View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Measurement and explanation of socioeconomic inequality in health with longitudinal data

Andrew M. Jones*[a] and Angel López Nicolás[b]

September 2003

[a] Department of Economics and Related Studies, University of York, York YO10 5DD, UK

[b] Departament d'Economia i Empresa and CRES, Universitat Pompeu Fabra, 08005-Barcelona, Spain

Abstract

This paper presents a method for the measurement of changes in health inequality and income-related health inequality over time in a population. For pure health inequality (as measured by the Gini coefficient) and income-related health inequality (as measured by the concentration index), we show how measures derived from longitudinal data can be related to cross section Gini and concentration indices that have been typically reported in the literature to date, along with measures of health mobility inspired by the literature on income mobility. We also show how these measures of mobility can be usefully decomposed into the contributions of different covariates. We apply these methods to investigate the degree of income-related mobility in the GHQ measure of psychological well-being in the first nine waves of the British Household Panel Survey (BHPS). This reveals that dynamics increase the absolute value of the concentration index of GHQ on income by 10%.

JEL classification: D63, I12, C21

Keywords: Health inequalities; mobility; Gini and concentration indices; mental health; BHPS.

* Corresponding author. Tel. +44 1904 433790. Fax: +44 1904 433759. E-mail: amj1@york.ac.uk. This paper derives from the project "The dynamics of income, health and inequality over the lifecycle" (known as the ECuity III Project), which is funded in part by the European Community's Quality of Life and Management of Living Resources programme (contract QLK6-CT-2002-02297). We are grateful to the EC, the Fundación BBVA and the Spanish DGES (project BEC2002-04294-C02-02) for financial support. We are grateful to Eddy van Doorslaer and Xander Koolman for their comments.

1. Introduction

Health economics has recently witnessed the development of analytic tools for the measurement and explanation of income-related health inequalities. The concentration index of health on income (Wagstaff et al. [1,2,3]; van Doorslaer et al. [4]; Humphries and van Doorslaer [5]) is nowadays the most popular measure of relative income-related health inequality, and the regression-based decomposition methods of Wagstaff et al. [6] are being used in a variety of settings and populations (e.g., Wildman [7], van Doorslaer and Jones [8]). Recent work by Bommier and Stecklov [9] argues that the concentration index is a more appropriate measure than inequality indices derived from a social welfare function if equity is defined according to a social justice approach. This approach defines "the health distribution in the ideal equitable society as one where access to health has not been determined by socioeconomic status or income" (Bommier and Stecklov, [9, p. 502]). While these methods have been applied to cross sectional information, it is evident that attention must be paid to the dynamics of health and their relation to socio-economic characteristics as revealed by longitudinal data (e.g., Adams et al. [10]; Benzeval et al. [11]; Contoyannis et al. [12, 13]; Hauck and Rice [14]). This paper presents a method for the analysis of health inequalities when longitudinal data is available.

We show that there are important features of income-related health inequality that cannot be revealed by cross sectional data. Our departure point is provided by measurement tools from the income distribution literature. In order to approach a measure of inequality in lifetime income, Shorrocks [15] considered inequality in the distribution of individual incomes averaged over a sequence of time periods. In particular, Shorrocks introduced the concept of income mobility to capture the degree to which income inequalities fade as the time window over which the population is analysed extends. This methodology, which has been used by a variety of authors (e.g., Jarvis and Jenkins [16]; Cantó Sánchez [17]) in the context of income inequality, can be a fruitful empirical tool for the analysis of pure health inequality over the lifecycle. When interest is focused on income-related health inequalities, the parallel question is whether taking a longitudinal perspective reduces or increases income-related health inequality and, if so, how can one measure the relevant change.

In this paper we use Shorrocks's [15] framework for the analysis of mobility with a view to developing a measurement tool for the change in income-related health inequality. Our analysis shows that, whenever there are systematic differences in health among individuals who are

upwardly (income) mobile and downwardly mobile, long-run income-related health inequality will differ from the picture that one might obtain when measurement is made either over a short time span or over a sequence of independent snapshots, which do not capture individual dynamics in income and health. More specifically, if healthy individuals are upwardly mobile and unhealthy individuals are downwardly mobile, income-related health inequality will tend to increase as time passes. Such changes can be measured by an index of *health-related income mobility*.

The analysis presented here is based on the familiar concentration index of health on income (Wagstaff et al. [1]).However it has been argued (e.g. Gakidou *et al*. [18]) that *all* health inequalities can to some extent be a cause of concern, not just those which display a systematic relationship with indicators of socio-economic status. Systematic health disparities have been shown to exist not only with respect to variables like income and education, but also with respect to place of residence, race, marital status, ethnic origin and a host of other characteristics of groups or individuals which health policy makers may find relevant. Consequently, it may be of interest to be able to compute measures of total or 'pure' inequality in health and decompose them into their sources, including socioeconomic factors like income. Although we present our analysis in terms of the concentration index for income-related inequality in health, all of our derivations could be applied to the Gini coefficient for pure inequality in health, simply by replacing income rank with health rank (e.g. Le Grand [19], and Wagstaff, et al. [2]).

One of the attractive features of the concentration index as a measure of income-related inequalities in health is the possibility to incorporate an econometric model for health and subsequently proceed to the decomposition of inequality into the contributions of each of the regressors (Wagstaff et al. [6]). By analogy, we show how health-related income mobility can be decomposed into the contributions of covariates in an econometric model. Both concentration indices and the new index of health-related income mobility require a cardinal measure of health outcome that can be aggregated across time in a meaningful way. We illustrate these methods by analysing the dynamics of income and mental health, as measured by the GHQ index of psychological well-being in the first nine waves of the British Household Panel Survey (BHPS), where GHQ is measured using an (additive) Likert scale. It is worth noting that the mobility index could be applied to other measures of health outcome such as (cardinal) QoL scores or specific indicators of morbidity such as days lost through illness or the number of symptoms reported.

The results of our empirical application reveal that over the long-run, represented here by a period of 9 years, adverse mental health is more concentrated among the poor. In particular, individual dynamics increase the absolute value of the concentration index of GHQ on income by 10%. A

simple econometric model for the GHQ score is able to isolate some of the contributors to this change, but an overwhelming proportion is attributable to unobserved individual heterogeneity.

In Section 2 we demonstrate the desirability of a longitudinal perspective in the analysis of income-related health inequalities by means of a series of simple examples. Section 3 presents the formal derivation of the measure of health-related income mobility and its decomposition through an econometric model for health. Section 4 presents and discusses the empirical results from the BHPS and Section 5 concludes.

2. The advantages of a longitudinal perspective for measuring income-related health inequality

Imagine that we are interested in a population of three individuals and we observe them over three periods. There are three possible health states with (cardinal) outcomes [1,2,3] and three possible income levels [10, 20,30]. First consider a baseline situation where there are no changes, neither in health nor in income: the three individuals always have the same level of health and income, as represented in Table 1.

Table 1: Case 1 values of health and income

Individual Period	1	2	3
1	1	2	3
	10	20	30
2	1	2	3
	10	20	30
3	1	2	3
	10	20	30

The cross sectional concentration indices of health on income in case 1 are:

Period 1: 0.2222

Period 2: 0.2222

Period 3: 0.2222

The joint distribution of individual average health and average income after three periods is given in Table 2. The concentration index of average health on average income in this distribution is 0.2222 again.

Table 2: Case 1 distribution of average income and health

Individual	Mean (health)	Mean (income)
1	1	10
2	2	20
3	3	30

Now consider a situation where there is perfect mobility both in income and health. That is, the three individuals each experience the three possible health states and income states over the three periods. This is shown in Table 3.

Table 3: Case 2 values of health and income

Individual Period	1	2	3
1	1	2	3
	10	20	30
2	3	1	2
	30	10	20
3	2	3	1
	20	30	10

The cross sectional concentration indices for case 2 are exactly the same as in the baseline situation. However, the joint distribution of individual average income and average health after three periods is shown in Table 4. The concentration index for this distribution is 0. Clearly, the cross sectional measures are unable to capture this important difference with respect to the baseline situation. In this case mobility, understood as the percentage by which the concentration index for the distribution of individual averages differs from the average of cross sectional measures, is 100 per cent.

Table 4: Case 2 distribution of average income and health

Individual	Mean (health)	Mean (income)
1	2	20
2	2	20
3	2	20

Now consider a third situation, in Table 5, whereby an individual has poor health over the three periods but, from a cross sectional point of view, does not fare too badly in terms of income, because he ranks second in the income distribution each period. Over time, however, the other two individuals experience changes in income.

Table 5: Case 3 values of health and income

Individual Period	1	2	3
1	1	2	3
	15	10	30
2	1	2	3
	15	30	10
3	1	2	3
	15	10	30

The cross sectional concentration indices suggest a lower degree of income-related health inequality than cases 1 and 2. Moreover, in period 2 there is even pro-poor income-related health inequality:

Period 1: 0.1111

Period 2: -0.1111

Period 3: 0.1111

When we analyse the distribution of average income and health over the three periods, in Table 6, the picture is completely different. The concentration index of average health on average income for this distribution is 0.2222. That is, income-related health inequality increases over time.

Table 6: Case 2 distribution of average income and health

Individual	Mean (health)	Mean (income)
1	1	15
2	2	16.667
3	3	23.333

It is important to stress the potential empirical relevance of this last scenario. As far as income is concerned, individual 1 is representative of a section of the population with steady, albeit low, income. The other individuals are representative of the population who receive an income flow which, albeit more irregular, results in a greater level of average income over the life cycle. Cross sectional measures of income-related inequality would completely miss the positive association between income and health over the life cycle that this case attempts to capture.

These three scenarios illustrate the advantages of taking an inter-temporal perspective. The preceding discussion suggests how longitudinal information permits the calculation of income-related inequality measures over a long time span. In the following section we show the relation of one of these measures, the concentration index of health on income, to its cross sectional counterparts and formally derive an index of health-related income mobility inspired by the results in Shorrocks [15].

3. Health-related income mobility and income-related health inequality

Suppose that longitudinal data is available and we can observe a population of N individuals over T time periods. Consider the following variables:

y_{it} : a cardinal measure of the health of individual i (i=1,…,n) at time t (t=1,…,T).

$y_i^T=(1/T)\Sigma_t y_{it}$: average health of individual i after T periods.

R_i^t: relative rank of individual i in the distribution of N incomes in period t.

R_i^T: relative rank of individual i in the distribution of average incomes after T periods.

Define the following means:

$$i) \quad \bar{y}^t = \frac{\sum_i y_{it}}{N}$$

$$ii) \quad \bar{\bar{y}}^T = \frac{\sum_t \sum_i y_{it}}{NT} = \frac{\sum_t \bar{y}_t}{T};$$

That is, i) is the within-period average health status and ii) is the overall average health status in T periods. This formulation is based on a balanced cohort of individuals in which all individuals are observed in all T periods. In reality there are likely to be exits from the population due, for example, to deaths. In this case the derivations presented here remain valid so long as the population is maintained at a constant size, by assigning a zero value of health to those who have died (see Shorrocks [15, p.383].

We can now write the concentration index of health on income for each sub-period, t, as (e.g., Kakwani [20]),

$$CI^t = \frac{2}{\bar{y}_t}\text{cov}\left(y_{it}, R_i^t\right) = \frac{2}{N\bar{y}_t}\sum_i \left(y_{it} - \bar{y}_t\right)\left(R_i^t - 1/2\right) \tag{1}$$

The corresponding expression for the Gini coefficient of health would be obtained replacing the income rank by the health rank (Lerman and Yitzhaki [21]; Lambert [22]). Similarly, we can define the concentration index for the distribution of average health after T periods as,

$$CI^T = \frac{2}{\bar{\bar{y}}^T}\text{cov}\left(y_i^T, R_i^T\right) = \frac{2}{N\bar{\bar{y}}^T}\sum_i \left(y_i^T - \bar{\bar{y}}^T\right)\left(R_i^T - 1/2\right) = \frac{2}{NT\bar{\bar{y}}^T}\sum_i \left(\sum_t y_{it} - \sum_t \bar{y}_t\right)\left(R_i^T - 1/2\right) \tag{2}$$

Note that,

$$R_i^T - 1/2 = \left(R_i^t - 1/2\right) - \left(R_i^t - R_i^T\right) \tag{3}$$

So,

$$CI^T = \frac{2}{NT\bar{\bar{y}}^T}\sum_i\sum_t \left(y_{it} - \bar{y}^t\right)\left(R_i^t - 1/2\right) - \frac{2}{NT\bar{\bar{y}}^T}\sum_i\sum_t \left(y_{it} - \bar{y}^t\right)\left(R_i^t - R_i^T\right) \tag{4}$$

Noting that the first term in the expression above is easily related to the within-period concentration indices, this can be re-written as,

$$CI^T = \sum_t w_t CI^t - \frac{2}{NT\bar{\bar{y}}^T}\sum_i\sum_t \left(y_{it} - \bar{y}^t\right)\left(R_i^t - R_i^T\right) \qquad where \quad w_t = \frac{\bar{y}^t}{T\bar{\bar{y}}^T} \tag{5}$$

Equation (5) is a key result. It shows that the concentration index for average health after T periods can be written down as the sum of two terms. The first term is a weighted sum of the concentration indices for each of the sub-periods (with weights equal to the share of "total" health in each period). If the income ranking remains constant over time, the standard decomposition result tells us that the concentration index for the average over time is equal to the (weighted) average of the concentration indices. However income ranks may change over time. The second term in equation (5) captures the difference between period specific income ranks and ranks for average income over all periods and their relationship to health. This expression is related to the family of measures developed by Shorrocks. In fact, in the case of the Gini coefficient, the results in Shorrocks [15, p. 382] guarantee that the second component has zero as a lower bound. The lower bound would be attained when individuals never change their rank in the health distribution as time passes. For the concentration index, however, the second term could be either positive or negative. This term will be different from zero if the following two conditions hold:

i) The income rank of individuals is sensitive to the length of the time window over which measurement is taken, i.e. there is income mobility, as defined by Shorrocks.

ii) These changes in income rank are associated with systematic differences in health.

Table 7: Concentration and mobility indices for the hypothetical scenarios

Case 1	Period	CI^t	Term 1	Term 2	CI^T	M^T
	1	0.2222	0.2222	0	0.2222	0
	2	0.2222	0.2222	0	0.2222	0
	3	0.2222	0.2222	0	0.2222	0
Case 2	Period	CI^t	Term 1	Term 2	CI^T	M^T
	1	0.2222	0.2222	0	0.2222	0
	2	0.2222	0.2222	0.1111	0.1111	0.5
	3	0.2222	0.2222	0.2222	0	1
Case 3	Period	CI^t	Term 1	Term 2	CI^T	M^T
	1	0.1111	0.1111	0	0.1111	0
	2	-0.1111	0	-0.2222	0.2222	0.8947
	3	0.1111	0.037	-0.1851	0.2221	-5

Table 7 illustrates the decomposition, given by equation (5), for the three hypothetical scenarios described in Section 2. In case 1 there is no income mobility, so the second term is zero regardless of whether there is health mobility. In case 2, the health level of each individual moves in parallel to their change in income rank. Consequently the second term counteracts the first term and income-related health inequality is zero when measurement is taken over the three periods. In case 3 there

are no health changes, but the second term still is not zero. This is because individuals who are downwardly (income) mobile, in the sense that, in the long-run, their income rank is lower than in the short-run (even if their absolute income does not change) - individuals of type 1 - have a lower than average level of health in the short-run, compared to individuals who are upwardly mobile - individuals of type 3. This exacerbates income-related health inequality when a lifetime perspective is taken, as reflected in the negative sign of the second term. Case 3 illustrates an important situation. Even if individuals do not experience health changes, long-run income-related inequality can be greater than that obtained with snapshot cross-sectional estimates, as long as the patterns of income mobility are systematically related to health. Averaging the short-run measures of inequality will tend to underestimate the long-run picture when individuals whose short-run income position is better than their long-run position tend to have lower than average health.

It is useful to measure how much the longitudinal perspective alters the picture that would emerge from a series of cross sections, in the same spirit as Shorrocks' [15] index of income mobility. We may define an index of *health-related income mobility* as,

$$M^T = 1 - \frac{CI^T}{\sum_t w_t CI^t} = \frac{2}{N\sum_t \bar{y}^t CI^t}\left(\sum_i \sum_t \left(y_{it} - \bar{y}^t\right)\left(R_i^t - R_i^T\right)\right) \tag{6}$$

This is simply one minus the ratio by which the concentration index for the joint distribution of longitudinal averages differs from the weighted average of the cross sectional concentration indices due to the systematic association between health and changes in the income rank of an individual. Note that in situations where income-related inequality tends to fade either solely due to health mobility or solely due to income mobility, the index in equation (6) would be zero. In these cases the second term in equation (5) would be zero and the information contained in the series of cross sectional concentration indices would be sufficient to capture the dynamics of interest.

The last column in Table 7 contains the values of the mobility index for each of the three hypothetical cases. Obviously in case 1 the index is equal to zero. In case 2, income-related health inequality vanishes after three periods due to the perfect association between the health level and the income rank. Thus the mobility index equals one in absolute value. Since the association between health and changes in the income rank of the individuals in case 3 makes the level of lifetime income-related health inequality greater than that which we could infer from the cross sectional information, the index is negative. The particular numbers chosen for the example result

in a 500% increase in income-related health inequality when the longitudinal perspective is adopted ($CI^T = 0.2221$), rather than using the weighted average of the cross section concentration indices (=0.037). For the purposes of next section, it is convenient to write equation (6) in a slightly modified way. In particular note that we can write,

$$M^T = 1 - \frac{CI^T}{\sum_t w_t CI^t} = \frac{2}{N\sum_t \bar{y}^t CI^t}\left(\sum_i\sum_t \left(y_{it} - \bar{y}^t\right)\left(R_i^t - R_i^T\right)\right) =$$
$$\frac{2}{N\sum_t \bar{y}^t CI^t}\left(\sum_i\sum_t \left(y_{it} - \bar{\bar{y}}^T\right)\left(R_i^t - R_i^T\right) - \sum_i\sum_t \left(\bar{y}^t - \bar{\bar{y}}^T\right)\left(R_i^t - R_i^T\right)\right) =$$
$$\frac{2}{N\sum_t \bar{y}^t CI^t}\sum_i\sum_t \left(y_{it} - \bar{\bar{y}}^T\right)\left(R_i^t - R_i^T\right) \tag{7}$$

4. Decomposition of the mobility index by factors

This section shows how, departing from an econometric model for the level of health, the mobility measure can be decomposed into the contributions of different regressors. This extends the regression-based decomposition of the concentration index by factors presented in Wagstaff et al. [6]. Assume that the level of health can be adequately represented by the following linear regression model,

$$y_{it} = \alpha + \sum_{k=1}^{K} \beta_k x_{itk} + u_{it} = \hat{\alpha} + \sum_{k=1}^{K} \hat{\beta}_k x_{itk} + \hat{u}_{it} \tag{9}$$

The first order conditions for the least squares estimation of the model imply that,

$$\bar{\bar{y}} = \hat{\alpha} + \sum_{k=1}^{K} \hat{\beta}_k \bar{\bar{x}}_k \tag{10}$$

Now substitute equations (9) and (10) into equation (7) to obtain,

$$M^T = \frac{2}{N\sum_t \bar{y}^t CI^t}\sum_i\sum_t \left(y_{it} - \bar{\bar{y}}^T\right)\left(R_i^t - R_i^T\right) =$$

$$\frac{2}{N\sum_t \bar{y}^t CI^t}\left[\sum_{k=1}^{K}\hat{\beta}_k \sum_i\sum_t \left(x_{itk} - \bar{\bar{x}}_k\right)\left(R_i^t - R_i^T\right) + \sum_i\sum_t \hat{u}_{it}\left(R_i^t - R_i^T\right)\right] \tag{11}$$

Define the x_k-related income mobility index after T periods as,

$$M_{x_k}^T = \frac{2}{N\sum_t \bar{x}^t CI_{x_k}^t}\sum_i\sum_t \left(x_{itk} - \bar{\bar{x}}_k\right)\left(R_i^t - R_i^T\right)$$

Equation (11) can be written as the following expression,

$$M^T = \frac{2}{N\sum_t \bar{y}^t CI^t}\left[\sum_{k=1}^{K}\hat{\beta}_k \sum_i\sum_t \left(x_{itk} - \bar{\bar{x}}_k\right)\left(R_i^t - R_i^T\right) + \sum_i\sum_t \hat{u}_{it}\left(R_i^t - R_i^T\right)\right] =$$

$$M^T = \sum_{k=1}^{K}\hat{\beta}_k \frac{\sum_t \bar{x}_k^t CI_{x_k}^t}{\sum_t \bar{y}^t CI^t} M_{x_k}^T + \frac{2}{N\sum_t \bar{y}^t CI^t}\sum_i\sum_t \hat{u}_{it}\left(R_i^t - R_i^T\right) = \sum_{k=1}^{K}\hat{\beta}_k \frac{\sum_t \bar{x}_k^t CI_{x_k}^t}{\sum_t \bar{y}^t CI^t} M_{x_k}^T + residual$$

$$M^T = \hat{M}^T + residual$$

(12)

Note that $\hat{\beta}_k \frac{\sum_t \bar{x}_k^t CI_{x_k}^t}{\sum_t \bar{y}^t CI^t}$ takes the form of an "elasticity" of health with respect to x_k, evaluated at the income-related inequality weighted means of x_k and health. The presence of the term $CI_{x_k}^t$ influences the sign of the elasticity. For example, assuming that there is pro-rich inequality in health ($CI^t > 0$), a regressor that has a positive association with health, reflected in a positive $\hat{\beta}_k$, but which has a pro-poor distribution ($CI_{x_k}^t < 0$), will have a negative "elasticity". In this sense, the expression gives an *inequality-weighted elasticity*.

So, the explained part of the mobility index, i.e. $\hat{M}^T$, can be decomposed into a sum of contributions from the regressors in the model. Each contribution is the inequality weighted elasticity of health with respect to the corresponding variable multiplied by the index of x_k-related

income mobility. As we shall see in the empirical application, this expression retains the interpretability of the familiar decompositions for the cross sectional concentration index.

5. An empirical application

5.1 The BHPS data

To illustrate the methods proposed above we estimate a model of psychological well-being based on scores from the General Health Questionnaire (GHQ) as measured in the British Household Panel Survey (BHPS). This application was chosen because the BHPS is a recent panel data set with good quality income and socio-economic variables and because the GHQ measure can be modelled conveniently in a linear regression framework. We use nine waves of data (1991-1999), to capture income and health mobility.

The BHPS is a longitudinal survey of private households in Great Britain (England, Scotland and Wales). It was designed to be a nationally representative survey of over 5,000 households and gives around 10,000 individual interviews of each adult (16+) household member. The BHPS is a repeated panel, with respondents questioned each year. The initial sample, collected in 1991, was selected using a two-stage stratified sampling procedure, designed to give each address an approximately equal probability of selection (Taylor et al. [23]). The first stage consisted of selecting 250 postcode sectors, with probabilities proportional to their size. The second stage selected delivery points. If multiple addresses were found the interviewer selected a particular household. The first wave of the survey was carried out between 1st September 1990 and 30th April 1991. The same individuals are re-interviewed in successive waves and, if new households are formed, the original sample members are interviewed along with all adults in the new household. Information at both the household and the individual level is collected, covering questions on neighbourhood, income, employment, health and caring, demographics, and values and opinions.

For our illustrative application a subset of individuals who had a full interview at each of the nine waves, between 1991 and 1999, is used. This balanced sample mimics the balanced population used in the derivations of the indices in section 3. A more elaborate analysis might take account of entry and exit in the sample and use inverse probability weights to ensure that each wave is cross-sectionally representative. Instead our results show the impact of comparing short-run and long-run measures of health for the cohort of individuals who appear in wave 1 and can be followed through

all subsequent waves. Selecting the balanced cases gives an initial sub-sample of 6,080 individuals. From these we have dropped those who do not report the GHQ score in all waves (1,097 individuals) and those whose full household income is reported to be either below £2000 or above £77,000 during any of the 9 waves (426 individuals). Thus the estimating sample contains 4,557 individuals observed during 9 waves, that is a sample of 41,103 observations altogether. The final sample contains individuals with missing values at particular waves for some of the dummy variables that are used in the regression model: one individual does not report marital status, 23 individuals do not report social class, 22 do not report job status and 7 individuals do not report their education level. Rather than dropping the individuals for which such item non-response arises, we have set the values of the corresponding dummies to zero and have kept them in the sample.

The BHPS self-completion questionnaire incorporates a reduced version of the General Household Questionnaire (Goldberg and Williams [24]; Bowling [25]). The GHQ was developed as a screening instrument for psychiatric illness and is now often used as an indicator of psychological well-being (Hauck and Rice [14]; Weich et al. [26]; Wildman [7]). The shortened GHQ includes 12 elements: concentration, sleep loss due to worry, perception of role, capability in decision making, whether constantly under strain, perception of problems in overcoming difficulties, enjoyment of day-to-day activities, ability to face problems, loss of confidence, self-worth, general happiness, and whether suffering depression or unhappiness. Responses are given on a 4-point scale ranging from 0 to 3, with 0 being the best score. For our dependent variable we use the Likert scale, which sums the individual components (Likert [27]). This gives an overall scale that runs from 0 to 36. To make the interpretation of results more intuitive and consistent with the discussion in sections 3 and 4, we have re-scaled this measure in order to make it increasing in good health. Therefore we use GHQ'=36-GHQ rather than the original GHQ score.

5.2 Mobility indices and decompositions

The CI^t column in Table 8 presents the concentration index of the GHQ' score on income in each of the 9 waves of the BHPS. These indices are all positive, indicating that there is 'pro-rich' inequality in health in all periods. Also, note the variability in the magnitude of the concentration indices across waves. The CI^T column presents the sequence of concentration indices using averages over one, two, etc. periods. The discrepancy between the short-run and long-run measures is illustrated on Figure 1. After nine periods, the degree of pro-rich health inequality is slightly

smaller than at the start of the period. What drives this reduction? The columns headed 'Term 1' and 'Term 2' refer to equation (5) and provide the answer.

Insert Figure 1 around here.

Term 1 is simply the weighted average of the cross sectional concentration indices up to the corresponding wave. Note that there is a slight downward trend in this term, as the cross sectional concentration indices for the middle periods are smaller than at the start and end of the nine years. Therefore, the decrease in pro-rich inequality in health within some of the periods is contributing to the reduction. However, the weighted average of the cross sectional concentration indices is smaller in absolute value than the indices that use the distribution of longitudinal averages. Accordingly, the series of estimates in the Term 2 column are negative. This suggests that downwardly (income) mobile individuals tend to have below average levels of health compared to upwardly mobile individuals. As we have discussed before, this makes long-run income-related health inequality greater than what we could infer from the cross sectional measures. In this case, this effect increases long-run income-related health inequality by 10%, as reflected by the mobility index (M^I) of –0.102 in the last column.. Figure 2 illustrates the health-related income mobility index and shows that, by the ninth wave, the short run measure under estimates long run inequality by around 10%.

Insert Figure 2 around here.

Table 8: Concentration and mobility indices for GHQ scores.

Wave	CI^t	Term1	Term2	CI^T	M^T
1	0.01216	0.01216	0.00000	0.01216	0.00000
2	0.01020	0.01119	-0.00043	0.01162	-0.03805
3	0.01144	0.01127	-0.00047	0.01174	-0.04179
4	0.00641	0.01006	-0.00055	0.01061	-0.05446
5	0.00972	0.00999	-0.00081	0.01080	-0.08072
6	0.01146	0.01023	-0.00059	0.01083	-0.05804
7	0.00903	0.01006	-0.00076	0.01082	-0.07555
8	0.00885	0.00991	-0.00095	0.01086	-0.09576
9	0.01311	0.01027	-0.00105	0.01132	-0.10213

5.3 Decomposition of mobility by factors

Next we use a regression model for the level of GHQ' score in order to decompose the health-related mobility index into the contributions of different covariates. The specification of the model is drawn from the analysis of the BHPS in Hauck and Rice [14], although the estimation is kept simple (a linear regression model that takes no account of potential error components structures and no dynamics). The intention of the regression model is simply to capture the linear association between the GHQ' score and a range of socioeconomic characteristics. It should not be taken as a structural model or used to infer a direction of causality. Table 9 gives the names and definitions of the regressors and full details of the OLS regression results are given in a table in the Appendix. The coefficients are presented along with Huber-White robust standard errors that are adjusted for clustering within-individuals due to the use of panel data.

Table 9: Names and definitions of regressors

Variable Name	Definition	Mean
Log(income)	Logarithm of equivalised annual household income, using √(household size)	9.33
Age	Age in years at 1st December of current wave	46.43
Age2	Age squared/100	24.21
Age3	Age cubed/10,000	13.87
Widowed	1 if widowed, 0 otherwise	0.07
Never married	1 if never married, 0 otherwise	0.13
Divorced	1 if divorced or separated, 0 otherwise	0.07
Children	Number of children in household	0.62
Professional	1 if Registrar General's SC is professional occupation, 0 otherwise	0.03
Managerial	1 if Registrar General's SC is managerial or technical occupation, 0 otherwise	0.20
Skilled	1 if Registrar General's SC is skilled manual occupation, 0 otherwise	0.13
Unskilled	1 if Registrar General's SC is partly skilled or unskilled occupation, 0 otherwise	0.03
Unemployed	1 if unemployed, 0 otherwise	0.03
Retired	1 if retired, 0 otherwise	0.19
Carer	1 if family carer, 0 otherwise	0.09
Student	1 if student, 0 otherwise	0.01
Other SC	1 if Registrar General's SC is other, 0 otherwise	0.44
Degree	1 if highest academic qualification is degree or higher degree, 0 otherwise	0.11
HND/HNCT	1 if highest academic qualification is HND or HNCT, 0 otherwise	0.07
O/CSE	1 if highest academic qualification is O level or CSE, 0 otherwise	0.31
No qualification	1 if no academic qualifications, 0 otherwise	0.34
Non-white	1 if ethnic origin is other than white, 0 otherwise	0.02
Female	1 if female, 0 otherwise	0.55

Table 10: Decomposition of the mobility index by factors

	CI(x) after 9 periods	Mobility(x)	Elasticity(x)	Contrib(x)
Log (income)	0.0308	0.1413	0.7481	0.1057
Age	-0.0544	-0.1564	4..643	-0.7261
Age2	-0.1204	-0.1402	-10.158	1.4242
Age3	-0.1902	-0.1309	5.212	-0.6822
Widowed	-0.5581	-0.0918	0.0675	-0.0062
Never married	0.0891	-0.4745	-0.0072	0.0034
Divorced	-0.2298	0.109	0.0725	0.0079
Children	-0.1353	0.1112	-0.0252	-0.0028
Professional	0.4998	-0.0978	-0.0102	0.001
Managerial	0.4007	-0.0595	-0.0303	0.0018
Skilled	0.083	0.113	0.0248	0.0028
Unskilled	-0.2345	-0.1387	-0.0087	0.0012
Unemployed	-0.2914	0.0819	0.0269	0.0022
Retired	-0.3963	-0.0501	-0.2196	0.011
Carer	-0.3261	0.0254	-0.0118	-0.0003
Student	-0.0314	0.8006	-0.0027	-0.0022
Other SC	-0.3054	-0.0361	0.03213	-0.0116
Degree	0.431	-0.1777	-0.0011	0.0002
HND/HNCT	0.3244	-0.1542	0.0370	-0.0057
O/CSE	0.0136	-0.0849	0.0035	-0.0003
No qualification	-0.2891	-0.1558	0.1046	-0.0163
Non-white	0.0069	1.8253	0.0004	0.0008
Female	-0.0626	-0.1339	0.1344	-0.018
Residual				-0.1926
Total index				-0.1021

Table 10 presents the contribution of each of the regressors to the index of health-related income mobility after 9 periods. The first column presents the long-run concentration index of the corresponding regressor on income. A positive CI indicates that the variable has a pro-rich distribution and a negative value that it has a pro-poor distribution. For example, the index of 0.431 indicates that those with degrees are more concentrated among those with higher incomes, while the index of –0.2891 indicates that those with no qualifications are more concentrated among lower incomes. The second column contains the x_k-related income-related mobility index. This indicates whether taking the long-run average of the regressor makes its distribution more or less unequal in absolute terms (whether pro-rich or pro-poor). A negative mobility index implies that the weighted average of short-run indices underestimates the degree of long-run inequality (whether pro-rich or pro-poor), while a positive value indicates that the weighted average overestimates the long-run inequality. For example, the mobility index of –0.1777 for 'Degree' indicates that the short-run measure underestimates long-run pro-rich inequality by 17.8%. The mobility index of –0.1558 for 'No qualification' indicates that the short-run measure underestimates long-run pro-poor inequality by 15.6%. The third column contains the inequality-weighted "elasticity" of health with respect to the regressor, based on the estimates from the econometric model for health. Following equation (12), the product of the mobility index and "elasticity" gives the actual contribution of the regressor to the index of health-related income mobility which is shown in the final column.

Starting with equivalent income, note that this variable has a positive income mobility (M^T_x=0.1413) so that taking a long-run average reduces the degree of (pro-rich) income inequality and, at the same time, a positive elasticity (0.7481). Thus, the dynamics of income would, ceteris paribus, reduce the degree of pro-rich income-related health inequality in the long-run. At 10%, this is the largest effect among the regressors in this model.

Now consider the effect of the marital status dummy variables, for which the omitted category is married or living as a couple. There is more pro-poor inequality in widowhood in the long-run, so its incidence makes health more concentrated among the rich in the long-run. Divorce, on the other hand, is less concentrated among the poor in the long-run, so it also makes health less concentrated among the rich in the long-run. Bachelorhood is more concentrated among the rich in the long-run, so it makes health less concentrated among the rich in the long-run. Having children is positively associated with health, and since there is less pro-poor inequality in the number of children in the long-run, fertility contributes to making health more concentrated among the rich in the long-run.

The coefficients for the dummy variables for social class suggest that there are no significant differences in GHQ' score among professionals, managers and skilled non-manual workers (the omitted category). The coefficient for unskilled workers suggests a positive association with GHQ' score, and because this condition is more pro-poor unequally distributed in the long-run, it contributes to reduce income-related pro-rich health inequality in the long-run. Unemployment is negatively associated to the GHQ' score but it is less concentrated among the poor in the long-run. It contributes with a reduction in income-related health inequality in the long-run. Being retired is positively associated with health and, because there is more pro-poor inequality in retirement in the long-run, it contributes with a reduction in income-related health inequality. The effects of family care and student are not statistically significant.

As for the education variables, note that having an university degree is not significantly different from having A levels (the omitted category). The coefficients for HND/HNCT and O/CSE suggest a positive association with health. Since there is more pro-rich inequality in these qualifications in the long-run, they contribute to make health more pro-rich unequally distributed. Having no qualifications acts in the opposite direction although it generates the same final effect: it is negatively associated with health and it is more pro-poor unequally distributed in the long-run so it contributes to more income-related health inequality in the long-run.

The coefficient on the dummy variable for non-white ethnic origin (with the omitted category set to white) reveals a negative association with health. But note that the long-run concentration index of non-white ethnic origins on income is positive, while the cross sectional counterparts are negative. This implies a high degree of upward income mobility in this group. Given these features, this condition contributes to make health less concentrated among the rich.

Women tend to have poorer health than men, and at the same time are downwardly income mobile. Gender is therefore one of the factors that contribute to make health more concentrated among the rich in the long-run.

Finally, the contribution of the residuals of the econometric model outweighs the contributions of the regressors, an indication of the presence of systematic unobserved heterogeneity which will have to be tackled with more sophisticated econometric specifications.

6. Conclusion

This paper presents a method for the analysis of health inequalities when longitudinal data is available and we show that there are important features of income-related health inequality that cannot be revealed by cross sectional data. We use Shorrocks's [15] framework for the analysis of mobility with a view to developing a measurement tool for the change in income-related health inequality when a longitudinal perspective is adopted. Our analysis shows that, whenever there are systematic differences in health among individuals who are upwardly (income) mobile and downwardly mobile, long-run income-related health inequality will differ from the picture that one might obtain when measurement is made either over a short time span or over a sequence of independent snapshots. If healthy individuals are upwardly mobile and unhealthy individuals are downwardly mobile, income-related health inequality will tend to increase as time passes and vice versa. Such changes can be measured by an index of health-related income mobility.

The analysis is based on the familiar concentration index of health on income. One of the attractive features of the concentration index as a measure of income-related inequalities in health is the possibility to incorporate an econometric model for health and subsequently proceed to the decomposition of inequality into the contributions of each of the regressors (Wagstaff et al. [6]). By analogy, we show how health-related income mobility can be decomposed into the contributions of covariates in an econometric model. We illustrate these methods by analysing the dynamics of income and mental health as measured by the GHQ index of psychological well-being in the first nine waves of the British Household Panel Survey (BHPS). The results reveal that over the long-run, represented here by a period of 9 years, adverse mental health is more concentrated among the poor. In particular, individual dynamics increase the absolute value of the concentration index of health on income by 10%. A simple econometric model for mental health is able to isolate some of the contributors to this change, but an overwhelming proportion is attributable to unobserved individual heterogeneity.

References

1. Wagstaff, A., van Doorslaer, E., Paci, P., 1989. Equity in the finance and delivery of health care: some tentative cross-country comparisons. *Oxford Review of Economic Policy* 5, 89-112.

2. Wagstaff, A., Paci, P., van Doorslaer, E., 1991. On the measurement of inequalities in health. *Social Science and Medicine* 33, 545-557.

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

4. Van Doorslaer, E., Wagstaff, A., Bleichrodt, H., *et al*, 1997. Income-related inequalities in health: some international comparisons. *Journal of Health Economics* 16, 93-112.

5. Humphries, K, E van Doorslaer, 2000, Income-related inequalities in health in Canada, *Social Science and Medicine*, 50, 663-671

6. Wagstaff, A, van Doorslaer, E., Watanabe, N., 2003. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam, *Journal of Econometrics*, 112: 207-223.

7. Wildman, J., 2003. Income related inequalities in mental health in Great Britain: analysing the causes of health inequality over time. *Journal of Health Economics* 22, 295-312

8. Van Doorslaer, E, Jones, A., 2003. "The determinants of inequalities in self reported health: a validation of a new approach to measurement. *Journal of Health Economics*.22, 61-87.

9. Bommier, A., Stecklov, G., 2002. Defining health inequality: why Rawls succeeds where social welfare theory fails. *Journal of Health Economics* 21, 497-513.

10. Adams, P., Hurd, M.D., McFadden, D., Merrill, A., Tiberiro, T., 2003. Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics* 112, 3-56.

11. Benzeval, M., Taylor, J., Judge, K., 2000. Evidence on the relationship between low income and poor health: is the Government doing enough? *Fiscal Studies* 21, 375-399.

12. Contoyannis, P., Jones, A.M., Rice, N., 2004, Simulation-based inference in dynamic panel probit models: an application to health. *Empirical Economics* forthcoming.

13. Contoyannis, P., Jones, A.M., Rice, N., 2004, The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics,* forthcoming.

14. Hauck, K., Rice, N., 2003, Health mobility in the UK: a longitudinal analysis of psychological well-being. *University of York,* mimeo.

15. Shorrocks, A., 1978, Income inequality and income mobility. *Journal of Economic Theory* 19, 376-393.

16. Jarvis, S., Jenkins, S., 1998. How much income mobility is there in Britain? *The Economic Journal* 108, 428-443.

17. Cantó Sánchez, O. 2000. Income mobility in Spain: How much is there?. *Review of Income and Wealth* 46(1), 85-102.

18. Gakidou, E, Murray, C., Frenk, J., 2000. Defining and measuring health inequality. *Bulletin of the World Health Organisation* 78(1), 42-52.

19. Le Grand, J., 1989. An international comparison of ages-at-death. In J. Fox (ed.) *Health inequalities in European countries*. Gower, Aldershot.

20. Kakwani, N.C., 1980. *Income Inequality and Poverty. Methods of Estimation and Policy Implications.* Oxford University Press, Oxford.

21. Lerman, R.I., Yitzhaki, S., 1989. Improving the accuracy of estimates of Gini coefficients. *Journal of Econometrics* 42, 43-47.

22. Lambert, P., 1993. The distribution and redistribution of income. A mathematical analysis. 2nd Edition. Manchester University Press. Manchester.

23. Taylor, M.F. (ed.) with Brice, J., Buck, N., and Prentice, E., 1998, *British Household Panel Survey User Manual,* University of Essex, Colchester.

24. Goldberg, D., Williams, P., 1988, *A user's guide to the General Health Questionnaire.* Nfer-Nelson, Windsor.

25. Bowling, A., 1991, *Measuring health. A review of quality of life measurement scales.* OUP, Milton Keynes.

26. Weich, S., Lewis, G., Jenkins, S.P., 2001, Income inequality and the prevalence of common mental disorders in Britain. *British Journal of Psychiatry* 178, 222-227.

27. Likert, R., 1952, A technique for the development of attitude scales. *Educational and Psychological Measurement* 12, 313-315.

Appendix: Full regression results (robust standard errors adjusted for clustering)

Variable Name	Coefficient	Std. error	t-ratio	p-value
Log(income)	0.572	0.088	6.54	0.000
Age	-0.545	0.072	-7.58	0.000
Age2	1.017	0.147	6.91	0.000
Age3	-0.572	0.094	-6.07	0.000
Widowed	-0.482	0.222	-2.18	0.030
Never married	-0.237	0.172	-1.38	0.169
Divorced	-1.084	0.228	-4.75	0.000
Children	0.067	0.059	1.13	0.257
Professional	-0.166	0.233	-0.71	0.475
Managerial	-0.104	0.137	-0.76	0.446
Skilled	0.527	0.145	3.63	0.000
Unskilled	0.379	0.252	1.50	0.133
Unemployed	-0.701	0.248	-2.82	0.005
Retired	0.793	0.19	4.17	0.000
Carer	0.088	0.187	0.47	0.639
Student	0.324	0.286	1.13	0.258
Other SC	-0.632	0.158	-4.00	0.000
Degree	-0.007	0.186	-0.04	0.971
HND/HNCT	0.479	0.216	2.22	0.027
O/CSE	0.217	0.147	1.48	0.138
No qualification	-0.311	0.169	-1.84	0.066
Non-white	-0.538	0.355	-1.51	0.130
Female	-1.129	0.109	-10.34	0.000
Constant	29.022	1.262	22.99	0.000

Figure 1: Short-run and long-run concentration indices

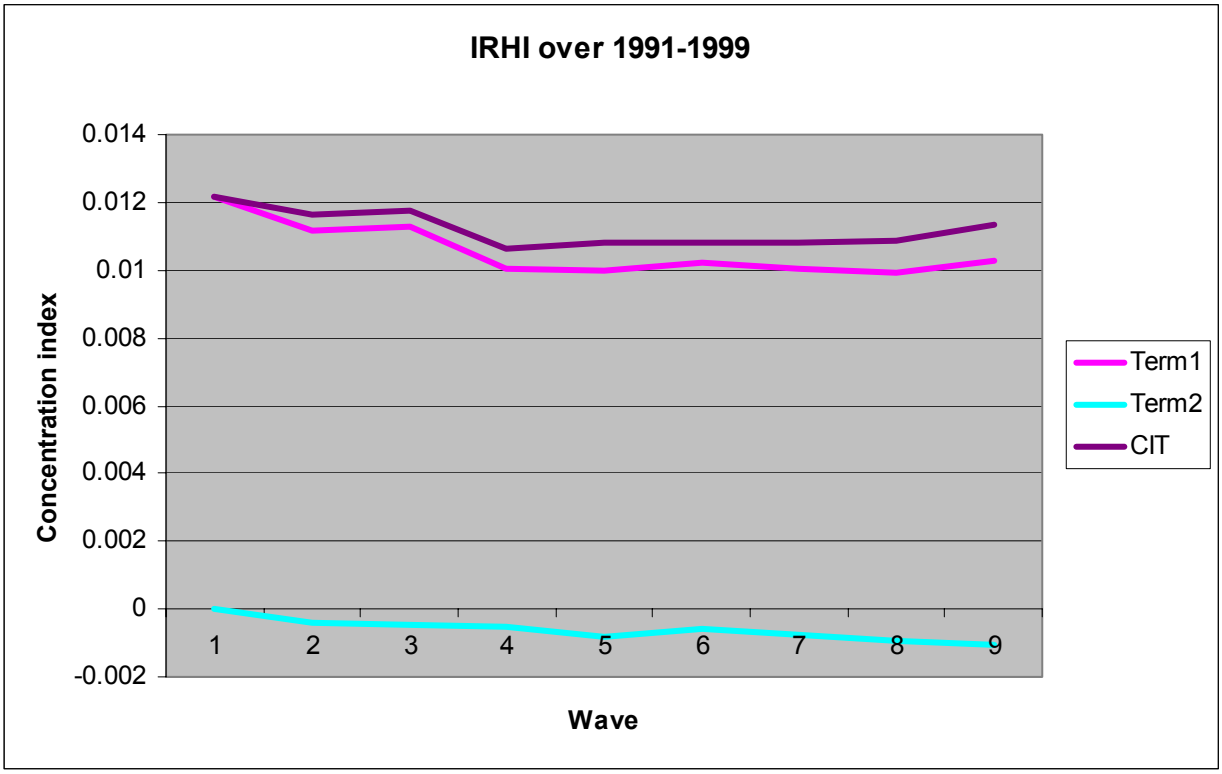
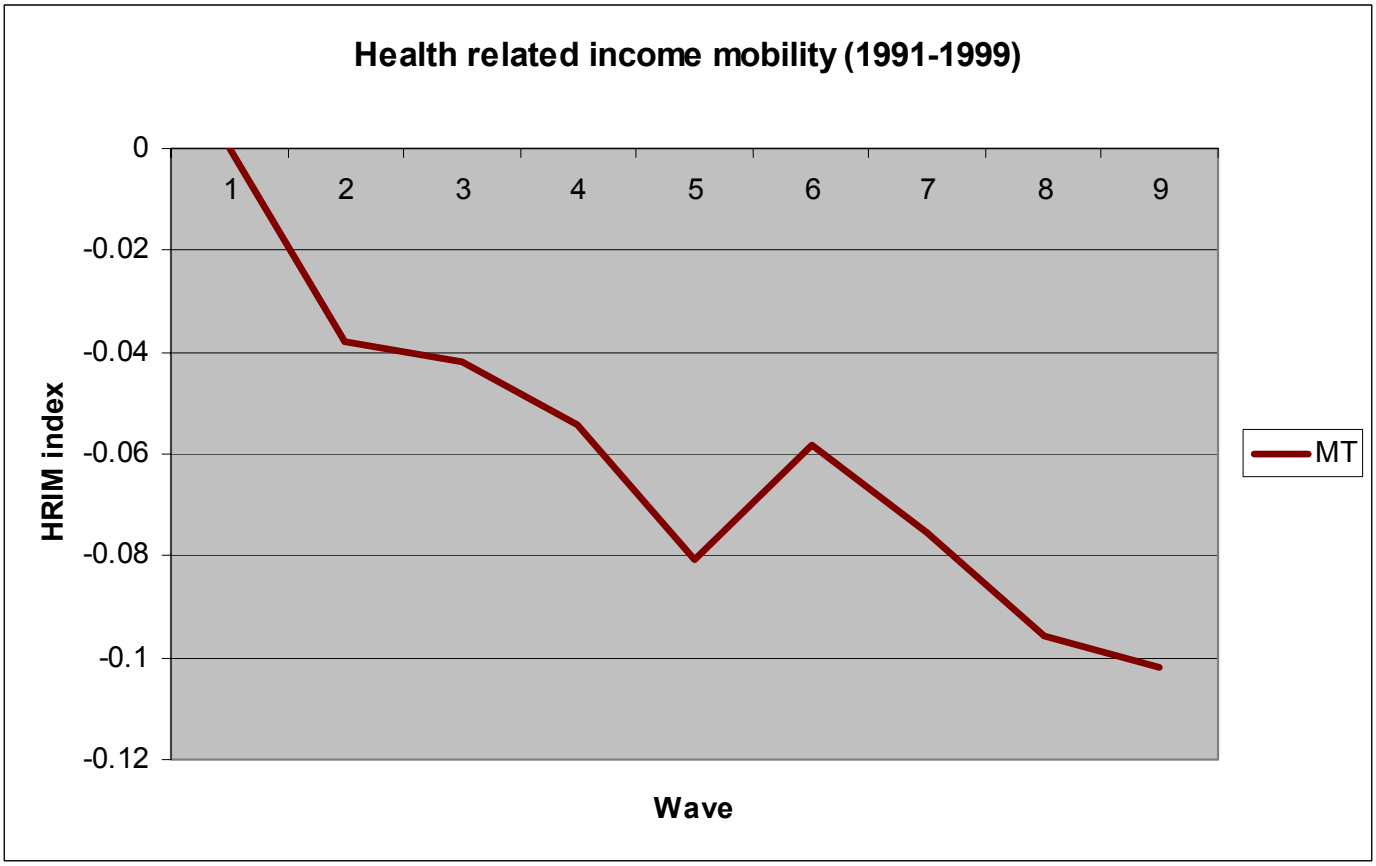
Figure 2: Index of health-related income mobility