On quasi approaches to inequality ordering of ordinal variables

Sandip Sarkar[1] **and Sattwik Santra**[2]

[1,2] *Centre for Studies in Social Sciences, Calcutta, India*

This version 21st October 2016

Abstract

In the context of distributions of ordinal variables and their inequality orderings, this paper provides three interesting contributions. Firstly it explores the association between certain types of transfers of population mass to the dominance rankings of distributions. Secondly it relates the inequality orderings of certain family of inequality indices to the dominance relationships of underlying distributions. Finally, the paper characterizes the class of inequality indices which allows comparison between ordinal distributions having different median categories. The same is illustrated using data on male and female educational attainments in India.

Keywords: *Ordinal Variables; Inequality; Polarization; Quasi Approach.*
JEL Classification: *D30; D63; I24; I30.*

1 Introduction

Recently, studies on income inequality have attracted considerable attention across the world (see Stiglitz, 2012; Piketty, 2014; Atkinson and Stiglitz, 2015). It is widely accepted that high inequality can have various undesirable socio-economic consequences (Stewart, 2004; Stiglitz, 2012). However, both academicians and policy makers have reached the consensus that merely focusing on inequalities based on income and wealth is not sufficient to understand disparities in human wellbeing (Sen, 1987; Sen et al., 1999). Recent developments in this field, thus, have emphasized the importance of non-income dimensions of individual well-being to formulate relevant policy prescriptions[1]. Inclusion of non-income dimensions however, entails overcoming a certain theoretical hurdle. While some of the non-income dimensions are cardinal in nature, many of these dimensions are ordinal (examples include self reported health status, happiness, educational attainments etc.) and measuring inequality for ordinal variables has become an important requirement.

In this paper we present a study on the quasi approaches to inequality ordering[2] of ordinal variables. We generalize a number of important results that establish an association between dominance orderings, inequality orderings and certain kinds of transfers of population mass. In addition to this, we propose a new methodology that enables comparisons across dominance and inequality orderings over a wider range of distributions compared to the existing literature.

In the literature on quasi approaches to inequality ordering, it has become customary to formulate indices that respond to certain characteristics of the underlying distributions (for which the indices are computed) in a particular way. The context of ordinal variables is no exception either and one fundamental requirement is that the inequality ranking of two distributions should

[1]For example, in 2011 the United Nations General Assembly adopted a unanimous resolution that states "*Recognizing that the gross domestic product indicator by nature was not designed to and does not adequately reflect the happiness and well-being of people in a country...Invites Member States to pursue the elaboration of additional measures that better capture the importance of the pursuit of happiness and well-being in development with a view to guiding their public policies;...Invites those Member States that have taken initiatives to develop new indicators...*(United Nations General Assembly Resolution 65/309)."

[2]A quasi order is a binary relation that is reflexive and transitive but not necessarily complete.

be unanimous for any choice of ordinal scale. This is however, not guaranteed if we consider the inequality indices like Gini coefficient, Theil etc. Similarly, in this context, measures of central tendency like mean, conveys no information (for illustrations, see the section "*Problems with the mean*" by Allison and Foster (2004), pp 507). In order to remedy such problems, researchers use the cumulative distribution as the domain of inequality measures and additionally consider order statistics such as the median instead of the mean, respectively (Berry and Mielke Jr, 1992; Allison and Foster, 2004; Apouey, 2007; Naga and Yalcin, 2008; Zheng, 2008; Chakravarty and Maharaj, 2015)[3]. An yet another desirable property of inequality indices for cardinal variables, relates to the sensitivity of an inequality index to mean preserving transfers that does not affect the relative rank of those affected by such transfers (Atkinson, 1970). Following our earlier logic since mean has no meaningful interpretation, it is not straightforward to extend such a result in case of ordinal variables. For ordinal variables, one of the most fundamental contributions in this direction by Allison and Foster (2004) considered median as the reference frame and viewed inequality as the spread away from the median. Building on this notion, they introduced a dominance condition based on first order inverse stochastic dominance[4] which they referred to as S dominance. On the basis of the ideas of median preserving spread, Kobus (2015) demonstrated the equivalence between finite sequences of median preserving spread and S dominance. Based on second order inverse stochastic dominance, Chakravarty and Maharaj (2015) introduced a cumulative version of the S dominance, which we refer as SS dominance for the sake of brevity. SS dominance requires that any changes in the distributions is valued more if it takes closer to the median state. While introducing SS dominance, the authors also show that SS dominance of one distribution

[3]However, following this procedure, it becomes difficult to differentiate the concepts of inequality and polarization meaningfully (See Kobus, 2015, pp 277). In the context of cardinal variables there are clear differences between inequality and polarization. For example, inequality declines as a result of rank preserving Pigou Dalton progressive transfer however, polarization may remain unchanged or even increase as a result of such transfers (Esteban and Ray, 1994). Nonetheless, this is the most widely accepted norm and many authors have developed polarization/inequality indices on the basis of that idea (see Allison and Foster, 2004; Apouey, 2007; Naga and Yalcin, 2008; Kobus and Miłoś, 2012; Chakravarty and Maharaj, 2015, for examples).

[4]The concept of stochastic dominance is well known. A related concept is inverse stochastic dominance. This was introduced in Muliere and Scarsini (1989). Note that for first and second orders, inverse stochastic dominance is equivalent to the stochastic dominance of the respective order.

over the other has a direct implication on the inequality ordering of the distributions for a particular class of additively separable inequality measures. However, additive inequality indices represent only a subset of the whole class of inequality indices and as noted in Dasgupta et al., "*... additive separability is a strong condition to impose on a general welfare function*" (Dasgupta et al., 1973, pp. 180). We show that the result of Chakravarty and Maharaj (2015) may be generalized further and for the equivalence, we do not need the additivity restriction. Thus eventually we can associate a bigger class of inequality measure with SS dominance. Furthermore, we establish an analogy between SS dominance and an arbitrary sequence of transfers which is referred to as Transfer Below and Above Median. This particular type of transfer is already existing in the literature on polarization (Apouey, 2007; Chakravarty and Maharaj, 2015).

Inequality orderings for two ordered response data following S and SS dominance is applicable only when both the distributions share a common median category. So far we have surveyed, we find that the current literature on inequality orderings for ordered response data remains almost mute on comparing distributions that do not share a common median category. The only contribution till date is by Naga and Yalcin (2010). In their paper, the authors extend the median preserving spread relation of Allison and Foster (2004). They first define a criteria for equivalence between distributions[5] and using this notion, they construct an equivalence class for any given distribution where members of the class are equivalent to each other as per the criteria but may differ in their median categories. A partial ordering based on median preserving spread is used to order two distributions belonging to two different equivalence classes (but having the same median categories) and this ordering is extended to all members of the equivalence classes (with different median categories).

In this paper, we adopt a new strategy of comparing distributions having different median categories through the use of "counterfactual" distributions. For a given class of inequality measures endowed with certain characteristics, the construction of "counterfactual distributions" to ascertain the inequality rankings of the underlying original distributions is quite common in the field of inequality ordering[6]. We show that whenever the original distributions

[5]We discuss some shortcomings of this approach in the relevant section.

[6]For example, consider comparing inequality between the distributions of two cardinal variables X and Y, where one Lorenz dominates the other. If the inequality index used, is

are not comparable because of their difference in median categories, there exist certain counterfactual versions of the distributions which have same median categories. We also show that for a family of inequality measures which follows certain well known properties, conclusions drawn on the basis of the dominance ordering of the counterfactual distributions up to the second order have similar implications on the inequality ordering of the original distributions. Thus we present a generalization over Naga and Yalcin (2010) whose partial ordering is based on the first order dominance ordering of Allison and Foster (2004). Moreover we show that, if the inequality indices are restricted to a certain family of additive inequality indices, then the relationship between the dominance ordering of the counterfactual distributions and the inequality ordering turns out to be an equivalence. We illustrate inequality orderings for ordinal variables with different median categories using data on male and female educational attainments in India.

The rest of the paper is organized as follows. The following section starts with an introduction to the notations and definitions used throughout the paper. Section 3 discusses the quasi approaches to ordering of distributions when the distributions under scrutiny have the same median category. We extend this in section 4 to incorporate distributions having different median categories and this is followed by a short empirical application in section 5. Finally, the last section concludes.

2 Preliminaries

2.1 Notations

We use the following notations throughout this paper:
Let n denote the finite number belonging to $\mathbb{Z}_{2+}$ where $\mathbb{Z}_{i+}$ denotes the set of positive integers that are greater than or equal to i.
By O we denote an n category ordinal variable that takes values from the

Lorenz consistent i.e. the inequality measure satisfies the axioms of Pigou-Dalton Transfer, Symmetry, Scale invariance and Replication Invariance (see pp 17 of Chakravarty, 2009, for definitions of these axioms) and the sample sizes are n_X and n_Y with $n_X \neq n_Y$, the way to conclude about the inequality rankings of the distributions is to construct counterfactual versions of the distributions by replicating their profiles n_Y and n_X times, respectively (which makes sample sizes of both distributions equal to $n_X n_Y$)(Foster, 1985). Since the index of inequality satisfies replication invariance axiom, the inequality orderings of the counterfactual distributions coincide with those of the original distributions of X and Y.

ordered vector $c = \{c_1, c_2, .., c_n\}$, such that $\forall i > j \iff c_i > c_j$ where "$>$" is some strict ordering relation and $i, j \in \{1, 2, ..n\}$. For the ordinal variable, the proportion of individuals belonging to the i^{th} category is denoted by $x_i \in [0, 1]$ and the vector $x = \{x_1, x_2, .., x_n\}$ denotes the Probability Distribution Function (PDF).
$X_i = \sum_{j=1}^{i} x_j$, is the cumulative proportions of individual belonging in i^{th} category such that $0 \leq X_k \leq 1 \; \forall k \in \{1, 2, .., n\}$ and $X_n \equiv 1$.
Thus the vector $X = \{X_1, X_2, .., X_n\}$ is the Cumulative Distribution Function (CDF). In the rest of this paper, we identify the distribution of an ordinal variable by its CDF.
$\mathfrak{C}^n$ denotes the set of CDFs of all n category ordinal variables.
$m(X)$ is the median category of the distribution X, such that $X_{m(X)} \geq 0.5$ and either $X_{m(X)-1} < 0.5$ or $m(X) = 1$. Note that by definition, $m(X)$ is unique. Given the CDF X, we define:
$\underline{X} = \{X_1, X_2, ..., X_{m-1}\}$ or $\underline{X}_i \equiv X_i$.
$\overline{X} = \{X_m, X_{m+1}, ..., X_n\}$ or $\overline{X}_i \equiv X_{m+i-1}$.
$\underline{X}^* = \{0.5 - X_{m-1}, 0.5 - X_{m-2}, ..., 0.5 - X_1\}$ or $\underline{X}^*_i \equiv 0.5 - \underline{X}_{m-i} \; \forall i \in \{1, 2, ..., m-1\}$.
$\overline{X}^* = \{X_m - 0.5, X_{m+1} - 0.5, ..., X_n - 0.5\}$ or $\overline{X}^*_i \equiv \overline{X}_i - 0.5 \; \forall i \in \{1, 2, .., n - m + 1\}$.
Note that, either $\overline{X}$ or $\underline{X}$ may be a null set and this would imply that the corresponding $\overline{X}^*$ or $\underline{X}^*$ to also be a null set.
Given any vector $V = \{v_1, v_2, v_3, .., v_n\} \in \mathbb{R}^n$, where $v_i \leq v_{i+1} \forall i \in \{1, 2, .., n\}$, we denote the cumulative (running) sum of V by $CS(V)$ where $CS(V) \equiv \{v_1, v_1 + v_2, v_1 + v_2 + v_3, .., \sum_{i=1}^{n} v_i\}$ i.e. $CS(V)_k \equiv \sum_{i=1}^{k} v_i$ where $k \in [1, .., n]$.
Thus $CS(A) = CS(B) \iff A = B$, also $CS(A \pm B) = CS(A) \pm CS(B)$.
For any given $\alpha \in \mathbb{R}$, the column vector of $m \geq 0$ rows with each of its elements equal to α is denoted by α_m. Note that α_0 is a null matrix. Unless stated otherwise, all vectors throughout this paper are considered as column vectors.

For two vectors $A, B \in \mathbb{R}^n$, $A \geq B \iff A_i \geq B_i \forall i \in \{1, 2, .., n\}$ and $A > B \iff A_i > B_i \forall i \in \{1, 2, .., n\}$. We say $\lim_{j \to \infty} A^j = B$ if and only if given any $\delta > 0, \exists \; n_\delta$ such that $|A^j_i - B_i| < \delta \; \forall j \geq n_\delta$ and $\forall i \in \{1, 2, ., n\}$.

2.2 Definitions

In this section we discuss some definitions that are necessary for the subsequent theoretical foundations of this paper. We begin with the concept of Inverse Stochastic Dominance (ISD). ISD has wide applications in the field of welfare economics, poverty and inequality ordering (Shorrocks, 1983; Foster and Shorrocks, 1988a,b; Allison and Foster, 2004). We revisit the definition of ISD in an ordinal setup.

Definition 1. ***Inverse Stochastic Dominance (First and Second Order)****: For any two ordered vectors $A = \{A_1, A_2, .., A_n\}$ and $B = \{B_1, B_2, .., B_n\}$ defined on $\mathbb{R}^n$ $(n \geq 2)$ such that $A_i \geq A_{i-1}$ and $B_i \geq B_{i-1} \forall i \in \{2, 3, .., n\}$.; A is said to be first order inverse stochastic dominate B, which is denoted by $A \succ_{FISD} B$ if and only if $A \geq B$ and $A \neq B$. Also A is said to be second order inverse stochastic dominate B which is denoted by $A \succ_{SISD} B$ if and only if $CS(A) \geq CS(B)$ and $A \neq B$.*

Remarks: First order ISD is a sufficient condition for second order ISD. However, the reverse is not true. Second order ISD is a cumulative version of first order ISD. Note that the notion of ISD can also be extended naturally to orders $n > 2$ (n being an integer). However, we skip these extensions since our applications are limited to only these two dominance criteria. ISD relationships of arbitrary order satisfy transitivity. For example $A \succ_{F/SISD} B$ and $B \succ_{F/SISD} C$ implies $A \succ_{F/SISD} C$ (see Chakravarty, 2009, for further reading). Furthermore, ISD relationship is also transitive in limits, i.e., if $B^1 \succ_{F/SISD} (\prec_{F/SISD}) B^2 \succ_{F/SISD} (\prec_{F/SISD}) B^3 ...$ and $\lim_{j \to \infty} B^j = A$ then $B^j \succ_{F/SISD} (\prec_{F/SISD}) A$ for any given finite $j \in \{1, 2, ...\}$[7].

Allison and Foster (2004) in their seminal contribution used ISD for inequality ordering of ordinal variables. They introduced the concept of S dominance. For any two distribution $X, Y \in \mathfrak{C}^n$; X S dominates Y when both X and Y have the same median category and there is evidence of first order ISD of X over Y below median and first order ISD of Y over X above median. Formally:

Definition 2. ***S Dominance:*** *For all $X, Y \in \mathfrak{C}^n$, and given $m(X) = m(Y) = m$, X S dominates Y (which is denoted by $X \succ_S Y$) if and only if $X \neq Y$ and exactly one of the following conditions hold:*

[7] A version of the proof is available with the authors on request.

1) $\underline{Y} \succ_{FISD} \underline{X}$ *and* $\overline{X} = \overline{Y}$ *or equivalently* $\underline{X}^* \succ_{FISD} \underline{Y}^*$ *and* $\overline{X}^* = \overline{Y}^*$
2) $\overline{X} \succ_{FISD} \overline{Y}$ *and* $\underline{X} = \underline{Y}$ *or equivalently* $\overline{X}^* \succ_{FISD} \overline{Y}^*$ *and* $\underline{X}^* = \underline{Y}^*$.
3) $\underline{Y} \succ_{FISD} \underline{X}$ *(i.e.* $\underline{X}^* \succ_{FISD} \underline{Y}^*$*) and* $\overline{X} \succ_{FISD} \overline{Y}$ *(i.e.* $\overline{X}^* \succ_{FISD} \overline{Y}^*$*).*

Remarks: The S Dominance relation is transitive, both sequentially and in limits (since the underlying *FISD* relationship is transitive in limits).

Given that X and Y have same median category, S Dominance condition $X \succ_S Y$ can equivalently be expressed as the case where a mass of population from Y moves towards the median category resulting in X. For example, consider the following CDF's: $X = \{0.1, 0.3, \mathbf{0.7}, 0.8, 1\}$ and $Y = \{0.2, 0.3, \mathbf{0.6}, 0.8, 1\}$ with their PDF's being $x = \{0.1, 0.2, \mathbf{0.4}, 0.1, 0.2\}$ and $y = \{0.2, 0.1, \mathbf{0.3}, 0.2, 0.2\}$, respectively. The median category for both the distributions in this case is 3. It is quite straightforward to check that $X \succ_S Y$ following definition 2. An alternative interpretation if we consider the PDF's, is that, X is obtained from Y following a shift of mass of 0.1 from category 1 to 2 (below median) and also from category 4 to 3 (above median). Clearly, both these movements can be considered as movements of population masses towards the median. Such movements were referred as "*Median Preserving Spread*" by Kobus (2015). Formally:

Definition 3. ***Median Preserving Spread (MPS)*** *For all* $X, Y \in \mathfrak{C}^n$ *and given* $m(X) = m(Y) = m$*, we say that* X *is obtained from* Y *by a median-preserving spread, if and only if for a given category* $i \neq m$*, there is a shift in population proportion of* δ *such that* $0 < \delta \leq y_i$ *and* $x = y + (0'_{i-1}, -\delta, \delta, 0'_{n-i-1})'$ *if* $i \leq m-1$ *or* $x = y + (0'_{i-2}, \delta, -\delta, 0'_{n-i})'$ *if* $i > m$.

Remarks: MPS may be equivalently expressed in terms of CDF. Using the notations of definition 3:
If $i \leq m-1$ then $\underline{\mathrm{X}} = \underline{\mathrm{Y}} + (0'_{i-1}, -\delta, 0'_{m-i-1})'$ and $\overline{X} = \overline{Y}$ or equivalently, $\underline{\mathrm{X}}^* = \underline{\mathrm{Y}}^* + (0'_{m-i-1}, \delta, 0'_{i-1})'$ and $\overline{X}^* = \overline{Y}^*$.
In case where $i > m$, then $\overline{X} = \overline{Y} + (0'_{i-m-1}, \delta, 0'_{n-i+1})'$ and $\underline{\mathrm{X}} = \underline{\mathrm{Y}}$ or equivalently, $\overline{X}^* = \overline{Y}^* + (0'_{i-m-1}, \delta, 0'_{n-i+1})'$ and $\underline{\mathrm{X}}^* = \underline{\mathrm{Y}}^*$.

In fact if two distributions are related through a S Dominance relation, then one distribution may be derived from the other by a sequence of MPS (Kobus, 2015).

In some cases certain distributions cannot be ranked according to S dominance criteria. In such cases, a weaker version of S Dominance is available with the literature by Chakravarty and Maharaj (2015). Based on the concept of SISD, these authors have introduced a cumulative version of S dominance. For the sake of brevity, we refer to this as *SS* dominance. Formally:

Definition 4. ***SS Dominance:*** *For all $X, Y \in \mathfrak{C}^n$, and given any $m(X)$=$m(Y)$ =m, X SS dominates Y (which is denoted by $X \succ_{SS} Y$); if and only if $X \neq Y$ and exactly one of the following conditions is satisfied:*
1) $\underline{X}^* \succ_{SISD} \underline{Y}^*$ and $\overline{X}^* = \overline{Y}^*$ (i.e. $\overline{X} = \overline{Y}$)
2) $\underline{X}^* = \underline{Y}^*$ (i.e. $\underline{X} = \underline{Y}$) and $\overline{X}^* \succ_{SISD} \overline{Y}^*$ (or equivalently $\overline{X} \succ_{SISD} \overline{Y}$).
3) $\underline{X}^* \succ_{SISD} \underline{Y}^*$ and $\overline{X}^* \succ_{SISD} \overline{Y}^*$ (i.e. $\overline{X} \succ_{SISD} \overline{Y}$).

Remarks: Note that unlike definition 2, $\underline{\mathrm{X}}^* \succ_{SISD} \underline{\mathrm{Y}}^*$ does not have an equivalent representation in terms of $\underline{\mathrm{X}}$ and $\underline{\mathrm{Y}}$ but like the S dominance relationship, *SS* dominance also exhibits transitivity. This follows from transitivity of *SISD* relationship and from the fact that $A \succ_{SISD} B$ and $B = C \implies A \succ_{SISD} C$. Like S dominance, *SS* Dominance is also transitive, in limits (because of the transitivity of the underlying *SISD* relationship under limits).

Analogous to the correspondence between S dominance and MPS, this paper establishes an association of *SS* dominance to a sequence of certain type of transfer of population mass either below or above the median category. This type of transfer of population mass has already been explored in the literature (Apouey, 2007; Chakravarty and Maharaj, 2015) but in a different context. Next, we restate this very form of transfer:

Definition 5. ***Transfer below and not below median category (TBNBM):*** *For all $X, Y \in \mathfrak{C}^n$ and given $m(X) = m(Y) = m$, X is said to be obtained from Y following TBNBM if and only if given any two categories i and j such that $1 < i \leq j < n$ and either $j + 1 \leq m$ or $i - 1 \geq m$, there are shifts of population proportions δ ($0 < \delta \leq min\{y_i, y_j\}$ if $i \neq j$ and $0 < 2\delta \leq y_i$ if $i = j$) from i to i-1 and from j to j+1. In terms of PDF: $x = y + (0'_{i-2}, \delta, -\delta, 0'_{j-i-1}, -\delta, \delta, 0'_{n-j-1})'$. The transfer is said to occur below median, if and only if the highest category involved in transfer is $j + 1 \leq m$. Similarly, we say that the transfer is not below the median category if and only if the lowest category involved in the transfer is $i - 1 \geq m$.*

Remarks: In terms of CDF, the above transfer may be represented as:
$\underline{\mathrm{X}} = \underline{\mathrm{Y}} + (0'_{i-2}, \delta, 0'_{j-i}, -\delta, 0'_{m-j-1})'$ and $\overline{X} = \overline{Y}$ if $j + 1 \leq m$ and
$\overline{X} = \overline{Y} + (0'_{i-m-1}, \delta, 0'_{j-i}, -\delta, 0'_{n-j})'$ and $\underline{\mathrm{X}} = \underline{\mathrm{Y}}$ if $i - 1 \geq m$. Equivalently,

$\underline{\mathrm{X}}^* = \underline{\mathrm{Y}}^* + (0'_{m-j-1}, \delta, 0'_{j-i}, -\delta, 0'_{i-2})'$ and $\overline{X}^* = \overline{Y}^*$ if $j+1 \leq m$ and $\overline{X}^* = \overline{Y}^* + (0'_{i-m-1}, \delta, 0'_{j-i}, -\delta, 0'_{n-j})'$ and $\underline{\mathrm{X}}^* = \underline{\mathrm{Y}}^*$ if $i-1 \geq m$.

For an illustration, consider the CDF of the following ordinal distributions $Y = \{0.05, 0.2, 0.35, 0.4, \mathbf{0.6}, 0.65, 0.8, 0.95, 1\}$ and $X = \{0.1, 0.2, 0.3, 0.4, \mathbf{0.6}, 0.7, 0.8, 0.9, 1\}$, with their PDF's being $y = \{0.05, 0.15, 0.15, 0.05, \mathbf{0.2}, 0.05, 0.15, 0.15, 0.05\}$ and $x = \{0.1, 0.1, 0.1, 0.1, \mathbf{0.2}, 0.1, 0.1, 0.1, 0.1\}$, respectively. In this case X is obtained from Y following TBNBM (definition 5). Below the median, a population mass of 0.05 is transferred from category 2 to 1 and from category 3 to 4. On the other hand, above the median, the transfer of the same mass of 0.05 takes place from categories 7 to 6 and from category 8 to 9. In this case it is quite straightforward to check that $X \succ_{SS} Y$.

So far, the definitions we have discussed characterizes different properties of distributions. Now we move on to some properties of functions that we use to characterize the family of inequality indices. The first is the notion of S concave/convex function. Formally it is defined as follows:

Definition 6. ***S concave/convex function****: A function $F : \mathbb{R}^n \to \mathbb{R}$ is called S concave if for all $X \in \mathbb{R}_n$, $F(Q_n X) \geq F(X)$ where Q_n is any bistochastic matrix of order n that is not a permutation matrix. The function is strictly S concave if $F(Q_n X) > F(X)$. The function F is S convex if and only if $-F$ is S concave.*[8]

We define a weaker version of S concave/convex functions as Piecewise S Concave/Convex functions, as:

Definition 7. ***Piecewise S (PS) Concave/Convex functions:*** *A function $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be PS concave at partition $\{n_1, n_2\}$ with $n_1, n_2 \in \mathbb{Z}_{1+}$ if and only if the following conditions are satisfied:*
1) $F(Q_{n_1} X_1, X_2) \geq F(X_1, X_2)$ whenever $X_1 \neq \{\phi\}$
2) $F(X_1, Q_{n_2} X_2) \geq F(X_1, X_2)$ whenever $X_2 \neq \{\phi\}$
where $X_1 \in \mathbb{R}^{n_1}$ and $X_2 \in \mathbb{R}^{n_2}$ $\forall$ n_1, n_2 such that $n_1 + n_2 = n$. Q_{n_1} and Q_{n_2} are bistochastic matrix of order n_1 and n_2, respectively that are not permutation matrices.
A function F is said to be strictly PS concave at the partition if and only if

[8] A square matrix is said to be bistochastic if all its entires are non-negative and the sum of all rows and columns individually equates to unity. If a bistochastic matrix has exactly one positive entry in each row and column, then it is called a permutation matrix.

strict inequality holds in both the above conditions. F is (strictly) PS convex at some partition if and only if $-F$ is (strictly) PS concave at the partition.

Note that it is quite straightforward to show that all S concave (convex) functions are PS concave (convex) for any arbitrary valid partition, however, the converse is not true.

Given the notations and definitions we are now in a position to introduce the various results of the paper to which we now turn.

3 Quasi approaches when median categories are same

In this section we formally establish the relationship between the inequality orderings of a class of inequality indices having certain functional characteristics and the dominance relationships (S and SS) of the underlying distributions.

The conceptual framework of this paper characterizes inequality of an ordinal variable as the distance between an observed distribution to that of a bipolar distribution (formalized by Apouey, 2007). A distribution is said to be bipolar if half of the individuals belong to the first category and the rest to the last category. Hence, the CDF of a bipolar distribution assumes the form $X^{BP} = \{0.5, 0.5, .., 0.5, 1\}$. The distance between an observed distribution $X = \{X_1, X_2, .., X_n\} \in \mathfrak{C}^n$ to that of X^{BP} is given by $\{|0.5 - X_1|, |0.5 - X_2|, ., |0.5 - X_n|\} \equiv \{\underline{X}^*, \overline{X}^*\}$. Consequently, the domain of the inequality measures considered here, is given by $\mathbb{D}^n \equiv [0, 0.5]^n$. Thus we denote the inequality index defined on the distribution X by $I(\underline{X}^*, \overline{X}^*)$ where $I : \mathbb{D}^n \longrightarrow \mathbb{R}$. For the purpose of analytical tractability we impose an assumption on the index function *I()* as follows:

Assumption 3.1. *The inequality index I, is bounded between m and M, where $m, M \in \mathbb{R}$ and $M > m$.*[9]

The class of all inequality indices satisfying assumption 3.1 is denoted by $\mathbb{I}$, i.e., $\mathbb{I} = \{I | I : \mathbb{D}^n \longrightarrow \mathbb{R}$ and $m \leq I() \leq M\}$.

[9]Ideally one should restrict $m = 0$, which ensures that inequality index always takes non negative values. Nevertheless, our consideration is more general.

With the above formalizations, we first establish the equivalence of inequality orderings for the family of strictly monotonically decreasing inequality indices with the S dominance ordering of the underlying distributions.

Theorem 3.1. *For all $X, Y \in \mathfrak{C}^n$, $X \neq Y$ and $m(X) = m(Y)$. The following statements are equivalent:*
1) $X \succ_S Y$
2) X is obtained from Y following arbitrary sequences of MPS.
3) $I(\underline{X}^*, \overline{X}^*) < I(\underline{Y}^*, \overline{Y}^*)$ $\forall I \in \mathbb{I}$ such that I is strictly monotonically decreasing in its arguments.

Proof: The equivalence between statements 1 and 2 has already been established by Kobus (2015) for finite sequences of *MPS*. A finite sequence of MPS can be readily expressed as an infinite sequence of *MPS* by splitting the amount of population transfer from any category into (say) an infinite AP series that adds up to the amount of population transfer from the category and redefining each element of the AP series as an MPS. Thus, that S dominance implies an arbitrary sequence of *MPS* is trivially true. So here we first show that infinite sequence of MPS implies S dominance.
$2 \implies 1$
An *MPS* from any category $i \neq m$ and $i \in \{1, 2, ..., n\}$ can be expressed as: $\underline{X}^* = \underline{Y}^* + (0'_{m-i-1}, \delta, 0'_{i-1})'$ and $\overline{X}^* = \overline{Y}^*$ if $i < m$ or, $\overline{X}^* = \overline{Y}^* + (0'_{i-m-1}, \delta, 0'_{n-i+1})'$ and $\underline{X}^* = \underline{Y}^*$ if $i > m$ where $\delta > 0$. Let the sequence of *MPS* be denoted by $MPS^1, MPS^2, ...$ such that X is obtained from Y by this sequence of *MPS*. Recursively define $Z^0 = Y$ and Z^s as the distribution obtained from Z^{s-1} by MPS^s. Note that Z^s is a valid distribution for all $s = \{1, 2, ...\}$ and $\lim_{s \to \infty} Z^s = X$. From the definition of S dominance (definition 2) if $X \succ_S Y$ then the relationship between X and Y can be expressed as:
$\underline{X}^* = \underline{Y}^* + \underline{\epsilon}$ and $\overline{X}^* = \overline{Y}^* + \bar{\epsilon}$
where $\underline{\epsilon}_i \geq 0 \forall i \in \{1, 2, .., m-1\}$ and $\bar{\epsilon}_j \geq 0$ $\forall j \in \{1, 2, .., n-m+1\}$ with strict inequality for at least one i or j. Also $\bar{\epsilon}_{n-m+1} = 0$ and $\bar{\epsilon}_j$ and $\underline{\epsilon}_i$ are constrained such that X is a legitimate distribution. Thus, $Z^s \succ_S Z^{s-1}$ and since S dominance is transitive in limits (see remarks to definition 2), $X \succ_S Y$ holds.

Next we establish the equivalence between 1 and 3 in order to complete the proof. For the proof, we define the following vectors:

$$\tilde{X} = \left(\underline{X}^{*\prime}, \overline{X}^{*\prime}\right)' \text{ and } \tilde{Y} = \left(\underline{Y}^{*\prime}, \overline{Y}^{*\prime}\right)'$$

$1 \implies 3$

Following the remarks to Definition 2, it is quite straightforward to show that $\exists\ \epsilon$ where $\epsilon_i \geq 0\ \forall i \in \{1, 2, .., n\}$ such that $X \succ_S Y \iff \tilde{X} = \tilde{Y} + \epsilon$, with strict inequality for at least one i. Hence for any strictly monotonically decreasing function $I()$ we have $I(\underline{\mathrm{X}}^*, \overline{X}^*) < I(\underline{\mathrm{Y}}^*, \overline{Y}^*)$.

$3 \implies 1$

We prove this by contradiction. We begin with the assumption that $X \not\succ_S Y$. Since $\tilde{X}$ and $\tilde{Y}$ are of the same order, they can be written as $\tilde{X} = \tilde{Y} + \epsilon$. Following the remarks to definition 2, $X \neq Y$ and $X \not\succ_S Y \implies \exists \epsilon_k$ such that $\epsilon_k < 0$ for some $k \in \{1, 2, .., n\}$. Now, for any $Z \in \mathfrak{C}^n$ with $m(Z) = m$, define $I(\mathrm{Z})$ as follows:

$$I(\underline{Z}^*, \overline{Z}^*) = -\sum_{i=1}^{m-1} \alpha_i \underline{Z}_i^* - \sum_{j=1}^{n-m+1} \beta_j \overline{Z}_j^* \tag{1}$$

where $\alpha_i, \beta_j > 0 \forall i, j$ and $\sum_{i=1}^{m-1} \alpha_i + \sum_{j=1}^{n-m+1} \beta_j = 1$. Note that the restrictions on α and β ensure that $I()$ is bounded and strictly monotonically decreasing.

Given $\epsilon_k < 0$, we choose the corresponding α or β (i.e. α_k if $k \leq m-1$ or β_{k-m+1} if $k \geq m$) high enough in order to get a contradiction i.e. $I(\underline{\mathrm{X}}^*, \overline{X}^*) > I(\underline{\mathrm{Y}}^*, \overline{Y}^*)$.

Q.E.D.

We denote the class of all strictly monotonically decreasing inequality indices satisfying assumption 3.1 by $\mathbb{I}_1$ i.e.: $\mathbb{I}_1 = \{I() \in \mathbb{I} | I(..., a, ...) \gtreqless I(..., b, ...) \iff a \lesseqgtr b\}$.

We now move on to characterize the inequality orderings of the class of inequality indices that are strictly decreasing and strictly PS convex when the underlying distributions can be ordered by SS dominance. A similar result was introduced by Chakravarty and Maharaj (2015) for class of additive inequality indices. We relax the additivity assumption and establish our result for a more general class of inequality indices. Formally:

Theorem 3.2. *For all $X, Y \in \mathfrak{C}^n$, $X \neq Y$ and $m(X) = m(Y)$. The following conditions are equivalent:*
1) $X \succ_{SS} Y$
2) $I(\underline{X}^*, \overline{X}^*) < I(\underline{Y}^*, \overline{Y}^*)\ \forall I \in \mathbb{I}_1$ such that I is strictly PS convex at partition $\{m-1, n-m+1\}$.

Proof:

$1 \implies 2$

Given $X \succ_{SS} Y \iff$ exactly one of the following cases holds (following Definition 4):

Case 1) $\underline{X}^* \succ_{SISD} \underline{Y}^*$ and $\overline{X}^* = \overline{Y}^*$

Case 2) $\underline{X}^* = \underline{Y}^*$ and $\overline{X}^* \succ_{SISD} \overline{Y}^*$.

Case 3) $\underline{X}^* \succ_{SISD} \underline{Y}^*$ and $\overline{X}^* \succ_{SISD} \overline{Y}^*$.

We provide a proof only for Case 3, the rest of the cases can be proved following similar reasoning.

Following Marshall and Olkin's theorem (lemma 7.1) we can write that there exist two bi-stochastic matrix Q_1 and Q_2 (of appropriate orders) that are not a permutation matrices, such that:

$$\underline{X}^* \geq Q_1 \underline{Y}^* \; and \; \overline{X}^* \geq Q_2 \overline{Y}^* \tag{2}$$

Choose -$I()$ as any strict monotonically increasing and a strict PS concave function at partition $\{m-1, n-m+1\}$. Following the strict monotonicity of -$I()$ we can write:

$$-I(\underline{X}^*, \overline{X}^*) \geq -I(Q_1 \underline{Y}^*, \overline{X}^*) \geq -I(Q_1 \underline{Y}^*, Q_2 \overline{Y}^*) \tag{3}$$

Furthermore, since -$I()$ is strict PS concave at partition $\{m-1, n-m+1\}$, following Definition 7, we can also write:

$$-I(Q_1 \underline{Y}^*, Q_2 \overline{Y}^*) > -I(Q_1 \underline{Y}^*, \overline{Y}^*) > -I(\underline{Y}^*, \overline{Y}^*) \tag{4}$$

Combining 3 and 4 we obtain $-I(\underline{X}^*, \overline{X}^*) > -I(\underline{Y}^*, \overline{Y}^*) \implies I(\underline{X}^*, \overline{X}^*) < I(\underline{Y}^*, \overline{Y}^*)$ where $I()$ is any strict monotonically decreasing and PS convex function at partition $\{m-1, n-m+1\}$.

$2 \implies 1$

We prove this by contradiction. We assume that $X \not\succ_{SS} Y$. Given $X \neq Y$ exactly one of the following conditions hold:

Case 1) $\underline{X}^* \not\succ_{SISD} \underline{Y}^*$ and $\overline{X}^* \not\succ_{SISD} \overline{Y}^*$.

Case 2) $\underline{X}^* \succ_{SISD} \underline{Y}^*$ and $\overline{X}^* \not\succ_{SISD} \overline{Y}^*$.

Case 3) $\underline{X}^* \not\succ_{SISD} \underline{Y}^*$ and $\overline{X}^* \succ_{SISD} \overline{Y}^*$.

Case 4) $\underline{X}^* \not\succ_{SISD} \underline{Y}^*$ and $\overline{X}^* = \overline{Y}^*$.

Case 5) $\underline{X}^* = \underline{Y}^*$ and $\overline{X}^* \not\succ_{SISD} \overline{Y}^*$.

We derive contradictions only for Case 3. Similar logic yields contradiction for rest of the cases.

Following lemma 7.3, we have $\underline{X}^* \nsucc_{SISD} \underline{Y}^* \implies \exists u$ such that $\sum_{i=1}^{m-1} u(\underline{X}_i^*) < \sum_{i=1}^{m-1} u(\underline{Y}_i^*)$, where $u()$ is strictly increasing and strictly concave. On the other hand, since $\overline{X}^* \succ_{SISD} \overline{Y}^*$ for any strictly concave and strictly increasing function $v(.)$, we have $\sum_{j=1}^{n-m+1} v(\overline{X}_j^*) > \sum_{j=1}^{n-m+1} v(\overline{Y}_j^*)$ (following Marshall and Olkin's theorem (lemma 7.1)).

For all $Z \in \mathfrak{C}^n$ such that $m(Z) = m$, define the function:

$$I(\underline{Z}^*, \overline{Z}^*) = \theta \sum_{i=1}^{m-1} -u(\underline{Z}_i^*) + (1-\theta) \sum_{j=1}^{n-m+1} -v(\overline{Z}_j^*) \tag{5}$$

where $0 < \theta < 1$.

Define $I_1(\underline{Z}^*) \equiv \sum_{i=1}^{m-1} u(\underline{Z}_i^*)$ and $I_2(\overline{Z}^*) \equiv \sum_{j=1}^{n-m+1} v(\overline{Z}_j^*)$. Following lemma 7.2, $I_1(.)$ and $I_2(.)$ are strict monotonically increasing and strict S concave functions. Since, $I() \equiv -I_1() - I_2()$, following definition 7, it is straightforward to show that $I()$ is strictly decreasing and strictly PS convex function at partition $\{m-1, n-m+1\}$. In order to get a contradiction (i.e., $I(\underline{X}^*, \overline{X}^*) > I(\underline{Y}^*, \overline{Y}^*)$) we choose a value of θ that is sufficiently close to 1.
Q.E.D.

We denote the class of all strictly monotonically decreasing and strictly PS convex inequality indices at partition $\{n_1, n_2\}$ satisfying assumption 3.1 by $\mathbb{I}_1$ i.e.: $\mathbb{I}_{2\{n_1,n_2\}} = \{I() \in \mathbb{I}_1 | I()$ is strictly PS convex with partition $\{n_1, n_2\}\}$. The relationship between S dominance and MPS is already established in the literature by Kobus (2015). However, an association between certain types of transfers and *SS* dominance, has not been formulated in the literature so far. In the next theorem we show the equivalence between *SS* dominance with MPS and TBNBM.

Theorem 3.3. *For all $X, Y \in \mathfrak{C}^n$ and $m(X) = m(Y)$. The following conditions are equivalent:*
1) $X \succ_{SS} Y$
2) X is limit of Y following arbitrary sequence of TBNBM and/or MPS.

Proof:
If the number of categories both below and above the median is less than

two then SS dominance is analytically equivalent to S dominance. In such cases $TBNBM$ is not applicable and the equivalence between SS dominance to a sequence of MPS is already provided in theorem 3.1.

From the definition of MPS, if X is obtained from Y through a MPS then the relationship between X and Y can be written in terms of CS() as:
i) $CS(\underline{\mathrm{X}}^*) = CS(\underline{\mathrm{Y}}^*) + (0'_{m-i-1}, \delta'_i)'$ and $CS(\overline{X}^*) = CS(\overline{Y}^*)$.
ii) $CS(\overline{X}^*) = CS(\overline{Y}^*) + (0'_{i-m-1}, \delta'_{n-i+2})'$ and $CS(\underline{\mathrm{X}}^*) = CS(\underline{\mathrm{Y}}^*)$.
where $0 < \delta \leq y_i$ denotes the shift of population proportion from category $i \neq m$ in accordance to definition 3.

From the definition of $TBNBM$ (definition 5) we know that if X is obtained from Y by $TBNBM$ such that the two categories involved in the transfer, (say i and j) are equal (i.e. $i = j$) where either $1 < i < m$ or $m < i < n$, and the shifts of population proportions are given by δ ($0 < 2\delta \leq y_i$) then, in terms of CS():
i) $CS(\underline{\mathrm{X}}^*) = CS(\underline{\mathrm{Y}}^*) + (0'_{m-i-1}, \delta, 0'_{i-1})'$ and $CS(\overline{X}^*) = CS(\overline{Y}^*)$ if $i + 1 \leq m$.
ii) $CS(\overline{X}^*) = CS(\overline{Y}^*) + (0'_{i-m-1}, \delta, 0'_{n-i+1})'$ and $CS(\underline{\mathrm{X}}^*) = CS(\underline{\mathrm{Y}}^*)$ if $i - 1 \geq m$.

From the definition of SS dominance (definition 4) if $X \succ_{SS} Y$ then the relationship between X and Y can be written in terms of CS() as:
$CS(\underline{\mathrm{X}}^*) = CS(\underline{\mathrm{Y}}^*) + \underline{\epsilon}$ and $CS(\overline{X}^*) = CS(\overline{Y}^*) + \bar{\epsilon}$.
where $\underline{\epsilon}_i \geq 0 \forall i \in \{1, 2, .., m-1\}$ and $\bar{\epsilon}_j \geq 0 \ \forall j \in \{1, 2, .., n-m+1\}$ with strict inequality for at least one i or j. Also $\bar{\epsilon}_{n-m+1} = \bar{\epsilon}_{n-m}$ and $\bar{\epsilon}_j$ and $\underline{\epsilon}_i$ are constrained such that X is a legitimate distribution.

$2 \implies 1$
Let the sequences of $TBNBM$ or MPS be denoted by $TBNBM/MPS^1$, $TBNBM/MPS^2, \ldots$. Define $Z^0 = Y$ and let Z^s be obtained from Z_{s-1} applying $TBNBM/MPS^s$ on Z^{s-1}. From the representation of $TBNBM$ or MPS in terms of CS we can write $CS(Z^s) = CS(Z^{s-1}) + \epsilon$ such that $\epsilon \geq 0_n$. Clearly, $Z^s \succ_{SS} Z^{s-1}$ and $\lim_{s \to \infty} Z^s = X$. From remarks to definition 4 thus, $X \succ_{SS} Y$.

$1 \implies 2$

$$X \succ_{SS} Y \iff \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-2} \\ x_{m-1} \\ x_m \\ x_{m+1} \\ x_{m+2} \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-2} \\ y_{m-1} \\ y_m \\ y_{m+1} \\ y_{m+2} \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} + \begin{bmatrix} -\underline{\epsilon}_{m-1} + \underline{\epsilon}_{m-2} \\ -2\underline{\epsilon}_{m-2} + \underline{\epsilon}_{m-1} + \underline{\epsilon}_{m-3} \\ \vdots \\ -2\underline{\epsilon}_2 + \underline{\epsilon}_1 + \underline{\epsilon}_3 \\ -2\underline{\epsilon}_1 + \underline{\epsilon}_2 \\ \underline{\epsilon}_1 + \bar{\epsilon}_1 \\ -2\bar{\epsilon}_1 + \bar{\epsilon}_2 \\ -2\bar{\epsilon}_2 + \bar{\epsilon}_1 + \bar{\epsilon}_3 \\ \vdots \\ -2\bar{\epsilon}_{n-m-1} + \bar{\epsilon}_{n-m-2} + \bar{\epsilon}_{n-m} \\ -\bar{\epsilon}_{n-m} + \bar{\epsilon}_{n-m-1}. \end{bmatrix},$$

where $\underline{\epsilon}_i \geq 0$ and $\bar{\epsilon}_j \geq 0$ with strict inequality for at least one i or j. Also, $\bar{\epsilon}_j$ and $\underline{\epsilon}_i$ are constrained such that $x \geq 0$. We write this more succinctly as: $x = y + T\eta$, where

$$T \equiv \begin{bmatrix} -1 & 1 & 0_{n-2} \\ & A & \\ 0_{n-2} & 1 & -1 \end{bmatrix} \text{ and } \eta_i = \begin{cases} \underline{\epsilon}_{m-i} & \text{if } i < m; \\ 0 & \text{if } i = m; \\ \bar{\epsilon}_{i-m} & \text{if } i > m. \end{cases}$$

and A=$[a_{ij}]_{(n-2)\times n}$ with

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \text{ or } i = j-2; \\ -2 & \text{if } i = j-1; \\ 0 & else \end{cases}$$

Define $z^0 = y,\ r^0 = \eta$ and for all $s = \{1, 2, ...\}$,
$z^s = z^{s-1} + T\tau^s$
$\tau^s = min(0.5\ z^{s-1}, r^{s-1})$
$r^s = r^{s-1} - \tau^s$.
Evidently, $x = z^s + Tr^s$, $r^{s-1} \geq r^s \geq 0$, $\tau^s \leq z^{s-1}$ and $\tau^s \leq r^{s-1}$, $\forall s = \{1, 2, ...\}$. Given the definition of τ^s, $T\tau^s$ denotes a set of *TBNBM* and/or *MPS* on z^{s-1}. This is because $2\tau_i^s \leq z_i^{s-1}$ $\forall i \in \{1, 2, .., n\}$ and $\forall s \in \{1, 2, ...\}$. So, the sequential addition of matrices given by:

$$T \begin{bmatrix} 0_{i-1} \\ \tau_i^s \\ 0_{n-i} \end{bmatrix}$$

to z^{s-1} $\forall i \in \{1,2,..,n\}$ represents a TBNBM/MPS(if $i=1$ or $i=n$) from definitions 5 and 3, respectively, in the sense that the distribution obtained after each such addition remains a legitimate distribution. In order to complete the proof we need to show that $\lim_{s\to\infty} r^s = 0_n$. We prove this by contradiction. Note that $\lim_{s\to\infty} r^s$ must exist, since $r^{s-1} \geq r^s \geq 0$ i.e., the sequence is non-increasing and bounded below (since every bounded non-increasing sequence has a limit point (see theorem 3.14, Rudin, 1971, pp 55). Since r^s is a convergent sequence, it must be Cauchy convergent (theorem 3.11, Rudin, 1971, pp 53) i.e. given any $\delta > 0$ $\exists s'$ such that $r^k - r^l < \delta$ $\forall k, l > s'$ and $l > k$. This implies, $\lim_{s\to\infty} \tau^s = 0$. If $\lim_{s\to\infty} r^s = r^L$ where $r^L \geq 0_n$, $r^L \neq 0_n$. Let $r^L_{i'}$ denote the maximal element of r^L with i' denoting the index of the element, such that either $r^L_{i'+1} < r^L_{i'}$ or $r^L_{i'-1} < r^L_{i'}$ is true. Such $r^L_{i'}$ always exists since $r^L_m = 0$. Thus, if $i' = 1$ or $i' = n$ the i' element of Tr^L is given by $-r^L_{i'} + r^L_{i'+1}$ or $-r^L_{i'} + r^L_{i'-1}$, respectively and $-2r^L_{i'} + r^L_{i'+1} + r^L_{i'-1}$ otherwise. Since $r^L_{i'}$ is the maximum with at least one of $r^L_{i'-1}$ or $r^L_{i'+1}$ strictly less than the maximum value, this implies that the i' element of Tr^L is strictly less than 0. Furthermore, $\lim_{s\to\infty} \tau^s = 0$ and $\lim_{s\to\infty} r^s_{i'} > 0 \implies \lim_{s\to\infty} z^s_{i'} = 0$. But this implies the i' element of x is negative, since $x = z^s + Tr^s$ $\forall s \in \{1,2,...\}$: a contradiction.
Q.E.D.

4 Quasi approaches when median categories are different

So far our analysis on inequality ordering of ordinal variables, is restricted to the case where the variables have same median category. In this section we relax this assumption. When the median category of two distributions do not coincide, we construct counterfactual versions of the original distributions (henceforth, we refer to the original distributions as the "base distributions") such that their median categories coincide. Furthermore, we show that for a class of indices satisfying some given properties, the dominance conditions (S/SS) of the counterfactual distributions have a direct implication on the inequality ordering of the original distributions. For the creation of the counterfactual distributions we follow the Slide Invariance (SI) and

Zero Frequency Independence (ZFI) properties meted out in the literature on polarization.

The first property *SI*, is adopted from Apouey (2007). It is applicable to distributions where the first or the last category has zero mass. Formally:

Property 4.1. ***Slide Invariance (SI)****: For all $X^l, X^r, Y^l, Y^r \in \mathfrak{C}^n$, if X^l, X^r and Y^l, Y^r are related as $X^r = \{0'_k, X^r_{k+1}, X^r_{k+2}, .., X^r_{n-1}, 1\}$, $X^l = \{X^l_{k+1}, X^l_{k+2}, .., X^l_{n-k-1}, 1'_{k+1}\}$, $Y^r = \{0'_k, Y^r_{k+1}, Y^r_{k+2}, .., Y^r_{n-1}, 1\}$ and $Y^l = \{Y^l_{k+1}, Y^l_{k+2}, .., Y^l_{n-k-1}, 1'_{k+1}\}$ for all $k \in \{1, 2, ..., n-1\}$, then $X^l(Y^l)$ is said to be obtained from $X^r(Y^r)$ by k slides to the left and likewise $X^r(Y^r)$ is said to be obtained from $X^l(Y^l)$, respectively by k slides to the right. Any inequality index $I()$, is said to be slide invariant if and only if $I(\underline{X}^{l*}, \overline{X}^{l*}) \mathfrak{R}\, I(\underline{Y}^{l*}, \overline{Y}^{l*}) \iff I(\underline{X}^{r*}, \overline{X}^{r*}) \mathfrak{R}\, I(\underline{Y}^{r*}, \overline{Y}^{r*})$.*

The second property (ZFI) states that any deletion or inclusion of a subgroup with zero population share does not change the level of polarization (Chakravarty and Maharaj, 2012; Chakravarty, 2015). We consider a restricted version of this property following which, inclusion(omission) of categories with zero population mass either at the bottom or at the top of two distribution does not change the inequality rankings of the distributions. We refer to this property as "*Insensitive to Terminal Unpopulated Categories*" (ITUC). Note that although ITUC is a technical property, a majority of inequality indices established in the literature on inequality ordering of ordinal variables, satisfy this property. Letting $\mathfrak{R}$ denote any one of the relationships $>$, $=$ and $<$, we can formally state ITUC as:

Property 4.2. ***Insensitive to Terminal Unpopulated Categories (IT UC):*** *For all $X, Y \in \mathfrak{C}^n$ where $k_1, k_2 \in \mathbb{Z}_{0+}$ with at least one of k_1 or k_2 strictly positive, two inequality measures $I^n_1()$ and $I^{n+k_1+k_2}_2()$ are said to be associated through Insensitive to Terminal Unpopulated Categories if and only if $I^n_1(\underline{X}^*, \overline{X}^*) \ \mathfrak{R}\ I^n_1(\underline{Y}^*, \overline{Y}^*) \iff I^{n+k_1+k_2}_2((0.5'_{k_1}\ \underline{X}^{*\prime})', (\overline{X}^{*\prime}\ 0.5'_{k_2})') \ \mathfrak{R}\ I^{n+k_1+k_2}_2((0.5'_{k_1}\ \underline{Y}^{*\prime})', (\overline{Y}^{*\prime}\ 0.5'_{k_2})')$.*

For the sake of simplicity we say that X^F is obtained from X by addition of unpopulated categories (ATUC) if and only if $X^F = \{0'_k, X'\}$ or $X^F = \{X', 1'_k\}$ $\forall k \in \mathbb{Z}_{1+}$.

In the analysis of the previous sections, the number of categories of the ordinal variables in question were fixed at n. However, the construction of the counterfactual distributions by ATUC, requires a higher number of categories compared to their base distributions. So in this section, whenever required, we explicitly indicate the dimension of domain as a superscript to the inequality indices as well as the class of the indices (i.e., instead of the notation $I()$ for an inequality index, we use the notation $I^n()$ and similarly instead of $\mathbb{I}$, we use $\mathbb{I}^n$).

Given the above definitions, we now discuss the main result of this section. For this, we define $\mathbb{I}_3 = \{I() | I() \in \mathbb{I}_1$ and $I()$ satisfies SI$\}$, $\mathbb{I}_{4\{m_1,m_2\}} = \{I() | I() \in \mathbb{I}_{2\{m_1,m_2\}}$ and $I()$ satisfies SI$\}$ and $\mathbb{I}_{5\{m_1,m_2\}}^{\{n,n+k\}} = \{$The tuples$(I_1^n(), I_2^{n+k}()) | I_1^n() \in \mathbb{I}_1^n$ and $I_2^{n+k}() \in \mathbb{I}_{4\{m_1,m_2\}}^{n+k}$ where $I_1^n()$ and $I_2^{n+k}()$ are linked through ITUC$\}$.

Now suppose the median of $X \in \mathfrak{C}^n$ is $m+k$ and that of $Y \in \mathfrak{C}^n$ is m; where k is a strictly positive integer. In such cases two possibilities exist. The first possibility is that there exists counterfactual version of X and Y obtained by slide such that the median categories of the counterfactuals coincide [10]. The second possibility is that either the distributions do not allow slide or that slide is not able to equalize the median categories of the counterfactuals. In this case, addition of a finite number of unpopulated categories to both X and Y followed by slide of at least one of the distributions should be able to generate counterfactuals with same median categories. First we provide some results for the first possibility relating the S/SS dominance ordering of the counterfactual distributions to the inequality ordering of the original distributions for the family of inequality indices which belongs to $\mathbb{I}_3$ and $\mathbb{I}_4$. Formally:

Theorem 4.1. *For all $X, Y \in \mathfrak{C}^n$ such that $m(X) = m+k$, $m(Y) = m$, and $k \in \{1, 2, ..., n-m\}$, let $X^S \in \mathfrak{C}^n$ be obtained from X by k_1 slides to the left and $Y^S \in \mathfrak{C}^n$ be obtained from Y by k_2 slides to the right ($k_1, k_2 \in \mathbb{Z}_{0+}$, $k_1 + k_2 = k$) such that $m(X^S) = m(Y^S)$. Then the following conditions hold:*
1) $X^S \succ_S (\prec_S)\ Y^S \implies I(\underline{X}^*, \overline{X}^*) < (>)\ I(\underline{Y}^*, \overline{Y}^*)\ \forall I \in \mathbb{I}_3$.
2) $X^S \succ_{SS} (\prec_{SS})\ Y^S \implies I(\underline{X}^*, \overline{X}^*) < (>)\ I(\underline{Y}^*, \overline{Y}^*)\ \forall I \in \mathbb{I}_{4\{m-1,n-m+1\}}$.

Proof: We provide the proof for statement 1. Statement 2 can be proved using similar logic by replacing S dominance with SS dominance, using theorem 3.2 in place of theorem 3.1 and considering $\mathbb{I}_{4\{m-1,n-m+1\}}$ instead of $\mathbb{I}_3$. Following theorem 3.1 we can write $X^S \succ_S (\prec_S)\ Y^S \implies I(\underline{X}^{S*}, \overline{X}^{S*}) < (>$

[10] For example, let $X = \{0, 0.1, \mathbf{0.6}, 1\}$ and $Y = \{0.3, \mathbf{0.55}, 0.7, 1\}$. In this case $m(X) = 3$, $m(Y) = 2$. The counterfactual distributions are $X^S = \{0.1, \mathbf{0.6}, 1, 1\}$ and Y.

) $I(\underline{Y}^{S*}, \overline{Y}^{S*})$ $\forall I() \in \mathbb{I}_3 \because I() \in \mathbb{I}_3 \implies I() \in \mathbb{I}_1$ and following SI, we can write $I(\underline{X}^{S*}, \overline{X}^{S*}) < (>) \; I(\underline{Y}^{S*}, \overline{Y}^{S*}) \iff I(\underline{X}^{*}, \overline{X}^{*}) < (>) \; I(\underline{Y}^{*}, \overline{Y}^{*})$.
Q.E.D.

We now move on to the second possibility where a slide of the distribution/s is not able to produce the counterfactual with same median categories. However, it is quite straightforward to show that in such cases counterfactual distributions with same median categories can always be produced by ATUC and slide. To illustrate this consider the distributions $X, Y \in \mathfrak{C}^n$ such that $m(X) = m + k$ and $m(Y) = m$ where $k \in \{1, 2, ..., n - m\}$. Denote by sl^X_{max} the maximum number of slides towards the left possible for X i.e. $sl^X_{max} = \{\#i \,|X_i = 0\}$ and likewise, $sr^Y_{max} = \{\#i \,|Y_i = 1 \text{ and } i < n\}$ indicates the maximum number of slides towards the right possible for Y (note that either or both these values may be zero). Clearly $m + k + sl^X_{max} > m + sr^Y_{max}$ (otherwise it would be possible to construct counterfactuals from X and Y by slide alone such that the counterfactuals have same median categories). Define k' as $k' = (m + k - sl^X_{max}) - (m + sr^Y_{max})$. The construction of the counterfactual distributions with same median categories can be done as follows. First we introduce k' unpopulated categories before the first category of both X and Y to obtain X^F and Y^F, respectively. Then we slide X^F to the left by $k' + sl^X_{max}$ and slide Y^F to the right by sr^Y_{max} to obtain X^{FS} and Y^{FS}, respectively[11]. It is easy to check that $m(X^{FS}) = m(Y^{FS}) = m + sr^Y_{max}$. Alternatively, we may also introduce k' unpopulated categories after the last category to obtain X^F and Y^F, respectively. Then we slide X^F to the left by sl^X_{max} and slide Y^F to the right by $k' + sr^Y_{max}$ to obtain X^{FS} and Y^{FS}, respectively, in which case $m(X^{FS}) = m(Y^{FS}) = m + k - sl^{max}_X$. The constructed counterfactual distributions however, may never exhibit S dominance of either distributions over the other. Formally:

Theorem 4.2. *For all $X, Y \in \mathfrak{C}^n$ with $m(X) = m + k$, $m(Y) = m$, $k \in \{1, 2, ..., n - m\}$ if X, Y is not transformable to X^S, Y^S by slide such that $m(X^S) = m(Y^S)$, then for the counterfactual distributions of X, Y obtained*

[11]For example, let $X = \{0, 0.1, 0.3, \mathbf{0.6}, 1\}$ and $Y = \{0.3, \mathbf{0.55}, 0.7, 0.8, 1\}$. In this case $m(X) = 4$, $m(Y) = 2$. Note that the only possible slide of X produces $X^S = \{0.1, 0.3, \mathbf{0.6}, 1, 1\}$ but fails to equalize the median category. In this case we add one unpopulated category before the first category of the distributions to yield $Y^F = \{0, 0.3, \mathbf{0.55}, 0.7, 0.8, 1\}$ and $X^F = \{0, 0, 0.1, 0.3, \mathbf{0.6}, 1\}$. Next we slide X^F to obtain $X^{FS} = \{0.1, 0.3, \mathbf{0.6}, 1, 1, 1\}$. The counterfactual distributions are X^{FS} and Y^F.

by ATUC and slide given by $X^{FS}, Y^{FS} \in \mathfrak{C}^{n+k'}$ $(k' \in \mathbb{Z}_{1+},\ k' \leq k)$ such that $m(X^{FS}) = m(Y^{FS})$, both $X^{FS} \not\succ_S Y^{FS}$ and $Y^{FS} \not\succ_S X^{FS}$ hold.

Proof: For the proof, we use the notations defined during the construction of X^{FS} and Y^{FS}.
The construction of X^{FS} and Y^{FS} ensure the inequalities $X_1^{FS} > 0,\ Y_1^{FS} = 0, X_{n-1}^{FS} = 1$ and $Y_{n-1}^{FS} < 1$. The inequalities $X_1^{FS} > 0$ and $Y_{n-1}^{FS} < 1$ hold since the construction of X^{FS} and Y^{FS} from X and Y, respectively, requires the inclusion of $k'(> 0)$ unpopulated categories to both X and Y either before the first category or after the last category. That $k' = (m+k-sl_{max}^X)-(m+sr_{max}^Y)$ implies that k' is the minimum number of unpopulated categories required in the construction of the counterfactuals versions of X and Y that have the same median categories. If $X_1^{FS} = 0$ or $Y_{n-1}^{FS} = 1$ then it necessarily implies that inclusion of fewer than k' unpopulated categories can yield counterfactual distributions from X and Y that have the same median categories which is a contradiction. That $Y_1^{FS} = 0$ and $X_{n-1}^{FS} = 1$ follows from the following facts. If the unpopulated categories are added before the first categories of both distributions then the construction of X^{FS} entails at least one left slide of X^F. This would imply $X_{n-1}^{FS} = 1$. Since Y^{FS} is obtained from Y^F by slides to the right, $Y_1^{FS} = 0$ is ensured since at least one unpopulated category is added before the first category of Y. Alternatively, in the case where the unpopulated categories are added after the last categories of both distributions, the construction of Y^{FS} entails at least one right slide of Y^F. This would imply $Y_1^{FS} = 0$. Since X^{FS} is obtained from X^F by slides to the left, $X_{n-1}^{FS} = 1$ is ensured since at least one unpopulated category is added after the last category of X.
These inequalities reduce to $X_1^{FS} > Y_1^{FS}$ and $X_{n-1}^{FS} > Y_{n-1}^{FS}$. The later set of inequalities prove the theorem as $1 < m(X^{FS}) = m(Y^{FS}) < (n + k')$.
Q.E.D.

Given the above theorem, it becomes apparent that for distributions having different median categories such that no counterfactual distributions having the same median categories can be generated from them with slide alone, dominance relationship may exist for *SS* dominance (and maybe higher orders) but not for *S* dominance. Even under such conditions, the relationship between the *SS* dominance ordering may be related to the inequality orderings of the underlying distributions for certain class of inequality indices. This is summarized in the next theorem:

Theorem 4.3. *For all $X, Y \in \mathfrak{C}^n$ with $m(X) = m + k$, $m(Y) = m$, $k \in \{1, 2, ..., n-m\}$ and there exist no X^S, Y^S obtained by slide on X and Y, respectively, such that $m(X^S) = m(Y^S)$, let $X^{FS}, Y^{FS} \in \mathfrak{C}^{n+k'}(k' \in \mathbb{Z}_{1+},\ k' \leq k)$ be obtained from X and Y, respectively, by ATUC and slide such that $m(X^{FS}) = m(Y^{FS}) = m'$, then for all $(I_1^n(), I_2^{n+k'}()) \in \mathbb{I}_{5\{m'-1,n+k'-m'+1\}}^{\{n,n+k'\}}$, $X^{FS} \succ_{SS} (\prec_{SS})\ Y^{FS} \implies I_1^n(\underline{X}^*, \overline{X}^*) < (>)\ I_1^n(\underline{Y}^*, \overline{Y}^*)$.*

Proof: For the proof, we use the notations defined during the construction of X^{FS} and Y^{FS}.
Let $X^{FS} = A, Y^{FS} = B, X^F = C$ and $Y^F = D$. Clearly $A, B, C, D \in \mathfrak{C}^{n+k'}$ and $m(A) = m(B) = m'$. Following theorem 3.2 we can write $A \succ_{SS} (\prec_{SS}) \ C \implies I_2^{n+k}(\underline{A}^*, \overline{A}^*) < (>)\ I_2^{n+k'}(\underline{B}^*, \overline{B}^*)\ \forall I_2^{n+k'}() \in \mathbb{I}_{2\{m'-1,n+k'-m'+1\}}^{n+k'}$ and following the definition of SI, we can write $I_2^{n+k'}(\underline{A}^*, \overline{A}^*) < (>)\ I_2^{n+k'}(\underline{B}^*, \overline{B}^*) \implies I_2^{n+k'}(\underline{C}^*, \overline{C}^*) < (>)\ I_2^{n+k'}(\underline{D}^*, \overline{D}^*)$.
Since $I_1^n()$ and $I_2^{n+k'}()$ are associated through ITUC, the following holds: $I_2^{n+k'}(\underline{C}^*, \overline{C}^*) < (>)\ I_2^{n+k'}(\underline{D}^*, \overline{D}^*) \implies I_1^n(\underline{X}^*, \overline{X}^*) < (>)\ I_1^n(\underline{Y}^*, \overline{Y}^*)$. Thus $X^{FS} \succ_{SS} (\prec_{SS})\ Y^{FS} \implies I_1^n(\underline{X}^*, \overline{X}^*) < (>)\ I_1^n(\underline{Y}^*, \overline{Y}^*)$.
Q.E.D.

Theorems 4.1 and 4.3 have a limitation: for the class of inequality indices meted out in the theorems, *S* and/or *SS* dominance rankings among the counterfactual distributions although sufficient to imply an inequality ordering of the original distributions but not necessarily does so. In the rest of the section we show that if we restrict the class inequality index further, then *S* and/or *SS* dominance among the counterfactual distributions have an equivalence with the inequality ordering of the original distributions. For the same, we define the class of functions to assume the following form:

$$I^n(\underline{Z}^*, \overline{Z}^*) = H\left(\sum_{i=1}^{m(Z)-1} \phi_i(\underline{Z}_i^*) + \sum_{j=1}^{n-m(Z)+1} \psi_j(\overline{Z}_j^*)\right) \tag{6}$$

where $H()$ is a strictly monotonically increasing function and $\phi_i(), \psi_j()$ are strictly decreasing functions with $\phi_i(0.5) = \psi_j(0.5)\ \forall i \in \{1, 2, ..., m(Z) - 1\}, \forall j \in \{1, 2, ..., n - m(Z) + 1\}$.

Remarks: Note that $I^n()$ is strictly decreasing function. Also note that $\phi_i(0.5) = \psi_j(0.5)\ \forall i \in \{1, 2, ..., m(Z) - 1\}, \forall j \in \{1, 2, ..., n - m(Z) + 1\}$ guarantees that $I^n(\underline{Z}^*, \overline{Z}^*)$ satisfies SI. Thus all functions $I^n()$ of the form given by equation 6 belongs to $\mathbb{I}_3^n$.

We also define a further restricted version of the class of inequality indices given by 6 to form:

$$I^n(\underline{Z}^*, \overline{Z}^*) = H\left(\sum_{i=1}^{m(Z)-1} \phi(\underline{Z}_i^*) + \sum_{j=1}^{n-m(Z)+1} \psi(\overline{Z}_j^*) \right) \tag{7}$$

where $H()$ is a strictly monotonically increasing function, $\phi()$ and $\psi()$ are strictly decreasing functions with $\phi(0.5) = \psi(0.5)$ and $\phi()$ and $\psi()$ are strictly convex functions.

Remarks: Note that as before, $I^n()$ is strictly decreasing function that satisfies SI. In addition to this, following 7.2 and definition 7, $I^n()$ is strictly PS convex with partition $\{m(Z), n - m(Z) + 1\}$. Furthermore since $H()$ is a strictly monotonically increasing function, any $I^n()$ and $I^{n+k}()$ $\forall k \in \mathbb{Z}_{1+}$, are related through ITUC. Thus all functions $I^n(\underline{Z}^*, \overline{Z}^*)$ of the form given by equation 7 belongs to $\mathbb{I}^n_{4\{m(Z)-1,n-m(Z)+1\}}$ and the tuples $(I^n(), I^{n+k}(\underline{Z}^*, \overline{Z}^*))$ $\forall k \in \mathbb{Z}_{1+}$ belongs to $\mathbb{I}^{n,n+k}_{5\{m(Z)-1,n-m(Z)+1\}}$.

We now establish the equivalence between dominance orderings of the counterfactual distributions to that of inequality ordering of the base distributions considering family of inequality measures specified in equations 6 and 7. Formally:

Theorem 4.4. *For all $X, Y \in \mathfrak{C}^n$ where $m(X) = m + k$, $m(Y) = m$ and $k \in \{1, 2, ..., n - m\}$, let $X^S \in \mathfrak{C}^n$ be obtained from X by k_1 slides to the left and $Y^S \in \mathfrak{C}^n$ be obtained from Y by k_2 slides to the right $(k_1, k_2 \in \mathbb{Z}_{0+},\ k_1 + k_2 = k)$ such that $m(X^S) = m(Y^S)$. Then the following conditions hold:*
1) $X^S \succ_S (\prec_S)\ Y^S \iff I(\underline{X}^*, \overline{X}^*) < (>)\ I(\underline{Y}^*, \overline{Y}^*)$ $\forall I$ having the form specified in equation 6.
2) $X^S \succ_{SS} (\prec_{SS})\ Y^S \iff I(\underline{X}^*, \overline{X}^*) < (>)\ I(\underline{Y}^*, \overline{Y}^*)$ $\forall I$ having the form specified in equation 7.

Proof:
ONLY IF part of both conditions 1 and 2
Follows from theorem 4.1 and the remarks to equations 6 and 7.
IF part of conditions 1 and 2
Follows from the last part (*ONLY IF*) of theorems 3.1 and 3.2, respectively,

along with the fact that the inequality indices used in the theorems have the functional form depicted in equations 6 and 7, respectively.
Q.E.D.

Theorem 4.5. *For all $X, Y \in \mathfrak{C}^n$ where $m(X) = m + k$, $m(Y) = m$, $k \in \{1, 2, ..., n - m\}$ and there exist no X^S, Y^S obtained by slide on X and Y, respectively, such that $m(X^S) = m(Y^S)$, let $X^{FS}, Y^{FS} \in \mathfrak{C}^{n+k'}(k' \in \mathbb{Z}_{1+},\ k' \leq k)$ be obtained from X and Y such that $m(X^{FS}) = m(Y^{FS}) = m'$, then $X^{FS} \succ_{SS} (\prec_{SS})\ Y^{FS} \iff I_1^n(\underline{X}^*, \overline{X}^*) < (>)\ I_1^n(\underline{Y}^*, \overline{Y}^*)\ \forall (I_1^n(), I_2^{n+k'}())$ having the form specified in equation 7.*

Proof:
ONLY IF part of the theorem
Follows from theorem 4.3 and the remarks to equation 7.
IF part of the theorem
We prove this by contradiction. Using the notations used in the construction of X^{FS} and Y^{FS}, let $X^{FS} = A, Y^{FS} = B, X^F = C$ and $Y^F = D$. The class of functions given by $I(\underline{Z}^*, \overline{Z}^*) = \theta \sum_{i=1}^{m(Z)-1} [u(\underline{Z}_i^*) - u(0.5)] + (1 - \theta) \sum_{j=1}^{p-m(Z)+1} [v(\overline{Z}_j^*) - v(0.5)]$ where $\theta \in (0, 1)$ and $u(), v()$ are strictly decreasing and strictly convex functions, satisfy the form laid down in equation 7 as well as the form of the inequality index used in the last part (ONLY IF) of theorem 3.2. Thus using the logic used in the last part (ONLY IF) of theorem 3.2, we can say that $A \not\succ_{SS} (\not\prec_{SS}) B \implies \exists I^{n+k'}()$ of the form specified above such that $I^{n+k'}(\underline{A}^*, \overline{A}^*) > (<) I^{n+k'}(\underline{B}^*, \overline{B}^*)$. Following the remarks to equation 7 it is evident that $I^{n+k'}()$ follows SI property, thereby implying $I^{n+k'}(\underline{C}^*, \overline{C}^*) > (<) I^{n+k'}(\underline{D}^*, \overline{D}^*)$. In addition to this, it is also ensured that there exists $I^n()$ of the form depicted in the equation such that $I^{n+k'}(\underline{C}^*, \overline{C}^*) > (<) I^{n+k'}(\underline{D}^*, \overline{D}^*) \implies I^n(\underline{X}^*, \overline{X}^*) > (<) I^n(\underline{Y}^*, \overline{Y}^*)$.
Q.E.D.

This concludes all the theoretical results of our analysis. At this juncture, it is worthwhile to mention that Naga and Yalcin (2010) devised a methodology of comparing distributions having different median categories through the construction of "equivalence classes" as mentioned in the introduction to this paper. To emphasize the generality of our analysis, we present certain limitations of the methodology developed by Naga and Yalcin (2010) which are circumvented by our approach. The first shortcoming concerns the way equivalence between distributions is defined in their paper. The authors formalizes the similarity between the distributions whose PDFs are given by $\{\mathbf{1.0}, 0, 0, 0\}, \{0, \mathbf{1.0}, 0, 0\}, \{0, 0, \mathbf{1.0}, 0\}$ and $\{0, 0, 0, \mathbf{1.0}\}$ under the pretext that all these distributions share the fundamental property: that every member of the distributions reports being at the respective median thus being most egalitarian distributions and (thus) exhibiting identical and the least amount of inequality. Building on this notion, the authors define two distributions to be equivalent if and only if the absolute difference between the CDFs of the distributions to that of the the most polarized distribution (where half of the population mass rests at the first and the last categories) are equal up to a permutation. In doing so, the equivalence is defined at the level of CDFs (in their deviation form from the most polarized distribution) but not at the level of PDFs and no rationale is provided for selecting CDFs over the PDFs. Thus although the distributions with PDFs given by $\{0.4, \mathbf{0.4}, 0.1, 0.1\}$ and $\{0.1, 0.1, \mathbf{0.4}, 0.4\}$ are equivalent as per their definition, the distributions with PDFs given by $\{0.4, \mathbf{0.4}, 0.1, 0.1\}$ and $\{0.1, \mathbf{0.4}, 0.1, 0.4\}$ fail to qualify as being equivalent: both these PDFs have the same distance in terms of the Euclidean norm from PDF of the most polarized distribution given by $\{0.5, 0, 0, 0.5\}$. Furthermore, the above definition of equivalence is completely ad hoc as there may be competing ways of defining equivalence between two distributions. As an example, consider the PDFs given by $\{0.4, \mathbf{0.4}, 0.1, 0.1\}$ and $\{0.1, \mathbf{0.4}, 0.4, 0.1\}$ related to each other by a mere permutation of the categories adjacent to the median. These distributions are not equivalent as per the definition of Naga and Yalcin (2010); an alternative definition where any permutation of population mass over the categories is defined to be equivalent might declare these two distributions equivalent. Thus the definition of equivalence puts restrictions on the inequality indices that acquiesce to the definition[12]. An yet another shortcoming of the methodology may be demonstrated with the distributions having the PDFs given by $\{0.2, 0.1, 0.1, \mathbf{0.1}, 0.4, 0.1\}$ and $\{0.1, 0.2, 0.1, \mathbf{0.1}, 0.3, 0.2\}$. These distributions are equivalent as per the definition thus implying that any "median independent inequality measure" must assign identical inequality values to these distributions. This however puts two restrictions on the

[12] Allison and Foster (2004) discuss an inequality index whose arguments are the values of the PDF and under certain parametric restrictions is insensitive to permutations of the PDF but fail to qualify for the families of "median independent inequality indices" as per Naga and Yalcin (2010).

functional form of the inequality indices. The first of these is that categories below and above the median must contribute equally to the indices. In an earlier article however, Naga and Yalcin (2008) have themselves acknowledged that it might be necessary for the researcher to accommodate differing judgments regarding inequality below and above the median. However, their definition of equivalence imposes a kind of symmetry on the inequality indices and thus it is not possible to assign different weights below and above the median. The second restriction on the inequality indices is that the indices may not weight categories based on their relative position in the distribution[13] - something which again might be necessary from the viewpoint of the researcher to accommodate differing judgments regarding inequality.
In the next section, we provide an application of the methodology of constructing counterfactuals from distributions that do not share a common median category, developed in the present section.

5 Empirical Illustrations

For the illustration, we compare the inequality in educational attainments (EA) among Indian males and females. For this exercise we consider data compiled by National Sample Survey Office (NSSO) as a part of their regular surveys on employment and unemployment. Data on EA provided by the NSSO is ordinal in nature. So far we have surveyed, there is no study on India that addresses intra gender inequality of EA considering the ordinal nature of the data.

We consider three NSSO round data 61st, 66th and 68th round. These data set that was collected for the period June 2004- July 2005, June 2009- July 2010 and June 2011- July 2012, respectively. The main variable of interest for this study is educational attainment (EA), which consists of the following categories: not literate, literate without formal schooling: through Non-formal Education Courses (NFEC) or Adult Education Centers (AEC) or Education Guarantee Scheme (EGS), Total Literacy Campaign (TLC), etc., literate with formal schooling: below primary, primary, middle, secondary, higher secondary, diploma/certificate course, graduate, postgraduate and above. Since the categories EGS/NFEC/AEC/TLC and others can not

[13] See Chakravarty and Maharaj (2015) for an inequality index that weighs categories based on their relative position in the distribution.

be ordered, we have combined all these categories as literate without formal schooling. For our study, we consider only the working age population (as defined by OECD) and thus restrict the age groups to 15-64 years.

In table 1 we present the cumulative distribution of the educational attainments for males and females across all the three NSSO survey rounds. It is readily observed that the educational attainment of males are better than that for females. For example, the percentages of illiterates is nearly double for the females compared to that of the males. On the other hand the better off categories like graduates and above, the percentages of female is much lower than male. However, it is clearly evident from this table that Indian educational attainment has improved substantially from the period 2004-05 to 2011-12.

Notice that the median category for the male and the female is different across all the three NSSO rounds. For the males the median is 4, 5 and 5 for 61st, 66th and 68th round, respectively. On the other hand the median category for the female for these NSSO rounds are 1,4, and 5, respectively. We consider our theory introduced in the theoretical section of this paper for addressing the issue of educational inequality. If we observe table 1 closely, the proposed two step algorithm fails to provide conclusive result for 2004-05. Nevertheless, for both 66th and 68th round there is clear evidence that males have lower inequality in education compared to that of the females.

For an illustration, of the two-stepped algorithm we consider the 68th round data. The median category for the males and females is 5 and 4, respectively. In the first step we introduce 1 (=6-5) unpopulated category before the first category of both the distributions. In the second step we slide the distribution of the male once towards the left. Thus the counterfactual distributions following this two stepped algorithm for the males and females are $M = \{0.187,\ 0.191,\ 0.282,\ 0.407,\ 0.603,\ 0.773,\ 0.88,\ 0.899,\ 0.976,\ 1, 1\}$ and $F = \{0, 0.372,\ 0.376,\ 0.463,\ 0.576,\ 0.727,\ 0.847,\ 0.926,\ 0.934,\ 0.984,\ 1\}$, respectively. Notice that median categories for both these counterfactual distributions (i.e., M and F) are equal to 5. Further, it is straightforward to show that the distribution M SS dominates F (see definition 4). The class of indices that shows higher inequality among the females is characterized in theorems 4.3 and 4.5 and exemplified in table 2.

6 Conclusion

In this paper we address some problems associated with inequality ordering of ordinal variables. Most of the papers written in this area is built on the ideas of median preserving spread and eventually on S dominance introduced in a seminal article by Allison and Foster (2004). Unlike others we do not limit our attention only on the S dominance but also focus on the cumulative version of S dominance introduced by Chakravarty and Maharaj (2015): SS dominance. We begin by exploring the association between certain types of transfers of population mass to that of SS dominance linked through such transfers. We also establish the relationship between inequality orderings of certain family of inequality indices to that of S and SS dominance. Finally, the paper characterizes the class of inequality indices which allows comparison between ordinal distributions having different median categories. We suggest a counterfactual based approach such that dominance ordering of certain counterfactual distributions have a direct relationship on the inequality ordering of the original distributions. We show that the proposed approach is more general to the only contribution in this direction by Naga and Yalcin (2010). To serve an example, the proposed methodology is also applied to data on educational attainments in India. Inequality orderings of working age population (as per definition of OECD) of males and the females is calculated where it is observed that in general, females have a higher level of inequality compared to males.

This paper is amenable to further extensions. Throughout the paper we assume that median category is unique. In a recent paper Kobus (2015) have redefined median such that ordinal variables may have more than one median. A future research direction is to extend this paper in the context of non-unique median categories. An yet another area that holds promise is to establish the equivalence between population transfers and families of inequality indices with higher order dominance orderings (i.e. orderings above the second order).

References

Allison, R. A. and Foster, J. E. (2004). Measuring health inequality using qualitative data. *Journal of health economics*, 23(3):505–524.

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

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

Atkinson, A. B. and Stiglitz, J. E. (2015). *Lectures on public economics.* Princeton University Press.

Berry, K. J. and Mielke Jr, P. W. (1992). Indices of ordinal variation. *Perceptual and motor skills*, 74(2):576–578.

Chakravarty, S. (2009). *Inequality Polarization and Poverty. Advances in Distributional Analysis.* Springer Verlag, New York.

Chakravarty, S. R. (2015). *Inequality, Polarization and Conflict: An Analytical Study*, volume 12. Springer.

Chakravarty, S. R. and Maharaj, B. (2012). Ethnic polarization orderings and indices. *Journal of Economic Interaction and Coordination*, 7(1):99–123.

Chakravarty, S. R. and Maharaj, B. (2015). Generalized gini polarization indices for an ordinal dimension of human well-being. *International Journal of Economic Theory*, 11(2):231–246.

Dasgupta, P., Sen, A., and Starrett, D. (1973). Notes on the measurement of inequality. *Journal of economic theory*, 6(2):180–187.

Esteban, J.-M. and Ray, D. (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851.

Foster, J. E. (1985). Inequality measurement. *Fair allocation*, 33:31–68.

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

Foster, J. E. and Shorrocks, A. F. (1988b). Poverty orderings and welfare dominance. *Social Choice and Welfare*, 5(2-3):179–198.

Kobus, M. (2015)). Polarization measurement for ordinal data. *The Journal of Economic Inequality*, 13(2):275–297.

Kobus, M. and Miłoś, P. (2012). Inequality decomposition by population subgroups for ordinal data. *Journal of Health Economics*, 31(1):15–21.

Marshall, A. W. and Olkin, I. (1979). *Theory of Majorization and its Applications*. Academic Press, New York.

Muliere, P. and Scarsini, M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory*, 49(2):314–323.

Naga, R. H. A. and Yalcin, T. (2008). Inequality measurement for ordered response health data. *Journal of Health Economics*, 27(6):1614–1625.

Naga, R. H. A. and Yalcin, T. (2010). Median independent inequality orderings. *University of Aberdeen Business School working paper series*, 3:1–25.

Piketty, T. (2014). *Capital in the twenty-first century*. Harvard University Press.

Rudin, W. (1971). *Principles of mathematical analysis*. McGraw-Hill Book Co.

Sen, A. (1987). The standard of living: lecture i, concepts and critiques. *The standard of living*, pages 1–19.

Sen, A. et al. (1999). Commodities and capabilities. *OUP Catalogue*.

Shorrocks, A. (1983). Ranking income distributions. *Economica*, 50:3–17.

Stewart, F. (2004). Development and security. *Conflict, Security & Development*, 4(3):261–288.

Stiglitz, J. (2012). *The price of inequality*. Penguin UK.

United Nations General Assembly Resolution 65/309 (2011). Happiness: towards a holistic approach to development, available at www.un.org/esa/socdev/ageing/documents/noteonhappinessfinalclean.pdf.

Zheng, B. (2008). Measuring inequality with ordinal data: a note. *Research on Economic Inequality*, 16:177–188.

Table 1: Cumulative distribution of educational attainments in the twenty major Indian States

Educational Categories	Round 61		Round 66		Round 68	
	Males	Females	Males	Females	Males	Females
Illiterate	0.264	$\boxed{0.510}$	0.195	0.397	0.187	0.372
Literate without formal schooling	0.289	0.530	0.200	0.402	0.191	0.376
Below Primary	0.377	0.601	0.282	0.484	0.282	0.463
Primary	$\boxed{0.527}$	0.715	0.419	$\boxed{0.607}$	0.407	$\boxed{0.576}$
Middle	0.723	0.842	$\boxed{0.623}$	0.755	$\boxed{0.603}$	0.727
Secondary	0.848	0.918	0.790	0.865	0.773	0.847
Higher Secondary	0.920	0.960	0.892	0.934	0.880	0.926
Diploma	0.933	0.966	0.908	0.941	0.899	0.934
Graduate	0.986	0.992	0.980	0.986	0.976	0.984
Post Graduate and above	1.000	1.000	1.000	1.000	1.000	1.000

Notes

[1] Authors' calculations based on data from: NSSO Employment-Unemployment Rounds 61, 66, 68 corresponding to years 2004-05,2009-10 and 2011-12, respectively.

[2] Age Group: 15-64 years.

[3] Median category is represented in the box.

[4] In this table the figures corresponds to the cumulative distribution function of educational attainments in twenty major states of India. Following are the descriptions of the educational categories: not literate -1, below primary (including literate without formal schooling through EGS/NFEC/AEC/TLC/ others) -2, primary -3, middle -4, secondary - 5, higher secondary -6, diploma/certificate course -7, graduate -8, postgraduate and above -9.

Table 2: Index of inequality for educational attainments in the twenty major Indian States

Educational Categories	Round 61		Round 66		Round 68	
Parameter	Males	Females	Males	Females	Males	Females
$\gamma = 1$	0.377	0.366	0.392	0.445	0.392	0.439
$\gamma = 0.81$ (Calibrated)	0.329	0.328	0.345	0.404	0.344	0.394
$\gamma = 0.5$	0.231	0.245	0.248	0.310	0.245	0.293
$\gamma = .3$	0.152	0.169	0.167	0.219	0.163	0.202

Notes

[1] Authors calculations based on CDF table 1.

[2] The mathematical form of the index may be written as follows:

$$I(X;\gamma) = 1 - \frac{2^{\gamma}}{n-1} \sum_{i=1}^{n-1} \left| X_i - 0.5 \right|^{\gamma}$$

This index is bounded between 0 and 1. The parameter γ measures the weight given to the median category. Whenever γ approaches zero, the relative weight given to the median category increases, and the relative contribution of the other categories is reduced. *Calibrated Parameter:* Apouey has argued that polarization is medium when all individuals are uniformly distributed over categories. The author has provided calibrated parameter γ^*, for which $I(X;\gamma^*) = 1/2$. The parameter γ^*, however, depends on n. Since we have n=10, in that case $\gamma^* = 0.81$.

7 Mathematical appendix

We use two results from Marshall and Olkin 1979 as the following two lemmas:

Lemma 7.1. *If $a' = \{a_1, a_2, .., a_n\}$ and $b' = \{b_1, b_2, .., b_n\}$ be any two vectors in $\mathbb{R}^n$, such that $a_1 \leq a_2 \leq ... \leq a_n$ and $b_1 \leq b_2 \leq ... \leq b_n$, the following conditions are equivalent:*
1)There exists a bi-stochastic matrix Q which is not a permutation matrix such that $a \geq Qb$.
2) $a_1 + a_2 + .. + a_k \geq b_1 + b_2 + .. + b_k$ for all $k \in \{1, 2, .., n\}$ with $>$ for at least one k, or equivalently $a \succ_{SISD} b$.
3) For all strictly concave, increasing and real valued function u defined on $\mathbb{R}$; $\sum_{i=1}^{n} u(x_i) > \sum_{i=1}^{n} u(y_i)$.

For the proof, see Marshall and Olkin (1979).

Lemma 7.2. *If $\theta()$ is strictly increasing and strictly concave function then the function $W : \mathbb{R}_n \rightarrow \mathbb{R}$ defined as $W = \sum_{i=1}^{n} \theta(X_i)$ is increasing and a strictly S concave function.*

For the proof, see (Marshall and Olkin, 1979, pp 64).

To prove theorem 3.2, we extend lemma 7.1 as follows:

Lemma 7.3. *Using the notations of lemma 7.1, $a \not\succ_{SISD} b$ and $a \neq b$, implies there exists a strictly increasing, strictly concave, continuous and real valued $u()$ such that*

$$\sum_{i=1}^{n} u(a_i) < \sum_{i=1}^{n} u(b_i) \tag{8}$$

Proof: From lemma 7.1 $a \not\succ_{SISD} b \implies$ there exists a real valued function $u()$ which is strictly concave, strictly increasing such that either $\sum_{i=1}^{n} u(a_i) < \sum_{i=1}^{n} u(b_i)$ or $\sum_{i=1}^{n} u(a_i) = \sum_{i=1}^{n} u(b_i)$. If strict inequality holds, the proof is complete. If $\sum_{i=1}^{n} u(a_i) = \sum_{i=1}^{n} u(b_i)$ then we show that there necessarily exists a real valued function $h()$ which is strictly concave and strictly increasing such that $\sum_{i=1}^{n} h(a_i) < \sum_{i=1}^{n} h(b_i)$. Clearly, $a \neq b \implies \exists$ a real scalar v such that $\#a_{i|a_i=v} > \#b_{i|b_i=v} \forall i \in \{1, 2, ..., n\}$. We select δ_1 and δ_2 with $v \in [\delta_1, \delta_2]$ such that $a_i = v \iff a_i \in [\delta_1, \delta_2]$ and $b_i = v \iff b_i \in [\delta_1, \delta_2] \forall i \in \{1, 2, ..., n\}$. Now define the following function

$$\begin{aligned} h(z) &= u(z) \; if \; z \notin [\delta_1, \delta_2] \\ &= \lambda u(z) + (1-\lambda)\left[\frac{u(\delta_2) - u(\delta_1)}{\delta_2 - \delta_1}(z - \delta_1) + u(\delta_1)\right] if \; z \in [\delta_1, \delta_2] \quad (9) \end{aligned}$$

where $\lambda \in (0, 1) \forall z \in \mathbb{R}$.

Following lemma 7.7, $h()$ is strictly increasing, strictly concave and $h(z) < u(z) \forall z \in (\delta_1, \delta_2)$. Thus $\sum_{i=1}^{n} h(a_i) < \sum_{i=1}^{n} h(b_i)$. **Q.E.D**

Lemma 7.4. *If $f(x)$ is strictly concave and $L(x)$ is a linear function then $\lambda f(x) + (1-\lambda)L(x)$ is strictly concave.*

Proof: $f(x)$ is strict concave $\implies$

$$f(\eta x + (1-\eta)y) > \eta f(x) + (1-\eta)f(y) \quad (10)$$

Linearity of $L()$ $\implies$

$$L(\eta x + (1-\eta)y) = \eta L(x) + (1-\eta)L(y) \quad (11)$$

10 $\times\lambda+$ 11 $\times(1-\lambda)$ $\implies$

$$\begin{aligned} \lambda f(\eta x + (1-\eta)y) + (1-\lambda)L(\eta x + (1-\eta)y) > \eta L(x) + (1-\eta)L(y) \quad > \\ \eta[\lambda f(x) + (1-\lambda)L(x)] + (1-\eta)[\lambda f(y) + (1-\lambda)L(y)] \end{aligned}$$

Hence, proved.

Lemma 7.5. *If f(x) is strictly concave and L(x) is a straight line then $f(x) = L(x)$ can be satisfied in at most 2 (distinct) points.*

Proof: If $f(x) = L(x)$ is satisfied in more than two points, we can take three points , call them x, y and z with $x < y < z$ and $f(x) = L(x)$, $f(y) = L(y)$ and $f(z) = L(z)$. We can also set a $\lambda \in [0, 1]$ such that $\lambda x + (1-\lambda)y = z$. Then by linearity of $L()$, $L(z) = \lambda L(x) + (1-\lambda)L(y)$. Strict concavity of $f()$ $\implies$ $f(z) > \lambda f(x) + (1-\lambda)f(y)$ which violates $f(x) = L(x)$, $f(y) = L(y)$ and $f(z) = L(z)$. **Q.E.D**

Lemma 7.6. *If f() is strictly concave, $L()$ is linear and $f(p) = L(p)$ then*
$f(x) > L(x)$ *for some* $x > p \implies f(x) < L(x) \forall x < p$
$f(x) > L(x)$ *for some* $x < p \implies f(x) < L(x) \forall x > p$

Proof: Suppose not, then $f(x) > L(x)$ for some $x > (<)p$ and $f(x) \geq L(x)$ for some $x < (>)p$.

Take any $x > (<)p$ such that $f(x) > L(x)$ and call it x_1. Take any $x < (>)p$ such that $f(x) > L(x)$ and call it x_2. Note that p can be expressed as $p = \lambda x_1 + (1-\lambda)x_2 : \lambda \in [0,1]$. Now, strict concavity of f() $\implies$

$$f(p) > \lambda f(x_1) + (1-\lambda) f(x_2)$$

and linearity of $L()$ $\implies$

$$L(p) = \lambda L(x_1) + (1-\lambda) L(x_2)$$

which clearly violates $f(p) = L(p)$, $f(x_1) > L(x_1)$ and $f(x_2) \geq L(x_2)$. **Q.E.D**

Lemma 7.7. *Given lemma 7.4, 7.5 and 7.6 we can show that if f() is strictly concave and*

$$\begin{aligned} h(x) &= f(x) \; if \; x \notin [a,b] \\ &= \lambda f(x) + (1-\lambda)\left[\frac{f(b)-f(a)}{b-a}(x-a) + f(a)\right] if \; x \in [a,b] \end{aligned} \quad (12)$$

where $\lambda \in [0,1]$, *then h(x) is strictly concave.*

Proof: Define a function $g : \mathbb{R} \longrightarrow \mathbb{R}$

$$g(x) = \lambda f(x) + (1-\lambda)\left[\frac{f(b)-f(a)}{b-a}(x-a) + f(a)\right] \quad (13)$$

and

$$L(x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a) \quad (14)$$

Note that

$$\begin{aligned} L(a) &= f(a) \\ L(b) &= f(b) \end{aligned} \quad (15)$$

By lemma 7.4, $g(x)$ is strictly concave.

$$\begin{aligned} h(x) &= f(x)\ if\ x \notin [a,b] \\ h(x) &= g(x)\ if\ x \in [a,b] \end{aligned} \tag{16}$$

Note that $\forall x \in (a,b)$, x can be expressed as $x = \eta b + (1-\eta)a \forall \eta \in (0,1)$. Following strict concavity of $f()$ we can write

$$f(x) > \frac{f(b)-f(a)}{b-a}(x-a) + f(a) \tag{17}$$

This implies $f(x) > L(x)$ and since $h(x) = \lambda f(x) + (1-\lambda)L(x)$ with $\lambda \in (0,1)$, hence we can write:

$$f(x) > h(x) > L(x)\ \forall x \in (a,b) \tag{18}$$

Now following lemma 7.6 we can also write $f(x) < L(x) \implies$

$$\forall x \notin [a,b], f(x) < g(x) < L(x) \tag{19}$$

Consider $x, y, z \in \mathbb{R}$ such that $x < z < y$ and $\exists \lambda \in (0,1)$; st $z = \lambda x + (1-\lambda)y$. We have to show that

$$h(z) > \lambda h(x) + (1-\lambda)h(y) \tag{20}$$

There may be five cases which we prove individually to prove this lemma.

Case 1: If $x, y, z \notin (a,b)$; $h(z) > \lambda h(x) + (1-\lambda)h(y)$. It follows, since $\forall x \notin (a,b) h(x) = f(x)$.

Case 2: If $x \in (a,b)$ while $y, z \notin (a,b)$ then $f(z) > \lambda f(x) + (1-\lambda)f(y) \implies f(z) > \lambda h(x) + (1-\lambda)f(y) \because f(x) > h(x) \forall x \in (a,b)$ following equation 18. Thus, $h(z) > \lambda h(x) + (1-\lambda)h(y) \because y, z \notin (a,b), h(z) = f(z)$ and $h(y) = f(y)$ following equation 12.

Using the same logic we can prove $h()$ is concave when $y \in (a,b)$ and $x, y \notin (a,b)$.

Case 3: If $x, z \in (a,b)$ while $y \notin [a,b]$. Following lemma 7.4 $g(x)$ is strictly concave $\implies g(z) > \lambda g(x) + (1-\lambda)g(y) \implies g(z) > \lambda g(x) + (1-\lambda)f(y) \because f(y) < g(y)$ by following equation 19. Since $x, z \in (a,b) \implies h(x) = g(x)$ and $h(z) = g(z)$. This implies equation 20 holds for this case.

Similar logic also holds if $y, z \in (a,b)$ while $x \notin [a,b]$.

Case 4 If $x, y, z \in [a, b]$; by lemma 7.4 $g()$ is strictly concave, which implies $g(z) > \lambda g(x) + (1 - \lambda)g(y)$. By construction $h(x) = g(x) \forall x \in [a, b]$. Hence, equation 20 is satisfied when $x, y, z \in [a, b]$.

Case 5: If $x, y \notin (a, b)$ while $z \in (a, b)$. Following lemma 7.4, $g(z) > \lambda g(x) + (1 - \lambda)g(y)$. This implies $g(z) > \lambda f(x) + (1 - \lambda)f(y) \because \forall x, z \notin (a, b), g(x) \geq f(x)$ and $g(y) \geq f(y)$ (following equation 19). This implies, following equation 16, that $h(z) > \lambda h(x) + (1 - \lambda)h(y)$.

Thus strict concavity of $h()$ is proved. **Q.E.D**