Stochastic dominance criteria have been important instruments in the welfare economist's toolkit since Atkinson's (1970) paper on the measurement of inequality. Many dominance criteria for the measurement of welfare, inequality and poverty have been proposed subsequently. Using dominance tools to compare poverty between distributions is of particular interest: if poverty dominance holds, the poverty ordering is robust both to the choice of the poverty measure and the choice of the poverty line (Atkinson, 1987; Foster and Shorrocks, 1988; Davidson and Duclos, 2000).

To state the dominance results which the present paper focuses on, we introduce the additive poverty measure

$$P(z;F) := \int_{x\leqslant z} \pi(x,z)dF(x), \tag{1}$$

where z is the poverty line, π is a function evaluating the contribution of the income receiver with income x to total poverty in the population, and F is an income distribution function. Throughout the paper, we will consider distribution functions with support contained in the real half-line $[0,\infty)$.[1]

Poverty dominance criteria relate dominance of certain functions to classes of poverty measures. We consider dominance criteria for absolute and relative poverty measures. A poverty measure is said to be *absolute* if its value is unaffected by adding an equal amount to all incomes and the poverty line. We will say that a poverty measure is *relative* if its value is unaffected by scaling all incomes and the poverty line by the same factor. We first introduce the sets of poverty measures and then the dominance functions we need.

With the notation $\pi^{(i)}$ for the i-th order derivative of $\pi(x,z)$ with respect to x, define Π^1 as the set of additive poverty measures such that $\pi^{(1)}(x,z) \leqslant 0$ for $x \in (0,z]$. Let Π^2 denote the set of poverty measures contained in Π^1 such that $\pi^{(2)}(x,z) \geqslant 0$ for $x \in (0,z]$ and $\pi(z,z) = 0$. Further, for a given integer $s \geqslant 3$, define Π^s recursively as the set of additive poverty measures contained in Π^{s-1} with $\pi(x,z)$ satisfying $(-1)^s\pi^{(s)}(x,z) \geqslant 0$ for $x \in (0,z]$ and $\pi^{(s-2)}(z,z) = 0$. Also,

[1] The requirement on the support of the distribution functions to be in $[0,\infty)$ could be removed. The statements of Proposition 1 and Theorem 1 below would then have to be adapted accordingly and some conditions on the left-hand tail of F would be needed. Since the requirement that income be positive is not unnatural, this does not seem to be a generalization worth pursuing from the econometric point of view.

let Π^s_A and Π^s_R denote the sets of absolute and relative poverty measures contained in Π^s respectively.

We now introduce the functions involved in the dominance comparisons. Define the stochastic dominance functions $D_s(x; F)$ recursively in the usual way as $D_1(x; F) := F(x)$ and $D_s(x; F) := \int_{y \leqslant x} D_{s-1}(y; F)dy$ for an integer $s \geqslant 2$. For our purposes, it is convenient also to introduce the functions

$$A_s(x; z; F) := D_s(z - x; F) \quad \text{and} \quad R_s(x; z; F) := \frac{D_s(xz; F)}{z^{s-1}}.$$

We are now in the position to state two basic results for which inferential methods will be developed in the present paper. The first result (cf. statement (2) below) concerns absolute poverty measures and the second one (cf. statement (3) below) concerns relative poverty measures (e.g. Davidson and Duclos, 2000).

Proposition 1 *Let F and G be distribution functions with supports in $[0, \infty)$. Then we have that $P(z_G - x; G) \geqslant P(z_F - x; F)$ for all $x \geqslant 0$ and for all $P \in \Pi^s_A$ if and only if*

$$A_s(x; z_G; G) \geqslant A_s(x; z_F; F) \text{ for all } x \geqslant 0. \tag{2}$$

Furthermore, we have that $P(xz_G; G) \geqslant P(xz_F; F)$ for all $x \in [0, 1]$ and for all $P \in \Pi^s_R$ if and only if

$$R_s(x; z_G; G) \geqslant R_s(x; z_F; F) \text{ for all } x \in [0, 1]. \tag{3}$$

Note that if the poverty lines z_F and z_G are equal, statements (2) and (3) are identical. We would then not have to distinguish between absolute and relative poverty measures and the statements (2) and (3) would hold if and only if $P(x; G) \geqslant P(x; F)$ for all $P \in \Pi^s$ and for all $x \in [0, z]$ with z denoting the common poverty line.

If the income distributions F and G are unknown, formal statistical tests are needed to check statements (2) and (3). Developing testing procedures for stochastic dominance has proved to be quite challenging. Much of the literature deals with the problem of testing whether one dominance function is above the other at a finite number of abscissae. See Davidson and Duclos (2000) and references therein. In this approach, however, the test results depend on the choice of the x-values at which the curves are compared. Moreover, the tests are possibly inconsistent.

Recently, Kolmogorov-Smirnov-type testing procedures have been proposed to test the hypothesis that $D_s(x; G) \geqslant D_s(x; F)$ for all $x \in I \subseteq (-\infty, \infty)$ (Barrett and Donald, 2003; Horváth, Kokoszka and

Zitikis, 2006; Linton, Maasoumi and Whang, 2005). Since these testing procedures do not require the researcher to choose grid points at which to check dominance, they do not suffer from the drawbacks which the former procedures are inextricably bound up with. The present paper develops procedures in much the same spirit to test whether statements (2) and (3) hold. Since in practice poverty lines are often estimated empirically, we allow the poverty lines to be random. This considerably complicates statistical inference as demonstrated previously by Davidson and Duclos (2000) and Zheng (2001). We deal with these issues in the present paper, covering also with a unified approach the cases when samples are independent and paired.

The large sample inferential theory is developed in the next section. Section 3 presents a bootstrap procedure to estimate critical values for our test statistic and Section 4 concludes.

2 Main results

We are interested in testing whether statements (2) and (3) hold. To economize on space, we introduce the generic notation K that would be A in the case of absolute poverty and R in the case of relative poverty. Note that the intervals of the x-values in statements (2) and (3) are different, that is,

$$I^A := [0, \infty) \text{ and } I^R := [0, 1].$$

With the above notation, both statements (2) and (3) can be be concisely stated as only one:

$$K_s(x; z_G; G) \geqslant K_s(x; z_F; F) \text{ for all } x \in I^K. \tag{4}$$

Let statement (4) be our null hypothesis, H_0, and let the alternative, H_1, be the complement to H_0. With the help of the quantity

$$S^K_{F-G} := \sup_{x \in I^K} (K_s(x; z_F; F) - K_s(x; z_G; G)),$$

we can reformulate the above defined hypotheses H_0 and H_1 as

$$H_0 : S^K_{F-G} \leqslant 0 \text{ vs. } H_1 : S^K_{F-G} > 0.$$

Our next task is to construct a test statistic and then develop an appropriate large sample based inferential theory.

Let $X_1, \ldots, X_n$ be independent copies of the random variable X with distribution function F. Furthermore, let $Y_1, \ldots, Y_m$ be independent copies of the random variable Y with distribution function G. The two samples may be independent or dependent. In the former case, the random variables X_i $(i = 1, .., n)$ are independent of all Y_i's. In the latter case, $n = m$ and the pairs $(X_1, Y_1), \dots, (X_n, Y_n)$ are independent bivariate vectors from the joint distribution $H(x, y)$ of X and Y.

The empirical distribution functions corresponding to the X's and Y's are defined in the usual way as, respectively,

$$\widehat{F}(x) := \frac{1}{n}\sum_{i}^{n} \mathbf{1}(X_i \leqslant x) \quad \text{and} \quad \widehat{G}(y) := \frac{1}{m}\sum_{i}^{m} \mathbf{1}(Y_i \leqslant y),$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. Assume now that there exists a number $\eta \in (0, 1)$ such that $n/(n + m) \to \eta$ when both n and m tend to infinity and let '$\Rightarrow$' stand for 'weakly converges to'. Then we have the following statement (that introduces two new notations, $\mathcal{F}$ and $\mathcal{G}$):

$$\sqrt{\frac{nm}{n+m}}\begin{pmatrix}\widehat{F} - F\\ \widehat{G} - G\end{pmatrix} \Rightarrow \begin{pmatrix}\sqrt{1-\eta}\ \mathcal{B}_1 \circ F\\ \sqrt{\eta}\ \mathcal{B}_2 \circ G\end{pmatrix} =: \begin{pmatrix}\mathcal{F}\\ \mathcal{G}\end{pmatrix} \tag{5}$$

with $\mathcal{B}_1$ and $\mathcal{B}_2$ denoting Brownian bridge processes. Obviously, the two coordinates $\mathcal{F}$ and $\mathcal{G}$ of the bivariate limiting process on the right-hand side of (5) are independent if the two samples are independent. If, however, the two samples are paired, then $\eta = 1/2$ and the covariance function of the bivariate limiting process on the right-hand side of (5) is given by

$$\begin{aligned}
&\mathbf{E}\left[\begin{pmatrix}\mathcal{F}(x_1)\\ \mathcal{G}(x_1)\end{pmatrix}\begin{pmatrix}\mathcal{F}(x_2)\\ \mathcal{G}(x_2)\end{pmatrix}'\right]\\
&= \frac{1}{2}\begin{pmatrix}\min(F(x_1), F(x_2)) - F(x_1)F(x_2) & H(x_1, x_2) - F(x_1)G(x_2)\\ H(x_2, x_1) - F(x_2)G(x_1) & \min(G(x_1), G(x_2)) - G(x_1)G(x_2)\end{pmatrix}.
\end{aligned}$$

In many practical applications, the poverty line is a functional of the (unknown) distribution function and so has to be estimated. Denote the estimator of z_F, the poverty line associated with the distribution F, as $\hat{z}_F$ and assume that the following representation holds:

$$\widehat{z}_F - z_F = \int \zeta_F(x)d(\widehat{F}(x) - F(x)) + o_{\mathbf{P}}(n^{-1/2}), \tag{6}$$

where ζ_F is a function such that the integral $\int \zeta_F^2(x)dF(x)$ is finite. An analogous assumption is made for z_G, the poverty line associated with the distribution function G (with the notation $\hat{z}_G$ and ζ_G).

We now present a few examples of poverty lines. First we note that, in principle, the poverty line z_F might be known, and thus its "estimator" $\widehat{z}_F$ can obviously be chosen to be z_F, in which case

$$\zeta_F(x) = 0.$$

The first *nontrivial* example is when z_F is set to a fraction k of the mean of F, in which case we have $\widehat{z}_F = k\bar{X}$ (with $\bar{X}$ the sample mean of the X's) and thus the representation (6) holds with

$$\zeta_F(x) = kx.$$

As a second example, consider a poverty line that is defined as a fraction k of the p^{th} quantile of F, i.e. $z_F = kF^{-1}(p)$, with F^{-1} the left-continuous inverse of F and $p \in (0,1)$. In this case, $\widehat{z}_F$ is a fraction k of the empirical quantile and the Bahadur representation (cf., e.g., Shao, 1999, p. 307) gives (6) with

$$\zeta_F(x) = -\frac{k}{F'(F^{-1}(p))}\mathbf{1}(x \leqslant F^{-1}(p)),$$

where F' is assumed to exist and to be strictly positive at the point $F^{-1}(p)$.

We are now ready to introduce the test statistic for H_0 vs. H_1:

$$\widehat{S}^K_{F-G} := \sqrt{\frac{nm}{n+m}} \sup_{x \in I^K} \left(K_s(x; \widehat{z}_F; \widehat{F}) - K_s(x; \widehat{z}_G; \widehat{G}) \right).$$

The following theorem provides a random variable (for $K = A$ and $K = R$) which dominates the limiting distribution of the test statistic under the null hypothesis, and thus determines the (conservative) critical values of the test.

Theorem 1 *Let F and G be distribution functions with supports in $[0,\infty)$. If the poverty lines z_F and z_G are unknown, assume, in addition to representation (6) for both z_F and z_G, that the functions $\frac{d}{dz}K_s(x;z;F)\big|_{z=z_F}$ and $\frac{d}{dz}K_s(x;z;G)\big|_{z=z_G}$ are continuous in x at every point in I^K. Then, under H_0 we have that, for every $c > 0$,*

$$\limsup_{n,m\to\infty} \mathbf{P}(\widehat{S}^K_{F-G} > c) \leqslant \mathbf{P}(\Lambda^K_{F-G} > c) \tag{7}$$

when n and m tend to infinity in such a way that $n/(n+m) \to \eta \in (0,1)$, and where

$$\Lambda^K_{F-G} := \sup_{x\in I^K} \Gamma^K_{F-G}(x)$$

with the Gaussian process

$$\begin{aligned}\Gamma^K_{F-G}(x) := \;& K_s(x; z_F; \mathcal{F}) + \frac{d}{dz}K_s(x;z;F)\Big|_{z=z_F} \int \zeta_F(y)d\mathcal{F}(y) \\ -\;& K_s(x; z_G; \mathcal{G}) - \frac{d}{dz}K_s(x;z;G)\Big|_{z=z_G} \int \zeta_G(y)d\mathcal{G}(y).\end{aligned}$$

Furthermore, under H_1, the test statistic $\widehat{S}^K_{F-G}$ converges in probability to infinity, that is,

$$\lim_{n,m\to\infty} \mathbf{P}(\widehat{S}^K_{F-G} > c) = 1$$

for every $c > 0$.

The proof of Theorem 1 is given in the Appendix. We now discuss the theorem and the assumptions imposed. The empirical estimation of the quantiles of the distribution of Λ^K_{F-G} is discussed in the next section.

We start with a note on the lim sup on the left-hand side of (7). In fact, it can be replaced by lim since the limit, as we shall show in a moment, does exist. Indeed, a little additional analysis of the proof of Theorem 1 reveals that under H_0 the test statistic $\widehat{S}^K_{F-G}$ converges in distribution to the random variable

$$\Lambda^K_{F-G}(E^K) := \sup_{x\in E^K} \Gamma^K_{F-G}(x)$$

if the set $E^K := \{x \in I^K : K_s(x; z_G; G) = K_s(x; z_F; F)\}$ is non-empty, and converges to $-\infty$ if this set is empty. Since $E^K \subseteq I^K$ and Λ^K_{F-G} equals $\Lambda^K_{F-G}(I^K)$, we have the inequality $\Lambda^K_{F-G} \geqslant \Lambda^K_{F-G}(E^K)$. This clarifies our earlier note that the quantiles of Λ^K_{F-G} provide *conservative* critical values.

We have already noted that the poverty lines z_F and z_G may in principle be known. In this case, the continuity assumption regarding $\frac{d}{dz}K_s(x; z; F)\big|_{z=z_F}$ and $\frac{d}{dz}K_s(x; z; G)\big|_{z=z_G}$ is unnecessary and the theorem holds with $\zeta_F(x) = 0$ and $\zeta_G(x) = 0$.

Suppose now that the poverty line z_F depends on the (unknown) population distribution function F and is therefore estimated using an estimator $\widehat{z}_F$ that satisfies the representation (6). Then Theorem 1 requires that the function

$$\kappa_s(x) := \frac{d}{dz}K_s(x; z; F)\Big|_{z=z_F}$$

be continuous at every point x of the interval I^K. For practical applicability of the theorem, we reformulate the continuity assumption on $\kappa_s(x)$ in terms of the distribution F. From the definition of the functions $K_s(x; z; F)$, note that $\kappa_s(x)$ is continuous at every point x of the interval I^K if the function $dD_s(x; F)/dx$ is continuous on $(-\infty, z_F + \epsilon)$ for some $\epsilon > 0$. Since for $s \geqslant 3$ the function $dD_s(x; F)/dx = D_{s-1}(x; F)$ is continuous for every F, the requirement that $\kappa_s(x)$ be continuous is always satisfied for $s \geqslant 3$. For $s = 2$, we have $D_{s-1}(x; F) = F(x)$, and so the continuity assumption is implied by the continuity of F on $(-\infty, z_F + \epsilon)$. For $s = 1$, the continuity assumption is implied by the continuity of F' on $(-\infty, z_F + \epsilon)$.

3 Estimating critical values

For the result in Theorem 1 to be of practical use, critical values are needed. Note, however, that the distribution of the random variable Λ^K_{F-G} depends on the (unknown) joint probability distribution of X and Y. Hence, we have to resort to simulation techniques to estimate the quantiles of the distribution of Λ^K_{F-G}. Here we propose a bootstrap technique. For this and other simulation methods, see Barrett and Donald (2003), and Linton, Maasoumi and Whang (2005).

If the two samples, that is, the X's and Y's are independent, then we let $X^*_1, \ldots, X^*_n$ be a simple random sample of size n from $\widehat{F}$, and let $Y^*_1, \ldots, Y^*_m$ be a simple random sample of size m from $\widehat{G}$. Denote the empirical distribution functions of the bootstrap samples as $\widehat{F}^*$ and $\widehat{G}^*$, respectively.

If, however, the samples are dependent, then we draw a simple random sample, say $(X^*_1, Y^*_1), \ldots, (X^*_n, Y^*_n)$ of size n from the joint empirical distribution function $\widehat{H}$ of the original pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$. Denote the empirical distribution functions of the first and second coordinates of the above bootstrap samples as $\widehat{F}^*$ and $\widehat{G}^*$, respectively. In both the independent and the paired case discussed above, denote the bootstrap version of the empirical poverty line $\hat{z}_F$ as $\widehat{z}^*_F$ and likewise $\widehat{z}^*_G$ for $\hat{z}_G$.

With these notations, define the statistic

$$B^K_{F-G} := \sqrt{\frac{nm}{n+m}} \sup_{x \in I^K} \Big(K_s(x; \widehat{z}^*_F; \widehat{F}^*) - K_s(x; \widehat{z}_F; \widehat{F}) \\ - K_s(x; \widehat{z}^*_G; \widehat{G}^*) + K_s(x; \widehat{z}_G; \widehat{G}) \Big).$$

Since, when both n and m tend to infinity,

$$\sqrt{\frac{nm}{n+m}} \begin{pmatrix} \widehat{F}^* - \widehat{F} \\ \widehat{G}^* - \widehat{G} \end{pmatrix} \Rightarrow \begin{pmatrix} \mathcal{F} \\ \mathcal{G} \end{pmatrix},$$

arguments similar to the ones used in the proof of Theorem 1 can be used to show that B^K_{F-G} converges in distribution to Λ^K_{F-G}. In view of this, the quantiles of the distribution of B^K_{F-G} will give asymptotically correct critical values for our testing problem. More formally, let $\mathbf{P}^*$ be the probability $\mathbf{P}$ conditioned on the random variables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, and let $H^*(c) := \mathbf{P}^*\{B^K_{F-G} \leqslant c\}$. The distribution function $H^*(c)$ is an empirical estimator of the theoretical distribution function $H(c) := \mathbf{P}(\Lambda^K_{F-G} \leqslant c)$. Hence,

$$c^*_\alpha := \inf\{c \geqslant 0 : \; H^*(c) \geqslant 1 - \alpha\}$$

is an empirical estimator of $c_\alpha := \inf\{c \geqslant 0 : \ H(c) \geqslant 1-\alpha\}$, which is the critical value provided by Theorem 1. Of course, for the desired result $c_\alpha^* \to_{\mathbf{P}} c_\alpha$ to hold, we need the continuity of the distribution function $H(c)$. This is implied by results of Tsirel'son (1975). Specifically, Tsirel'son shows that suprema of Gaussian processes have continuous distribution functions except in pathological cases such as degenerate processes. In summary, under the null hypothesis H_0 we have that

$$\limsup_{n,m\to\infty} \mathbf{P}(\widehat{S}^K_{F-G} > c^*_\alpha) \leqslant \alpha,$$

and under the alternative H_1 we have that $\mathbf{P}(\widehat{S}^K_{F-G} > c^*_\alpha) \to 1$. The asymptotic p-value $1 - H(\widehat{S}^K_{F-G})$ (or empirical size) of the test can be approximated as

$$\mathbf{P}^*\{B^K_{F-G} > \widehat{S}^K_{F-G}\}.$$

4 Conclusions

Dominance criteria are powerful tools to compare income distributions in terms of welfare, inequality or poverty. In this paper, we focus on poverty dominance. Dominance criteria corresponding to both absolute and relative poverty measures are considered. A consistent test is proposed to test the null of dominance against the alternative of nondominance. We present a simple to execute bootstrap procedure to estimate critical values for the test. Since in practice poverty lines are often functionals of the population distribution functions and samples may be dependent, our testing theory is developed to accommodate these situations.

References

Atkinson, A.B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2, 244-263.

Atkinson, A.B. (1987). On the measurement of poverty. *Econometrica*, 55, 749-764.

Barrett, G. and Donald, S. (2003). Consistent tests for stochastic dominance. *Econometrica*, 71, 71-104.

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

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

Horváth, L., Kokoszka, P. and Zitikis, R. (2006). Testing for stochastic dominance using the weighted McFadden type statistic. *Journal of Econometrics*. (Forthcoming)

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

Shao, J. (1999). *Mathematical Statistics*. Springer, New York.

Tsirel'son, B. S. (1975). The density of the distribution of the maximum of a Gaussian process. *Theory Probab. Appl.*, 20, 847-856.

Zheng, B. (2001). Statistical inference for poverty measures with relative poverty lines. *Journal of Econometrics*, 101, 337-356 .

Appendix

Proof of Theorem 1. Under H_0, the function $K_s(x; z_F; F)$ should not exceed $K_s(x; z_G; G)$ for any $x \in I^K$. Hence, the difference between the functions must be non-positive, and so we have the bound

$$\begin{aligned} &\sqrt{\frac{nm}{n+m}} \sup_{x\in I^K} \Big(K_s(x;\widehat{z}_F;\widehat{F}) - K_s(x;\widehat{z}_G;\widehat{G}) \Big) \\ &\leqslant \sqrt{\frac{nm}{n+m}} \sup_{x\in I^K} \Big(K_s(x;\widehat{z}_F;\widehat{F}) - K_s(x;z_F;F) - K_s(x;\widehat{z}_G;\widehat{G}) + K_s(x;z_G;G) \Big) . \end{aligned} \tag{8}$$

To prove Theorem 1, we have to show that the right-hand side of (8) converges in distribution to Λ^K_{F-G}. This we do with the help of the following weak convergence result that we shall prove below:

$$\begin{aligned} \sqrt{\frac{nm}{n+m}} \Big(& K_s(\,\cdot\,;\widehat{z}_F;\widehat{F}) - K_s(\,\cdot\,;z_F;F) \\ & - K_s(\,\cdot\,;\widehat{z}_G;\widehat{G}) + K_s(\,\cdot\,;z_G;G) \Big) \Rightarrow \Gamma^K_{F-G}. \end{aligned} \tag{9}$$

Note that the supremum of the limiting process in (9) with respect to $x \in I^K$ is exactly the random variable Λ^K_{F-G} defined in the formulation of the theorem.

We start the proof of (9) by writing

$$\begin{aligned} K_s(x;\widehat{z}_F;\widehat{F}) - K_s(x;z_F;F) = & \big[K_s(x;\widehat{z}_F;F) - K_s(x;z_F;F) \big] \\ & + K_s(x;z_F;\widehat{F}-F) \\ & + \big[K_s(x;\widehat{z}_F;\widehat{F}-F) - K_s(x;z_F;\widehat{F}-F) \big], \end{aligned} \tag{10}$$

where we have used the fact that the function $K_s(x; z_F; F)$ is linear in the third argument.

For the first term, $K_s(x; \widehat{z}_F; F) - K_s(x; z_F; F)$, on the right-hand side of (10), we can write

$$\begin{aligned} K_s(x; \widehat{z}_F; F) - K_s(x; z_F; F) &= \frac{d}{dz} K_s(x; z; F)\Big|_{z=z_F} (\widehat{z}_F - z_F) + R_{1,n}(x) \\ &= \frac{d}{dz} K_s(x; z; F)\Big|_{z=z_F} \int \zeta_F(x) d[\widehat{F}(x) - F(x)] + R_{2,n}(x), \end{aligned}$$

where the two remainder terms $R_{1,n}(x)$ and $R_{2,n}(x)$ are uniformly $o_{\mathbf{P}}(n^{-1/2})$. Indeed, the aforementioned order of $R_{1,n}(x)$ follows from the continuity of $\frac{d}{dz} K_s(x; z; F)\big|_{z=z_F}$. For the second equality and thus the claimed order of $R_{2,n}(x)$, we have used (6).

To show that the third term, $K_s(x; \widehat{z}_F; \widehat{F} - F) - K_s(x; z_F; \widehat{F} - F)$, on the right-hand side of (10) is asymptotically negligible, we consider the two cases $K = A$ and $K = R$ separately.

When $K = A$, we deal with the process $D_s(\widehat{z}_F - x; \widehat{F} - F) - D_s(z_F - x; \widehat{F} - F)$ indexed by $x \in I^A$. Since $\sqrt{n}(D_s(x - z; \widehat{F} - F))$ converges weakly to the x-indexed Gaussian process $D_s(x - z; \mathcal{B} \circ F)$, we have by the stochastic equicontinuity of the limiting process and by the fact that $|\widehat{z}_F - z_F| = o_{\mathbf{P}}(1)$ that

$$\sup_{x \in I^A} \Big| D_s(\widehat{z}_F - x; \widehat{F} - F) - D_s(z_F - x; \widehat{F} - F) \Big| = o_{\mathbf{P}}(n^{-1/2}).$$

This is the desired result.

When $K = R$, we rewrite $R_s(x; \widehat{z}_F; \widehat{F} - F) - R_s(x; z_F; \widehat{F} - F)$ as the sum

$$\begin{aligned} &\frac{1}{[z_F]^{s-1}} \Big(D_s(x\widehat{z}_F; \widehat{F} - F) - D_s(x z_F; \widehat{F} - F) \Big) \\ &\qquad + D_s(x\widehat{z}_F; \widehat{F} - F) \left(\frac{1}{[\widehat{z}_F]^{s-1}} - \frac{1}{[z_F]^{s-1}} \right). \end{aligned} \tag{11}$$

Again, since $\sqrt{n}(D_s(zx; \widehat{F} - F))$ converges weakly to the x-indexed Gaussian process $D_s(zx; \mathcal{B} \circ F)$, we have by the stochastic equicontinuity of the limiting process and by the fact that $|\widehat{z}_F - z_F| = o_{\mathbf{P}}(1)$ that the supremum over all $x \in I^R$ of the absolute value of the first term in (11) is $o_{\mathbf{P}}(n^{-1/2})$. As to the supremum over all $x \in I^R$ of the absolute value of the second term in (11), we have that it does not exceed

$$\begin{aligned}
&c\,|\widehat{z}_F - z_F| \sup_{x\in I^R} D_s(x\widehat{z}_F; \widehat{F} - F) + o_{\mathbf{P}}(n^{-1/2}) \\
&= O_{\mathbf{P}}(n^{-1/2})O_{\mathbf{P}}(n^{-1/2}) + o_{\mathbf{P}}(n^{-1/2}) \\
&= o_{\mathbf{P}}(n^{-1/2}).
\end{aligned}$$

Combining the above results in the case of F, and since similar results hold for the distribution function G, we have

$$\begin{aligned}
&K_s(x;\widehat{z}_F;\widehat{F}) - K_s(x;z_F;F) - K_s(x;\widehat{z}_G;\widehat{G}) + K_s(x;z_G;G) \\
&\qquad = K_s(x;z_F;\widehat{F}-F) + \frac{d}{dz}K_s(x;z;F)\Big|_{z=z_F}\int \zeta_F(x)d\big(\widehat{F}(x)-F(x)\big) \\
&\qquad\quad - K_s(x;z_G;\widehat{G}-G) - \frac{d}{dz}K_s(x;z;G)\Big|_{z=z_G}\int \zeta_G(x)d\big(\widehat{G}(x)-G(x)\big) \\
&\qquad\quad + R_{3,n}(x),
\end{aligned}$$

where the remainder term $R_{3,n}(x)$ is uniformly $o_{\mathbf{P}}(n^{-1/2})$. Now, multiply the above equation by $\sqrt{nm/(n+m)}$ and use (5) to obtain the earlier made claim in (9).

To complete the proof, we have to show that under the alternative H_1, the test statistic $\widehat{S}^K_{F-G}$ converges in probability to infinity. Note that, under H_1, there exists $x^* \in I^K$ such that $K_s(x^*;z_F;F) - K_s(x^*;z_G;G) > 0$. We can write the inequality

$$\widehat{S}^K_{F-G} \geqslant \sqrt{\frac{nm}{n+m}}\left(K_s(x^*;\widehat{z}_F;\widehat{F}) - K_s(x^*;\widehat{z}_G;\widehat{G})\right).$$

Since the right-hand side converges in probability to infinity, so does the left-hand side. ■