C.I.R.P.É.E

Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **10-39**

Testing for Welfare Comparisons when Populations Differ in Size

Jean-Yves Duclos

Agnès Zabsonré

Septembre/September 2010

Duclos: Département d'économique, PEP and CIRPÉE, Université Laval, Canada
jyduc@ecn.ulaval.ca
Zabsonré: CIRPÉE, Université Laval, Canada
wendgloumde-agnes.zabsonre.1@ulaval.ca

This work was carried out with support from SSHRC, FQRSC and the Poverty and Economic Policy (PEP) Research Network, which is financed by the Government of Canada through the International Development Research Centre and the Canadian International Development Agency, and by the Australian Agency for International Development. We are grateful to Araar Abdelkrim, Sami Bibi, John Cockburn, Paul Makdissi and participants at the 2010 SCSE and CEA annual conferences for useful comments and advice.

Abstract:
Assessments of social welfare do not usually take into account population sizes. This can lead to serious social evaluation flaws, particularly in contexts in which policies can affect demographic growth. We develop in this paper a little-known though ethically attractive approach to correcting the flaws of traditional welfare analysis, an approach that is population-size sensitive and that is based on critical-level generalized utilitarianism (CLGU). Traditional CLGU is extended by considering arbitrary orders of welfare dominance and ranges of “poverty lines” and values for the “critical level” of how much a life must be minimally worth to contribute to social welfare. Simulation experiments briefly explore the normative relationship between population sizes and critical levels. We apply the methods to household level data to rank Canada’s social welfare across 1976, 1986, 1996 and 2006 and to estimate normatively and statistically robust lower and upper bounds of critical levels over which these rankings can be made. The results show dominance of recent years over earlier ones, except when comparing 1986 and 1996. In general, therefore, we conclude that Canada’s social welfare has increased over the last 35 years in spite (or because) of a substantial increase in population size.

Keywords: CLGU, welfare dominance, FGT dominance, estimation of critical levels, welfare in Canada

JEL Classification: C12, D31, D63, I30

1 Introduction

Is the "value" of a society increasing with its population size? How can that question be dealt with in a normatively robust framework? What sort of statistical procedures can assess this empirically? What does the evidence actually suggest? To address these questions is the main objective of this paper.

Poverty and welfare comparisons are routinely made under the implicit assumptions that population sizes do not matter, or equivalently that population sizes are the same. Technically, this is implicitly or explicitly done by calling on the so-called population replication invariance axiom. The population replication invariance axiom says that an income distribution and its k-fold replication, with k being any positive integer, should be deemed equivalent from a social welfare perspective. Welfare and inequality comparisons can then be performed in *per capita* terms.

However, as Blackorby, Bossert, and Donaldson (2005) and others have argued, population size should probably matter when assessing social welfare. We may not be indifferent, for instance, to whether some income (or GDP) statistics are expressed in *per capita* or in total terms. When total income changes in a society, we may wish to know whether this is due to changes in population size or changes in *per capita* income; when *per capita* income changes, we may also wish to know whether this is associated with a change in population size. Generally speaking, our assessment of the welfare value of a change in the distribution of incomes may depend on how population size also changes.

In addressing these issues — which we believe to be important ones — our work adopts as a conceptual framework for social welfare comparisons the "critical-level generalized utilitarianism" (CLGU) principle of Blackorby and Donaldson (1984). CLGU essentially says that adding a person to an existing population will increase social welfare if and only if that person's income exceeds the value of a *critical level*. From a normative perspective, the critical level can be interpreted as the minimum income needed for someone to add "value" to humanity. (The critical level has been termed *the value of living* by Broome (1992b).) Social welfare according to CLGU is then defined as the sum of the differences between some transformation of individual incomes and the same transformation of the critical level.

CLGU is a social evaluation approach that is both normatively attractive and (surprisingly) little known; it has also not yet (to our knowledge) been tested and applied. There are, however, two major difficulties in implementing CLGU. First, it is difficult in practice to agree on a non-arbitrary value for the critical level. In a world of heterogenous preferences and opinions, it is indeed difficult to envisage a relatively wide consensus on something as fundamentally un-consensual as the "value of living". Second, it is also difficult to agree on which transformation to apply to individual incomes when computing social welfare. We get around these difficulties in this paper by applying stochastic dominance methods for making population comparisons under a CLGU framework. This avoids having to specify a particular form for the transformation of individual incomes. This also enables assessing the ranges of critical levels over which normatively robust CLGU comparisons can be made. In a poverty comparison context, it also makes it possible to derive the ranges of poverty lines over which robust CLGU comparisons can be obtained.

Although the paper's main objective in this paper is to compare welfare through CLGU, the use of CLGU for social evaluation purposes has important implications for the design of policy and for the analysis and monitoring of human development in general. According to CLGU , the socially optimal population size maximizes the product of population size and the difference between a single-individual "socially representative income" and the critical level. This results in policy prescriptions that optimize the trade-off between population size and some measure of *per capita* well-being in excess of the critical level.

For instance, the process of demographic transition (through a reduction of both fertility and mortality) in which a large part of humanity has recently engaged is often rationalized as one that maximizes *per capita* welfare under resource constraints. It is unlikely for developed countries that this process also maximizes social welfare in a CLGU perspective. As we will also see in our illustration, Canada's CLGU has robustly increased in the last 35 years despite a significant increase in population size. For developed countries, such a social evaluation perspective can thus provide a rationale for promoting policies that encourage fertility, such as the provision of relatively generous child benefits for families with many children.

Whether the current demographic transition is consistent with CLGU maximization in developing countries depends much on the value that is set for the critical level. A social planner would favor a population increase only if the additional persons enjoyed a level of income at least equal to that level. This would be more difficult to achieve in less developed countries, where average income is lower relative to the critical level, so a smaller population might then be desirable. Optimal policies would then aim to increase *per capita* income and raise social welfare by limiting demographic growth (particularly of the poor people). This could involve compulsory measures of birth control for the poor and measures for increasing the life years (only) of the more affluent.

The use of CLGU thus enables social evaluations to be made when the distributions and policy outcomes to be compared involve varying population sizes. These are certainly the most generally encountered cases in theory and in practice. This is also almost always the appropriate setting when making welfare comparisons across time.

A few papers have recently considered comparisons of populations of unequal sizes without using the replication-invariance axiom. One of the most recent is Aboudi, Thon, and Wallace (2010), who generalize the well-known concept of majorization and suggest that an income distribution should be deemed more equal than another one if the first distribution can be constructed from the second distribution through linear transformations of incomes. Pogge (2007) proposes the use of the Pareto criterion to compare social welfare in income distributions with different numbers of individuals. Considering only the most well-off persons in the larger population (such that their number be equal to the size of the smaller population), Pogge (2007) suggests that social welfare in the larger population should be greater than in the smaller population if every person in the larger population reduced to the size of the smaller one enjoys a level of well-being greater than that of every person in the smaller population. Other relatively recent interesting contributions include Broome (1992b), Mukherjee (2008) and Gravel, Marchant, and Sen (2008). Our paper differs from these earlier papers by focussing on how to rank distributions and outcomes normatively and empirically using CLGU-based dominance criteria.

The paper's normative setting is described in Section 2, where CLGU is introduced and motivated and social welfare dominance relations are defined. Section 2 also discusses how this relates to well-known poverty dominance criteria. This dominance context extends Blackorby and Donaldson (1984)'s focus on CLGU indices. It also builds on the theoretical contribution of Trannoy and Weymark (2009), who proposes a CLGU dominance criterion that is an extension of generalized Lorenz dominance and second-order welfare dominance.

Section 3 presents the statistical framework that is used for analyzing dominance relations, both in terms of estimation and inference. It also develops the apparatus necessary to estimate normatively robust ranges of critical levels. Section 4 provides the results of a few simulation experiments that show how and why population size may be of concern — normatively and statistically — for social welfare rankings.

Section 5 applies the methods to comparable Canadian Surveys of Consumer Finances (SCF) for 1976 and 1986 and Canadian Surveys of Labour and Income Dynamics (SLID) for 1996 and 2006. Canada's population size has increased by almost 50% between 1976 and 2006. We assess whether social welfare has increased or decreased over that period in Canada, allowing for variations in population size and income distributions and using ranges of "poverty lines" (or censoring points) and values of critical levels. Using asymptotic and bootstrap tests, we find that Canada's welfare has globally improved in the last 35 years despite the substantial increase in population size and the fact that new lives do not necessarily increase society's value in a CLGU framework. More surprisingly perhaps, Canada's smaller population in 1986 is nevertheless socially better than Canada's larger population in 1996 for a relatively wide range of critical levels and despite a significant increase in average and total income. Hence, not only can average and total utilitarianism present significant ethical weaknesses, but their social evaluation rankings can differ importantly from those derived from critical-level utilitarianism. Section 6 concludes.

2 CLGU: an alternative approach to assessing social welfare

2.1 Average and total utilitarianism

The most popular methods to assess social welfare in the context of variable population sizes are based on average utilitarianism. Using average utilitarianism as a social evaluation criterion implicitly assumes that population sizes should not matter. One consequence of this is that a population with only one individual will dominate any other population of arbitrarily larger size as long as those larger populations' average utility is (perhaps only slightly) smaller than the single person's utility level — see for instance Cowen (1989), Broome (1992a), Blackorby, Bossert, and Donaldson 2005, and Kanbur and Mukherjee (2007). This social evaluation framework would seem to be too biased against population size: it would say for instance that a society made of a single very rich person (Bill Gates for example) would be preferable to *any* other society of greater size but lower average utility.

An alternatively popular social evaluation criterion is total utilitarianism. Adopting total utilitarianism leads, however, to Parfit (1984)'s "repugnant conclusion". Parfit (1984)'s "repugnant conclusion" bemoans the implication that, with total utilitarianism, a sufficiently large population will necessarily be considered better than any other smaller population, even if the larger population has a very low average utility:

> *For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.* (Parfit 1984, p.388).

Such a social evaluation framework again seems to be too strongly biased, this time against average utility.

2.2 Critical-level generalized utilitarianism

Blackorby and Donaldson (1984) have proposed CLGU as an alternative to (and in order to address the flaws of) average and total utilitarianism. To see how CLGU is defined, consider two populations of different sizes. The smaller population of size M has a distribution of incomes (or some other indicator of individual welfare) given by the vector $\mathbf{u}$, and the larger population of size N has a distribution of incomes given by the vector $\mathbf{v}$, with $M < N$. Let $\mathbf{u} := (u_1, u_{2,}..., u_M)$, where u_i being the income of individual i, and $\mathbf{v} := (v_1, v_{2,}..., v_N)$ with v_j being the income of individual j. Let the level of social welfare in $\mathbf{u}$ and $\mathbf{v}$ be given by

$$W(\mathbf{u}; \alpha) = \sum_{i=1}^{M} (g(u_i) - g(\alpha)) \tag{1}$$

and

$$W(\mathbf{v}; \alpha) = \sum_{j=1}^{N} (g(v_j) - g(\alpha)), \tag{2}$$

where g is some increasing transformation of incomes and α is a "critical level". Note that social welfare in the two populations remains unchanged when a new individual with income equal to α is added to the population. The smaller population exhibits greater social welfare than the larger one given this if and only if $W(\mathbf{u}; \alpha) \geq W(\mathbf{v}; \alpha)$.

CLGU thus aggregates the differences between transformations of individual incomes and of a critical level. It can therefore avoid some of average utilitarianism's problems, since the addition of a new person will be socially profitable if that person's income is higher than the critical level, although that income may not necessarily be higher than average income. CLGU can also avoid the "repugnant conclusion" since it is socially undesirable to add individuals with incomes lower than the critical level, regardless of how many there may be of them. Overall, CLGU provides a relatively appealing and transparent basis on which to make social evaluations and avoid the flaws associated to average and total utilitarianism.

Suppose now that we may wish to focus on those income values below some censoring point z. This is a typical procedure in poverty analysis. Suppose that z^+ is the maximum possible level for such a censoring point (or maximum "poverty line" in a poverty context). Also denote $\mathbf{u}_\alpha := (\mathbf{u}, \alpha, ..., \alpha)$ as $\mathbf{u}$ "*expanded to size of population* $\mathbf{v}$" by adding $N - M$ α elements. For a poverty line z, the well-known FGT (Foster, Greer, and Thorbecke 1984) poverty indices with parameter $s - 1$ (order s in what follows) for distribution $\mathbf{v}$ are defined as

$$P^s_{\mathbf{v}}(z) = \frac{1}{N}\sum_{i=j}^{N}(z - v_j)^{s-1} I\left(v_j \leq z\right), \tag{3}$$

where $I(\cdot)$ is an indicator function with value set to 1 if the condition is true and to 0 if not. Similarly, the FGT indices for the expanded population $\mathbf{u}_\alpha$ are defined as

$$\begin{aligned} P^s_{\mathbf{u}_\alpha}(z) &= \frac{M}{N}\sum_{i=1}^{M}(z - u_i)^{s-1} I\left(u_i \leq z\right) \\ &\quad + \left(1 - \frac{M}{N}\right)(z - \alpha)^{s-1} I\left(\alpha \leq z\right). \end{aligned} \tag{4}$$

These expressions will be useful to test for CLGU dominance.

2.3 CLGU dominance

The welfare functions in (1) and (2) depend on g and α. One could choose a specific functional form for g and a specific value for α, but that would be inconvenient in the sense that the welfare rankings of $\mathbf{u}$ and $\mathbf{v}$ could then be criticized as depending on those choices. It is thus useful to consider making welfare rankings that are valid over classes of functions g and ranges of critical levels α. To do this, let s =1,2,..., stand for an order of "welfare dominance". Consider $\mathcal{C}^s$ as the set of functions $\mathbb{R} \longrightarrow \mathbb{R}$ that are s times continuously differentiable. Define the class $\mathcal{F}^s_{z^-,z^+}$ of functions as

$$\mathcal{F}^s_{z^-,z^+} := \left\{ g^z \in \mathcal{C}^s \left| \begin{array}{c} z \leq z^+, \\ g^z(x) = g^z(z) \text{ for all } x > z, \\ g^z(x) = g^z(z^-) \text{ for all } x < z^-, \\ \text{and where } -(1)^k \frac{d^k g^z(x)}{dx^k} \leq 0 \ \forall k = 1, ..., s. \end{array} \right. \right\} \tag{5}$$

Also denote W^s_{α,z^-,z^+} as the set of CLGU social welfare functions with $g^z \in \mathcal{F}^s_{z^-,z^+}$ and critical level α. For any vector of income $\mathbf{v} \in \mathbb{R}^N_+$, $N \geq 1$, this set is defined as:

$$W^s_{\alpha,z^-,z^+} := \left\{ W \left| W(\mathbf{v};\alpha) = \sum_{i=1}^{N} (g^z(v_i) - g^z(\alpha)) \text{ where } g^z \in \mathcal{F}^s_{z^-,z^+} \text{ and } \mathbf{v} \in \mathbb{R}^N \right. \right\}. \tag{6}$$

The first and third lines in (5) say that the censoring point z must be below some upper level z^+. The second line says that for social evaluation purposes we can set to z^- those incomes that are lower than z^- — this assumption is mostly made for statistical tractability reasons, to which we come back later. The fourth line on the derivatives of g^z imposes that the social welfare functions be Paretian (for $k = 1$), be concave and thus increasing with a transfer from a richer to a poorer person (for $k = 2$), be transfer-sensitive in the sense of Shorrocks (1987) (for $k = 3$), *etc.*. The greater the order s, the more sensitive is social welfare to the income levels of the poorest.

We can then define the (partial) CLGU dominance ordering $\succsim^{sW}_{\alpha,z^-,z^+}$ as

$$\mathbf{u} \succsim^{sW}_{\alpha,z^-,z^+} \mathbf{v} \Leftrightarrow W(\mathbf{u};\alpha) \geq W(\mathbf{v};\alpha) \, \forall W \in W^s_{\alpha,z^-,z^+}. \tag{7}$$

The welfare ordering (7) considers $\mathbf{u}$ to be better than $\mathbf{v}$ if and only if $W(\mathbf{u};\alpha)$ is greater than $W(\mathbf{v};\alpha)$ for all of the functions W that belong to W^s_{α,z^-,z^+}.

Similarly, define the (partial) FGT dominance ordering $\succsim^{sP}_{z^-,z^+}$ as

$$\mathbf{u}_\alpha \succsim^{sP}_{z^-,z^+} \mathbf{v} \Leftrightarrow P^s_{\mathbf{u}_\alpha}(z) - P^s_{\mathbf{v}}(z) \leq 0 \text{ for all } z^- \leq z \leq z^+. \tag{8}$$

This FGT ordering (8) considers $\mathbf{u}$ to be better than $\mathbf{v}$ if and only if the FGT curve $P^s_{\mathbf{u}_\alpha}(z)$ for $\mathbf{u}_\alpha$ is always below the FGT curve $P^s_{\mathbf{v}}(z)$ for $\mathbf{v}$ for all values of $z^- \leq z \leq z^+$.

Duclos and Zabsonré (2009) demonstrate that the two partial orderings are equivalent, for some α, z^- and z^+:

$$\mathbf{u} \succsim^{sW}_{\alpha,z^-,z^+} \mathbf{v} \Leftrightarrow \mathbf{u}_\alpha \succsim^{sP}_{z^-,z^+} \mathbf{v}. \tag{9}$$

This result is used as a foundation for the statistical and the empirical analysis of the rest of the paper. The current paper uses in fact a natural extension of (9) by focussing on dominance over a *range* of critical levels $\alpha \in [\alpha^-, \alpha^+]$:

$$\mathbf{u} \succsim^{sW}_{\alpha,z^-,z^+} \mathbf{v}, \ \forall \alpha \in [\alpha^-, \alpha^+] \Leftrightarrow \mathbf{u}_\alpha \succsim^{sP}_{z^-,z^+} \mathbf{v}, \ \forall \alpha \in [\alpha^-, \alpha^+]. \tag{10}$$

This provides us with a social ordering that is robust over a class s of functions g and over ranges $[z^-, z^+]$ and $[\alpha^-, \alpha^+]$ of censoring points and critical levels.

3 Statistical inference

This section develops methods to infer statistically the above dominance relations. For the purpose of statistical inference, we assume that the population data have been generated by a data generating process (DGP) from which a finite (but usually large) population is generated. For some (but not for all of the results), we will need to assume that this DGP is continuous, but this is different from saying that the populations must be continuous (or of infinite size) too. For purposes of inference on the populations, we will use data provided by a finite (typically relatively small) sample of observations drawn from the populations. We define F and G as the distribution functions of the DGP that generate the population vectors $\mathbf{u}$ and $\mathbf{v}$ respectively.

3.1 Testing dominance

The equivalence between FGT dominance and CLGU dominance conveniently allows focusing on FGT dominance. As above, let α denote the critical level and α^+ be the maximum possible value that we assume this critical level can take. For any poverty line z, define the FGT index of order s ($s \geq 1$) for the expanded population $\mathbf{u}_\alpha$ as

$$P^s_{F_\alpha}(z) = \int_0^z (z-u)^{s-1} dF_\alpha(u), \tag{11}$$

where $F_\alpha(z) := \frac{M}{N}F(z) + \frac{N-M}{N}I(\alpha \leq z)$ is the distribution of the expanded population $\mathbf{u}_\alpha$ and $F(z)$ is the distribution function of $\mathbf{u}$. The FGT index of the population $\mathbf{v}$ is similarly defined as

$$P^s_G(z) = \int_0^z (z-v)^{s-1} dG(v). \tag{12}$$

The task now is to introduce procedures to test for whether a population CLGU-dominates another one at order s, and this, over intervals of censoring points and critical levels. Two general approaches can be followed for that purpose. The first is based on the following formulation of hypotheses:

$$H^s_0 \ : \ P^s_G(z) - P^s_{F_\alpha}(z) \leq 0 \quad \text{for all } (z,\alpha) \in \left[z^-, z^+\right] \otimes \left[\alpha^-, \alpha^+\right], \tag{13}$$

$$H^s_1 \ : \ P^s_G(z) - P^s_{F_\alpha}(z) > 0 \quad \text{for some } (z,\alpha) \in \left[z^-, z^+\right] \otimes \left[\alpha^-, \alpha^+\right]. \tag{14}$$

This formulation leads to what are generally called "union-intersection" tests. It amounts to define a null of dominance and an alternative of non-dominance. (The null above is that $\mathbf{v}$ dominates $\mathbf{u}$, but that can be reversed.) It has been used and applied in several papers where a Wald statistic or a test statistic based on the supremum of the difference between the FGT indices is generally used to test for dominance — see for example Bishop, Formby, and Thistle (1992) and Barrett and Donald (2003) and Lefranc, Pistolesi, and Trannoy (2006). Davidson and Duclos (2006) discuss why this formulation leads to decisive outcomes only when it rejects the null of dominance and accepts non-dominance. This, however, fails to order the two populations. In those cases in which it is desirable to order the populations, it may be useful to use a second approach and reverse the roles of (13) and (14) by positing the hypotheses as

$$H_0^s \;:\; P_G^s(z) - P_{F_\alpha}^s(z) \geq 0 \quad \text{for some } (z,\alpha) \in \left[z^-, z^+\right] \otimes \left[\alpha^-, \alpha^+\right], \tag{15}$$
$$H_1^s \;:\; P_G^s(z) - P_{F_\alpha}^s(z) < 0 \quad \text{for all } (z,\alpha) \in \left[z^-, z^+\right] \otimes \left[\alpha^-, \alpha^+\right]. \tag{16}$$

This formulation leads to "intersection-union" tests, in which the null is the hypothesis of non-dominance and the alternative is the hypothesis of dominance. This test has been employed by Howes (1993) and Kaur, Prakasa Rao, and Singh (1994). Both papers use a minimum value of the t-statistic. An alternative test is based on empirical likelihood ratio (ELR) statistics, first proposed by Owen (1988) — see also Owen (2001) for a comprehensive account of the EL technique and its properties. Here, we follow the procedure of Davidson and Duclos (2006), which can also be found in Batana (2008), Chen and Duclos (2008) and Davidson (2009). Unlike these papers, we must, however, pay special attention to the value of the critical level and to the sizes of the two populations.

Let m and n be the sizes of the samples drawn from the populations $\mathbf{u}$ and $\mathbf{v}$ respectively and let $\tilde{w}_i^{\mathbf{u}}$ and $\tilde{w}_j^{\mathbf{v}}$ be the sampling weights associated to the observation of individual i in the sample of $\mathbf{u}$ and individual j in the sample of $\mathbf{v}$ respectively. Suppose also that $(u_i, \tilde{w}_i^{\mathbf{u}})$ and $\left(v_j, \tilde{w}_j^{\mathbf{v}}\right)$ are independently and identically distributed (iid) across i and j. For the purposes of asymptotic analysis, define $w_i^{\mathbf{u}}$ and $w_j^{\mathbf{v}}$ such that

$$w_i^{\mathbf{u}} = m\tilde{w}_i^{\mathbf{u}} \text{ and } w_j^{\mathbf{v}} = n\tilde{w}_j^{\mathbf{v}}. \tag{17}$$

These quantities can be used and interpreted as estimates of the population sizes of $\mathbf{u}$ and $\mathbf{v}$ respectively. They remain of the same order as m and n tend to infinity. We can then compute $\hat{P}_{F_\alpha}^s(z)$ and $\hat{P}_G^s(z)$, which are respectively the sample equivalents of $P_{F_\alpha}^s(z)$ and $P_G^s(z)$. They are given by

$$\hat{P}^s_{F_\alpha}(z) = \left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}} (z-u_i)_+^{s-1}\right) \Big/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right) + \left[1 - \left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\right) \Big/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right)\right] (z-\alpha)_+^{s-1} \quad (18)$$

and

$$\hat{P}^s_G(z) = \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}} (z-v_j)_+^{s-1}\right) \Big/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right), \quad (19)$$

where $(z-x)_+^{s-1} \equiv (z-x)^{s-1} I(x \le z)$ for any income value x.

We use the above to compute an ELR statistic. Let $p_i^{\mathbf{u}}$ and $p_j^{\mathbf{v}}$ be the empirical probabilities associated to observations i and j respectively. The ELR statistic is similar to an ordinary LR statistic, and is defined as twice the difference between the unconstrained maximum of an empirical loglikelihood function (ELF) and a constrained ELF maximum. Subject to the null (15) that $\mathbf{u}$ dominates $\mathbf{v}$ at *some* given value of z and α, the constrained ELF maximum $ELF(z,\alpha)$ is given by

$$ELF(z,\alpha) = \max_{p_i^{\mathbf{u}}, p_j^{\mathbf{v}}} \left[\sum_{i=1}^{m} \log p_i^{\mathbf{u}} + \sum_{j=1}^{n} \log p_j^{\mathbf{v}}\right] \quad (20)$$

subject to

$$\sum_{i=1}^{m} p_i^{\mathbf{u}} = 1, \sum_{j=1}^{n} p_j^{\mathbf{v}} = 1 \quad (21)$$

and

$$\sum_{i=1}^{m} p_i^{\mathbf{u}} w_i^{\mathbf{u}} (z-u_i)_+^{s-1} + \left(\sum_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}} - \sum_{i=1}^{m} p_i^{\mathbf{u}} w_i^{\mathbf{u}}\right)(z-\alpha)_+^{s-1} \le \sum_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}} (z-v_j)_+^{s-1}. \quad (22)$$

The unconstrained maximum ELF is defined as (20) subject to (21). Notice that (22) can also be rewritten as

$$\sum_{i=1}^{m} p_i^{\mathbf{u}} w_i^{\mathbf{u}} \left[(z-u_i)_+^{s-1} - (z-\alpha)_+^{s-1}\right] \le \sum_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}} \left[(z-v_j)_+^{s-1} - (z-\alpha)_+^{s-1}\right]. \quad (23)$$

In the spirit of Davidson and Duclos (2006), we compute the ELR statistic for all possible pairs of $(z, \alpha) \in [z^-, z^+] \otimes [\alpha^-, \alpha^+]$, so that we can inspect the value of that statistic when the null hypothesis in (15) is verified at each of these pairs separately. The final ELR test statistic is then given by

$$LR = \min_{(z,\alpha)\in[z^-,z^+]\otimes[\alpha^-,\alpha^+]} LR\left(z, \alpha\right), \tag{24}$$

where

$$LR\left(z, \alpha\right) = 2\left[ELF - ELF\left(z, \alpha\right)\right]. \tag{25}$$

When, in the samples, there is non-dominance of $\mathbf{u}$ on $\mathbf{v}$ at some value of z and α in $[z^-, z^+] \otimes [\alpha^-, \alpha^+]$, the constraint (23) does not matter and the constrained and unconstrained ELF values are the same. The resulting (unconstrained) empirical probabilities are given by

$$p_i^{\mathbf{u}} = \frac{1}{m} \text{ and } p_j^{\mathbf{v}} = \frac{1}{n}. \tag{26}$$

In the case where there is dominance in the samples, the constraint (23) binds and the probabilities obtained from the resolution of the problem are:

$$p_i^{\mathbf{u}} = \frac{1}{m - \rho\left[\nu - w_i^{\mathbf{u}}\left((z-u_i)_+^{s-1} - (z-\alpha)_+^{s-1}\right)\right]} \tag{27}$$

and

$$p_j^{\mathbf{v}} = \frac{1}{n + \rho\left[\nu - w_j^{\mathbf{v}}\left((z-v_j)_+^{s-1} - (z-\alpha)_+^{s-1}\right)\right]}. \tag{28}$$

The constants ρ and ν are the solutions to the following equations,

$$\begin{cases} \sum\limits_{i=1}^{m} p_i^{\mathbf{u}} w_i^{\mathbf{u}}\left[(z-u_i)_+^{s-1} - (z-\alpha)_+^{s-1}\right] = \sum\limits_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}}\left[(z-v_j)_+^{s-1} - (z-\alpha)_+^{s-1}\right] \\ \sum\limits_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}}\left[(z-v_j)_+^{s-1} - (z-\alpha)_+^{s-1}\right] = \nu, \end{cases} \tag{29}$$

with $p_i^{\mathbf{u}}$ and $p_j^{\mathbf{v}}$ given in (27) and (28). The solutions cannot be found analytically, so a numerical method must be used.

An alternative, though analogous, statistic is the t-statistic of Kaur, Prakasa Rao, and Singh (1994), which is the minimum of $t\left(z, \alpha\right)$ over $[z^-, z^+] \otimes [\alpha^-, \alpha^+]$, where

$$t\left(z, \alpha\right) = \frac{\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)}{\left[\widehat{\text{var}}\left(\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)\right)\right]^{1/2}}, \tag{30}$$

and $\widehat{\text{var}}\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)$ is the estimate of the asymptotic variance of $\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)$ for some pair (z, α). Denote that minimized t-statistic by t.

Testing the null of dominance makes sense only when there is dominance in the original samples.We can then proceed with asymptotic tests and/or bootstrap tests with either LR or t statistics, although for bootstrap tests we must first obtain the empirical probabilities of the ELF approach. Let LR_a and t_a denote the statistics in the case of asymptotic tests and let LR_b and t_b be the statistics for the bootstrap tests. For asymptotic tests and for a test of level β, the decision rule is to reject the null of non-dominance in favor of the alternative of dominance if t_a exceeds the critical value associated to β of the standard normal distribution. Note that LR and the square of t are asymptotically equivalent — see Section 8.3 in the Appendix for more details. We can therefore also use a decision rule of rejecting the null of non-dominance in favor of the alternative of dominance if LR_a exceeds the critical value associated to β of the chi-square distribution.

The bootstrap testing procedure is formally set up as follows:

Step 1: For two initial samples drawn from two populations, compute $LR\,(z, \alpha)$ and $t\,(z, \alpha)$ for every pair (z, α) in $[z^-, z^+] \otimes [\alpha^-, \alpha^+]$ as described above. If there exists at least one (z, α) for which $\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z) \geq 0$, then H^s_0 cannot be rejected; choose then a value equal to 1 for the p-value and stop the process. If not, continue to the next step.

Step 2: Search for the minima statistics, that is to say, find LR as the minimum of $LR\,(z, \alpha)$ and t as the minimum of $t\,(z, \alpha)$ over all pairs (z, α). Suppose that LR is obtained at $(\tilde{z}, \tilde{\alpha})$ and denote $\tilde{p}^{\mathbf{u}}_i$ and $\tilde{p}^{\mathbf{v}}_j$ the resulting probabilities given by (27) and (28) and evaluated at $(\tilde{z}, \tilde{\alpha})$.

Step 3: Use $\tilde{p}^{\mathbf{u}}_i$ and $\tilde{p}^{\mathbf{v}}_j$ to generate bootstrap samples of size m for $\mathbf{u}$ and of size n for $\mathbf{v}$ by resampling the original data with these probabilities. The bootstrap samples are thus drawn with unequal probabilities $\tilde{p}^{\mathbf{u}}_i$ and $\tilde{p}^{\mathbf{v}}_j$. It can result that, in some of the bootstrap samples, the estimated size of population $\mathbf{u}$ becomes larger than that of population $\mathbf{v}$. In such cases, the roles of F_α and G are subsequently reversed, that is, we consider F and G_α.[1]

Step 4: As is usual, consider 399 bootstrap replications, $b =$ 1,...,399. For each replication, use the bootstrap data and follow previous step 3. Compute the two statistics LR_b and t_b for every $b \leq$ 399 as in the original data.

Step 5: Compute the p-value of the bootstrap statistics as the proportion of LR_b that are greater than LR — the ELR statistic obtained with the original data — or as the proportion of t_b that are greater than t — the t-statistic obtained with the original data.

Step 6: Reject the null of non-dominance if the bootstrap p-value is lower than some specified nominal levels.

[1]This will not occur in samples where all observations have the same weights.

3.2 Estimating robust ranges of critical levels

To get around the problem of the absence of empirical/ethical consensus on an appropriate range of values for the critical level, we can search for evidence on the ranges of critical levels that *can* order distributions (see Blackorby, Bossert, and Donaldson 1996 and Trannoy and Weymark 2009 for a discussion). For this, consider again two populations $\mathbf{u}$ and $\mathbf{v}$ of sizes M and N respectively. Suppose that we have two samples drawn from $\mathbf{u}$ and $\mathbf{v}$ and assume for simplicity that they are independent and that their moments of order $2\,(s-1)$ are finite. Denote m and n the sizes of the two samples. For some fixed z^- and z^+, define α_s and α^s respectively as follows:

$$\alpha_s = \max\{\alpha | P^s_{F_\alpha}(z) \geq P^s_G(z) \text{ for all } z^- \leq z \leq z^+\} \tag{31}$$

and

$$\alpha^s = \min\{\alpha | P^s_{F_\alpha}(z) \leq P^s_G(z) \text{ for all } z^- \leq z \leq z^+\}. \tag{32}$$

In the light of how they are defined, we can refer to α_s as an "upper bound" of the critical level and α^s as a "lower bound" of the critical level. In order to have FGT dominance made robustly over ranges of censoring points, we can also define critical values for the maximum censoring point as:

$$z^+_s = \max\{z^+ | P^s_{F_\alpha}(z) \geq P^s_G(z) \text{ for all } z^- \leq z \leq z^+\} \tag{33}$$

and

$$z^{s+} = \max\{z^+ | P^s_{F_\alpha}(z) \leq P^s_G(z) \text{ for all } z^- \leq z \leq z^+\}, \tag{34}$$

where α is some fixed value of critical level. z^+_s is the maximum censoring point for which $\mathbf{v}$ dominates $\mathbf{u}$ and z^{s+} is the maximum censoring point for which $\mathbf{u}$ dominates $\mathbf{v}$.

Given the definitions (31) and (32) and assuming that α_s and α^s exist, it is useful to define the following assumptions. For α_s, suppose that

$$\begin{cases} \frac{M}{N} P^s_F(z) \geq P^s_G(z) & \text{for all } z \leq \alpha_s \\ \frac{M}{N} P^s_F(z) < P^s_G(z) & \text{for some } z \geq \alpha_s + \epsilon \text{ and } z^- \leq z \leq z^+, \end{cases} \tag{$\mathbf{VDU_s}$}$$

where ϵ is some arbitrarily small positive value. For α^s, consider first the case of s =1 and suppose that

$$\begin{cases} \frac{M}{N} P^1_F(z) + \frac{(N-M)}{N} I(\alpha^1 \leq z) \leq P^1_G(z) & \text{for all } z^- \leq z \leq z^+ \\ \frac{M}{N} P^1_F(z) + \frac{(N-M)}{N} > P^1_G(z) & \text{for some } z \leq \alpha^1 - \epsilon, \end{cases} \tag{$\mathbf{UDV_1}$}$$

where ϵ is again some arbitrarily small positive value. When $s \geq 2$, we modify the above assumptions slightly and define α^s as:

$$\begin{cases} \frac{M}{N} P_F^s(z) + \frac{(N-M)}{N} \left(z - \alpha^s\right)_+^{s-1} \leq P_G^s(z) & \text{for all } z^- \leq z \leq z^+ \\ \frac{M}{N} P_F^s(z^s) + \frac{N-M}{N} (z^s - \alpha^s)_+^{s-1} = P_G^s(z^s) & \text{for } \alpha^s < z^s \leq z^+ \end{cases} \qquad (\mathbf{UDV_s})$$

with $(z-x)_+^{s-1} = \max\left[(z-x)^{s-1}, 0\right]$. Suppose that z^s exists and is the crossing point between the FGT curves $P_{F_\alpha}^s$ and P_G^s. In most cases, we would expect z^s to coincide with z^+ — see Section 8.1 in the Appendix for more details. Assumptions $\mathbf{VDU}_s$ and $\mathbf{UDV}_s$ are useful for the estimation of α_s and α^s. In order to better understand their role, consider the case of $s = 1$. Figures 1 and 2 graph cumulative distributions functions adjusted for differences in population sizes.[2] It is supposed that the larger population $\mathbf{v}$ dominates the smaller population $\mathbf{u}$ for a range $[0, z^+]$ of censoring points $\alpha \in [0, \alpha_1]$. This is expressed by the fact that the cumulative distribution G of $\mathbf{v}$ is under the cumulative distribution F of $\mathbf{u}$ adjusted by the ratio $\frac{M}{N}$ up to $\alpha_1 > 0$. Figure 1 shows the case where the critical level is equal to 0. In this case, the larger population clearly dominates the smaller one. At the critical level value α_1, the two functions cross; $\mathbf{v}$ just dominates $\mathbf{u}$ when the critical level is equal to α_1. However, $\mathbf{v}$ does not dominate $\mathbf{u}$ when the critical level takes a value $\alpha_0 > \alpha_1$.

In Figure 3, $\mathbf{u}$ is assumed to dominate $\mathbf{v}$. The dominance of $\mathbf{u}$ over $\mathbf{v}$ is preserved when the critical level has a value at least equal to α^1. But this is not true for any critical level α_0 lower than α^1.[3]

Note that α_1 and α^1 are the crossing points of FGT curves. This suggests the application of the procedure of Davidson and Duclos (2000) for estimation and inference of the population values of α_1 and α^1. Consider the populations $\mathbf{u}$ and $\mathbf{v}$ with sample sizes equal to m and n respectively. Using assumption $\mathbf{UDV}_1$ (also see Figure 3) and assuming continuity of the DGP at α^1, we have that

$$\frac{M}{N} P_F^1(\alpha^1) + \frac{(N-M)}{N} - P_G^1(\alpha^1) = 0. \tag{35}$$

Denoting $\psi(z) = \frac{M}{N} P_F^1(z) + \frac{(N-M)}{N} I(\alpha^1 \leq z) - P_G^1(z)$, then $\psi(z) \leq 0$ for all $z^- \leq z \leq z^+$ and $\psi(\alpha^1) = 0$. Recall that

$$\hat{P}_F^1(z) = \frac{\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}} I(u_i \leq z)}{\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}}, \quad \hat{P}_G^1(z) = \frac{\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}} I(v_j \leq z)}{\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}}, \tag{36}$$

where $w_i^{\mathbf{u}}$ and $w_j^{\mathbf{v}}$ are given in the previous section. A natural estimator of α^1 would be $\widehat{\alpha}^1$ such that

$$\frac{\hat{M}}{\hat{N}} \hat{P}_F^1(\hat{\alpha}^1) + \frac{\left(\hat{N} - \hat{M}\right)}{\hat{N}} - \hat{P}_G^1(\hat{\alpha}^1) = 0, \tag{37}$$

[2]See Section 8.1 in the Appendix for the case of $s > 1$.

[3]The Appendix illustrates graphically two cases of dominance of $\mathbf{u}$ over $\mathbf{v}$ when $s > 1$.

where $\hat{M} = \frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}$ and $\hat{N} = \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}$ are respectively the estimators of the population sizes of $\mathbf{u}$ and $\mathbf{v}$.

For $s \geq 2$, denote $\phi(\alpha^s) = \frac{M}{N} P_F^s(z^s) + \frac{N-M}{N}(z^s - \alpha^s)_+^{s-1} - P_G^s(z^s)$. Recall that z^s is defined on page 13 and $z^s > \alpha^s$. Then $\phi'(\alpha^s) = -(s-1)\frac{N-M}{N}(z^s - \alpha^s)^{s-2} \neq 0$. A consistent estimator of α^s, $\hat{\alpha}^s$, can be obtained from

$$\frac{\hat{M}}{\hat{N}} \hat{P}_F^s(z^s) + \frac{\hat{N} - \hat{M}}{\hat{N}}(z^s - \hat{\alpha}^s)^{s-1} - \hat{P}_G^s(z^s) = 0, \tag{38}$$

where

$$\hat{P}_F^s(z^s) = \frac{\frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}} (z^s - u_i)_+^{s-1}}{\frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}} \quad \text{and} \quad \hat{P}_G^s(z^s) = \frac{\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}} (z^s - v_j)_+^{s-1}}{\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}}. \tag{39}$$

For $s \geq 2$, $\hat{\alpha}^s$ is given analytically by

$$\hat{\alpha}^s = z^s - \left[\frac{\hat{N}\hat{P}_G^s(z^s) - \hat{M}\hat{P}_F^s(z^s)}{\hat{N} - \hat{M}}\right]^{\frac{1}{s-1}}. \tag{40}$$

To derive the asymptotic distribution of $\hat{\alpha}^s$ for $s \geq 1$, assume that F and G are differentiable and denote $P_F^0(z) = F'(z)$ and $P_G^0(z) = G'(z)$. Also suppose that $(w_i^{\mathbf{u}})_{i=1}^m \sim iid\left(\mu_{w^{\mathbf{u}}}, \sigma_{w^{\mathbf{u}}}^2\right)$ and $\left(w_j^{\mathbf{v}}\right)_{j=1}^n \sim iid\left(\mu_{w^{\mathbf{v}}}, \sigma_{w^{\mathbf{v}}}^2\right)$. Assuming that $r = \frac{m}{n}$ remains constant as m and n tend to infinity, let

$$\Lambda^1 = \begin{pmatrix} \frac{1}{\Gamma_4{}^2} \operatorname{var}\left[w^{\mathbf{v}}(\alpha^1 - v)_+^0\right] + \\ \frac{r^{-1}}{\Gamma_4{}^2}\left[\operatorname{var} w^{\mathbf{u}}(\alpha^1 - u)_+^0 + \sigma_{w^{\mathbf{u}}}^2\right] + \\ \left[\frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^2} - \frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^2} + \frac{\Gamma_2}{\Gamma_4{}^2}\right]^2 \sigma_{w^{\mathbf{v}}}^2 - \\ \frac{2r^{-1}}{\Gamma_4{}^2}\left[E\left((w^{\mathbf{u}})^2(\alpha^1 - u)_+^0\right) - \Gamma_1(\alpha^1)\Gamma_2\right] + \\ 2\left[\frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^3} - \frac{\Gamma_2}{\Gamma_4^3} - \frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^3}\right] \times \\ \left[E\left((w^{\mathbf{v}})^2(\alpha^1 - v)_+^0\right) - \Gamma_3(\alpha^1)\Gamma_4\right] \end{pmatrix} \tag{41}$$

where $\Gamma_1(\alpha^1) = E\left[m^{-1}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}(\alpha^1 - u_i)_+^0\right]$, $\Gamma_2 = \mu_{w^{\mathbf{u}}}$, $\Gamma_3(\alpha^1) = E\left[n^{-1}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}(\alpha^1 - v_j)_+^0\right]$ and $\Gamma_4 = \mu_{w^{\mathbf{v}}}$ and, for $s \geq 2$,

$$\Lambda^s = \begin{pmatrix} \frac{1}{\Gamma_4{}^2}\operatorname{var}\left[w^{\mathbf{v}}\left(z^s - v\right)_+^{s-1}\right] + \\ \frac{r^{-1}}{\Gamma_4{}^2}\left[\operatorname{var} w^{\mathbf{u}}\left(z^s - u\right)_+^{s-1} + \left(z^s - \alpha^s\right)^{2s-2}\sigma_{w^{\mathbf{u}}}^2\right] + \\ \left[\frac{\Gamma_3}{\Gamma_4{}^2} - \frac{\Gamma_1}{\Gamma_4{}^2} + \frac{\Gamma_2}{\Gamma_4{}^2}\left(z^s - \alpha^s\right)^{s-1}\right]^2\sigma_{w^{\mathbf{v}}}^2 - \\ \frac{2r^{-1}}{\Gamma_4{}^2}\left(z^s - \alpha^s\right)^{s-1}\left[E\left(\left(w^{\mathbf{u}}\right)^2\left(z^s - u\right)_+^{s-1}\right) - \Gamma_1\Gamma_2\right] + \\ 2\left[\frac{\Gamma_1}{\Gamma_4{}^3} - \frac{\Gamma_2}{\Gamma_4{}^3}\left(z^s - \alpha^s\right)^{s-1} - \frac{\Gamma_3}{\Gamma_4{}^3}\right] \times \\ \left[E\left(\left(w^{\mathbf{v}}\right)^2\left(z^s - v\right)_+^{s-1}\right) - \Gamma_3\Gamma_4\right] \end{pmatrix} \tag{42}$$

where $\Gamma_1 = E\left[m^{-1}\sum_{i=1}^{m} w_i^{\mathbf{u}}\left(z^s - u_i\right)_+^{s-1}\right]$, $\Gamma_2 = \mu_{w^{\mathbf{u}}}$, $\Gamma_3 = E\left[n^{-1}\sum_{j=1}^{n} w_j^{\mathbf{v}}\left(z^s - v_j\right)_+^{s-1}\right]$, and $\Gamma_4 = \mu_{w^{\mathbf{v}}}$.

We can now state the following theorem.

Theorem 1
For $s = 1$, assume that there exists α^1 such that the conditions $\mathbf{UDV}_1$ on page 13 are satisfied and that $\frac{M}{N}P_F^0(\alpha^1) - P_G^0(\alpha^1) \neq 0$. Then, $\sqrt{n}(\hat{\alpha}^1 - \alpha^1) \xrightarrow{d} N(0, V^1)$, with

$$V^1 = \lim_{m,\, n \to \infty} \operatorname{var}\left(\sqrt{n}(\hat{\alpha}^1 - \alpha^1)\right) = \frac{\Lambda^1}{\left(\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}P_F^0(\alpha^1) - P_G^0(\alpha^1)\right)^2}$$

and Λ^1 given in (41).
For $s \geq 2$, suppose that there exists α^s such that conditions $\mathbf{UDV}_s$ on page 13 are satisfied and that $z^s > \alpha^s$. Then, $\sqrt{n}(\hat{\alpha}^s - \alpha^s) \xrightarrow{d} N(0, V^s)$, where

$$V^s = \lim_{m,\, n \to \infty} \operatorname{var}\left(\sqrt{n}(\hat{\alpha}^s - \alpha^s)\right) = \frac{\Lambda^s}{\left[(s-1)\left(1 - \frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}\right)\left(z^s - \alpha^s\right)^{s-2}\right]^2}$$

and Λ^s given in (42).

Proof: See Appendix.

Let us consider the critical value α_s and suppose that conditions $\mathbf{VDU}_s$ are satisfied. Assuming continuity of the DGP at α_s, we obtain that

$$\frac{M}{N}P_F^s(\alpha_s) - P_G^s(\alpha_s) = 0. \tag{43}$$

A consistent estimator of α_s is obtained from

$$\frac{\hat{M}}{\hat{N}}\hat{P}_F^s\left(\hat{\alpha}_s\right) - \hat{P}_G^s(\hat{\alpha}_s) = 0. \tag{44}$$

Using the same previous conditions when dealing with the asymptotic distribution of $\hat{\alpha}^s$, denote

$$\Lambda_s = \begin{pmatrix} \frac{r^{-1}}{\Gamma_4{}^2}\left[\mathrm{var}\, w^{\mathbf{u}}\left(\alpha_s - u\right)_+^{s-1}\right] + \\ \frac{1}{\Gamma_4{}^2}\,\mathrm{var}\left[w^{\mathbf{v}}\left(\alpha_s - v\right)_+^{s-1}\right] + \\ \left[\frac{\Gamma_3(\alpha_s)}{\Gamma_4{}^2} - \frac{\Gamma_1(\alpha_s)}{\Gamma_4{}^2}\right]^2 \sigma^2_{w^{\mathbf{v}}} + \\ 2\left[\frac{\Gamma_1(\alpha_s)}{\Gamma_4{}^3} - \frac{\Gamma_3(\alpha_s)}{\Gamma_4{}^3}\right] \times \\ \left[E\left(\left(w^{\mathbf{v}}\right)^2\left(\alpha_s - v\right)_+^{s-1}\right) - \Gamma_3(\alpha_s)\Gamma_4\right] \end{pmatrix} \tag{45}$$

where $\Gamma_1(\alpha_s) = E\left[m^{-1}\sum_{i=1}^{m} w_i^{\mathbf{u}}\left(\alpha_s - u_i\right)_+^{s-1}\right]$, $\Gamma_2 = \mu_{w^{\mathbf{u}}}$, $\Gamma_3(\alpha_s) = E\left[n^{-1}\sum_{j=1}^{n} w_j^{\mathbf{v}}\left(\alpha_s - v_j\right)_+^{s-1}\right]$, and $\Gamma_4 = \mu_{w^{\mathbf{v}}}$. The following theorem gives the asymptotic distribution of $\hat{\alpha}_s$.

Theorem 2

Suppose that conditions $\mathbf{VDU}_s$ on page 13 are satisfied and that for $s \geq 1$ there exists α_s such that $\frac{M}{N}P_F^s(\alpha_s) = P_G^s(\alpha_s)$ and $\frac{M}{N}P_F^s(z) > P_G^s(z)$ for all $z < \alpha_s$. Denote $\varphi(z) = \frac{M}{N}P_F^s(z) - P_G^s(z)$ and note that $\varphi(z) > 0$ for all $z < \alpha_s$ and $\varphi(\alpha_s) = 0$. Then, $\varphi'(\alpha_s) = (s-1)\left(\frac{M}{N}P_F^{s-1}(\alpha_s) - P_G^{s-1}(\alpha_s)\right) \neq 0$. We have that $\sqrt{n}(\hat{\alpha}_s - \alpha_s) \xrightarrow{d} N(0, V_s)$ where for $s = 1$,

$$V_1 = \lim_{m,\, n \to \infty} \mathrm{var}\left(\sqrt{n}(\hat{\alpha}_1 - \alpha_1)\right) = \frac{\Lambda_1}{\left(\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}P_F^0(\alpha_1) - P_G^0(\alpha_1)\right)^2}$$

and for $s \geq 2$,

$$V_s = \lim_{m,\, n \to \infty} \mathrm{var}\left(\sqrt{n}(\hat{\alpha}_s - \alpha_s)\right) = \frac{\Lambda_s}{(s-1)^2\left(\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}P_F^{s-1}(\alpha_s) - P_G^{s-1}(\alpha_s)\right)^2}$$

with Λ_s given in (45).

Proof: See Appendix.

4 Simulations of the effect of population size on social evaluation

We now briefly illustrate the impact of population sizes on welfare rankings using the CLGU dominance approach. To do this, we consider two populations of different sizes. The smaller population is of size M and has a distribution F and the larger one is of size N and has a distribution G. We define those distributions over the $[0, 1]$ interval.

Let population $\mathbf{v}$ have a uniform distribution on $[0, 1]$ and population $\mathbf{u}$ be piecewise-linear distributed, that is to say, be uniform over 20 equal segments belonging to the $[0, 1]$

interval. The upper limits of these segments are 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0 .95, and 1.00. Because $\mathbf{v}$ has a uniform distribution, these upper limits also correspond to the cumulative probabilities for $\mathbf{v}$ at these points. For the first case that we consider, the cumulative probabilities for $\mathbf{u}$ at the upper limit of each segment are respectively 0.15, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.82, 0.85, 0.87, 0.90, 0.95, 0.97 and 1.00. We suppose that $\frac{M}{N} = 2/3$.

$\mathbf{v}$ dominates $\mathbf{u}$ for low values of α. Figures 4 and 5, also show that α_1 =0.3 and $\alpha_2 =$ 0.6. The larger population $\mathbf{v}$ thus dominates the smaller population $\mathbf{u}$ at first order for any critical level at most equal to 0.3. Second-order dominance is obtained with any $\alpha \leq$ 0.6.

The second case we consider lets the smaller population $\mathbf{u}$ dominate the larger population $\mathbf{v}$. For this, the cumulative probabilities for $\mathbf{u}$ are set to 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.45, 0.55, 0.65, 0.70, 0.75, 0.80 and 1.00. We can then find the critical levels α^s. Figures 6 and 7 show that $\alpha^1 =$ 0.4 and $\alpha^2 =$ 0.2. Hence, the smaller population $\mathbf{u}$ dominates the larger one, at first order, for any critical level $\alpha \geq$ 0.4, and at second-order for any $\alpha \geq$ 0.2.

Table 1 and Table 2 show how the lower and upper bounds for the ranges of normatively robust critical levels vary with the order of dominance s. α_s (the upper bound) is increasing with s and α^s (the lower bound) is decreasing with s. In both cases, this says the ranges of normatively robust critical levels increase with the order of dominance.

Tables 1 and 2 also show how those bounds are affected by population size. As the ratio of the population sizes approaches 1 (the two distributions are left unchanged), the value of α_s increases whereas the value of α^s decreases. Conversely, if the ratio of the sizes is sufficiently small, α_s becomes small and that of α^s becomes large. The intuition is that the larger the difference in population sizes, the greater the importance of the critical level in ranking the distributions. *Ceteris paribus*, therefore, the larger the difference in population sizes, the more restricted are the ranges of critical levels over which it is possible to rank distributions.

5 Illustration using Canadian data

We now illustrate the use of the normative and statistical framework developed earlier. The data are drawn from the Canadian Surveys of Consumer Finances (SCF) for 1976 and 1986 and the Canadian Surveys of Labour and Income Dynamics (SLID) for 1996 and 2006. Empirical studies on poverty and welfare in Canada have mostly used these same data: see *inter alia* Chen and Duclos (2008), Chen (2008) and Bibi and Duclos (2009). We use equivalized net income as a measure of individual well-being. We rely for that purpose on the equivalence scale often employed by Statistics Canada. This equivalence scale applies a factor of 1 for the oldest person in the family, 0.4 for all other members aged at least 16 and 0.3 for the remaining members under age 16. In order to take into account the differences in spatial prices, we adjust incomes by the ratio of spatial "market basket measures" (see Human Resources and Social Development Canada 2006). We also use Statistics Canada's consumer price indices to convert dollars into 2002 constant dollars.

The sample sizes from 1976, 1986, 1996 and from 2006 are respectively 28,613, 36,389, 31,973 and 28,524. The use of the sampling weights leads to estimates of Canada's population size in 1976 of 22,230,000, of 25,384,000 for 1986, of 28,870,000 for 1996, and of 31,853,000 for 2006. We assign the value of 0 to all negative incomes — this concerns 1.9% of the observations for 1976 and less than 0.5% for the other years. The cumulative distribution for all four years is shown in Figure 8.

We now turn to testing dominance. The FGT dominance tests set the upper bound of the censoring point z^+ to \$70,500, with the implicit assumption that the range [\$9,500, \$70,500] will cover any censoring point that one would want to apply. The value of $z^- =$ \$9,500 is the minimum equivalent income that allows inferring dominance for most of the comparisons we will consider below. No more than 7.1% of the observations in any of the four distributions have equivalent incomes in excess of $z^+ =$ \$70,500. Setting such a relatively high value for z^+ is also useful to be able to interpret the FGT dominance rankings (almost) as welfare ones.

Table 3 presents the results of the dominance tests based on the range of censoring points $[z^-, z^+] =$ [\$9,500, \$70,500] and the range of critical levels $[\alpha^-, \alpha^+] =$ [\$5,000, \$15,000]. The lower limit α^- of the critical levels is set arbitrarily to \$5,000; the upper limit α^+ is close to Statistics Canada's Low-Income Cutoff, a popular poverty threshold in Canada.

In Table 3, we test the null hypothesis that the larger population does not dominate the smaller one. For expositional brevity, we focus on the first-order results. At a 5% significance level, recent years dominate earlier years for both asymptotic and bootstrap tests, except when comparing 1986 and 1996. The relatively large lower bound of $z^- =$ \$9,500 is needed to infer the dominance of 2006 over 1986 and over 1996; for the other dominance relations, however, z^- can be set lower, such as \$3,500 for the dominance of 1986 over 1976 and \$4,500 for the dominance of 1996 over 1976. Notice that all of the dominance relations of larger over smaller years remains unchanged when the lower bound α^- of the critical level becomes arbitrary close to 0 — see Duclos and Zabsonré (2009).

We now turn to the estimation of the upper bounds α_s of the ranges of those critical levels over which welfare dominance rankings can be made. For this procedure to be valid for dominance of a large over a smaller population, we need to have verified the hypothesis $\mathbf{VDU}_s$ for given s. Given the inference results of Table 3, we therefore focus on five dominance relationships: 1976 versus 1986, 1976 versus 1996, 1976 versus 2006, 1986 versus 1996 and 1996 versus 2006.

Table 4 shows the estimates $\hat{\alpha}_s$ for the dominance of 1986 and 1996 over 1976. Analogous estimates are given in Table 5 for the dominance of 2006 over 1976 and 1996 respectively. Table 4 shows for instance that 1986 dominates 1976 for all critical levels up to an upper bound of \$30,550, with a standard error of \$1,639. As can be seen, the estimates of α_s indicate that the dominance of 2006 over 1996 is stronger than the dominance of 2006 over 1976 and the dominance of 1986 over 1976. For instance, the use of any critical level lower than

$49,592 leads to the dominance of 2006 over 1996 at first-order. However, the dominance of 1996 over 1976 is obtained only for critical levels at most equal to $17,453 (with a standard error of $1,129). This also indicates that for values of α_s greater than $17,453 and for some of the CLGU welfare indices that are members of the first-order class $\mathcal{F}^1_{z^-,z^+}$ (see (5)), 1996 would not show more welfare than 1976.

We can also estimate the lower bounds α^s of critical levels over which smaller populations dominate larger ones. This is possible to do with our Canadian data only for the dominance of 1986 over 1996 and when $s \geq 2$. The case of $s = 1$ is indeed too demanding since $\hat{\alpha}^1$ does not exist; there are therefore first-order indices that would rank 1996 better and this, for any choice of critical level value. Considering α^s for the dominance of other smaller populations over larger ones is not possible because the $\mathbf{UDV}_s$ assumption posited in Section 3 is not satisfied for such relations. This is partly due to the fact that there are more individuals with equivalent income equal to 0 in the samples of earlier years than in the samples of more recent years. Consequently, the estimates of the absolute number of lower-income people in 1976 and 1996 exceeds those of 2006 and it becomes difficult to obtain dominance of 1976 and 1996 over 2006; the same applies for 1986 over 2006.

Figure 9 shows a plot of the estimated absolute number of people below z ("number of poor") in 1976 and 1996. As can be seen, if the censoring point z is no more than the critical level $\hat{\alpha}_1$, then there are more poor in 1976 than in 1996. For z equal to $\hat{\alpha}_1$, the number of poor is estimated to be the same at 8.38 millions for the two years.

Table 6 shows the estimates of α^s (for $s \geq 2$) for dominance of 1986 over 1996. The critical level α^1 cannot be found, given the initial non-dominance of 1986 over 1996 at first-order. The estimate of α^2 is $23,878, with a standard error of around $1,100. From the results of Table 6 we can therefore infer that social welfare in Canada has decreased robustly between 1986 and 1996 if lives need to enjoy a level of well-being of at least $25,100 ($23,878 plus two standard errors) to contribute positively to social welfare, as measured by second-order welfare indices. With these critical levels, Canada's smaller population in 1986 exhibits greater social welfare than Canada's larger population in 1996 for all of the social welfare indices that belong to W^2_{α,z^-,z^+}. If we restrict attention to the class of third-order indices, W^3_{α,z^-,z^+}, then Table 6 says that 1986 has greater social welfare than 1996 if the critical levels are higher than $21,592 ($19,592 plus twice the standard error of $1,000). For $s = 4$, the corresponding figure is around $19,539.

We can also bound the ranges of censoring points over which there is robust dominance of one year over another. For all critical level values no less than $31,000 and for all second-order welfare indices, Canada in 1986 is better than in 1996 for all censoring points up to $53,096. This upper bound of the censoring points increases as the order of dominance s increases; it reaches a value of $218,360 for $s = 4$. (The influence of s on α_s, α^s, and z^+_s is established in Duclos and Zabsonré (2009).) The link between the critical level α and the upper bound of the censoring points z^+ is also considered in Figure 10 for first-order dominance of 1986 over 1976. As the value of z^+ increases, the critical level α_1 weakly decreases — an analogous relationship holds true for higher orders of dominance and for dominance of smaller over larger populations. Thus, the greater the ranges of possible censoring points we wish to allow for, the lower the ranges of critical levels over which we can find dominance. (This result is also theoretically discussed in Duclos and Zabsonré (2009).)

Note that given the definition of $\mathbf{VDU}_s$ on page 13, any value of the critical level greater than z^+ does not affect the relation of dominance of a larger over a smaller population. That is to say, if $\alpha_s = z^+$, the larger population still dominates the smaller one even if α is arbitrarily larger than z^+ — setting α_s is then harmless. Take for instance the case of the first-order dominance of 1986 over 1976, for which $\hat{\alpha}_1 =$ \$30,550. For $z^+ <$ \$30,550, $\hat{\alpha}_1$ can thus be set to as high a level as needed; for $z^+ \geq$ \$30,550, we have $\hat{\alpha}_1 =$ \$30,550.

6 Conclusion

This paper develops and applies methods for assessing society's welfare in contexts in which both population sizes and the distributions of individual welfare can differ. This issue has important implications for monitoring human development and for thinking about public policy. The paper makes three main contributions to the literature. First, it is one of the first to use the critical-level generalized utilitarianism (CLGU) framework of Blackorby and Donaldson (1984), a framework that avoids some of the fundamental weaknesses of the more traditional total and average utilitarian frameworks. Second, it introduces and uses relationships that can order distributions over classes of CLGU social welfare functions, in the tradition of the stochastic dominance approach. Third, it is the first paper to analyze combined population-sizes and population-distributions rankings in a coherent statistical and inferential framework. This is done *inter alia* by developing tools for testing for CLGU dominance and for estimating the bounds of critical levels and welfare censoring points over which robust CLGU rankings can be made.

The paper is also the first to apply the CLGU framework to real data. This is done using Canadian Surveys of Consumer Finances (SCF) for 1976 and 1986, and Canadian Surveys of Labour and Income Dynamics (SLID) for 1996 and 2006. Asymptotic and bootstrap procedures are used to test for dominance relationships across these years, relationships that involve testing over classes of social welfare functions, ranges of censoring points as well as ranges of critical level values. It is found that recent years generally dominate earlier ones, suggesting that there has been a social welfare improvement in Canada in spite of the fact that population size has increased substantially and that new lives do not always increase society's welfare in a CLGU framework.

More surprisingly perhaps, Canada's smaller population in 1986 is socially better than Canada's larger population in 1996 in a CLGU framework for a relatively wide range of critical levels. Yet, comparisons of total and average income indicate the contrary. Total income in Canada indeed amounts to \$654 billion and \$789 billion respectively for 1986 and 1996; Canada's average income is respectively \$25,789 and \$27,334 for 1986 and 1996.

Hence, not only can the evaluating frameworks of average and total utilitarianism diverge in theory and in practice, but they can also give opposite social evaluation rankings to those of critical-level utilitarianism, an alternative social evaluation framework that has been shown to resolve nicely some of the ethical lacuna of average and total utilitarianism. This is an important lesson for anyone interested in the evaluation of policy and human development in the presence of demographic changes.

7 Figures and tables

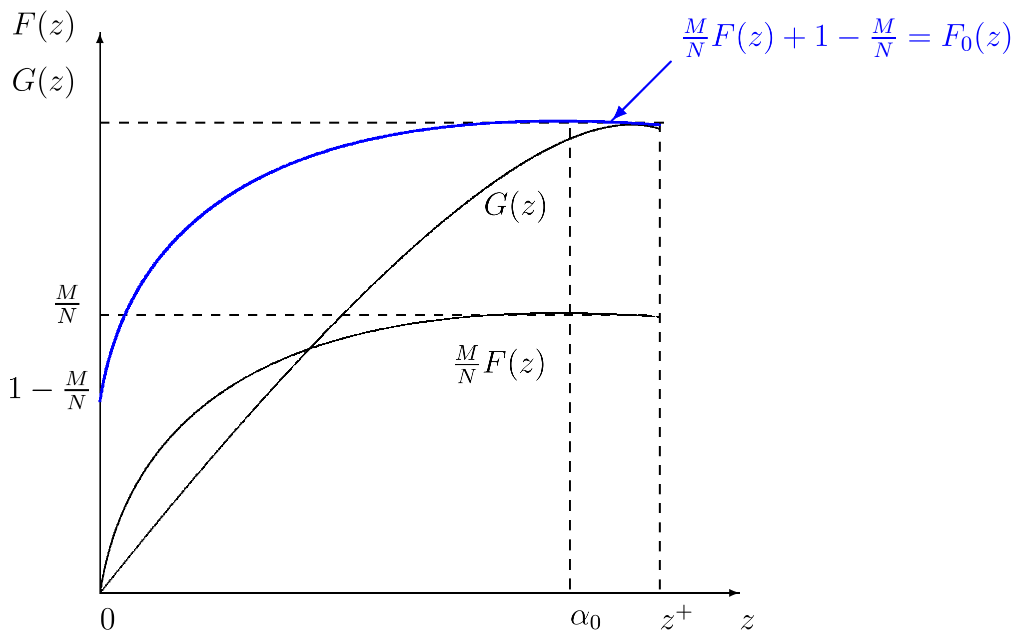

Figure 1: Poverty incidence curves with $\alpha = 0$ adjusted for differences in population sizes

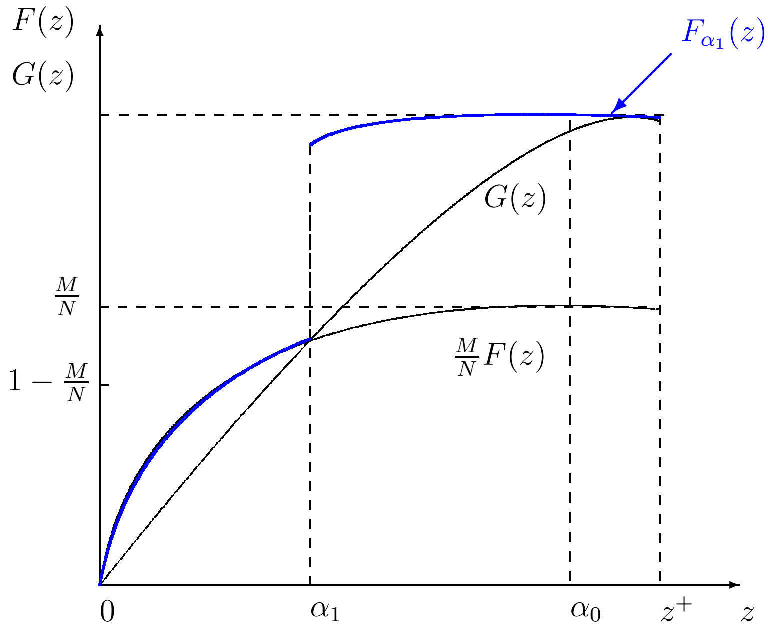

Figure 2: Poverty incidence curves with $\alpha = \alpha_1$ adjusted for differences in population sizes

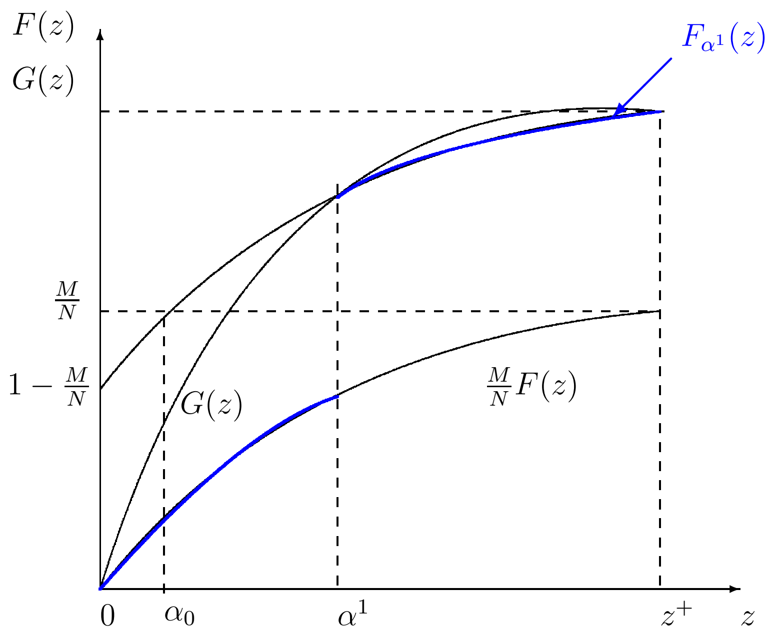

Figure 3: Poverty incidence curves with $\alpha = \alpha^1$ adjusted for differences in population sizes

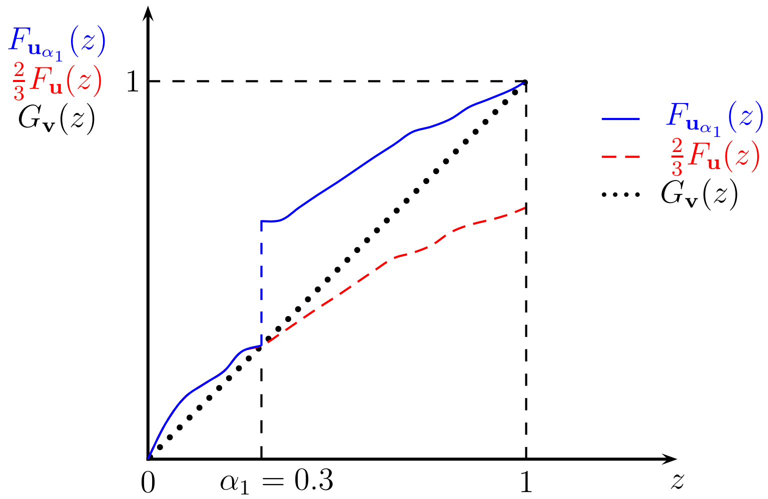

Figure 4: Population poverty incidence curves and dominance of the larger population

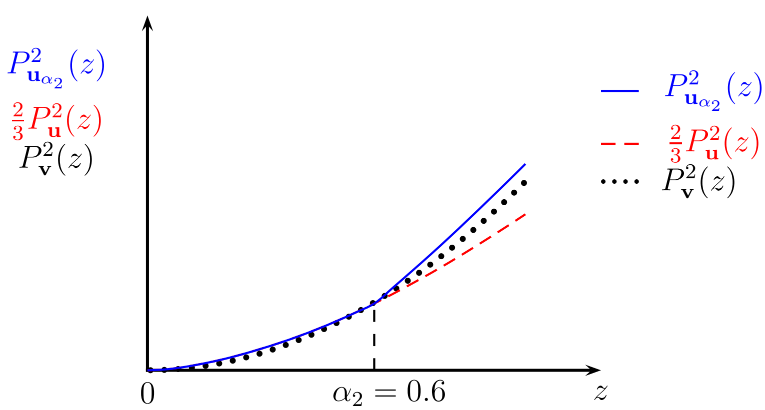

Figure 5: Population P^2 curves and dominance of the larger population

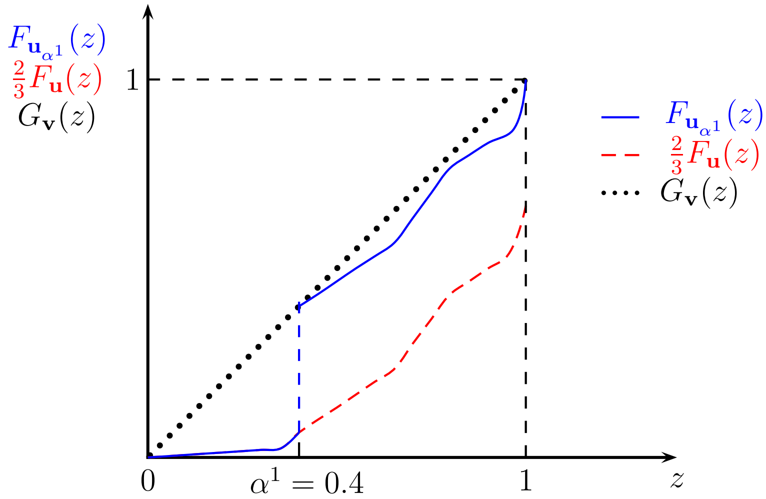

Figure 6: Population poverty incidence curves and dominance of the smaller population

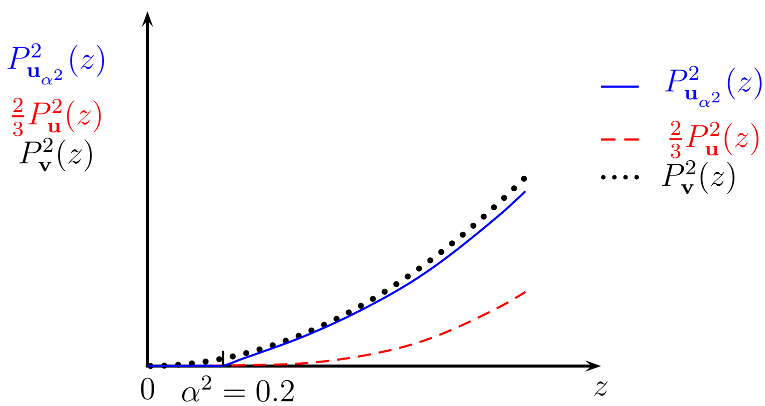

Figure 7: Population P^2 curves and dominance of the smaller population

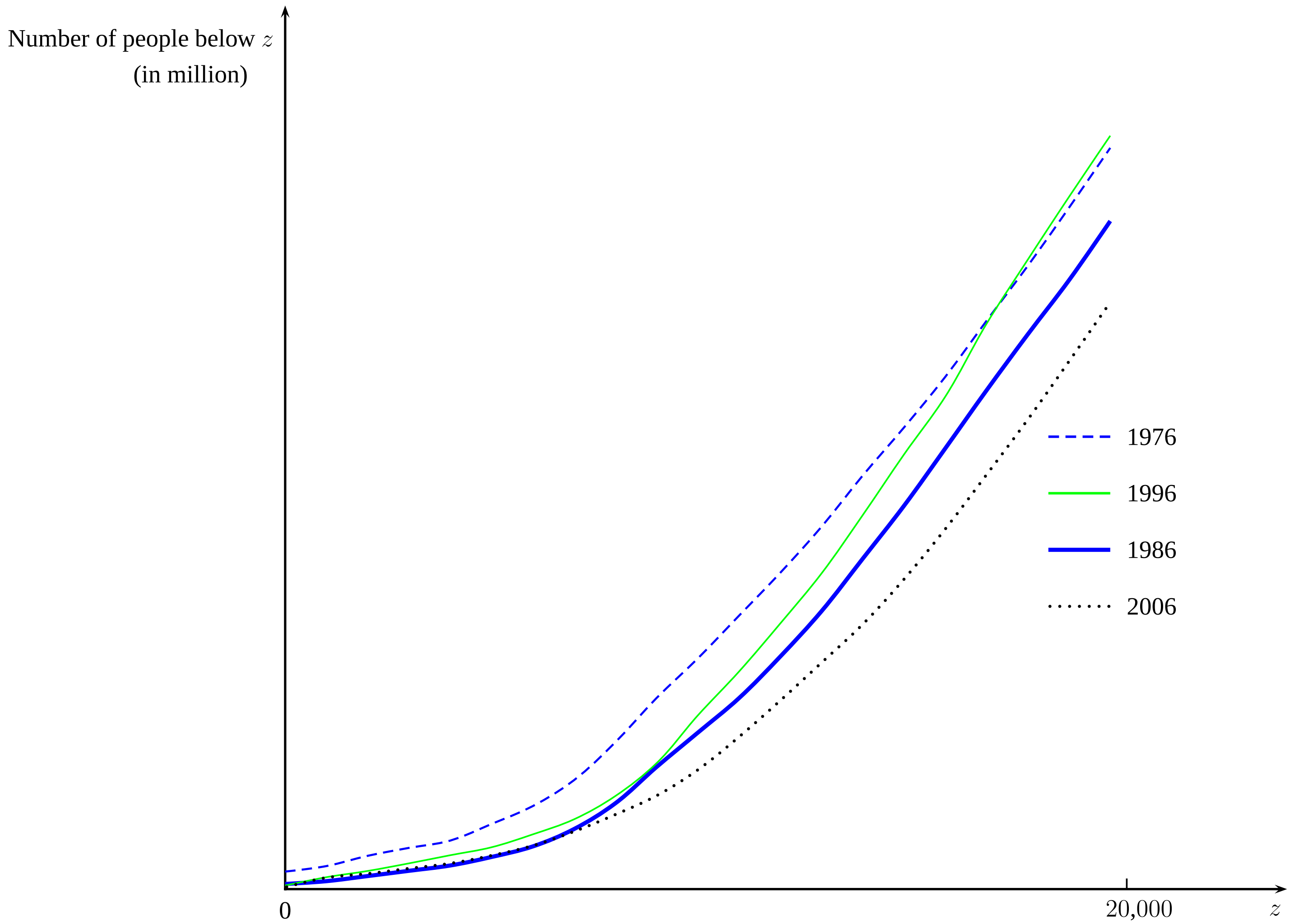

Figure 8: Canadian cumulative distributions

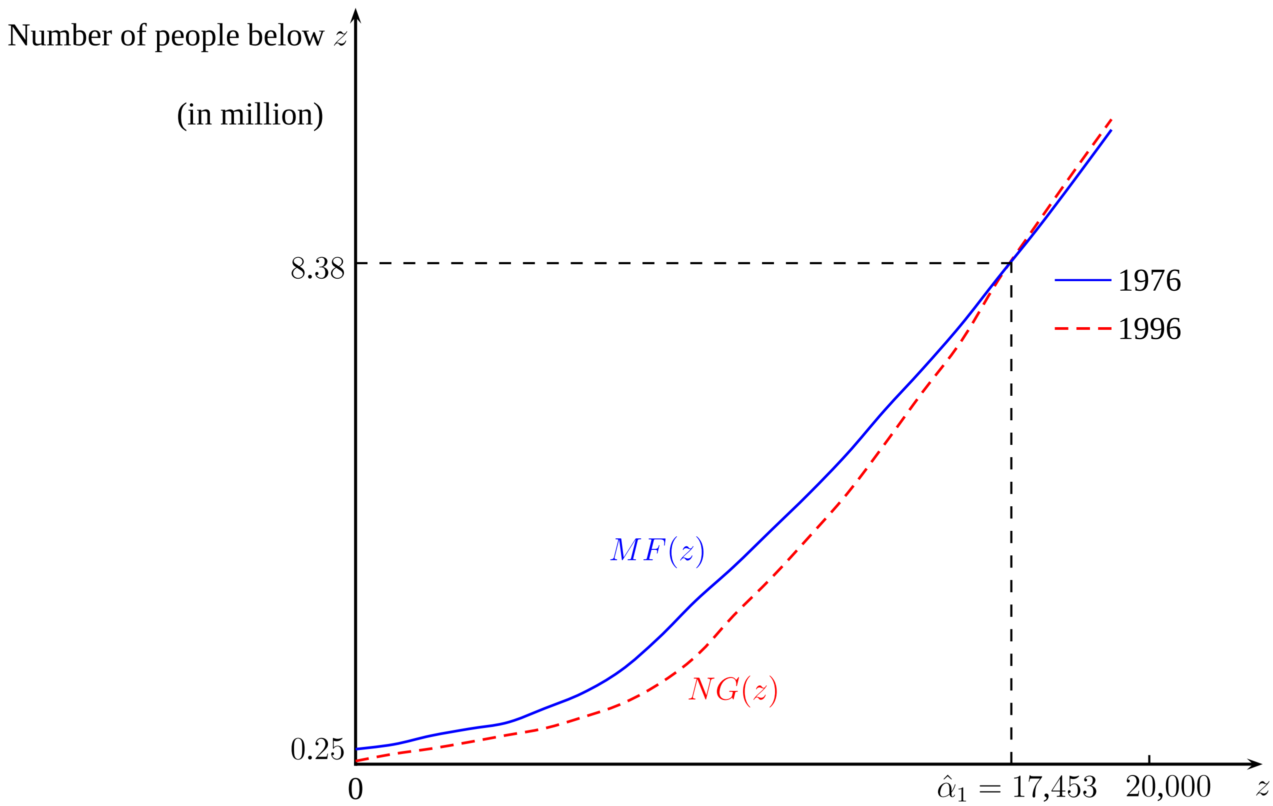

Figure 9: Cumulative distributions of 1976 and 1996 and the critical level

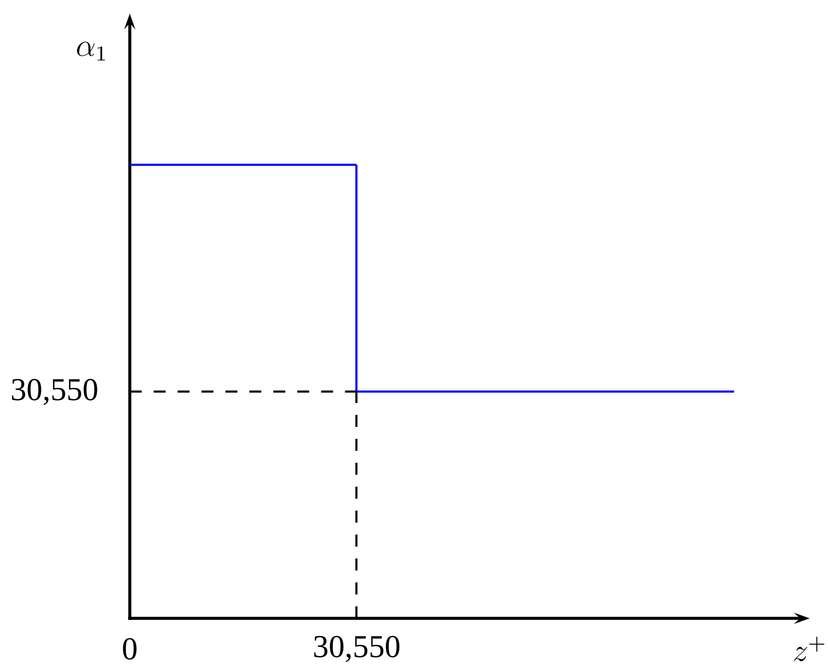

Figure 10: Relation between z^+ and α_1

Table 1: Population sizes and upper bounds of the critical level — large dominates small

α	$\frac{M}{N}=\frac{1}{4}$	$\frac{M}{N}=\frac{1}{2}$	$\frac{M}{N}=\frac{2}{3}$	$\frac{M}{N}=\frac{3}{4}$
α_1	0.05	0.2	0.3	0.5
α_2	0.05	0.3	0.6	0.85

Table 2: Population sizes and lower bounds of the critical level — small dominates large

α	$\frac{M}{N}=\frac{1}{4}$	$\frac{M}{N}=\frac{1}{2}$	$\frac{M}{N}=\frac{2}{3}$	$\frac{M}{N}=\frac{3}{4}$
α^1	0.95	0.85	0.4	0.3
α^2	0.5	0.35	0.2	0.15

Table 3: First-order dominance tests

Dominance tests	Asymptotic p-value		Bootstrap p-value	
	LR	t	LR	t
1986 dominates 1976	0.000	0.000	0.000	0.000
1996 dominates 1976	0.000	0.005	0.002	0.000
2006 dominates 1976	0.000	0.000	0.000	0.000
1996 dominates 1986	0.500	0.500	-	-
2006 dominates 1986	0.000	0.027	0.000	0.000
2006 dominates 1996	0.000	0.019	0.000	0.000

Table 4: Estimates of the upper bound of range of critical levels over which the larger population dominates the smaller one

	1986 dominates 1976		1996 dominates 1976	
s	$\hat{\alpha}_s$	$\hat{\sigma}_s$	$\hat{\alpha}_s$	$\hat{\sigma}_s$
$s = 1$	30,550	1,639	17,453	1,129
$s = 2$	48,294	2,153	30,708	2,104
$s = 3$	69,958	3,854	41,263	2,653
$s = 4$	92,847	5,678	52,203	3,464

Note: All amounts are in 2002 constant dollars; $z^{+} =$ *\$100,000.*

Table 5: Estimates of the upper bound of the range of critical levels over which the larger population dominates the smaller one

	2006 dominates 1976		2006 dominates 1996	
s	$\hat{\alpha}_s$	$\hat{\sigma}_s$	$\hat{\alpha}_s$	$\hat{\sigma}_s$
$s = 1$	33,103	536	49,592	1,674
$s = 2$	49,628	1,382	90,278	8,772
$s = 3$	68,704	2,289	140,544	16,691
$s = 4$	88,770	3,464	192,319	24,773

Note: All amounts are in 2002 constant dollars; $z^{+} =$ *\$100,000 (for 1976) and* $z^{+} =$ *\$200,000 (for 1996).*

Table 6: Estimates of lower bound of the critical level

	1986 dominates 1996	
s	$\hat{\alpha}^s$	$\hat{\sigma}^s$
$s = 1$	-	-
$s = 2$	23,878	1,098
$s = 3$	19,592	1,003
$s = 4$	17,609	965

Note: All amounts are in 2002 constant dollars; $z^{+} =$ *\$30,000.*

8 Appendix

8.1 Graphical illustrations of higher orders of dominance

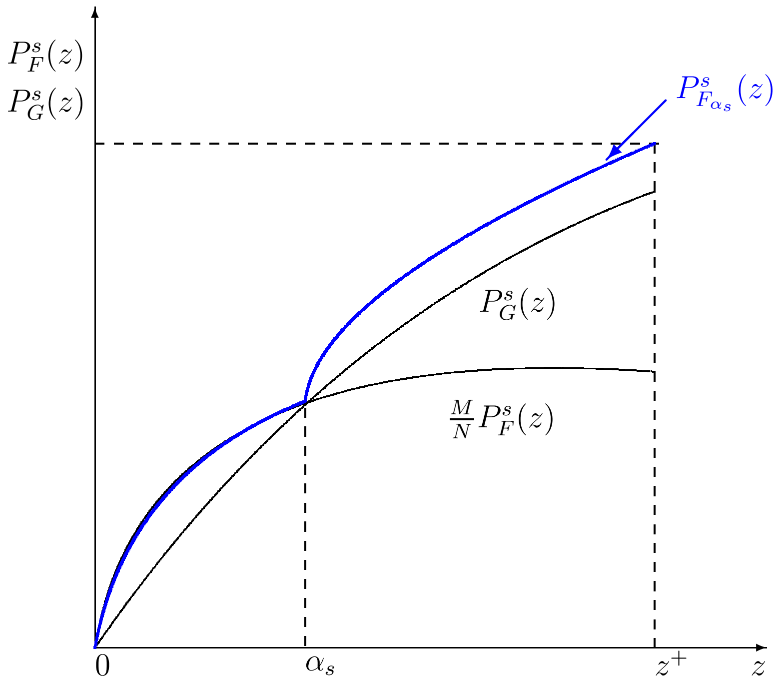

Figure 11: P^s curves and dominance of the larger population

Figures 11, 12 and 13 display FGT curves (adjusted for differences in population sizes) for a given order of dominance $s \geq 2$. In Figure 11, the larger population (with cumulative distribution G) dominates the smaller one with cumulative distribution F. The three curves $\frac{M}{N}P_F^s(z)$, $P_G^s(z)$ and $P_{F_{\alpha_s}}^s(z)$ cross at the same point since we assume that $\mathbf{VDU_s}$ is satisfied and because $P_{F_{\alpha_s}}^s(z)$ coincides with $\frac{M}{N}P_F^s(z)$ when $\alpha_s = z$.

In Figures 12 and 13, we show two cases for the dominance of the smaller population over the larger one. In the first case, a censoring point z^s is introduced. As defined in Section 3, z^s is the censoring income value at which $P_{F_{\alpha_s}}^s$ and P_G^s intersect. Figure 12 is a more general case; Figure 13 occurs when z^s is equal to z^+.

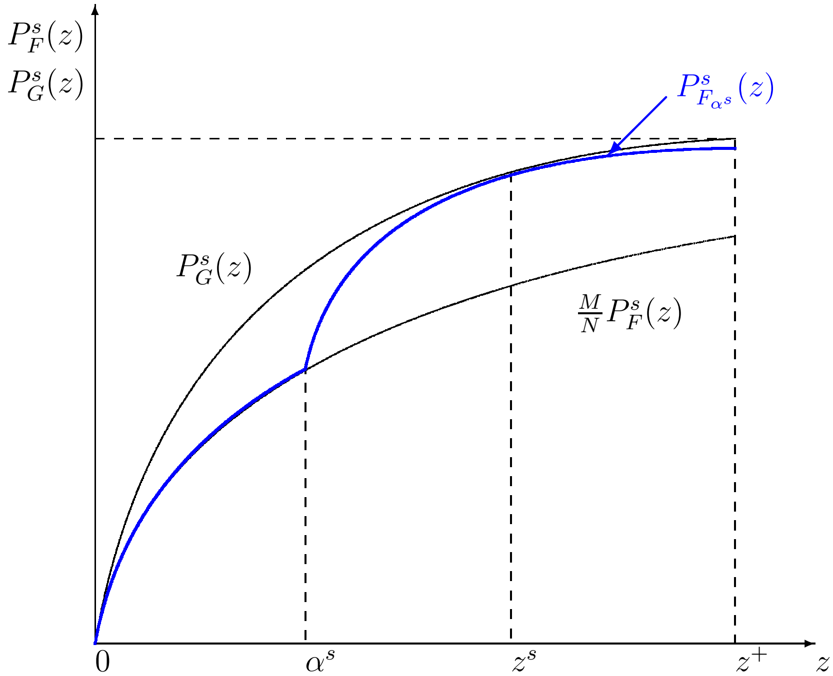

Figure 12: P^s curves and dominance of the smaller population (case 1)

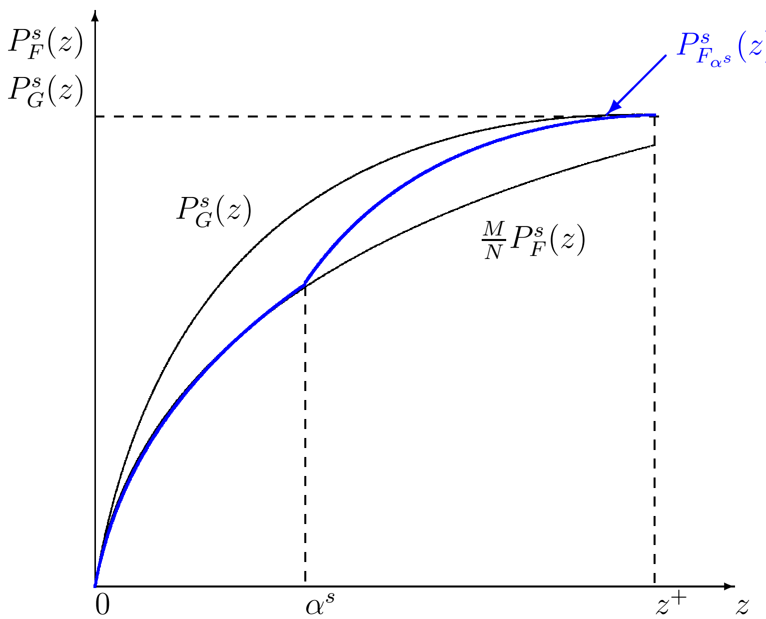

Figure 13: P^s curves and dominance of the smaller population (case 2)

8.2 Proof of Theorems (1) and (2)

The proof is similar to that of Theorem 3 in Davidson and Duclos (2000) (henceforth DD) on page 1460. Let $\psi(z) = \frac{M}{N} P_F^1(z) + \frac{(N-M)}{N} I(\alpha^1 \le z) - P_G^1(z)$ and then $\psi(\alpha^1) = \frac{M}{N} P_F^1(\alpha^1) + \frac{(N-M)}{N} - P_G^1(\alpha^1)$. An estimator of $\psi(z)$ is $\hat{\psi}(z) = \frac{\hat{M}}{\hat{N}} \hat{P}_F^1(z) + \frac{(\hat{N}-\hat{M})}{\hat{N}} I(\alpha^1 \le z) - \hat{P}_G^1(z)$. We have that $\hat{M} = \frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}$ and $\hat{N} = \frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}$ where $(w_i^{\mathbf{u}})_{i=1}^m \sim iid\left(\mu_{w^{\mathbf{u}}}, \sigma^2_{w^{\mathbf{u}}}\right)$ and $\left(w_j^{\mathbf{v}}\right)_{j=1}^n \sim iid\left(\mu_{w^{\mathbf{v}}}, \sigma^2_{w^{\mathbf{v}}}\right)$, because $(\tilde{w}_i^{\mathbf{u}})_{i=1}^m$ and $\left(\tilde{w}_j^{\mathbf{v}}\right)_{j=1}^n$ are assumed to be iid.

According to (37), $\psi(\alpha^1) \equiv 0$. So, using a Taylor expansion for $\psi(\hat{\alpha}^1)$, there exists $\tilde{\alpha}^1$ such that $|\tilde{\alpha}^1 - \alpha^1| < |\hat{\alpha}^1 - \alpha^1|$ and

$$\psi\left(\hat{\alpha}^1\right) \approx \left(\hat{\alpha}^1 - \alpha^1\right) \psi'\left(\tilde{\alpha}^1\right).$$

For m and $n \to \infty$ such that $r = \frac{m}{n}$ remains constant, we have that $\hat{\alpha}^1 \to \alpha^1$ and $\tilde{\alpha}^1 \to \alpha^1$. Then, for large samples, $\psi'(\tilde{\alpha}^1) \neq 0$ because $\psi'(\alpha^1) \neq 0$ by assumption, and then $(\hat{\alpha}^1 - \alpha^1) \approx \frac{\psi(\hat{\alpha}^1)}{\psi'(\tilde{\alpha}^1)}$. We can use the following result as in DD:

$$\hat{\psi}\left(\alpha^1\right) + \psi\left(\hat{\alpha}^1\right) = o\left(n^{-1/2}\right). \tag{46}$$

Therefore,

$$\left(\hat{\alpha}^1 - \alpha^1\right) \approx -\frac{\hat{\psi}(\alpha^1)}{\psi'(\tilde{\alpha}^1)}. \tag{47}$$

As defined above, we have

$$\begin{aligned} \hat{\psi}\left(\alpha^1\right) &= \frac{\hat{M}}{\hat{N}} \hat{P}_F^1\left(\alpha^1\right) + \left(1 - \frac{\hat{M}}{\hat{N}}\right) - \hat{P}_G^1\left(\alpha^1\right) & (48) \\ &= \frac{m^{-1}\sum_{i=1}^m w_i^{\mathbf{u}} (\alpha^1 - u_i)_+^0}{n^{-1}\sum_{j=1}^n w_j^{\mathbf{v}}} + \left(1 - \frac{m^{-1}\sum_{i=1}^m w_i^{\mathbf{u}}}{n^{-1}\sum_{j=1}^n w_j^{\mathbf{v}}}\right) & (49) \\ &\quad - \frac{n^{-1}\sum_{i=j}^n w_j^{\mathbf{v}} (\alpha^1 - v_j)_+^0}{n^{-1}\sum_{j=1}^n w_j^{\mathbf{v}}}. \end{aligned}$$

Let

$$\hat{\Gamma}(\alpha^1) = \begin{pmatrix} m^{-1}\sum_{i=1}^m w_i^{\mathbf{u}} (\alpha^1 - u_i)_+^0 \\ m^{-1}\sum_{i=1}^m w_i^{\mathbf{u}} \\ n^{-1}\sum_{i=j}^n w_j^{\mathbf{v}} (\alpha^1 - v_j)_+^0 \\ n^{-1}\sum_{j=1}^n w_j^{\mathbf{v}} \end{pmatrix} \tag{50}$$

with $\Gamma(\alpha^1) = E\left[\hat{\Gamma}(\alpha^1)\right]$. Note that all of the elements of $\hat{\Gamma}(\alpha^1)$ are sums of iid observations. Let $\Gamma_i(\alpha^1)$ be the ith element of $\Gamma(\alpha^1)$. We can then write

$$\psi\left(\alpha^1\right) = \frac{\Gamma_1(\alpha^1)}{\Gamma_4(\alpha^1)} + \left(1 - \frac{\Gamma_2(\alpha^1)}{\Gamma_4(\alpha^1)}\right) - \frac{\Gamma_3(\alpha^1)}{\Gamma_4(\alpha^1)}.$$

Because $\Gamma_2(\alpha^1) = E\left[m^{-1}\sum_{i=1}^{m} w_i^{\mathbf{u}}\right] = \mu_{w^{\mathbf{u}}} \equiv \Gamma_2$ and $\Gamma_4(\alpha^1) = E\left[n^{-1}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right] = \mu_{w^{\mathbf{v}}} \equiv \Gamma_4$, we can rewrite $\psi\left(\alpha^1\right)$ as

$$\psi\left(\alpha^1\right) = \frac{\Gamma_1(\alpha^1)}{\Gamma_4} + \left(1 - \frac{\Gamma_2}{\Gamma_4}\right) - \frac{\Gamma_3(\alpha^1)}{\Gamma_4}, \tag{51}$$

and similarly for $\hat{\psi}\left(\alpha^1\right)$ by replacing Γ by $\hat{\Gamma}$ in (51). Let the gradient of $\psi\left(\alpha^1\right)$ with respect to $\Gamma(\alpha^1)$ be given by the 4×1 vector H:

$$H(\alpha^1) = \begin{pmatrix} \Gamma_4^{-1} \\ -\Gamma_4^{-1} \\ -\Gamma_4^{-1} \\ -\frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^2} + \frac{\Gamma_2}{\Gamma_4{}^2} + \frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^2} \end{pmatrix}. \tag{52}$$

Then, using a usual Taylor's approximation, we have

$$\hat{\psi}\left(\alpha^1\right) - \psi\left(\alpha^1\right) = H(\alpha^1)'\left[\hat{\Gamma}(\alpha^1) - \Gamma(\alpha^1)\right] + O\left(n^{-1/2}\right) \tag{53}$$

where H' is the transpose of H. Therefore,

$$\text{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^1\right)\right) = nH(\alpha^1)' \operatorname{var}\left(\hat{\Gamma}(\alpha^1)\right) H(\alpha^1). \tag{54}$$

We can now give the expression of $\operatorname{var}\left(\hat{\Gamma}(\alpha^1)\right)$, which, for simplicity, we sometimes write $\operatorname{var}\left(\hat{\Gamma}\right)$:

$$\operatorname{var}\left(\hat{\Gamma}\right) = \begin{bmatrix} \operatorname{var}\left(\hat{\Gamma}_1\right) & \operatorname{cov}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) & 0 & 0 \\ \operatorname{cov}\left(\hat{\Gamma}_2, \hat{\Gamma}_1\right) & \operatorname{var}\left(\hat{\Gamma}_2\right) & 0 & 0 \\ 0 & 0 & \operatorname{var}\left(\hat{\Gamma}_3\right) & \operatorname{cov}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) \\ 0 & 0 & \operatorname{cov}\left(\hat{\Gamma}_4, \hat{\Gamma}_3\right) & \operatorname{var}\left(\hat{\Gamma}_4\right) \end{bmatrix}. \tag{55}$$

The elements of $\operatorname{var}\left(\hat{\Gamma}(\alpha^1)\right)$ can be estimated consistently using the sample covariance matrix of the elements of $\Gamma(\alpha^1)$. Using equations (52), (54) and (55) together gives

$$\operatorname{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^1\right)\right) = \left(\begin{array}{c} 2n\left[\frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^3} - \frac{\Gamma_2}{\Gamma_4{}^3} - \frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^3}\right]\operatorname{cov}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) + \\ n\left[\frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^2} - \frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^2} + \frac{\Gamma_2}{\Gamma_4{}^2}\right]^2 \operatorname{var}\left(\hat{\Gamma}_4\right) + \\ \frac{n}{\Gamma_4{}^2}\left[\operatorname{var}\left(\hat{\Gamma}_1\right) + \operatorname{var}\left(\hat{\Gamma}_2\right)\right] - \\ \frac{2n}{\Gamma_4{}^2}\operatorname{cov}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) + \\ \frac{n}{\Gamma_4{}^2}\operatorname{var}\left(\hat{\Gamma}_3\right) \end{array}\right). \tag{56}$$

Note that

$$\begin{aligned} \operatorname{cov}\left(\hat{\Gamma}_1(\alpha^1), \hat{\Gamma}_2\right) &= m^{-2}\sum_{i=1}^{m}\sum_{j=1}^{m}\operatorname{cov}\left[w_i^{\mathbf{u}}\left(\alpha^1 - u_i\right)_+^0, w_j^{\mathbf{u}}\right] \\ &= m^{-2}\sum_{i=1}^{m}\operatorname{cov}\left[w_i^{\mathbf{u}}\left(\alpha^1 - u_i\right)_+^0, w_i^{\mathbf{u}}\right] \\ &\quad + m^{-2}\sum_{i\neq j}^{m}\underbrace{\operatorname{cov}\left[w_i^{\mathbf{u}}\left(\alpha^1 - u_i\right)_+^0, w_j^{\mathbf{u}}\right]}_{=0} \\ &= m^{-2}\sum_{i=1}^{m}\operatorname{cov}\left[w_i^{\mathbf{u}}\left(\alpha^1 - u_i\right)_+^0, w_i^{\mathbf{u}}\right] \\ &= m^{-1}E\left[\left(w^{\mathbf{u}}\right)^2\left(\alpha^1 - u\right)_+^0\right] \\ &\quad - m^{-1}\Gamma_1(\alpha^1)\Gamma_2 \end{aligned} \tag{57}$$

and, in the same manner,

$$\operatorname{cov}\left(\hat{\Gamma}_3(\alpha^1), \hat{\Gamma}_4\right) = n^{-1}E\left[\left(w^{\mathbf{v}}\right)^2\left(\alpha^1 - v\right)_+^0\right] - n^{-1}\Gamma_3(\alpha^1)\Gamma_4, \tag{58}$$

$$\operatorname{var}\left(\hat{\Gamma}_1(\alpha^1)\right) = m^{-1}\operatorname{var}\left[w^{\mathbf{u}}\left(\alpha^1 - u\right)_+^0\right], \tag{59}$$

$$\operatorname{var}\left(\hat{\Gamma}_3(\alpha^1)\right) = n^{-1}\operatorname{var}\left[w^{\mathbf{v}}\left(\alpha^1 - v\right)_+^0\right], \tag{60}$$

$$\operatorname{var}\left(\hat{\Gamma}_2\right) = m^{-1}\sigma_{w^{\mathbf{u}}}^2, \tag{61}$$

$$\operatorname{var}\left(\hat{\Gamma}_4\right) = n^{-1}\sigma_{w^{\mathbf{v}}}^2. \tag{62}$$

Putting (57), (58), (59), (60), (61) and (62) together gives

$$\operatorname{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^{1}\right)\right)=\left(\begin{array}{c}\frac{1}{\Gamma_4{}^2}\operatorname{var}\left[w^{\mathbf{v}}\left(\alpha^{1}-v\right)_{+}^{0}\right]+\\ \frac{r-1}{\Gamma_4{}^2}\left[\operatorname{var}w^{\mathbf{u}}\left(\alpha^{1}-u\right)_{+}^{0}+\sigma_{w^{\mathbf{u}}}^{2}\right]+\\ \left[\frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^2}-\frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^2}+\frac{\Gamma_2}{\Gamma_4{}^2}\right]^{2}\sigma_{w^{\mathbf{v}}}^{2}-\\ \frac{2r-1}{\Gamma_4{}^2}\left[E\left(\left(w^{\mathbf{u}}\right)^{2}\left(\alpha^{1}-u\right)_{+}^{0}\right)-\Gamma_1(\alpha^1)\Gamma_2\right]+\\ 2\left[\frac{\Gamma_1(\alpha^1)}{\Gamma_4{}^3}-\frac{\Gamma_2}{\Gamma_4{}^3}-\frac{\Gamma_3(\alpha^1)}{\Gamma_4{}^3}\right]\\ \times\left[E\left(\left(w^{\mathbf{v}}\right)^{2}\left(\alpha^{1}-v\right)_{+}^{0}\right)-\Gamma_3(\alpha^1)\Gamma_4\right]\end{array}\right). \qquad (63)$$

By (53), $\hat{\psi}\left(\alpha^{1}\right)$ is a linear combination of sums of iid variables. We can thus apply the Central Limit Theorem, which gives that

$$\sqrt{n}\hat{\psi}\left(\alpha^{1}\right)\xrightarrow{d}N(0,\operatorname{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^{1}\right)\right).$$

Using (47), the asymptotic variance of $\hat{\alpha}^{1}$ is given by

$$\lim_{m,\,n\to\infty}\operatorname{var}\left(\sqrt{n}(\hat{\alpha}^{1}-\alpha^{1})\right)=\frac{\lim\limits_{m,\,n\to\infty}\operatorname{var}\left(\sqrt{n}\hat{\psi}\left(\alpha^{1}\right)\right)}{\left(\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}P_F^0(\alpha^1)-P_G^0(\alpha^1)\right)^{2}}. \qquad (64)$$

It remains to show that

$$\hat{\psi}\left(\alpha^{1}\right)+\psi\left(\hat{\alpha}^{1}\right)=o\left(n^{-1/2}\right).$$

Through equations (44) and (37), we know that $\hat{\psi}\left(\hat{\alpha}^{1}\right)=\psi\left(\alpha^{1}\right)=0$. Then rewrite

$$-\left[\hat{\psi}\left(\alpha^{1}\right)+\psi\left(\hat{\alpha}^{1}\right)\right]=\hat{\psi}\left(\hat{\alpha}^{1}\right)-\hat{\psi}\left(\alpha^{1}\right)-\left[\psi\left(\hat{\alpha}^{1}\right)-\psi\left(\alpha^{1}\right)\right].$$

Using Theorem 2 of DD, we have that $\hat{\alpha}^{1}-\alpha^{1}=O\left(n^{-1/2}\right)$. Simplifying the notation, denote

$$\begin{aligned}\hat{\Psi}\left(\hat{\alpha}^{1},\alpha^{1}\right)&\equiv\hat{\psi}\left(\hat{\alpha}^{1}\right)-\hat{\psi}\left(\alpha^{1}\right)-\left[\psi\left(\hat{\alpha}^{1}\right)-\psi\left(\alpha^{1}\right)\right]\\&=\hat{\psi}\left(\alpha^{1}+O\left(n^{-1/2}\right)\right)-\hat{\psi}\left(\alpha^{1}\right)-\left[\psi\left(\alpha^{1}+O\left(n^{-1/2}\right)\right)-\psi\left(\alpha^{1}\right)\right].\end{aligned}$$

Consequently,

$$\begin{aligned}p\lim\hat{\Psi}\left(\hat{\alpha}^{1},\alpha^{1}\right)&=p\lim\left[\hat{\psi}\left(\alpha^{1}+O\left(n^{-1/2}\right)\right)-\psi\left(\alpha^{1}+O\left(n^{-1/2}\right)\right)\right]\\&\quad-p\lim\hat{\psi}\left(\alpha^{1}\right)+\psi\left(\alpha^{1}\right)\\&=p\lim\hat{\psi}\left(\alpha^{1}\right)-\psi\left(\alpha^{1}\right)-p\lim\hat{\psi}\left(\alpha^{1}\right)+\psi\left(\alpha^{1}\right)\\&=0.\end{aligned}$$

The second equality comes from the fact that, asymptotically, $\hat{\alpha}^1 = \alpha^1 + O\left(n^{-1/2}\right) \to \alpha^1$. Appling Bienaymé-Chebyshev's inequality to $\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)$, we can write that for any $\varepsilon > 0$,

$$\Pr\left(\sqrt{n}\left|\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right| > \varepsilon\right) \leq \frac{\mathrm{var}\left(\sqrt{n}\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right)}{\varepsilon^2}. \tag{65}$$

We can then compute $\mathrm{Avar}\left(\sqrt{n}\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right)$:

$$\begin{aligned} \mathrm{Avar}\left(\sqrt{n}\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right) &= \mathrm{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^1 + O\left(n^{-1/2}\right)\right)\right) \\ &\quad + \mathrm{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^1\right)\right) \\ &\quad -2\,\mathrm{Acov}\left(\sqrt{n}\hat{\psi}\left(\alpha^1 + O\left(n^{-1/2}\right)\right), \sqrt{n}\hat{\psi}\left(\alpha^1\right)\right). \end{aligned} \tag{66}$$

Using the expression of $\mathrm{Avar}\left(\sqrt{n}\left(\hat{\psi}\left(\alpha^1\right)\right)\right)$ given in (54), note that $\mathrm{Avar}\left(\sqrt{n}\hat{\psi}\left(\alpha^1 + O\left(n^{-1/2}\right)\right)\right)$ can be written in a similar way. Thus,

$$\begin{aligned} \mathrm{Avar}\left(\sqrt{n}\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right) &= nH\left(\alpha^1 + O\left(n^{-1/2}\right)\right)' \mathrm{var}\left(\hat{\Gamma}(\alpha^1 + O\left(n^{-1/2}\right))\right) \\ &\quad \times H\left(\alpha^1 + O\left(n^{-1/2}\right)\right) \\ &\quad + nH\left(\alpha^1\right)' \mathrm{var}\left(\hat{\Gamma}(\alpha^1)\right) H\left(\alpha^1\right) \\ &\quad -2nH\left(\alpha^1 + O\left(n^{-1/2}\right)\right)' \mathrm{cov}\left(\hat{\Gamma}(\alpha^1 + O\left(n^{-1/2}\right)), \hat{\Gamma}(\alpha^1)\right) \\ &\quad \times H\left(\alpha^1\right). \end{aligned}$$

Asymptotically, $H\left(\alpha^1 + O\left(n^{-1/2}\right)\right) \approx H\left(\alpha^1\right)$ and $\hat{\Gamma}(\alpha^1 + O\left(n^{-1/2}\right) \approx \hat{\Gamma}(\alpha^1)$. Hence,

$$\mathrm{cov}\left(\hat{\Gamma}(\alpha^1 + O\left(n^{-1/2}\right)), \hat{\Gamma}(\alpha^1)\right) \approx \mathrm{var}\left(\hat{\Gamma}(\alpha^1)\right).$$

We thus have

$$\mathrm{Avar}\left(\sqrt{n}\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right) = 0.$$

We obtain that

$$\lim_{m,\, n \to \infty} \Pr\left(\sqrt{n}\left|\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right)\right| > \varepsilon\right) = 0. \tag{67}$$

Because $\hat{\Psi}\left(\hat{\alpha}^1, \alpha^1\right) = \hat{\psi}\left(\alpha^1\right) + \psi\left(\hat{\alpha}^1\right)$, then

$$\hat{\psi}\left(\alpha^1\right) + \psi\left(\hat{\alpha}^1\right) = o\left(n^{-1/2}\right).$$

Using (47), the asymptotic variance of $\hat{\alpha}^1$ is given by

$$\lim_{m,\,n\to\infty} \operatorname{var}\left(\sqrt{n}(\hat{\alpha}^1-\alpha^1)\right) = \frac{\lim\limits_{m,\,n\to\infty} \operatorname{var}\left(\sqrt{n}\hat{\psi}\left(\alpha^1\right)\right)}{\left(\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}} P_F^0(\alpha^1) - P_G^0(\alpha^1)\right)^2} \tag{68}$$

where $\lim\limits_{m,\,n\to\infty} \operatorname{var}\left(\sqrt{n}\hat{\psi}\left(\alpha^1\right)\right)$ is given in (63).

We use a similar procedure to derive the asymptotic variance of $\hat{\alpha}^s$ for $s \geq 2$. Recall that, for $s \geq 2$, α^s satisfies the following equation

$$\frac{M}{N} P_F^s(z^s) + \frac{N-M}{N}(z^s-\alpha^s)^{s-1} - P_G^s(z^s) = 0. \tag{69}$$

Denote $\phi\left(\alpha^s\right) = \frac{M}{N} P_F^s(z^s) + \frac{N-M}{N}(z^s-\alpha^s)^{s-1} - P_G^s(z^s)$. Using a Taylor expansion, there exists $\tilde{\alpha}^s$ such that $|\tilde{\alpha}^s - \alpha^s| < |\hat{\alpha}^s - \alpha^s|$ and

$$\phi\left(\hat{\alpha}^s\right) \approx \left(\hat{\alpha}^s - \alpha^s\right) \phi'\left(\tilde{\alpha}^s\right)$$

where $\phi'\left(\tilde{\alpha}^s\right) = -\left(s-1\right) \frac{N-M}{N}(z^+ - \tilde{\alpha}^s)^{s-2}$ and $\phi'\left(\tilde{\alpha}^s\right) \neq 0$ by assumption. From (47), we obtain that

$$\left(\hat{\alpha}^s - \alpha^s\right) \approx -\frac{\hat{\phi}\left(\alpha^s\right)}{\phi'\left(\tilde{\alpha}^s\right)}.$$

Notice that $\hat{\phi}\left(\alpha^s\right) = \frac{\hat{M}}{\hat{N}} \hat{P}_F^s(z^s) + \frac{\hat{N}-\hat{M}}{\hat{N}}(z^s-\alpha^s)^{s-1} - \hat{P}_G^s(z^s)$ and suppose that $z^s \leq z^+$ is known. Applying the previous results and assuming that the moments of order $2\left(s-1\right)$ exist, we find that

$$\begin{aligned}
\operatorname{Avar}\left(\sqrt{n}\hat{\phi}\left(\alpha^s\right)\right) &= \lim_{m,\,n\to\infty} \operatorname{var}\left(\sqrt{n}\hat{\phi}\left(\alpha^s\right)\right) \\
&= \left(\begin{array}{c}
\frac{1}{\Gamma_4{}^2} \operatorname{var}\left[w^{\mathbf{v}}\left(z^s-v\right)_+^{s-1}\right] + \\
\frac{r^{-1}}{\Gamma_4{}^2}\left[\operatorname{var} w^{\mathbf{u}}\left(z^s-u\right)_+^{s-1} + \left(z^s-\alpha^s\right)^{2s-2}\sigma_{w^{\mathbf{u}}}^2\right] + \\
\left[\frac{\Gamma_3}{\Gamma_4{}^2} - \frac{\Gamma_1}{\Gamma_4{}^2} + \frac{\Gamma_2}{\Gamma_4{}^2}\left(z^s-\alpha^s\right)^{s-1}\right]^2 \sigma_{w^{\mathbf{v}}}^2 - \\
\frac{2r^{-1}}{\Gamma_4{}^2}\left(z^s-\alpha^s\right)^{s-1}\left[E\left(\left(w^{\mathbf{u}}\right)^2\left(z^s-u\right)_+^{s-1}\right) - \Gamma_1\Gamma_2\right] + \\
2\left[\frac{\Gamma_1}{\Gamma_4{}^3} - \frac{\Gamma_2}{\Gamma_4{}^3}\left(z^s-\alpha^s\right)^{s-1} - \frac{\Gamma_3}{\Gamma_4{}^3}\right] \times \\
\left[E\left(\left(w^{\mathbf{v}}\right)^2\left(z^s-v\right)_+^{s-1}\right) - \Gamma_3\Gamma_4\right]
\end{array}\right)
\end{aligned}$$

where $\Gamma_1 = E\left[m^{-1}\sum_{i=1}^m w_i^{\mathbf{u}}\left(z^s-u_i\right)_+^{s-1}\right]$ and $\Gamma_2 = \mu_{w^{\mathbf{u}}}$; $\Gamma_3 = E\left[n^{-1}\sum_{j=1}^n w_j^{\mathbf{v}}\left(z^s-v_j\right)_+^{s-1}\right]$ and $\Gamma_4 = \mu_{w^{\mathbf{v}}}$.

Therefore, the asymptotic variance of $\hat{\alpha}^s$ is given by

$$\lim_{m,\, n\to\infty} \operatorname{var}\left(\sqrt{n}(\hat{\alpha}^s - \alpha^s)\right) = \frac{\lim\limits_{m,\, n\to\infty} \operatorname{var}\left(\sqrt{n}\hat{\phi}\left(\alpha^s\right)\right)}{\left[(s-1)\left(1-\frac{\mu_{w^{\mathbf{u}}}}{\mu_{w^{\mathbf{v}}}}\right)\left(z^s-\alpha^s\right)^{s-2}\right]^2}.$$

Similar arguments can be used to establish the asymptotic distribution of $\hat{\alpha}_s$. ■

8.3 Asymptotic equivalence of statistics

Proposition 1

Suppose that $r = \frac{m}{n}$ *remains constant as* m *and* n *tend to infinity. For* $s \geq 1$ *and for any pair* (z, α) *in the interior of* $[z^-, z^+] \otimes [\alpha^-, \alpha^+]$*, such that* $P_G^s(z) = P_{F_\alpha}^s(z)$*, the statistic* $LR\left(z, \alpha\right)$ *tends to the square of the asymptotic* t*-statistic where*

$$t^2\left(z,\alpha\right) = \frac{\left(\Delta P^s\left(z,\alpha\right)\right)^2}{\operatorname{Avar}\left(\sqrt{n}\left(\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)\right)\right)} + O\left(n^{-1/2}\right) \tag{70}$$

with $\Delta P^s\left(z,\alpha\right) = \underset{m,\, n\to\infty}{p\lim} \sqrt{n}\left(\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)\right) = O\left(1\right)$*.*

Proof

(Based on Davidson (2009)). We know that

$$\hat{P}_{F_\alpha}^s(z) = \left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\left(z-u_i\right)_+^{s-1}\right) \Bigg/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right) \tag{71}$$

$$+\left[1-\left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\right) \Bigg/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right)\right]\left(z-\alpha\right)_+^{s-1} \tag{72}$$

and

$$\hat{P}_G^s(z) = \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\left(z-v_j\right)_+^{s-1}\right) \Bigg/ \left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right). \tag{73}$$

Let $\hat{\Gamma}_1 = \frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}\left(z-u_i\right)_+^{s-1}$, $\hat{\Gamma}_2 = \frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}$, $\hat{\Gamma}_3 = \frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\left(z-v_j\right)_+^{s-1}$ and $\hat{\Gamma}_4 = \frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}$. Thus

$$\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z) = \frac{\hat{\Gamma}_3}{\hat{\Gamma}_4} - \frac{\hat{\Gamma}_1}{\hat{\Gamma}_4} - \left(1-\frac{\hat{\Gamma}_2}{\hat{\Gamma}_4}\right)\left(z-\alpha\right)_+^{s-1}.$$

Also denote $\hat{\Gamma}^s = \left(\hat{\Gamma}_1, \hat{\Gamma}_2, \hat{\Gamma}_3, \hat{\Gamma}_4\right)'$ and $\Gamma_i = E\left[\hat{\Gamma}_i\right]$ for $i = 1, ..., 4$. We use a Taylor approximation to compute the variance of $\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)$. Let

$$H^s = \begin{bmatrix} -\Gamma_4^{-1} \\ (z-\alpha)_+^{s-1}\Gamma_4^{-1} \\ \Gamma_4^{-1} \\ \Gamma_4^{-2}\left(\Gamma_1 - (z-\alpha)_+^{s-1}\Gamma_2 - \Gamma_3\right) \end{bmatrix}$$

and

$$\operatorname{var}\left(\hat{\Gamma}^s\right) = \begin{bmatrix} \operatorname{var}\left(\hat{\Gamma}_1\right) & \operatorname{cov}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) & 0 & 0 \\ \operatorname{cov}\left(\hat{\Gamma}_2, \hat{\Gamma}_1\right) & \operatorname{var}\left(\hat{\Gamma}_2\right) & 0 & 0 \\ 0 & 0 & \operatorname{var}\left(\hat{\Gamma}_3\right) & \operatorname{cov}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) \\ 0 & 0 & \operatorname{cov}\left(\hat{\Gamma}_4, \hat{\Gamma}_3\right) & \operatorname{var}\left(\hat{\Gamma}_4\right) \end{bmatrix}.$$

Therefore,

$$\begin{aligned}
& \operatorname{Avar}\left(\sqrt{n}\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)\right) \\
= & \lim_{m,\, n\to\infty} n\left(H^s\right)' \operatorname{var}\left(\hat{\Gamma}^s\right)\left(H^s\right) \\
= & \lim_{m,\, n\to\infty} \left\{ \frac{n}{\Gamma_4^2}\left[\operatorname{var}\left(\hat{\Gamma}_1\right) - 2(z-\alpha)_+^{s-1}\operatorname{cov}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) + \left[(z-\alpha)_+^{s-1}\right]^2 \operatorname{var}\left(\hat{\Gamma}_2\right)\right] \right. \\
& + \frac{n}{\Gamma_4^2}\operatorname{var}\left(\hat{\Gamma}_3\right) + \frac{n\left[\Gamma_1 - (z-\alpha)_+^{s-1}\Gamma_2 - \Gamma_3\right]^2}{\Gamma_4^4}\operatorname{var}\left(\hat{\Gamma}_4\right) \\
& \left. + \frac{2n\left[\Gamma_1 - (z-\alpha)_+^{s-1}\Gamma_2 - \Gamma_3\right]}{\Gamma_4^3}\operatorname{cov}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) \right\}.
\end{aligned}$$

But $P^s_G(z) = P^s_{F_\alpha}(z) \Longrightarrow \frac{\Gamma_3}{\Gamma_4} - \frac{\Gamma_1}{\Gamma_4} - \left(1 - \frac{\Gamma_2}{\Gamma_4}\right)(z-\alpha)_+^{s-1} = 0$. Then $\Gamma_1 - (z-\alpha)_+^{s-1}\Gamma_2 - \Gamma_3 = -(z-\alpha)_+^{s-1}\Gamma_4$. Consequently, $\operatorname{Avar}\left(\sqrt{n}\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)\right)$ becomes

$$\begin{aligned}
& \operatorname{Avar}\left(\sqrt{n}\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)\right) \\
= & \lim_{m,\, n\to\infty} \left\{ \frac{n}{\Gamma_4^2}\left[\operatorname{var}\left(\hat{\Gamma}_1\right) - 2(z-\alpha)_+^{s-1}\operatorname{cov}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) + \left[(z-\alpha)_+^{s-1}\right]^2 \operatorname{var}\left(\hat{\Gamma}_2\right)\right] \right. \\
& \left. + \frac{n}{\Gamma_4^2}\left[\operatorname{var}\left(\hat{\Gamma}_3\right) - 2(z-\alpha)_+^{s-1}\operatorname{cov}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) + \left[(z-\alpha)_+^{s-1}\right]^2 \operatorname{var}\left(\hat{\Gamma}_4\right)\right] \right\}.
\end{aligned}$$

Now consider the ELR statistic. Recall that $ELR = 2\left[ELF - ELF\left(z, \alpha\right)\right]$. We have that

$$ELF = \sum_{i=1}^{m} \log\left(\frac{1}{m}\right) + \sum_{j=1}^{n} \log\left(\frac{1}{n}\right). \tag{74}$$

For ease of exposition, denote $\mathbf{u}_{i\alpha} = (z - u_i)_+^{s-1} - (z - \alpha)_+^{s-1}$ and $\mathbf{v}_{j\alpha} = (z - v_j)_+^{s-1} - (z - \alpha)_+^{s-1}$. Using the results of the empirical probabilities $p_i^{\mathbf{u}}$ and $p_j^{\mathbf{v}}$ obtained in (27) and (28) respectively, we have

$$\begin{aligned} ELF\left(z, \alpha\right) &= \sum_{i=1}^{m} \log\left(\frac{1}{m - \rho\left(\nu - w_i^{\mathbf{u}} \mathbf{u}_{i\alpha}\right)}\right) + \sum_{j=1}^{n} \log\left(\frac{1}{n + \rho\left(\nu - w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}\right)}\right) \\ &= \sum_{i=1}^{m} \log\left(\frac{1}{m}\right) + \sum_{i=1}^{m} \log\left(\frac{1}{1 - \frac{\rho\left(\nu - w_i^{\mathbf{u}} \mathbf{u}_{i\alpha}\right)}{m}}\right) \\ &\quad + \sum_{j=1}^{n} \log\left(\frac{1}{n}\right) + \sum_{j=1}^{n} \log\left(\frac{1}{1 + \frac{\rho\left(\nu - w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}\right)}{n}}\right). \end{aligned} \tag{75}$$

Hence, as $ELR = 2\left[ELF - ELF\left(z, \alpha\right)\right]$, we find that

$$\frac{1}{2} ELR = \sum_{i=1}^{m} \log\left(1 - \frac{\rho\left(\nu - w_i^{\mathbf{u}} \mathbf{u}_{i\alpha}\right)}{m}\right) + \sum_{j=1}^{n} \log\left(1 + \frac{\rho\left(\nu - w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}\right)}{n}\right).$$

We now look at the Lagrange multipliers. First define $\hat{\Gamma}_{1\alpha}$ and $\Gamma_{1\alpha}$ respectively as $\hat{\Gamma}_{1\alpha} = \frac{1}{m} \sum_{i=1}^{m} w_i^{\mathbf{u}} \mathbf{u}_{i\alpha}$ and $\Gamma_{1\alpha} = E\left(\hat{\Gamma}_{1\alpha}\right)$. Suppose that $r = \frac{m}{n}$ remains constant as m and n tend to infinity. Because $(w_i^{\mathbf{u}} \mathbf{u}_{i\alpha})$ for $i = 1, ..., m$ remain constant as m tends to infinity, and because $(w_i^{\mathbf{u}} \mathbf{u}_{i\alpha})$ $i = 1, ..., m$ are iid, $\hat{\Gamma}_{1\alpha}$ is a root-n consistent estimator of $\Gamma_{1\alpha}$. The same applies for $\hat{\Gamma}_{2\alpha} = \frac{1}{n} \sum_{j=1}^{n} w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}$ and $\Gamma_{2\alpha} = E\left(\hat{\Gamma}_{2\alpha}\right)$. Recall that ν is given in (29) by

$$\nu = \sum_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}. \tag{76}$$

Rewrite ν as

$$\nu = \frac{1}{n} \sum_{j=1}^{n} \left(n p_j^{\mathbf{v}}\right) \left(w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}\right). \tag{77}$$

Because the terms $\left(np_j^{\mathbf{v}}\right)$ and $\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)$ for $j=1,...,n$ remain constant as n tends to infinity, and because $\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)$ $j=1,...,n$ are iid, $\sqrt{n}$ multiplied by the quantity of the right-hand side of the above equation is of constant variance and hence is of order 1 in probability. Let $\bar{\nu}$ be the limit in probability of ν as n tends to infinity. It follows that $\nu=\bar{\nu}+O\left(n^{-1/2}\right)$. In fact and as will be clearer below, $\nu=\hat{\Gamma}_{2\alpha}+O\left(n^{-1/2}\right)$. We now turn to ρ. Because the first relation of (29) gives that

$$\sum_{i=1}^{m} p_i^{\mathbf{u}} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}=\sum_{j=1}^{n} p_j^{\mathbf{v}} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}, \tag{78}$$

this allows displaying ρ by solving for ρ in (78). Using a Taylor expansion on the values of $p_i^{\mathbf{u}}$ and $p_j^{\mathbf{v}}$, we obtain that

$$\rho=\frac{\left(\frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)-\left(\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}{\frac{1}{m^2}\sum\limits_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2-\frac{\nu}{m}\left(\frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)+\frac{1}{n^2}\sum\limits_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2-\frac{\nu}{n}\left(\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}. \tag{79}$$

Because $\left(\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2\right)_{i=1}^{m}$ and $\left(\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2\right)_{j=1}^{n}$ are iid, $p\lim\left[\frac{1}{m}\sum\limits_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2\right]=E\left[\left(w^{\mathbf{u}}\mathbf{u}_{\alpha}\right)^2\right]$ and $p\lim\left[\frac{1}{n}\sum\limits_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2\right]=E\left[\left(w^{\mathbf{v}}\mathbf{v}_{\alpha}\right)^2\right]$. Therefore, $\frac{1}{m}\sum\limits_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2$ and $\frac{1}{n}\sum\limits_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2$ are of order 1 in probability. See for example Green (2003). This gives that ρ is $O\left(n^{1/2}\right)$.

Using this, we can show that $\nu=\hat{\Gamma}_{2\alpha}+O\left(n^{-1/2}\right)$. Indeed, rewrite ν as

$$\nu=\sum_{j=1}^{n}\frac{w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}}{n+\rho\left(\nu-w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}.$$

Then,

$$\begin{aligned}\nu-\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} &= -\frac{\rho}{n}\left[\frac{1}{n}\sum_{j=1}^{n}\frac{w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\left(\nu-w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}{1+\frac{\rho}{n}\left(\nu-w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}\right]\\ &= -\frac{\rho}{n}\left[\frac{1}{n}\sum_{j=1}^{n} np_j^{\mathbf{v}} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\left(\nu-w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)\right].\end{aligned}$$

But $p\lim\left[\frac{1}{n}\sum\limits_{j=1}^{n} np_j^{\mathbf{v}} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\left(\nu-w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)\right]=\bar{\nu}^2-p\lim\left[\frac{1}{n}\sum\limits_{j=1}^{n} np_j^{\mathbf{v}}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2\right]$. The second term is of order 1 in probability, because it is defined as a weighted average of iid variables.

Hence, $\frac{\rho}{n}\left[\frac{1}{n}\sum_{j=1}^{n} np_j^{\mathbf{v}} w_j^{\mathbf{v}} \mathbf{v}_{j\alpha}\left(\nu - w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)\right]$ is $O\left(n^{-1/2}\right)$ and $\nu - \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} = O\left(n^{-1/2}\right)$. Similarly, the relation (78) allows us writing that $\nu - \frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} = O\left(n^{-1/2}\right)$. Using such relations, we can write that $\nu = \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} + O\left(n^{-1/2}\right)$ and $\nu = \frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} + O\left(n^{-1/2}\right)$. Consequently, ρ becomes

$$\rho = \frac{\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} - \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}}{\frac{1}{m^2}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{\nu^2}{m} + \frac{1}{n}\sum_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{\nu^2}{n}}. \tag{80}$$

Using again a Taylor expansion applied on the $\log$ function, we obtain that

$$\begin{aligned}
\sum_{i=1}^{m}\log\left(1 - \frac{\rho\left(\nu - w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)}{m}\right) &= \sum_{i=1}^{m}\left(-\frac{\rho\left(\nu - w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)}{m} - \frac{\rho^2\left(\nu - w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2}{2m^2}\right) + O\left(n^{-1/2}\right) \\
&= -\rho\nu + \rho\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} - \frac{\rho^2\nu^2}{2m} - \frac{\rho^2}{2m^2}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 \\
&\quad + \frac{\rho^2\nu}{m^2}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} + O\left(n^{-1/2}\right)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{n}\log\left(1 + \frac{\rho\left(\nu - w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}{n}\right) &= \sum_{j=1}^{n}\left(\frac{\rho\left(\nu - w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)}{n} - \frac{\rho^2\left(\nu - w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2}{2n^2}\right) + O\left(n^{-1/2}\right) \\
&= \rho\nu - \rho\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} - \frac{\rho^2\nu^2}{2n} - \frac{\rho^2}{2n^2}\sum_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 \\
&\quad + \frac{\rho^2\nu}{n^2}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} + O\left(n^{-1/2}\right).
\end{aligned}$$

Then the expression of ELR becomes

$$\frac{1}{2}ELR = \rho\left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} - \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right) + \frac{\rho^2}{2}\left[\frac{2\nu}{m}\left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right) - \frac{\nu^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2\right] + \frac{\rho^2}{2}\left[\frac{2\nu}{n}\left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right) - \frac{\nu^2}{n} - \frac{1}{n^2}\sum_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2\right] + O\left(n^{-1/2}\right).$$

Using the fact that $\nu = \frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} + O\left(n^{-1/2}\right)$ and $\nu = \frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha} + O\left(n^{-1/2}\right)$, this expression is equivalent to

$$\frac{1}{2}ELR = \rho\left(\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} - \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right) - \frac{\rho^2}{2}\left[\frac{1}{m^2}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{\nu^2}{m} + \frac{1}{n^2}\sum_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{\nu^2}{n}\right] + O\left(n^{-1/2}\right). \tag{81}$$

Using (80), the ELR is simply given by

$$ELR = \frac{\left(\frac{1}{m}\sum\limits_{i=1}^{m} w_i^{\mathbf{u}}\mathbf{u}_{i\alpha} - \frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2}{\left[\frac{1}{m^2}\sum\limits_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{\nu^2}{m} + \frac{1}{n^2}\sum\limits_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{\nu^2}{n}\right]} + O\left(n^{-1/2}\right).$$

Hence, dividing the numerator and the denominator by $\left(\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\right)^2$ allows writing ELR as

$$ELR = \frac{\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)^2}{\left[\frac{1}{m^2}\sum\limits_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{1}{m}\nu^2 + \frac{1}{n^2}\sum\limits_{j=1}^{n}\left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{1}{n}\nu^2\right] \Big/ \left(\frac{1}{n}\sum\limits_{j=1}^{n} w_j^{\mathbf{v}}\right)^2} + O\left(n^{-1/2}\right). \tag{82}$$

This last expression is equivalent to

$$ELR = \frac{n\left(\hat{P}^s_G(z) - \hat{P}^s_{F_\alpha}(z)\right)^2}{n\left[\frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{1}{m}\nu^2 + \frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{1}{n}\nu^2\right] \Big/ \left(\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}\right)^2} + O\left(n^{-1/2}\right). \tag{83}$$

Denote $\mathbf{u}_i = (z-u_i)_+^{s-1}$, $\mathbf{u}_\alpha = (z-\alpha)_+^{s-1}$ and $\mathbf{v}_j = (z-v_j)_+^{s-1}$. Then

$$\begin{aligned}
\frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{1}{m}\nu^2 &= \frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 - \frac{1}{m}\left(\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}\mathbf{u}_{i\alpha}\right)^2 + O\left(n^{-1/2}\right) \\
&= \frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\mathbf{u}_i\right)^2 - \frac{1}{m}\left(\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}\mathbf{u}_i\right)^2 \\
&\quad -2\mathbf{u}_\alpha\left(\frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\right)^2\mathbf{u}_i - \frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\mathbf{u}_i\right)\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}\right) \\
&\quad +\mathbf{u}_\alpha^2\left[\frac{1}{m^2}\sum_{i=1}^m \left(w_i^{\mathbf{u}}\right)^2 - \frac{1}{m}\left(\frac{1}{m}\sum_{i=1}^m w_i^{\mathbf{u}}\right)^2\right] + O\left(n^{-1/2}\right).
\end{aligned}$$

The same thing applies for $\left(\frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{1}{n}\nu^2\right)$ and we obtain that

$$\begin{aligned}
\frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\mathbf{v}_{j\alpha}\right)^2 - \frac{1}{n}\nu^2 &= \frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\mathbf{v}_j\right)^2 - \frac{1}{n}\left(\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}\mathbf{v}_j\right)^2 \\
&\quad -2\mathbf{u}_\alpha\left(\frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\right)^2\mathbf{v}_j - \frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\mathbf{v}_j\right)\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}\right) \\
&\quad +\mathbf{u}_\alpha^2\left[\frac{1}{n^2}\sum_{j=1}^n \left(w_j^{\mathbf{v}}\right)^2 - \frac{1}{n}\left(\frac{1}{n}\sum_{j=1}^n w_j^{\mathbf{v}}\right)^2\right] + O\left(n^{-1/2}\right).
\end{aligned}$$

Therefore the dominator of ELR is simply the following expression:

$$\frac{n}{\left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right)^2}\left(\widehat{\text{var}}\left[\frac{1}{m}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_i\right)\right] - 2\mathbf{u}_\alpha\widehat{\text{cov}}\left[\frac{1}{m}\sum_{i=1}^{m}\left(w_i^{\mathbf{u}}\mathbf{u}_i\right), \frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\right] + \mathbf{u}_\alpha^2\widehat{\text{var}}\left[\frac{1}{m}\sum_{i=1}^{m} w_i^{\mathbf{u}}\right]\right)$$
$$+\frac{n}{\left(\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right)^2}\left(\widehat{\text{var}}\left[\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_j\right] - 2\mathbf{u}_\alpha\widehat{\text{cov}}\left[\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\mathbf{v}_j, \frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right] + \mathbf{u}_\alpha^2\widehat{\text{var}}\left[\frac{1}{n}\sum_{j=1}^{n} w_j^{\mathbf{v}}\right]\right)$$
$$+O\left(n^{-1/2}\right).$$

Using the notation on pages 38 and 44, this expression can be rewritten as

$$\frac{n}{\Gamma_4^2}\left[\widehat{\text{var}}\left(\hat{\Gamma}_1\right) - 2\left(z-\alpha\right)_+^{s-1}\widehat{\text{cov}}\left(\hat{\Gamma}_1, \hat{\Gamma}_2\right) + \left[\left(z-\alpha\right)_+^{s-1}\right]^2\widehat{\text{var}}\left(\hat{\Gamma}_2\right)\right]$$
$$+\frac{n}{\Gamma_4^2}\left[\widehat{\text{var}}\left(\hat{\Gamma}_3\right) - 2\left(z-\alpha\right)_+^{s-1}\widehat{\text{cov}}\left(\hat{\Gamma}_3, \hat{\Gamma}_4\right) + \left[\left(z-\alpha\right)_+^{s-1}\right]^2\widehat{\text{var}}\left(\hat{\Gamma}_4\right)\right] + O\left(n^{-1/2}\right).$$

Notice that the above expression is exactly the estimator of the variance of $n^{1/2}\left(\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)\right)$ when using the condition that $P_G^s(z) = P_{F_\alpha}^s(z)$. Hence, as we can see, ELR coincides asymptotically with $t^2\left(z,\alpha\right)$ given that $\Delta P^s\left(z,\alpha\right) = \underset{m,\, n\to\infty}{p\lim}\sqrt{n}\left(\hat{P}_G^s(z) - \hat{P}_{F_\alpha}^s(z)\right) = O\left(1\right)$. ■

References

ABOUDI, R., D. THON, AND S. WALLACE (2010): “Inequality comparisons when the populations differ in size,” *Journal of Economic Inequality*, 8, 47–70.

BARRETT, G. AND S. DONALD (2003): “Consistent Tests for Stochastic Dominance,” *Econometrica*, 71, 71–104.

BATANA, Y. M. (2008): “Comparaisons Multidimensionnelles de Bien-être et de Pauvreté: Méthodes, Inférence et Applications,” Ph.D. Thesis, Université Laval.

BIBI, S. AND J.-Y. DUCLOS (2009): “La pauvreté au Québec et au Canada,” Working Paper 09-22, CIRPEE.

BISHOP, J., J. FORMBY, AND P. THISTLE (1992): “Convergence of the South and Non-South Income Distributions, 1969-1979,” *American Economic Review*, 82, 262–72.

BLACKORBY, C., W. BOSSERT, AND D. DONALDSON (1996): “Quasi Ordering and Population Ethics,” *Social Choice and Welfare*, 13, 129–150.

——— (2005): *Populations Issues in Social Choice Theory, Welfare Economics and Ethics*, New York: Cambridge University Press.

BLACKORBY, C. AND D. DONALDSON (1984): "Social Criteria for Evaluating Population Change," *Journal of Public Economics*, 25, 13–33.

BROOME, J. (1992a): *Counting the Cost of Global Warming*, White Horse, Cambridge.

——— (1992b): "The Value of Living," *Recherches Economiques de Louvain*, 58, 125–142.

CHEN, W.-H. (2008): "Comparing Low Income of Canada's Regions: A Stochastic Dominance Approach," Tech. Rep. 06, Statistics Canada, Ottawa.

CHEN, W.-H. AND J.-Y. DUCLOS (2008): "Testing for Poverty Dominance: An Application to Canada," Tech. Rep. 3829, IZA Discussion Paper.

COWEN, T. (1989): "Normative Population Theory," *Social Choice and Welfare*, 6, 33–43.

DAVIDSON, R. (2009): "Testing for Restricted Stochastic Dominance: Some Further Results," *Review of Economic Analysis*, 1, 34–59.

DAVIDSON, R. AND J.-Y. DUCLOS (2000): "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68, 1435–64.

——— (2006): "Testing for Restricted Stochastic Dominance," IZA Discussion Paper No 2047, IZA.

DUCLOS, J.-Y. AND A. ZABSONRÉ (2009): "Welfare comparisons with populations of different sizes: a theoretical analysis," mimeo, Université Laval.

FOSTER, J., J. GREER, AND E. THORBECKE (1984): "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761–776.

GRAVEL, N., T. MARCHANT, AND A. SEN (2008): "Comparing societies with different numbers of individuals: A characterization of the Average Utilitarian criterion," Working paper.

GREEN, W. H. (2003): *Econometric Analysis*, New York University: Upper Saddle River, New Jersey.

HOWES, S. (1993): "Asymptotic Properties of Four Fundamental Curves of Distributional Analysis," mimeo, STICERD, London School of Economics.

HUMAN RESOURCES AND SOCIAL DEVELOPMENT CANADA (2006): "Low Income in Canada: 2000-2002 using the Market Basket Measure," Tech. Rep. SP-628-05-06E, Ottawa.

KANBUR, R. AND D. MUKHERJEE (2007): "Premature Mortality and Poverty Measurement," *Bulletin of Economic Research*, 59, 0307–3378.

KAUR, A., B. L. S. PRAKASA RAO, AND H. SINGH (1994): "Testing for Second-Order Stochastic Dominance of Two Distributions," *Econometric Theory*, 10, 849–66.

LEFRANC, A., N. PISTOLESI, AND A. TRANNOY (2006): "Equality of opportunity: Definitions and testable conditions, with an application to income in France," IDEP Working Papers 0609, Institut d'Économie Publique.

MUKHERJEE, D. (2008): "Poverty measures incorporating variable rate of alleviation due to population growth," *Social Choice and Welfare*, 31, 97–107.

OWEN, A. B. (1988): "Empirical Likelihood Ratio Confidence Intervals for a Single Functional," *Biometrika*, 75, 237–249.

——— (2001): *Empirical Likelihood*, Chapman and Hall/CRC.

PARFIT, D. (1984): *Reasons and Persons*, Oxford: Oxford University Press.

POGGE, T. (2007): *John Rawls; his life and theory of justice*, Oxford University Press.

SHORROCKS, A. (1987): "Transfer Sensitive Inequality Measures," *Review of Economic Studies*, LIV, 485–497.

TRANNOY, A. AND J. A. WEYMARK (2009): "Dominance for Critical-Level Generalized Utilitarianism," in *Arguments for a Better World: Essays in Honor of Amartya Sen, Volume I: Ethics, Welfare, and Measurement*, Oxford University Press: Kaushik, B. and Kanbur, R., 262–281.