

CAHIER DE RECHERCHE #1208E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER #1208E
Department of Economics
Faculty of Social Sciences
University of Ottawa

Measuring Socioeconomic Health Inequalities in Presence of Multiple Categorical Information

Paul Makdissi^{*} and Myra Yazbeck[†]

July 2012

^{*} CIRPÉE and Department of Economics, University of Ottawa, 55 Laurier E. (10125), Ottawa, Ontario, Canada, K1N 6N5; Email: paul.makdissi@uottawa.ca.

[†] CIRPÉE and Department of Epidemiology Biostatistics and Occupational Health, Purvis Hall, McGill University, Montreal, Canada, H3A 1A2; Email: myra.yazbeck@mcgill.ca

Abstract

While many of the measurement approaches in health inequality measurement assume the existence of a ratio-scale variable, most of the health information available in population surveys is given in the form of categorical variables. Therefore, the well-known inequality indices may not always be readily applicable to measure health inequality as it may result in the arbitrariness of the health concentration index's value. In this paper, we address this problem by changing the dimension in which the categorical information is used. We therefore exploit the breadth of this information, define a new ratio-scaled health status variable and develop positional stochastic dominance conditions that can be implemented in a context of categorical variables. We also propose a parametric class of population health and socioeconomic health inequality indices. Finally we provide a twofold empirical illustration using the Joint Canada/United States Surveys of Health 2004 and the National Health Interview Survey 2010.

Key words: *Population Health, Health Inequality, Stochastic Dominance*

JEL Classification: D63, I10

Résumé

Bien que plusieurs approches à la mesure des inégalités de santé font comme hypothèse qu'il existe une variable d'échelle de ratio, l'information que l'on retrouve dans les enquêtes sur la santé de la population prend souvent la forme de variables catégoriques. Dans ce contexte, les indices d'inégalité connus peuvent ne pas être directement applicables à la mesure des inégalités de santé puisqu'ils produiront des valeurs arbitraires. Dans cet article, nous apportons une solution à ce problème en changeant la dimension dans laquelle l'information catégorique est utilisée. Nous exploitons l'étendue de cette information afin de définir une nouvelle variable d'échelle de ratio et nous développons des tests de dominance stochastique qui peuvent être ainsi implémentés dans le contexte de variables catégoriques. Nous proposons aussi une classe paramétrique d'indices de santé de la population et d'indices d'inégalité socioéconomique de santé. Finalement, nous proposons deux illustrations basées sur l'Enquête conjointe Canada/États-Unis sur la santé et le National Health Interview Survey 2010.

Mots clés: *Santé de la population, inégalité de santé, dominance stochastique*

Classification JEL: D63, I10

1 Introduction

Measuring health inequalities is essential for the implementation and the monitoring of health policies. Most of the current literature on health inequality is using the concentration index as an accepted measure of health inequality. However, the use of the concentration index presents four important measurement problems. The first results from the fact that the concentration indices do not account for the average level of health in the population considered (Wagstaff, 2002). Thus, a policy that improves the average level of health, while keeping the relative distribution of health constant, will be deemed neutral when using the concentration index. To overcome this problem, Wagstaff (2002) proposes the use of an achievement index that captures simultaneously the average level of health status and the socioeconomic inequality of its distribution. The second measurement problem is the well known *mirror problem* pointed out by Clarke *et al.* (2002). It results in the absence of consistency between the rankings of health attainments and health shortfalls when the concentration index is used. Erreygers (2009) suggested a corrected version of the concentration index that accounts for this inconsistency. Yet, Lambert and Zheng (2011) show that no index of relative inequality can really avoid this problem.¹

As for the third measurement problem, it resides in the concentration index's *blindness to health status* (Makdissi and Yazbeck, 2011). More specifically, any index that belongs to Wagstaff's class of health achievement indices or extended concentration indices may produce spurious conclusions. For instance, it may react favourably when a health transfer is made from an individual at a lower rank in the health distribution to a person at a higher rank, provided that the former has a slightly higher income. Therefore, the use of such indices without accounting for *blindness to health status* may be misleading if utilized in the evaluation of health policy performance. Makdissi and Yazbeck (2011) suggest a solution by proposing a parametric class of indices that accounts for pure health

¹We believe that this problem does not have serious repercussions. Our position is such that the variable of interest is inequality in health attainments. Therefore, inequality in shortfalls must be accounted for in proportion of their contribution to total attainments inequality.

inequality and socioeconomic health inequality aversion simultaneously.

The last measurement problem, is the *arbitrariness of the concentration index* (for details see Erreygers, 2006 and Zheng, 2008). It results from a common misuse of non-ratio-scaled variables while computing inequalities indices that are developed for ratio-scaled variables. Indeed, a large body of the health inequality measurement literature is based on the accumulated knowledge in income inequality measurement where the 0 has a well defined meaning.² In population surveys, most of the available information on health status is given in the form of categorical variables where the meaning of the 0 and the scale of the variable are not well defined.³ As a result, the well-known income inequality indices can not always be readily applied to measure inequalities in health status. In this perspective, this paper highlights possible issues resulting from this common practice and proposes an alternative method to avoid arbitrary rankings.

There are two possible paths an analyst can take to solve this fourth measurement problem. The first consists of finding a way to transform the available information so that it becomes ratio scaled. In this case, one has to modify the dimension in which the information provided by the categorical variable is exploited. Consequently, the focus would be on the breadth of the information rather than the depth. An alternative path consists of constructing a unit invariant inequality measure that is robust to non-ratio scaled variables such as interval variables and ordinal variables. To our best knowledge this path has lead to a series of impossibility theorems (for more details see Beckman, Smith and Zheng, 2009 and Zheng, 2008).

Our objective in this paper is to address this fourth measurement problem by following the first path. In doing so, we redefine health status variable so that it can readily be used to compute inequality indices and captures the information on socioeconomic inequality in health attainments. Our approach finds its inspiration in the work of Alkire and Foster (2011) on the *width* of poverty approach, yet differs from it on three respects. First, it limits the counting procedure at the indi-

²Income is a good example of such as variable.

³The grading provided by such variables can, at best, provide a ranking of individuals where being ranked second is not necessarily twice as bad as being ranked first.

vidual level, while Alkire and Foster’s counting procedure is used at the individual and population level. Second, it uses a different aggregation method. Whereas Alkire and Foster use counting aggregation procedure, this approach uses the information produced by the counting approach at the individual level and transfers it in a rank dependent social decision function. Third, in addition to proposing a different structure for the indices, it provides the associated stochastic dominance criteria.

The remaining of this paper is organized as follows. The next section presents the theoretical measurement framework. Section 3 presents the rank dependant ethical principles and develop two positional stochastic dominance conditions. Section 4 present a parametric class of population health and their associated socioeconomic health inequality indices. Section 5 presents a brief empirical illustration using the Joint Canada/United States Surveys of Health 2004 and the National Health Interview Survey 2010. The last section summarizes our results.

2 Theoretical Framework

Let $F(y)$ represents the cumulative distribution of income y and $p = F(y)$ be the socioeconomic status of an individual whose income y , where the health information for a given individual with socioeconomic status p is categorical and provides information on K health attributes. Also, let $\mathbf{H}(\mathbf{p}) = (h_1(p), h_2(p), \dots, h_K(p))$ represent this information for an individual at socioeconomic rank p . Assume that, for an individual at socioeconomic rank p , there exist a methodology $\phi(\mathbf{H}(\mathbf{p}))$ that transforms the information on the K health attributes into a scalar that reflects an individual’s health status level. The use of such transformations is a common practice in epidemiology where similar methodologies are often used to compute health related quality of life (HRQL) indices such the Quality of Well Being index (Kaplan, Bush and Berry, 1976), the Health Utility Index (Torrance, Feeny, Furlong et al., 1996), the EQ-5D (The Euro-Qol Group, 1990) and SF-6D (Brazier, Usherwood, Harper et al., 1998). All these indices use different methodologies to assign numerical

values for each category, and then aggregate these values into a scalar. Once this scalar is obtained, a measure of population health can be written in the following general form:

$$A(\mathbf{H}) = \int_0^1 \nu(p)\phi(\mathbf{H}(\mathbf{p}))dp, \quad (1)$$

where $\nu(p) \geq 0$ is a weight function associated with the socioeconomic status p .⁴ If this weight function is decreasing in the socioeconomic status p , then the population health measure A is said to display aversion to socioeconomic health inequality. This means that an index of socioeconomic health inequality of the following form can be associated to this population health measure (see Atkinson, 1970):

$$I(\mathbf{H}) = 1 - \frac{A}{\mu_\phi}, \quad (2)$$

where $\mu_\phi = \int_0^1 \phi(\mathbf{H}(\mathbf{p}))dp$ is the average health status.

The use of such methodologies may often be problematic as it can result in an arbitrary assignment of numerical values to health categories. Thus, the rankings of individual health statuses produced by the aggregation procedure $\phi(\mathbf{H}(\mathbf{p}))$ will not be robust to monotonically increasing transformations $g_k(h_k(p))$. To overcome this problem, we change the dimension in which we measure health status and derive a measure for the width of health problems (see Alkire and Foster, 2011) where the population health and the socioeconomic health inequality measures will be based on widths of health problems rather than depth. The associated aggregation procedure $\phi(\mathbf{H}(\mathbf{p}))$ based on this new measure will produce the same value for any monotonic transformation $g(h_k(p))$ of any health attribute, providing hence a consistent ranking.

In order to obtain information on the width of the health problems, we count the dimensions in which an individual is considered to have a health status that is below a specified acceptable threshold. More formally, for each health attribute, we assume that there exist a specific threshold category, τ_k , below which the person is considered to have a health problem in that attribute. Given

⁴This weight function is such that $\int_0^1 \nu(p)dp = 1$.

the following transformation

$$\Upsilon(\mathbf{H}(\mathbf{p})) = (\iota(h_1(p) < \tau_1), \iota(h_2(p) < \tau_2), \dots, \iota(h_K(p) < \tau_K)), \quad (3)$$

where $\iota(x < x_0)$ is an indicator function that takes a value of 1 if $x < x_0$ and is 0 otherwise, it is trivial that $\Upsilon(\mathbf{H}(\mathbf{p})) = \Upsilon(g(\mathbf{H}(\mathbf{p})))$ for any increasing monotonic transformation $g(\cdot)$ since $g_k(h_k(p)) < g_k(\tau_k)$ if and only if $h_k(p) < \tau_k$. Let $\Theta = (\theta_1, \theta_2, \dots, \theta_K)$ be a vector of weights for health attributes where the taxicab norm⁵ $\|\Theta\|_1$ is equal to K .⁶ The width of health problems is then given by $s(\mathbf{H}(\mathbf{p}))$ where $s(\mathbf{H}(\mathbf{p})) = \Upsilon(\mathbf{H}(\mathbf{p}))\Theta'$ is the weighted sum of health attributes in which an individual has a health problem. Once the width of the health problem is measured, one has to choose an aggregation procedure that is invariant to any monotonic transformation applied to numerical values of health categories. In this paper, we suggest the following aggregation procedure that meets this criterion⁷:

$$\phi(\mathbf{H}(\mathbf{p})) = \frac{K - s(\mathbf{H}(\mathbf{p}))}{K}. \quad (4)$$

The resulting measure in (4) is a quantification of an individual's health achievement as it represents the weighted proportion of health attributes without problems. The expression as laid in equation (4) is invariant to any monotonic transformation on the values associated to health categories. It clearly provides a solution for the arbitrariness problem mentioned earlier, however this consistency is not without a cost at the level of the depth of the health problem in each health attribute.⁸ Using equations (1), (2) and (4), we can thus define a class of population health and socioeconomic inequalities indices in the width of health problems. This class of indices will be invariant to any monotonic transformation $g(\cdot)$ that one may apply to $\mathbf{H}(\mathbf{p})$. The resulting population health index

⁵Note that the taxicab norm of a vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is defined by $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$.

⁶One specific example of a weight vector Θ is a vector of ones, e . We will use the weight vector e in the empirical illustration since it is consistent with the weighting of different health attribute used in many HRQL indices (see Kopec and Willison, 2003).

⁷Other aggregation procedure may be chosen as well.

⁸It is important to mention that this measure does not account for the *mirror problem* present in socioeconomic health inequality measures as discussed earlier (see footnote 1 for more details). A decomposition of this linear class of indices will reveal that the contribution of any shortcoming in health will be a function of its distribution through socioeconomic status and of its relative average importance compared to average health achievement.

for this class is given by:

$$A(\mathbf{H}) = 1 - \frac{1}{K} \int_0^1 \nu(p) s(\mathbf{H}(\mathbf{p})) dp, \quad (5)$$

and the associated socioeconomic health inequalities index is given by:

$$I(\mathbf{H}) = \frac{\int_0^1 (\nu(p) - 1) s(\mathbf{H}(\mathbf{p})) dp}{K - \int_0^1 s(\mathbf{H}(\mathbf{p})) dp}. \quad (6)$$

3 Rank dependant ethical principles and positional stochastic dominance

The objective of this section is to describe the ethical principles underlying measurement of socioeconomic health inequalities. In doing so, we adapt the definitions used by Makdissi and Mussard (2008) to the context of socioeconomic health inequalities. We also establish the necessary and sufficient conditions for positional stochastic dominance.

Let us define classes of population health indices Ω^σ , $\sigma \in \{1, 2, \dots\}$, and classes of socioeconomic health inequalities Ξ^σ , $\sigma \in \{2, \dots\}$, where σ represents the order of ethical principles obeyed by indices that belong to these classes. The ethical principles' role is to determine the mathematical properties of the weighting function $\nu(p)$. To describe the normative implications underlying the functional form of $\nu(\cdot)$, we define a set of positional ethical principles. The first principle stipulates that the ordering of two distributions of health levels is equivalent to making the health levels “parade” simultaneously alongside each other, and verifying if one parade weakly dominates the other – this exercise was first suggested by Pen (1971) in the context of income distributions.

Definition 1 *A health distribution $\tilde{\mathbf{H}}$ is obtained from the distribution \mathbf{H} by a Pen increment if there is an improvement in the category of one health attribute at position p_0 , everything else being held constant. A population health measure satisfies the Pen Parade Principle if:*

$$A(\mathbf{H}) \leq A(\tilde{\mathbf{H}}). \quad (7)$$

Since $v(p) \geq 0$ for all $p \in [0, 1]$, all $A(\cdot) \in \Omega^1$ satisfy this normative principle.

The second principle is the Pigou-Dalton Principle of Transfers (Pigou, 1912 and Dalton, 1920). Careful attention should be paid when adapting this transfer principle to the context of socioeconomic inequalities in the width of health problems. In this specific context, there is a change in the numerical value of the function $\phi(\mathbf{H}(\mathbf{p}))$ if and only if there is a change in at least one health attribute k at socio-economic rank p and that this change involves a movement across the threshold τ_k . For expositional simplicity we will present a variation in health, δ , in terms of the measure of width of well functioning health attributes $\phi(\mathbf{H}(\mathbf{p}))$.

Definition 2 Consider a health distribution $\tilde{\mathbf{H}}$ that is obtained from the distribution \mathbf{H} by a progressive Pigou-Dalton transfer of health δ , such that $\Delta\phi(\mathbf{H}(\mathbf{p}_i)) = \delta$ and $\Delta\phi(\mathbf{H}(\mathbf{p}_j)) = -\delta$, and where $p_i < p_j$. A population health measure and its associated socioeconomic health inequality index satisfy the Pigou-Dalton Principle of Transfer if:

$$A(\mathbf{H}) \leq A(\tilde{\mathbf{H}}), \quad (8)$$

and

$$I(\mathbf{H}) \geq I(\tilde{\mathbf{H}}). \quad (9)$$

A population health measure $A(\mathbf{H})$ and its associated socioeconomic health inequality index $I(\mathbf{H})$ satisfy this normative principle if and only if, $v'(p) \leq 0$ for all $p \in [0, 1]$. This condition will be verified in the definition of the classes of indices $\Omega^2 \subset \Omega^1$ and Ξ^2 .

In the literature on income distribution, it is frequent to impose more structure on the social welfare function. This is achieved by selecting transfer principles that the social welfare function should obey. There is a wide range of transfer principles. The Principle of Transfer Sensitivity (Kolm, 1976), is one of the most widely used principle in income distribution analysis. It postulates that an income transfer, valued to be δ , from a higher-income individual to a lower-income one yields a better impact on social welfare insofar as incomes (y) are the lowest possible. Hence, a decision maker respecting the Principle of Transfer Sensitivity prefers a transfer from y_2 to y_1

rather than from y_4 to y_3 (where $y_4 > y_2$ and $y_3 > y_1$), given that the rank of the individuals remains unchanged after such transfers and that $y_2 - y_1 = y_4 - y_3$. Despite its wide applicability, there is no consensus on the desirability of this principle as it is insensitive to individuals' rank. This, becomes an important issue in the context of socioeconomic health inequalities. To overcome this problem, a homologous principle was developed in a rank-dependent framework by Mehran (1976) and Kakwani (1980). It assumes that an income transfer from a high-income individual to a low-income individual yields a better impact on social welfare, insofar as individuals' ranks are the lowest possible. A decision maker respecting the Principle of Positional Transfer Sensitivity prefers a transfer from y_2 to y_1 rather than from y_4 to y_3 , given that the rank of the individuals remains unchanged after such transfers and that $p_2 - p_1 = p_4 - p_3$. A similar structure can be imposed in the context of socioeconomic health inequality. The Principle of Positional Transfer sensitivity can be easily adapted to the context of this paper since the position of any individual is defined by his socioeconomic status. Imposing this ethical principle will allow to characterize the sets of indices Ω^3 and Ξ^3 . In a more general perspective, Fishburn and Willig (1984) proposed a class of Generalized Transfer Principles that states that as σ increases, the weight that is associated to transfer in the bottom of the income distribution increases. These principles were adjusted to fit the rank dependant context by Aaberge (2009). To adapt the formal generalization of these principles to the socioeconomic health inequality measurement context, we follow closely Makkidissi and Mussard (2008). We first define $\Delta_{p,\gamma}A(\delta, \mathbf{H})$.

Definition 3 *The variation of population health index and its associated socioeconomic health inequality index induced by a progressive Pigou-Dalton health transfer δ from the person at rank $p + \gamma$, $\gamma > 0$ to the one at rank p is expressed as:*

$$\Delta_{p,\gamma}A(\delta, \mathbf{H}) := A(\tilde{\mathbf{H}}) - A(\mathbf{H}), \quad (10)$$

and

$$\Delta_{p,\gamma}I(\delta, \mathbf{H}) := I(\tilde{\mathbf{H}}) - I(\mathbf{H}), \quad (11)$$

Let us now define $\Delta_{p,\Gamma}^\sigma A(\delta, \mathbf{H})$ and $\Delta_{p,\Gamma}^\sigma I(\delta, \mathbf{H})$. These terms are recursively deduced as follows:

$$\Delta_{p,\Gamma^2}^2 A(\delta, \mathbf{H}) := \Delta_{p,\gamma_1} A(\delta, \mathbf{H}) - \Delta_{p+\gamma_2,\gamma_1} A(\delta, \mathbf{H}), \quad (12)$$

and

$$\Delta_{p,\Gamma^2}^2 I(\delta, \mathbf{H}) := \Delta_{p,\gamma_1} I(\delta, \mathbf{H}) - \Delta_{p+\gamma_2,\gamma_1} I(\delta, \mathbf{H}), \quad (13)$$

where $\Gamma^2 = (\gamma_1, \gamma_2)$, $\gamma_i > 0$,

⋮

$$\Delta_{p,\Gamma^\sigma}^\sigma A(\delta, \mathbf{H}) := \Delta_{p,\Gamma^{\sigma-1}}^{\sigma-1} A(\delta, \mathbf{H}) - \Delta_{p+\gamma_\sigma,\Gamma^{\sigma-1}}^{\sigma-1} A(\delta, \mathbf{H}), \quad (14)$$

and

$$\Delta_{p,\Gamma^\sigma}^\sigma I(\delta, \mathbf{H}) := \Delta_{p,\Gamma^{\sigma-1}}^{\sigma-1} I(\delta, \mathbf{H}) - \Delta_{p+\gamma_\sigma,\Gamma^{\sigma-1}}^{\sigma-1} I(\delta, \mathbf{H}), \quad (15)$$

where $\Gamma^\sigma = (\gamma_1, \gamma_2, \dots, \gamma_\sigma)$, $\Gamma^1 = \gamma_1$ and $\gamma_i > 0$.

Definition 4 *A population health measure $A(\mathbf{H})$ and its associated socioeconomic health inequality index $I(\mathbf{H})$ satisfy the Principle of σ th-degree Positional Transfer Sensitivity if,*

$$\Delta_{p,\Gamma^{\sigma-1}}^{\sigma-1} A(\delta, \mathbf{H}) \geq \Delta_{p',\Gamma^{\sigma-1}}^{\sigma-1} A(\delta, \mathbf{H}) \text{ and } \Delta_{p,\Gamma^{\sigma-1}}^{\sigma-1} I(\delta, \mathbf{H}) \leq \Delta_{p',\Gamma^{\sigma-1}}^{\sigma-1} I(\delta, \mathbf{H}) \text{ for all } p' > p.$$

The Positional Transfer Principle of order σ requires some assumptions on the weight function. Let $v^{(i)}(\cdot)$ be the i th derivative of the function $v(\cdot) =: v^{(0)}(\cdot)$. Consider the following sets of rank dependant population health measures:

$$\begin{aligned} \Omega^\sigma &:= \left\{ A(\mathbf{H}) \in \Omega^1 : v^{(i)} \text{ is continuous and } (\sigma - 1)\text{-time differentiable almost} \right. \\ &\quad \left. \text{everywhere over } [0, 1], (-1)^i v^{(i)}(p) > 0 \forall p \in [0, 1], v^{(i)}(1) = 0, \forall i = 1, 2, \dots, \sigma - 1 \right\}, \end{aligned} \quad (16)$$

and the following sets of socioeconomic health inequality measures:

$$\begin{aligned} \Xi^\sigma &:= \left\{ I(\mathbf{H}) : v^{(i)} \text{ is continuous and } (\sigma - 1)\text{-time differentiable almost} \right. \\ &\quad \left. \text{everywhere over } [0, 1], (-1)^i v^{(i)}(p) > 0, v^{(i)}(1) = 0, \forall p \in [0, 1], \forall i = 1, 2, \dots, \sigma - 1 \right\}. \end{aligned} \quad (17)$$

Thus, a population health measure $A(\mathbf{H}) \in \Omega^1$ satisfies the Pen Parade Principle. A population health measure $A(\mathbf{H}) \in \Omega^2 \subset \Omega^1$ and its associated socioeconomic health inequality index $I(\mathbf{H}) \in \Xi^2$, satisfies the Pigou-Dalton Principle of Transfers. Also, $A(\mathbf{H}) \in \Omega^3 \subset \Omega^2 \subset \Omega^1$ and $I(\mathbf{H}) \in \Xi^3 \subset \Xi^2$ satisfy the Principle of Positional Transfer Sensitivity. Further, $A(\mathbf{H}) \in \Omega^\sigma \subset \Omega^{\sigma-1} \subset \dots \subset \Omega^2 \subset \Omega^1$ and $I(\mathbf{H}) \in \Xi^\sigma \subset \Xi^{\sigma-1} \subset \dots \subset \Xi^3 \subset \Xi^2$ for all $\sigma \in \{3, 4, \dots\}$, satisfy the Principle of σ th-degree Positional Transfer Sensitivity (see Aaberge (2009), Theorem 3.2A).

Having established that the proposed class of indices satisfy the rank dependant ethical principles, let us now introduce the concept of σ -Generalized Concentration curves and σ -Modified Concentration curves that are necessary to establish positional stochastic dominance conditions

Definition 5 *Let the second-order generalized concentration curve be defined by $GC^2(p) = \int_0^p s(\mathbf{H}(\mathbf{u}))du$.*

The σ -generalized concentration curve, $GC^\sigma(p) = \int_0^p GC^{\sigma-1}(u)du$, is given for all integers $\sigma \in \{3, 4, \dots\}$.

At this point it is interesting to highlight that the second-order generalized concentration curve at p is equal to the standard generalized concentration curve of the width of health problems at p multiplied by the average width of health problems in the population, μ_s .

Definition 6 *Let the second-order modified concentration curve be defined by $\tilde{C}^2(p) = \int_0^p \frac{s(\mathbf{H}(\mathbf{u}))}{\mu_\phi} du$.*

The σ -modified concentration curve, $\tilde{C}^\sigma(p) = \int_0^p \tilde{C}^{\sigma-1}(u)du$, is given for all integers $\sigma \in \{3, 4, \dots\}$.

Note that there is a difference between the σ -modified concentration curves, $\tilde{C}^\sigma(p)$, and the σ -concentration curves, $C^\sigma(p)$ introduced by Makdissi and Mussard (2008) where $\tilde{C}^\sigma(p) = \frac{\mu_s}{\mu_\phi} C^\sigma(p)$.⁹

In what follows we present and prove our positional stochastic dominance conditions.

Theorem 1 *$A(\tilde{\mathbf{H}}) \geq A(\mathbf{H})$ for all $A(\cdot) \in \Omega^\sigma$, $\sigma \in \{2, \dots\}$ if and only if*

$$GC_H^\sigma(p) - GC_{\tilde{H}}^\sigma(p) \geq 0, \forall p \in [0, 1] \quad (18)$$

⁹Also note that the Makdissi and Mussard (2008) second order concentration curve corresponds to the usual health concentration curve that is often used in the health inequality literature.

Proof. We start by determining under which conditions $\Delta A_{\tilde{\mathbf{H}},\mathbf{H}} = A(\tilde{\mathbf{H}}) - A(\mathbf{H})$ is non negative.

Using equation (6), we get

$$\Delta A_{\tilde{\mathbf{H}},\mathbf{H}} = \frac{1}{K} \int_0^1 \nu(p) s(\mathbf{H}(\mathbf{p})) dp - \frac{1}{K} \int_0^1 \nu(p) s(\tilde{\mathbf{H}}(\mathbf{p})) dp. \quad (19)$$

From (19), we infer that $\Delta A_{\tilde{\mathbf{H}},\mathbf{H}} \geq 0$ if and only if

$$\int_0^1 \nu(p) [s(\mathbf{H}(\mathbf{p})) - s(\tilde{\mathbf{H}}(\mathbf{p}))] dp = \int_0^1 \nu(p) [GC_H^1(p) - GC_{\tilde{H}}^1(p)] \geq 0. \quad (20)$$

Let us now integrate by parts $\int_0^1 \nu(p) GC_H^1(p) dp$:

$$\int_0^1 \nu(p) GC_H^1(p) dp = \nu(p) GC_H^2(p) \Big|_0^1 - \int_0^1 \nu'(p) GC_H^2(p) dp. \quad (21)$$

Using the definition of Ω^2 in (16) and Definition 5, we can deduce that the first term on right hand side of the equation is nil. This means that:

$$\int_0^1 \nu(p) GC_H^1(p) dp = - \int_0^1 \nu'(p) GC_H^2(p) dp. \quad (22)$$

Assume that at $\sigma - 1$ we have

$$\int_0^1 \nu(p) GC_H^1(p) dp = (-1)^{\sigma-2} \int_0^1 \nu^{(\sigma-2)}(p) GC_H^{\sigma-1}(p) dp, \quad (23)$$

where $\nu^{(i)}(p)$ denotes the i th derivative of $\nu(p)$. Integrating by parts equation (23)

$$\int_0^1 \nu(p) GC_H^1(p) dp = (-1)^{\sigma-2} \left\{ \nu^{(\sigma-2)}(p) GC_H^\sigma(p) \Big|_0^1 - \int_0^1 \nu^{(\sigma-1)}(p) GC_H^\sigma(p) dp \right\}. \quad (24)$$

Using the definition of Ω^σ in (16) and Definition 5, we can deduce that the first term on right hand side of the equation is nil. once again. This means that:

$$\int_0^1 \nu(p) GC_H^1(p) dp = (-1)^{\sigma-1} \int_0^1 \nu^{(\sigma-1)}(p) GC_H^\sigma(p) dp. \quad (25)$$

Given that equations (20) and (22) conform to the relation depicted in equation (23) it follows that equation (25) is true for all integer $\sigma \in \{2, \dots\}$. Using this result, we can state that, if $A \in \Omega^\sigma$, then $\Delta A_{\tilde{\mathbf{H}},\mathbf{H}} \geq 0$ if and only if

$$(-1)^{\sigma-1} \int_0^1 \nu^{(\sigma-1)}(p) [GC_H^\sigma(p) - GC_{\tilde{H}}^\sigma(p)] \geq 0. \quad (26)$$

Note that $(-1)^{\sigma-1}\nu^{(\sigma-1)}(p)$ is positive. This implies that if $GC_H^\sigma(p) - GC_{\tilde{H}}^\sigma(p) \geq 0$ for all $p \in [0, 1]$, then $\Delta A_{\tilde{\mathbf{H}}, \mathbf{H}} \geq 0$. This proves for sufficiency of the condition.

Having provided the sufficiency condition let us now prove the necessity of the condition. Consider the set of functions $A \in \Omega^\sigma$ for which the $(\sigma - 1)$ th derivative of $v(p)$ ($v^{(\sigma-1)}(p)$ being $v(p)$ itself) is of the following form:

$$v^{(\sigma-1)}(p) = \begin{cases} (-1)^{\sigma-1} \epsilon & p \leq \bar{p} \\ (-1)^{\sigma-1} (\bar{p} + \epsilon - p) & \bar{p} < p \leq \bar{p} + \epsilon \\ 0 & p > \bar{p} + \epsilon \end{cases} . \quad (27)$$

Since $v(p)$ is differentiable almost everywhere except at \bar{p} and $\bar{p} + \epsilon$, it satisfies the conditions in (16). Thus the population health indices, whose weight functions $v(p)$ have the particular form presented above for $v^{(\sigma-1)}(p)$, belong to Ω^σ . This yields:

$$v^{(\sigma)}(p) = \begin{cases} 0 & p \leq \bar{p} \\ (-1)^\sigma & \bar{p} < p \leq \bar{p} + \epsilon \\ 0 & p > \bar{p} + \epsilon \end{cases} . \quad (28)$$

Imagine now that $GC_H^\sigma(p) - GC_{\tilde{H}}^\sigma(p) < 0$ on an interval $[\bar{p}, \bar{p} + \epsilon]$ for ϵ that can be arbitrarily close to 0. For $v(p)$ defined as in (27), expression (26) leads to a decrease in population health. Hence, it cannot be that $GC_H^\sigma(p) - GC_{\tilde{H}}^\sigma(p) < 0$ for $p \in [\bar{p}, \bar{p} + \epsilon]$. ■

Theorem 2 $I(\tilde{\mathbf{H}}) \leq I(\mathbf{H})$ for all $I(\cdot) \in \Xi^\sigma$, $\sigma \in \{2, 3, \dots\}$ if and only if

$$\tilde{C}_H^\sigma(p) - \tilde{C}_{\tilde{H}}^\sigma(p) \geq 0, \forall p \in [0, 1] \quad (29)$$

Proof. By construction:

$$\Delta I_{\tilde{\mathbf{H}}, \mathbf{H}} = I(\tilde{\mathbf{H}}) - I(\mathbf{H}) \quad (30)$$

$$= 1 - \frac{A(\tilde{\mathbf{H}})}{\mu_{\phi\tilde{\mathbf{H}}}} - 1 + \frac{A(\mathbf{H})}{\mu_{\phi\mathbf{H}}} \quad (31)$$

$$= \frac{A(\mathbf{H})}{\mu_{\phi\mathbf{H}}} - \frac{A(\tilde{\mathbf{H}})}{\mu_{\phi\tilde{\mathbf{H}}}} \quad (32)$$

$$= 1 - \frac{1}{K} \int_0^1 \nu(p) \frac{s(\mathbf{H}(\mathbf{p}))}{\mu_{\phi\mathbf{H}}} dp - 1 + \frac{1}{K} \int_0^1 \nu(p) \frac{s(\tilde{\mathbf{H}}(\mathbf{p}))}{\mu_{\phi\tilde{\mathbf{H}}}} dp \quad (33)$$

$$= \frac{1}{K} \int_0^1 \nu(p) \frac{s(\tilde{\mathbf{H}}(\mathbf{p}))}{\mu_{\phi\tilde{\mathbf{H}}}} dp - \frac{1}{K} \int_0^1 \nu(p) \frac{s(\mathbf{H}(\mathbf{p}))}{\mu_{\phi\mathbf{H}}} dp \quad (34)$$

$$= \frac{1}{K} \int_0^1 \nu(p) \tilde{C}_H^s(p) dp - \frac{1}{K} \int_0^1 \nu(p) \tilde{C}_H^s(p) dp \quad (35)$$

From (35), we deduce that $\Delta I_{\tilde{\mathbf{H}}, \mathbf{H}} \leq 0$ if and only if

$$\int_0^1 \nu(p) \left[\tilde{C}_H^\sigma(p) - \tilde{C}_H^\sigma(p) \right] dp \leq 0. \quad (36)$$

Equation (36) has the same structure than equation (20) except that it has the opposite inequality sign. Sufficiency and necessity are proved by the result in Theorem 1. ■

The positional dominance tests presented in Theorem 1 and 2 are valid for a given weight vector Θ . In a situation where there is uncertainty about the weights to be assigned to each health status, these dominance tests ought to be performed for all potential values of Θ . In epidemiology setting $\Theta = [1, 1, 1, \dots]$ is a common practice (see Kopec and Willison, 2003).¹⁰

4 A parametric class

Applying the positional stochastic dominance conditions in Theorems 1 and 2 does not produce complete orderings between all possible distributions of health. This is why it is beneficial to use a parametric class of indices when complete orderings are required. For this purpose, we will adapt Wagstaff's (2002) parametric class of health achievement and inequality indices to the context of distributions of the width of health problems. An interesting aspect of this adjustment resides in its

¹⁰We will be using this particular form of weights in the empirical illustration section.

ability to turn this parametric class into a class that is invariant to any monotonic transformation $g(\cdot)$ that may be applied to $\mathbf{H}(\mathbf{p})$. Therefore, the obtained indices's values will no longer be arbitrary. The population health indices are given by

$$A_\eta(\mathbf{H}) = 1 - \frac{1}{K} \int_0^1 \eta(1-p)^{\eta-1} s(\mathbf{H}(\mathbf{p})) dp, \quad \eta > 1, \quad (37)$$

where the parameter η can be interpreted as a parameter of socioeconomic health inequality aversion (see Yitzhaki, 1983). The socioeconomic health inequality index that is associated with the population health index described in equation (37) is given by

$$I_\eta(\mathbf{H}) = \frac{\int_0^1 (\eta(1-p)^{\eta-1} - 1) s(\mathbf{H}(\mathbf{p})) dp}{K - \int_0^1 s(\mathbf{H}(\mathbf{p})) dp}, \quad \eta > 1. \quad (38)$$

Proposition 1 $A_\eta(\mathbf{H})$ satisfies the Pen Principle and the Pigou-Dalton Principle. In addition, it satisfies all the Principle of σ th-degree Positional Transfer Sensitivity, $\sigma \in \{3, 4, \dots\}$ if $\eta > \sigma - 1$.

Proposition 2 $I_\eta(\mathbf{H})$ satisfies the Pigou-Dalton Principle. In addition, it satisfies all the Principle of σ th-degree Positional Transfer Sensitivity, $\sigma \in \{3, 4, \dots\}$ if $\eta > \sigma - 1$.

Up to this point, all the indices that have been presented in this paper are blind to health status (see Makdissi and Yazbeck, 2011). As all these indices are linear in health status, they will react favourably when a health transfer is made from an individual at a lower rank in the health distribution to a person at a higher rank, provided that the former has a slightly higher income. Following Makdissi and Yazbeck (2011) we can perform the following transformation on the population health measure:

$$A_{\eta,\varepsilon}(\mathbf{H}) = \begin{cases} 1 - \frac{1}{K} \int_0^1 \eta(1-p)^{\eta-1} \frac{(s(\mathbf{H}(\mathbf{p})))^{1-\varepsilon}}{1-\varepsilon} dp & \text{for } \nu \geq 1, \varepsilon \geq 0 \text{ and } \varepsilon \neq 1 \\ 1 - \frac{1}{K} \int_0^1 \eta(1-p)^{\eta-1} \ln s(\mathbf{H}(\mathbf{p})) dp & \text{for } \nu \geq 1, \varepsilon = 1 \end{cases} \quad (39)$$

Using the population health measure in (39) solves the problem of blindness to health status if we set $\varepsilon > 0$. However, since $s(\mathbf{H}(\mathbf{p}))$ is not necessarily monotonic in p , the stochastic dominance condition developed in Section 3 will not hold to this class of indices.

5 Empirical illustration

The objective of this section is to illustrate how the theoretical results of this paper may be implemented empirically. We conduct a two-fold empirical application. The first focuses on between country comparisons. It compares population health and socioeconomic health inequality between the United States of America and Canada. The second looks at within country comparisons. It compares geographical differences in population health and socioeconomic health inequality within the United States.

5.1 The Data

5.1.1 The Joint Canada United States Health Survey(JCUSH)

We use the Joint Canada/United States Surveys of Health 2004 to conduct the between country comparison. This survey has 8688 observations of which 3505 are Canadian residents and 5183 are US residents. It covers individuals between the age 18 and 85 years and provides information about their clinical condition as well as their demographic characteristics and their socioeconomic status. We use information on household income to infer the socioeconomic rank of the individual. As for the information on the individual's health status, it is based on eight categorical variable covering eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. We fix the vector of threshold as described in details in Table 1.

5.1.2 National Health Interview Survey(NHIS)

We utilize the public use files of the National Health Interview Survey (NHIS) for 2010. The NHIS has monitored the health of the United States of America since 1957. It is a cross-sectional household interview survey that is representative of households and non-institutional group quarters. The NHIS contains data on a broad range of health topics that are collected through personal household interviews. We use the representative sub-sample for which the Quality of Life questionnaire has

Table 1: Attributes Threshold Definitions (JCUSH)

Attribute	Threshold met if:
Vision	the problem cannot be corrected by glasses or contact lenses.
Hearing	the problem hearing problem cannot be corrected with a hearing aid.
Speech	the individual cannot be fully understood in his native language.
Ambulation	the individual has some mobility problems.
Dexterity	the individual has some dexterity.
Emotion	the individual is “somewhat unhappy”.
Cognition	the individual has some difficulty thinking.
Pain	pain prevents some activities.

been filled and focus on adult population only.¹¹ The total sample size after these restriction is 6005 observations. In this survey, information on household income is used to infer the socioeconomic rank of the individual. As for the information on the health status of the individual, it is based on eleven categorical variables: vision, hearing, ambulation, communication, cognition, dexterity, learning capacity, anxiety, depression, pain and exhaustion. We fix the vector of threshold as detailed in Table 2.

5.2 Results

5.2.1 Empirical application using JCUSH survey.

Using the information presented in the data section we compute the indices developed in the theoretical section. Table 3 presents the estimated values of the population health indices presented

¹¹To remain consistent with the previous sample restriction using JCUSH.

Table 2: Attributes Threshold Definitions (NHIS)

Attribute	Threshold met if positively answered the following questions. Do you have:
Vision	difficult vision even with Glasses?
Hearing	difficult hearing even with a hearing Aid?
Ambulation	difficulty walking or climbing steps?
Communication	difficulty in understanding or being understood in your native language?
Cognition	difficulty remembering or concentrating?
Dexterity	difficulty with self care (washing all over, dressing)?
Learning Capacity	learning the rules of a new game? OR following instructions (use a new cell phone get to a new place)?
Pain	frequent pain?
Attribute (Emotional)	Threshold met if answer is different from "never" to the following questions
Anxiety	How often do you feel worried, nervous, anxious.
Depression	How often do you feel depressed.
Exhaustion	How often did you feel very tired exhausted (in past 3 months)

in equation (37) for Canada and the U.S. We consider four different values of η between 1.5 and 3. For all estimated values of the population health indices, Canada has a higher level of population health than the United States of America. We check the robustness of this conclusion by performing

the stochastic dominance test suggested in Theorem 1 (see left hand panel in Figure 1). This test indicates that Canada outperforms the United States' population health for all population health indices that belongs to Ω^2 (i.e., display inequality aversion).

Table 3: Population Health Indices for Canada and the U.S.A.

	$\eta = 1.5$	$\eta = 2$	$\eta = 2.5$	$\eta = 3$
Canada	0.93183	0.91869	0.90987	0.90453
USA	0.92158	0.90640	0.89596	0.88917

Figure 1: US vs Canada

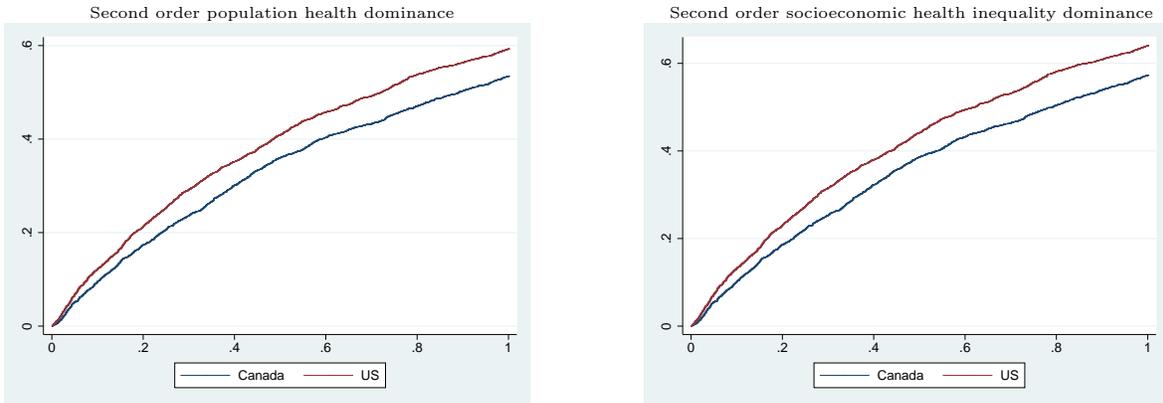


Table 4 presents the estimated values of the socioeconomic health inequality presented in equation (38) for both countries. As earlier, we consider four different values of η between 1.5 and 3. For all the estimated values of those inequality indices, Canada has a lower level of socioeconomic health inequality than the United States of America. To check the robustness of this conclusion, we perform the stochastic dominance test suggested in Theorem 2 (see right hand panel in Figure 1) This test indicates that Canada outperforms the United States on socioeconomic health inequality for all indices that belongs to Ξ^2 .

Table 4: Socioeconomic Health Inequality Indices for Canada and the U.S.A.

	$\eta = 1.5$	$\eta = 2$	$\eta = 2.5$	$\eta = 3$
Canada	0.00141	0.01549	0.02494	0.03066
USA	0.00462	0.02102	0.03230	0.03963

5.2.2 Empirical application using NHIS.

Having performed between countries comparisons, we now turn to compare the geographical differences in population health and socioeconomic health inequality within the United States.

Table 5: Population Health Indices for US Regions

	$\eta = 1.5$	$\eta = 2$	$\eta = 2.5$	$\eta = 3$
Northeast	0.74863	0.74037	0.73403	0.72911
Midwest	0.74045	0.73708	0.73530	0.73448
South	0.74162	0.73324	0.72753	0.72348
West	0.75685	0.75275	0.75010	0.74834

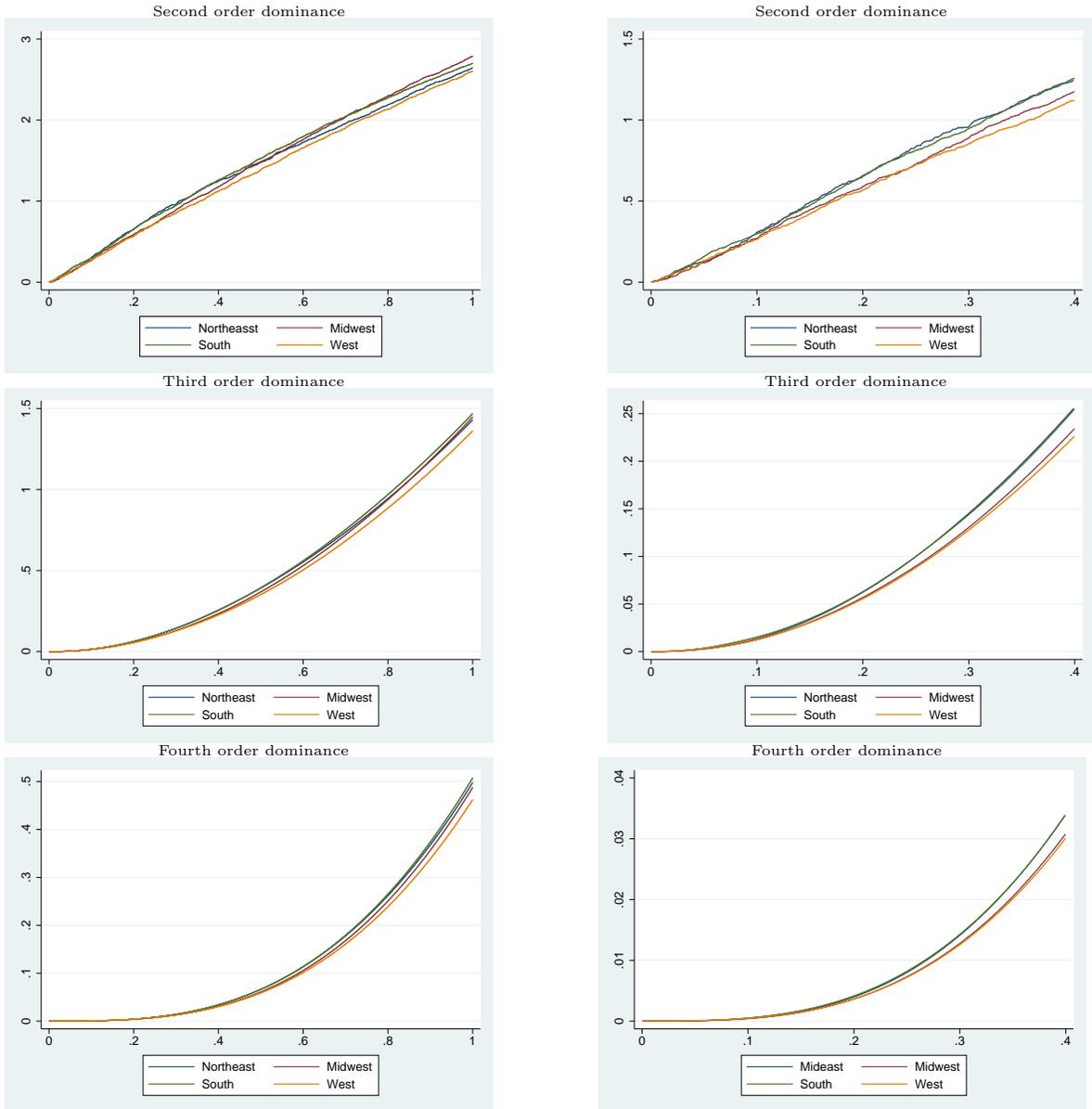
Table 5 presents the estimated values of the population health indices as presented in equation (37) for US regions. As in the previous section, we consider four different values of η between 1.5 and 3. We notice that for all estimated values of the population health indices, the West has a higher level of population health than all other regions and that this the only robust ranking between the regions. We then perform the stochastic dominance test as suggested in Theorem 1 to check the robustness of the only consistent result across values of η in Table 5 (see Figure 2). The second order dominance test indicates that the West outperforms only the South at the population health level, when all population health indices that belongs to Ω^2 are considered. The third order dominance test point to an additional restricted robust ranking: the Midwest outperforms only the South in population health when we consider all population health indices that belongs to Ω^3 . The fourth order dominance test does not add any robust ranking.

Table 6: Socioeconomic Health Inequality Indices for US Regions

	$\eta = 1.5$	$\eta = 2$	$\eta = 2.5$	$\eta = 3$
Northeast	0.01433	0.02520	0.03355	0.04003
Midwest	0.00805	0.01260	0.01498	0.01608
South	0.01713	0.02823	0.03580	0.04117
West	0.00837	0.01374	0.01721	0.01952

Table 6 presents the estimated values of the socioeconomic health inequality indices as presented in equation (38) for US regions. Again, we consider four different values of η between 1.5 and 3. For

Figure 2: Population Health Dominance Between US Regions

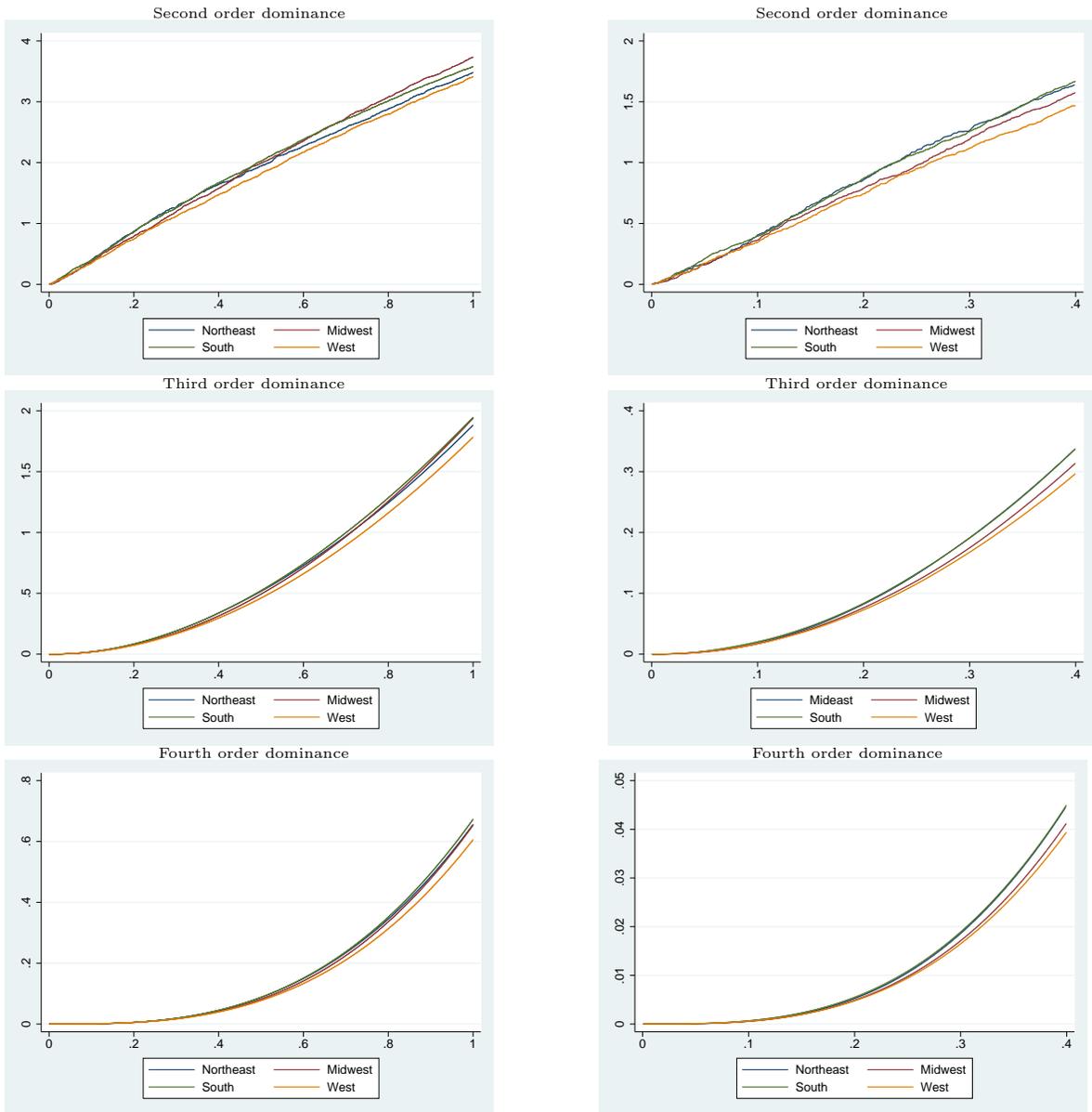


all estimated values of the socioeconomic health inequality indices, the Midwest has a lower level of socioeconomic health inequality, the West has the second lowest values followed by the South and then the Northeast. We also perform the stochastic dominance test suggested in Theorem 2 to check the robustness of this ranking from Table 6, these results are summarized in Figure 3. The second order dominance test indicates that the West outperforms only the South in socioeconomic health inequality when we consider all indices that belongs to Ξ^2 . The third order dominance test point to an additional restricted robust ranking: the Midwest outperforms only the South in socioeconomic health inequality when we consider all population health indices that belongs to Ξ^3 . The fourth order dominance test does not add any robust ranking.

6 Conclusion

In this paper we establish that the concentration index values may be arbitrary when categorical variables are used. To address this issue we change the dimension in which categorical data is exploited and focus on the breadth of the information it can provide. This allows us to obtain a ratio-scaled variable for which the arbitrariness will no longer be an issue. We then develop positional stochastic dominance conditions that can be implemented in a context of categorical variables. We also propose a parametric class of population health and socioeconomic health inequality indices and provide two empirical illustrations.

Figure 3: Socioeconomic Health Inequality Dominance Between US Regions



References

- [1] Aaberge, R. (2009), Ranking Intersecting Lorenz Curves, *Social Choice and Welfare*, 33, 235-259.
- [2] Alkire, S. and J. Foster (2011), Counting and multidimensional poverty measurement, *Journal of Public Economics*, 95, 476-487.
- [3] Atkinson, A.B. (1970), On the measurement of inequality, *Journal of Economic Theory*, 2, 244-263.
- [4] Beckman, S., W.J. Smith and B. Zheng (2009), Measuring inequality with interval data, *Mathematical Social Science*, 58, 25-34.
- [5] Brazier, J., T. Usherwood, R. Harper et al. (1998), Deriving a preference-based single index from UK SF-36 Health Survey, *Journal of Clinical Epidemiology*, 51, 1115-1128.
- [6] Clarke, P.M., U.G. Gerdtham, M. Johannesson, K. Bingenfors and L. Smith (2002) On the measurement of relative and absolute income-related health inequality. *Social Science and Medicine*, 55, 1923-1928.
- [7] Erreygers, G. (2006), Beyond the health Concentration Index: an Atkinson alternative for the measurement of the socioeconomic inequality of health, in *Paper Presented at the Conference Advancing Health Equity*, Helsinki, WIDER-UN.
- [8] Erreygers, G. (2009), Correcting the concentration index, *Journal of Health Economics*, 28, 504-515.
- [9] Fishburn, P.C. and R.D. Willig (1984), Transfer principles in income redistribution, *Journal of Public Economics*, 25, 323-328.

- [10] Kaplan, R.M., J.W. Bush and C.C. Berry (1976), Health status: types of validity and the index of well-being, *Health Service Research*, 1976, 478-507.
- [11] Kolm, J.-C. (1976), Unequal Inequalities I, *Journal of Economic Theory*, 12, 416-442.
- [12] Kopec, J.A. and K.D. Willison (2003), A comparative review of four preference-weighted measures of health-related quality of life, *Journal of Clinical Epidemiology*, 56, 317-325.
- [13] Lambert, P. and B. Zheng (2011), On the consistent measurement of attainment and shortfall inequality, *Journal of Health Economics*, 30, 214-219.
- [14] Le Grand, J. (1989), An international comparison of distributions of ages-at-death, in Fox, J. (Ed.), *Health Inequalities in European Countries*, Gower, Aldershot.
- [15] Le Grand, J. and M. Rabin (1986), Trends in British health inequality 1931-1983, in Culyer, A.J. and B. Jonsson (Eds.), *Public and Private Health Services*, Blackwell, Oxford.
- [16] Makdissi, P. and S. Mussard (2008), Analyzing the impact of indirect tax reforms on rank dependant social welfare functions: a positional dominance approach, *Social Choice and Welfare*, 30, 385-399.
- [17] Makdissi, P. and M. Yazbeck (2011), A class of health achievement and health inequality indices, Working Paper #1207E, Department of Economics, University of Ottawa
- [18] The EuroQol Group (1990), EuroQol: a new facility for the measurement of health-related quality of life, *Health Policy*, 16, 199-208.
- [19] Torrance, G.W., D.H. Feeny, W.J. Furlong et al. (1996), A multiattribute utility function for comprehensive health status classification system: Health Utility Index Mark 2, *Medical Care*, 34, 702-722.

- [20] Wagstaff, A. (2002), Inequality aversion, health inequalities and health achievement, *Journal of Health Economics*, 21, 627-641.
- [21] Wagstaff, A., , P. Paci and E. van Doorslaer (1991), On the measurement of inequalities in health, *Social Science and Medicine*, 33, 545-557.
- [22] Wagstaff, A., E. van Doorslaer and P. Paci (1989), Equity in the finance and delivery of health care: some tentative cross-country comparisons, *Oxford Review of Economic Policy*, 5, 89-112.
- [23] Yitzhaki, S. (1983), On an extension of the Gini index, *International Economic Review*, 24, 617-628.
- [24] Zheng, B. (2008), Measuring inequality with ordinal data, *Research on Economic Inequality*, 16, 177-188.
- [25] Zoli, C, (1999), Intersecting generalized Lorenz curves and the Gini index, *Social Choice and Welfare*, 16, 183-196.