

Measuring Health Inequality and Health Opportunity

Buhong Zheng

Department of Economics
University of Colorado at Denver and HSC
Denver, CO, USA
buhong.zheng@cudenver.edu

November 2006

Abstract: The paper considers the measurement of health opportunity with categorical data of health status. A society's health opportunity is represented by an income-health matrix that relates socioeconomic class with health status; each row of the matrix corresponds to a socioeconomic class and contains the respective probability distribution of health. In the first part of the paper, we formally demonstrate an important limitation in applying standard inequality criteria to distributions of health: without specifying the cardinal value for each health status, it is impossible to employ Lorenz dominance in measuring health inequality. In the second part of the paper, we argue that it is a more sensible approach to measure inequality of health opportunity. By introducing a monotone assumption on the income-health matrix, we derive a sequence of welfare-dominance conditions for health-opportunity comparisons. We then obtain dominance conditions for Lorenz curve-based inequality-rankings of health opportunities. Finally, we apply the results to compare health opportunities in the US and Canada using the newly released JCUSH data.

Key Words: health inequality, health opportunity, Lorenz dominance, income-health matrix, super-Lorenz dominance, income mobility.

Acknowledgment: An earlier version of the paper was presented in the UNU-WIDER conference on Advancing Health Equity held in Helsinki, Finland (September 29 - 30, 2006). I thank Eddy van Doorslaer, Guido Erreygers, James Foster, Tom van Ourti, Tony Shorrocks and James Smith for the helpful comments and suggestions.

Measuring Health Inequality and Health Opportunity

I. Introduction

Distributive justice concerns the fair distribution of social well-beings among the citizens of a society. A particularly important component of well-being is health since it contributes to an individual's capability to function (Sen, 1985). It has been argued that a society should be more concerned with "inequalities in health than in income" (Anand, 2004). In recent decades, issues on health distribution and its measurement have received immense interests from economists and, as a result, a large literature has been generated (see Wagstaff and van Doorslaer, 2000, for a recent survey). It seems that the broad literature of measurement of health distributions falls roughly into two areas. The first area of the research closely follows the measurement of income inequality and is solely concerned with how unevenly health condition varies across individuals in a population. Not surprisingly, almost all measures of income inequality find their way into the measurement of health inequality. The second (and more recent) area of the research does not measure inequality of health per se but rather tries to incorporate socioeconomic dimension into the measurement. Specifically, the research seeks to gauge "to what extent are there inequalities in health that are systematically related to socioeconomic status?" (Wagstaff et al., 1991). Many standard measures of income inequality, also not surprisingly, fail to reflect this consideration. In this sense, the measurement of health distributions can be quite different from the measurement of income distributions.

The measurement of health inequality differs from the measurement of income inequality in yet another important dimension – the use of categorical data of health status. While cardinal indicators of health such as mortality rate and life expectancy at birth are sometimes available for studies with specific aims, a comprehensive indicator of health status that is of a cardinal nature is rarely available. As a result, research and policy analyses have to rely upon self-assessed health status data. For example, both the US National Health Interview Survey (NHIS) and the Canadian National Population Health Survey (NPHS) adopt a five-category health rating: poor, fair, good, very good, and excellent. To compute an inequality index or apply an inequality dominance criterion, it is common to assign a cardinal number to each health category. For example, one can assign values 1 through 5 to the five health levels. Alternatively, one can assign such values through a more sophisticated channel (e.g., Wagstaff and van Doorslaer, 1994). But no matter in what way the values are assigned, the arbitrariness involved in the procedure remains (e.g., one can argue for the use of values 101 to 105 instead of 1 to 5). In the case of categorical data, Allison and Foster (2004) convincingly demonstrated the drawback of using the mean health status – the standard practice in income inequality measurement – as the reference point for inequality measurement. They proposed the median health rating as the reference point and characterized a quite different inequality ordering criterion that is robust to different cardinal scales used for health status.

In this paper, first of all, we further document the difficulty in applying the traditional inequality methods to measuring and ranking inequality across health distributions. Lorenz dominance is probably the most popular and most important measure of income inequality. This is so because if two distributions can be rank ordered by the Lorenz criterion, then all summary indices possessing certain property, namely the Pigou-Dalton principle of transfers, will rank the two distributions in the same way. Lorenz dominance has been used extensively in ranking health distributions. A key result in the first part of the paper is that if health status is measured dichotomously, i.e., people are rated healthy or unhealthy, then Lorenz dominance will fail to rank any pair of health distributions. This result holds not just for the usual relative Lorenz criterion but also for the absolute and intermediate Lorenz criteria. When health status is evaluated using a system with more than two categories of health such as that used by NHIS, the same impossibility verdict remains if the cardinal value of each health status can vary freely (while the ordinal rankings of the health values remain unchanged). The difficulty in employing Lorenz dominance surely transmits into summary indices of inequality such as the Gini coefficient and the Atkinson family: for two non-degenerate distributions of health we can always find a set of cardinal health values and a class of inequality indices so that either distribution can be regarded as more unequal than the other.

This impossibility result may not be surprising at all to many researchers working in the area as the limitations with using categorical data have long been known. It nevertheless raises an important question: is the exercise of measuring health inequality meaningful after all? Of course, if a reasonable cardinal valuation of health can be established – such as the McMaster Health Utility Index Mark 3 (HUI3) – then health inequality can still, to some extent, be meaningfully measured by the Lorenz ordering criterion. But even here, the use of Lorenz curve and other inequality devices may not be completely justified. This is the case because the idea of Lorenz dominance rests upon the Pigou-Dalton principle of transfers: a rich-to-poor transfer of income reduces income inequality. Translated into the measurement of health distributions, this requires that a transfer of health level from a healthy person to someone less healthy is always preferred. Clearly, the health version of the Pigou-Dalton principle cannot be unconditionally justified as many people in the field have been aware (e.g., Bleichrodt and van Doorslaer, 2006). This concern has diverted researchers’ attention to the second area of health inequality measurement – how much inequality in health that we observe is due to unequal socioeconomic structure? Wagstaff et al. (1989) proposed the use of the concentration curve and its associated index as measures of health inequality.¹ If health status is perfectly rank correlated with socioeconomic status, then the concentration curve is the same as Lorenz curve; if the correlation is not perfect, then the two curves will diverge. Clearly, the concentration-curve

¹For recent contributions to the concentration measurement of health inequality, see, among others, van Doorslaer and Jones (2003), Fleurbaey (2005), Brommier and Stecklov (2002), and Bleichrodt and van Doorslaer (2006).

approach also demands a cardinal (ratio-scale) valuation of health status.

In the second part of the paper, we take up the same issue that first motivated Wagstaff et al. (1989) and provide dominance conditions that are free from any cardinal valuation of health status. We start from the observation that the relationship between health status and socioeconomic class is a stochastic one, because an individual’s health is a function of many factors and socioeconomic status is only one of them, albeit an important one. The stochastic relationship can be conveniently and intuitively represented by an income-health matrix that is akin to the transition matrix used in income mobility research. Suppose we consider a society with three socioeconomic classes (poor, middle, and rich) and three health statuses (poor, fair, and good). An example of income-health matrix is:

$$\begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} .$$

In the matrix, rows represent socioeconomic classes while columns indicate health statuses. Each row summarizes the health distribution of all members in that socioeconomic class; each column documents the composition of socioeconomic classes in that health status. For example, for a randomly selected individual from the “poor” socioeconomic class, the first row states that 40 percent probability that the person will have “poor” health, 30 percent “fair” health and 30 percent “good” health. In this sense, each row can be described as the distribution of “odds” or “opportunity” of health for a person born into that class. In this paper, we define rows of the income-health matrix as the profiles of “health opportunity” and the entire matrix displays the “health opportunity” of the society.²

The issue then becomes the comparison between different income-health matrices. If a society can choose between two health opportunities which are represented by two different income-health matrices, which one should it choose? Here we consider two sets of criteria: welfare and equity. Namely, we want to determine which matrix yields higher welfare in terms of health and which one is more equitable. Since an income-health matrix resembles a transition matrix in the measurement of income mobility, many approaches and results from the mobility measurement research can be considered for ranking health opportunities even though a transition matrix is usually a square matrix while an income-health matrix need not be square. An approach that is particularly appealing is the one characterized by Dardanoni (1993). Following Atkinson (1983), Dardanoni (1993) measures mobility by considering “how economic mobility influences social welfare.” In this paper, we ask a very similar

²Recently, Erreygers (2006) also employed the matrix in measuring health inequality but in a different sense. While we view each row of the matrix as a health opportunity profile, Erreygers treated each row as a risk profile and then derived an Atkinson-style inequality index. It is useful to point out that his index depends upon a specific utility function and a cardinal valuation of each health status – our results require none of them.

question: “how health opportunity influences social welfare of health?” In deriving the various dominance conditions, we opt for a broader class of social welfare functions than Dardanoni’s and also extend some of his findings. The dominance conditions Dardanoni derived are free from cardinal utility values assigned to each economic class, a feature our dominance conditions also possess. We then turn our exploration to inequality orderings of health opportunities. Here we are able to derive a sufficient condition and a set of necessary conditions for relative Lorenz dominance, and a necessary-and-sufficient condition for absolute Lorenz dominance.

The rest of the paper is organized as follows. The next section addresses measurement of health inequality and presents some possibility and impossibility results. Section III defines the income-health matrix and investigates the social welfare rankings of health opportunities. Section IV investigates the inequality rankings of health opportunities. Section V applies the results to health opportunities in the US and Canada using the newly released Joint Canada/US Survey of Health (JCUSH) data. Section VI concludes the paper.

II. The Measurement of Health Inequality

Consider a society having m health categories with $2 \leq m < \infty$. Denote h_i the cardinal value of health level i , $i = 1, 2, \dots, m$. Assume health statuses are listed in increasing order, i.e., from the poorest health to the best health and $h_1 \leq h_2 \leq \dots \leq h_m$. Also assume that all health values are positive, i.e., $h_1 > 0$, which may be interpreted as an assumption of “life is worth living.”³ The proportion of people in each health class is denoted p_i with $\sum_{i=1}^m p_i = 1$. A *health distribution* is defined as $\mathbf{p} = (p_1, p_2, \dots, p_m)$ and the cumulative population proportion is $\tilde{p}_i = \sum_{j=1}^i p_j$. A health distribution is *degenerate* if $p_i = 0$ for some i ; a non-degenerate distribution has people filling into each health status. The average health level of distribution $\mathbf{p}$ is denoted $\bar{h}(\mathbf{p}) = \sum_{i=1}^m p_i h_i$. For all societies of interest, the health classifications are the same.

II.1. Generalized, relative and absolute Lorenz dominances: definitions

In the literature of income distributions, welfare comparisons between distributions are often carried out using generalized Lorenz dominance while inequality comparisons are performed with either relative Lorenz dominance or absolute Lorenz dominance.⁴

Generalized Lorenz dominance compares the generalized Lorenz curves between two health distributions. For each cumulative population share, $\tilde{p}_l$, the corresponding

³For welfare rankings, this assumption is not needed but for inequality orderings all health levels must be strictly positive in order to have a meaningful dominance condition.

⁴For a recent and thorough introduction to the various Lorenz dominances and the literature of income distributions in general, see Lambert (2001).

generalized Lorenz ordinate is the (normalized) cumulated total health level, i.e.,

$$GL(\mathbf{p}; \tilde{p}_l) = \sum_{i=1}^l p_i h_i.$$

The generalized Lorenz curve is the linear segment of the m Lorenz coordinates and the origin $(0, 0)$. It follows that for a cumulative population share p that is not in $\{0, \tilde{p}_1, \tilde{p}_2, \dots, 1\}$, its generalized Lorenz ordinate is simply an extrapolation from the nearest generalized Lorenz ordinate below p . For example, if $\tilde{p}_1 < p < \tilde{p}_2$, then

$$GL(\mathbf{p}; p) = p_1 h_1 + (p - p_1) h_2.$$

Distribution $\mathbf{p}$ generalized Lorenz dominates distribution $\mathbf{q}$ if and only if the generalized Lorenz curve of $\mathbf{p}$ lies nowhere below and somewhere strictly above that of $\mathbf{q}$.

Relative and absolute Lorenz dominances, respectively, compare relative and absolute Lorenz curves. The difference between the two approaches lies in the different ways that health distributions are normalized to have equal mean-health levels. Relative Lorenz curve remains unchanged to an equal-proportional change in all health levels (e.g., if all h_i s are increased by 10%) while absolute Lorenz curve remains the same when all health values are increased by an equal-absolute-amount (e.g., if all h_i s are increased by 0.1).⁵ Relative Lorenz ordinates $\{RL(\mathbf{p}; \tilde{p}_l)\}$ for distribution $\mathbf{p}$ are defined by

$$RL(\mathbf{p}; \tilde{p}_l) = \frac{\sum_{i=1}^l p_i h_i}{\bar{h}(\mathbf{p})},$$

and distribution $\mathbf{p}$ relative Lorenz dominates $\mathbf{q}$ if and only if the relative Lorenz curve of $\mathbf{p}$ lies nowhere below and somewhere strictly above that of $\mathbf{q}$. Absolute Lorenz ordinates $\{AL(\mathbf{p}; \tilde{p}_l)\}$ for distribution $\mathbf{p}$ are defined by

$$AL(\mathbf{p}; \tilde{p}_l) = \sum_{i=1}^l p_i h_i - \left(\sum_{i=1}^l p_i \right) \bar{h}(\mathbf{p}),$$

and distribution $\mathbf{p}$ absolute Lorenz dominates $\mathbf{q}$ if and only if the absolute Lorenz curve of $\mathbf{p}$ lies nowhere below and somewhere strictly above that of $\mathbf{q}$. Note that both relative and absolute Lorenz curves are derived from generalized Lorenz curve: relative Lorenz curve is obtained by scaling down generalized Lorenz curve by $\bar{h}(\mathbf{p})$,

⁵It is well-known that the two types of Lorenz dominance reflect two value judgements in measuring inequality. It is possible to consider a “balanced” or “intermediate” value judgement between the relative view and the absolute view. In a recent paper, Zheng (2006a) showed that under certain reasonable conditions, the only intermediate Lorenz curve is a weighted geometric mean between the relative Lorenz curve and the absolute relative Lorenz curve. All results derived in this section for relative and absolute Lorenz dominances remain valid for an intermediate Lorenz dominance.

while absolute Lorenz curve is obtained by shifting down generalized Lorenz curve by $\bar{h}(\mathbf{p})$.

Clearly, the location of any Lorenz curve (generalized, relative or absolute) depends on the values that h_j s are assigned and the dominance evaluation may hinge on the specific values used. Given the arbitrariness in assigning such values, it is important to investigate whether there are conditions under which dominance relations hold with respect to all possible values of health levels. In what follows, we explore the possibility and impossibility for generalized Lorenz dominance, relative and absolute Lorenz dominances.

II.2. Generalized Lorenz dominance: limited possibility

For two health distributions $\mathbf{p}$ and $\mathbf{q}$, $\mathbf{p}$ generalized Lorenz dominates $\mathbf{q}$ if and only if

$$GL(\mathbf{p}; r) \geq GL(\mathbf{q}; r)$$

for all $r \in [0, 1]$ with the strict inequality holding for some $r \in (0, 1]$. A necessary condition is $GL(\mathbf{p}; 1) \geq GL(\mathbf{q}; 1)$ – the two ending points of the two generalized Lorenz curves – or

$$\sum_{i=1}^m p_i h_i \geq \sum_{i=1}^m q_i h_i. \quad (2.1)$$

Using Abel's partial summation formula (Rudin, 1976, p. 70), (2.1) can be written as

$$\sum_{i=1}^m \left\{ \sum_{j=i}^m p_j \right\} (h_i - h_{i-1}) \geq \sum_{i=1}^m \left\{ \sum_{j=i}^m q_j \right\} (h_i - h_{i-1}) \quad (2.1a)$$

where $h_0 \equiv 0$. Abel's lemma states that for any $0 < h_1 \leq h_2 \leq \dots \leq h_m$, a sufficient condition for (2.1) is⁶

$$\sum_{j=i}^m p_j \geq \sum_{j=i}^m q_j \text{ for } i = 1, 2, \dots, m. \quad (2.2)$$

Condition (2.2) is obviously also necessary to maintain (2.1a) or (2.1). This is because if $\sum_{j=l}^m p_j < \sum_{j=l}^m q_j$ for some l ($1 \leq l \leq m$), then by choosing $h_1 = \dots = h_{l-1}$, $h_l = \dots = h_m$ and $h_l > h_{l-1}$ we will have the desired contradiction.

Since $\sum_{j=1}^m p_j = \sum_{j=1}^m q_j = 1$, (2.2) is equivalent to

$$\tilde{p}_j = \sum_{i=1}^j p_i \leq \sum_{i=1}^j q_i = \tilde{q}_j \text{ for } j = 1, 2, \dots, m-1, \quad (2.2a)$$

i.e., the cumulative frequency of health status is no greater in $\mathbf{p}$ than in $\mathbf{q}$.

⁶The result in (2.2) has been proved by Allison and Foster (2004, Theorem 1). We furnish a proof here since Abel's partial summation and Abel's lemma will be used repeatedly in the rest of the paper.

Condition (2.2a) also ensures that the entire generalized Lorenz curve of $\mathbf{p}$ lies nowhere below that of $\mathbf{q}$. This assertion can be directly verified as follows (though it will also become apparent shortly).

Since the generalized Lorenz curve of $\mathbf{p}$ is a linear segment of $(0, 0)$, $[\tilde{p}_1, GL(\mathbf{p}; \tilde{p}_1)]$, $[\tilde{p}_2, GL(\mathbf{p}; \tilde{p}_2)]$, ..., and $[1, \bar{h}(\mathbf{p})]$, it follows that $GL(\mathbf{p}; r) \geq GL(\mathbf{q}; r)$ for all $r \in (0, 1)$ holds if and only if $GL(\mathbf{p}; r) \geq GL(\mathbf{q}; r)$ holds for $r = \{\tilde{p}_1, \tilde{q}_1, \tilde{p}_2, \tilde{q}_2, \dots, \tilde{q}_{m-1}, 1\}$. Suppose $r \in \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_{m-1}\}$, say $r = \tilde{p}_l$. Then $GL(\mathbf{p}; r) = \sum_{i=1}^l p_i h_i$. If $\tilde{p}_l$ lies between $\tilde{q}_{k-1}$ and $\tilde{q}_k$ such that $\tilde{q}_{k-1} < \tilde{p}_l \leq \tilde{q}_k$, then

$$GL(\mathbf{q}; r) = \sum_{i=1}^{k-1} q_i h_i + (\tilde{p}_l - \tilde{q}_{k-1}) h_k.$$

Since $\tilde{p}_l \leq \tilde{q}_l$ by (2.2a), we must have $l \geq k$; otherwise $l \leq k-1$ and thus $\tilde{p}_l \leq \tilde{q}_{k-1}$ – a contradiction. Abel's lemma then entails $GL(\mathbf{p}; r) \geq GL(\mathbf{q}; r)$.

Suppose now $r \in \{\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_{m-1}\}$, say $r = \tilde{q}_l$. Then $GL(\mathbf{q}; r) = \sum_{i=1}^l q_i h_i$. If $\tilde{q}_l$ lies between $\tilde{p}_k$ and $\tilde{p}_{k+1}$ such that $\tilde{p}_k \leq \tilde{q}_l < \tilde{p}_{k+1}$, then

$$GL(\mathbf{p}; r) = \sum_{i=1}^k p_i h_i + (\tilde{q}_l - \tilde{p}_k) h_{k+1}.$$

Since $\tilde{p}_l \leq \tilde{q}_l$ by (2.2a), we must have $k \geq l$; otherwise $k+1 \leq l$ and thus $\tilde{p}_{k+1} \leq \tilde{q}_l$ – also a contradiction. Abel's lemma again entails $GL(\mathbf{p}; r) \geq GL(\mathbf{q}; r)$.

To ensure that the generalized Lorenz curve of $\mathbf{p}$ lies somewhere strictly above that of $\mathbf{q}$, i.e., $GL(\mathbf{p}; r) > GL(\mathbf{q}; r)$ for some r , we need to further require that the inequality in (2.2a) hold strictly for some $j = 1, 2, \dots, m-1$. The above derivations and discussions can be summarized formally into the following proposition.

Proposition 2.1. For any two health distributions $\mathbf{p}$ and $\mathbf{q}$, $\mathbf{p}$ generalized Lorenz dominates $\mathbf{q}$ if and only if condition (2.2a) holds for all $j = 1, 2, \dots, m$ and holds strictly for some $j = 1, 2, \dots, m-1$.

Condition (2.2a), however, is essentially a first-order stochastic dominance condition as derived in Allison and Foster (2004, Theorem 1). When generalized Lorenz dominance – which is a second-order stochastic dominance condition – is required to hold for all possible values of h_i s, the second-order condition simply collapses to the first-order condition. This observation holds for any higher-order stochastic dominance condition because such a condition would require a dominance of the mean between the two distributions which leads precisely to condition (2.2a).⁷ On the other hand, from the literature of stochastic dominance, (2.2a) implies any higher-order stochastic dominance. In this sense, it is futile to consider any higher-order stochastic dominance than first-order in measuring welfare of health distributions.

⁷For a detailed and up-to-date exposition on stochastic dominance, one may consult Levy (2006). In this paper, due to space limitation, we are unable to expand the discussions on stochastic dominance.

In this sense, we can only have limited possibility in welfare measurement as the following corollary documents.

Corollary 2.1. For any two health distributions $\mathbf{p}$ and $\mathbf{q}$, $\mathbf{p}$ welfare dominates $\mathbf{q}$ in any degree of stochastic dominance if and only if condition (2.2a) holds for all $j = 1, 2, \dots, m$ and holds strictly for some $j = 1, 2, \dots, m - 1$.

II.3. Lorenz dominances: impossibility

For two health distributions $\mathbf{p}$ and $\mathbf{q}$, $\mathbf{p}$ relative Lorenz dominates $\mathbf{q}$ if and only if

$$RL(\mathbf{p}; r) \geq RL(\mathbf{q}; r)$$

for all $r \in [0, 1]$ with the strict inequality holding for some $r \in (0, 1)$. Distribution $\mathbf{p}$ absolute Lorenz dominates $\mathbf{q}$ if and only if

$$AL(\mathbf{p}; r) \geq AL(\mathbf{q}; r)$$

for all $r \in [0, 1]$ with the strict inequality holding for some $r \in (0, 1)$. In what follows we first consider relative Lorenz dominance.

The impossibility result for relative Lorenz dominance stems from the following lemma in which only two health statuses are considered.

Lemma 2.1. Suppose there are only two health statuses (e.g., healthy and unhealthy) with $0 < h_1 < h_2$. Then for any non-degenerate health distributions $\mathbf{p}$ and $\mathbf{q}$, the two relative Lorenz curves must be either cross or identical; there is no possibility of dominance.

Proof. Since there are only two health statuses, $\mathbf{p} = (p_1, p_2)$ and $\mathbf{q} = (q_1, q_2)$. In order for the Lorenz curve of $\mathbf{p}$ to dominate that of $\mathbf{q}$, it must be the case that $RL(\mathbf{p}; r) \geq RL(\mathbf{q}; r)$ at $r = \{p_1, q_1\}$.

If $p_1 < q_1$, then $RL(\mathbf{p}; r) \geq RL(\mathbf{q}; r)$ at $r = p_1$ means

$$\frac{p_1 h_1}{p_1 h_1 + p_2 h_2} \geq \frac{p_1 h_1}{q_1 h_1 + q_2 h_2},$$

which requires $p_1 h_1 + p_2 h_2 \leq q_1 h_1 + q_2 h_2$ since $p_1 h_1 > 0$ by the non-degenerate assumption. But $p_1 < q_1$ implies $p_2 > q_2$ and, consequently, we must have

$$p_1 h_1 + p_2 h_2 = h_1 + p_2(h_2 - h_1) > h_1 + q_2(h_2 - h_1) = q_1 h_1 + q_2 h_2$$

for any h_1 and h_2 satisfying $0 < h_1 < h_2$ – a contradiction. It follows that we can only have $p_1 \geq q_1$. Condition $RL(\mathbf{p}; r) \geq RL(\mathbf{q}; r)$ at $r = p_1$ becomes

$$\frac{p_1 h_1}{p_1 h_1 + p_2 h_2} \geq \frac{q_1 h_1 + (p_1 - q_1) h_2}{q_1 h_1 + q_2 h_2}.$$

Expanding the inequality, we arrive at

$$(p_1 - q_1) p_2 h_1 \geq (p_1 - q_1) p_2 h_2.$$

Clearly, for $0 < h_1 < h_2$, the above weak inequality can hold only as an equality and only when either $p_1 = q_1$ or $p_2 = 0$. But $p_2 = 0$ means a degenerate distribution, i.e., only one health status for all individuals and hence is excluded. It follows that the only case remains is $p_1 = q_1$ and thus $\mathbf{p} = \mathbf{q}$. This means that the only case where two distributions do not cross is when they are identical. $\square$

The lemma is of interest in its own right: if a population is classified dichotomously into healthy group and unhealthy group, then there is no way to rank any distributions using relative Lorenz dominance if none of the distributions is degenerate. In the case where two distributions are not identical and neither one is degenerate, the two relative Lorenz curves will always cross. For a health-evaluating system with more than two categories, however, the impossibility presented in this lemma will disappear. A simple example can verify that this is indeed the case. In fact, many empirical studies have demonstrated the possibility. But when the non-uniqueness of the health value is considered, the possibility again turns into impossibility. This result is formally stated and proved as follows.⁸

Proposition 2.2. For a given health-evaluating system with m health statuses and two non-degenerate health distributions $\mathbf{p}$ and $\mathbf{q}$, the relative Lorenz curve of $\mathbf{p}$ lies nowhere below that of $\mathbf{q}$ for all possible values of health statuses satisfying $0 < h_1 \leq h_2 \leq \dots \leq h_m$ if and only if $\mathbf{p} = \mathbf{q}$.

Proof. The key to the proof is to group all m health classes into just two groups and then apply Lemma 2.1.

We first consider the lowest-health group versus the rest of the population by choosing $h_2 = \dots = h_m$. The distribution $\mathbf{p}$ becomes $(p_1, \sum_{i=2}^m p_i)$ which is non-degenerate. Do the same for distribution $\mathbf{q}$ and the distribution becomes $(q_1, \sum_{i=2}^m q_i)$. Let $h_2 > h_1$, Lemma 2.1 then implies that if the relative Lorenz curve of $\mathbf{p}$ lies nowhere below that of $\mathbf{q}$ then $p_1 = q_1$.

In general, if $p_i = q_i$ for $i = 1, 2, \dots, s$, then we must have $p_{s+1} = q_{s+1}$. This is proved by choosing $h_1 = h_2 = \dots = h_{s+1} < h_{s+2} = \dots = h_m$ and grouping the first $s+1$ health classes together against the remaining $m-s-1$ classes. Again by applying Lemma 2.1, we have $\sum_{i=1}^{s+1} p_i = \sum_{i=1}^{s+1} q_i$ which, together with $\sum_{i=1}^s p_i = \sum_{i=1}^s q_i$, implies $p_{s+1} = q_{s+1}$.⁹ Thus we have $p_i = q_i$ for $i = 1, 2, \dots, m$ or $\mathbf{p} = \mathbf{q}$. The sufficiency of the proposition is obvious. $\square$

⁸A general result can be established for all health distributions (including degenerate distributions). Since in reality a health distribution is unlikely to be degenerate (otherwise an entire health status is null; none of the health distributions that we have come across is degenerate), we only include in the paper the result for non-degenerate distributions. The general result, however, is available from the author upon request.

⁹If one desires strict inequality among the health values, i.e., $h_1 < h_2 < \dots < h_m$, then the described grouping process cannot be done. In this case, for a given s , we can construct a sequence of health values $\{h_1^{(k)}, h_2^{(k)}, \dots, h_m^{(k)}\}$ such that $h_1^{(k)} < h_2^{(k)} < \dots < h_m^{(k)}$ with $\lim_{k \rightarrow \infty} h_1^{(k)} = \lim_{k \rightarrow \infty} h_{s+1}^{(k)}$ and $\lim_{k \rightarrow \infty} h_{s+2}^{(k)} = \lim_{k \rightarrow \infty} h_m^{(k)}$. Thus we can group the population into two groups in the limit. The continuity of Lorenz curve in health values then carries through the proof.

For absolute Lorenz dominance, similar results can be established with similar proofs.

Lemma 2.2. Suppose there are only two health statuses (e.g., healthy and unhealthy) and $0 < h_1 < h_2$. Then for any two non-degenerate health distributions $\mathbf{p}$ and $\mathbf{q}$, the two absolute Lorenz curves must be either cross or identical; there is no possibility of dominance.

Proposition 2.3. For any health-evaluating system with m health statuses and two non-degenerate health distributions $\mathbf{p}$ and $\mathbf{q}$, the absolute Lorenz curve of $\mathbf{p}$ lies nowhere below that of $\mathbf{q}$ for all possible values of health statuses $h_1, h_2, \dots, h_m$ satisfying $0 < h_1 \leq h_2 \leq \dots \leq h_m$ if and only if $\mathbf{p} = \mathbf{q}$.

Why (limited) possibility for generalized Lorenz dominance but impossibility for relative and absolute Lorenz dominances? In other words, why it is possible to compare welfare across health distributions but not to compare inequality? The key point here is that generalized Lorenz curve does not hold the average health level constant but both relative and absolute Lorenz curves employ some normalization process so that the mean health levels become the same across different distributions. It is well known that if two distributions have the same mean, then generalized Lorenz dominance and relative (absolute) Lorenz dominance become the same. If we require equal-mean between the two distributions, say $\mathbf{p}$ and $\mathbf{q}$, then

$$\bar{h}(\mathbf{p}) = \sum_{i=1}^m p_i h_i = \sum_{i=1}^m q_i h_i = \bar{h}(\mathbf{q}).$$

Clearly, for the above equality to hold for all possible values of h_i s, $\mathbf{p}$ and $\mathbf{q}$ must be identical.

III. The Measurement of Health Opportunity: Welfare

The difficulties in measuring health inequality may be disappointing, but a more fundamental question is why do we want to measure health inequality. It is easy to be convinced about the welfare implications of health distribution as revealed through generalized Lorenz dominance. After all, if a society can enhance the overall level of health for its citizens, then it is always a good thing to have. But why do we prefer to have a more equal health distribution? And can the health level in a population be really equalized? Rawls's First Principle of Justice stipulates that all basic freedoms must be distributed equally throughout society. As noted by Brommier and Stecklov (2002), however, "Rawls himself labels health as a natural good and explicitly rules out health as a basic freedom" and from the list of goods and services desired to be equalized. The reason for this exclusion is that the health condition of an individual depends upon many factors. Among these factors some are the obligations of the individual (e.g., lifestyle and health consciousness), some may be the responsibility of the society (e.g., easy access to medical services) and some cannot be controlled

by anyone (e.g., unexpected illness and genetic defects). Clearly, both the first and third groups of factors cannot be equalized through any institutional efforts. What could be equalized in an equitable society is the second group of factors – mainly the factors related to socioeconomic status and income. Consequently, we can measure to what extent a society reduces health inequality that is due to unequal socioeconomic structure. The research in this direction leads us into the second area of health distribution measurement.

III.1. Income-health matrix and health opportunity profiles

Suppose the society can be further divided into n ($2 \leq n < \infty$) socioeconomic classes and the proportion of people in class i is π_i with $\sum_{i=1}^n \pi_i = 1$. The *socioeconomic structure* is $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ and the cumulative population proportion is $\tilde{\pi}_i = \sum_{j=1}^i \pi_j$. In general, the number of socioeconomic classes needs not be the same as the number of health statuses.

It is important to recognize that the relationship between an individual's income level and his health status is a stochastic one. That is, for a randomly selected individual from a socioeconomic class, we can only describe his health with a probability distribution – the “odds” for his health to be at all possible statuses. Since both socioeconomic class and health status are discrete in nature, it is appropriate to depict this stochastic relationship with a matrix. Like the ranking of health statuses, we also sort socioeconomic classes in increasing order, i.e., from the lowest to the highest. Denote α_{ij} the probability that an individual in socioeconomic class i will have health status j , an income-health matrix is $A_{n \times m} = |\alpha_{ij}|$ with $\alpha_{ij} \geq 0$ and $\sum_{j=1}^m \alpha_{ij} = 1$ for all $i = 1, 2, \dots, n$. Recalling that $\mathbf{p} = (p_1, p_2, \dots, p_m)$ is the health distribution of the society, we also have $p_j = \sum_{i=1}^n \pi_i \alpha_{ij}$.

Since each row of A represents the health prospect for the corresponding socioeconomic class, we define the row $(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{im})$ as the *health opportunity profile* for socioeconomic class i . Thus the income-health matrix A represents the *health opportunity* of the entire society. Assuming that all other factors in health determination are uncorrelated with socioeconomic status (e.g., people are biologically similar between the rich and the poor; potential genetic defects are nondiscriminatory and present equal threats to people of all socioeconomic origins), then the differences among the health opportunity profiles reflect nothing but the unequal socioeconomic structure.¹⁰

It is useful to point out that inequality in health opportunity and inequality in health status are completely different concepts; different matrices of income-health can be associated with the same distribution of health status. A perfectly equal health opportunity can go along with a very unequal health distribution (for a given set of health values).

¹⁰In practice, this uncorrelation assumption may not be satisfied. In this case, to isolate the effect of socioeconomic status on health, one needs to adjust the income-health matrix using appropriate econometric methods such as ordered probit regressions. In the absence of this adjustment, we are simply measuring socioeconomic inequality in health as most people in the area have attempted to measure. I thank Tom van Ourti for this remark.

III.2. Measuring health opportunity: welfare dominances

Suppose a society can choose between two health opportunities (or two different income-health matrices), which opportunity should the society choose? Or put it more vividly, if a person is “parachuted from a very great height” (Lambert, 2001, p. 85) into the society without knowing which socioeconomic class he will land, which health opportunity should he prefer? The United States and Canada have two very different health-care provision systems – the US uses mainly private provision for health care while Canada and many other countries adopt public provision systems. Imagine that the US adopts the Canadian health-care system, what will be the implications in terms of welfare and equality in health opportunity? This is a meaningful and timely question since the rising cost of health care in the US has made the Canadian health-care system very appealing. It has been reported (e.g., OECD, 2005) that Canadians live longer and are healthier than Americans and yet Canadians pay roughly half as much per capita as Americans do; the US also spends higher percentage of its GDP on health care than Canada does. The questions the existing literature did not address are the distributional aspects of health opportunities presented in the two countries. In this section, we develop tools for welfare comparisons between different health opportunities as represented by different income-health matrices. The methods for inequality comparisons will be developed in the following subsection.

The income-health matrix resembles the transition matrix employed in the studies of income mobility. A transition matrix – usually a square matrix – documents the dynamic relationship between incomes in two time periods. The welfare implications of different transition matrices have been investigated by Atkinson (1983), Dardanoni (1993) and others. Assuming the economy stays at a steady-state (the income distribution in each period remains the same) and employing a class of additive asymmetric social welfare functions, Dardanoni (1993) derived a set of powerful dominance conditions. Given the similarity between a transition matrix and an income-health matrix, we could apply the Atkinson-Dardanoni approach to evaluate health opportunities. In this paper, however, we opt for a more general approach which extends Dardanoni’s approach in two directions. First, we do not require the steady-state assumption (Formby et al. (2003) have already extended Dardanoni’s results by dropping his steady-state assumption) for an income-health matrix. This is very natural for health opportunity since a steady state in the health context would require health distribution to be identical to income distribution, i.e., $p_i = \pi_i$ and $m = n$, a condition that is neither justified nor required. Second, we establish our results by not relying upon additive social welfare functions. We invoke the generalized Lorenz dominance criterion which is founded upon a broad class of social welfare functions defined directly over health opportunity profiles.

But another important assumption employed by Dardanoni (1993) will be retained for health opportunity measurement. It is the monotone-matrix assumption. A transition matrix is monotone if each row of the matrix weakly stochastically (in first-order) dominates the row above it. In intergenerational income mobility the

assumption requires that, in the words of Dardanoni, “each child at time t is better off, in terms of stochastic dominance, by having a parent from state $i + 1$ than by having a parent from state i .” Translated into health opportunity, the monotone assumption is simply “richer people, in the sense of stochastic dominance, are healthier.” For its central role in our characterization, we formally state this assumption as follows.

The assumption of monotone income-health matrix. The income-health matrix is monotone in that

$$\sum_{j=1}^k \alpha_{sj} \geq \sum_{j=1}^k \alpha_{tj} \quad (3.1)$$

for all $s < t$ and $k = 1, 2, \dots, m$.

An immediate implication of this assumption is that an individual from a higher socioeconomic class will have better expected health than an individual from a lower class regardless of the cardinal valuations of health statuses. Algebraically,

$$e_s =: \sum_{j=1}^m \alpha_{sj} h_j \leq \sum_{j=1}^m \alpha_{tj} h_j := e_t, \quad (3.2)$$

and the vector of expected health levels $\mathbf{e} = (e_1, e_2, \dots, e_n)$ is always ranked in the non-decreasing order for all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$.

The monotone assumption – stated in (3.2) – finds strong supports in the fields of public health and epidemiology. There, the assumption is referred to as *the social gradient in health outcomes* – “the lower a group’s social position, the worse the average health status of its members” (Peter, 2004). The gradient phenomenon has long been observed and it has been confirmed in almost all health categories across cultures and over time.¹¹ For example, in his famous study of poverty in London at the end of 19th century, Rowntree (1901) described such a clear gradient in infant mortality along the social working classes that he defined. Interestingly, Rowntree also documented the health of a group of boys and girls using the categorical health statuses as we use today. He presented the relationships between the health status and social working class for the boys and girls, respectively, in two 3×4 matrices (Rowntree, 1901, p. 253) – the very same income-health matrix as we examine in this paper!¹² For both matrices, the monotone assumption (3.1) is satisfied! More recently, along with numerous other studies, Marmot et al. (1978) and Marmot and Shipley (1996) proved the existence of social gradients in all major causes of death in British civil servants. Using the self-rated categorical health status data, Marmot et al. (1991) also showed a sharp gradient among the civil servants in their Whitehall

¹¹An exception is breast cancer mortality in women in which the gradient is reversed – higher mortality for women in higher social classes (Marmot, 2004). Historically, cardiovascular disease also exhibited the reversed social gradient but the gradient changed its direction some 60 years ago.

¹²I owe Guido Erreygers for this reference.

II study. In the economics literature, Smith (1999) provided some thorough and convincing evidences for the social gradient in health using the PSID data.¹³

With the monotone assumption, it is easy to derive the welfare implications of different health opportunity scenarios. Suppose for a given socioeconomic structure, there are two alternative income-health matrices, $A_{n \times m} = |\alpha_{ij}|$ and $B_{n \times m} = |\beta_{ij}|$, to be considered. Denote the two vectors of expected health under the two matrices as $\mathbf{e}(A)$ and $\mathbf{e}(B)$, respectively, then we can state the following proposition.¹⁴

Proposition 3.1. For a given socioeconomic structure $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ and for all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$, the vector of expected health $\mathbf{e}(A)$ generalized Lorenz dominates vector $\mathbf{e}(B)$ if and only if

$$\sum_{i=1}^l \sum_{j=1}^k \pi_i \alpha_{ij} \leq \sum_{i=1}^l \sum_{j=1}^k \pi_i \beta_{ij} \quad (3.3)$$

for all $k = 1, 2, \dots, m$ and all $l = 1, 2, \dots, n$ with the inequality holding strictly for some pairs of (k, l) .

Proof. For a given socioeconomic structure $\boldsymbol{\pi}$, a generalized Lorenz ordinate of $\mathbf{e}(A)$ is

$$GL(A; \tilde{\pi}_l) = \sum_{i=1}^l \sum_{j=1}^m \pi_i \alpha_{ij} h_j.$$

The generalized Lorenz curve of $\mathbf{e}(A)$ dominates that of $\mathbf{e}(B)$ if and only if $GL(A; \tilde{\pi}_l) \geq GL(B; \tilde{\pi}_l)$ for all $l = 1, 2, \dots, n$ with the inequality holding strictly for some l . Thus for each l , we have

$$\sum_{i=1}^l \sum_{j=1}^m \pi_i \alpha_{ij} h_j \geq \sum_{i=1}^l \sum_{j=1}^m \pi_i \beta_{ij} h_j. \quad (3.4)$$

Since $h_1 \leq h_2 \leq \dots \leq h_m$, Abel's lemma and the necessary-side argument presented in Proposition 2.1 entail that the necessary and sufficient condition for (3.4) is

$$\sum_{i=1}^l \sum_{j=k}^m \pi_i \alpha_{ij} \geq \sum_{i=1}^l \sum_{j=k}^m \pi_i \beta_{ij} \quad (3.5)$$

for $k = 1, 2, \dots, m$. Further, since $\sum_{i=1}^l \sum_{j=1}^m \pi_i \alpha_{ij} = \sum_{i=1}^l \pi_i \sum_{j=1}^m \alpha_{ij} = \sum_{i=1}^l \pi_i$ and also $\sum_{i=1}^l \sum_{j=1}^m \pi_i \beta_{ij} = \sum_{i=1}^l \pi_i$, (3.5) is equivalent to (3.3). $\square$

¹³Smith (1999) emphasized that the casual relationship goes both ways – higher income level yields better health and better health produces more income. Marmot (2004) and Chandola et al. (2003) argued that although at the individual level health and social-status can influence each other, for social groups as a whole, the direction “is likely to be from social environment to illness, not the other way” (Marmot, 2004, p. 46). In international comparisons, Pritchett and Summers (1997) also concluded that richer countries lead to better health, not the other way around.

¹⁴The dominance condition presented here is the same as the central result derived by Dardanoni (1993, Theorem 1).

Since $p_j = \sum_{i=1}^n \pi_i \alpha_{ij}$, the distribution of health $\mathbf{p}$ in terms of the income-health matrix is $(\sum_{i=1}^n \pi_i \alpha_{i1}, \sum_{i=1}^n \pi_i \alpha_{i2}, \dots, \sum_{i=1}^n \pi_i \alpha_{im})$. Note that the generalized Lorenz dominance condition characterized in Proposition 2.1, $\sum_{j=1}^k \sum_{i=1}^n \pi_i \alpha_{ij} \leq \sum_{j=1}^k \sum_{i=1}^n \pi_i \beta_{ij}$, is simply condition (3.3) with $l = n$. This observation gives rise to the following corollary which reveals that generalized Lorenz dominance of health opportunity is a stronger notion than generalized Lorenz dominance of health distribution.

Corollary 3.1. For any two income-health matrices A and B , if A generalized Lorenz dominates B in the sense of (3.3), then the health distribution of A also generalized Lorenz dominates that of B in the sense of (2.1).

It is well-known that generalized Lorenz dominance is equivalent to welfare dominance by all social welfare functions satisfying, in addition to other regularity conditions, the monotonicity axiom and the Pigou-Dalton principle of transfers. Within the context of income distributions, the monotonicity axiom states that, all else the same, an increase in any individual's income increases the aggregate social welfare. The Pigou-Dalton principle of transfers requires that, all else the same, a transfer of income from a rich person to a poor person also increase social welfare. When applying the same set of axioms to characterizing generalized Lorenz dominance with a distribution of expected health values, the monotonicity axiom is pretty straightforward: if any socioeconomic class's expected health improves then the health situation of the society as a whole improves as well. It is the principle of transfers that may make people frown on: the principle makes a transfer of health from a healthier class to a less healthy class a welfare improvement. In measuring health opportunity, however, this axiom is not as unreasonable as it might appear when the monotone assumption is assumed. With the monotone assumption, a socioeconomic class with a better expected health will also have a higher income level; a transfer of expected health will not reverse the health rankings. Without the monotone assumption, the unconditional Pigou-Dalton principle in the context of health distribution is indeed problematic.

Both the principle of transfers and the monotonicity axiom can be interpreted in terms of the income-health matrix. Dardanoni (1993) proposed the notion of dynamic Pigou-Dalton transfer while Formby et al. (2003) defined the concept of simple increment on a transition matrix. Within the context of income-health matrix, a simple increment is performed within a given socioeconomic class by reducing the probability of being in a poor health status by δ and increasing the probability of being in a better health status by δ . That is,

$$(\alpha_{i1}, \dots, \alpha_{ik} - \delta, \dots, \alpha_{il} + \delta, \dots, \alpha_{im})$$

which improves the prospect (or opportunity) of health for economic class i (i.e., increasing the expected level of health). The monotonicity axiom that Formby et al. (2003) considered would require that such a simple increment increase the social

welfare of expected health. Dardanoni's (1993) dynamic Pigou-Dalton transfer occurs between two socioeconomic classes i and j ($i < j$) consisting of a simple increment in class i and a simple decrement in class j such that

$$(\alpha_{i1}, \dots, \alpha_{ik} - \delta, \dots, \alpha_{il} + \delta, \dots, \alpha_{im})$$

and

$$(\alpha_{j1}, \dots, \alpha_{jk} + \sigma, \dots, \alpha_{jl} - \sigma, \dots, \alpha_{jm})$$

with $\pi_i \delta = \pi_j \sigma$. After the transfer, the lower class (i) faces a better prospect of health while the higher class (j) lives with a worse-than-before prospect (but still better than class i). Stated in terms of expected health, the transfer shifts some expected health from class j to class i while maintaining the rank of expected health unaltered for all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$. The Pigou-Dalton principle of transfers states that such a transfer also improves the social welfare.¹⁵

As in the measurement of income distributions, generalized Lorenz dominance can be modified in two directions: it can be strengthened to a more robust condition or weakened to a condition with more ranking power. A condition that is more robust than generalized Lorenz dominance is rank dominance; rank dominance implies generalized Lorenz dominance but not vice versa. In the context of health opportunity measurement, rank dominance simply requires that the expected health level of *each* socioeconomic class under income-health matrix A be greater than or equal to that under B , i.e.,

$$\sum_{j=1}^m \alpha_{ij} h_j \geq \sum_{j=1}^m \beta_{ij} h_j \quad (3.6)$$

for all $i = 1, 2, \dots, n$ with the inequality holding strictly for some i . As for the social welfare functions involved, only those satisfying the monotonicity axiom are required. In this sense, rank dominance may have appeals to people who are uncomfortable with or are not completely convinced by the Pigou-Dalton principles of transfers.

A dominance condition that is more powerful (i.e., able to rank more pairs of distributions) than generalized Lorenz dominance is obtained by further requiring that a dynamic Pigou-Dalton transfer become more effective in enhancing social

¹⁵Bleichrodt and van Doorslaer (2006) recently examined the axiomatic foundations for the various health inequality measures. They concluded that, *inter alia*, the Pigou-Dalton principle of transfers may be problematic within the context of health inequality measurement. For an inequality index of health, for example, a transfer of health from a healthier person to a less healthy person cannot be justified if the healthier person is relatively poorer. These problems, however, do not appear in our characterizations because of the monotone matrix assumption that we have introduced and because the social welfare is defined over the expected health vector rather than the actual health levels. Under the monotone assumption, a representative person from a higher socioeconomic class will always have better expected health than a representative person from a lower class; and the transfer of health always occurs between two *expected* health levels (i.e., changing the health opportunity profiles). Such a transfer will leave all ranks unchanged – both the socioeconomic status and the expected level of health.

welfare when the transfer occurs in lower socioeconomic classes. In the literature of income distributions, such an additional requirement is referred to as the Kolm principle of diminishing transfers (Kolm, 1976) or the “transfer sensitivity axiom” (Shorrocks and Foster, 1987). Note that since the comparisons occur between two pairs of equal-distanced socioeconomic classes, the transfer sensitivity axiom invoked is the positional version characterized by Zoli (2001). The resulting dominance by incorporating the positional transfer sensitivity axiom is to compare the cumulative generalized Lorenz ordinates

$$\sum_{t=1}^l \sum_{i=1}^t \sum_{j=1}^m \pi_i \alpha_{ij} h_j \geq \sum_{t=1}^l \sum_{i=1}^t \sum_{j=1}^m \pi_i \beta_{ij} h_j. \quad (3.7)$$

for all $l = 1, 2, \dots, n$ with the inequality holding strictly for some l . For ease of reference, we refer to this dominance as *generalized super-Lorenz dominance*.

Similar to Proposition 3.1, we can prove the following proposition on these two additional dominance criteria.

Proposition 3.2. For all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$, and two vectors of expected health $\mathbf{e}(A)$ and $\mathbf{e}(B)$, (1) $\mathbf{e}(A)$ rank dominates $\mathbf{e}(B)$ if and only if

$$\sum_{j=1}^k \pi_i \alpha_{ij} \leq \sum_{j=1}^k \pi_i \beta_{ij} \quad (3.8)$$

for all $k = 1, 2, \dots, m$ and all $i = 1, 2, \dots, n$ with the inequality holding strictly for some pairs of (k, l) ; and (2) $\mathbf{e}(A)$ generalized super-Lorenz dominates $\mathbf{e}(B)$ if and only if

$$\sum_{t=1}^l \sum_{i=1}^t \sum_{j=1}^k \pi_i \alpha_{ij} \leq \sum_{t=1}^l \sum_{i=1}^t \sum_{j=1}^k \pi_i \beta_{ij} \quad (3.9)$$

for all $k = 1, 2, \dots, m$ and all $l = 1, 2, \dots, n$ with the inequality holding strictly for some pairs of (k, l) .

These seemingly complex conditions are in fact all related and connected to the income-health matrix through recursive cumulations. Given two income-health matrices A and B , these conditions are generated as follows. First we multiply each row of A and B by the respective population proportion π_i ; the resulting matrices become $|\pi_i \alpha_{ij}|_{n \times m}$ and $|\pi_i \beta_{ij}|_{n \times m}$, respectively. Then we cumulate elements of $|\pi_i \alpha_{ij}|$ and $|\pi_i \beta_{ij}|$ row-wise eastbound. The resulting matrices are $|\sum_{k=1}^j \pi_i \alpha_{ik}|$ and $|\sum_{k=1}^j \pi_i \beta_{ik}|$, respectively. Denote the matrices A_1 and B_1 . Condition (3.8) then requires $A_1 \leq B_1$ – all elements of A_1 are pair-wise no greater than B_1 . From A_1 and B_1 , we cumulate the matrices column-wise southbound, the resulting matrices are denoted A_2 and B_2 , respectively. Condition (3.3) requires $A_2 \leq B_2$. Cumulating A_2 and B_2 again column-wise southbound, we obtain A_3 and B_3 . Condition (3.9) requires $A_3 \leq B_3$. If we further cumulate A_3 and B_3 row-wise eastbound, we would

obtain another interesting condition similar to that established by Dardanoni (1993, Theorem 5).¹⁶ The alternating cumulations between eastbound and southbound can of course generate more sophisticated dominance conditions if one wishes. The more advanced dominance conditions would have a stronger flavor of the Rawlsian maximin principle: what happens at a lower health status and in a lower socioeconomic class would affect more on the outcomes of dominance comparisons.¹⁷

We now illustrate these dominance conditions with a simple numerical example. Assume a society has three socioeconomic classes and three health statuses. Also assume the population proportions in the three classes are the same. Consider two income-health matrices:

$$A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}.$$

The rank dominance condition (3.8) is not satisfied since (use notations from the previous paragraph) there is no dominance between A_1 and B_1 :

$$A_1 = \begin{bmatrix} 0.4 & 0.7 & 1 \\ 0.3 & 0.7 & 1 \\ 0.2 & 0.5 & 1 \end{bmatrix} \text{ and } B_1 = \begin{bmatrix} 0.5 & 0.8 & 1 \\ 0.3 & 0.8 & 1 \\ 0.2 & 0.4 & 1 \end{bmatrix}.$$

The generalized Lorenz dominance condition (3.3) holds since it can be checked that

¹⁶The difference between Dardanoni's condition and (3.9) lies in that his condition allows the transfer at lower-socioeconomic classes to be between lower health statuses than those in higher classes; our condition (3.9) requires the transfers to be between the same health statuses in both higher and lower classes. For example, Dardanoni's condition would rank

$$A' = \begin{bmatrix} \alpha_{11} - \delta & \alpha_{12} + \delta & \alpha_{13} & \alpha_{14} \\ \alpha_{21} + \delta & \alpha_{22} - \delta & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} + \delta & \alpha_{34} - \delta \\ \alpha_{41} & \alpha_{42} & \alpha_{43} - \delta & \alpha_{44} + \delta \end{bmatrix}$$

as welfare-improving over $A_{4 \times 4} = |\alpha_{ij}|$. Our condition (3.9) cannot rank A with A' ; we can only compare A with, for example,

$$A'' = \begin{bmatrix} \alpha_{11} - \delta & \alpha_{12} + \delta & \alpha_{13} & \alpha_{14} \\ \alpha_{21} + \delta & \alpha_{22} - \delta & \alpha_{23} & \alpha_{24} \\ \alpha_{31} + \delta & \alpha_{32} - \delta & \alpha_{33} & \alpha_{34} \\ \alpha_{41} - \delta & \alpha_{42} + \delta & \alpha_{43} & \alpha_{44} \end{bmatrix}$$

which is welfare-improving over A .

¹⁷Once the income-health matrix is obtained, these conditions can be easily computed and compared using spreadsheets such as what we have done in all illustrations and applications. In the measurement of income mobility, Formby et al. (2004) developed statistical inference to test mobility dominances, the same procedure can be modified and extended to test the dominance conditions in the measurement of health opportunity.

$A_2 \leq B_2$:

$$A_2 = \begin{bmatrix} 0.4 & 0.7 & 1 \\ 0.7 & 1.4 & 2 \\ 0.9 & 1.9 & 3 \end{bmatrix} \text{ and } B_2 = \begin{bmatrix} 0.5 & 0.8 & 1 \\ 0.8 & 1.6 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Given this, condition (3.9) – $A_3 \leq B_3$ – will also hold since it is implied by (3.3).

IV. The Measurement of Health Opportunity: Inequality

With the monotone-matrix assumption, standard Lorenz dominance criteria can be applied to rank vectors of expected health $\mathbf{e}(A)$ and $\mathbf{e}(B)$. Again, we consider two types of Lorenz dominances based on relative and absolute Lorenz curves. In the context of health opportunity, the relative Lorenz curve of expected health remains unchanged if all expected health levels across socioeconomic classes are increased/decreased proportionally. In contrast, the absolute Lorenz curve of expected health remains unchanged if all expected health levels across socioeconomic classes are increased/decreased by the same absolute amount.

We consider relative Lorenz dominance first. For an income-health matrix A , relative Lorenz ordinates $\{RL(A; \tilde{\pi}_l)\}$ are defined by

$$RL(A; \tilde{\pi}_l) = \frac{\sum_{i=1}^l \sum_{j=1}^m \pi_i \alpha_{ij} h_j}{\sum_{i=1}^n \sum_{j=1}^m \pi_i \alpha_{ij} h_j}.$$

For $0 < h_1 \leq h_2 \leq \dots \leq h_m$, $RL(A; \tilde{\pi}_l)$ can be rewritten using Abel's partial summation formula as

$$\begin{aligned} RL(A; \tilde{\pi}_l) &= \frac{\sum_{k=1}^m \left\{ \sum_{j=k}^m \sum_{i=1}^l \pi_i \alpha_{ij} \right\} (h_k - h_{k-1})}{\sum_{k=1}^m \left\{ \sum_{j=k}^m \sum_{i=1}^n \pi_i \alpha_{ij} \right\} (h_k - h_{k-1})} \\ &= \frac{\sum_{k=1}^m \alpha_l^k \tilde{h}_k}{\sum_{k=1}^m \alpha_n^k \tilde{h}_k}, \end{aligned} \tag{4.1}$$

where $\alpha_l^k = \sum_{j=k}^m \sum_{i=1}^l \pi_i \alpha_{ij}$ which is the sum of all elements in the upper-right $l \times (m - k + 1)$ submatrix of $|\pi_i \alpha_{ij}|$ and $\tilde{h}_k = h_k - h_{k-1}$.

Using (4.1) and notation α_l^k , we can derive the following result on relative Lorenz dominance.

Proposition 4.1. For all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$, the vector of expected health $\mathbf{e}(A)$ relative Lorenz dominates vector $\mathbf{e}(B)$ if

$$\alpha_l^s \beta_n^t + \alpha_l^t \beta_n^s \geq \alpha_n^t \beta_l^s + \alpha_n^s \beta_l^t \tag{4.2}$$

for all $s \leq t = 1, 2, \dots, m$ and all $l = 1, 2, \dots, n - 1$ with the inequality holding strictly for some pairs of (s, t, l) . If $m = 2$, the condition is also “only if.”

Proof. That vector $\mathbf{e}(A)$ relative Lorenz dominates vector $\mathbf{e}(B)$ means that for each $l = 1, 2, \dots, n-1$,

$$\frac{\sum_{k=1}^m \alpha_l^k \tilde{h}_k}{\sum_{k=1}^m \alpha_n^k \tilde{h}_k} \geq \frac{\sum_{k=1}^m \beta_l^k \tilde{h}_k}{\sum_{k=1}^m \beta_n^k \tilde{h}_k} \quad (4.3)$$

for all possible values $\tilde{h}_k \geq 0$. Expanding (4.3), we have for each $l = 1, 2, \dots, n-1$,

$$\sum_{s=1}^m \alpha_l^s \beta_n^s \tilde{h}_s^2 + \sum_{s < t} (\alpha_l^s \beta_n^t + \alpha_l^t \beta_n^s) \tilde{h}_s \tilde{h}_t \geq \sum_{s=1}^m \alpha_n^s \beta_l^s \tilde{h}_s^2 + \sum_{s < t} (\alpha_n^s \beta_l^t + \alpha_n^t \beta_l^s) \tilde{h}_s \tilde{h}_t. \quad (4.4)$$

Clearly, a sufficient condition for relative Lorenz dominance between $\mathbf{e}(A)$ and $\mathbf{e}(B)$ is (4.2) since $\tilde{h}_s \tilde{h}_t \geq 0$. When $m = 2$, the condition is also necessary since (4.4) then becomes

$$(\alpha_l^1 \beta_n^2 + \alpha_l^2 \beta_n^1) \tilde{h}_1 + \alpha_l^2 \beta_n^2 \tilde{h}_2 \geq (\alpha_n^1 \beta_l^2 + \alpha_n^2 \beta_l^1) \tilde{h}_1 + \alpha_n^2 \beta_l^2 \tilde{h}_2$$

for all values of $\tilde{h}_1 \geq 0$ and $\tilde{h}_2 \geq 0$. Here we have used the fact $\alpha_l^1 = \beta_l^1 = l$ for $l = 1, 2, \dots, m$. $\square$

Condition (4.2) does not appear to be necessary for Lorenz dominance when $m > 2$. A set of necessary conditions, however, can be derived by considering specific values of h_i (or $\tilde{h}_i$). For example, for each l , a necessary condition would be (by letting $\tilde{h}_i = \tilde{h}_j$ for all i and j)

$$\sum_{k=1}^m \alpha_l^k \sum_{k=1}^m \beta_n^k \geq \sum_{k=1}^m \alpha_n^k \sum_{k=1}^m \beta_l^k;$$

another necessary condition would be (by letting $\tilde{h}_k > 0$ and $\tilde{h}_i \rightarrow 0$ for all $i \neq k$)

$$\alpha_l^k \beta_n^k \geq \alpha_n^k \beta_l^k \text{ for all } k = 1, 2, \dots, m.$$

These necessary conditions are useful in the case where the sufficient condition fails, they can be used to rule out Lorenz dominance if any one of them is not satisfied.

The meaning of (4.2) can be best apprehended with $m = n = 2$. It is easy to verify that (4.2) then amounts to requiring

$$\begin{aligned} \alpha_{22} - \alpha_{12} &\leq \beta_{22} - \beta_{12}, \\ \text{and } \frac{\alpha_{22}}{\alpha_{12}} &\leq \frac{\beta_{22}}{\beta_{12}}. \end{aligned} \quad (4.5)$$

Since $\alpha_{22} \geq \alpha_{12}$ and $\beta_{22} \geq \beta_{12}$ by the monotone-matrix assumption, (4.5) states that for matrix A to be more equitable in health opportunity than B , the probabilities of attaining better health (status 2) between the two socioeconomic classes are closer in A than in B – both relatively and absolutely. Such an idea of convergence in probabilities remains in other more general conditions. For example, one condition for $m = n = 3$ is

$$\alpha_{22} + \alpha_{23} + \alpha_{32} + \alpha_{33} - 2(\alpha_{12} + \alpha_{13}) \leq \beta_{22} + \beta_{23} + \beta_{32} + \beta_{33} - 2(\beta_{12} + \beta_{13}),$$

which is similar to the first condition of (4.5). But, in general, some of the conditions become more difficult to interpret and here we will not attempt to provide a complete interpretation.

Now we turn to characterizing absolute Lorenz dominance. For an income-health matrix A , absolute Lorenz ordinates $\{AL(A; \tilde{\pi}_l)\}$ are defined by

$$AL(A; \tilde{\pi}_l) = \sum_{i=1}^l \sum_{j=1}^m \pi_i \alpha_{ij} h_j - \left\{ \sum_{i=1}^l \pi_i \right\} \sum_{i=1}^n \sum_{j=1}^m \pi_i \alpha_{ij} h_j.$$

For $0 < h_1 \leq h_2 \leq \dots \leq h_m$, $AL(A; \tilde{\pi}_l)$ can be rewritten using Abel's partial-sum formula

$$\begin{aligned} AL(A; \tilde{\pi}_l) &= \sum_{k=1}^m \sum_{j=k}^m \left\{ \sum_{i=1}^l \pi_i \alpha_{ij} - \tilde{\pi}_l \sum_{i=1}^n \pi_i \alpha_{ij} \right\} (h_k - h_{k-1}) \\ &= \sum_{k=1}^m (\alpha_l^k - \tilde{\pi}_l \alpha_n^k) \tilde{h}_k, \end{aligned} \quad (4.6)$$

where $\alpha_l^k = \sum_{j=k}^m \sum_{i=1}^l \pi_i \alpha_{ij}$, $\tilde{\pi}_l = \sum_{i=1}^l \pi_i$ and $\tilde{h}_k = h_k - h_{k-1}$.

Using (4.6) and notation α_l^k , we can state the following result on absolute Lorenz dominance.

Proposition 4.2. For all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$, the vector of expected health $\mathbf{e}(A)$ absolute Lorenz dominates vector $\mathbf{e}(B)$ if and only if

$$\tilde{\pi}_l \alpha_n^k - \alpha_l^k \leq \tilde{\pi}_l \beta_n^k - \beta_l^k \quad (4.7)$$

for all $k = 2, \dots, m$ and $l = 1, 2, \dots, n-1$ with the inequality holding strictly for some pairs of (k, l) .

Note that the condition for absolute Lorenz dominance is both necessary and sufficient. It is also much easier to interpret than the condition for relative Lorenz dominance. For $m = n = 2$ and $\pi_1 = \pi_2$, condition (4.7) simply states $\alpha_{22} - \alpha_{12} \leq \beta_{22} - \beta_{12}$ – the probability difference in attaining better health between the two socioeconomic classes is smaller in A than in B . The same interpretation holds for the general condition. Note that (4.7) can be equivalently expressed as

$$\tilde{\pi}_l (\alpha_n^k - \alpha_l^k) - (1 - \tilde{\pi}_l) \alpha_l^k \leq \tilde{\pi}_l (\beta_n^k - \beta_l^k) - (1 - \tilde{\pi}_l) \beta_l^k. \quad (4.7a)$$

Since $\tilde{\pi}_l (\alpha_n^k - \alpha_l^k)$ is the probability that a person comes from a socioeconomic class above l and has a health status at least k , and $(1 - \tilde{\pi}_l) \alpha_l^k$ is probability that a person comes from a socioeconomic class at or below l and has a health status at least k , it follows that (4.7a) states that the difference between the two probabilities is smaller in A than in B . It is useful to note that this condition requires only the absolute

difference to be smaller while relative Lorenz dominance (4.2) requires also the relative differences to be smaller.

The axiom behind both relative and absolute Lorenz dominances is the principle of dynamic Pigou-Dalton transfers; the principle of simple increment is not relevant for inequality measurement. If Lorenz dominances fail to rank a pair of income-health matrices, one can invoke super-Lorenz dominances by considering a positional transfer sensitivity axiom. Income-health matrix A relative super-Lorenz dominates B if

$$RSL(A; \tilde{\pi}_l) := \sum_{k=1}^l RL(A; \tilde{\pi}_k) \geq \sum_{k=1}^l RL(B; \tilde{\pi}_k) := RSL(B; \tilde{\pi}_l)$$

for $l = 1, 2, \dots, n$ with the inequality holding strictly for some l and $RL(\cdot; \tilde{\pi}_k)$ is defined in (4.1). Income-health matrix A absolute super-Lorenz dominates B if

$$ASL(A; \tilde{\pi}_l) := \sum_{k=1}^l AL(A; \tilde{\pi}_k) \geq \sum_{k=1}^l AL(B; \tilde{\pi}_k) := ASL(B; \tilde{\pi}_l)$$

for $l = 1, 2, \dots, n$ with the inequality holding strictly for some l and $AL(\cdot; \tilde{\pi}_k)$ is defined in (4.6).

Similar to the derivations of conditions (4.2) and (4.7), we can prove the following refined conditions on super-Lorenz dominances. In both conditions, $\tilde{\alpha}_l^k = \sum_{j=k}^m \sum_{v=1}^l \sum_{i=1}^v \pi_i \alpha_{ij}$, $\tilde{\beta}_l^k = \sum_{j=k}^m \sum_{v=1}^l \sum_{i=1}^v \pi_i \beta_{ij}$ and $\hat{\pi}_l = \sum_{i=1}^l \tilde{\pi}_i$.

Proposition 4.3. For all possible health values h_i s such that $0 < h_1 \leq h_2 \leq \dots \leq h_m$, and two vectors of expected health $\mathbf{e}(A)$ and $\mathbf{e}(B)$, (1) $\mathbf{e}(A)$ relative super-Lorenz dominates $\mathbf{e}(B)$ if

$$\tilde{\alpha}_l^s \beta_n^t + \tilde{\alpha}_l^t \beta_n^s \geq \alpha_n^t \tilde{\beta}_l^s + \alpha_n^s \tilde{\beta}_l^t \quad (4.8)$$

for all $s \leq t = 1, 2, \dots, m$ and all $l = 1, 2, \dots, n-1$ with the inequality holding strictly for some pairs of (s, t, l) . If $m = 2$, the condition is also “only if;” and (2) $\mathbf{e}(A)$ absolute super-Lorenz dominates $\mathbf{e}(B)$ if and only if

$$\hat{\pi}_l \alpha_n^k - \tilde{\alpha}_l^k \leq \hat{\pi}_l \beta_n^k - \tilde{\beta}_l^k \quad (4.9)$$

for all $k = 2, \dots, m$ and $l = 1, 2, \dots, n-1$ with the inequality holding strictly for some pairs of (k, l) .

We can illustrate the inequality dominance conditions with the same income-health matrices A and B employed in the previous section. For relative Lorenz dominance, a sufficient condition is $|\alpha_l^s \beta_m^t + \alpha_l^t \beta_m^s| \geq |\alpha_m^t \beta_l^s + \alpha_m^s \beta_l^t|$ for $l = 1, 2$. It is easy to compute that the matrices $|\alpha_l^s \beta_m^t + \alpha_l^t \beta_m^s|$ for $l = 1, 2$ are, respectively,

$$\begin{bmatrix} 6 & 3.8 & 1.9 \\ 3.8 & 2.4 & 1.2 \\ 1.9 & 1.2 & 0.6 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 12 & 7.9 & 3.8 \\ 7.9 & 5.2 & 2.5 \\ 3.8 & 2.5 & 1.2 \end{bmatrix}.$$

The matrices $|\alpha_m^t \beta_l^s + \alpha_m^s \beta_l^t|$ for $l = 1, 2$ are, respectively,

$$\begin{bmatrix} 6 & 3.6 & 1.7 \\ 3.6 & 2.1 & 1.0 \\ 1.7 & 1.0 & 0.4 \end{bmatrix} \text{ and } \begin{bmatrix} 12 & 7.8 & 3.4 \\ 7.8 & 5.0 & 2.2 \\ 3.4 & 2.2 & 0.9 \end{bmatrix}.$$

Clearly, $|\alpha_l^s \beta_m^t + \alpha_l^t \beta_m^s| \geq |\alpha_m^t \beta_l^s + \alpha_m^s \beta_l^t|$ holds for $l = 1, 2$. Thus, A is more equitable than B as judged by the relative Lorenz criterion.

Absolute Lorenz dominance relies on the comparisons between the two 2×2 matrices $|\tilde{\pi}_l \alpha_n^k - \alpha_l^k|$ and $|\tilde{\pi}_l \beta_n^k - \beta_l^k|$. It is also easy to verify that both matrices are, respectively, $1/3$ of the following matrices:

$$\begin{bmatrix} 0.3 & 0.2 \\ 0.3 & 0.4 \end{bmatrix} \text{ and } \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.8 \end{bmatrix}.$$

Since $|\tilde{\pi}_l \alpha_n^k - \alpha_l^k| \leq |\tilde{\pi}_l \beta_n^k - \beta_l^k|$, it follows that A is more equitable than B as judged by the absolute Lorenz criterion. Both conditions in Proposition 4.3 are also satisfied since they are implied by the two conditions in Propositions 4.1 and 4.2, respectively.

Finally, one may be interested in knowing the maximal and minimal elements for welfare dominances and inequality orderings of income-health matrices. For a given socioeconomic structure, what is the income-health matrix that yields the most social welfare? Is there a matrix that guarantees the most equitable distribution of expected health? For welfare rankings, the highest social welfare of expected health is achieved when people from all socioeconomic classes have the best health (h_m). This happens when the income-health matrix has ones in the last column and zeros everywhere else. On the other hand, the lowest social welfare of expected health occurs when the income-health matrix has ones in the first column and zeros everywhere else – all socioeconomic classes have the worst health (h_1). For both relative and absolute Lorenz dominances, the most equitable distribution of expected health, as alluded before, occurs when all expected levels of health are the same. Translating into the income-health matrix, this is the case when all rows of the matrix are identical. But since different income-health matrices may lead to different health distributions $\mathbf{p}$, the most equitable income-health matrix may not be unique. In fact there are multiple such most-equitable matrices – as long as all rows of the matrix equal $(\frac{p_1}{n}, \frac{p_2}{n}, \dots, \frac{p_m}{n})$ for all possible $\mathbf{p}$ such that $\sum_{i=1}^m p_i = 1$. The least equal distribution of expected health happens when the highest socioeconomic class has the best health status while all others have the worst health status. The corresponding income-health matrix has ones in its first column except the last entry, one in the southeast-corner entry and zeros in the rest of the matrix. It is easy to verify that these characterizations of maximums and minimums are consistent with all the dominance conditions derived above.

V. Health Opportunities in the US and Canada

Much has been written about the different health-care systems adopted in the United States and Canada; the distributions of health status in both countries have been repeatedly measured with standard measures of inequality (e.g., van Doorslaer et al, 1997; Kunst et al., 1995). Little, however, has been said about the different health opportunities presented in the two countries. As we have argued, the distribution of health opportunity is a much more fundamental issue to a society than the distribution of health status. We need to know not just the number of people in each health category but also, and more importantly, what health opportunity a society can offer and how the opportunity varies across socioeconomic classes.

In this illustrative application, we apply the dominance conditions developed in the paper to evaluate and compare health opportunities in the US and Canada using the newly released Joint Canada/US Survey of Health (JCUSH) data. The JCUSH was designed specifically to enable accurate comparisons between the two countries in health status and health opportunity. The survey was a collaboration between Statistics Canada and the National Center for Health Statistics (NCHS) of the US Centers for Disease Control and Prevention (CDC). The JCUSH was administered between November 2002 and March 2003 from Statistics Canada’s regional offices. The sampled population includes noninstitutionalized residents in both countries age eighteen or older. The sample consists of 3,505 Canadians and 5,183 US residents, and the sample weights for the two countries are derived, respectively, from the 1996 Canada Census and the October 2002 US Current Populations Survey (CPS).

The JCUSH response rates were relatively low (66% and 50% in Canada and the US, respectively) compared with other health surveys such as the US National Health Interview Survey (NHIS) and the Canadian National Population Health Survey (NPHS). The missing responses including “don’t know”, “not stated”, and “refusal” were also high due to the survey method used. Despite these limitations, since the survey uses a common questionnaire and identical data collection and processing methods, the JCUSH is still probably the most suitable data for the purpose of comparing health status and health opportunity in the US and Canada.

The first step in our investigation of health opportunity in the two countries is to construct the respective income-health matrix. For our purpose, we drop all observations with missing income information or health status. The JCUSH collects two types of aggregate incomes – household total income and personal total income. All persons in the sample are assigned into an income quintile according to the adjusted household equivalent income (total household income divided by the square root of household size) and personal income. That is, an individual may belong to two different quintiles, depending upon which income is considered. We considered both types of incomes in forming socioeconomic classes and found that the qualitative results are almost identical. Here, we follow the approach used in Sanmartin et al. (2004) and report only the results using the adjusted household income; the results using personal income are available upon request. Since we use income quintiles as socioeconomic

classes, the socioeconomic structure for both countries is $\boldsymbol{\pi} = (0.2, 0.2, 0.2, 0.2, 0.2)$. Since there are five health statuses (poor, fair, good, very good, and excellent), the five health classes are naturally established. We rank income classes from the lowest to the highest and health statuses from poor to excellent. The income-health matrices are computed directly from the JCUSH and are reported below (A is the matrix for the US and C the matrix for Canada):

$$A = \begin{bmatrix} 11.23 & 19.78 & 31.81 & 22.37 & 14.81 \\ 3.00 & 10.26 & 32.04 & 34.29 & 20.40 \\ 2.31 & 7.19 & 25.58 & 38.37 & 26.55 \\ 0.52 & 4.15 & 22.28 & 40.76 & 32.30 \\ \mathbf{1.44} & 4.32 & 17.70 & 36.40 & 40.14 \end{bmatrix},$$

and

$$C = \begin{bmatrix} 8.77 & 18.17 & 33.8 & 25.54 & 13.72 \\ 4.69 & 8.81 & 31.39 & 35.94 & 19.18 \\ 1.83 & 5.91 & 26.48 & 39.31 & 26.48 \\ 1.14 & 5.33 & 22.48 & 45.52 & 25.52 \\ 0.65 & 3.24 & 17.71 & 41.25 & 37.15 \end{bmatrix}.$$

First we check whether the assumption of monotone-matrix is satisfied. It is satisfied by the Canadian matrix (C) with the exception of the last two elements in row 4 but the violation is well within the range of statistical errors. The US matrix (A) satisfies the monotone assumption except the last row of the matrix. The violation appears to be significant statistically. But the number 1.44 seems to be unusually large for the richest quintile; it means that a person from the richest class is almost three times as likely to have poor health as a person from the class below. Sanmartin et al. (2004, Table A-7) cautioned about the interpretation of the numbers of the highest quintile at the lowest two health statuses in both countries. We do not view this unusually large number as an evidence against the monotone-matrix assumption.¹⁸ In fact, in our other investigations using both the NHIS and NPHS data, the monotone-matrix assumption always holds.

With this acknowledgment, we proceed with our analysis. Since $p_j = \sum_{i=1}^n \pi_i \alpha_{ij}$, we can obtain the health distributions $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5)$ for the US and Canada as follows.

	<i>Health Distribution</i>
US	(3.70, 9.14, 25.88, 34.44, 26.84)
Canada	(3.42, 8.29, 26.37, 37.51, 24.41)

Applying generalized Lorenz dominance (or the first-order condition (2.2a)) to these distributions, it is easy to verify by using condition (2.2a) that neither health distribution dominates the other. A closer examination of the dominance comparison reveals

¹⁸In a private communication, Claudia Sanmartin of Statistics Canada agreed that the abnormality is likely caused by the small sample sizes at the high end of income distributions and the lowest health status.

that the Canadian distribution dominates that of the US up to the third health status (good) and then reverses the direction of dominance. This implies that the Canadian sub-distribution of the three lowest health statuses welfare dominates that of the US while the US welfare dominates Canada in the top two levels of health distribution (poor health people are better off in Canada while good health people are better off in the US).

Now consider welfare dominance of health opportunity between the two income-health matrices. Corollary 3.1 suggests that, if no generalized Lorenz dominance between the two health distributions, then there would also be no generalized Lorenz dominance between the two distributions of health opportunity. This is indeed the case as the difference between $A_2 = |\sum_{i=1}^l \sum_{j=1}^k \pi_i \alpha_{ij}|$ and $C_2 = |\sum_{i=1}^l \sum_{j=1}^k \pi_i \gamma_{ij}|$ is

$$C_2 - A_2 = \begin{bmatrix} -2.46 & -4.07 & -2.08 & 1.09 & 0 \\ -0.77 & -3.83 & -2.49 & 2.33 & 0 \\ -1.25 & -5.59 & -3.35 & 2.41 & 0 \\ -0.63 & -3.79 & -1.35 & 9.17 & 0 \\ -1.42 & -5.66 & -3.21 & 12.16 & 0 \end{bmatrix},$$

which yields no dominance. This no-dominance situation remains unchanged with generalized super-Lorenz dominance (3.9). It is however interesting to see that a clear dominance emerges when a condition similar to the one that Dardanoni (1993, Theorem 5) derived is applied. Recalling that this condition compares the matrices that are derived from A_3 and C_3 – the two matrices corresponding to generalized super-Lorenz dominance – by cumulating A_3 and C_3 row-wise eastbound. Denote the resulting matrices A_4 and C_4 , then

$$C_4 - A_4 = \begin{bmatrix} -2.46 & -6.53 & -8.61 & -7.52 & -7.52 \\ -3.23 & -11.1 & -15.7 & -12.3 & -12.3 \\ -4.48 & -17.9 & -25.9 & -20.1 & -20.1 \\ -5.11 & -22.4 & -31.7 & -16.7 & -16.7 \\ -6.53 & -29.5 & -41.9 & -14.8 & -14.8 \end{bmatrix}.$$

Thus, according to the Dardanoni-type dominance condition, the Canadian matrix of health opportunity welfare dominates the US matrix. But this dominance conclusion is hinged upon an extra condition that the health value h_i is concave in i – the difference between two adjacent health values, $h_i - h_{i-1}$, decreases as i increases. Put in words, this concavity condition assumes that the difference between the “excellent” health and the “very good” health is smaller than the difference between the “very good” health and a mere “good” health, and so forth. It is an open question whether such an assumption can be justified. Without it, however, we are forced to draw the conclusion that there is no welfare ranking between the Canadian and the US health opportunities.¹⁹

¹⁹It is interesting to note that even with this assumption, there is still no comparison between the two health distributions.

Can we tell which country provides more equal health opportunity? Fortunately, we can – with a higher-order condition. We first consider Lorenz dominance by applying conditions (4.2) and (4.7), respectively, to the two matrices. There is no relative Lorenz dominance or absolute Lorenz dominance between the two matrices.²⁰ For example, for absolute Lorenz dominance, the difference matrix between the US absolute Lorenz matrix $|\tilde{\pi}_l \alpha_n^k - \alpha_l^k|$ and the Canadian matrix $|\tilde{\pi}_l \gamma_n^k - \gamma_l^k|$ is (when the former is subtracted from the latter)

$$\begin{bmatrix} -2.17 & -2.94 & -1.44 & -1.34 \\ -0.22 & -1.58 & -1.22 & -2.55 \\ -0.42 & -2.22 & -1.45 & -4.91 \\ 0.50 & 0.74 & 1.22 & -0.56 \end{bmatrix}.$$

When the super-Lorenz dominance criteria (4.8) and (4.9) are considered, however, the Canadian matrix dominates the US matrix – both relatively and absolutely. For example, the difference matrix between the US absolute super-Lorenz matrix and the Canadian matrix is (when the former is subtracted from the latter)

$$\begin{bmatrix} -2.17 & -2.94 & -1.44 & -1.34 \\ -2.39 & -4.52 & -2.66 & -3.89 \\ -2.81 & -6.74 & -4.11 & -8.80 \\ -2.31 & -6.00 & -2.89 & -9.36 \end{bmatrix}.$$

Therefore, we can conclude that Canada exhibits more equal health opportunity than the US. This conclusion is based upon the super-Lorenz dominance criterion which allows greater weights to be given to inequality occurring in the lower part of the socioeconomic structure.

The findings in our brief study are consistent with the general impressions of health opportunities in the two countries as well as studies based upon other health survey data. Our results, however, are derived from a clearly defined notion of health opportunity and based upon a set of axiomatically characterized dominance conditions. Since the main purpose of our study is illustrative, caution must be taken in interpreting the findings. A more thorough study along the direction and use different health data will surely shed more lights on the US and Canadian health inequality and health opportunity, but such a mission is clearly beyond the scope of the present paper.

²⁰If we replace the last row of the US matrix with the corresponding row estimated from the 2002 NHIS data, then both relative Lorenz dominance and absolute Lorenz dominance prevail: Canada dominates the US in equality of health opportunity. Such a change, however, does not seem to affect the welfare comparisons.

VI. Summary and Conclusion

Disparity in health among the citizens is a natural cause for concern in an equitable society. Recently, however, it has been argued (e.g., Brommier and Stecklov, 2002) that disparity in accessing health services rather than inequality in health itself should be more a cause for concern.²¹ The main objective of the paper has been to measure health inequality caused by unequal access to health services and other factors related to uneven socioeconomic structure. Specifically, we attempted to characterize a situation where socioeconomic factors are singled out for comparisons and established dominance conditions to facilitate the comparisons. To that end, we introduced the income-health matrix which documents the stochastic relation between the socioeconomic structure and health distribution of a society. Each row of the matrix is characterized as the health opportunity profile for the corresponding socioeconomic class and the entire matrix represents the health opportunity for the society.

Since an income-health matrix resembles the transition matrix used in income mobility measurement, approaches developed there can be adapted to measuring health opportunity. The particular approach that we followed is the one pioneered by Atkinson (1983) and further developed by Dardanoni (1993). Following Dardanoni (1993), we also considered a monotone-matrix assumption that stipulates “richer people tend to be healthier” within the context of health opportunity. With this assumption, we were able to develop a set of welfare dominance conditions as well as a set of inequality ordering conditions. This development is in sharp contrast with the measurement of health distributions where we obtained only a limited possibility for welfare dominance and an impossibility for inequality dominance. We then applied our results to compare health opportunity in the US and Canada. Using the newly released JCUSH data, we concluded that there is no definite verdict on welfare comparisons along the line of generalized Lorenz dominance unless we further invoke a condition of “diminishing marginal value of health.” In that case, Canada welfare dominates the US in health opportunity. As for inequality in health opportunity, Canada was shown to be more equal than the US in the sense of super-Lorenz dominance which puts more weights to inequalities occurring in lower socioeconomic classes than the familiar Lorenz dominance.

All of the conditions derived so far in this paper for measuring health opportunities are dominance conditions. In many empirical investigations, however, summary indices are often desired. For all dominance conditions that we have developed, with the exception of relative (super-) Lorenz dominance, we can define an index measure based upon the dominance condition.²² Consider the generalized Lorenz dominance condition (3.3). Denote $\check{H}$ the matrix with ones in the last column and zeros elsewhere, and $\hat{H}$ the matrix with ones in the first column and zeros elsewhere. Note

²¹See also Daniels (1985) and Daniels et al. (2000).

²²The sufficient condition for relative Lorenz dominance involves both distributions of interest in each side of condition (4.2). The nonseparable nature of the two distributions prevents the construction of a summary measure in the same way as in other cases.

that $\check{H}$ yields the highest welfare while $\hat{H}$ produces the lowest welfare. If we further denote $S(H)$ as the sum of all elements of the generalized-Lorenz-matrix associated with matrix H , then a summary measure of welfare for any income-health matrix H is

$$w(H) = \frac{S(\hat{H}) - S(H)}{S(\hat{H}) - S(\check{H})} = \frac{m \sum_{i=1}^n (n-i+1)\pi_i - S(H)}{(m-1) \sum_{i=1}^n (n-i+1)\pi_i}.$$

It is easy to see that $w(\cdot)$ lies between zero and one and satisfies both the principle of simple increment and the principle of dynamic Pigou-Dalton transfers. For the US income-health matrix from the JCUSH data, $w(A) = 0.639$; for the Canadian matrix $w(C) = 0.642$. Thus, according to this summary index, Canada exhibits higher welfare than the US in health even though the generalized Lorenz criterion indicates a non-comparison.

In the same manner, we can define a summary inequality index of health opportunity. Denote $\check{\check{H}}$ the matrix with all rows equal, and $\check{H}$ the matrix with ones in the first column except the last one and zeros elsewhere except the very southeast-corner element which is one. Also note that $\check{\check{H}}$ leads to a perfect equality of expected health while $\check{H}$ leads to the most unequal distribution of expected health. If we define, with a slight abuse of notation, $S(H)$ as the sum of all elements of the absolute-Lorenz-matrix associated with matrix H , then a summary absolute measure of inequality in health opportunity can be defined as

$$I(H) = \frac{S(H)}{S(\check{H}) - S(\check{\check{H}})} = \frac{S(H)}{(m-1)\pi_n \sum_{i=1}^{n-1} (n-i)\pi_i}.$$

where we have used the fact $S(\check{\check{H}}) = 0$. Clearly, $I(\cdot)$ also lies between zero and one and satisfies the principle of dynamic Pigou-Dalton transfers. For the US income-health matrix, $I(A) = 0.30$; for the Canadian matrix $I(C) = 0.28$. Thus, according to this summary index, Canada provides more equal health opportunity than the US even though the absolute Lorenz criterion indicates a non-comparison.²³

Both summary inequality measures of health opportunity as well as the dominance conditions derived in the body of the paper do not depend on the cardinal valuation of health statuses. It follows that these measures and conditions provide the most robust assessments of health opportunities for empirical applications. The applicability of these measures and conditions, however, hinges upon the satisfaction of the monotone-matrix assumption. Although the notion implied by the monotone assumption, i.e., the social gradient in health, is well documented and strongly supported in the literatures of public health and epidemiology, the assumption may

²³It is useful to note that, in the case where socioeconomic classes are exogenously determined (e.g., when they are defined through absolute income levels rather than income percentiles), both summary measures are decomposable. The overall inequality of health opportunity is the weighted sum of the inequality levels of population subgroups with the weights being the subgroup population shares. If socioeconomic classes are formed endogenously, the decomposition is more complicated. A general exposition of decomposition in mobility measurement is the subject of Zheng (2006b).

be less likely fulfilled as the number of socioeconomic classes and/or the number of health categories increase. In fact, there are reasons to believe that even the well-established gradient patterns in the epidemiology literature may become fuzzier when ten or twenty socioeconomic classes are considered.²⁴ The question then for empirical applications is how many socioeconomic classes and health categories should one consider? In theory, the socioeconomic classes should be broadly defined so that they are fundamentally different from each other in socioeconomic status. For example, the traditional definition of three social classes (poor, middle and rich) is very telling about the differences among them. Imagine, however, if a society is divided into ten or more socioeconomic classes, then the differences between these classes would be much less distinct. Consequently, the inequality in health opportunities that we try to measure would reflect mostly the variations in individual health states rather than the differences between social groups. In this sense, we should not partition the society into too many socioeconomic classes. As a matter of fact, browsing through the literature of social inequality in health, it seems that most studies use five or fewer socioeconomic classes. With these understandings, we feel that the fulfillment of the monotone-matrix assumption may not be a serious concern for empirical applications of the methods advanced in the paper.

On the other hand, the monotone matrix assumption itself *can* be somewhat relaxed if we are willing to impose further conditions on health values h_i s. In the paper, the assumption we made is essentially a first-order condition, i.e., $\sum_{j=1}^k \alpha_{sj} \geq \sum_{j=1}^k \alpha_{tj}$ for all $s < t$ and $k = 1, 2, \dots, m$. The requirement on h_i s was $h_1 \leq h_2 \leq \dots \leq h_m$. If we further require that h_i s be also concave in i , then we can use a second-order monotone condition: $\sum_{l=1}^k \sum_{j=1}^l \alpha_{sj} \geq \sum_{l=1}^k \sum_{j=1}^l \alpha_{tj}$ for all $s < t$ and $k = 1, 2, \dots, m$. Clearly, the second-order condition is less demanding than the first-order condition and, hence, is more likely to be satisfied. Another potentially fruitful line of extension – though it will not be pursued here – is to specify a range of values for each health status. In the paper, each h_i can take any value between zero and infinity. If we can somehow narrow this range and impose further restrictions on how h_i and h_j are related, then the resulting monotone assumption will be more likely to be satisfied and the ranking conditions of health opportunities will be able to rank more pairs of income-health matrices. A similar idea has been successfully applied to characterizing the *sequential dominance condition* in inequality and poverty rankings with heterogeneous populations (e.g., Atkinson and Bourguignon, 1987; Atkinson, 1992; Jenkins and Lambert, 1993). We certainly hope that a similar condition can be obtained for measuring and ranking health opportunities.

²⁴In the limit, at the individual level (each person is defined as a class), there will not be no gradient but a random variation in health.

References

- Allison, R. A. and J. E. Foster (2004) Measuring health inequality using qualitative data, *Journal of Health Economics* 23, 505-524.
- Anand, S. (2004): The concern for equity in health, in: S. Anand, F. Peter, and A. Sen, eds., *Public Health, Ethics and Equity* (Oxford University Press, New York).
- Atkinson, A. B. (1983): The measurement of economic mobility, in: *Social Justice and Public Policy* (MIT Press, Cambridge, MA).
- Atkinson, A. B. (1992): Measuring poverty and differences in family composition, *Economica* 59, 1-16.
- Atkinson, A. B. and F. Bourguignon (1987): Income distribution and differences in needs, in: G. R. Feiwel, ed., *Arrow and the Foundations of the Theory of Economic Policy* (Macmillan, London).
- Bleichrodt, H. and E. van Doorslaer (2006): A welfare economics foundation for health inequality measurement, *Journal of Health Economics* 25, 945-957.
- Brommier, A. and G. Stecklov (2002): Defining health inequality: why Rawls succeeds where social welfare theory fails, *Journal of Health Economics* 21, 497-513.
- Daniels, N. (1985): *Justice Health Care* (Cambridge University Press, New York).
- Daniels, N., B. Kennedy and I. Kawachi (2000): Justice is good for our health, in: J. Cohen and J. Rogers, eds., *Is Inequality Bad for Our Health* (Beacon Press, Boston, MA).
- Dardanoni, V. (1993): On measuring social mobility, *Journal of Economic Theory* 61, 372-394.
- Erreygers, G. (2006): Beyond the health concentration index: an Atkinson alternative for the measurement of the socioeconomic inequality of health, paper presented in the UNU-WIDER conference on *Advancing Health Equity*, Helsinki, Finland.
- Fleurbaey, M. (2005): Health, equity and fairness, mimeo, Université de Pau.
- Formby, J. P., W. J. Smith and B. Zheng (2003): Economic growth, welfare and the measurement of social mobility, in: Y. Amiel and J. Bishop, eds., *Research in Economic Inequality*, Vol 9 (JAI Press, Amsterdam).
- Formby, J. P., W. J. Smith and B. Zheng (2004): Mobility measurement, transition matrices and statistical inference, *Journal of Econometrics* 120, 181-205.
- Jenkins, S. and P. Lambert (1993): Ranking distributions when needs differ, *Review of Income and Wealth* 39, 337-356.
- Kolm, S. (1976): Unequal Inequalities II, *Journal of Economic Theory* 13, 82-111.
- Kunst, A. , J. Geurts and J. van den Berg (1995): International variation in socioeconomic inequalities in self-reported health, *Journal of Epidemiology and Community Health* 49, 117-123.
- Lambert, P. (2001): *The Distribution and Redistribution of Income*, 3rd edition (The Manchester University Press, Manchester and New York).

- Levy, H. (2006): *Stochastic Dominance* (Springer-Verlag, Berlin).
- Lindeboom, M. and E. van Doorslaer (2004): Cut-point shift and index shift in self-reported health, *Journal of Health Economics* 23, 1083-1099.
- Marmot, M. (2003): Understanding social inequalities in health, *Perspectives in Biology and Medicine* 46 (supplement), S9-S23.
- Marmot, M. (2004): Social causes of social inequalities in health, in: S. Anand, F. Peter, and A. Sen, eds., *Public Health, Ethics and Equity* (Oxford University Press, New York).
- Marmot, M., et al. (1978): Employment grade and coronary heart disease in British civil servants, *Journal of Epidemiology and Community Health* 32, 244-249.
- Marmot, M., et al. (1991): Health inequalities among British civil servants: the Whitehall II study, *Lancet* 337, 1387-1393.
- Marmot, M. and M. Shipley (1996): Do socio-economic differences in mortality persist after retirement? 25 year follow up of civil servants from the first Whitehall study, *British Medical Journal* 313, 1177-1180.
- Organization for Economic Cooperation and Development (2005): *Health at a Glance: OECD Indicators 2005* (OECD, Paris).
- Peter, F. (2004): Health equity and social justice, in: S. Anand, F. Peter, and A. Sen, eds., *Public Health, Ethics and Equity* (Oxford University Press, New York).
- Pritchett, L. and L. Summers (1996): Wealthier is healthier, *Journal of Human Resources* 31, 841-868.
- Rowntree, B. (1901): *Poverty: A Study of Town Life* (Macmillan, London and New York).
- Rudin, W. (1976): *Principles of Mathematical Analysis*, 3rd edition (McGraw-Hill, New York).
- Sanmartin, C. and others (2004): *Joint Canadian/United States survey of health, 2002-03*, Statistics Canada.
- Sen, A. (1985): *Commodities and Capabilities* (North-Holland, Amsterdam).
- Shorrocks, A. and J. Foster (1987): Transfer sensitive inequality measures, *Review of Economic Studies* 54, 485-497.
- Smith, J. P. (1999): Healthy bodies and thick wallets: the dual relation between health and economic status, *Journal of Economic Perspectives* 13, 145-166.
- van Doorslaer, E., A. Wagstaff and others (1997): Income-related inequalities in health: some international comparisons, *Journal of Health Economics* 16, 93-112.
- van Doorslaer, E. and A. Jones (2003): Inequalities in self-reported health: validation of a new approach to measurement, *Journal of Health Economics* 22, 61-67.
- Wagstaff, A., P. Paci and E. van Doorslaer (1991): On the measurement of inequalities in health, *Social Science and Medicine* 33, 545-557.

Wagstaff, A. and E. van Doorslaer (1994): Measuring inequalities in health in the presence of multiple-category morbidity indicators, *Health Economics* 3, 281-291.

Wagstaff, A. and E. van Doorslaer (2000): Equity in health care finance and delivery, in: Culyer, A. J. and J. P. Newhouse, eds., *Handbook of Health Economics* (Elsevier, Amsterdam).

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.

Zheng, B. (2006a): Unit consistency and inequality orderings, mimeo, Department of Economics, University of Colorado at Denver and HSC.

Zheng, B. (2006b): Transition matrix based decomposable mobility measures, work in progress.

Zoli, C. (2001): Inverse stochastic dominance, inequality measurement and Gini indices, *Journal of Economics*, Supplement No.9, 119-161.