University of Toronto
Department of Economics

Working Paper 751

The mode and the measurement of inequality with single and multiple ordinal variables

By Gordon Anderson and Gaston Yalonetzky

April 22, 2023

The mode and the measurement of inequality with single and multiple ordinal variables

Gordon Anderson
University of Toronto, IFS

Gaston Yalonetzky
University of Leeds, OPHI, III

January 2023

Abstract

The extensive use of multivariate ordered categorical data in the social sciences presents challenges for the measurement of socioeconomic inequities. The ambiguities inherently associated with artificial attribution of scale to ordinal categories, preclude the use of standard distance-based inequality and polarization measures. These issues have been surmounted in the univariate world by employing notions of likelihood distance (the increasing likelihood that an outcome between two categories will occur the bigger is their categorical gap) and aggregating outcome distances from the median category as a reference point, adapting the transfer principle using Hammond transfers, or with probability-based notions of status. Unfortunately, the median category is not always the ideal reference point of complete commonality and is not uniquely defined in multivariate environments. However, as the most frequently observed outcome, the modal category provides a natural measure of the extent of commonality or equity in the population, thus providing a readymade reference point from which to measure likelihood outcome distance. We provide axiomatic foundations and characterise classes of modally focused inequality measures for univariate and multivariate ordered categorical environments together with their asymptotic distributions for inference purposes. We also identify the partial ordering induced by our proposed mode-clustering transfers which provides a useful robustness test for inequality indices in the spirit of stochastic dominance conditions. In an empirical illustration we study the evolution of inequality in educational attainment and experience among men and women in Canada.

JEL Codes: D63, I14, I31.

Keywords: Inequality, mode, ordinal variables, multivariate analysis

1. Introduction.

The extensive use of survey information in multidimensional cardinal and ordered categorical formats, and the growing interest in understanding the extent of inequality across societies, calls for measures appropriate for such environments. When the outcomes have cardinal measure, measuring distance between them does not present a problem.[1] However, when outcomes are ordered categories, distance measurement is problematic. One solution has been to attribute cardinal scale to the ordinal categories (Cantril 1965), but the arbitrary nature of the scaling and the scale-dependency of traditional inequality measures for cardinal variables, such as the Gini index, are each cause for concern (Bond and Lang 2019, Schroder and Yitzhaki 2017). Mendelsohn (1987) resolved this issue by positing the notion of likelihood distance from a quantile reference point, a construct employed by Allison and Foster (2004) and subsequent contributions to the paradigm of median-preserving spreads,[2] as well as more recently by Cowell and Flachaire (2017) in their proposal to measure status inequality with ordinal variables.

These proposals for ordinal variables capture the intuition of inequality as the extent to which observations are not located at a common reference point. In the context of prospect theory and decision making, Kahneman and Tversky stressed the importance of specifying an appropriate point of reference (Kahneman and Tversky 1979, Tversky and Kahneman 1991), a sentiment that is, arguably, equally pertinent in inequality measurement. From a likelihood perspective, the most likely candidate for a category in which all could potentially reside is the mode, the most frequently observed category, which provides an ideal, intuitively plausible, likelihood-based reference point for analysis. This likelihood approach is especially relevant in multidimensional spaces where unique median points of reference are difficult to conceptualize. All that remains is to quantify the extent to which the realized outcomes diverge from the reference point and relate these divergences to ethically meaningful distributional transformations.

In ordered categorical spaces the absence of cardinal measure and the ambiguities inherently associated with artificial attribution of scale to ordered categories, precludes employment of standard distance-based inequality measures such as the Gini (Gini 1921) and its polarization analogue (Esteban and Ray 1994). Following Mendelsons' (1987) foundational work on quantifying distance between categories in likelihood terms and given the importance of reference-dependent analysis, the Median Preserving Spread (MEPS) approach has become a popular workhorse in inequality measurement with single ordinal variables[3] (Allison and Foster 2004) which has been extended to multidimensional spaces. Typically, this work identifies the

[1] Alternatively, for cardinal data one can implement measures based on the Atkinson-Kolm-Sen paradigm.

[2] See e.g., Apouey (2007), Abul Naga and Yalcin (2008), Kobus and Milos (2012), Lazar and Silber (2013), Kobus (2015), Lv et al. (2015). These echo the seminal notion of mean-preserving spreads for continuous variables (Rothschild and Stiglitz, 1970).

[3] MEPS overcomes the scaling problems associated with measuring distance between categories when they have been artificially attributed cardinal value (Bond and Lang 2019, Schroder and Yitzhaki 2017).

median category[4] as the reference point and employs Mendelson's notions of likelihood distance from it. However, from a likelihood perspective, given that inequality is measured as the extent to which agents do not universally reside in a unique single outcome, the median outcome may not be the best reference category since it is not always the most frequently inhabited and thus most likely point of commonality. Furthermore, extending the median category condition to multidimensional space usually results in an array of categories defining a probability contour on that space, of which the uniquely defined quantile category proposed by Kobus and Kurek (2019) may or may not be a member.

An alternative is to employ the modal category as a reference point which is, under broad appropriate circumstances, readily conceptualized as unique in any multivariate measurement paradigm. Furthermore, since the mode is the most common outcome by definition, its frequency provides a natural, albeit incomplete, likelihood-oriented measure of the extent of commonality or equality of outcome inherent in a distribution; and its complement can be readily seen as an elementary measure of the extent of inequality or lack of commonality of outcome. Consider, for instance, ordered-categorical distributions $\boldsymbol{p} = (0.4,0,0.2,0.2,0.2)$ and $\boldsymbol{q} = (0.45,0,0.15,0.2,0.2)$, where $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a transfer of 0.05 from the common median (the third category) to the common mode (the first category). Any inequality index satisfying the MEPS principle, would deem $\boldsymbol{q}$ more unequal than $\boldsymbol{p}$ as the left tail of **q** is thicker than **p**'s, *ceteris paribus*; i.e., there is a higher proportion of units further away from the common median. However, at the same time, there is a higher proportion of people in the most popular category, namely the mode. Moreover, the right tail has become slimmer; that is, fewer units are that far away from the mode as before. Thus, building on this intuition we can propose an inequality measurement criterion based on the mode as a reference point.

Our main contribution is the proposal of axiomatic foundations and indices for the measurement of inequality with ordinal variables using the mode as the reference point for both univariate and multivariate settings. The axiomatically characterised indices are sensitive to transfers which decrease the likelihood distance from the mode. Thus, though the median and modal approaches to inequality measurement employ the same likelihood distance measures and agree in their identification of fully egalitarian distributions, they fundamentally differ in their notion of maximum inequality construct. With the median approach, extreme inequality exists when all the probability mass resides evenly and exclusively at the extreme categories, whereas extreme inequality in the modal approach arises when the distribution becomes uniform, for only then the modal category ceases to exist as such.

We also identify the (incomplete) partial ordering induced by our proposed mode-clustering transfers which provides a useful robustness test for inequality indices in the spirit of stochastic dominance conditions. In fact, this partial ordering's implementation conditions are remarkably similar to those of the MEPS partial ordering derived by Allison and Foster (2004), only differing

[4] The median is defined in a way that lower categories have jointly less than a 50% chance of occurring whilst the rest, including the median itself, have jointly at least a 50% likelihood of occurrence.

in their reference point. Additionally, we provide analytical formulas for the standard errors of our flagship measures in the spirit of Abul Naga and Stapenhurst (2020).

Finally, we illustrate the proposed measures with a study of the evolution of inequality in educational attainment among men and women in Canada. Even though in most comparisons, the trends in inequality between 2006 and 2016 coincide between mode-focused and median-focused indices, we do find a couple of instances in the female sample in which the choice of reference point bears practical consequences.

The rest of the paper proceeds as follows. Section 2 introduces the axiomatic foundations for inequality indices of ordinal variables focusing on the mode, chiefly the mode-clustering transfers principle, together with their respective motivation. Section 3 provides a simple axiomatic characterisation of inequality indices satisfying the key axioms, followed by an exemplary class of readily implementable measures, with particular emphasis on a `flagship' member. Section 4 derives the incomplete partial ordering induced by the mode-clustering transfers principle. This ordering relates to a stochastic dominance condition which is useful as a robustness test for inequality comparisons. Section 5 provides the analytical formulas for the standard errors of the flagship index introduced in section 3. Section 6 extends the framework to the multivariate environment. Section 7 provides the empirical illustration on the evolution of educational inequality by gender in Canada between 2006 and 2016. Finally, section 8 concludes with some remarks.

2. Axiomatic foundations for inequality measures based on the mode.

2.1 Notation

Let $\boldsymbol{p} \equiv (p_1, \dots p_K)$ be a distribution of frequencies where $K > 1$ is the number of ordered categories, such that $0 \leq p_i \leq 1$ for all $i = 1, \dots, K$ and $\sum_{i=1}^{K} p_i = 1$, with each category labelled by a natural number between 1 and K, just for reference purposes (i.e., these labels are not cardinal scales). Also let $\boldsymbol{P} \equiv (P_1, \dots P_K)$ be the cumulative distribution function (CDF), such that $P_i = \sum_{j=1}^{i} p_j$ for all $i = 1, \dots, K$. The set of all possible distributions with K categories is $\mathcal{S}^K$. Additionally, let n be a population size.

As is well known for ordinal variables, the only two measures of central tendency that can be defined independently of arbitrary scales are the median and the mode. Following Kobus (2015), let the single median category m^e exist and be defined whenever (1) $P_{m^e-1} < 0.5$ and $P_{m^e} \geq 0.5$ for $m^e > 1$ or (2) $m^e = 1$ if $P_1 > 0.5$.[5] Meanwhile, let the single modal category m^o exist and be defined whenever $p_{m^0} > p_i$ for all natural numbers $i \in \{1, \dots, K\}/\{m^o\}$. In the case of multimodal distributions, which are also possible, we have a set of categories $\{m_1^0, \dots, m_J^0\}$ with

[5] There are also distributions with multiple median categories, but we do not consider them here for ease of presentation. Further details can be found in Kobus (2015).

$J \leq K$ such that: $p_{m_1^o} = \cdots = p_{m_J^o} = \max\{p_1, \dots, p_K\}$. For ease of presentation, we focus mostly on unimodal distributions.

Finally, let $\mathcal{E}^K \in \mathcal{S}^K$ be the set of K degenerate distributions, i.e. egalitarian distributions characterised by $p_i = 1$ for one $i \in \{1, \dots, K\}$, $\boldsymbol{u}^{\boldsymbol{K}}$ be a uniform distribution with K categories, and $I: \mathcal{S}^K \to \mathbb{R}_+$ be an inequality index for ordinal variables.

2.2 Motivation and key axioms

The median-preserving spread (MEPS) principle (Allison and Foster, 2004; Apouey 2007; Abul Naga and Yalcin, 2008; Kobus and Milos, 2012; Kobus, 2015) states that an inequality index should increase (or minimally not decrease) whenever frequency mass is moved away from the median category and toward the tails without changing the median category. Even though inequality indices respecting the MEPS principle only rank degenerate distributions as egalitarian, they will generally signal higher inequality as more frequency mass moves toward the tails irrespective of the formation of high frequencies in them. In this sense, borrowing the intuition and terminology from the polarization literature (starting with Esteban and Ray, 1994), we would say that the MEPS principle prioritises *alienation* from the median category irrespective of *identification* within alternative categories.

Even though the MEPS principle is admittedly appealing, inequality measurement criteria based on alternative intuitions can also be proposed. Again, borrowing concepts from the polarization literature, we could conceive inequality for ordinal variables as the opposite of *identification* within one category, coupled with *alienation* from the most common category. In that case, the mode would be a more suitable measure of central tendency for an assessment of dispersion around it.

Consider, for instance, $\boldsymbol{p} = (0.4,0,0.2,0.2,0.2)$ and $\boldsymbol{q} = (0.45,0,0.15,0.2,0.2)$, where $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a transfer of 0.05 from the common median ($m^e = 3$) to the common mode ($m^o = 1$). If I satisfies the MEPS principle, then we would get $I(\boldsymbol{q}) > I(\boldsymbol{p})$ as the left tail of **q** is thicker than **p**'s, ceteris paribus; i.e., there is a higher proportion of units further away, or *alienated* from, the common median. However, at the same time identification with the common mode has also increased. That is, there is a higher proportion of people in the most popular category. Moreover, the right tail has become slimmer; that is, fewer units are alienated from the mode. Thus, building on this intuition we can propose an inequality measurement criterion based on the mode as a reference point. Then, an index satisfying such a property (to be formally defined below) would stipulate that $I(\boldsymbol{q}) < I(\boldsymbol{p})$.

In order to formalise this intuition into a measurement proposal definition 1 introduces a mode-preserving spread:

Definition 1: Mode clustering transfers (MOCT): $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a mode clustering transfer if, for any pair of categories i, j such that either $1 \leq i < j \leq m^o$ or $m^o \leq j < i \leq K$, $q_i = p_i - \frac{1}{n}, q_j = p_j + \frac{1}{n}, q_l = p_l \ \forall l \neq i, j$, $q_{m^o} \geq p_{m^o}$ and $m^o(\boldsymbol{q}) = m^o(\boldsymbol{p})$.

Note from definition 1 that the two distributions have the same mode, but $\boldsymbol{q}$ features higher identification with it together with lower alienation away from it. Then we state the corresponding principle:

Axiom 1 Mode clustering transfers (MOCT) principle: $I(\boldsymbol{q}) \leq I(\boldsymbol{p})$ if $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a sequence of mode clustering transfers.

Take any distribution and implement a sequence of MOCT; that is, clustering transfers around the mode. Then we invariably end up with a degenerate distribution featuring complete identification with the mode and nil alienation from it. Hence, for any K, the set of K egalitarian distributions relevant to inequality measurement criteria based on MOCT is identical to the set based on MEPS. Furthermore, we can reasonably ask inequality indices to satisfy the mode-centered version of the equality principle:

Axiom 2 Equality principle: $I(\boldsymbol{p}) = 0$ if and only if $\boldsymbol{p} \in \mathcal{E}^K$, for any $K > 1$.

Likewise, within $\mathcal{S}^K$ we can identify a set of distributions with maximal inequality. We know from the definition of a MOCT that the frequency of the invariant mode cannot decrease after an equality-inducing transformation (i.e. a MOCT). Therefore, maximal inequality must be associated with the lowest possible frequency in the mode, which is attained only in the case of a uniform distribution. Therefore, the set of distributions with maximal inequality contains one element and we can ask indices to satisfy the following axiom if we would like their values to be bound by the real interval $[0,1]$:

Axiom 3 Maximality principle: $I(\boldsymbol{p}) = 1$ if and only if $\boldsymbol{p} = \boldsymbol{u^K}$ for any $K > 1$.

3. Indices of inequality for single ordinal variables respecting the MOCT principle.

We can provide a simple characterisation of inequality indices for ordinal variables satisfying the three axioms introduced in the previous section:

Proposition 1: For every natural number $K > 1$, $I: \mathcal{S}^K \rightarrow \mathbb{R}_+$ satisfies the equality principle, the MOCT principle and the maximality principle if and only if for every $\boldsymbol{p} \in \mathcal{S}^K$:

$$I(\boldsymbol{p}) = \frac{h(\boldsymbol{p}) - h(\boldsymbol{q})}{h(\boldsymbol{u^K}) - h(\boldsymbol{q})}$$

Where $h: \mathcal{S}^K \rightarrow \mathbb{R}_+$ satisfies the MOCT principle and $\boldsymbol{q} \in \mathcal{E}^K$.

Proof: Sufficiency: Since, for every natural number $K > 1$, $\boldsymbol{q} \in \mathcal{E}^K$ has the mode with the highest possible frequency ($m^o(\boldsymbol{q}) = 1$) and $\boldsymbol{u^K}$ has the mode with the lowest possible frequency ($m^o(\boldsymbol{u^K}) = \frac{1}{K}$), then it is easy to show that any distribution $\boldsymbol{p} \in \mathcal{S}^K$ can be obtained from $\boldsymbol{u^K}$ through a finite sequence of MOCT, while at least one of the K possible $\boldsymbol{q} \in \mathcal{E}^K$ can be obtained from any distribution $\boldsymbol{p} \in \mathcal{S}^K$ through a finite sequence of MOCT. Therefore, since h satisfies the MOCT principle, we have $h(\boldsymbol{u^K}) \geq h(\boldsymbol{p})$, $h(\boldsymbol{p}) \geq h(\boldsymbol{q})$ and $h(\boldsymbol{u^K}) > h(\boldsymbol{q})$, which in turn implies $1 = I(\boldsymbol{u^K}) \geq I(\boldsymbol{p}) \geq I(\boldsymbol{q}) = 0$, for every $\boldsymbol{p} \in \mathcal{S}^K$. Thus $I(\boldsymbol{p})$ satisfies the equality principle and the maximality principle. Also, it is straightforward to note that if h satisfies the MOCT principle then so will $I(\boldsymbol{p})$ as $\frac{1}{h(\boldsymbol{u^K})-h(\boldsymbol{q})}$ and $\frac{-h(\boldsymbol{q})}{h(\boldsymbol{u^K})-h(\boldsymbol{q})}$ remain constant for a given K.

Necessity: Without loss of generality, consider $I(\boldsymbol{p}) = ah(\boldsymbol{p}) + b$ such that $h: \mathcal{S}^K \to \mathbb{R}_+$, $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$. Since I satisfies the MOCT principle then h must satisfy it as well. Moreover, satisfaction of the equality principle requires $I(\boldsymbol{q}) = ah(\boldsymbol{q}) + b = 0 \to b = -ah(\boldsymbol{q})$. Meanwhile, satisfiaction of the maximality principle requires $I(\boldsymbol{u^K}) = ah(\boldsymbol{u^K}) - ah(\boldsymbol{q}) = 1 \to a = \frac{1}{h(\boldsymbol{u^K})-h(\boldsymbol{q})}$. Finally, replacing both a and b we obtain $I(\boldsymbol{p}) = \frac{h(\boldsymbol{p})-h(\boldsymbol{q})}{h(\boldsymbol{u^K})-h(\boldsymbol{q})}$. ■

Potentially, numerous functional forms for h are admissible. For instance, with $h(\boldsymbol{p}) = \sum_{i=1}^{m^o-1} w_i P_i + \sum_{i=m^o}^{K} w_i(1-P_i)$ such that $w_i > 0$ for all $i = 1, \dots, K$, we obtain the following subclass of inequality indices (noting that $h(\boldsymbol{q} \in \mathcal{E}^K) = 0$):

$$I(\boldsymbol{p}) = K \frac{\sum_{i=1}^{m^o-1} w_i P_i + \sum_{i=m^o}^{K} w_i(1-P_i)}{\sum_{i=1}^{m^o-1} w_i i + \sum_{i=m^o}^{K} w_i(K-i)} \quad (1)$$

Where $w_i > 0$ for all $i = 1, \dots, K$.[6] Furthermore, a convenient special case obtains when $w_1 = \cdots = w_K$:

$$I(\boldsymbol{p}) = 2K \frac{\sum_{i=1}^{m^o-1} P_i + \sum_{i=m^o}^{K} (1-P_i)}{m^0(m^o-1)+(K-m^0+1)(K+m^0-2)} \quad (2)$$

4. Partial orderings for single ordinal variables

As mentioned, the class of inequality indices characterised in proposition 1 admits several members depending on the choice of functional form for h. Then, as usual in other areas of distributional analysis, we can ask whether some distributions can be ordered robustly by all members of class I in proposition 1. Theorem 1 provides the stochastic dominance condition for inequality measures for ordinal variables based on MOCTs.

---

[6] This is a popular choice among proposals for indices based on the MEPS paradigm (e.g., see Kobus and Milos, 2012; Lazar and Siber, 2013; Chakravarty and Maharaj, 2015). Note that the weights w_i do not need to add up to 1.

Theorem 1: For any $\boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}^K$, the following three statements are equivalent:

(i) $I(\boldsymbol{p}) \geq I(\boldsymbol{q})$ for every $I: \mathcal{S}^K \rightarrow \mathbb{R}_+$ satisfying the MOCT principle.
(ii) $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a finite sequence of MOCT.
(iii) $P_i \geq Q_i$ for all $i = 1, \dots, m^o - 1$ and $Q_i \geq P_i$ for all $i = m^o, \dots, K$.

Proof: $(ii) \rightarrow (i)$ is straightforward.

$(i) \rightarrow (iii)$: Since $(i) \rightarrow (iii)$ if and only if $\sim (iii) \rightarrow \sim (i)$, we prove the latter. First, consider that (i) implies that $I(\boldsymbol{p}) \geq I(\boldsymbol{q})$ for all I in equation 1, namely for all possible weighting vectors $w_1, \dots, w_K$, with strictly positive elements, since the class in equation 1 is a subset of all indices considered in statement (i). Second, let $\sim (iii)$. Then for all categories $i < m^o$ for which $P_i < Q_i$ and $i \geq m^o$for which $Q_i < P_i$, we could respective weights w high enough in order to render $I(\boldsymbol{p}) < I(\boldsymbol{q})$ for at least one set of weights in equation 1.

$(iii) \rightarrow (ii)$. Using the language and definitions of Gravel et al. (2020) we note that a MOCT from category i to j such that $1 \leq i < j \leq m^o$ is essentially an *increment*, whereas a MOCT from category i to j such that $m^o \leq j < i \leq K$ is a *decrement*. Hence, we know from Gravel et al. (2020) that $P_i \geq Q_i$ for all $i = 1, \dots, m^o - 1$ implies that $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a finite sequence of increments involving categories 1 through m^o (because it could be that $P_{m^o-1} > Q_{m^o-1}$). This can be ascertained by moving sequentially $P_i - Q_i$ to category $i + 1$ in $\boldsymbol{p}$ starting with $i = 1$ all the way to $i = m^o - 1$. A mentioned, these increments are all MOCT because they occur to the left of the mode. Likewise, we know that $Q_i \geq P_i$ for all $i = m^o, \dots, K$ implies that $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a finite sequence of decrements involving categories m^o through K. This can be ascertained by moving sequentially $Q_i - P_i$ from category $i + 1$ to category i in $\boldsymbol{p}$ starting with $i = K - 1$ all the way to $i = m^o$. A mentioned, these decrements are all MOCT because they occur to the right of the mode. Hence if $P_i \geq Q_i$ for all $i = 1, \dots, m^o - 1$ and $Q_i \geq P_i$ for all $i = m^o, \dots, K$; then $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a finite sequence of MOCT.■

Note the striking resemblance between the implementable condition (iii) and its testable counterpart for the MEPS principle derived by Allison and Foster (2004) and Kobus (2015). The key difference between the two is the measure of central tendency. The MEPS implementable condition relies on the median, whereas ours relies on the mode.

5. Statistical Inference.

The following establishes the sampling distributions of the inequality index estimates. To facilitate inference, note that given $\boldsymbol{p} \in \mathcal{S}^K$ and a K-dimensional unit vector $\mathbf{1}_{\boldsymbol{K}}$, the vector of CDF values $\boldsymbol{P}$ is given by $\boldsymbol{P} = \boldsymbol{pL}$ where $\boldsymbol{L}$ is an upper-diagonal matrix of ones. The corresponding vector of Survival Function values $\boldsymbol{P^C}$ with typical element P_k^C is given by $\boldsymbol{P^C} = \mathbf{1}_{\boldsymbol{K}} - \boldsymbol{P}$. Inference in this context is straightforward. Following Rao (2009) and Abul Naga and Stapenhurst (2015), given an independent random sample of size n, $\widehat{\boldsymbol{p}}$, the estimator of the vector of outcome probabilities $\boldsymbol{p}$ is multivariate normal:

$$\sqrt{n}(\widehat{\boldsymbol{p}}-\boldsymbol{p})\sim N(\mathbf{0}_{\boldsymbol{K}},\mathbf{V})$$

where:

$$\boldsymbol{V}=\begin{bmatrix} p_1 & 0 & 0 & . & 0 \\ 0 & p_2 & 0 & . & 0 \\ 0 & 0 & p_3 & . & 0 \\ . & . & . & . & 0 \\ 0 & 0 & 0 & . & p_K \end{bmatrix}-\begin{bmatrix} p_1 \\ p_2 \\ . \\ . \\ p_K \end{bmatrix}\begin{bmatrix} p_1 & p_2 & . & . & p_K \end{bmatrix}$$

So that: $\sqrt{n}(\widehat{\boldsymbol{P}}-\boldsymbol{P})\sim N(\underline{0},\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L})$.

Given a focus category m^0(which could be either the median or the mode), $I(\boldsymbol{p})$ from equation (2) may be written as $\varphi\mathbf{1}_{\boldsymbol{K}}\boldsymbol{C}$ where:

$$\boldsymbol{C}=\begin{bmatrix} P_1 \\ . \\ P_{m^0-1} \\ 1-P_{m^o} \\ . \\ 1-P_k \end{bmatrix}=\boldsymbol{P}'\boldsymbol{D}+(\mathbf{0}_{\boldsymbol{m^o-1}},\mathbf{1}_{\boldsymbol{K-m^o+1}})',$$ where $\boldsymbol{D}$ is a diagonal matrix with the first m^o elements equal to 1 and the rest equal to -1; and $\varphi=\frac{2K}{m^0(m^o-1)+(K-m^0+1)(K+m^0-2)}$.

Then $I(\boldsymbol{p})$ may be shown to be asymptotically normal such that

$$\sqrt{n}\big(I(\widehat{\boldsymbol{p}})-I(\boldsymbol{p})\big)\sim N\big(\underline{0},\varphi^2\mathbf{1}_{\boldsymbol{K}}\boldsymbol{D}\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L}'\boldsymbol{D}\mathbf{1}_{\boldsymbol{K}}'\big)$$

Modal Determination.

A category is a Local Mode when its mass is greater than its immediate neighbors', and it is a Universal mode when its mass is greater than that of all other categories. To examine this, consider k a candidate modal category, then for it to be a local mode $p_{\mathrm{k}}-p_{\mathrm{k-1}}>0$ and $p_{\mathrm{k}}-p_{\mathrm{k+1}}>0$ and for it to be a universal mode $p_{\mathrm{k}}-p_{\mathrm{j}}>0$ for all $j=1,..,K\ j\neq k$. Letting $\boldsymbol{d}_{\boldsymbol{k,j}}$ be a K x 1 vector with its k'th element 1 and its j'th element -1 with all other elements set to 0, this amounts to jointly testing $\boldsymbol{p}\boldsymbol{d}_{\boldsymbol{k,k-1}}'>0$ and $\boldsymbol{p}\boldsymbol{d}_{\boldsymbol{k,k+1}}'>0$ in the case of local modality and $\boldsymbol{p}\boldsymbol{d}_{\boldsymbol{k,j}}'>0$ for all $j=1,..,K\ j\neq k$ in the universal case. Noting that generically since:

$$\sqrt{n}\boldsymbol{d}_{\boldsymbol{k,j}}(\widehat{\boldsymbol{p}}'-\boldsymbol{p}')\sim N\big(\underline{0},\boldsymbol{d}_{\boldsymbol{k,j}}\,\boldsymbol{V}\boldsymbol{d}_{\boldsymbol{k,j}}'\big)$$

these may be examined using the Studentized Maximum Modulus Distribution (Stoline and Ury 1979). Alternatively, the situation can be quickly assessed by forming the $2\ x\ 1$ or $K-1\ x\ 1$ vector $\underline{dp}=\left[\boldsymbol{d}_{\boldsymbol{k,j}}\widehat{\boldsymbol{p}}'\right]$ where $j\ =\ k-1\ ,k+1$ in the local mode instance and $j=1,..,K\ j\neq k$ in the universal mode instance. Let s(.) be a vector element summation function and consider $AMB=s\left(\underline{dp}\right)/s\left(\left|\underline{dp}\right|\right)$, if $AMB\ =\ 1$ the corresponding modality hypothesis would never be rejected.

6. Multivariate extension.

Extending the Median Preserving Spread to many dimensions can be problematic since in that space there is rarely a unique multidimensional median category (the standard $F(.) < 0.5, F(.) \geq 0.5$ condition usually defines a contour of points). Kobus and Kurek (2019) resolve the problem by defining the unique "median category" in terms of the medians of the marginal distributions but that may result in a point that is not actually on the contour[7]. However, the modal based proposal can be extended for inequality comparisons with multiple ordinal variables because, in the absence of multidimensional uniformity and multiple equal probability modes, the multidimensional modal category is usually unique. However, the extension is not straightforward because if we have D ordinal variables, then there are up to 2^D directions toward which people can become more alienated from the mode. For instance, with $D = 1$ as in the previous sections, further alienation can occur when either people initially at the mode or below move further downward toward the bottom category or people initially at the mode or above move further upward toward the top category. Meanwhile with $D = 2$, someone at the mode could move away from it in four (2^2) possible directions if the mode is not in a bivariate category defined by one or two extreme univariate categories. Indeed, they could move toward the joint bottom bivariate category, the joint top bivariate category, or either of the two extreme lopsided categories.

For ease of presentation, we present the bivariate extension here, but the results can be extended to $D > 2$. Let K_1 and K_2 be the numbers of categories of the first and second ordinal variables, respectively. Then $\boldsymbol{p} = (p_{11}, \dots, p_{K_1 K_2})$ denotes the bivariate distribution, with $0 \leq p_{ij} \leq 1$ for all $i = 1, \dots, K_1$ and $j = 1, \dots, K_2$ and $\sum_{i=1}^{K_1} \sum_{j=1}^{K_2} p_{ij}$. If there is a single bivariate mode then it will be characterised by $p_{\boldsymbol{m^o}} > p_{ij}$for all $(i, j) \in \{1, \dots, K_1\} \times \{1, \dots, K_2\}/\{\boldsymbol{m^o}\}$, where

[7] For example, consider a discrete uniform ordered categorical bivariate distribution $f(x_i, y_j)$ with K categories in each dimension (for convenience assume K is odd), each ordered by their respective subscripts so that, given a value ordering $<_V$, $i' < i'' \leftrightarrow x_{i'} <_V x_{i''}$ and similarly, $j' < j'' \leftrightarrow y_{j'} <_V y_{j''}$ and $f(x_i, y_j) = 1/K^2 \;\forall\, i, j = 1, .., K$ so that $F(x_i, y_j) = (i * j)/K^2$. In two dimensions an array of pairs (x_i^*, y_j^*) satisfy the median category condition, that surrounding lower valued outcomes have a strictly less than 0.5 chance of occurring whereas equally or higher valued outcomes have at least a 0.5 chance of occurring, i.e., $F(x_{i-1}^*, y_j^*) < 0.5$ & $0.5 < F(x_i^*, y_j^*)$ and $F(x_i^*, y_{j-1}^*) < 0.5$ & $0.5 < F(x_i^*, y_j^*)$ and $F(x_{i-1}^*, y_{j-1}^*) < 0.5$ & $0.5 < F(x_i^*, y_j^*)$. It is this array which describes the median contour in the two dimensional space of the Joint distribution. When $K = 5$ the (i, j) location of these outcomes is (3,5), (4,4) and (5,3). Kobus and Kurek (2019) define their unique multidimensional median category $KK = (x_{m_i}, y_{m_j})$ as the outcome where m_i is the median category of the marginal distribution $f(x_i) = \sum_{j=1}^{K} f(x_i, y_j)$ and m_j is the median category of the marginal distribution $f(y_j) = \sum_{i=1}^{K} f(x_i, y_j)$. In the current example $m_i = m_j = (K + 1)/2$ so that KK would not coincide with any median contour outcome (for $K = 5$ it would be outcome (3,3) with a cumulative density value of 0.36 which is surrounded from above by outcomes (4,3), (4,4), (3,4) with correspondingly higher probability values 0.48, 0.64, 0.48 which do not comply with the median contour condition).

$\boldsymbol{m^o} = \{m_1, m_2\}$ is a two-element vector containing the categories from each variable which, combined, yield the mode's joint category. [8]

Since there are up to 4 directions of alienation in the bivariate case, we need to define four cumulative sums of relative frequencies: (1) the bivariate CDF, $\boldsymbol{P} \equiv (P_{11}, \dots, P_{K_1K_2})$ such that $P_{ij} = \sum_{u=1}^{i} \sum_{v}^{j} p_{uv}$ for all $i = 1, \dots, K_1$ and $j = 1, \dots, K_2$; (2) the reverse bivariate CDF, $\boldsymbol{P'} \equiv (P'_{11}, \dots, P'_{K_1K_2})$ such that $P'_{ij} = \sum_{u=1}^{i} \sum_{v=1}^{j} p_{K_1-u+1,K_2-v+1}$ for all $i = 1, \dots, K_1$ and $j = 1, \dots, K_2$; (3) the first lopsided CDF $\boldsymbol{L} \equiv (L_{11}, \dots, L_{K_1K_2})$ such that $L_{ij} = \sum_{u=1}^{i} \sum_{v=1}^{j} p_{u,K_2-v+1}$ for all $i = 1, \dots, K_1$ and $j = 1, \dots, K_2$; and (4) the second lopsided CDF $\boldsymbol{L'} \equiv (L'_{11}, \dots, L'_{K_1K_2})$ such that $L'_{ij} = \sum_{u=1}^{i} \sum_{v=1}^{j} p_{K_1-u+1,v}$ for all $i = 1, \dots, K_1$ and $j = 1, \dots, K_2$.

Continuing with the adaptation of the notation to the bivariate case, the set of all possible distributions with K_1 categories in the first variable and K_2 in the second one is $\mathcal{S}^{K_1K_2}$. Likewise, let $\mathcal{E}^{K_1K_2} \in \mathcal{S}^{K_1K_2}$ be the set of K_1K_2 degenerate distributions, i.e. egalitarian distributions characterised by $p_{ij} = 1$ for one $(i,j) \in \{1, \dots, K_1\} \times \{1, \dots, K_2\}$, $\boldsymbol{u^{K_1K_2}}$ be a bivariate uniform distribution with K_1K_2 categories, and $I: \mathcal{S}^{K_1K_2} \to \mathbb{R}_+$ be an inequality index for two ordinal variables.

The next key step is to adapt the definition of a MOCT to the bivariate case bearing in mind the up to four possible alienation paths:

Definition 2: Bivariate mode clustering transfers (BMOCT): $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a mode clustering transfer if, for any cuadruplet of categories i, j, k, l such that either (1) $1 \leq i < j \leq m_1$ and $1 \leq k < l \leq m_2$, or (2) $m_1 \leq j < i \leq K_1$ and $m_2 \leq l < k \leq K_2$, or (3) $1 \leq i < j \leq m_1$ and $m_2 \leq l < k \leq K_2$,, or (4) $m_1 \leq j < i \leq K_1$ and $1 \leq k < l \leq m_2$: $q_{ik} = p_{ik} - \frac{1}{n}, q_{jl} = p_{jl} + \frac{1}{n}, q_{\boldsymbol{t}} = p_{\boldsymbol{t}}$ $\forall \boldsymbol{t} \neq \{i,k\} \cup \{j,l\}$, $q_{\boldsymbol{m^o}} \geq p_{\boldsymbol{m^o}}$ and $\boldsymbol{m^o}(\boldsymbol{q}) = \boldsymbol{m^o}(\boldsymbol{p})$.

Finally, we just reinstate the three axioms for the univariate case, but bearing in mind that: (1) now we have a BMOCT principle referring to definition 2, whereby $I(\boldsymbol{q}) \leq I(\boldsymbol{p})$ if $\boldsymbol{q}$ is obtained from $\boldsymbol{p}$ through a sequence of bivariate mode clustering transfers; (2) the equality principle now states that $I(\boldsymbol{p}) = 0$ if and only if $\boldsymbol{p} \in \mathcal{E}^{K_1K_2}$, for any $K_1 > 1$ $K_2 > 1$; (3) the maximality principle now states that $I(\boldsymbol{p}) = 1$ if and only if $\boldsymbol{p} = \boldsymbol{u^{K_1K_2}}$ for any $K_1 > 1$ and $K_2 > 1$.

Likewise, we will have a version of proposition 1, whereby I satisfies the BMOCT principle and the adaptations of the equality and maximality principle if and only if

$$I(\boldsymbol{p}) = \frac{h(\boldsymbol{p}) - h(\boldsymbol{q})}{h(\boldsymbol{u^{K_1K_2}}) - h(\boldsymbol{q})}$$

[8] Note that $\boldsymbol{m^o}$ does not generally necessarily coincide with $\{m_1^o, m_2^0\}$, namely the univariate modes of the two variables.

Where $h: \mathcal{S}^{K_1K_2} \to \mathbb{R}_+$ satisfies the BMOCT principle and $\boldsymbol{q} \in \mathcal{E}^{K_1K_2}$. Then an example of a class of indices from proposition 2 is:

$$I(\boldsymbol{p}) = \frac{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} w_{ij}P_{ij} + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{K_2+1-m_2} w_{ij}{P'}_{ij} + \sum_{i=1}^{m_1}\sum_{j=1}^{K_2+1-m_2} w_{ij}L_{ij} + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{m_2} w_{ij}{L'}_{ij} - 4w_{m_1m_2}p_{\boldsymbol{m^o}}}{U}$$

Where $w_{ij} > 0$ for all $(i,j) \in \{1, \dots, K_1\} \times \{1, \dots, K_2\}$, and:

$$U = \frac{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} w_{ij}ij + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{K_2+1-m_2} w_{ij}ij + \sum_{i=1}^{m_1}\sum_{j=1}^{K_2+1-m_2} w_{ij}ij + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{m_2} w_{ij}ij - 4w_{m_1m_2}}{K_1K_2}$$

And a special case is given by $w_{11} = \cdots = w_{K_1K_2}$:

$$I(\boldsymbol{p}) = \frac{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} P_{ij} + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{K_2+1-m_2} {P'}_{ij} + \sum_{i=1}^{m_1}\sum_{j=1}^{K_2+1-m_2} L_{ij} + \sum_{i=1}^{K_1+1-m_1}\sum_{j=1}^{m_2} {L'}_{ij} - 4p_{\boldsymbol{m^o}}}{U}$$

With

$$U = \frac{m_1m_2(m_1+1)(m_2+1)/4 + (K_1+1-m_1)(K_2+1-m_2)(K_1+2-m_1)(K_2+2-m_2)/4 + m_1(K_2+1-m_2)(m_1+1)(K_2+2-m_2)/4 + m_2(K_1+1-m_1)(m_2+1)(K_1+2-m_1)/4 - 4}{K_1K_2}$$

7. Empirical illustration.

An individuals' human resources, the agglomeration of their experience, and embodied human capital (their education and training,) combined with their efforts (their hours and intensity of work) are the primary drivers of their incomes and inequalities in the distribution of those resources will be a fundamental source of income inequality in incomes and an impediment to growth and social cohesion (Galor 2011, Milanovic 2011). Like many other developed economies, Canada has experienced "a Grand Gender Convergence in Incomes" (Goldin 2014) though it is not so clear that it has been the result of a convergence in human resources since gender-based distributions of human resources appear to be going in different directions (Anderson 2022). If there were a Gender convergence in human capital stocks, there should be an increasing similarity across gender in human-resource inequality within each gender. Here differences in those inequalities across the gender divide in Canada are examined over the 2006-2016 decade. Both experience and embodied human capital are at best ordered categorical variates, bereft of cardinal measure which, given the arbitrary nature of scaling and weighting, renders their combination to cardinally measure individual human resource levels awkward and questionable. Inequalities in human resource levels are thus best studied in terms of the joint ordered categorical distribution of experience and embodied human capital in a given population.

To examine the progress of human resource inequalities, we compute and compare the index of equation (2) using both the mode and the median as reference point (the latter becomes an inequality measure respecting the MEPS principle). Data on the age, education and training status of individuals has been drawn from the Census of Canada Individual Files for the years 2006 and 2016. Everyone over the age of 19 who received an income and reported age and educational status were included in the study resulting in 312405 female and 296133 male observations in 2006 and 326676 female and 283670 male observations in 2016. Experience is proxied for by age group category with 20-29, 30-39, 40-49, 50-59, 60-69 and ≥70 being the designated experience groups. Education and training embodied human capital levels are based on 5 ordered categories:- EDU1: No certificate, diploma or degree, EDU2: Secondary (high) school diploma or equivalency certificate, EDU3: Trades certificate or diploma, Certificate of Apprenticeship or Certificate of Qualification, Program of 3 months to 2 years (College, CEGEP and other non-university certificates or diplomas), EDU4: Program of more than 2 years (College, CEGEP and other non-university certificates or diplomas), University certificate or diploma below bachelor level or Bachelor's degree, and EDU5: University certificate or diploma bachelor level and above, Degree in medicine, dentistry, veterinary medicine or optometry, Master's degree or Earned doctorate.

Marginal analysis

Table 1 reports the gender-based Education and Training pdf's and cdf's $p(i)$ and $P(i)$ for the observation years 2006 and 2016 highlighting their respective Modal and Median categories.

Table 1 Education and Training Marginal Density and Cumulative Distribution Functions 2006-2016

	Marginal Density Distribution p(category i)					**Marginal Cumulative Distribution P(category i)**				
Girls	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
2006	0.21037	0.27925	0.31087#	0.05704	0.14246	0.21037	0.48962	0.80049*	0.85754	1.00000
2016	0.16808	0.30349#	0.27514	0.03985	0.21344	0.16808	0.47157	0.74671*	0.78656	1.00000
Boys.	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
2006	0.21387	0.25909	0.35024#	0.04464	0.13216	0.21387	0.47296	0.82320*	0.86784	1.00000
2016	0.20104	0.33734#	0.23106	0.03140	0.19916	0.20104	0.53838*	0.76944	0.80084	1.00000

#Modal Education *Median education

Noting that, for two distributions F_A and F_B, satisfaction of the First Order Stochastic Dominance condition $F_A(i) \leq F_B(i)\ \forall\ i = 1,..,K-1$ and $F_A(i) < F_B(i)$ some i indicates unambiguous superiority of F_A over F_B in terms of the overall level of Embodied Human Capital (at every level there is at least as great a proportion of the population at a higher level in Population A than in Population B). Thus, comparisons between observation years with respect to a particular gender and comparisons between genders within an observation year can be made.

Table 1a. First-order dominance comparisons

Comparison.	$\frac{\sum_{i=1}^{K}(F_B(i) - F_A(i))}{\sum_{i=1}^{K}\lvert F_B(i) - F_A(i)\rvert}$
Girls 2016 (A) vs 2006 (B)	1.00000
Boys 2016 (A) vs 2006 (B)	0.34255
Girls (A) vs Boys (B) 2006	0.37333
Girls (A) vs Boys (B) 2016	1.00000

Table 1a reports the dominance comparisons in terms of the sum of CDF differences $\sum_{i=1}^{K}(F_B(i) - F_A(i))$ divided by the sum of absolute CDF differences which, if positive and equal to one, suggests First Order Dominance of A over B. It indicates that, whereas Girls experienced an unequivocal improvement in embodied human capital stocks over the period the same is not true for Boys. In a similar vein there is no unequivocal superiority of one gender-based distribution over the other in 2006; however, the improvement in the girls' distribution over the period resulted in their unequivocal superiority of their embodied human capital stock distribution boys in 2016. To see how this has changed Inequality in the distribution of embodied human capital stocks in the two populations, Table 1b reports modally focused and median focused inequality measures together with their standard errors for the respective populations.

Table 1b. Inequality Measures

	2006		2016	
	Modal Focus	Median Focus	Modal Focus	Median Focus
Girls Measure (Standard Error)	0.3703 (0.0015)	0.3703 (0.00152)	0.3542 (0.0017)	0.4139 (0.0015)
Boys Measure (Standard Error)	0.3470 (0.0016)	0.3470 (0.0016)	0.3500 (0.0017)	0.3641 (0.0022)

When the Median and Modal focus categories differ, MEPS-compliant inequality measures are greater than MOCT-compliant measures within gender and female Embodied Human Capital Stock distributions appear to be more diverse than the corresponding male distributions. Note that the gender differences are significant but while the MOCT-compliant measure records a diminished difference the MEPS-compliant measure records an increase.

A similar analysis can be performed with respect to the distribution of experience in the respective populations. Table 2 reports the corresponding gender-based Experience pdf's and cdf's for the observation years 2006 and 2016.

Table 2 Experience Marginal Density and Cumulative Distribution Functions 2006-2016

	Marginal probability mass Distribution						**Marginal Cumulative Distribution**					
Girls	20-29	30-39	40-49	50-59	60-69	> 69	20-29	30-39	40-49	50-59	60-69	> 69
2006	0.16891	0.17486	0.21694#	0.17505	0.12034	0.14390	0.16891	0.34377	0.56071*	0.73576	0.85610	1.00000

2016	0.16737	0.16248	0.16473	0.19208#	0.16086	0.15248	0.16737	0.32985	0.49458	0.68666*	0.84752	1.00000
Boys.	20-29	30-39	40-49	50-59	60-69	> 69	20-29	30-39	40-49	50-59	60-69	> 69
2006	0.18066	0.17816	0.22322#	0.18716	0.11954	0.11127	0.18066	0.35881	0.58203*	0.76919	0.88873	1.00000
2016	0.19039	0.16895	0.16848	0.19588#	0.15376	0.12254	0.19039	0.35934	0.52782*	0.72370	0.87746	1.00000

#Modal Experience *Median experience

Table 2a. First-order dominance comparisons

Comparison.	$\frac{\sum_{i=1}^{K}\left(F_B(i)-F_A(i)\right)}{\sum_{i=1}^{K}\lvert F_B(i)-F_A(i)\rvert}$
Girls 2016 (A) vs 2006 (B)	1.00000
Boys 2016 (A) vs 2006 (B)	0.83073
Girls (A) vs Boys (B) 2006	1.00000
Girls (A) vs Boys (B) 2016	1.00000

Table 2a shows that both genders appear to have gained experience over the decade (a natural consequence of increasing life expectancy with survival functions increasing almost at every experience level). Girls experience levels have increased unequivocally whereas boys 2016 outcomes are only almost dominant over 2006 (Leshno and Levy 2004). Girls unequivocally have more experience than boys in both observation periods.

Table 2b. Inequality Measures

	2006		2016	
	Modal Focus	Median Focus	Modal Focus	Median Focus
Girls Measure (Standard Error)	0.3903 (0.0017)	0.3903 (0.0017)	0.4149 (0.0017)	0.4149 (0.0017)
Boys Measure (Standard Error)	0.3706 (0.0017)	0.3706 (0.0017)	0.4093 (0.0018)	0.3998 (0.0017)

Experience inequality patterns are very similar to Embodied Human Capital Patterns with Girls distributions more diverse than the corresponding Boys distributions and again MOCT-compliant measures record a greater similarity between the genders in 2016 than do the MEPS-compliant measures and both types of measures record an increase in inequality over the decade.

Multidimensional Analysis.

However, Human Resources are a complex combination of experience and embodied human capital and rather than considering marginal distributions independently and formulating a weighted view of the respective outcomes, inequalities in human resources should be examined in the context of their joint distribution in a multidimensional analysis. One of the challenges of MEPS-based inequality measurement in multidimensional environments is that the "median category" is not uniquely defined; it is in fact a contour. Kobus and Kurek (2019) resolve the problem by defining the median category in terms of the medians of the respective marginal distributions.

Table 3. Girls Education and Training and Experience Joint Mass and Cumulative Distribution Functions 2006-2016

2006	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
20-29	0.01807	0.05517	0.05427	0.00767	0.03372	0.01807	0.07325	0.12752	0.13519	0.16891
30-39	0.01822	0.03897	0.06664	0.01050	0.04054	0.03629	0.13043	0.25134	0.26951	0.34377
40-49	0.02939	0.06283	0.07952#	0.01286	0.03234	0.06568	0.22265	0.42308&	0.45411	0.56071*
50-59	0.03202	0.05403	0.05574	0.01135	0.02192	0.09770	0.30869	0.56486*	0.60724*	0.73576
60-69	0.04079	0.03211	0.03042	0.00806	0.00895	0.13849	0.38160	0.66818*	0.71863	0.85610
> 69	0.07188	0.03614	0.02429	0.00660	0.00499	0.21037	0.48962	0.80049*	0.85754	1.00000
2016	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
20-29	0.01369	0.05881	0.04362	0.00515	0.04608	0.01369	0.07250	0.11613	0.12128	0.16737
30-39	0.01354	0.03642	0.05109	0.00667	0.05477	0.02724	0.12246	0.21718	0.22900	0.32985
40-49	0.01658	0.04041	0.05440	0.00746	0.04589	0.04381	0.17945	0.32855	0.34783	0.49458
50-59	0.02715	0.06376#	0.05968	0.00796	0.03353	0.07096	0.27035	0.47914&	0.50638*	0.68666*
60-69	0.03320	0.05795	0.04075	0.00662	0.02235	0.10416	0.36150	0.61104*	0.64490	0.84752
>69	0.06392	0.04614	0.02559	0.00600	0.01082	0.16808	0.47157	0.74671*	0.78656	1.00000

Boys Education and Training and Experience Joint Mass and Cumulative Distribution Functions 2006-2016

2006	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
20-29	0.02645	0.06727	0.05650	0.00667	0.02378	0.02645	0.09372	0.15021	0.15688	0.18066
30-39	0.02443	0.04441	0.06749	0.00829	0.03353	0.05088	0.16255	0.28654	0.30150	0.35881
40-49	0.03904	0.05572	0.08735#	0.01005	0.03106	0.08992	0.25731	0.46864&	0.49366	0.58203*
50-59	0.03724	0.04727	0.06831	0.00927	0.02507	0.12715	0.34182	0.62146*	0.65575*	0.76919
60-69	0.03742	0.02450	0.04003	0.00593	0.01166	0.16457	0.40373	0.72341*	0.76362	0.88873
> 69	0.04930	0.01993	0.03056	0.00443	0.00706	0.21387	0.47296	0.82320*	0.86784	1.00000
2016	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5	EDU 1	EDU 2	EDU 3	EDU 4	EDU 5
20-29	0.02361	0.08549	0.03969	0.00414	0.03746	0.02361	0.10910	0.14879	0.15292	0.19039
30-39	0.02234	0.05254	0.04543	0.00508	0.04356	0.04595	0.18398	0.26910	0.27831	0.35934
40-49	0.02459	0.05240	0.04460	0.00613	0.04076	0.07054	0.26097	0.39068&	0.40603	0.52782*
50-59	0.04124	0.06342#	0.04940	0.00623	0.03559	0.11178	0.36563	0.54474*	0.56632*	0.72370
60-69	0.03727	0.05070	0.03334	0.00554	0.02690	0.14906	0.45360	0.66606*	0.69319	0.87746
>69	0.05198	0.03280	0.01860	0.00428	0.01488	0.20104	0.53838*	0.76944	0.80084	1.00000

#Modal Focus Category *Median category contour &Median Focus Category

Table 3 reports the joint probability masses and respective marginal distributions and identifies the Modal categories, the Median contours and the Kobus-Kurek "Median category". As will be seen the median category so defined is never on the median contour but is always below it with a cumulative mass lower than 0.5. Table 3a reports the first-order dominance comparisons, for the ALEP substitutability case, of the respective joint masses recording the gender differences in the observation years and the progress of human resource stock acquisition over the period.

Table 3a. Joint distribution comparisons

Comparison.	$\frac{\sum_{j=1}^{J}\sum_{i=1}^{K}\left(p_B(i,j)-p_A(i,j)\right)}{\sum_{j=1}^{J}\sum_{i=1}^{K}\lvert p_B(i)-p_A(i)\rvert}$
Girls 2016 (A) vs 2006 (B)	1.00000
Boys 2016 (A) vs 2006 (B)	0.61956
Girls (A) vs Boys (B) 2006	0.95910
Girls (A) vs Boys (B) 2016	1.00000

While Girls have made a clear unequivocal advance in human resource stocks boys somewhat strikingly have not so that while girls only almost first order dominate boys in 2006, they unequivocally dominate them in 2016. As for the multidimensional modally and median focused inequality measures reported in Table 3b, we see a stridently different set of results from the marginal analysis with girls having substantially higher inequality measures than boys in 2006 and lower measures in 2016 with the absolute differences widening over the period.

Table 3b. Inequality Measures

	2006		2016	
	Modal Focus	Median Focus	Modal Focus	Median Focus
Girls	0.2761 (0.0008)	0.2761 (0.0008)	0.3242 (0.0008)	0.2757 (0.0008)
Boys	0.2433 (0.0008)	0.2433 (0.0008)	0.4571 (0.0009)	0.3290 (0.0009)

8. Conclusions.

The important takeaway from these results is that the choice of reference point for inequality measurement with ordinal variables has both conceptual and empirical implications, as so does a marginal versus joint analysis. Different measures have different motivations and answering the question as to which measure should be employed will depend upon the purpose of analysis. As usual, it is a normative call. The median-focused measure is concerned with differences from a notion of “the middle” of the collection of outcomes with mass polarized at the extremes resulting in maximal inequality, whereas the modally focused measure is concerned with differences from “the most likely” complete commonality outcome with mass uniformly distributed across the outcomes resulting in maximal inequality. In joint analyses identifying the unique median joint outcome is problematic whereas identifying a unique mode is less so except in the unusual case of multiple modes with identical mass values.

As for the application, irrespective of reference points, the inequality indices nearly universally pointed to increases in educational inequality in both genders. However, the indices did not always agree in their trends. What is clear throughout is that gender-based human resource stock inequalities in Canada have changed significantly over the period.

References.

Abul Naga, R. and Yalcin, T. (2008) “Inequality measurement for ordered response health data. Journal of Health Economics”. 27 (6), 1614–1625.

Abul Naga, R. and C. Stapenhurst (2015) “Estimation of inequality indices of the cumulative distribution function”, Economics Letters, 130, 109-112.

Allison, R. A. and J. E. Foster (2004). “ Measuring health inequality using qualitative data”. Journal of Health Economics 23, 505- 24.

Anderson G.J. and T.W. Leo. (2021) "Sufficient conditions for j'th order stochastic dominance for discrete cardinal variables, and their formulae," Economics Letters, Volume 209, 110144.

Anderson G.J. and G. Yalonetzky (2022) "Mode-preserving spreads and the measurement of polarizaton with single and multiple ordinal variables."

Anderson G.J. (2022) "Has Canada's 21st Century Grand Gender Convergence Stalled? Male and Female Income and Human Resource Stock Distributions Viewed Through an Equal Opportunity Lens" Forthcoming Review of Income and Wealth.

Apouey, B. (2007) "Measuring health polarization with self-assessed health data." Health Economics, 16 (9), 875–894.

Bond T.N. and K. Lang (2019) "The Sad Truth about Happiness Scales," Journal of Political Economy 127, no. 4: 1629-1640.

Chakravarty, S. and Maharaj. B (2015) "Generalized Gini polarization indices for an ordinal dimension of human well-being", International Journal of Economic Theory, 11: 231-246.

Esteban, J and Ray, D. (1994) "On the measurement of polarization." Econometrica, 62 (4), pp. 819-51

Galor, O. (2011) Unified Growth Theory. Princeton University Press.

Gini, C. (1921). Measurement of Inequality of Incomes. The Economic Journal 31 (121), 124-126.

Goldin, C. (2014) "A Grand Gender Convergence: Its Last Chapter" American Economic Review, 104(4), 1–30

Gravel, N., Magdalou, B., Moyes, P., (2021). Ranking distributions of an ordinal variable. Economic Theory 71, 33–80.

Hammond, P.J. (1976) "Equity, Arrow's Conditions and Rawls's Difference Principle" Econometrica 44 793-803

Kahneman D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision Under Risk" Econometrica 47, 263-291.

Kobus, M. (2015). Polarization measurement for ordinal data. Journal of Economic Inequality 13, 275-297.

Kobus, M. and P. Milos (2012) "Inequality decomposition by population subgroups for ordinal data". Journal of Health Economics, 31, 15–21.

Kobus M. and Kurek R. (2019) "Multidimensional polarization for ordinal data" The Journal of Economic Inequality 17: 301-317.

Lazar, A. and J. Silber (2013) “On the cardinal measurement of health inequality when only ordinal information is available on individual health status”. Health Economics, 22, 106–113.

Leshno, M. and Levy, H. (2002). “Preferred by “All" and Preferred by “Most" Decision Makers: Almost Stochastic Dominance.” Management Science, 48(8), 1074-1085.

Mendelson, H. (1987). Quantile- preserving spread. Journal of Economic Theory 42, 334-51.

Milanovic, B. (2011). The Haves and the Have-Nots: A Brief and Idiosyncratic History of Global Inequality. New York: Basic Books.

Rothstchild, M. and Stiglitz, J. (1970). Increasing risk: I. A definition”, Journal of Economic Theory, 2(3): 225-223.

Schroder, C. and Yitzhaki, S. (2017). Revisiting the evidence for cardinal treatment of ordinal variables. European Economic Review, 92:337-358.

Stoline M.R. and H.K. Ury (1979). “Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons among Means", Technometrics, 21, 87-93

Tversky A. and D. Kahneman (1991) “Loss Aversion in Riskless Choice: A Reference-Dependent Model” The Quarterly Journal of Economics 106, 1039-1061.