



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Health Economics

journal homepage: [www.elsevier.com/locate/econbase](http://www.elsevier.com/locate/econbase)

Short communication

## Measurement of horizontal inequity in health care utilisation using European panel data

Teresa Bago d'Uva<sup>a,b,\*</sup>, Andrew M. Jones<sup>c</sup>, Eddy van Doorslaer<sup>a,b</sup><sup>a</sup> Department of Applied Economics, Room H13-09, Erasmus School of Economics, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands<sup>b</sup> Netspar, The Netherlands<sup>c</sup> University of York, United Kingdom

## ARTICLE INFO

## Article history:

Received 2 August 2007

Received in revised form

25 September 2008

Accepted 29 September 2008

Available online xxx

## Keywords:

Inequality

Inequity

Health care utilisation

Panel data

ECHP

## ABSTRACT

Measurement of inequity in health care delivery has focused on the extent to which health care utilisation is or is not distributed according to need, irrespective of income. Studies using cross-sectional data have proposed various ways of measuring and standardizing for need, but inevitably much of the inter-individual variation in needs remains unobserved in cross-sections. This paper exploits panel data methods to improve the measurement by including the time-invariant part of unobserved heterogeneity into the need-standardization procedure. Using latent class hurdle models for GP and specialist visits estimated on 8 annual waves of the *European Community Household Panel* we compute indices of horizontal equity that partition total income-related variation in use into a need- and a non-need related part, not only for the observed but also for the unobserved but time-invariant component. We also propose and compare a more conservative index of horizontal inequity to the conventional statistic. We find that many of the cross-country comparative results appear fairly robust to the panel data test, although the panel-based methods lead to significantly higher estimates of horizontal inequity for most countries. This confirms that better estimation and control for need often reveals more pro-rich distributions of doctor utilisation.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

An equitable system of health care delivery appears to remain a core objective in most of the OECD member states with comprehensive and universal coverage and proposed health system reforms usually quote equity preservation or improvement as an important goal (Van Doorslaer et al., 2006). Because in many countries horizontal equity is being interpreted as the principle of equal treatment for equal need, health economists have typically approached the measurement of inequity using inequality measures (Wagstaff and Van Doorslaer, 2000a). In most empirical work, horizontal inequity is measured as the degree to which utilisation is still related to income after differences in needs across the income distribution have been appropriately standardised for (Wagstaff and Van Doorslaer, 2000b). Several cross-country comparisons have adopted variants of these methods to compare across countries in the European Union (Van Doorslaer et al., 2004), in the OECD (Van Doorslaer et al., 2006) and in Asia (Lu et al., 2007).

Invariably, these comparative studies have relied on cross-sectional surveys and have adjusted for needs by comparing *actual* utilisation distributions (by income) with *need-predicted* utilisation using some regression-based standardization

\* Corresponding author. Tel.: +31 104081477; fax: +31 104089141.  
E-mail address: [bagoduva@few.eur.nl](mailto:bagoduva@few.eur.nl) (T. Bago d'Uva).

procedure. This means that adjustment can only be made for need differences that are observed in general, self-reported, health questions which are common across a large number of surveys. Typically, only a small fraction of the inter-individual variation in utilisation measures like doctor visits can be explained by these models. And while an individual's demographic and self-reported health characteristics are known to be very powerful predictors of health care utilisation, nonetheless most of the inter-individual variation remains unexplained.

This paper aims to go beyond the earlier approaches in at least three ways. First, the availability of the full eight waves of the *European Community Household Panel* (1994–2001), and the development of appropriate models for analysing panel utilisation data (Bago d'Uva, 2006; Bago d'Uva and Jones, this issue) provides an opportunity to further examine the unexplained variation in use. We use estimation results of the preferred specification to model our data in Bago d'Uva and Jones (this issue), the latent class hurdle model. This specification captures the time-invariant components of individual unobserved heterogeneity, by modelling the latent class membership probabilities as functions of individual characteristics. Not only does this explain a greater share of the variation, it also allows for the possibility to partition the contribution of individual unobserved heterogeneity into need and non-need factors. In that way, some of the explicitly modelled individual heterogeneity can be included in the computation of the inequity index in much the same way as the observed heterogeneity.

Secondly, by using a multi-year period to assess the degree to which there are any deviations between actual and needed utilisation distributions, we move from a short to a more robust long-run perspective. Jones and López Nicolás (2004) have proposed short-run and long-run measures of income-related health inequality to examine the phenomenon of health-related income mobility. Adapting their method to our analyses of inequity in health care use, we are able to adopt a long-run perspective that accounts for this phenomenon. We find that the upwardly mobile in the income distribution are more likely to use health care, especially specialist services.

Finally, we propose a new measure of horizontal inequity in health care use that differs from the standard measure in the way that the variation left unexplained by the regression models is regarded, and which we label “conservative” index. The “conventional” index of horizontal inequity is defined as a residual and labels as inequity all income-related inequality in use that is not demonstrably related to needs. That means that all the residual income-related variation, not explained by either the need or the non-need variables, is assumed to be inequitable. But some of this residual variation may in fact be due to unobservable need differences. An alternative is to treat only the income-related inequality that is demonstrably related to non-need variables as our index of inequity. The difference between the two indices depends on the degree to which the income-related variation that is not due to included need and non-need variables is pro-rich or pro-poor. Our comparison shows that the “conservative” index tends to give higher estimates.

In what follows, we first explain how we will proceed with measuring inequity using panel data, then we describe the data we have used and the results obtained for the European countries considered. In the final section we discuss what we can and cannot conclude from this study.

## 2. Methods for measurement of inequality

### 2.1. Cross-sectional and longitudinal measures of inequality

We measure income-related inequality in the utilisation of health care (GP and specialist visits) in each wave  $t$ , using the concentration index  $C_t$  (Wagstaff et al., 1991; Kakwani et al., 1997) of the number of doctor visits. The concentration index takes on a positive/negative/zero value when there is pro-rich/pro-poor/no inequality.

Until recently, research on health equity by economists was focused on measures of socioeconomic inequalities in health and health care that were designed for use with cross sectional data. Jones and López Nicolás (2004) explore what more can be gained by using the longitudinal perspective provided by panel data. This allows the analysis to take a longer term view of the outcomes of interest, which will be of interest to policy makers whose ethical concern is with inequalities in long-run health. For example, the “fair innings” perspective suggests that equity should be defined in terms of a person's lifetime experience of health (Williams and Cookson, 2000). A similar argument may apply to taking a longer view of health care utilisation that smooths short-term fluctuations in health care use. Nevertheless, application of these methods with panel data, such as the ECHP, limits the long-run perspective to the limited window of 8 years spanned by the panel.

Jones and López Nicolás (2004) using British data (and Hernández-Quevedo et al., 2006, applying the same methodology to data from 14 European countries) find that measuring longer term health inequality by simply taking a (weighted) average of inequality measured in each year, similar to considering repeated cross-sections, masks relevant features and leads to underestimation of long-run health inequality. This indicates that in a longer run perspective, inequalities could be a greater concern than a cross-section ‘snap shot’ suggests.

Similarly, long-run income-related inequality in health care utilisation can be analysed with panel data. By analogy, we define the long-run concentration index of health care utilisation,  $C^I$ , as the concentration index for the average number of visits across periods, using as ranking variable the average income across periods. Jones and López Nicolás (2004) show that, only when the income ranking remains constant over time, the long-run concentration index equals the (weighted) average of the short-run concentration indices, defined as

$$\sum_t w_t C_t, \quad \text{where } w_t = \frac{\bar{y}_t}{T\bar{y}}, \quad (1)$$

where  $\bar{y}_t$  is the average across individuals of the number of visits to a doctor in period  $t$ , and  $\bar{\bar{y}}_T$  is equal to the average of  $\bar{y}_t$  across  $t$ . This measure, obtained using repeated cross-sections, differs from  $CI_T$  to the extent that income ranks change over time and that these changes are associated with systematic differences in health care utilisation. For example, if individuals that are upwardly income mobile tend to use more health care, the level of long-run inequality given by  $CI_T$  will be larger than the weighted average of the short-run indices. Relying only on the latter, will then result in an underestimation of the measure of interest.

## 2.2. Measurement of horizontal inequity: beyond the conventional approach

Income-related inequality in health care does not imply inequity in health care. In particular, variation in the use of health care attributable to differences in morbidity may be seen as unavoidable and hence a legitimate source of inequality (see e.g., Van Doorslaer et al., 2004). The horizontal version of the egalitarian principle requires that people in equal need of care are treated equally, regardless of socioeconomic factors such as income, level of education, place of residence, race, etc.<sup>1</sup>

In the recent literature on inequity, it has become the norm to measure horizontal inequity (HI) as the difference between the concentration index of actual use and that of predicted need-expected utilisation, obtained from an econometric model for health care use (e.g. Wagstaff and Van Doorslaer, 2000b; Van Doorslaer et al., 2000; Van Doorslaer et al., 2004, hereafter DKJ). In this paper, we refer to this as the “conventional” HI index. It measures the extent to which the difference between actual utilisation of health care and the use that would be expected on the basis of need is systematically related to the income rank of individuals. Need for medical care for each individual has been measured as the predicted use from regressions on need and non-need indicators, but neutralising the impact of non-need variables by setting those equal to their means (Schokkaert and Van de Voorde, 2000; Gravelle, 2003; DKJ, 2004). One important complication in the measurement of HI is that the dependent variable in health care demand models is typically specified as a nonlinear function of the regressors. For example, DKJ (2004) estimate logistic and truncated and generalized negative binomial regression models, which are intrinsically nonlinear. As long as the model is linear, predicting need-expected use by setting the non-need variables equal to their mean (or, in fact, any constant value), achieves complete neutralisation of these variables. This does not hold for a nonlinear model and so the estimated HI is contingent on the values chosen for the non-need variables.

### 2.2.1. The “conventional” HI index

The first step in the computation of the “conventional” HI index is the prediction of need-expected utilisation based on the actual values of the  $x^N$  variables. These predictions are contingent on the level of the non-need variables ( $x^{NN}$ ) selected. By analogy with the linear case, we use sample means of the non-need variables,  $\bar{x}_t^{NN}$ . For a nonlinear functional form  $G(\cdot)$ , the need-expected level of care for individual  $i$  in period  $t$  is predicted as

$$\hat{y}_{it}^N = E [y|x_{it}^N, \bar{x}_t^{NN}] = G \left( \sum_N \hat{\beta}_N x_{it}^N + \sum_{NN} \hat{\beta}_{NN} \bar{x}_t^{NN} \right) \quad (2)$$

The “conventional” HI index is then obtained by subtracting the concentration index of  $\hat{y}_{it}^N$ , from the concentration index of  $y_{it}$ , the observed number of visits for individual  $i$  in period  $t$ :

$$HI_t^{conv} = CI_t - CI(\hat{y}_{it}^N). \quad (3)$$

We calculate short-run HI indices for each wave  $t$ ,  $HI_t^{conv}$ , and long-run measures using the average predicted number of visits across periods ( $\hat{y}_i^N$ ),  $LR^{conv} = HI_t^{conv} = CI_T - CI(\hat{y}_i^N)$ . Additionally, we compute “average” short-run HI indices,  $AvgSR^{conv}$ , as the difference between the weighted average of the short-run indices, as in Eq. (1), and what is obtained when that formula is applied to the concentration indices of predicted utilisation.  $AvgSR^{conv}$  will differ from  $LR^{conv}$  when changes in income ranks over time are associated with systematic differences in observed or need-expected health care utilisation. For example, if upwardly income mobile individuals tend to use more health care, conditional on need, then the long-run inequity will be underestimated by  $AvgSR^{conv}$ . Note that the estimated vector of coefficients  $\hat{\beta}$ , which embodies the implicit vertical equity norm of the country’s system, and so determines what an individual needs, is country-specific but not wave-specific. We therefore assume that the norm of what constitutes needed and non-needed care use is constant across the period considered.

The predictions of need-expected number of doctor visits,  $\hat{y}_{it}^N$ , result from a regression model, and therefore entail a prediction error. DKJ (2004) argue that the contribution of the residuals can be attributed to either justifiable or unjustifiable sources of inequity. In their “conventional” approach, all residual variation is classified as unjustifiable. In other words, any residual variation in health care use that is left unexplained, and that is systematically related to income rank, is assumed to be determined by non-need factors. This approach therefore adopts a rather narrow definition of need, since it considers as legitimate health care use only what is shown (by the regression) to be systematically associated with need factors. However, DKJ (2004) note that their assumption is not indisputable, as some or all of the unexplained variation may capture

<sup>1</sup> In this study, we focus solely on the issue of horizontal inequity, similar to most of the previous analyses of inequity in health care use. An exception is Sutton (2002) who studied both issues of vertical and horizontal inequity.

unmeasured need.<sup>2</sup> A natural alternative to the “conventional” HI index is one that equals the concentration index of the non-need-expected health care use only. This is a more “conservative” approach in that only the inequality that results from the observed systematic association between income rank and non-need factors is considered inequitable, whilst the residual variation is considered justifiable. So, the definition of need in this alternative method is broader than in the “conventional” method. The treatment of unexplained variation is highly relevant since count data like reported doctor visits are notably difficult to predict, especially in the tails of the distribution.

### 2.2.2. The “conservative” HI index

The computation of the “conservative” HI index requires the prediction of non-need-expected utilisation, based on the actual values of the  $x^{NN}$  variables, and on the sample means of the need variables  $x^N, \bar{x}_t^N$ :

$$\hat{y}_{it}^{NN} = E [y | \bar{x}_t^N, x_{it}^{NN}] = G \left( \sum_N \hat{\beta}_N \bar{x}_t^N + \sum_{NN} \hat{\beta}_{NN} x_{it}^{NN} \right) \quad (4)$$

The “conservative” HI index then equals the concentration index of  $\hat{y}_{it}^{NN}$ :

$$HI_t^{cons} = CI(\hat{y}_{it}^{NN}) \quad (5)$$

Again, we can define short-run HI indices for each wave,  $HI_t^{cons}$ , and long-run HI indices,  $LR^{cons} = HI_t^{cons} = CI(\hat{y}_t^{NN})$ , where  $\hat{y}_t^{NN}$  is the average of  $\hat{y}_{it}^{NN}$  across periods. We compute “average” short-run HI indices,  $AvgSR^{cons}$ , using Eq. (1) for  $\hat{y}_{it}^{NN}$  instead of actual utilisation.<sup>3</sup>

### 2.3. Need adjustment with the latent class hurdle model

We now turn to the methods for predicting  $\hat{y}_{it}^N$  and  $\hat{y}_{it}^{NN}$ . The availability of panel data makes it possible to control for unobserved individual heterogeneity when modelling the number of doctor visits. We use latent class (or finite mixture) hurdle models (LCH), as developed by Bago d'Uva (2006) and estimated on the ECHP data in Bago d'Uva and Jones (this issue). This framework offers an alternative representation of heterogeneity, where individuals are drawn from a finite number,  $C$ , of latent classes, and so individual effects are approximated using a discrete distribution, imposing no distribution on the unobserved individual effects. Within each latent class, a hurdle model is considered, with the NegBin2 as the underlying distribution for the two stages.<sup>4</sup> Class membership probabilities,  $\pi_{ij}$ , are modelled as functions of individual time-invariant characteristics,  $z_i$ , using a multinomial logit (Clark et al., 2005; Clark and Etilé, 2006; Bago d'Uva and Jones, this issue). LCH models were estimated separately for each country and for GP and specialist visits in Bago d'Uva and Jones (this issue) and shown to outperform alternative specifications. They show it fit the data substantially better than the standard hurdle model, which does not account for the panel feature of the data.

In order to calculate the “conventional” and “conservative” HI indices, we require predictions of need-expected and non-need expected number of visits. Computation of predictions from the highly nonlinear LCH models is not straightforward. While, in the estimation stage, these models control for unobserved heterogeneity, as desired, in the prediction stage, it becomes necessary to define whether the individual unobserved heterogeneity represents need, non-need, or a combination of both. As noted by Van Ourti (2004), the unobserved individual heterogeneity may reflect need factors (such as unobserved health) as well as non-need factors (such as health care preferences). The horizontal equity norm is that there is equal treatment for equal need. Therefore, the key condition that the predictions for need-based utilisation have to meet is that they vary only with need. Conversely, for the “conservative” HI index, it is required that predictions of non-need health care utilisation vary only with non-need factors. In the case of panel data models, these conditions mean that different assumptions regarding the nature of the individual unobserved heterogeneity require different procedures to predict utilisation.

We use estimates from Bago d'Uva and Jones (this issue), who specify the class membership probabilities as functions of the time-invariant individual characteristics,  $z_i$ , which, similar to  $x_{it}$ , can include both *need* and *non-need* factors. In this specification, unlike in one with constant class membership, individual unobserved heterogeneity is not restricted to be solely non-need (as in Van Ourti, 2004) or need, and can therefore be standardised for need and non-need. This feature makes the more flexible specification with variable  $\pi_{ij}(\cdot)$  preferable for predicting (non-)need expected utilisation. Furthermore, the models estimated by Bago d'Uva and Jones (this issue) do indeed show significant associations with both types of factors

<sup>2</sup> DKJ (2004) also decompose the “conventional” HI index (the contributions of different types of factors, including the residual contribution). The decomposition analysis makes the consequences of the assumptions transparent, as it makes it possible to assess whether the residuals make a pro-rich or a pro-poor contribution to HI. It should however be noted that in a nonlinear setting the decomposition requires a linear approximation which means that the residual contribution arises both from a prediction error and from a linear approximation error.

<sup>3</sup> Similar to the “conventional” HI index, this index is contingent on the values used for the need variables in the computation of  $\hat{y}_{it}^{NN}$ , which means that their effect is not completely neutralized.

<sup>4</sup> The NegBin2 derives from a Poisson distribution with gamma variance. This is the version of the NegBin most commonly used (Cameron and Trivedi, 2005), the default version in popular statistical packages such as Stata (also in the estimation of the zero-truncated model, i.e., the second part of the hurdle model) and TSP, and has been used in most applications of the latent class framework and the hurdle model (for example, Gerdtham, 1997; Jiménez-Martín et al., 2002; Gerdtham and Trivedi, 2001; DKJ, 2004).

(especially with need) in almost all cases. The unobserved individual heterogeneity that we are able to partition into need and non-need, using this econometric specification, should not be confused with the variation that is unexplained by the regression. Inevitably, this variation cannot be partitioned and so has to be assumed to capture just non-need, as in the “conventional” approach, or just need, as in the “conservative” approach.

2.3.1. *Need-expected utilisation (for “conventional” HI)*

The need-based predictions of health care use can be computed as

$$\hat{y}_{it}^N = \sum_j^C \hat{\pi}_{ij} (z_i^N, \bar{z}^{NN}) E_j [y_{it} | x_{it}^N, \bar{x}_t^{NN}], \tag{6}$$

where  $x_{it}^N$  and  $\bar{x}_t^{NN}$  are as defined above,  $\bar{z}^{NN}$  the sample averages of non-need variables characteristics that enter the class membership probabilities,  $z_i^N$  the actual values of the need variables in  $\pi_{ij}$ , and  $E_j[\cdot]$  represents the expected number of visits, conditional on belonging to class  $j$ , as assumed by the LCH. Since the class membership probabilities are computed for fixed values of the non-need variables, we are assuming that, across individuals, only the variation in  $\pi_{ij}(\cdot)$  that is related to need is legitimate. All individuals with the same need are attributed the same class membership probabilities, regardless of the value of the non-need variables. Similarly, conditional on the latent class, predictions only vary according to need. Therefore, the resulting predictions, unconditional on the latent class,  $\hat{y}_{it}^N$ , vary only with the observed need factors, in line with the horizontal equity norm.

2.3.2. *Non-need-expected utilisation (for “conservative” HI)*

The non-need-based predictions of health care use can be computed as

$$\hat{y}_{it}^{NN} = \sum_j^C \hat{\pi}_j (\bar{z}^N, z_i^{NN}) E_j [y_{it} | \bar{x}_t^N, x_{it}^{NN}], \tag{7}$$

where  $\bar{x}_t^N$  and  $\bar{z}^N$  are the sample averages of the need variables and  $x_{it}^{NN}$  and  $z_i^{NN}$  are the actual values of the non-need variables, and  $E_j[\cdot]$  is as defined above. The class membership probabilities are computed for fixed values of the need variables and therefore, all variation in  $\pi_{ij}(\cdot)$  that is determined by non-need variables is considered illegitimate. Regardless of the need variables, all the individuals that are equal in terms of non-need variables have the same class membership probabilities and the same predicted use, conditional on the latent class. Consequently, the resulting unconditional predictions,  $\hat{y}_{it}^{NN}$ , vary only with observed non-need factors.

3. Data

The data are taken from the *European Community Household Panel User Database (ECHP-UDB)*. The standardised questionnaire allows for cross-country as well as longitudinal comparisons. We use all 8 waves available for 10 EU member states: Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Netherlands, Portugal and Spain. Austria joined the survey in 1995 (wave 2) and Finland only in 1996 (wave 3).<sup>5</sup>

We analyse health care utilisation over the previous year, represented by the number of visits to a GP and the number of visits to a specialist. These data are available from wave 2 onwards. The ECHP income variable is total net household income in PPPs, and deflated by national CPIs and divided by the OECD modified equivalence scale to account for household size and composition.

In the analysis of inequity in health care, we use an identical set of covariates across countries to represent need and non-need factors. As need indicators, we use demographics and 1-year lagged health measures based on two questions: (a) responses to a question on self-assessed general health status as either very good, good, fair, bad or very bad<sup>6</sup>; and (b) responses to “Do you have any chronic physical or mental health problem, illness or disability? (yes/no)” and, if so, “Are you hampered in your daily activities by this physical or mental health problem, illness or disability? We use one dummy variable to indicate whether the individual has some limitation. Gender and age are represented by the variables: Age, Age<sup>2</sup>, a dummy variable for males (Male), Age × Male and Age<sup>2</sup> × Male.

Apart from income, the following non-need variables are included: (i) the highest level of general or higher education completed, i.e. recognised third level education (ISCED 5–7), second stage of secondary level of education (ISCED 3) or less than second stage of secondary education (ISCED 0–2); (ii) marital status, distinguishing between married and unmarried (including cohabiting); (iii) activity status includes employed, self-employed and not working. In addition, we include time

<sup>5</sup> In the United Kingdom, Luxembourg and Germany, the ECHP was held only in waves 1–3, after which it was replaced by national panel surveys. For this reason we drop these three countries. We do not include France because the French data do not contain detailed information on the number of specialist visits.

<sup>6</sup> Across countries, the proportion of individuals that responds ‘very bad’ is very small, so we collapsed the categories ‘bad and very bad’. For Portugal, the category ‘very good’ also contains a small proportion of individuals, so we collapsed the categories ‘good’ and ‘very good’ for this country.

dummies to allow for time trends in mean utilisation. The sample used is an unbalanced panel of individuals observed for up to 5 waves in the case of Finland, up to 6 waves in the case of Austria and up to 7 for the remaining countries. For the analysis of inequity we use cross-sectional individual sampling weights.

#### 4. Results

This section presents an analysis of horizontal inequity in the utilisation of primary (GP) and secondary (specialist) care for 10 EU member states. Predictions of (non-)need-expected health care utilisation were obtained as explained in Section 2.3, from LCH models estimated separately for each country and for GP and specialist visits by Bago d'Uva and Jones (this issue). We then use the methods described in Section 2.2 to compute short-run indices (by wave), and their weighted average, and long-run HI indices, for each country, according to the “conventional” and the “conservative” versions of HI. We focus on the indices based on all waves available (i.e., long-run and weighted averages of short-run indices). Standard errors for statistical inference are based on 200 replications of the index computation, each using a random draw from the sampling distribution of the parameter estimates (multivariate normal distribution with mean vector and covariance matrix equal to estimated parameter vector and estimated covariance matrix).<sup>7</sup>

##### 4.1. GP visits

Table 1 presents results of long-run (LR) and average of short-run (AvgSR) inequity in the number of GP visits for the 10 countries considered in the analysis, ranked by the values of the LR “conservative” HI index. Controlling for unequal need distributions, following the “conventional” HI approach, we can distinguish two groups of countries: for Finland, Portugal and Austria, the indices are positive (except for a non-significant negative LR index for Austria); all other countries display pro-poor inequity in the use of GPs. For Finland, all SR indices (on average, 0.028) as well as the LR index (0.028) are larger than for any other country, showing the largest pro-rich inequity. Averaged across waves, short-run pro-poor inequity is largest for Belgium (−0.053), followed by Ireland (−0.045) and Spain (−0.039). In the long-run, pro-poor inequity is larger for Ireland and Belgium (−0.055, −0.054) and Spain (−0.038). Across countries, except for Finland, there are significant differences between the AvgSR and the LR measures of inequity. The LR indices are larger than the AvgSR ones for 5 of the 10 countries but the differences are generally small. Across countries, the results obtained here for the 1996 wave of the ECHP (not shown) are of the same sign as the ones obtained in DKJ (2004) for the same wave. For 6 of the 9 countries in common in both studies, the 1996 inequity indices obtained here are larger (more pro-rich or less pro-poor) than the ones in the previous paper (and not within the confidence intervals of our 1996 indices).

Table 1 also shows the AvgSR and the LR HI indices for the new “conservative” approach. Again, Finland displays the highest levels of pro-rich inequity in GP visits (AvgSR, 0.033; LR, 0.035), followed by Portugal (not significantly smaller) and Austria (significantly smaller than Finland and Portugal). The remaining countries exhibit pro-poor inequity, and this is greater for Spain, Ireland, Belgium and Italy and less pronounced, but still significant, for Greece, Denmark and The Netherlands. Except for Greece and Italy, measured inequity increases when the long-run perspective is adopted, but the discrepancies are very small and only significant for Portugal and Finland.

The “conservative” and “conventional” HI indices have the same index signs and therefore provide the same answers to the question of whether there is pro-rich or pro-poor inequity in GPs visits in the countries studied here. Similarly, the ranking of countries according to HI is generally robust to the approach chosen. An exception is Italy that belongs to the group of countries with smaller pro-poor “conventional” inequity, while it is among the ones with largest pro-poor “conservative” inequity. However, the “conservative” approach generally results in larger HI indices than the “conventional” approach. For Greece, this is only true for the LR indices; for Spain, the LR and AvgSR HI indices are equal, regardless of the approach; and for Italy the “conventional” approach gives higher AvgSR and LR indices. This suggests that the differences between the “conventional” HI indices, that assume that the unexplained variation is all non-need, and the “conservative” HI indices, that only identify as non-need the variation that is demonstrably associated with non-need factors, are mostly due to pro-poor contributions. This means that the assumptions regarding the need/non-need nature of the residuals are not entirely innocuous. In particular, if the residuals were pro-poor and due to need rather than non-need factors, horizontal inequity in GP visits would be underestimated by the “conventional” HI method for most of the countries studied in this paper, although the differences are only significant in a minority of cases.

##### 4.2. Specialist visits

Horizontal inequity indices in the number of specialist visits according to both methods are shown in Table 2. We first analyse the results given by the “conventional” approach and compare them with the ones obtained by DKJ (2004). The LR and AvgSR HI indices are positive and significant, indicating pro-rich inequity in specialist visits, for all countries. The highest levels of pro-rich inequity are observed for Portugal (LR, 0.204; AvgSR, 0.199) and Finland (LR, 0.134; AvgSR, 0.143). Pro-rich

<sup>7</sup> We thank an anonymous referee for suggesting this procedure, as an alternative to bootstrapping the entire process, including model estimation, which would have been prohibitive given the computation time required.

**Table 1**  
Average of short-run (AvgSR) and long-run (LR) inequity for number of GP visits.

	"Conventional" HI						"Conservative" HI						"Conservative" HI – "Conventional" HI			
	AvgSR		LR		LR-AvgSR		AvgSR		LR		LR-AvgSR		Avg SR		LR	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Spain	-0.039*	(0.001)	-0.038*	(0.001)	0.001*	(0.0001)	-0.039*	(0.004)	-0.038*	(0.004)	0.001	(0.001)	0.000	(0.004)	0.000	(0.004)
Ireland	-0.045*	(0.002)	-0.054*	(0.002)	-0.010*	(0.0003)	-0.036*	(0.007)	-0.035*	(0.008)	0.001	(0.002)	0.009	(0.006)	0.019*	(0.008)
Italy	-0.027*	(0.001)	-0.024*	(0.001)	0.003*	(0.0001)	-0.031*	(0.003)	-0.033*	(0.003)	-0.002*	(0.001)	-0.004	(0.003)	-0.009*	(0.003)
Belgium	-0.053*	(0.002)	-0.055*	(0.002)	-0.002*	(0.0003)	-0.035*	(0.005)	-0.033*	(0.006)	0.002	(0.001)	0.018*	(0.004)	0.022*	(0.005)
Greece	-0.016*	(0.002)	-0.025*	(0.002)	-0.009*	(0.0002)	-0.019*	(0.005)	-0.021*	(0.005)	-0.002 <sup>†</sup>	(0.001)	-0.003	(0.005)	0.004	(0.005)
Denmark	-0.020*	(0.004)	-0.016*	(0.004)	0.004*	(0.0004)	-0.016*	(0.007)	-0.014 <sup>†</sup>	(0.008)	0.002	(0.002)	0.004	(0.006)	0.002	(0.007)
Netherlands	-0.028*	(0.001)	-0.022*	(0.001)	0.006*	(0.0002)	-0.015*	(0.004)	-0.013*	(0.005)	0.002	(0.001)	0.013*	(0.004)	0.009 <sup>†</sup>	(0.005)
Austria	0.009*	(0.002)	-0.001	(0.002)	-0.010*	(0.0001)	0.010 <sup>†</sup>	(0.005)	0.011 <sup>†</sup>	(0.006)	0.001	(0.001)	0.001	(0.005)	0.012*	(0.006)
Portugal	0.018*	(0.001)	0.019*	(0.001)	0.001*	(0.0002)	0.022*	(0.004)	0.024*	(0.005)	0.002*	(0.001)	0.004	(0.004)	0.004	(0.004)
Finland	0.028*	(0.003)	0.028*	(0.003)	0.000	(0.0004)	0.033*	(0.006)	0.035*	(0.006)	0.002*	(0.001)	0.005	(0.005)	0.008	(0.006)

\* Significant at 5%, assuming normality.

<sup>†</sup> Significant at 10%, assuming normality.

**Table 2**  
Average of short-run (AvgSR) and long-run (LR) inequity for number of specialist visits.

	"Conventional" HI						"Conservative" HI						"Conservative" HI – "Conventional" HI			
	AvgSR		LR		LR-AvgSR		AvgSR		LR		LR-AvgSR		AvgSR		LR	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Netherlands	0.026*	(0.002)	0.050*	(0.002)	0.024*	(0.0003)	0.037*	(0.007)	0.047*	(0.008)	0.010*	(0.001)	0.011†	(0.007)	–0.003	(0.008)
Belgium	0.034*	(0.003)	0.040*	(0.003)	0.007*	(0.0004)	0.052*	(0.007)	0.066*	(0.008)	0.014*	(0.002)	0.018*	(0.006)	0.025*	(0.007)
Denmark	0.052*	(0.008)	0.088*	(0.008)	0.035*	(0.0007)	0.057*	(0.013)	0.078*	(0.014)	0.021*	(0.004)	0.005	(0.011)	–0.009	(0.012)
Austria	0.090*	(0.003)	0.089*	(0.003)	–0.002*	(0.0002)	0.082*	(0.006)	0.088*	(0.006)	0.007*	(0.001)	–0.009†	(0.005)	0.000	(0.006)
Spain	0.083*	(0.001)	0.089*	(0.001)	0.005*	(0.0001)	0.092*	(0.005)	0.105*	(0.006)	0.013*	(0.001)	0.009†	(0.005)	0.017*	(0.006)
Greece	0.073*	(0.002)	0.070*	(0.002)	–0.003*	(0.0003)	0.100*	(0.007)	0.106*	(0.007)	0.006*	(0.001)	0.027*	(0.006)	0.036*	(0.007)
Italy	0.096*	(0.001)	0.094*	(0.001)	–0.001*	(0.0001)	0.103*	(0.004)	0.114*	(0.005)	0.012*	(0.001)	0.007†	(0.004)	0.020*	(0.005)
Finland	0.134*	(0.004)	0.143*	(0.004)	0.009*	(0.0007)	0.142*	(0.008)	0.152*	(0.009)	0.009*	(0.001)	0.009	(0.008)	0.009	(0.008)
Ireland	0.123*	(0.004)	0.128*	(0.004)	0.006*	(0.0007)	0.142*	(0.012)	0.157*	(0.013)	0.015*	(0.002)	0.019†	(0.010)	0.028*	(0.011)
Portugal	0.199*	(0.002)	0.204*	(0.003)	0.005*	(0.0005)	0.180*	(0.006)	0.195*	(0.006)	0.015*	(0.001)	–0.019*	(0.006)	–0.009	(0.006)

\* Significant at 5%, assuming normality.

† Significant at 10%, assuming normality.

inequity is smallest in Belgium (LR, 0.034; AvgSR, 0.040) and the Netherlands (LR, 0.026; AvgSR, 0.050). Comparing the LR and AvgSR indices in Table 2, we see that, for 7 of the 10 countries considered, LR pro-rich inequity is understated by the AvgSR measure (for the exceptions – Austria, Greece and Italy – the LR indices are only slightly smaller than the AvgSR ones). In all cases, the estimated differences are statistically significant at 5%.

Like DKJ (2004), we find pro-rich inequity for all countries. Also, apart from Finland, not included in the previous study, the countries with the highest pro-rich inequity in both studies are Portugal and Ireland, while inequity is lowest for Belgium and the Netherlands. However, for 7 of the 9 overlapping countries, the results obtained here for 1996 using the “conventional” method (not shown) are larger than the ones obtained in the previous paper (and, except for Ireland, these smaller indices are not within the confidence intervals of our 1996 indices). This suggests that the panel-based model used here to predict need-expected use is better capable to account for need (observed and unobserved) than cross-section-based models. Except for Ireland, the LR “conventional” HI indices are also larger than the ones obtained for 1996 in DKJ (2004), and significantly so, except for Denmark. On the whole, the longer run picture provided in this paper indicates significantly greater levels of horizontal inequity than the previous evidence.

The “conservative” HI index results, also presented in Table 2, confirm that across waves (not shown here) and in the LR, there is pro-rich inequity in specialist visits. The largest inequity indices are obtained for Portugal (AvgSR, 0.180; LR, 0.195), followed by Ireland (AvgSR, 0.142; LR, 0.157) and Finland (AvgSR, 0.142; LR, 0.152). The lowest indices are for The Netherlands (AvgSR, 0.037; SR, 0.046), Belgium (AvgSR, 0.052; LR, 0.066) and Denmark (AvgSR, 0.057; LR, 0.078). For all countries, the LR measure of inequity is significantly larger than the weighted average of SR indices: the AvgSR indices underestimate LR inequity by a proportion ranging from 6% (Finland and Greece) to 27% (Denmark). This suggests that, not only richer individuals tend to use more specialist care in the short-run (controlling for need) but also, that individuals moving up in the income distribution over time tend to visit a specialist more often than those moving in the opposite direction. This aspect of inequity in the use of specialists cannot be identified with cross-sections, or when the panel feature of the data is not accounted for.

Both approaches tell us that there is pro-rich inequity in all countries studied. However, the – supposedly – more “conservative” method does not produce lower but higher LR HI indices for 6 of the 10 countries (in 5 significantly so) and also higher average SR indices, in all but two cases (significant at 10% in 7 cases). The two methods do result in rather similar rankings of countries according to LR HI. Exceptions are Belgium, Ireland and Greece that move up the ranking most when going from the LR “conventional” HI to the LR “conservative” HI method.

The question that arises is why this is the case. Recall that the two methods differ in their treatment of the unexplained variation. In the “conventional” method, this variation is assumed to be non-need, while in the new “conservative” method, it is assumed to reflect need. The fact that the “conservative” HI indices are, on the whole, larger than the “conventional” ones indicates that the differences between the two represent pro-poor contributions to inequity. If it is the case that the residuals are pro-poor and that they capture mainly justifiable variation in the use of specialists, then the “conventional” approach underestimates inequity in specialist use in most countries, as that fraction of health care use that is attributed to non-need makes a pro-poor contribution to horizontal inequity.<sup>8</sup>

## 5. Conclusion

Achieving equitable access to health care for all citizens, irrespective of their incomes, remains an important public health policy goal in Europe’s largely publicly funded health care systems. Key to the monitoring of the extent to which various systems are successful in attaining this goal is the appropriate and reliable measurement of the degree of inequity in the distribution of health care. Over the last two decades, Europe has invested heavily in the collection of comparable data to enable proper analysis and comparison of EU member countries’ relative performance. The completion of the *European Community Household Panel* surveys provides an opportunity to adopt a longitudinal perspective in this comparison.

This paper has explored the possibilities to exploit the European panel data in order to obtain more reliable estimates of inequity than what a single cross-section can provide. We believe it makes three contributions. First, we account for time-invariant individual unobserved heterogeneity – and partition also this heterogeneity (as captured by class membership probabilities) into need and not need related – by using latent class models for the prediction of health care use. We find that, in almost all cases, this extension leads to higher (i.e. less pro-poor or more pro-rich) index values. This suggests that much of the variation associated with income that remains unexplained in cross-sectional models not accounting for heterogeneity derives from unobserved need heterogeneity.

Secondly, we find that, for almost all countries, long-run indices are higher than the average of short-run indices, especially in the case of specialist visits. This suggests that upward income mobility contributes to more pro-rich (or less pro-poor) inequity while downward mobility does the opposite. In other words, given the same needs, doctors are not only consulted more often by those with higher incomes, but also by those with faster growing incomes. This is of interest, as it suggests a hitherto unappreciated role of income mobility in the generation of patterns of income-related inequity in use.

<sup>8</sup> The results of decomposition analysis in DKJ (2004) lend some support to the argument that a pro-poor residual contribution is mainly due to unobserved need. In that paper, observed need variables always have a pro-poor contribution to inequity in specialist visits, while the contribution of income and other non-need variables is mostly pro-rich.

Thirdly, we have proposed a new approach to the measurement of inequity that is not inherently related to the availability of panel data. While we have labeled this a more “conservative” approach than the conventional one, it nonetheless tends to give higher estimates of inequity. This is a consequence of the fact that the difference between the two approaches can be either pro-poor or pro-rich distributed, and we find it generally to be more pro-poor.

The rankings of countries by level of long-run inequity in the use of primary and specialised care are generally in accordance with those obtained for 1996 in DKJ (2004). The general result of pro-poor inequity in GP visits in most (7 out of 10) EU countries and pro-rich inequity in specialists in all countries studied does not change with the longitudinal perspective employed here or with the new “conservative” measurement of horizontal inequity. Similarly, the country rankings are fairly robust to variations in the approach. All in all, our re-assessment based on the full panel largely extends and strengthens the earlier cross-sectional findings. It appears then, that the earlier conclusion of people in equal need not all being treated equally everywhere in Europe also holds after a tougher test. The results do suggest, however, that better control for need, adoption of a longer run perspective and even using an arguably more conservative index all lead to more pro-rich inequity for specialist care and less pro-poor inequity in GP care than measured in previous work.

## Acknowledgements

The *European Community Household Panel Users' Database*, version of December 2003, was supplied by Eurostat. This paper derives from the project “The dynamics of income, health and inequality over the lifecycle” (known as ECuity III Project), which is funded in part by the European Commission's Quality of Life and Management of Living Resources programme (contract QLK6-CT-2002-02297). Teresa Bago d'Uva was funded by Fundação para a Ciência e Tecnologia, under PhD grant SFRH/BD/10551/2002. Teresa Bago d'Uva and Eddy van Doorslaer also acknowledge funding from the Netspar project: “Health, income and work across the life cycle”. We are grateful to Geert Dhaene, Tom Van Ourti, participants at the Health Economics Workshop in Leuven, 2006, and two anonymous referees for helpful comments.

## References

- Bago d'Uva, T., 2006. Latent class models for health care utilisation. *Health Economics* 15, 329–343.
- Bago d'Uva, T., Jones, A.M. Health care utilisation in Europe: new evidence from the ECHP. *Journal of Health Economics* (this issue), doi:10.1016/j.jhealeco.2008.11.002.
- Cameron, A.C., Trivedi, P.K., 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.
- Clark, A., Etilé, F., 2006. Don't give up on me baby: spousal correlation in smoking behaviour. *Journal of Health Economics* 25, 958–978.
- Clark, A., Etilé, F., Postel-Vinay, F., Senik, C., Van der Straeten, K., 2005. Heterogeneity in reported well-being: evidence from twelve European countries. *The Economic Journal* 115, C118–C132.
- Gerdtham, U.G., 1997. Equity in health care utilisation: further tests based on hurdle models and swedish micro data. *Health Economics* 6, 303–319.
- Gerdtham, U.G., Trivedi, P.K., 2001. Equity in swedish health care reconsidered: new results based on the finite mixture model. *Health Economics* 10, 565–572.
- Gravelle, H., 2003. Measuring income related inequality in health: standardization and the partial concentration index. *Health Economics* 12, 803–819.
- Hernández-Quevedo, C., Jones, A.M., López-Nicolás, Á., Rice, N., 2006. Socioeconomic inequalities in health: a comparative longitudinal analysis using the European Community Household Panel. *Social Science and Medicine* 63, 1246–1261.
- Jones, A.M., López Nicolás, A., 2004. Measurement and explanation of socioeconomic inequality in health with longitudinal data. *Health Economics* 13, 1015–1030.
- Jiménez-Martín, S., Labeaga, J.M., Martínez-Granado, M., 2002. Latent class versus two-part models in the demand for physician services across the European Union. *Health Economics* 11, 301–321.
- Kakwani, N.C., Wagstaff, A., Van Doorslaer, E., 1997. Socioeconomic inequalities in health: measurement, computation and statistical inference. *Journal of Econometrics* 77 (1), 87–104.
- Lu, J.F., Leung, G.M., Kwon, S., Tin, K.Y., Van Doorslaer, E., O'Donnell, O., 2007. Horizontal equity in health care utilisation—evidence from three high-income Asian economies. *Social Science and Medicine* 64 (1), 199–212.
- Schokkaert, E., Van de Voorde, C., 2000. Risk selection and the specification of the risk adjustment formula. Centre for Economic Studies Discussion Paper, 2000/011. University of Leuven, Leuven.
- Sutton, M., 2002. Vertical and horizontal aspects of socio-economic inequity in general practitioner contacts in Scotland. *Health Economics* 11, 537–549.
- Van Doorslaer, E., et al., 2000. Equity in the delivery of health care in Europe and the US. *Journal of Health Economics* 19 (5), 553–583.
- Van Doorslaer, E., Koolman, X., Jones, A.M., 2004. Explaining income-related inequalities in doctor utilisation in Europe. *Health Economics* 13, 609–628.
- Van Doorslaer, E., Masseria, C., Koolman, X., the OECD Health Equity Group, 2006. Inequalities in access to medical care by income in developed countries. *Canadian Medical Association Journal* 174, 177–183.
- Van Ourti, T., 2004. Measuring horizontal inequity in Belgian health care using a Gaussian random effects two part count data model. *Health Economics* 13, 705–724.
- Wagstaff, A., Van Doorslaer, E., 2000a. Equity in health care finance and delivery. In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*. North-Holland, pp. 1803–1862.
- Wagstaff, A., Van Doorslaer, E., 2000b. Measuring and testing for inequity in the delivery of health care. *Journal of Human Resources* 35 (4), 716–733.
- Wagstaff, A., Paci, P., Van Doorslaer, E., 1991. On the measurement of inequalities in health. *Social Science and Medicine* 33, 545–557.
- Williams, A., Cookson, R., 2000. Equity in health. In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*, vol. 1B, Chapter 35. Elsevier Science BV, Amsterdam, pp. 1863–1910.