Network for Studies on Pensions, Aging and Retirement

Philip Clarke

Tom van Ourti

Correcting the Bias in the Concentration Index When Income is Grouped

Discussion Paper 01/2009 - 004

January 20, 2009

Correcting the bias in the concentration index when income is grouped

Philip Clarke[a,*] and Tom Van Ourti[b,c]

[a] *School of Public Health, The University of Sydney, New South Wales 2006, Australia*

[b] *Erasmus School of Economics, Erasmus University Rotterdam, PB 1738, 3000 DR Rotterdam, The Netherlands*

[c] *Tinbergen Institute, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands*

20 January 2009

Abstract

The problem introduced by grouping income data when measuring socioeconomic inequalities in health (and health care) has been highlighted in a recent study in this journal. We re-examine this issue and show there is a tendency to underestimate the concentration index at an increasing rate when lowering the number of income categories. This bias results from a form of measurement error and we propose two correction methods. Firstly, the use of instrumental variables (IV) can reduce the error within income categories. Secondly, through a simple formula for correction that is based only on the number of groups. We compare the performance of these methods using data from 15 European countries and the United States. We find that the simple correction formula reduces the impact of grouping and always outperforms the IV approach. Use of this correction can substantially improve comparisons of the concentration index both across countries and across time.

KEY WORDS: concentration index, errors-in-variables, instrumental variables, categorical data, first-order correction

JEL Classification: C2, D31, I19

Acknowledgements

We are grateful for comments received from Teresa Bago d'Uva, Hans van Kippersluis, and participants at seminars given at Australian National University, Tilburg University and Erasmus University Rotterdam. We also acknowledge funding from the NETSPAR project "Income, health and work across the life cycle" and thank EUROSTAT for access to the ECHP. Tom Van Ourti is a Postdoctoral Fellow of the Netherlands Organisation for Scientific Research – Innovational Research Incentives Scheme – Veni. Philip Clarke is supported by a Sydney University Fellowship. Part of this work was undertaken during a stay at the Melbourne Institute of Applied Economic and Social Research, and Economics RSSS at the Australian National University, the hospitality of which is gratefully acknowledged. The usual caveats apply and all remaining errors are our responsibility.

[*] Corresponding author. Tel: +61-2-93515424; fax: Fax: +61 2 9351 7420.

E-mail addresses: philipc@health.usyd.edu.au (Philip Clarke), vanourti@few.eur.nl (Tom Van Ourti).

1 Introduction

The concentration index has become the standard measure to quantify income-related inequalities in health economics (Wagstaff and van Doorslaer, 2000; O'Donnell *et al.*, 2008). It can be applied to grouped/aggregated data (Wagstaff *et al.*, 1989; Kakwani *et al.*, 1997; Wagstaff, 2002) but has mainly been applied to micro data sets that contain information on an individual's income and his/her health (care) status (van Doorslaer *et al.*, 1992; van Doorslaer *et al.*, 1997; van Doorslaer *et al.*, 2000; van Doorslaer and Koolman, 2004; van Doorslaer *et al.*, 2004; Bago d'Uva *et al.*, 2008a). Micro datasets are generally preferred to grouped datasets as the former result in consistent estimation of the concentration index, as point estimates from grouped datasets ignore information on within group association between income (rank) and health (care) status (Kakwani *et al.*, 1997). In addition, micro datasets are needed to apply regression based decomposition techniques for the concentration index (Wagstaff *et al.*, 2003).

Despite these advantages, estimating the concentration index using individual level data is not always preferable as more than one individual may report the same level of income. This can arise for several reasons. First, different individuals may receive the same income if they are for example, both on the same pay scale, or both receiving the same state pension, also if household income is used, all persons within a household will be assigned the same income. Second, when answering a survey, a respondent may round off his/her reported income instead of reporting an exact amount or more generally income might be misreported (Moore *et al.*, 2000). Finally, income data is sometimes only collected or reported in a limited number of categories, often because of confidentiality reasons.

While it is technically straightforward to compute the concentration index based on individual level data, this relies on being able to uniquely rank all individuals in the population by income which is not possible when they report the same level of income. A straightforward solution is to aggregate the micro data to a grouped data set – with the income levels defining the groups – and next to apply the grouped data estimator for the concentration index (Kakwani *et al.*, 1997; O'Donnell *et al.*, 2008). So in the first case (outlined above) if each individual's reported income level equals his/her actual income level, one cannot improve upon the (point) estimate of the grouped data estimator. In the two other cases of *rounding off* and having *a limited number of income categories*, reported income levels cannot equal *unobserved* actual income levels; and the (point) estimate of the concentration index would improve if actual income levels were observed and used to construct the fractional income rank.

There are many such examples of inequality studies that have involved surveys where the income variable used to rank individuals is reported in categories. These include – among others – Gerdtham *et al.* (1999) who use Swedish health survey data with an income measure with six categories; van Doorslaer *et al.* (2000) use Finnish and Danish data with categorical income data; Wagstaff (2002) and Meheus and van Doorslaer (2008) use aggregate data in wealth quintiles; Humphries and van Doorslaer (2000) and Wagstaff and van Doorslaer (2004) use Canadian data with income deciles; and van Doorslaer *et al.* (2006) use Canadian and Australian data with a limited number of income categories.

The issue of dealing with income grouping when measuring the concentration index with micro level data has also been highlighted in a recent study in this journal by Chen and Roy (2009). However, the focus of this study is confined to calculating potential bounds on the concentration index and the implications of existing estimators for efficiency of statistical inference. To date the broader question of the consequences (and solutions) of grouping the income variable over ranges in terms of bias in the estimated inequality measure has not been addressed.

In this paper our main focus is on the third case where income is measured categorically since it allows us to abstract from misclassification error between the income categories, while allowing for any other type of misreporting within the income categories. We show that grouping the data creates a form of the classical errors-in-variables problem (for example, Cameron and Trivedi, 2006) in which an individual's ranking is measured with error within, but not between groups. Initially, we illustrate the degree to which this form of measurement error affects the point estimate of the concentration index. While it is known that estimation on individual level data does not ignore within group association between income (rank) and health (care) status, there is hardly any information on the magnitude and direction of the bias resulting from using grouped data. The extent of this bias and ways to overcome this problem have been extensively explored for the Gini coefficient in the context of income inequality measurement (e.g. see Gastwirth, 1972; Rasche, 1980; Lerman and Yitzhaki, 1989; Sarabia *et al.*, 1999; Van Ourti and Clarke, 2008). However, findings from this literature can only be extrapolated to the concentration index if concentration curves are globally convex or concave, which is unlikely to hold. We then propose and apply two procedures to reduce the degree of bias in the grouped data estimator. The first we term the *IV approach* which involves finding an instrumental variable to reduce the error in ordering individuals within each of the income categories. The second, which we refer to as the *overall correction approach*, was put forward by Van Ourti and Clarke (2008) in the context of the Gini coefficient. They derived a correction factor to reduce the degree of bias in the grouped data estimator for the Gini, but this approach could also be applied to the concentration index as it uses only information on the number of income groups. We show using data from 15 European countries and the US that it not only performs well in reducing the degree of bias in the concentration index across several health and health care indicators, but also that it outperforms the *IV approach* in our empirical application. This arises because it is difficult to find a suitable instrument from variables typically collected in health surveys.

The remainder of the paper is organised as follows. In the next section, we discuss estimators for the concentration index in case of individual and grouped/categorical income data. The third section describes the data which comes from the European Community Household Panel (ECHP) and the Medical Expenditure Panel Survey (MEPS). Using these datasets, we then illustrate the impact of income grouping upon the point estimate of the concentration index. In the fifth section, we present our IV approach for reducing the bias alongside the overall correction approach. In the sixth section we compare both approaches and show that the overall correction approach outperforms the IV approach in our empirical application. Next, we use the ECHP and ten waves of the MEPS to illustrate the performance of the preferred overall correction approach in correcting income-related inequality comparisons across countries and over time. The final section concludes and discusses the wider relevance and applicability of our correction methods.

2 Background

The concentration index is defined as twice the area between the concentration curve (which represents the cumulative proportion of the population ranked by income, starting with the lowest income against cumulative proportions of health/health care) and the diagonal. The bounds of this measure are –1 and +1 with a negative (positive) value representing pro-poor (pro-rich) inequality.[1] Kakwani *et al.* (1997) have shown that the concentration index can be calculated using a so-called convenient OLS-regression.

Let us start with the situation where we have micro data at hand and every individual $i = 1,\ldots,n$ reports a variable of interest m_i such as health (care) and his/her actual income level y_i with $y_c \le y_d$ for $c < d$. It is easy to show that the concentration index of m_i – $C(m_i)$ – equals $\hat{\alpha}_1$ in the underneath 'convenient' OLS-regression:

$$2\sigma_R^2 \frac{m_i}{\bar{m}} = \alpha_0 + \alpha_1 R_i^y + \varepsilon_i \tag{1}$$

where $\bar{m}$ is the average of m_i, R_i^y is the fractional rank of y_i, $\sigma_R^2 = n^{-1}\sum_{i=1}^{n}\left(R_i^y - 0.5\right)^2$ $= \left(12n^2\right)^{-1}\left(n^2 - 1\right)$ is the variance of R_i^y, ε_i is an error term with mean zero, and α_0, α_1 are parameters to be estimated.[2] It is common practice to use the fractional rank as proposed by Lerman and Yitzhaki (1989), i.e. $R_i^y = n^{-1}\left(i - 0.5\right)$, but in order to allow for individuals having the same actual income level, it is better to use:

$$R_i^y = \frac{p(y_i) + 0.5\left[q(y_i) - p(y_i)\right]}{n} \tag{2}$$

where $q(y_i) = \sum_{k=1}^{n} 1(y_k \le y_i)$ and $p(y_i) = \sum_{k=1}^{n} 1(y_k < y_i)$ equal the proportion of individuals having at least income y_i, including and excluding y_i respectively (Van Ourti, 2004; Chen and Roy, 2009).

In case one uses micro data, but income is only recorded in a limited number of categories, equation (1) still applies, but the calculation of the fractional income rank differs: there are K different income categories r_i such that $r_i = j$ if $\psi_{j-1} < y_i \le \psi_j$ with $j = 1,\ldots,K$ and ψ_{j-1} and ψ_j are the bounds of each income category j. The fractional income rank then becomes:

[1] Erreygers (2008) has shown that for any variable of interest with a finite upper value or a strictly positive lower value, the bounds of the concentration index need not be -1 and +1. All the results in this paper should also apply to his corrected concentration index, to Wagstaff's (2005) normalized concentration index, and to the generalized concentration index (Wagstaff *et al.*, 1991) since mean health and its lower and upper bound are not affected by income grouping.

[2] We note that one should use the population formula of the variance of the fractional rank (and not a small-sample adjustment) since OLS is used as an arithmetic (and not as a statistical) device.

$$R_i^y = R_j^r = \frac{q(\psi_{j-1}) + 0.5\left[q(\psi_j) - q(\psi_{j-1})\right]}{n} = \frac{\sum_{k=1}^{j-1} n_k + 0.5(n_j - n_{j-1})}{n} \quad (3)$$

where n_j is the number of individuals in income category j.

Finally, in case of grouped data, we can still apply equation (3), but must replace equation (1) by equation (4). The grouped data estimator for $C(m_j;K)$ now equals $\hat{\beta}_1$:

$$2\sigma_{R^K}^2 \frac{m_j}{\bar{m}} \sqrt{n_j} = \beta_0 \sqrt{n_j} + \beta_1 R_j^r \sqrt{n_j} + \zeta_j \quad (4)$$

where m_j is the average variable of interest within income category j, $\sigma_{R^K}^2 = n^{-1} \sum_{j=1}^{K} n_j \left(R_j^r - 0.5\right)^2$ is the variance of R_j^r, ζ_j is an error term with mean zero, and β_0, β_1 are parameters to be estimated.[3]

The first goal of this paper is to show how income grouping impacts upon the point estimate of the concentration index. Figure 1 provides some intuition on the effect of grouping. The solid lines show a hypothetical Lorenz (panel a) and concentration curve (panel b). Consider the effect of grouping the population into tertiles:[4] both the Lorenz and concentration curve are now composed of straight dotted lines as the within group variation in income or health (care) no longer contributes to the respective curve (Lambert, 2001). In case of the Lorenz curve, grouping of the income variable always leads to an underestimation as the straight dotted lines that approximate the Lorenz curve are bound to lie inside the original curve (see panel a in Figure 1 in which the difference between the grouped and original curves is shaded grey). Hence, grouping of the income variable leads to a downward bias of the Gini coefficient. Lerman and Yitzhaki (1989) have shown that the magnitude of the bias is non-negligible for relatively small numbers of income categories. Van Ourti and Clarke (2008) show that the bias increases at an increasing pace when lowering the number of income categories. However, with the concentration curve this bias need not be downward, see for example panel b in figure 1. Here the inflections in the concentration curve mean that the straight line lies both outside and inside the original concentration curve and so the concentration index based on income grouping will be greater (or lesser) depending on

[3] We note that the point estimate of the grouped data estimator equals that of the individual level estimator with a limited number of income categories, but the variances will differ since the grouped data estimator neglects the variability of the variable of interest within the income categories.

[4] For simplicity, we assume that everyone within the first and third tertile has the same value for the variable of interest, i.e. income for the Gini and health (care) for the concentration index. There is only variation within the second tertile.

the degree to which it compresses inequalities above or below the dotted line (error also indicated in grey). In general, there is no a priori guidance on the sign or magnitude of this bias on the concentration index since the only prerequisite is that the concentration curve should be increasing in income rank (for example, the concentration curve unlike the Lorenz curve can lie above the diagonal). The only corollary is that sign and magnitude of the bias are bound to depend on the marginal distribution of the variable of interest (conditional on income rank). As far as we are aware, the magnitude of the bias has not been analysed empirically, except for Kakwani *et al.* (1997) who showed that using income deciles instead of individual level data resulted in a small downward estimate (around 1 percent) of a concentration index of a binary ill-health indicator in the 1980 and 1981 Dutch Health Interview surveys. The magnitude of the bias for other numbers of income groups, other health (care) indicators, and other countries is not known.

[Figure 1 about here]

3 Data

To explore how categorical income data impacts on the point estimate of the concentration index, we use one wave of data from 15 countries participating in the European Community Household Panel (ECHP) and the waves conducted between 1996 and 2005 of the *Medical Expenditure Panel Survey* (MEPS) from the United States.

The ECHP was designed and coordinated by EUROSTAT (2003), and consists of a representative panel of non-institutionalised households providing data on socioeconomic, demographic and health characteristics of individuals aged 16 or older. We use the second (1995) wave (which was the first to include all health-related information) for 13 EU member states: Austria (AT), Belgium (BE), Denmark (DK), France (FR), Germany (DE), Greece (GR), Ireland (IRL), Italy (IT), Luxembourg (LU), Netherlands (NL), Portugal (PT), Spain (ES) and the United Kingdom (UK). In the case of Finland (FI) and Sweden (SE), we use later waves since these countries joined the ECHP in wave 3 (1996) and wave 4 (1997) respectively.

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health status and health care use for the U.S. civilian non-institutionalized population. We used micro-data files for 10 annual waves of the MEPS survey conducted between 1996 and 2005 which was made publically available through the Agency for Healthcare Research and Quality (2008).

[Table 1 about here]

Summary statistics of the variables in the ECHP are given in table 1; and table 2 provides summary statistics of the variables in the MEPS. The key variables for this study are household income, self-reported health and health care use which are defined as follows:

Income: The ECHP collects information on disposable (i.e. after-tax) household income, which is all net monetary income received by the household members during the *previous* year. It includes income from work (employment and self-employment), private income (from investments and property and private transfers to the household), pensions and other direct social transfers received. The definition of income in MEPS was similar to that of the ECHP. We measure all incomes in national currencies. For both datasets, the income variable was further divided by the OECD modified equivalence scale in order to account for household size and composition (EQINC).[5]

Self-reported health (SRH): In the ECHP it was measured as the response to an ordered 5-point scale (ranging from very good to very poor) on the question "How is your health in general?" In MEPS the 5-point scale ranged from excellent to poor. While reporting heterogeneity in SRH between populations and cultures has been raised in previous studies (for example, Lindeboom and van Doorslaer, 2004; Bago d'Uva *et al.*, 2008b) it should not concern us here as this paper is about illustrating the bias resulting from categorical income data. In addition, many studies in the existing income-related health inequality literature transform SRH responses on a common cardinal scale in order to undertake international comparisons of SRH. We follow this approach by transforming the ordinal SRH responses onto the cardinal HUI scale (Feeny *et al.*, 1995) which takes values between 0 (death) and 1 (perfect health). This involves attaching mean HUI scores for each SRH category in the 1994 Canadian National Population Health Survey (available in table 1 in Jones and van Doorslaer, 2003) to the SRH categories in the ECHP and MEPS.

Health care use variables: Both the ECHP and the MEPS contain information on health care utilization including: (i) the *number of nights in hospital* during the last

[5] The OECD modified equivalence scale gives a weight of 1.0 to the first adult, 0.5 to the second and each subsequent person aged 15 and over, and 0.3 to each child aged under 15.

12 months (NIGHTS), (ii) the *number of dental visits during the last 12 months* (DENT) and (iii) the *number of visits to a GP or specialist during the last 12 months* (PHYS).

[Table 2 about here]

4 Empirical difference between the individual level and grouped data estimators of the concentration index

Figure 2 gives an overview of the impact of income grouping upon estimates of concentration indices of self-reported health, hospital nights, dental visits, and physician visits. We summarize the impact of grouping individual level microdata into equally sized groups, i.e. 'income-tiles', by calculating $100\left\{\left[C\left(m_j,K\right)/C\left(m_i\right)\right]-1\right\}$ for 50 to 2 'income-tiles', $K=50,49,\ldots,3,2$.[6] We do this separately for each of the 15 European ECHP countries and the 2000 wave of the MEPS (United States), and arrive at three main findings.[7]

First, we observe both downward (<0) and upward (>0) impacts of income grouping indicating that the underlying concentration curves are neither globally convex nor globally concave. Second and unsurprisingly, we find a much larger and 'random' impact of income grouping across K in those cases where the concentration index based on individual level microdata was insignificant (i.e. $p>0.1$). This is the case for PHYS in the MEPS 2000 data (highlighted in red) and Sweden (highlighted in green) as well as NIGHTS and SRH in France (highlighted in blue) and DENT for Germany.[8] This follows from our 'relative' indicator of the impact of income grouping that gets inflated if the denominator is very small. Third, the pattern of the impact of income grouping across the number of income categories K shows a similar shape for different variables and countries, despite the large differences in the concentration indices across health (care) variables and countries (see rows (C) in table 1 and 2). There seems to be a

[6] The impact of income grouping will *in general* be similar for income groupings of unequal size, but one cannot preclude that it deviates for very peculiar shapes of the marginal distribution of m_i over R_i^y. We revisit the case of unequal income groups in more detail in section 5.

[7] One might wonder why we have not resorted to Monte Carlo simulations. We follow the reasoning of Van Ourti and Clarke (2008) that parametric distribution functions are of limited value in estimating Lorenz curves precisely (a.o. Schader and Schmid, 1994); and hence are unlikely to deliver robust conclusions on the *empirically relevant* bias from income grouping on the Gini index. In case of the concentration index, this reasoning is even more compelling as it concerns a bivariate distribution.

[8] We observe the significance level of the individual level concentration indices since we create the problem of income grouping. This is not possible for the applied researcher using grouped data since she/he only observes the significance level of the concentration indices resulting from grouped data.

common concave shape revealing that income grouping 'on average', and at least for the variables studied here, underestimates the concentration index at an increasing rate when lowering the number of income categories. This similarity does not arise for an 'absolute' indicator of the impact of income grouping; and justifies our 'relative' measure as it suggests that the shape of the underlying marginal distribution of the variable of interest (conditional on income rank) is similar across countries, except for the spread. The impact of income grouping seems only empirically relevant – i.e. dominating the randomness across countries – for 10 or less income groups, and markedly so for 5 or less groups. In the extreme case of 2 income groups, the median degree of downward underestimation across all countries is 26% for SRH, 30% for NIGHTS, 27% for DENT and 20% for PHYS.

[Figure 2 about here]

5 Correcting the impact of income grouping

In order to improve upon estimates based on grouped income data, it has been common practice in the literature on income inequality measurement to fit parametric Lorenz curves to grouped data and to estimate the Gini coefficient from the parameters (among others, Rasche *et al.*, 1980; Villaseñor and Arnold, 1989; Sarabia *et al.*, 1999). Unfortunately, this approach requires the property of the Lorenz curve that it is globally convex – which may not always be the case for concentration curves as we illustrated in figure 1b – and was criticized by Schader and Schmid (1994) to be unreliable. Alternatively, income inequality researchers have derived non-parametric bounds for the Gini index (among others Gastwirth, 1972; Ogwang, 2003). Chen and Roy (2009) have extended this approach to the concentration index. Since this approach is based on the lowest and highest potential within-income-group-correlations between health (care) and income, the bounds can be very wide so this approach suffers from empirical uncertainty. Finally, as mentioned by Chen and Roy (2009), Gundgaard (2005, 2006) and Gundgaard and Lauridsen (2006a-b) propose random sorting within income groups to address the impact of income grouping. This approach is however inappropriate since, compared to the estimator in equation (1)-(3), a purely random ordering within income categories will only increase the sampling variability of the concentration index.[9]

We take a fourth route by re-interpreting the problem as a classical errors-in-variables problem (for example Wooldridge, 2003, section 4.4.2) since income grouping is equivalent to error in the ranking variable (R_i^y in equation 1) within (but not between) groups. We propose an IV-approach to remove the impact of income grouping upon estimates of the concentration index. As it uses an instrument to create income-variation within the income categories, it cannot be used with pure grouped data, as the instrument must vary within the group. Our second approach applies the procedure of Van Ourti and Clarke (2008) which was originally derived to deal with the impact of income grouping upon the Gini index. Contrary to the IV approach it can be applied to both grouped data and micro data with a limited number of income categories.

5.1 IV regression as a tool to reduce the impact of income grouping

Thanks to the convenient regression approach of Kakwani *et al.* (1997), it is straightforward to apply 2SLS to equation (1) where the fractional income rank is defined as in (3). This approach has three advantages. First, conditional upon finding a suitable instrument, it will always reduce the impact of income grouping upon the point estimate of the concentration index. Second, 2SLS, like OLS, provides standard errors for the 'income-grouping-corrected' concentration index that can be used for statistical inference. Third, 2SLS allows using more than one instrument to create income-variation within income categories. Hence, overidentifying restriction tests, such as the J-statistic of Hansen (1982), can be used to assess the appropriateness of instruments.

As we are using 2SLS, the normal conditions for a good instrument must apply: (i) sufficient correlation between the instrument and R_i^y, and (ii) no correlation between the instrument and ε_i in equation 1. When employing the standard IV approach to address errors-in-variables the first condition can be tested, but the second must be maintained. However, despite the technical similarity, our approach differs from the standard IV approach as there is no measurement error at the level of the income

[9] Without sufficient replications of this random ordering, the point estimate will also be different.

categories, but only within the categories.[10] It follows that neither condition can be observed within income groups, so we suggest the following procedure for deciding on a suitable instrument.

First, the instrument(s) should be defined in such a way that if we were to use individual level income as an instrument (i.e. the unobserved actual income value), our 2SLS would give the same point estimate as the individual level estimator in equation (1)–(2). This can be achieved by making sure that the instrument(s) preserves the ranking across income categories (i.e. all people in a higher income category continue to be assigned a greater rank than individuals in any lower income category), and by ranking the individuals within the income categories by 'another variable' that is correlated with the income rank. Note that if this variable only takes a limited number of values (e.g. years of education), one should apply equation (3) to this 'other variable' to calculate the ranking within the income categories.

Second, the sign of the correlation between the rank of this 'other variable' and the fractional income rank based on the income categories is used to decide whether we rank the individuals by this 'other variable' in an increasing or decreasing manner within the income categories. So for example if the fractional income rank based on income categories is positively correlated with years of education the individuals within each category will be re-ranked in order of increasing years of education. It follows that our approach basically imposes that the correlation within the income categories is similar to that between the income categories and so it is likely to perform well only if the within income categories correlations between the fractional rank of unobserved actual income and the fractional rank of the 'other variable' have the same sign and magnitude as that of the between correlations. Finally, we note that – due to its reliance upon 2SLS – our IV approach can easily deal with income groupings of unequal size.

5.2 Overall correction approach to reduce the impact of income grouping

Similarly to the IV approach, the overall correction approach is derived from studying the estimator of the concentration index within a measurement error framework. The main difference with the IV approach is that it is a first-order-correction approach based on the number (and relative size) of income groups only. While this increases the

[10] In addition, the measurement error is not correlated with the rank of the variable measured with error, which holds since the distribution of the measurement error is similar within each income category.

likelihood of some remaining second order bias it is simple to implement and as Van Ourti and Clarke (2008) show that it has good empirical performance when applied to the Gini index. It is precisely this reliance upon the number (and relative size) of income groups that makes it a potential candidate for reducing the impact of income grouping upon the concentration index. The intuition of their approach – applied to the concentration index – consists of comparing the estimator in equation (1) with the one based on a grouping of the data in equation (4), and by next exploiting the properties of the fractional rank and those of OLS as an *arithmetic* tool. For clarity, we first present the procedure of Van Ourti and Clarke (2008) for income groupings of equal size, i.e. $n_1 = n_2 = \ldots = n_K = n/K$, and next generalize for income groups of unequal size.

Let us start from the observation that $C\left(m_j;K\right)$ differs from $C\left(m_i\right)$ if the fractional income rank R_i^y is associated with m_i within the income groups (see also Figure 1b). The same insight emerges from the difference between the LHS and RHS of equations (1) and (4).[11] The RHS difference is addressed by defining an equation that describes the measurement error

$$R_i^j = R_i^y + \delta_i^j \tag{5}$$

where R_i^j is the fractional rank of group j – $K^{-1}\left(j-0.5\right)$ – assigned to each individual i and δ_i^j is the measurement error with zero mean[12]. The LHS difference is addressed by multiplying through with the ratio of the variances of the fractional income ranks at the individual and grouped data level, i.e.

$$\frac{\sigma_R^2}{\sigma_{R^K}^2} = \frac{K^2\left(n^2-1\right)}{n^2\left(K^2-1\right)} \tag{6}$$

After some algebra (consult Van Ourti and Clarke (2008) for more details), one gets an equation that expresses the concentration index estimated from individual level data as a function of the concentration index estimated from equally-sized groupings of these data, i.e.

---

[11] Strictly speaking, one cannot compare equations (1) and (4) as these are based on respectively n and K observations. Nevertheless it is easy to derive an equation that gives the same OLS point estimate for β_1 as (4), but that is defined on n observations. With the assumption of income groups of equal size, equation (4) reduces to $\left[K^{-2}\left(K^2-1\right)\right]\left[\left(6\overline{m}\right)^{-1} m_j\right] = \beta_0 + \beta_1\left[K^{-1}\left(j-0.5\right)\right] + \zeta_j$. If we next use individual level data and assign the fractional rank of group j to each individual, i.e. $R_i^j = K^{-1}\left(j-0.5\right)$, one gets $\left[K^{-2}\left(K^2-1\right)\right]\left[\left(6\overline{m}\right)^{-1} m_i\right] = \beta_0 + \beta_1 R_i^j + \psi_i$, which is comparable to equation (1).

[12] As explained before, the measurement error δ_i^j is not correlated with R_i^j.

$$C(m_i)=\frac{K^2}{K^2-1}\left[\frac{n^2-1}{n^2}C(m_j;K)-\frac{12}{n}\sum_{i=1}^{n}\delta_i^j\varepsilon_i\right] \tag{7}$$

Assuming $n\to+\infty; K<+\infty$ (i.e. the number and relative size of groups is fixed in the population), equation (7) reduces to $C(m_i)=\left(K^2-1\right)^{-1}K^2\left[C\left(m_j;K\right)-12\operatorname{cov}\left(\delta_i^j,\varepsilon_i\right)\right]$, and thus provides a first-order-correction term $\left(K^2-1\right)^{-1}K^2$ and an expression for the remaining second-order bias $-12\operatorname{cov}\left(\delta_i^j,\varepsilon_i\right)K^2\left(K^2-1\right)^{-1}$.

Equation (6) shows that the first-order-correction can be interpreted as a "grouped data" adjustment of the variance of the fractional rank.[13] The remaining second order bias is a function of the covariance between the measurement error and the error from equation (1); and the smaller its value, the better the overall-correction-approach/first-order-correction performs in empirical applications. Although the exact value and sign of this covariance cannot be known as ε_i is unobservable, we can make some approximate statements. First, the remaining second order bias will be zero if m_i is uniformly distributed over R_i^y (i.e. a uniform marginal distribution) within income categories since the variance of ε_i equals zero in this case. Second, this covariance will be smaller the higher the number of groups K.[14] Finally, the covariance is *likely* to be negative for an asymmetric unimodal distribution (i.e. left or right skewed). For example, an extreme long right tail is likely to result in a negative covariance as can be seen from equation (1) and (5).

The generalisation to groups of unequal size – equation (4) – is straightforward:

$$C(m_i)=\frac{1}{\sigma_{R^K}^2}\left[\frac{n^2-1}{12n^2}C(m_j;K)-\frac{1}{n}\sum_{i=1}^{n}\delta_i^j\varepsilon_i\right] \tag{8}$$

Equation (8) and (7) are similar (for $n\to+\infty; K<+\infty$ equation (8) reduces to $C(m_i)=\left(12\sigma_{R^K}^2\right)^{-1}\left[C\left(m_j;K\right)-12\operatorname{cov}\left(\delta_i^j,\varepsilon_i\right)\right]$), except that it is impossible in equation (8) to come up with an exact expression for $\sigma_{R^K}^2$. Nevertheless, the first-order-correction remains easy to calculate, and the interpretation of the first- and second-order terms remain unchanged.

[13] Note that the first order correction also equals $\left[12\operatorname{cov}\left(R_i^j,R_i\right)\right]^{-1}$ which is intuitive as a high covariance between the grouped and individual fractional rank implies a low first-order correction term.

[14] While the covariance decreases with K, the effect upon the remaining second order bias depends on its reduction relative to $K^2\left(K^2-1\right)^{-1}$.

6 Empirical applications

In this section we provide four empirical applications of the performance of these correction methods in removing the impact associated with the grouping of the income variable. In the first section we compare the IV and overall correction approaches using MEPS 2000 and ECHP data and find the overall correction approach outperforming the IV approach. We then test the performance of the overall correction approach in removing the impact of income grouping upon inequality rankings across countries using the ECHP, and across time using MEPS data.

6.1 Comparative performance of the IV and Overall Correction Approach

While the IV approach has the potential to completely remove the impact of grouping of the income variable, in practice there are typically a limited number of potential candidates for constructing instruments from variables routinely collected in health (care) surveys. One way of examining the performance of this approach is to use a set of commonly collected variables to construct instruments for the MEPS and ECHP data sets and then see to what degree we are able to reduce the impact of income grouping. We use the same waves of MEPS and ECHP used to quantify the magnitude of the impact of income grouping reported in section 4.

In the MEPS and ECHP data sets we identified three variables for constructing instruments. First, the modified OECD equivalence scale (*eqscale*) is by definition correlated with equivalent income; and hence with the fractional income rank. Second, we have information on the number of years of education completed in MEPS, and on the highest educational degree obtained in the ECHP (*educ*). Given the overwhelming evidence on the returns to education (Card, 1999), it should correlate with the fractional income rank. Third, we have the age of each individual (*age*) which should correlate with the fractional income rank given the evidence on the life-cycle behaviour of incomes (e.g. King and Dicks-Mireaux, 1982). Since most datasets on which the concentration index has been applied are restricted to the adult population, we have excluded all individuals younger than 16 from our analyses.

We defined instruments from these variables along the lines set out in section 5.1, i.e. we construct a fractional rank that preservers the ranking across income categories and that ranks the individuals within income categories by these variables.

The signs of the correlations between the fractional rank of these variables and the fractional income rank based on the income categories are used to determine whether the individuals within income categories are ranked increasingly or decreasingly by these variables. While these correlations are always statistically significant (at the 1 percent level) for the MEPS data[15] and are significant (again 1 percent level) in almost all ECHP countries[16] and hence there is potential for using these instruments (most likely for *educ*), their ultimate performance will depend on whether the two conditions for a good instrument apply, i.e. whether the ranks of *eqscale*, *educ* and *age* correlate with the fractional income rank *conditional* on the income categories; and whether they do not correlate with the error term of equation (1), also *conditional* on the income categories. It seems likely that the first (second) condition is more likely to hold the lower (higher) the number of income groups. Unfortunately, as we explained in section 5.1, neither condition can be tested by researchers applying our approach to data with income recorded in categories. Despite this unavoidable uncertainty, the instruments have been used to examine the performance of the IV approach relative to the overall correction approach in reducing the impact of grouping. As we have previously shown the concentration index has low statistical significance (i.e. $p > 0.1$) in five cases (see figure 2). We have not applied the correction methods to these cases as there would seem to be little benefit from correcting an index with such wide random variations across incomes groups.

[Figure 3 about here]

Figure 3 shows the degree to which each of the three instruments and the overall correction is able to reduce the impact of grouping, defined as $B_j = 100\left\{\left[C\left(m_j, K\right) / CI\left(m_i\right)\right] - 1\right\}$. Figure 3a–3c reports these statistics for the United States using the MEPS from the year 2000 for SRH, NIGHTS and DENT[17] and 3d-3g provides a graphical summary of the performance across ECHP countries using the median degree of error by income group. The figures show that the impact of the IV

---

[15] For respectively 50, 10, 5 and 2 income categories, the correlation coefficients are (-0,131;-0,131;-0,129;-0,118) for *eqscale* (0,471;0,469;0,462;0,403) for *educ*, and (0,043;0,042;0,040;0,037) for *age*.

[16] The correlation coefficients for all instruments were statistically significant in all ECHP countries, except for Finland and Greece for the one based on *eqscale*; and for France, Germany, Luxembourg and Spain for the one based on age. For respectively 50, 10, 5 and 2 income categories, the correlation coefficients ranged between (-0,185↔0,191;-0,185↔0,194;-0,180↔0,188;-0,139↔0,187) for *eqscale*; between (0,207↔0,392;0,206↔0,393;0,204↔0,395;0.182↔0.349) for *educ*; and between (-0.172↔0.140;-0.174↔0.136;-0.177↔0.128;-0.163↔0.078).

[17] Results for PHYS are not shown as *C* was not significant (see section 4).

approach varies considerably by both instrument and the variable of interest. For example, IV(age) results in a downward bias when applied to NIGHTS in MEPS, but an upward bias in the majority of ECHP countries. While in some cases (i.e. Figure 3c) the IV approach appears to remove a considerable proportion of the bias, the benefits of the IV correction are not universal – in Figure 3d, IV(eqscale) and IV(age) increase the bias relatively the original grouped *C*. In contrast the overall correction approach appears to always reduce the bias relative to the original index when there are low numbers of income groups.

We also provide additional summary measures of the performance of each of the 4 corrections by calculating the mean squared error (MSE) across all 49 groups, i.e. $\sum_{j=2}^{50} B_j^2 / 49$, and over ranges (2-5,6-15,16-50). To illustrate this for a single country we report these MSE's for SRH, NIGHTS and DENT in the MEPS 2000 in Table 3. The left most column (under the heading *C*) reports the MSE resulting from grouping the data (i.e. corresponding to 'original *C*' in figure 3). The remaining columns report the MSE associated with each of the correction methods. Based on comparisons of MSE of each correction method, it is clear that IV(educ) always reduces the error compared with the original grouped *C* when there are less than five groups, while the overall correction approach has a lower MSE across the entire range of groups.

[Table 3 about here]

To further examine the performance of these correction methods across countries in the MEPS and ECHP Table 4 reports the number of countries where the MSE for the 'corrected' *C* is lower than that of the original (biased) estimate. So for example, using the IV approach with *educ* to correct the *C* of dental visits, results in a lower MSE across all groups in 6 of the 15 countries, while the overall correction approach produces a lower MSE in 13 countries.[18] Use of other instruments (i.e. IV(eqscale)) tends to results in a higher MSE and for several of the comparisons there is no improvement across any of the countries.

Three conclusions emerge in each of these analyses. First, for 5 or less income groups, the overall correction considerably improves upon the original bias and removes a major share of the impact of grouping. In several countries, it also improves matters for higher numbers of income groups. For example, figure 3 and table 3 show that it

[18] Note that the number of countries included in the analysis differs for each variable since the number of insignificant *C*'s differs across health (care indicators) and since data is lacking for some countries (see also table 1).

always improves upon the original bias in the MEPS. Second, also for the IV approach using *educ* to construct the instrument, we find improvements upon the original bias for a sufficiently small number of income groups, especially for DENT (see table 4). However, it worsens matters for a higher numbers of income groups for all three health (care) variables. The other IV approaches often only worsen matters compared to the "original *C*", for example both IV(*eqscale*) and IV(*age*) always underperform compared to the original *C* for SRH and NIGHTS. Finally, the overall correction approach always outperforms IV(*educ*) in this empirical illustration.

While it should be acknowledged that only three variables to construct instruments have been tested, most (health) surveys contain relatively few variables that could be used with the IV approach. More generally, the poorer performance of the IV approach reveals that the impact of grouping can only be reduced through the judicious use of instruments and that using poor instruments can lead to an increase rather a reduction of the impact of grouping.[19] The related issue of instrument validity across countries and time is another potential drawback of the IV approach. Figure 2 illustrated the similarity of the impact of grouping across 15 ECHP countries and the US. Consequently, one would hope that the IV approach reduces the impact of income grouping in a fairly consistent way across these 16 countries, i.e. that instruments are valid across countries. Unfortunately, we found that instrument invalidity matters a great deal. For example, while the instrument constructed from *educ* reduced the impact of grouping in the US for less than 5 income groups, it was always worsening matters in at least half of the ECHP countries. Further given the simplicity and general applicability of the overall correction approach, it is likely to be the superior method for reducing the impact of grouping in most practical applications where income is recorded in groups. To a large extent this must also follow from the fact that health (care) variables tend to be fairly uniformly distributed across R_i^y within income groups.

[Table 4 about here]

[19] We hinted in section 5.1 that using several instruments jointly is an advantage of the IV approach. Combining the three instruments (constructed from *educ*, *eqscale*, and *age*) in an IV regression is infeasible due to too high collinearity between these instruments. Alternatively, one can replace the single instrument by K variables; each variable equalling the instrument for one income category and taking zero for all other income categories. The intuitive reasoning behind this set of instruments is that it allows using overidentifying restrictions tests, such as the J-statistic of Hansen (1982). Sensitivity analyses for the MEPS show that this approach works in all cases where the validity of the instruments is confirmed by the J-statistic, but that the added flexibility comes at the cost of a smaller reduction of the impact of income grouping.

6.2 *Performance of the Overall Correction Approach across countries and time*

While figure 3 and tables 3-4 show that the overall correction method reduces a substantial part of the impact of grouping, we believe it is worthwhile to present two case studies to determine the degree to which it can reduce the effect of income groupings on comparisons across countries (involving 15 ECHP countries) and over time (involving 10 years of data for the United States). We have analysed how the income-related health (care) inequality ranking is affected if one were to use the *C* based on grouped income data for one country (or one time period) while all other countries (or time periods) are based on individual level data. We prefer this over case studies where the concentration indices of all countries are based on a different number of income groupings since our design leads to a more conservative assessment of the overall correction approach.

[Table 5 about here]

Table 5 presents the results of our first case study involving 15 ECHP countries. We calculate for each country the change in the income-related inequality ranking from grouping the data for the country under study (and using the full sample *C* for the other countries). The column '*C*' shows the sum across all 15 countries of all these changes. For example, we observe 8 changes in the income-related SRH inequality for 3 income groups. The column 'OCA' shows the sum across all 15 countries of the remaining changes in income-related inequality ranking after applying our overall correction approach: cells in dark grey imply that there are more changes in the country ranking relative to *C* after applying the overall correction approach, while light grey cells imply that one comes closer to the country ranking based on individual level data. Table 5 reveals that changes in the country ranking of income-related inequalities occur frequently for less than 5 income groups; and that the overall correction approach generally manages to reduce the impact of income grouping. For more income groupings, there are hardly any changes in the country ranking (except for nights); and consequently the overall correction approach is of limited value.

[Table 6 about here]

Since the *C*'s might differ less within than between countries, we did a second case study involving only one country, but instead exploiting the time dimension. We used 10 waves of MEPS and report the findings of this analysis in table 6 which has the same setup as table 5. Compared to the cross-country case study, we find somewhat more changes in the ranking of income-related inequalities (this time across time), but only so for relatively low numbers of income groups. The overall correction approach seems even more promising here as it always improves matters (except once for dental care), which reflects the lower heterogeneity in the C's based on individual level data (see rows "C" in table 2).

7 Concluding remarks and discussion

This paper discusses and illustrates the bias in the point estimate of the concentration index resulting from categorical income data. Despite the prevailing use of grouped income data and contrary to the literature on income inequality, little is known on the signs and magnitude of this bias. In addition, the issue is conceptually different since the underlying concentration curves need not be convex; and thus the bias can also be upward. We exploit the MEPS and ECHP data that have individual level data on health (care) indicators and income to illustrate the impact of grouping by constructing hypothetical income groups. More specifically, we compare the individual and grouped data estimator for 50 to 2 equally-sized income categories. We find upward biases in some cases, but the overall tendency is to underestimate the concentration index at an increasing rate when lowering the number of income categories. Grouping reveals similar patterns for different health (care) variables and seems empirically relevant between 2 and 10 income groups implying that the impact of censoring (such as top-coding) is generally unimportant for the value of the concentration index. We also find that it can have substantial effects on income-related inequality rankings within and between countries, in particular when using a small number of income groups.

We have proposed a measurement error framework to reduce the impact of income grouping upon the point estimate of the concentration index. An IV approach is easy to apply since concentration indices can be calculated from a so-called 'convenient regression' and involves finding an instrumental variable to reduce the error in ordering individuals within each of the income categories. We have also put forward the correction factor derived by Van Ourti and Clarke (2008) and termed here the 'overall correction approach' as it only uses information on the number (and relative size) of income groups. Their correction factor was derived for reducing the impact of income grouping upon the Gini index, but will work equally well for the concentration index if health (care) is uniformly distributed *within* (not between) income categories. In addition, it can be applied to both grouped data and micro data with categorical incomes, while the IV approach can only be used for micro data with income recorded in categories. We compare both approaches using the MEPS and ECHP data.

We find that the IV approach manages to reduce the impact of income grouping in some cases, but that it makes matters worse otherwise. This follows from the usual IV problem of being unable to test whether the conditions for a good instrument apply. Instead, the overall correction approach reduces the impact of income grouping to a large extent and outperforms the IV approaches in our empirical application. We confirm the usefulness of the overall correction approach by illustrating its performance in removing the impact of income grouping upon inequality rankings across countries using the ECHP, and across time using MEPS data. Overall, we conclude that the overall correction approach is likely to be the superior method for reducing the impact of grouping in most practical applications where income is recorded in groups. Our empirical analyses suggests that it is likely to perform well in cross-country analyses for 5 or less income groups, and that it is likely to perform better for higher numbers of income groups for analyses across time.

Although this paper deals with a specific issue in the field of income-related health inequality measurement, we believe the approach has wider applicability. First, concentration indices have a long history outside health economics for analysing distributional issues in taxation (see for example, Lambert, 2001), and our correction methods may help reduce the impact of grouping in such analyses. Second, there are several examples in the health economics literature where concentration indices are calculated with categorical non-income data as the ranking variable, for example, Burström *et al.* (2005) and Clarke *et al.* (2002) used occupational classes as the ranking variable. If this categorical ranking variable derives from grouping a 'continuous' variable, the findings in this paper apply. However, if the ranking variable is genuinely categorical (i.e. there is no underlying 'continuous' (latent) variable), the concentration index is no longer biased. Nevertheless, the value of the point estimate will be influenced by the number of groups and might affect inequality comparisons if the ranking variable displays different number of categories across countries/time. Our correction methods may increase the understanding as it suggests that all else being equal socio-economic ranking variables that involve the classification of individual into a small number of groups (e.g. occupational classes or levels of educational attainment) will tend to produce estimates of inequality that are lower than using continuous measures such as equivalent income. Third, we have assumed absence of misclassification across income groups and so both the IV and overall correction approach are only intended to deal with biases from income grouping. In doing so, we have abstracted from the case where individuals might be classified into the wrong income group based on their misreported income. van Praag *et al.* (1983) have shown that misclassification bias and bias due to income groupings might be offsetting each other for other inequality measures. Future research should analyse the relative importance of both biases for the concentration index, but our results nevertheless show that the bias from income groupings can be considerable. Finally, as far as we are aware the econometrics literature on categorical variables where the errors are only within the categories is very limited. Manski and Tamer (2002) have also examined the problems arising from the use of categorical variables in regression analysis and propose methods for dealing with these type of data, which could be explored in future work.

References

Agency for Healthcare Research and Quality. *Medical Expenditure Panel Survey*. Agency for Healthcare Research and Quality, Rockville, MD, 2008. http://www.ahrq.gov/info/customer.htm

Bago d'Uva T, Jones AM, van Doorslaer E. Measurement of horizontal inequity in health care utilization using European panel data. *Tinbergen Institute discussion paper series* 2007-059/3, Erasmus University Rotterdam.

Bago d'Uva T, Jones AM, van Doorslaer E. Measurement of horizontal inequity in health care utilization using European health data. *Journal of health economics* 2008a; doi:10.1016/j.jhealeco.2008.09.008.

Bago d'Uva T, van Doorslaer E, Lindeboom M, O'Donnell M. Does reporting heterogeneity bias the measurement of health disparities? *Health Economics* 2008b; 17(3): 351-375.

Burström K, Johannesson M, Diderichsen F. Increasing socio-economic inequalities in life expectancy and QALYs in Sweden 1980-1997. *Health Economics* 2005; 14: 831-850.

Card D. The causal effect of education on earnings. In: Ashenfelter O, Card D (eds), *Handbook of labor economics-volume 3*. North Holland: Amsterdam, 1999; 1801-1863.

Chen Z, Roy K. Calculating the concentration index with repetitive values of indicators of economic welfare. *Journal of Health Economics* 2009; 28(1) 169-175.

Clarke PM, Gerdtham UG, Johannesson M, Bingefors K, Smith L. On the measurement of relative and absolute income-related health inequality. *Social Science and Medicine* 2002; 55: 1923–1928.

Cameron AC, Trivedi PK. *Microeconometrics: methods and applications*. Cambridge University Press: Cambridge, 2005.

Erreygers G. Correcting the Concentration Index. *Journal of health economics* 2008; doi:10.1016/j.jhealeco.2008.02.003.

EUROSTAT, *ECHP UDB Description of variables: Data dictionary, codebook and differences between countries and waves*. European Commission: Luxembourg, 2003.

Feeny D, Furlong W, Boyle M, Torrance G. Multi-attribute health status classification systems: Health Utilities Index. *Pharmacoeconomics* 1995; 7: 490-502.

Gastwirth J. Robust Estimation of the Lorenz Curve and Gini Index. *Review of economics and statistics* 1972; 54: 306-316.

Gerdtham U-G, Johannesson M, Lundberg L, Isacson D. A note on validating Wagstaff and van Doorslaer's health measure in the analysis of inequalities in health. *Journal of Health Economics* 1999; 18: 117-124.

Gundgaard J. Income related inequality in prescription drugs in Denmark. *Pharmacoepidemiology and Drug Safety* 2005; 14(5): 307-317.

Gundgaard J. Income-related inequality in utilization of health services in Denmark: evidence from Funen County. *Scandinavian Journal of Public Health* 2006; 34(5): 462-471.

Gundgaard J, Lauridsen J. Decomposition of sources of income-related health inequality applied on SF-36 summary scores: a Danish health survey. *Health and Quality of Life Outcomes* 2006a; 4, 53.

Gundgaard J, Lauridsen J. A decomposition of income-related health inequality applied to EQ-5D. *European Journal of Health Economics* 2006b; 7(4): 231-237.

Hansen L. Large sample properties of generalized method of moments estimators. *Econometrica* 1982; 50: 1029-1054.

Hernández-Quevedo C, Jones AM, López-Nicolás A, Rice N. Socioeconomic inequalities in health: A comparative longitudinal analysis using the European Community Household Panel. *Social Science and Medicine* 2006; 63: 1246–1261.

Humphries KH, van Doorslaer E. Income-related health inequality in Canada. *Social Science and Medicine* 2000; 50: 663-671.

Jones AM, van Doorslaer E. Inequalities in self-reported health: validation of a new approach to measurement. *Journal of health economics* 2003; 22: 61-87.

Kakwani N, Wagstaff A, van Doorslaer E. Socioeconomic inequalities in health: measurement, computation and statistical inference. *Journal of econometrics* 1997; 77: 87-103.

King MA, Dicks-Mireaux LDL. Assets holdings and the life-cycle. *The Economic Journal* 1982; 92: 247-267.

Lambert PJ. *The distribution and redistribution of income: third edition*. Manchester University Press: Manchester, 2001.

Lerman RI, Yitzhaki S. Improving the accuracy of estimates of Gini coefficients. *Journal of econometrics* 1989; 42: 43-47.

Lindeboom M, van Doorslaer E. Cut-point shift and index shift in self-reported health. *Journal of Health Economics* 2004; 23: 1083–1099.

Meheus F, van Doorslaer E. Achieving better measles immunization in developing countries: does higher coverage imply lower inequality? *Social Science and Medicine* 2008; 66(8): 1709-1718.

Moore JC, Stinson LL, Welniak EJ. Income Measurement Error in Surveys: A Review. *Journal of Official Statistics* 2000; 16: 331-361.

O'Donnell O, van Doorslaer E, Wagstaff A, Lindelow M. *Analyzing health equity using household survey data: a guide to techniques and their implementation*. The World Bank: Washington DC, 2008.

Ogwang T. Bounds of the Gini index using sparse information on mean incomes. *Review of Income and Wealth* 2003; 49: 415-423.

Rasche RH, Gaffney J, Koo AYC, Obst N. Functional forms for estimating the Lorenz curve. *Econometrica* 1980; 48: 1061-1062.

Sarabia J-M, Castillo E, Slottje DJ. An ordered family of Lorenz curves. *Journal of Econometrics* 1999; 91: 43-60.

Schader M, Schmid F. Fitting Parametric Lorenz Curves to Grouped Income Distributions – A Critical Note. *Empirical Economics* 1994; 19: 361-370.

van Doorslaer E, Koolman X. Explaining the differences in income-related health inequalities across European countries. *Health Economics* 2004; 13: 609-628.

van Doorslaer E, Koolman X, Jones AM. Explaining income-related inequalities in doctor utilisation in Europe. *Health Economics* 2004; 13: 629-647.

van Doorslaer E, Masseria C, Koolman X. Inequalities in access to medical care by income in developed countries. *Canadian medical association journal* 2006; 174: 177-183.

van Doorslaer E, Wagstaff A, Bleichrodt H, Calonge S, Gerdtham U-G, Gerfin M, Geurts J, Gross L, Häkkinen U, Leu R, O'Donnell O, Propper C, Puffer F, Rodriguez M, Sundberg G, Winkelhake O. Socioeconomic inequalities in health: some international comparisons. *Journal of Health Economics* 1997; 16: 93-112.

van Doorslaer E, Wagstaff A, Calonge S, Christiansen T, Gerfin M, Gottschalk P, Janssen R, Lachaud C, Leu RE, Nolan B, O'Donnell O, Paci P, Pereira J, Pinto CG, Propper C, Retie J, Rochaix L, Rodriguez M, Rutten F, Upward R, Wolfe B. Equity in the delivery of health care: some cross-country comparisons. *Journal of Health Economics* 1992; 11: 389-411.

van Doorslaer E, Wagstaff A, van der Burg H, Christiansen T, De Graeve D, Duchesne I, Gerdtham U-G, Gerfin M, Geurts J, Gross L, Häkkinen U, John J, Klavus J, Leu RE, Nolan B, O'Donnell O, Propper C, Puffer F, Schellhorn M, Sundberg G, Winkelhake O. Equity in the delivery of health care in Europe and the US. *Journal of Health Economics* 2000; 19: 553-583.

Van Ourti T. Measuring horizontal inequity in Belgian health care using a Gaussian random effects two part count data model. *Health Economics* 2004; 13: 705-724.

Van Ourti T, Clarke P. The bias of the Gini coefficient due to grouping: revisiting first-order corrections. *Tinbergen Institute Discussion paper series* 2008-095/3. Erasmus University Rotterdam.

van Praag B, Hagenaars A, van Eck W. The influence of classification and observation errors on the measurement of income inequality. *Econometrica* 1983; 51: 1093-1108.

Villaseñor J, Arnold B. Elliptical Lorenz curves. *Journal of econometrics* 1989; 40: 327-338.

Wagstaff A. Inequality aversion, health inequalities and health achievement. *Journal of health economics* 2002; 21: 627-641.

Wagstaff A. The bounds of the Concentration Index when the variable of interest is binary, with an application to immunization inequality. *Health Economics* 2005; 14(4): 429-432.

Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. *Social Science and Medicine* 1991; 33(5): 545–557.

Wagstaff A, van Doorslaer E. Equity in health care finance and delivery. In: Culyer A, Newhouse J (eds), *Handbook of Health Economics*. North Holland: Amsterdam, 2000; 1804–1862.

Wagstaff A, van Doorslaer E. Overall versus socioeconomic health inequality: a measurement framework and two empirical illustrations. *Health Economics* 2004; 13: 297-301.

Wagstaff A, van Doorslaer E, Paci P. Equity in the Finance and Delivery of Health Care: Some Tentative Cross-Country Comparisons. *Oxford Review of Economic Policy* 1989; 5: 89–112.

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

Wooldridge JM. *Econometric analysis of cross section and panel data*. The MIT Press: London, 2002.

Figure 1: hypothetical example of the bias in the Gini and concentration indexes resulting from categorical income data

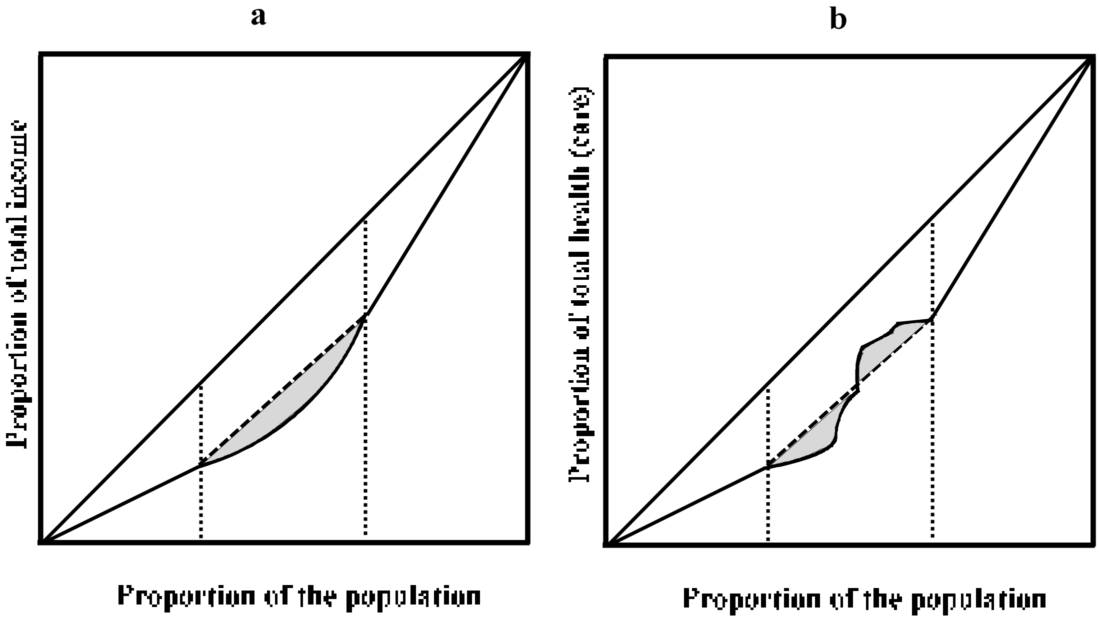
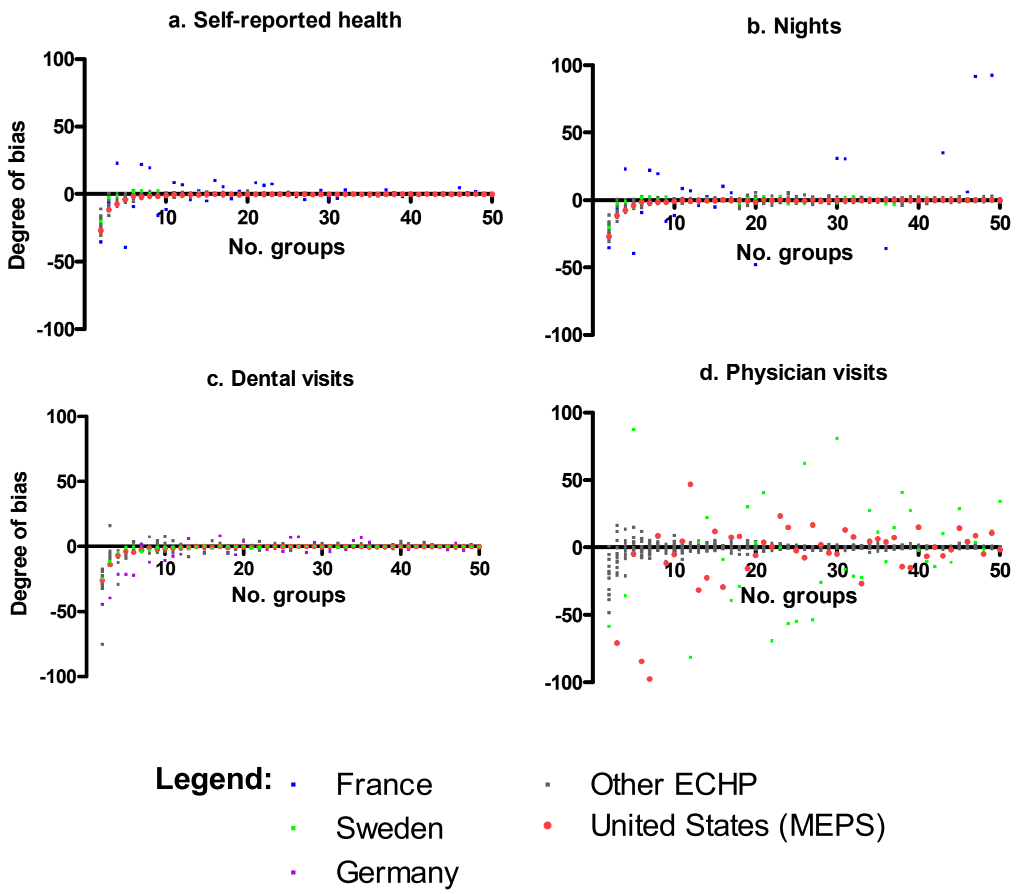
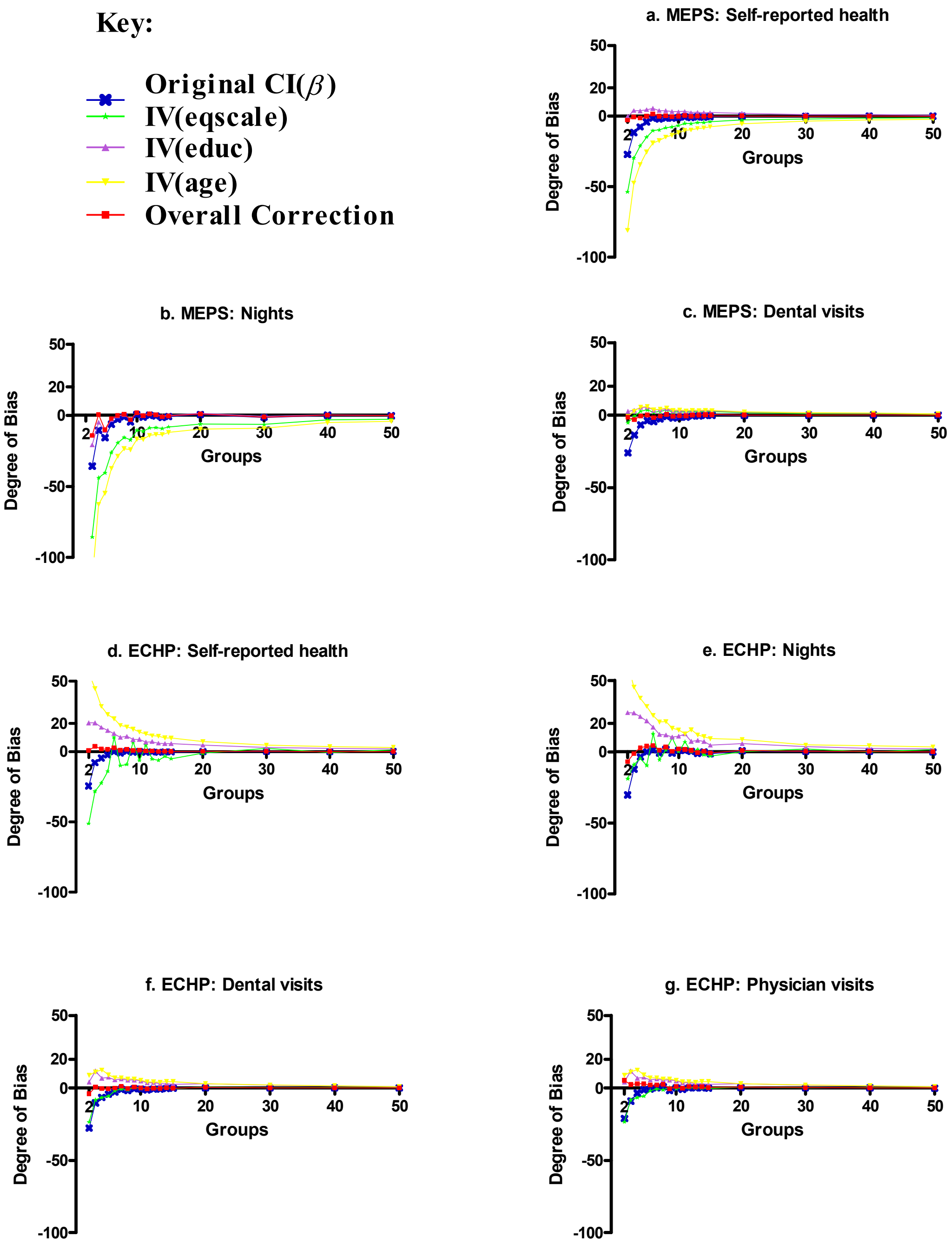


Figure 2: The impact of income grouping on the concentration index based on ECHP and MEPS data



Note: (i) Highlighted dots refer to the impact of income grouping of concentration indices based on individual micro-data that are insignificant at the 10 percent level: these are France for self-reported health and nights, dental visits for Germany and Sweden/US for Physician visits. (ii) The following countries have been excluded due to lack of data: NIGHTS (Germany); DENT (France); PHYS (France).

Figure 3: The impact of income grouping on the concentration index and various correction methods using data from the United States (using MEPS 2000) and ECHP countries



Notes: (i) Median degree of error reported for ECHP countries in figures d. to g. (ii) The following countries have been excluded from comparison due to either lack of data or a non-significant concentration index: SRH (France); NIGHTS (France, Germany); DENT (France, Germany); PHYS (France, Sweden, United States)

Table 1: summary statistics of variables in ECHP

Country	AT	BE	DK	FI	FR	DE	GR	IRL	IT	LU	NL	PT	ES	SE	UK
Year	1995	1995	1995	1996	1995	1995	1995	1995	1995	1995	1995	1995	1995	1995	1995
N	7347	6236	5473	7463	13662	8691	11978	8377	17459	1945	9044	11646	15693	5198	7514
Equivalent income (EQINC)															
Mean	214405	619721	136224	86723	94441	31678	1744296	8365	17222	897792	29469	935378	1167541	133855	9505
Std.dev	123694	529787	86236	47837	99145	18740	1325929	8419	11742	553552	20833	734345	762602	62295	6762
Self-reported health (SRH)															
Mean	0.903	0.908	0.914	0.901	0.892	0.901	0.905	0.919	0.891	0.903	0.909	0.866	0.891	0.912	0.904
Std.dev	0.066	0.052	0.060	0.056	0.080	0.058	0.075	0.047	0.070	0.059	0.049	0.087	0.075	0.057	0.063
C	0.006	0.006	0.007	0.003	0.001	0.004	0.009	0.004	0.004	0.006	0.003	0.013	0.006	0.004	0.007
Hospital nights (NIGHTS)															
Mean	2.029	1.369	1.216	1.084	1.104	NA	0.939	1.019	1.157	1.458	0.944	0.775	1.084	0.379	1.125
Std.dev	8.729	8.475	8.004	6.649	6.223	NA	5.923	6.061	6.551	8.691	5.410	6.193	7.338	3.219	7.623
C	-0.054	-0.274	-0.238	-0.138	-0.003	NA	-0.116	-0.070	-0.058	-0.126	-0.145	-0.201	-0.106	0.004	-0.208
Dental visitis (DENT)															
Mean	1.522	1.404	1.736	1.534	NA	2.030	0.715	0.721	1.004	1.645	1.682	0.509	0.845	1.091	1.421
Std.dev	2.372	2.421	1.746	2.421	NA	2.848	2.119	1.607	2.702	2.885	1.743	1.449	2.342	0.786	1.818
C	0.076	0.077	0.084	0.074	NA	0.006	0.161	0.196	0.126	0.033	0.058	0.294	0.110	0.004	0.085
Physician visits (PHYS)															
Mean	6.561	6.807	3.715	3.171	NA	7.822	3.694	3.950	4.957	5.098	4.615	4.065	5.526	1.793	4.829
Std.dev	9.627	9.984	5.745	4.155	NA	11.443	6.498	6.800	7.966	5.837	7.473	5.893	9.751	1.486	6.907
C	-0.026	-0.113	-0.070	0.015	NA	-0.033	-0.097	-0.097	-0.039	-0.036	-0.050	-0.027	-0.066	0.004	-0.087

Note: country abbreviations are Austria (AT), Belgium (BE), Denmark (DK), Finland (FI), France (FR), Germany (DE), Greece (GR), Ireland (IRL), Italy (IT), Luxembourg (LU), Netherlands (NL), Portugal (PT), Spain (ES), Sweden (SE), and the United Kingdom (UK). *C* is the concentration index calculated from individual level data using equation (1) and (2). Finally note that the ECHP reports Italian incomes in amounts of 000's Lira.

Table 2: summary statistics of variables in MEPS

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
N	15471	23929	16547	16958	17212	23323	26938	23177	23389	23167
Equivalent income (EQINC)										
Mean	25971	25971	28324	29276	30110	30535	30348	29270	30128	31259
Std.dev	20187	20708	23638	22805	23678	24405	25033	24955	25309	26146
Self-reported health (SRH)										
Mean	0.886	0.885	0.883	0.888	0.886	0.884	0.882	0.881	0.881	0.881
Std.dev	0.086	0.086	0.087	0.081	0.082	0.083	0.086	0.086	0.086	0.086
C	0.014	0.012	0.011	0.012	0.012	0.012	0.013	0.013	0.013	0.012
Hospital nights (NIGHTS)										
Mean	0.577	0.581	0.604	0.514	0.582	0.569	0.569	0.587	0.616	0.577
Std.dev	4.076	3.598	4.064	3.506	4.405	3.734	4.492	4.187	5.138	3.892
C	-0.2266	-0.239	-0.2145	-0.223	-0.2533	-0.2677	-0.2479	-0.2439	-0.2472	-0.2163
Dental visitis (DENT)										
Mean	1.042	0.953	0.931	0.983	0.928	0.955	0.937	0.890	0.881	0.873
Std.dev	1.883	1.814	1.778	1.859	1.827	1.770	1.725	1.669	1.675	1.678
C	0.202	0.220	0.207	0.206	0.205	0.212	0.231	0.244	0.249	0.238
Physician visits (PHYS)										
Mean	3.613	3.486	3.518	3.371	3.442	3.588	3.552	3.510	3.542	3.467
Std.dev	6.208	6.257	6.392	6.015	5.895	6.418	6.041	6.050	6.250	6.385
C	-0.0063	0.0009	0.0001	-0.0076	0.0018	-0.0097	0.0028	-0.0002	0.0047	0.0149

Note: *C* is the concentration index calculated from individual level data using equation (1) and (2).

Table 3: Mean squared error of the impact of income grouping on the concentration index and various correction methods based on MEPS 2000 data

		Correction Method			
Groups (range)	*C*	*IV(educ)*	*IV(eqscale)*	*IV(age)*	*Overall Correction Approach*
Self-reported health (SRH)					
16-50	0.04	1.58	3.90	14.21	0.02
6-15	1.35	11.67	48.22	159.30	0.23
2-5	236.58	12.34	1107.09	2654.96	2.65
All groups	19.61	4.66	103.21	260.26	0.28
Hospital nights (NIGHTS)					
16-50	0.25	0.19	14.78	34.35	0.18
6-15	421.95	147.91	2907.17	5360.62	79.54
2-5	35.44	12.54	279.03	530.89	6.97
All groups	0.25	0.19	14.78	34.35	0.18
Dental visitis (DENT)					
16-50	0.04	2.08	1.36	2.58	0.03
6-15	3.92	10.18	5.37	12.76	0.56
2-5	235.02	18.58	11.97	18.47	3.17
All groups	20.02	5.08	3.04	5.95	0.39

Note: Excludes MSE of PHYS as it was not significant.

Table 4: Summary of comparisons of mean squared error for IV and overall correction across all countries in ECHP and MEPS

	No. of countries where correction produces lower MSE			
Groups (range)	*IV(educ)*	*IV(eqscale)*	*IV(age)*	*Overall Correction Approach*
Self-reported health (SRH) (15 countries)				
16-50	0	0	0	1
6-15	0	0	0	5
2-5	3	0	0	13
All groups	3	0	0	13
Hospital nights (NIGHTS) (14 countries)				
16-50	1	0	0	4
6-15	1	0	0	4
2-5	4	0	0	10
All groups	3	0	0	10
Dental visitis (DENT) (14 Countries)				
16-50	1	0	0	7
6-15	1	6	0	8
2-5	6	4	3	13
All groups	6	3	3	13
Physician visits (PHYS)(13 countries)				
16-50	0	0	0	3
6-15	0	0	0	3
2-5	2	0	0	9
All groups	2	0	0	9

Note: The following countries have been excluded from comparison due to either lack of data or a non-significant concentration index: SRH (France); NIGHTS (France. Germany); DENT (France. Germany); PHYS (France. Sweden. United States)

Table 5: Overall Correction Approach: A Case Study using comparisons across 15 ECHP countries

	Self-reported health (SRH)		*Hospital nights (NIGHTS)*		*Dental visitis (DENT)*		*Physician visits (PHYS)*	
Groups	*C*	OCA	*C*	OCA	*C*	OCA	*C*	OCA
50	1	1	1	1	0	0	0	0
40	1	1	1	1	1	1	1	0
30	1	1	1	1	1	0	1	0
20	2	1	0	0	0	1	1	2
15	0	1	1	1	0	0	2	1
14	1	1	2	1	2	1	1	0
13	1	1	1	1	1	0	3	1
12	1	2	1	2	0	0	2	2
11	1	0	0	0	2	2	1	1
10	3	2	1	1	2	0	2	2
9	0	2	2	2	1	0	3	2
8	1	1	3	3	2	2	1	2
7	1	2	3	2	3	4	2	0
6	0	2	4	2	2	1	2	2
5	2	2	2	3	4	0	3	2
4	4	1	2	3	6	2	1	2
3	5	4	4	3	6	0	6	3
2	24	6	22	8	24	3	13	6

Notes: (i) 50-2: change in rank from income grouping (*C*)/ applying overall correction approach (OCA) to each country while using the *C* from the full sample for all other countries. (ii) The following countries have been excluded from comparison due to either lack of data or a non-significant concentration index: SRH (France); NIGHTS (France. Germany); DENT (France. Germany); PHYS (France. Sweden).

Table 6: Overall Correction Term: A Case Study using comparisons across 10 years of MEPS data

	Self-reported health (SRH)		*Hospital nights (NIGHTS)*		*Dental visitis (DENT)*	
Groups	*C*	OCA	*C*	OCA	*C*	OCA
50	0	0	0	0	0	0
40	0	0	0	0	0	0
30	0	0	1	1	0	0
20	1	1	2	1	0	0
15	1	1	1	1	2	1
14	2	0	3	0	0	0
13	2	1	4	2	0	0
12	2	2	2	2	2	0
11	2	0	3	1	0	1
10	2	1	5	2	2	0
9	2	1	3	1	2	1
8	2	0	6	5	3	0
7	4	2	6	4	4	1
6	6	3	12	6	8	1
5	10	2	10	7	13	0
4	20	1	23	15	18	1
3	38	2	33	8	26	2
2	45	7	45	14	45	2

Note: 50-2: change in rank from income grouping (*C*)/ applying overall correction approach (OCA) to each year while using the *C* from the full sample for all other years. PHYS has been omitted as the concentration index is not significant for several of the years.