

Similarity, dissimilarity and exceptionality: generalizing Gini's transvariation to measure "differentness" in many distributions

Gordon Anderson¹ · Oliver Linton² · Jasmin Thomas¹

Received: 1 March 2017 / Accepted: 15 June 2017 / Published online: 20 July 2017
© Sapienza Università di Roma 2017

Abstract Following the work of Gini, Dagum and Tukey, this paper extends Gini's Transvariation measure for comparing two distributions to the simultaneous comparison of many distributions. In so doing, it develops measures of absolute and relative similarity, dissimilarity and exceptionality together with techniques for assessing particular aspects of variations across those distributions. These techniques are exemplified in a study of differences between the income distributions of males and females drawn from Metis, Inuit, North American Indian and Non-Aboriginal constituencies in Canada in the first decade of the twenty-first century. While the distributions were becoming increasingly similar (interpreted as improving equality of opportunity), this was occurring primarily at the center of the distribution. At the extremes, the distributions were diverging, suggesting that such improvements in equality of opportunity were not for all.

Keywords Relative transvariation · Dissimilarity · Exceptionality

1 Introduction

The similarity or dissimilarity of a collection of $K (>2)$ distributions, $f(x)$, $g(x)$, $h(x) \dots$, of the random vector x (where x could be a scalar or vector) is of interest in many fields. Wherever the degree of heterogeneity of types is an issue, measures and tests of the degree of similarity–dissimilarity would be of great use especially in the era of big data. Notions of dissimilarity in many distributions are closely related to Gini's Transvariation Measure for two discrete distributions [34,35], which was generalized to many discrete distributions by Dagum [14, 15]. In economics this interest ranges from concerns regarding the similarity–dissimilarity of the outcome distributions of agents in different circumstance classes in the equality of opportunity and social justice literature to the similarity–dissimilarity of technologies in the

✉ Gordon Anderson
anderson@chass.utoronto.ca

¹ Economics Department, University of Toronto, Toronto, Canada

² Economics Department, Cambridge University, Cambridge, UK

empirical growth literature to the commonality of valuation distributions in first price auction theory literature. In economics and sociology literatures such measures would have relevance for social classification [10, 12], social mobility [36] and segregation [7, 13, 18, 19, 21, 27, 30, 32, 43]. In the political science literature on polarization [33], conflict, and diversity [22–24, 51] the subject matter has much to do with increasing or diminishing dissimilarities of the latent subgroup distributions of an overall population distribution [20].

Perception of similarities and differences from particular viewpoints is also of interest. In the population ethics literature [8, 11, 16, 17] versions of the constituency principle¹ argue that only the views of a particular population or constituency (essentially those affected by the comparison) are relevant in assessing goodness (or badness) of states. Underlying this principle is the notion that “better-ness” from the point of view of one population may differ from the corresponding view of another, apparently unaffected, population. Discourse in this literature was pursued in the context of successive populations, but similar ideas can be explored in the context of contemporaneous populations or constituencies. For example, it is this doctrine that underlays the Focus Axiom in poverty analysis wherein only the incomes of the poor constituency should matter in the poverty calculation. As a principle, it is strong in the sense that only the views of one constituency are of account, the views of all other constituencies have zero weight in the calculus. Weaker versions of the constituency principle would see primacy given to the views of the affected population with lower but non-zero weight attached to other constituencies. Extending the poverty measurement analogy, higher orders of the Foster et al. [29] family of poverty measures attach increasing weight to constituencies further below the poverty line. This in turn leads to the idea of intensifying the focus, or magnifying in some sense, particular aspects of the distribution of interest to those constituencies with primacy.

Here these ideas will be explored and extended to comparisons of continuous distributions, to the idea of relative dissimilarity and to the idea of comparisons of dissimilarity of particular features of distributions all of which could potentially be multivariate. In Sect. 2 the concepts will be developed in the context of Gini’s original two-distribution world. Various extensions to the many distribution case are considered in Sect. 3. The comparison techniques will be exemplified in a study examining differences by gender in Aboriginal and non-Aboriginal constituency income distribution in Canada in the twenty first century in Sect. 4. Section 5 concludes.

2 Dissimilarity in a two-constituency world

Gini’s Transvariation Measure, “*GINITD*”, for comparing two discrete distributions f_i and g_i defined on i outcomes $i = 1, \dots, I$ where f_i, g_i are respectively the probabilities of outcome “ i ” occurring is given by:

$$GINITD = \frac{1}{2} \sum_{i=1}^I |f_i - g_i|$$

It can be shown that *GINITD* is bounded by 0 and 1. In addition, when f and g have mutually exclusive support, *GINITD* = 1 (since $\sum_i |f_i - g_i| = \sum_i f_i + \sum_i g_i = 2$), and when f and g are identical *GINITD* = 0. A convenient way of writing *GINITD* in what follows is:

¹ The constituency principle avers that the goodness of alternative states should be judged from the point of view of some identified constituency of individuals who alone are judged to be the relevant and interested parties to the outcome of the comparison exercise.

$$GINITD = \frac{1}{2} \sum_{i=1}^I |f_i - g_i| = \frac{1}{2} \sum_{i=1}^I (\max(f_i, g_i) - \min(f_i, g_i)) = \frac{1}{2}(\theta_U - \theta_L),$$

where $\theta_U = \sum_{i=1}^I \max(f_i, g_i)$ and $\theta_L = \sum_{i=1}^I \min(f_i, g_i)$.

Noting that $\theta_U = 2 - \theta_L$ this can be rewritten as:

$$GINITD = 1 - \theta_L \tag{1}$$

The continuous version of Gini’s Transvariation Measure, *GINITC*, is similarly related to the integral of absolute differences between two PDFs, $f(x)$ and $g(x)$, each defined on the real line and is similarly related to the overlap measure $\theta_L (= \int \min\{f(x), g(x)\}dx)$:

$$GINITC = 0.5 \int_{-\infty}^{\infty} |f(x) - g(x)| dx = 1 - \theta_L \tag{1a}$$

In both continuous and discrete cases θ can be shown to be asymptotically normally distributed [2,6] and, since $\theta \sim_a N(\theta, V)$, then both *GINITD* and *GINITC* $\sim_a N(1 - \theta, V)$ thus facilitating inference. The distribution and properties of the latter parameter, θ , which is a measure of the degree of overlap or commonality between the distributions, have been provided for the two distribution multivariate case in Anderson et al. [2]. In essence, when Kernel estimates of $f(x)$ and $g(x)$ are employed it is shown that:

$$\sqrt{n}(\hat{\theta}_L - \theta_L) - a_n \sim N(0, v)$$

Thus from the relationship between *GINIT* and θ_L in Eq. (1) above it follows that:

$$\sqrt{n}(\widehat{GINIT} - GINIT) + a_n \sim N(0, v)$$

Here a_n , a bias correction term and v , the variance are dependent on the “contact set”, the set of points where $f(x) = g(x)$, when this set is empty or of measure 0 the bias term is 0 and v is given by:

$$v = h(n)\{p_f(1 - p_f) + p_g(1 - p_g) + (p_{f,g} - p_f p_g)\},$$

where p_f is the estimated probability that $f(x) - g(x) < -c$, p_g is the estimated probability that $f(x) - g(x) > c$ and $p_{f,g}$ is the estimated probability that $-c < f(x) - g(x) < c$ for some tuning parameter c . (Note the last term is omitted if the random variables under f and g are not jointly distributed).

A convenient way of thinking about this comparison is to note that it can be written as:

$$\begin{aligned} GINITC &= 0.5 \int_{-\infty}^{\infty} |f(x) - g(x)| dx = 0.5 \int_{-\infty}^{\infty} \left| \frac{f(x)}{g(x)} - 1 \right| g(x) dx \\ &= 0.5 E_{g(x)} \left(\left| \frac{f(x)}{g(x)} - 1 \right| \right) \end{aligned}$$

similarly:

$$\begin{aligned} GINITD &= 0.5 \sum_{i=1}^I \left| \frac{f_i(x)}{g_i(x)} - 1 \right| g_i(x) \\ &= 0.5 E_{g(x)} \left(\left| \frac{f(x)}{g(x)} - 1 \right| \right) \end{aligned} \tag{2}$$

GINIT then has the interpretation of half of the first moment or expected value of $\left| \frac{f(x)}{g(x)} - 1 \right|$ under distribution $g(x)$.² In essence it is a first moment comparator and the sample average provides an immediate estimator with a method of moments interpretation.

2.1 Importance weighted comparisons in a two distribution world

Often interest will focus on particular aspects of the differences between $f(x)$ and $g(x)$ in various regions of x , for example differences at the extremes of a reference distribution or differences at the center of a reference distribution. Alternatively, interest may focus on the differences between $f(x)$ and $g(x)$ in high frequency or low frequency regions of x according to the reference or target distribution. For convenience suppose the reference distribution $h(x)$ has positive support over the whole domain of x then, for some chosen value λ , importance weighted versions of *GINITD* and *GINITC* may be written as:

$$\begin{aligned}
 GINITCI &= \int_{-\infty}^{\infty} |f(x) - g(x)| \frac{h(x)^\lambda}{E_{h(x)}(h(x)^\lambda)} dx \\
 &= E_{g(x)} \left(\left| \frac{f(x)}{g(x)} - 1 \right| \frac{h(x)^\lambda}{E_{h(x)}(h(x)^\lambda)} \right) \\
 &= \frac{1}{E_{h(x)}(h(x)^\lambda)} E_{g(x)} \left(\left| \frac{f(x)}{g(x)} - 1 \right| h(x)^\lambda \right) \\
 GINITDI &= \sum_{i=1}^I |f_i - g_i| \frac{h_i^\lambda}{E_{h(x)}(h_i^\lambda)} \\
 &= E_{g(x)} \left| \frac{f_i}{g_i} - 1 \right| \frac{h_i^\lambda}{E_{h(x)}(h_i^\lambda)} \\
 &= \frac{1}{E_{h(x)}(h_i^\lambda)} E_{g(x)} \left(\left| \frac{f_i}{g_i} - 1 \right| h_i^\lambda \right)
 \end{aligned}$$

Whereas the Gini Transvariation Measure can be seen as cumulating the absolute difference between the functions over all possible outcomes, the importance weighted versions can be seen as cumulating the ‘‘importance’’ weighted absolute difference between the distributions.³ Generally $h(x)$ will be related to $g(x)$ or $f(x)$ and, when $h(x)$ is a regular distribution or a cumulative distribution, the weighting function $h^*(x) = \frac{h(x)^\lambda}{E_{h(x)}(h(x)^\lambda)}$ when applied to the regular distribution $f(x)$ can be seen to transform it to another distribution $f^*(x) = h^*(x)f(x)$ which retains all of the properties of a regular probability distribution

² Note also the close relationship to Hellinger Distance [39,49]:

$$\begin{aligned}
 H^2(f, g) &= \frac{1}{2} \int \left(\sqrt{f(x)} - \sqrt{g(x)} \right)^2 dx = 1 - \int \sqrt{f(x)g(x)} dx \\
 &= 1 - \int \sqrt{\frac{f(x)}{g(x)}} g(x) dx = \int \left(1 - \sqrt{\frac{f(x)}{g(x)}} \right) g(x) dx \\
 &= E_{g(x)} \left(1 - \sqrt{\frac{f(x)}{g(x)}} \right)
 \end{aligned}$$

³ This should not be confused with standardizing by a location measure, see for example [28].

($f^*(x) \geq 0, 0 \leq F^*(x) \leq 1$) with aspects of $f(x)$ amplified or muted according to the choice of $h(x)$ and λ . Thus the importance weighted version may be written as in Eq. (2) in terms of transformed distributions, retaining all of the properties of that measure (boundedness between 0 and 1, association with the overlap measure including its asymptotic normality, etc.). As a result, the Importance Weighted Gini Transvariation Measure may be written as:

$$\begin{aligned} GINITCI &= 0.5 \int_{-\infty}^{\infty} |f^*(x) - g^*(x)| dx = 0.5 \int_{-\infty}^{\infty} \left| \frac{f^*(x)}{g^*(x)} - 1 \right| g^*(x) dx \\ &= 0.5 E_{g^*(x)} \left(\left| \frac{f^*(x)}{g^*(x)} - 1 \right| \right) \end{aligned}$$

similarly:

$$GINITDI = 0.5 \sum_{i=1}^I \left| \frac{f_i^*(x)}{g_i^*(x)} - 1 \right| g_i^*(x) = 0.5 E_{g^*(x)} \left(\left| \frac{f^*(x)}{g^*(x)} - 1 \right| \right) \tag{2a}$$

In the following $\lambda = -0.5$ is considered,⁴ a choice inspired by, and in the spirit of, entropic measures of variation such as in [31,46,52,56]. For example, by letting $h(x) = g(x)$, the continuous version of Thiel’s Entropic Measure (TE) can be related to the continuous version of Pearson’s Chi Squared Dissimilarity Measure ($PCHI$) as follows:⁵

$$\begin{aligned} TE &= \int f(x) \ln \left(\frac{f(x)}{g(x)} \right) dx = \int f(x) \ln \left(\frac{f(x) + g(x) - g(x)}{g(x)} \right) dx \\ &= \int f(x) \ln \left(1 + \frac{f(x) - g(x)}{g(x)} \right) dx \approx \int f(x) \left(\frac{f(x) - g(x)}{g(x)} \right) dx \\ &= \int \left(\frac{(f(x) - g(x))^2}{g(x)} \right) dx = PCHI \end{aligned}$$

Note that the argument under the integral sign in $GINITCI$ can be seen to be the square root of the argument under the integral sign of $PCHI$. Intuition for Pearson-type measures may be gleaned from noting that:

$$\begin{aligned} PCHI &= \int \left(\frac{(f(x) - g(x))^2}{g(x)} \right) dx = \int \left(\frac{(f(x) - g(x))^2}{(g(x))^2} \right) g(x) dx \\ &= \int \left(\frac{(f(x))^2 + (g(x))^2 - 2f(x)g(x)}{(g(x))^2} \right) g(x) dx \\ &= \int \left(\frac{(f(x))^2}{(g(x))^2} + 1 - 2 \frac{f(x)}{g(x)} \right) g(x) dx \\ &= \int \left(\frac{f(x)}{g(x)} - 1 \right)^2 g(x) dx \\ &= E_{g(x)} \left[\left(\frac{f(x)}{g(x)} - 1 \right)^2 \right] \end{aligned} \tag{3}$$

⁴ It is interesting to note that Tukey [57] proposed visualizing distributions in terms of a “Rootgram” since it increased the visual importance of low frequency outcomes and muted the importance of high frequency outcomes. Thus rescaling can be seen as viewing distributions relative to the Rootgram of the target distribution standardized to make it conformable with a regular probability distribution. See also [37].

⁵ A similar algebra can readily demonstrate the relationship in the discrete paradigm.

PCHI in turn has the interpretation of the expected value under $g(x)$ of $\left| \frac{f(x)}{g(x)} - 1 \right|^2$ yielding an interpretation of the statistic as a second moment measure of $\left| \frac{f(x)}{g(x)} - 1 \right|$ with the sample average of $\left| \frac{f(x)}{g(x)} - 1 \right|^2$ providing an estimator of the dissimilarity parameter with a method of moments interpretation.⁶

Importance weighting can be seen as weighting the differences between $f(x)$ and $g(x)$ by some monotonic function of the “target” function so that for $\lambda < 0$, a given difference from a small target plays a bigger role in the calculation than the same order of difference in a correspondingly larger target (for $\lambda > 0$ the reverse is true). In essence the differences are weighted with respect to some reference distribution (in this case h) and the statistic can thus be considered a “relative” dissimilarity measure where differences are measured relative to the target measure $h(x)$.

Choice of the target distribution is a matter of some consequence, some special cases are obvious. When the target measure $h(x)$ is set equal to either $g(x)$ or $f(x)$, *PCHI* has the form of a classic Pearson goodness-of-fit test of distributional form or independence where the target distribution is the null hypothesized distribution. Suppose g and f are empirical distributions, such that $f(x)$ first-order stochastically dominates $g(x)$ then measures with a target $g(x)$ are dissimilarity measures relative to the poor distribution and measures with a target $f(x)$ are dissimilarity measures relative to the non-poor distribution. When $h(x)$ is the sample size weighted sum of $f(x)$ and $g(x)$, the statistic has the form of a goodness-of-fit test for homogeneity of distributions, in essence f and g are being compared to some “average” distribution. Generally, when the distributions f and g are not identical, the dissimilarity of the distributions $f(x)$ and $g(x)$ in terms of $f(x)$ will not be the same as the dissimilarity of those distributions in terms of $g(x)$.⁷

3 Indices of dissimilarity in a many distribution world

These ideas can be extended to the comparison of many distributions in either continuous or discrete paradigms. (Intuition suggests that the above statistics will work for parametric, semi parametric and non-parametric representations of the distributions.)⁸ Consider first continuous scalar x with continuous support on the interval $\{a, b\}$ and discrete scalar x with support on the set of non-negative integers with a set U_K of K distributions $f(x), g(x), h(x), \dots$ under consideration:

For x continuous:

$$DISC = \frac{\int_a^b \max(f(x), g(x), h(x), \dots) dx - \int_a^b \min(f(x), g(x), h(x), \dots) dx}{K}$$

⁶ Yalonetsky [60] in comparing the Overlap with the Pearson Measure can be construed as comparing the performance of the two moment estimators.

⁷ For this to be the case $f(x)$ and $g(x)$ have to be such that:

$$E_g \left(\left(\frac{f(x)}{g(x)} \right)^2 \right) - 2E_g \left(\left(\frac{f(x)}{g(x)} \right) \right) = E_f \left(\left(\frac{g(x)}{f(x)} \right)^2 \right) - 2E_f \left(\left(\frac{g(x)}{f(x)} \right) \right)$$

A sufficient condition for this is that $f(x)$ and $g(x)$ should be reflective of each other around some point x^* so that $f(x^* + \delta) = g(x^* - \delta)$ for all δ since then for every $\frac{f(x_1)}{g(x_1)}$ there will be a corresponding $\frac{g(x_2)}{f(x_2)}$ of equal value with the same importance weight in the calculation.

⁸ Dagum [14, 15] first mooted the idea of extending Gini’s Transvariation to many distributions in a discretized paradigm.

For x discrete:

$$DIS_D = \frac{\sum_{x=0}^{\infty} \max(f(x), g(x), h(x), \dots) - \sum_{x=0}^{\infty} \min(f(x), g(x), h(x), \dots)}{K}$$

Note that $0 \leq DIS \leq 1$. In particular, for complete dissimilarity (perfect segmentation in the terminology of Ref. [61]) $DIS = 1$, and for complete similarity (all distributions identical) $DIS = 0$. This index may be construed as the distributional analogue of a range statistic for a variable measuring as it does the cumulative distances between the upper and lower extremes of the collection of distributions over the values of x . The asymptotic distribution of Eq. (3) and bootstrapping techniques for standard errors are developed in [5].

Conceptually these indices depend upon two components θ_U and θ_L given by:

$$\begin{aligned} \theta_U &= \int_a^b \max(f(x), g(x), h(x), \dots) dx \\ &\left\{ = \sum_{x=0}^{\infty} \max(f(x), g(x), h(x), \dots) \text{ for } x \text{ discrete} \right\} \\ \theta_L &= \int_a^b \min(f(x), g(x), h(x), \dots) dx \\ &\left\{ = \sum_{x=0}^{\infty} \min(f(x), g(x), h(x), \dots) \text{ for } x \text{ discrete} \right\} \end{aligned} \tag{4}$$

When dissimilarity is complete (perfect segmentation of all distributions) $\theta_U = K$ and $\theta_L = 0$. When similarity is complete (perfect overlap of all distributions) $\theta_U = 1$ and $\theta_L = 1$. When U is the set of outcome distributions for K inheritance classes, $M = 1 - DIS$ can be interpreted as a mobility index with $M = 1$ representing perfect mobility (outcome distributions identical for all inheritance classes) and $M = 0$ representing complete immobility (inheritance class outcome distributions have no common points of support). When U is the set of outcome distributions for K social classes whose alienation is ordinal measured on x , DIS can be considered a polarization measure. If dissimilarity from a base or target distribution, for convenience refer to it as $g(x)$, is of interest, following Tukey [57] for interpretational purposes, one could contemplate:

For x continuous:

$$DIS_C = \frac{\int_a^b \max(f(x), g(x), h(x), \dots) g(x)^{-0.5} dx - \int_a^b \min(f(x), g(x), h(x), \dots) g(x)^{-0.5} dx}{K}$$

For x discrete:

$$DIS_D = \frac{\sum_{x=0}^{\infty} \max(f(x), g(x), h(x), \dots) g(x)^{-0.5} - \sum_{x=0}^{\infty} \min(f(x), g(x), h(x), \dots) g(x)^{-0.5}}{K} \tag{5}$$

So for example $g(x)$ could be a weighted average of all the other distributions i.e. a mixture where the weights reflected proportions of the population under the corresponding distribution (commonly used in the segregation literature (see for example [43] and references therein). Alternatively, $g(x)$ could be the “poorest” or “richest” distribution in a collection of distributions or the “median” distribution. In effect “inequality in distribution” or degrees

of difference from a distribution $g(x)$ ⁹ is being measured where $g(x)$ would be the basis of the “importance” weighting function $g^*(x)$ as in $h^*(x)$ above.

These may be compared with the “many distribution” analogue to the Pearson goodness-of-fit test [1] which for K distributions $f_k(x)$ $k = 1, \dots, K$ may be written in the continuous case as:

$$PT = \frac{1}{K - 1} \sum_{k=1}^K \int \frac{(f_k(x) - g(x))^2}{g(x)} dx,$$

where $g(x)$, the target distribution, is usually of the form:

$$g(x) = \sum_{k=1}^K w_k f_k(x) \quad \text{where} \quad \sum_{k=1}^K w_k = 1$$

Following Eq. (3) above this multivariate Pearson Test has a similar Method of Moments interpretation:¹⁰

$$\frac{1}{K - 1} \sum_{k=1}^K E_{g(x)} \left[\left(\frac{f_k(x)}{g(x)} - 1 \right)^2 \right].$$

3.1 Specific dominance based indices

Stochastic dominance tests have been employed in a wide variety of situations, particularly in the empirical wellbeing and finance literatures, where there is a specific theoretical rationale for comparing distributions at a particular order of dominance in order to see how different they are in that particular aspect. In empirical welfare analysis different orders of dominance correspond to specific types of wellbeing measure. In finance different orders of dominance correspond to specific types of risk class.

The condition for the dominance of $G()$ by $F()$ at order J is given by:

$$F_J(x) \leq G_J(x) \quad \forall x \in [a, b] \quad \text{and} \quad F_J(x) < G_J(x) \quad \text{for some } x \in [a, b]$$

where

$$F_i(x) = \int_a^x F_{i-1}(z) dz \quad \text{and} \quad F_0(x) = f(x) \quad (\text{similarly for } G_i(x)) \quad (6)$$

Essentially the condition requires that the functions $F_J(x)$ and $G_J(x)$ do not cross so that one, the dominating distribution, is “unambiguously” below the other.¹¹

A variety of tests have been proposed for these dominance conditions at various orders. It may readily be seen that successive orders of dominance attach increasing weight/importance to lower values of x . If concern was with differences at high values of x one would work with the condition for the dominance of $G()$ by $F()$ at order J given by:¹²

$$F_J(x) \leq G_J(x) \quad \forall x \in [a, b] \quad \text{and} \quad F_J(x) < G_J(x) \quad \text{for some } x \in [a, b],$$

⁹ For example, in the equality of opportunity literature, concern is for inequality of outcome distributions across inheritance classes where the concern may be differences from the median inheritance class outcome distribution or the richest or poorest inheritance class distributions.

¹⁰ All of the above indices can be shown to satisfy piecewise continuity, scale invariance, scale independence and normalization axioms for inequality indices. See for example [42, 54, 55].

¹¹ Also related here is the Integrated Squared Difference approach [38].

¹² This is the condition used in the finance literature for risk loving behavior.

where

$$F_i(x) = \int_x^b F_{i-1}(z)dz \quad \text{and} \quad F_0(x) = f(x) \quad (\text{similarly for } G_i(x)) \quad (7)$$

Essentially the condition requires that the functions $F_J(x)$ and $G_J(x)$ do not cross so that one, the dominating distribution, is “unambiguously” below the other.

For the indices of similarity for the collection of k distributions of the variable x , $f(x), g(x), h(x), \dots$ at the J th order we construct $U_J(x) = \max_x (F_J(x), G_J(x), H_J(x), \dots)$ and $L_J(x) = \min_x (F_J(x), G_J(x), H_J(x), \dots)$. The maximum range of variation of the collection of distributions at the J th order is given by the J th order distance measure:

$$DISS(J) = \int_a^b (U_J(x) - L_J(x))dx$$

To provide a complete ordering of a collection of distributions at a particular order one could also contemplate “proximity to the boundary” indices (see [4,5]) for a particular distribution say $f(x)$ of the form:

$$PROXU_J(F_J(x)) = 1 - \frac{\int_a^b (U_J(x) - F_J(x))dx}{DISS(J)}$$

$$PROXL_J(F_J(x)) = 1 - \frac{\int_a^b (F_J(x) - L_J(x))dx}{DISS(J)},$$

where $PROXU_J(F_J(x)) \{PROXL_J(F_J(x))\}$ measures proximity to the upper {lower} boundary.

3.2 Distribution separation measures

The degree of separation between the classes reflecting increased within class homogeneity or better identified group classification in a polarization sense is of interest. To fix ideas, suppose class distributions are ordered by some location measure, then one minus the average overlap of contiguous classes is a useful measure, as is the overlap of distributions at the extremes. These measures are closely associated with the measures of polarization in Ref. [20],¹³ and the segmentation terminology of Ref. [61] in the following sense.

When the k th pair of contiguous lower and upper class distributions, $f_L(x)$ and $f_U(x)$ respectively, do not overlap (i.e. $\theta_{LUk} = \int \min(f_L(x), f_U(x))dx = 0$), they are perfectly segmented in the terminology of Ref. [61]. When they overlap perfectly (i.e. $\theta_{LUk} = \int \min(f_L(x), f_U(x))dx = 1$) separate distributions cannot be identified. Thus $1 - \theta_{LUk}$ provides a measure of the polarization or distance between two contiguous groups and $1 - \theta = K^{-1} \sum_k (1 - \theta_{LUk})$, the average over all pairs of contiguous groups, provides an index of polarity in the collection of groups. When all pairs are perfectly segmented the average of these overlaps will be 0 and $1 - \text{average } \theta$ will be 1. When all classes overlap perfectly and cannot be separately identified it will be 0. In contrast the extent of overlap of the extreme distributions yields a lower bound to the average overlap, reflects the extent to which the extremes have polarized, and provides a measure of potential for polarization in the collection of classes.

¹³ These measures can be shown to be the average polarization distance between all pairs of observations [3].

3.3 “Leave one out” measures and tests of exceptionality

Where concern regarding exceptionality (the distinct difference of one distribution from a collection of others) is of interest a “leave one out test” i.e. a change in the dissimilarity or overlap statistic is likely to have particular power. If concern is about differences in the lower regions of the distributions, higher orders of dominance comparators are the relevant instrument since they implicitly weight lower components of the indices more heavily. If concern is about differences in the higher regions of the distributions, higher orders of dominance comparators for the counter-cumulative densities are the relevant instrument since they implicitly weight higher components of the indices more heavily.

In general, when some of the groups are segmented or non-overlapping in distribution with all other groups they may be thought to be exceptional. Leaving them out of the calculus, but not reducing K , would reduce DIS by at most the proportion of the number of segmented groups. If the segmented groups were first order dominated by and first order dominated other distributions in the collection, that is they were internal to the collection, they would be unexceptional and their omission would have no effect on DIS (if K was not reduced). In general, for exceptionality measures, interest focuses on $\frac{DIS - DIS_{OK_1}}{K}$ where DIS_O relates to the many group dissimilarity measure with $K - K_1$ classes omitted. If the omitted classes were perfectly segmented, the biggest contribution to the transvariation of K distributions that the $K - K_1$ constituencies could make would be $\frac{K - K_1}{K}$, thus $DIF_O = \frac{K \left(\frac{DIS - DIS_{OK_1}}{K} \right)}{K - K_1}$ provides a $[0, 1]$ bounded measure of the degree of segmentation of a group of $K - K_1$ subgroups in a K constituency comparison.

4 An application: comparing the incomes of Aboriginal and non-Aboriginal constituencies in Canada in the twenty first century

Canada’s constitution recognizes three Aboriginal peoples: North American Indians, Inuit, and Metis.¹⁴ It is well known that Aboriginal peoples in Canada have, on average, lower incomes than non-Aboriginal people. For example, in comparison with non-minority native-born workers with similar characteristics, Aboriginal women faced income and earnings gaps of 10–20% between 1995 and 2005, while Aboriginal men faced gaps of 20–50% [47]. Less commonly known, however, is that there are also significant disparities between Aboriginal identity groups within Canada (NAEDB [44]; NAEDB [45]). The Right Honourable Paul Martin, a former Prime Minister, argued that Canada faced “a moral imperative” to close the income gaps that exist between Aboriginal and non-Aboriginal people—“the descendants of the people who first occupied this land deserve an equal chance to work for and to enjoy the benefits of our collective prosperity” [59, p. vi], prompting a variety of government policies directed at improving the lot of aboriginal peoples.¹⁵

To check the response to government policies, income gaps have been tracked and documented over time. The previous literature tended to emphasize average incomes; here the

¹⁴ In this paper, we maintain the terminology used by Statistics Canada in their codebooks to identify Aboriginal individuals belonging to these three groups.

¹⁵ These policies include, for example, the Affordable Housing Initiative (2001–2007), which helped fund the “construction and renovation of affordable housing units” [9, p. 67] and the Urban Aboriginal Strategy aimed at improving the socio-economic status of urban Aboriginals [9]. Other policies include Aboriginal Head Start in Urban and Northern Communities [41], Aboriginal Head Start on Reserve, and the First Nations and Inuit Child Care Initiative [40].

focus is on differences in the distributions of incomes over the 8 constituencies of Inuit, Metis, North American Indian and non-Aboriginal of each gender. In the spirit of Paul Martin's "equal chance" declaration, the issue is very much one of equality of opportunity, that income distributions should be independent of Aboriginal status, i.e. income distributions conditional on Aboriginal status should be identical. An index of the degree of variation in these conditional income distributions provides a measure of the extent to which equality of opportunity prevails in the sense of similar outcomes across Aboriginal and non-Aboriginal communities by gender. Usually this is examined in the context of a $k \times k$ transition matrix (see for example [26, 50, 53–55, 58]). Here the outcome state is continuous i.e. incomes have not been categorized with the consequent loss of information. Furthermore, an adaptation of Tukey's "Rootgram" approach facilitates extra focus on particular aspects of the distribution, magnifying differences within regions of interest in the comparison distributions. By setting λ at -0.5 and setting $h(x)$ to be equal to one of the probability density or cumulative densities of interest, differences in low frequency regions in the former case and differences in the lower income strata in the latter case are magnified. To exemplify this, the extent to which Canadian Aboriginal and non-Aboriginal groups have similar-dissimilar income distributions and how those differences have progressed over time is studied.

4.1 Data

Data from the 2001 and 2006 Census and the 2011 National Household Survey are employed to construct various weighted and unweighted transvariation estimates for the three observation years.¹⁶ Statistics Canada's 2001 and 2006 censuses provide detailed data on the social and demographic characteristics of the Canadian population. Each census is composed of two parts: the long-form and the short-form, with fewer questions the short-form census requires less time to complete than the long-form census. In 2001 and 2006, the short-form census was delivered to 100% of households, while the long-form census was distributed to one-fifth of Canadian households. Both surveys were mandatory. The long-form census had a response rate of about 94% in 2006 [48]. In 2011, the short-form census was delivered as usual, but the 2011 National Household Survey (NHS) replaced the long-form census, the main differences being that it was voluntary and distributed to 33% of households. The response rate, unsurprisingly, dropped precipitously to 69% [48].

4.2 Results

The relative positions of Aboriginal and non-Aboriginal constituencies in the income distribution have long been an issue especially with respect to their relative progress. 8 identifiable subgroups were examined for the years 2000, 2005 and 2010: the groups were male or female members of Inuit, Metis, North American Indian and non-Aboriginal Societies.¹⁷ Nominal incomes for individuals over the age of 15 were employed in the analysis. Note there was some top and bottom coding. Values greater than the 99th percentile in each geographical

¹⁶ As noted by Feir and Hancock [25], some caution is required when using Aboriginal data from the Censuses and the NHS. First, a number of reserves in Canada are not enumerated. Second, the structure of the ethnic origin question has changed a number of times. Finally, there is the potential impact of intra-generational ethnic mobility. Each plausibly causing exogenous variation in the size and characteristics of the Aboriginal population.

¹⁷ The incomes reported in the 2001, 2006, and 2011 surveys are based on individual earnings in 2000, 2005, and 2010, respectively.

Table 1 Summary statistics, total income, Aboriginal and non-Aboriginal, 2000, 2005 and 2010

	2000			2005			2010		
	Income	Standard deviation	N	Income	Standard deviation	N	Income	Standard deviation	N
Non-Aboriginal									
M	35,684	31,653.00	294,529	46,404	74,208.60	311,705	50,974	71,253.00	327,386
F	22,671	19,712.90	303,297	28,623	31,047.40	325,264	34,011	35,075.40	341,434
North American Indian									
M	19,609	20,024.10	4640	23,345	31,182.40	5,535	29,625	38,665.90	6655
F	15,546	14,260.70	5076	20,098	20,905.60	6184	24,780	25,479.60	7306
Metis									
M	26,796	24,153.00	2371	33,900	38,165.70	3676	42,877	50,645.30	3915
F	17,770	15,916.90	2297	23,214	23,562.40	3720	29,556	27,207.00	4196
Inuit									
M	22,052	20,710.90	344	27,052	28,775.40	389	33,513	36,323.10	429
F	19,379	18,557.50	344	24,603	25,912.10	430	31,715	33,663.30	482

Table 2 Relative male–female total income, Aboriginal and non-Aboriginal, 2000, 2005, and 2010

	2000	2005	2010
Non-Aboriginal	1.6	1.6	1.5
North American Indian	1.3	1.2	1.2
Metis	1.5	1.5	1.5
Inuit	1.1	1.1	1.1

Table 3 Coefficients of variation, total income, Aboriginal and non-Aboriginal, 2000, 2005, 2010

	2000	2005	2010
Non-Aboriginal			
M	0.89	1.60	1.40
F	0.87	1.08	1.03
North American Indian			
M	1.02	1.34	1.31
F	0.92	1.04	1.03
Metis			
M	0.90	1.13	1.18
F	0.90	1.02	0.92
Inuit			
M	0.94	1.06	1.08
F	0.96	1.05	1.06

region and gender were top coded and some negative values were down coded to a low threshold.

Table 1 presents basic statistics for income by identity and gender. The data show that male incomes are higher than female incomes for all four groups across all 3 years. Non-Aboriginal people face the largest gap between male and female incomes (Table 2). Metis have the next highest male–female relative income measure, followed by North American Indians and, finally, the Inuit. One surprising result is the consistency of the relative male–female income gap across this decade for all four groups. There were slight declines for non-Aboriginal people and North American Indians, but the Metis and the Inuit had the same level of relative male–female income in all 3 years.

Coefficients of variation (Table 3) show that in 2000, non-Aboriginal people of both genders had lower coefficients of variation than all three Aboriginal groups. In 2005, the reverse holds true. The data for males in 2010 suggest that non-Aboriginal people continued to have higher coefficients of variation, but the gap was declining. For females, the picture is less clear, as the highest coefficient of variation was among the Inuit, while the lowest coefficient of variation was among the Metis.

The data also show that coefficients of variation in 2005 are higher than coefficients of variation in 2000, while coefficients of variation in 2010 are lower for non-Aboriginal people, North American Indians, and female Metis, but higher for male Metis and the Inuit. The log income distributions for the 8 constituencies together with upper and lower boundaries for the years 2000, 2005 and 2010 respectively are illustrated in Figs. 1, 2 and 3.

Table 4 reports the Unweighted Transvariation Measures for 8 individual constituent distributions (combined), 4 female constituent distributions, 4 male constituent distributions and 4 gender integrated constituent distributions. In addition it reports the corresponding Transvari-

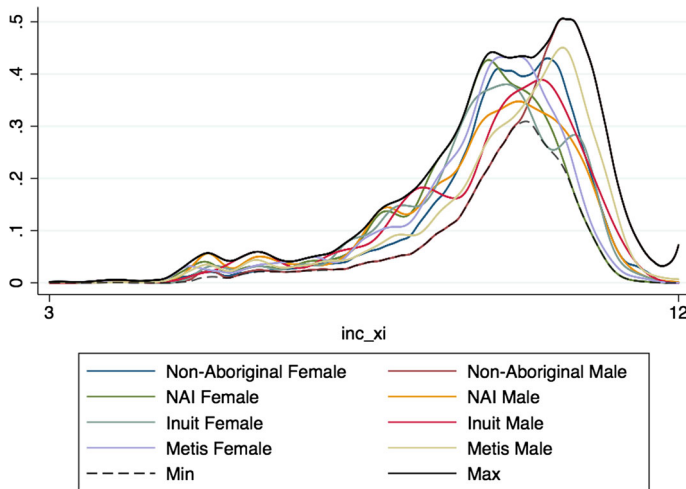


Fig. 1 Distribution of log income, Aboriginal and Non-Aboriginal by Gender in 2000

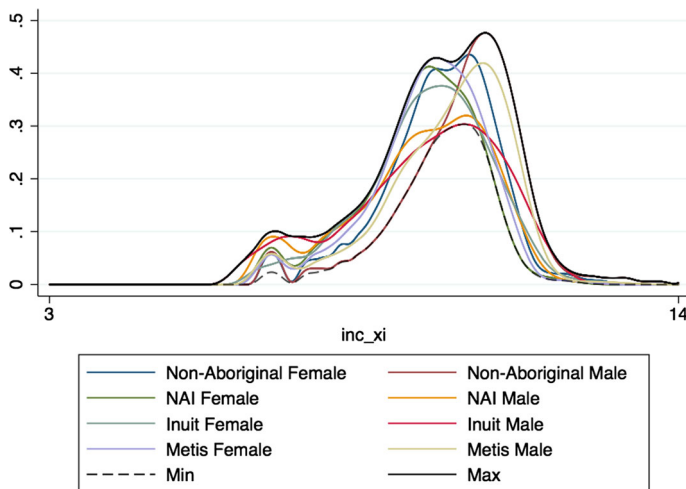


Fig. 2 Distribution of log income, Aboriginal and Non-Aboriginal by Gender in 2005

ation Measures for the Aboriginal constituencies. Firstly, from the full sample results, note that there is substantially more distributional variability in the males than in the females over all 3 years and there is a general trend in the reduction of distributional variability over time. Such trends in Aboriginal group results are less discernable. Over the whole period, the variation across Aboriginal categories in their distributions is much lower than for the full collection, male distributional differences have tightened somewhat in 2010. But considering the 8 distributions as a collection, there has been a considerable tightening, signalling the closing gaps between the extreme distributions. This overall may be seen as an improvement of equality of opportunity indexed by $1 - DIS$.

Turning to the exceptionality results Table 5 reports the exceptionality measures for each of the constituencies. In the context of 8 constituencies the non-Aboriginal groups are the most

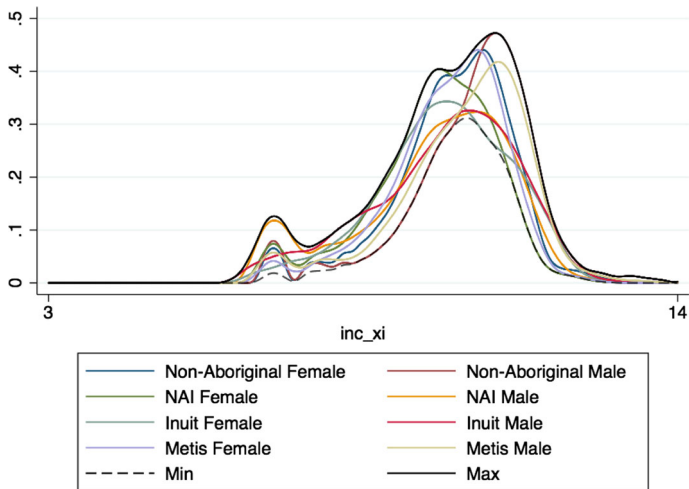


Fig. 3 Distribution of log income, Aboriginal and Non-Aboriginal by Gender in 2010

exceptional, that is to say their distribution contributes most to the transvariation measure, generating the biggest drop in transvariation when it is left out, the impact is greater when male, female and male–female combined distributions are considered but its omission is ubiquitously more important than the omission of any other group.

With regard to trends, the overriding trend in the 8 constituency analysis is a reduction in the exceptionality of the non-Aboriginal group and the increasing exceptionality of the North American Indian group. These results are a consequence of the non-Aboriginal group being the primary constituency in the lower boundary (i.e. the highest income group) and North American Indians being the primary constituency in the upper boundary (i.e. the lowest income group). Thus it can be deduced that improvements in equality of opportunity have been achieved largely as a result of reductions in the exceptionality of the non-Aboriginal constituencies.

4.3 Importance weighted transvariation measures

Importance weighting presents an opportunity to assess aspects of transvariation with respect to particular aspects of distributions relevant to particular groups. First importance weighting with respect to the full societal distribution is explored, then with respect to the Aboriginal communities, then with respect to non-Aboriginal communities. Two types of weighting are entertained, weighting with respect to the probability density function and weighting with respect to the cumulative density function. The former will emphasize variations in low density regions of the focus distribution which, given the foregoing figures, will be in the low and high income regions of the focus distribution. The latter emphasizes the low income regions of the focus distribution. The results are reported in Table 6.

The full societal PDF focus reveals increased variation in the tails of the distribution in the middle of the decade, which become attenuated at the end of the decade. The gender specific analysis reveals males having the same trend, whereas females have the opposite trend, which nets out when the genders are aggregated. The cumulative density results are similar suggesting this is in essence a lower tail issue. Turning to the non-Aboriginal and Aboriginal focus, an increasing trend in distributional variation in the tails is observed in virtually all

Table 4 Unweighted transvariation measures, Aboriginal and Non-Aboriginal by Gender in 2000, 2005, 2010

	2000	2005	2010
Full sample			
MNA, FNA, MI, FI, MM, FM, MNA, FNAI	0.1083	0.1002	0.0949
FNA, FI, FM, FNAI	0.1114	0.0948	0.1098
MNA, MI, MM, MNAI	0.1533	0.1482	0.1310
MFCNA, MFCI, MFCM, MFCNAI	0.1225	0.1180	0.1167
Aboriginal sample			
MI, FI, MM, FM, MNAI, FNAI	0.1063	0.1024	0.1041
FI, FM, FNAI	0.0527	0.0356	0.0567
MI, MM, MNAI	0.0628	0.0664	0.0610
MFCI, MFCM, MFCNAI	0.0474	0.0496	0.0556

MNA male Non Aboriginal, *FNA* female Non Aboriginal, *MI* male Inuit, *FI* female Inuit, *MM* male Metis, *FM* female Metis, *MNAI* male North American Indian, *FNAI* female North American Indian, *MFCNA* male–female combined Non Aboriginal, *MFCI* male–female combined Inuit, *MFCM* male–female combined Metis, *MFCNAI* male–female combined North American Indian

Table 5 Exceptionality of Aboriginal and non-Aboriginal people, change in transvariation measures, total income, 2000, 2005, 2010

Excluded group	2000	2005	2010
Panel A: Exceptionality of Non-Aboriginal people			
Non-Aboriginal, male and female	0.1143	0.0939	0.0673
Non-Aboriginal, female	0.2876	0.2722	0.2693
Non-Aboriginal, male	0.4248	0.3936	0.3410
Non-Aboriginal, male and female combined	0.3476	0.3233	0.2998
Panel B: Exceptionality of North American Indians			
North American Indian, male and female	0.0404	0.0341	0.0438
North American Indian, female	0.2611	0.2178	0.2506
North American Indian, male	0.3778	0.3453	0.3080
North American Indian, male and female combined	0.3119	0.2824	0.2795
Panel C: Exceptionality of Metis			
Metis, male and female	0.0123	0.0073	0.0063
Metis, female	0.2384	0.1975	0.2256
Metis, male	0.3076	0.3029	0.2646
Metis, male and female combined	0.2504	0.2418	0.2376
Panel D: Exceptionality of Inuit			
Inuit, male and female	0.0283	0.0246	0.0292
Inuit, female	0.2571	0.2179	0.2778
Inuit, male	0.3297	0.3210	0.2884
Inuit, male and female combined	0.2619	0.2593	0.2685

Table 6 Weighted transvariation measures, total income, Aboriginal and non-Aboriginal, 2000, 2005, 2010

	2000	2005	2010
Panel A: Population weighting			
PDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0128	0.0182	0.0164
FNA, FI, FM, FNAI	0.0160	0.0114	0.0157
MNA, MI, MM, MNAI	0.0196	0.0344	0.0301
MFCNA, MFCI, MFCM, MFCNAI	0.0168	0.0168	0.0171
CDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0187	0.0761	0.0700
FNA, FI, FM, FNAI	0.0254	0.0759	0.0809
MNA, MI, MM, MNAI	0.0296	0.1172	0.1026
MFCNA, MFCI, MFCM, MFCNAI	0.0264	0.0941	0.0901
Panel B: Non-Aboriginal weighting			
PDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0225	0.0386	0.0416
FNA, FI, FM, FNAI	0.0286	0.0371	0.0464
MNA, MI, MM, MNAI	0.0358	0.0596	0.0601
MFCNA, MFCI, MFCM, MFCNAI	0.0305	0.0474	0.0533
CDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0204	0.0253	0.0313
FNA, FI, FM, FNAI	0.0206	0.0185	0.0322
MNA, MI, MM, MNAI	0.0309	0.0435	0.0511
MFCNA, MFCI, MFCM, MFCNAI	0.0239	0.0278	0.0373
Panel C: Aboriginal weighting			
PDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0190	0.0767	0.0717
FNA, FI, FM, FNAI	0.0262	0.0764	0.0828
MNA, MI, MM, MNAI	0.0305	0.1180	0.1045
MFCNA, MFCI, MFCM, MFCNAI	0.0273	0.0948	0.0919
CDF			
MNA, FNA, MI, FI, MM, FM, MNAI, FNAI	0.0101	0.0727	0.0666
FNA, FI, FM, FNAI	0.0127	0.0726	0.0771
MNA, MI, MM, MNAI	0.0148	0.1130	0.0984
MFCNA, MFCI, MFCM, MFCNAI	0.0129	0.0904	0.0862

MNA male non Aboriginal, *FNA* female non Aboriginal, *MI* male Inuit, *FI* female Inuit, *MM* male Metis, *FM* female Metis, *MNAI* male North American Indian, *FNAI* female North American Indian, *MFCNA* male–female combined non Aboriginal, *MFCI* male–female combined Inuit, *MFCM* male–female combined Metis, *MFCNAI* male–female combined North American Indian

categories of analysis suggesting that while there is increasing equality of opportunity overall it is occurring mainly in the middle of the distribution and there is increasing variation in incomes (i.e. increasing inequality of opportunity) in the tails of the distribution. That is to say that, while average Metis, Inuit, North American Indian and non-Aboriginal people are becoming increasingly similar in their outcomes the low and high income generators are not.

In essence while some progress in equality of opportunity has been achieved over the decade it has not been the experience of low-income earners in the Aboriginal communities.

5 Conclusions

A methodology for analyzing the degree of variation in a collection of distributions has been presented which facilitates exploration of particular aspects of the differences between many distributions. The techniques were employed to study differences between the income distributions of males and females drawn from Metis, Inuit, North American Indian and non-Aboriginal constituencies in Canada in the first decade of the twenty first century. It was found that, while the distributions were becoming more alike (which may be interpreted as increasing equality of opportunity) this was occurring in the middle of the income distributions and that, at their extremes the distributions were diverging, suggesting that such improvements in equality of opportunity were not for all, that high and, in particular, low Aboriginal income earners were experiencing diminishing equality of opportunity over the period. An exceptionality analysis revealed that these results were largely driven by non-Aboriginal income earners contributing the most to the variation in income distributions, followed by the North American Indian constituency, the former at the high end of the collection of distributions, the latter at the bottom.

References

1. Anderson, T., Goodman, L.: Statistical inference about Markov chains. *Ann. Math. Stat.* **28**(1), 89–110 (1957)
2. Anderson, G., Linton, O., Whang, Y.-J.: Nonparametric estimation and inference about the overlap of two distributions. *J. Econom.* **171**(1), 1–23 (2012)
3. Anderson, G.: Polarization measurement and inference in many dimensions when subgroups can not be identified. *Economics: The Open-Access, Open-Assessment E-Journal*. **5**(2011–11), 1–19 (2011). doi:[10.5018/economics-ejournal.ja.2011-11](https://doi.org/10.5018/economics-ejournal.ja.2011-11)
4. Anderson, G.J., Leo, T.W.: On providing a complete ordering of non-combinable alternative prospects. Mimeo University of Toronto Discussion Paper (2016)
5. Anderson, G., Post, T., Whang, Y.-J.: Somewhere between Utopia and Dystopia: choosing from multiple incomparable prospects. University of Toronto Economics Discussion paper (2017)
6. Anderson, G.J., Ge, Y., Leo, T.W.: Distributional overlap: simple, multivariate, parametric and non-parametric tests for alienation, convergence and general distributional difference issues. *Econom. Rev.* **29**, 247–275 (2009)
7. Ballester, C., Vorsatz, M.: Random-walk-based segregation measures. *Rev. Econ. Stat.* **96**(3), 383–401 (2014). doi:[10.1162/RESTa00399](https://doi.org/10.1162/RESTa00399)
8. Blackorby, C., Bossert, W., Donaldson, D.: Quasi orderings and population ethics. *Soc. Choice Welf.* **13**, 129–150 (1996)
9. Bonesteel, S.: Canada's relationship with Inuit: a history of policy and program development. In: Anderson, E. (ed.) Prepared for Indian and Northern Affairs Canada (2006)
10. Brieman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Wadsworth International Group, Belmont (1984)
11. Broome, J.: The welfare economics of population. *Oxf. Econ. Pap.* **48**, 177–193 (1996)
12. Calò, D.: On a transvariation based measure of group separability. *J. Classif.* **23**, 143–167 (2006)
13. D'Ambrosio, C., Bossert, W., La Ferrara, E.: A generalized index of fractionalization. *Economica* **78**, 723–750 (2011)
14. Dagum, C.: Nonparametric and Gaussian Transvariation Theory: Its Economic Applications. Econometric Research Program, Princeton University, Research Memorandum 99 (1968)
15. Dagum, C.: Multivariate Transvariation Theory Among Several Distributions and Its Economic Applications. Econometric Research Program, Princeton University, Research Memorandum 100 (1968)

16. Dasgupta, P.: Lives and wellbeing. *Soc. Choice Welf.* **5**, 103–126 (1988)
17. Dasgupta, P.: Savings and fertility: ethical issues. *Philos. Public Aff.* **23**, 99–127 (1994)
18. Deutsch, J., Silber, J.: Analyzing the Impact of Income Sources on Changes in Polarization. Mimeo, New York (2008)
19. Deutsch, J., Silber, J.: On the Decomposition of Income Polarization by Population Subgroups. Mimeo, New York (2008)
20. Duclos, J.-Y., Esteban, J., Ray, D.: Polarization: concepts, measurement, estimation. *Econometrica* **72**, 1737–1772 (2004)
21. Echenique, F., Fryer Jr., R.G.: A measure of segregation based on social interactions. *Q. J. Econ.* **122**(2), 441–485 (2007)
22. Esteban, J., Ray, D.: Polarization, fractionalization and conflict. *J. Peace Res.* **45**, 163–182 (2008)
23. Esteban, J., Schneider, G.: Polarization and conflict: theoretical and empirical issues (introduction to the special issue). *J. Peace Res.* **45**, 131–141 (2008)
24. Fearon, J., Laitin, D.: Ethnicity, insurgency, and civil war. *Am. Polit. Sci. Rev.* **97**, 75–90 (2003)
25. Feir, D., Hancock, R.L.A.: Answering the call: a guide to reconciliation for quantitative social scientists. *Can. Public Policy* **42**(3), 350–365 (2016)
26. Fields, G.S.: Income mobility. In: Blume, L., Durlauf, S. (eds.) *The new Palgrave dictionary of economics*. Palgrave-MacMillan (2008)
27. Flückiger, Y., Silber, J.: *The Measurement of Segregation in the Labour Force*. Springer Science and Business Media, Berlin (2012)
28. Foster, J., Shneyerov, A.: Path independent inequality measures. *J. Econ. Theory* **91**(199), 222 (2000)
29. Foster, J., Greer, J., Thorbecke, E.: A class of decomposable poverty measures. *Econometrica* **52**, 761–766 (1984)
30. Frankel, D.M., Volij, O.: Measuring school segregation. *J. Econ. Theory* **146**(1), 1–38 (2011)
31. Georgescu-Roegen, N.: *The Entropy Law and the Economic Process*. Harvard University Press, Cambridge (1971)
32. Gentzkow, M., Shapiro, J.M.: Ideological segregation online and offline. *Q. J. Econ.* **126**(4), 1799–1839 (2011)
33. Gentzkow, M., Shapiro, J.M., Taddy, M.: *Measuring Polarization in High-Dimensional Data: Method and Application to Congressional Speech*. Mimeo, New York (2015)
34. Gini, C.: Il concetto di transvariazione e le sue prime applicazioni. *Giornale degli Economisti e Rivista di Statistica*. **5**, 1–35 (1916)
35. Gini, C.: *Transvariazione*. Libreria Goliardica, Rome (1959)
36. Goldthorpe, J.H.: *Social Mobility and Class Structure in Modern Britain*, 2nd edn. Clarendon Press, Oxford (1987)
37. Handcock, M., Morris, M.: *Relative Distribution Methods in the Social Sciences. Statistics for Social Science and Public Policy*. Springer, Berlin (1999)
38. Hall, P., Yatchew, A.: Unified approach to testing functional hypotheses in semiparametric contexts. *J. Econom.* **127**, 225–252 (2005)
39. Hogg, R., Tanis, E.: *Probability and Statistical Inference*. Prentice Hall, Upper Saddle River (1997)
40. Inuit Tapiriit Kanatami [ITK]: *Assessing the Impact of the First Nations and Inuit Child Care Initiative Across Inuit Nunangat*. Inuit Tapiriit Kanatami (2014)
41. Kay-Raining Bird, E.: Health, education, language, dialect, and culture in First Nations, Inuit, and Metis communities in Canada: an overview. *Can. J. Speech Lang. Pathol.* **35**(2), 110–124 (2011)
42. Kobus, M., Milo's, P.: Inequality decomposition by population subgroups for ordinal data. *J. Health Econ.* **31**(2012), 15–21 (2012)
43. Mele, A.: Poisson indices of segregation. *Reg. Sci. Urban Econ.* **43**, 65–85 (2013)
44. National Aboriginal Economic Development Board (2012) *The Aboriginal Economic Benchmarking Report*
45. National Aboriginal Economic Development Board (2015) *The Aboriginal Economic Progress Report 2015*
46. Pearson, K.: On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Mag. Ser. 5* **50**(302), 157–175 (1900)
47. Pendakur, K., Pendakur, R.: Aboriginal income disparity in Canada. *Can. Public Policy* **37**(1), 61–83 (2011)
48. Penney, C.: Aboriginal data as a result of changes to the 2011 Census of population. Strategic Research, Aboriginal Affairs and Northern Development Canada (2013). https://www.aadnc-aandc.gc.ca/DAM/DAM-INTER-HQ-AI/STAGING/texte-text/rs_re_brief_2011NHS_print_1380028890150_eng.pdf

49. Pollard, D.E.: *A User's Guide to Measure Theoretic Probability*. Cambridge University Press, Cambridge (2002)
50. Ramos, X., Van de Gaer, D.: Empirical approaches to inequality of opportunity: principles, measures, and evidence. Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit, No. 6672 (2014)
51. Reynal-Querol, M.: Ethnicity, political systems, and civil wars. *J. Confl. Resolut.* **46**, 29–54 (2002)
52. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423 (1948)
53. Shorrocks, A.F.: Income mobility and the Markov assumption. *Econ. J.* **46**, 566–578 (1976)
54. Shorrocks, A.F.: The measurement of mobility. *Econometrica* **46**, 1013–1024 (1978)
55. Shorrocks, A.F.: Income inequality and income mobility. *J. Econ. Theory* **46**, 376–393 (1978)
56. Theil, H.: *Economics and Information Theory*. North Holland, Amsterdam (1967)
57. Tukey, J.W.: *Exploratory Data Analysis*. Addison-Wesley, Reading (1977)
58. Van de Gaer, D., Schokkaert, E., Martinez, M.: Three meanings of intergenerational mobility. *Economica* **68**(272), 519–537 (2001)
59. Weinstein, J.: *Quiet Revolution West: The Rebirth of Metis Nationalism*. Fifth House Publishers, Calgary (2007)
60. Yalonetzky, G.: A comparison between the Pearson-based dissimilarity index and the multiple-group overlap index. OPHI discussion paper series 16a (2010)
61. Yitzahki, S.: Economic distance and overlapping of distributions. *J. Econom.* **61**, 147–159 (1994)