Journal of Health Economics 34 (2014) 84–95

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Measuring socioeconomic health inequalities in presence of multiple categorical information☆

Paul Makdissi[a,*], Myra Yazbeck[b]

[a] Department of Economics, University of Ottawa, 120 University (9008), Ottawa, Ontario, Canada, K1N 6N5
[b] School of Economics, 612- Colin Clark (39) Building, University of Queensland, Brisbane, QLD 4072, Australia

ARTICLE INFO

Article history:
Received 16 July 2012
Received in revised form
26 November 2013
Accepted 26 November 2013
Available online 8 December 2013

JEL classification:
D63
I10

Keywords:
Population health
Socioeconomic health inequality
Health achievement
Categorical variables
Stochastic dominance

ABSTRACT

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

1. Introduction

Measuring health inequalities is essential for the implementation and the monitoring of health policies. It is well known that a large body of the health inequality measurement literature is based on the accumulated knowledge in income inequality measurement where the 0 has a well defined meaning. In population surveys, most of the available information on health status is given in the form of categorical variables (i.e., for which 0 does not have a well defined meaning). As a result, the well-known income inequality indices cannot always be readily applied to measure inequalities in health status. To overcome this problem, this paper exploits the multi-dimensionality of the information available in health surveys and offers to researchers who wish to use categorical variables a new approach to socioeconomic health inequality.

Most of the current literature on health inequality is using the concentration index as an accepted measure of health inequality. However, the use of the concentration index presents three important measurement problems. The first results from the fact that the concentration indices do not account for the average level of health in the population considered (Wagstaff, 2002). Thus, a policy that improves the average level of health, while keeping the relative distribution of health constant, will be deemed neutral when using the concentration index. To overcome this problem, Wagstaff (2002) proposes the use of an achievement index that captures simultaneously the average level of health status and the socioeconomic inequality of its distribution. The second measurement problem is the well known mirror problem pointed out by Clarke et al. (2002). It results in the absence of consistency between the rankings of health attainments and health shortfalls when the concentration index is used. Erreygers (2009a,b) suggested a corrected version of the concentration index that accounts for this inconsistency and highlights that this index is not an index of relative inequality.[1]

☆ We thank Serge Coulombe, Jean-Yves Duclos, Bernard Fortin, Lilia Karnizova, Guy Lacroix, Aggey Semenov, Nigel Rice and two anonymous referees for useful comments.

* Corresponding author. Tel.: +1 613 562 5800x4897; fax: +1 613 562 5999.
E-mail addresses: paul.makdissi@uottawa.ca (P. Makdissi), m.yazbeck@uq.edu.au (M. Yazbeck).

[1] Subsequently, Lambert and Zheng (2011) show that no index of relative inequality can really avoid this problem.

0167-6296/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhealeco.2013.11.008

The last measurement problem, is the *arbitrariness of the concentration index* (for details see Erreygers, 2006 and Zheng, 2008). It results from a common misuse of non-ratio-scale variables while computing inequalities indices that are developed for ratio-scale variables. One can group health variables according to their measurement scales as follows: nominal, ordinal, cardinal, ratio-scale, and fixed (Erreygers and Van Ourti, 2011). While the use of nominal health variables (e.g., type of illness) allows the researcher to classify individuals in different (illness) groups, it does not allow her to rank them. With ordinal (or categorical) health variables such as self assessed health status, one can rank individuals but the interpretation of the differences between individuals' health variables is meaningless. It is important to emphasize that while self reported health status is the most commonly used health variable in the measurement of health inequality, it fails to provide a meaningful cardinal difference when used to compare individuals' health status (i.e. compute inequality indices).[2] For cardinal variables, the differences between individuals are meaningful, however, since the zero point is arbitrary, their ratio is meaningless. Body temperature is a good example of this group of health variables. Ratio-scale variables (e.g., such as body length and life expectancy) are cardinal variables for which the zero correspond to a situation of complete absence. Such measurement scales are unique to a proportional scaling factor. Finally, a variable is fixed if the measurement scale is unique. As pointed out by Erreygers and Van Ourti (2011), when health variables are fixed or ratio scaled, the concentration index can be readily used. However, this is not the case with cardinal health variables. Wagstaff (2005) and Erreygers (2009a) offer alternative measurement methods that accounts for the problem encountered when one uses the concentration index with cardinal health variables, yet, their method cannot be applied to ordinal or nominal variables.[3]

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

The objective of this paper is to address this third measurement problem by following the first path. We transform the available information (i.e. exploit its multi-dimensionality) so that it becomes ratio scaled, then redefine the health status variable so that it (1) can readily be used to compute inequality indices (2) captures the information on socioeconomic inequality in health attainments. Our objective can be related to that of Allison and Foster (2004) and Zheng (2011) in its intent to overcome the *arbitrariness of the health inequality measure*, however it differs from their work with respect to the dimensions considered (i.e. health dimensions and socioeconomic dimension). Allison and Foster (2004) offer a solution in the dimension of pure health inequality. They propose a stochastic dominance approach to identify robust rankings of health distributions. Using Allison and Foster's approach, one can account for the depth of health status, however, this comes at the cost of overlooking the socioeconomic dimension of health inequality. Zheng (2011) offers a solution by grouping individuals into socioeconomic ranks and imposing monotonicity of health in socioeconomic ranks. In doing so, he adapts the income mobility matrix to the context of socioeconomic health inequality in which he sorts individuals into increasingly ranked socioeconomic classes.[4] He then assumes that the health distribution of individuals from higher socioeconomic rank, stochastically dominates the health distribution of individuals in lower socioeconomic rank.[5] In such a context the health status variable does not need a cardinal interpretation. Thus, using Zheng's approach one can use non-ratio scaled variables when analysing socioeconomic health inequalities. Nevertheless, this comes at the cost of (1) overlooking heterogeneity within socioeconomic classes and (2) potentially imposing Pigou-Dalton principle when it is not justified. Indeed, the literature has not yet offered a complete solution for the *arbitrariness of the concentration index* in the dimension of socioeconomic health inequality. In this perspective, this paper contributes to this literature by proposing a new method for the measurement of socioeconomic health inequalities when the categorical information covers multiple dimensions of the health status. Our approach finds its inspiration in the work of Alkire and Foster (2011) on the *width* of poverty approach, yet differs from it in three respects. First, it limits the counting procedure at the individual level, while Alkire and Foster's counting procedure is used at the individual and population level. Second, it uses a different aggregation method. Whereas Alkire and Foster use counting aggregation procedure, this approach uses the information produced by the counting approach at the individual level and transfers it in a rank dependent social decision function. Third, in addition to proposing a different structure for the indices, it provides the associated stochastic dominance criteria.

The remaining of this paper is organized as follows. The next section presents the theoretical measurement framework. Section 3 presents the rank dependent ethical principles and develops two positional stochastic dominance conditions. Section 4 presents a brief empirical illustration using the Joint Canada/United States Surveys of Health 2004 and the National Health Interview Survey 2010. The last section summarizes our results.

2. Theoretical framework

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

[2] The same problem remains present when using regression analysis to impute cardinal health measure when the numerical scale associated with the ordinal self assessed health status in the regression is not unique.

[3] The relative merits of both solutions have been discussed in details in Wagstaff (2009) and Erreygers (2009b).

[4] Each element of the income-health matrix represents the probability that an individual in a socioeconomic class will have a particular health status.

[5] The underlying assumption is that people from low socioeconomic rank and high socioeconomic rank are biologically similar.